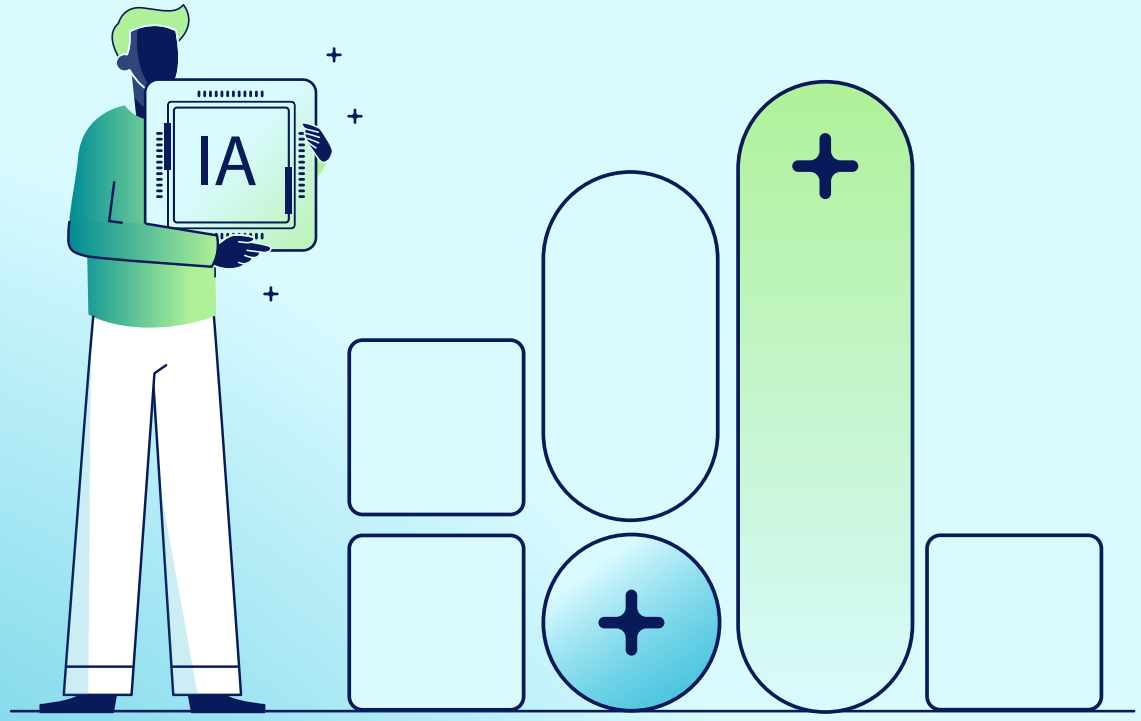
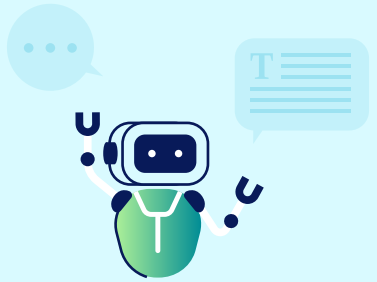
netskope

Le guide stratégique pour sécuriser l'IA

Un guide pratique pour sécuriser l'IA de bout en bout, partout

Le guide stratégique pour sécuriser l'IA





Sommaire

Introduction

Les technologies d'intelligence artificielle (IA) se sont rapidement imposées comme des outils utiles et importants pour de nombreuses entreprises. Avec de nouvelles capacités et de nouveaux cas d'utilisation émergeant en permanence, l'IA fait désormais partie intégrante de la pile technologique de la plupart des entreprises.

En outre, l'ascension rapide de l'IA a été marquée par des niveaux d'investissement élevés. Selon les analystes d'IDC, le marché mondial des dépenses informatiques liées à l'IA devrait atteindre près de 750 milliards de dollars d'ici 2028, les dépenses spécifiques à l'IA générative s'élevant à un peu plus de 300 milliards de dollars.[1]

