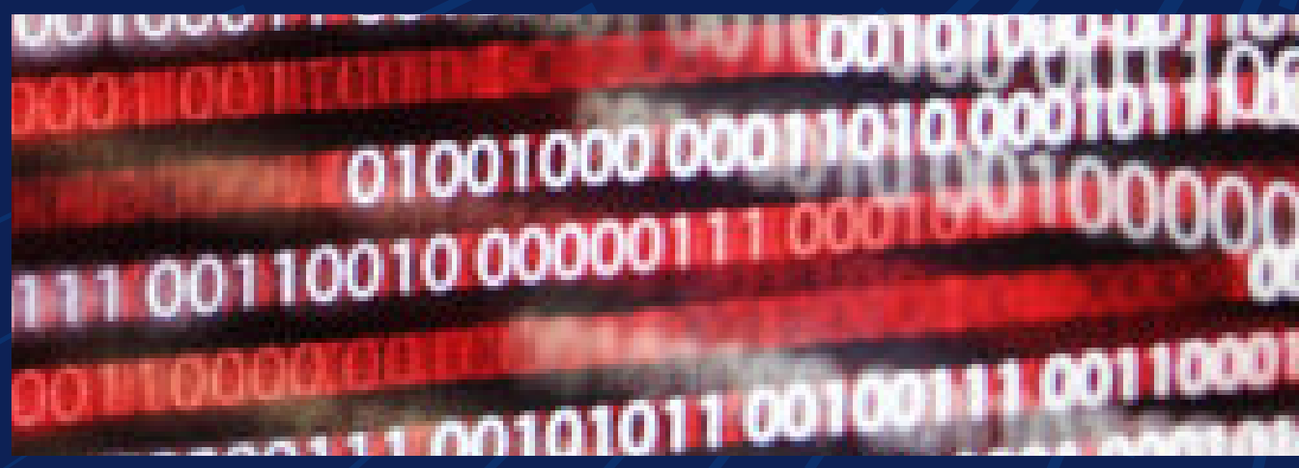
Combattre les idées reçues en matière de cybersécurité :

Bousculer les mythes liés à la sécurité de l'IA



L'IA transforme les secteurs d'activité, mais lorsqu'il s'agit de sécuriser l'IA, bon nombre d'organisations s'engouffrent dans des mythes qui la font sembler plus complexe qu'elle ne l'est réellement. La vérité ? La protection des systèmes d'IA n'impose nullement de repartir de zéro : transposer les principes de cybersécurité existants aux problématiques propres à l'IA permet déjà de gagner du terrain.

Chez Dell Technologies, nous comprenons l'architecture sous-jacente de l'IA et pouvons vous aider à adapter vos solutions actuelles afin qu'elles s'intègrent à ce nouveau cadre. Brisons les mythes les plus courants autour de la sécurité de l'IA et découvrons les vérités qui vous aideront à sécuriser efficacement vos systèmes.

Mythe n° 1 : « Les systèmes d'IA sont trop complexes pour être sécurisés. »

La vérité : il est vrai que l'IA soulève de nouveaux risques en matière de cybersécurité, comme l'injection d'invites, la manipulation des données et la divulgation d'informations sensibles, pour n'en citer que quelques-uns. Quant aux systèmes d'IA agentique, ils offrent également une surface d'attaque plus large et peuvent être exploités pour manipuler les résultats ou augmenter les privilèges d'accès.

Cela dit, bien qu'il soit essentiel de reconnaître ces vulnérabilités et de mettre en œuvre des mesures de sécurité pour protéger les systèmes d'IA contre les menaces traditionnelles et spécifiques à l'IA, il n'en reste pas moins possible de maîtriser les risques et de sécuriser les modèles d'IA. Il est important de garder à l'esprit que les systèmes d'IA ont besoin d'immenses quantités de données en amont et qu'ils génèrent à leur tour d'immenses quantités de données en aval. Dès lors, la protection des données occupe une place centrale dans les stratégies de sécurité clés, de même que :

- les principes Zero-Trust, tels que la gestion des identités, l'accès basé sur les rôles et la vérification continue ;
- les tests de pénétration réguliers et la gestion des vulnérabilités pour identifier les faiblesses ;
- la journalisation et les audits pour valider les entrées et sorties de données.

Mythe n° 2 : « Aucun de mes outils existants n'est capable de sécuriser l'IA. »

La vérité : la sécurité de l'IA ne consiste pas à faire table rase de l'existant, mais à travailler plus intelligemment avec les outils dont vous disposez déjà. La plupart des outils de cybersécurité existants peuvent être adaptés pour sécuriser efficacement les systèmes d'IA. Fondamentalement, l'IA n'est autre qu'une charge applicative de plus dans l'arsenal qui stimule votre activité, même si elle présente des caractéristiques uniques. Les pratiques de cybersécurité fondamentales, telles que la gestion des identités, la segmentation et la surveillance des réseaux, la protection des terminaux et la protection des données, restent essentielles pour protéger les environnements d'IA. La clé consiste à adapter ces pratiques pour relever les défis spécifiques à l'IA, notamment en protégeant les données d'entraînement, en sécurisant les algorithmes et en atténuant les risques comme les données d'entrée contradictoires.

Une défense robuste commence par une bonne hygiène sur le plan de la cybersécurité, avec l'application de correctifs système, un contrôle d'accès et une solide gestion des vulnérabilités. L'important est d'adapter ces pratiques aux risques propres à l'IA. En intégrant à votre approche de sécurité actuelle des stratégies axées sur l'IA et en adoptant les bons outils, la sécurité de l'IA devient à la fois facile à gérer et efficace.

Il est cependant important de souligner que la modernisation du matériel peut jouer un rôle essentiel dans la lutte contre les cyberattaques. Par exemple, les PC IA modernes créent une première ligne de défense solide contre un vecteur d'attaque majeur : les points de terminaison. Avec la fin du support de Windows 10, les PC obsolètes s'exposent à un risque. De plus, Windows 11 nécessite le module TPM (Trusted Platform module) version 2.0, une puce de sécurité qui facilite le chiffrement, le démarrage sécurisé et la protection contre les attaques au niveau du firmware. La plupart des PC d'ancienne génération ne disposent pas du module TPM ou ne prennent en charge qu'une version plus ancienne. Dell propose des PC IA professionnels sécurisés qui intègrent ces améliorations.

Il en va de même pour les infrastructures d'IA telles que les serveurs et le stockage. La Dell AI Factory comprend du matériel optimisé pour la sécurité de l'IA et contient un certain nombre de fonctions de sécurité intégrées allant d'une chaîne logistique sécurisée à l'immuabilité des données, en passant par l'isolement et le chiffrement.

Mythe n° 3 : « La sécurité de l'IA se limite à une simple protection des données. »

La vérité : la sécurité de l'IA va au-delà d'une simple protection des données. Elle suppose de protéger l'ensemble de l'écosystème de l'IA, à savoir les modèles, les API, les sorties, les systèmes et les appareils. Les risques d'utilisation abusive ou d'exploitation de l'IA augmentent à mesure que s'intensifie son intégration dans les applications stratégiques. En l'absence de mesures de sécurité robustes, les modèles d'IA peuvent être altérés pour générer des résultats nuisibles ou trompeurs, les API peuvent être exploitées pour obtenir un accès non autorisé à des systèmes sensibles, et les sorties peuvent exposer accidentellement des informations privées ou confidentielles.

Une sécurité complète de l'IA nécessite une approche multicouche, qui consiste à protéger les modèles contre les attaques qui tentent de manipuler les données d'entrée afin d'induire en erreur les systèmes d'IA, à sécuriser les API à l'aide de méthodes d'authentification solides pour empêcher toute utilisation non autorisée et **à surveiller les sorties en continu** pour identifier les schémas inhabituels ou suspects susceptibles de signaler une attaque ou un dysfonctionnement. Une sécurité efficace de l'IA non seulement garantit l'intégrité et la fiabilité des systèmes d'IA, mais renforce également la confiance des utilisateurs et des parties prenantes en atténuant les risques d'utilisation malveillante ou de conséquences involontaires.

Mythe n° 4 : « L'IA n'a pas besoin de surveillance humaine. »

La vérité : la gouvernance et la surveillance humaine sont essentielles pour garantir des systèmes d'IA qui respectent les règles d'éthique et qui soient prévisibles et conformes aux valeurs humaines. Les systèmes d'IA avancés, en particulier l'IA agentique qui intègre des capacités de prise de décision autonomes, soulèvent des problématiques particulières qui exigent des protections robustes. Sans surveillance adéquate, ces systèmes peuvent s'écarter des objectifs prévus ou présenter des comportements involontaires potentiellement à risque.

Pour résoudre ce problème, il est essentiel de définir des limites claires, de mettre en œuvre des mécanismes de contrôle en couches et d'impliquer l'humain en continu dans les processus décisionnels stratégiques. Des audits réguliers, une transparence des opérations d'IA et des tests approfondis peuvent renforcer la responsabilité et la confiance, contribuant ainsi à prévenir les abus et à promouvoir un déploiement responsable des technologies d'IA.

Pratiques d'excellence pour renforcer la sécurité de l'IA

Pour combler les failles de sécurité spécifiques à l'IA, les organisations doivent adopter une approche proactive et stratégique. Vous trouverez ici 10 pratiques d'excellence à adopter pour sécuriser vos systèmes d'IA :

Architecture de sécurité en couches :
Utiliser la segmentation, les pare-feux et l'authentification forte pour protéger votre infrastructure, vos logiciels et vos données sur chaque couche.

Sécurisation de la chaîne logistique :
Mettre en œuvre un solide programme de gestion des fournisseurs. Auditer les fournisseurs et les composants tiers, valider leur intégrité et utiliser un code signé pour éviter les vulnérabilités dans le cycle de développement de l'IA.

Protection des données et modèles d'entraînement :
Se prémunir des données falsifiées, des entrées contradictoires et des autres menaces en surveillant l'intégrité des données et en appliquant des outils de validation robustes.

Intensification des contrôles d'accès :
Appliquer les principes du moindre privilège, mettre en œuvre un contrôle d'accès basé sur les rôles (RBAC), changer régulièrement les informations d'identification et vérifier les autorisations dans le cadre d'un audit pour empêcher tout accès non autorisé.

Sécurisation des API :
Utiliser des protocoles d'authentification stricts (comme Oauth 2.0), appliquer un chiffrement HTTPS et mettre à jour régulièrement les API pour éliminer les failles de sécurité potentielles.

Surveillance et validation des résultats générés par l'IA :
Utiliser la détection des anomalies, la journalisation et les alertes pour surveiller les schémas inhabituels ou les comportements nuisibles dans les résultats générés par l'IA.

Planification de la résilience :
Sauvegarder régulièrement les données et tester les plans de reprise après sinistre afin de minimiser les interruptions de service et de garantir une reprise rapide en cas de violation de sécurité.

Mise en œuvre de mécanismes de chiffrement robustes :
Chiffrer les données sensibles au repos et en transit à l'aide d'algorithmes puissants, gérer les clés de chiffrement en toute sécurité et les remplacer régulièrement.

Audits de sécurité et tests de pénétration réguliers :
Évaluer fréquemment les vulnérabilités des systèmes et utiliser des tests de pénétration pour identifier les risques avant qu'ils ne puissent être exploités.

Formation du personnel aux pratiques d'excellence en matière de sécurité de l'IA :
Former régulièrement votre équipe au développement sécurisé, à la reconnaissance des menaces et au maintien de pratiques de sécurité rigoureuses pour prévenir les violations de sécurité.

DELL Technologies

Proposition de valeur de Dell : des solutions pratiques pour la sécurité de l'IA.

La sécurité de l'IA peut sembler complexe, mais elle n'est pas aussi déconcertante qu'elle n'y paraît. La vérité ? La sécurisation de l'IA n'est pas si différente de la sécurisation de vos charges applicatives existantes. L'important est de comprendre l'architecture et d'appliquer les stratégies appropriées. C'est là que Dell Technologies entre en jeu.

Nous démystifions la sécurité de l'IA en tirant parti de vos solutions actuelles et en les intégrant de manière transparente dans des architectures axées sur l'IA. Nous gérons toutes sortes de menaces, comme l'injection d'invites, l'utilisation abusive d'API et les attaques malveillances sans vous obliger à entreprendre une refonte complète de votre infrastructure.

L'expertise de Dell consiste à briser les mythes autour de la sécurité de l'IA et à démontrer qu'elle est réalisable. Que vous en soyez aux premières étapes de votre parcours d'adoption de l'IA ou que vous cherchiez à améliorer vos défenses, nous vous aiderons à protéger vos investissements, à sécuriser vos systèmes et à bâtir un avenir numérique résilient, efficacement et en toute confiance. Ensemble, simplifions la sécurité de l'IA.

Produits et solutions Dell qui peuvent vous aider

Solution Dell proposée	Description
Dell AI Factory	La Dell AI Factory sécurise les charges applicatives d'IA via une chaîne logistique sécurisée, qui garantit une infrastructure fiable, de l'étape de développement jusqu'à la phase de déploiement. Avec des fonctionnalités telles que l'immuabilité, l'isolement et le chiffrement des données, elle protège les modèles et les datasets sensibles, assure une protection contre les cybermenaces et garantit des opérations d'IA évolutives, efficaces et transparentes dans des environnements dynamiques axés sur les données.
Cyber-résilience	PowerProtect sécurise les charges applicatives d'IA grâce à des fonctionnalités avancées, telles que l'immuabilité et l'isolement, pour garantir l'intégrité des données et leur protection face aux cybermenaces. La solution prend en charge un chiffrement et une détection des anomalies de bout en bout, tout en assurant une récupération rapide afin de minimiser les interruptions de service.
Dell Trusted Workspace (sécurité des points de terminaison)	Une combinaison de fonctionnalités intégrées et complémentaires en option conçues pour sécuriser les PC IA professionnels et leurs charges applicatives d'IA. Inspirées des pratiques de chaîne logistique sécurisées, les fonctionnalités intégrées incluent SafeBIOS et SafeID avec TPM. Les modules complémentaires Secured Component Verification et SafeID with ControlVault sont proposés en option, de même que les logiciels partenaires CrowdStrike et Absolute conçus pour optimiser la sécurité de l'espace de travail.
Services de conseil en sécurité IA	Une suite de services conçue pour vous aider à développer et mettre en œuvre une stratégie de sécurité de l'IA complète. La suite comprend, entre autres offres, des services de conseil, un vCISO spécialiste de l'IA et une planification de la sécurité des données.
Managed Security Operations for AI	Procure une excellente visibilité sur l'ensemble de la pile pour détecter les menaces et réagir rapidement. Intègre diverses fonctionnalités : Managed Detection and Response, Managed AI Guard, Penetration Testing for AI, Incident Response and Recovery Services, etc.
Intégration de logiciels de sécurité	Concevoir, installer et configurer des outils de sécurité qui protègent la gestion des accès, les applications, les réseaux, les Clouds, etc.

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