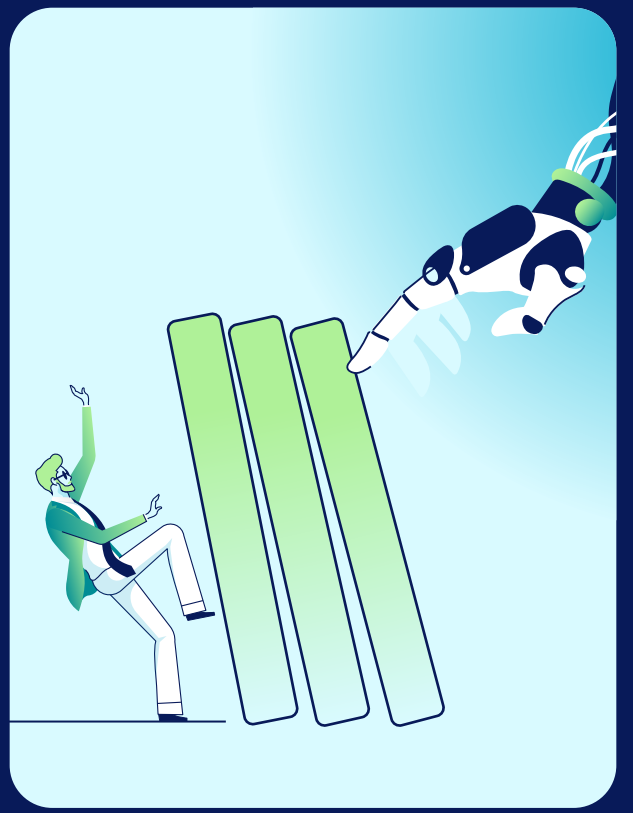
+ Le marché mondial des dépenses informatiques liées à l'IA devrait atteindre près de 750 milliards de dollars d'ici 2028, les dépenses spécifiques à l'IA générative s'élevant à un peu plus de 300 milliards de dollars.

Les professionnels de la sécurité sont conscients des risques que présentent les applications de l'IA dans leur environnement, et ces risques ne cessent de croître. Au niveau d'adoption le moins élevé, les données sont partagées avec des applications tierces dans le cloud. Du point de vue de la sécurité, cette pratique soulève des questions quant à la nature des données que les employés intègrent dans ces systèmes et aux contrôles mis en place pour les gérer. L'arrivée de nouveaux protocoles standard, comme le Model Context Protocol ou MCP, facilite encore plus le partage de données avec les applications d'IA, ce qui aggrave ces risques.[2]

Les problèmes de sécurité sont appelés à s'intensifier à mesure que la technologie de l'IA d'entreprise évolue. Les systèmes d'IA agentique, par exemple, peuvent fonctionner de manière autonome pour atteindre des objectifs spécifiques ou exécuter des tâches définies sans nécessiter d'intervention humaine constante. Les analystes de Gartner prévoient que, d'ici 2028, 25 % des failles dans les entreprises seront liées à l'utilisation abusive d'agents d'IA.[3]

Compte tenu de l'évolution rapide des risques auxquels sont confrontés les professionnels de la sécurité, il n'est pas surprenant qu'ils recherchent de l'aide pour faire face à ce nouveau paysage. Dans cet eBook, nous décrivons les principaux problèmes de sécurité auxquels les entreprises sont confrontées aujourd'hui et les solutions que Netskope peut leur apporter.

+ Gartner prévoit que, d'ici 2028, 25 % des failles dans les entreprises seront liées à l'utilisation abusive d'agents d'IA.



[1] IDC Market Forecast, Worldwide Artificial Intelligence IT Spending Forecast, 2024-2028, Rick Villars et al., octobre 2024, Doc n° US52635424.

[2] Rapport Netskope 2025 sur les menaces liées au cloud https://www.netskope.com/netskope-threat-labs/cloud-threat-report/cloud-and-threat-report-2025

[3] Les principales prédictions de Gartner pour 2025.

Les problèmes de sécurité liés à l'IA

Les trois principaux problèmes auxquels sont confrontées les équipes de sécurité aujourd'hui

1 Extension de la surface d'attaque

À mesure que l'utilisation de l'IA évolue, passant d'outils d'IA générative pure (comme ChatGPT) à des capacités d'IA intégrées dans les applications d'entreprise et à des applications d'IA créées par le secteur privé, la surface d'attaque continue de s'étendre. Chaque étape introduit de nouveaux risques :

- Les outils publics d'IA générative comportent des risques d'exposition involontaire de données sensibles.
- Les fonctions d'IA intégrées dans les applications SaaS existantes peuvent ouvrir de nouvelles voies pour la fuite ou la manipulation de données.
- Les LLM hébergés de manière privée et les applications d'IA personnalisées introduisent de nouveaux vecteurs, tels que des contrôles d'accès mal configurés ou des vulnérabilités dans les pipelines de données.
- L'interconnexion des applications d'IA avec les sources de données, grâce à des protocoles innovants comme le MCP, accroît les vulnérabilités face aux risques d'exfiltration de données.

2 Exposition et exfiltration de données sensibles

Le risque le plus immédiat lié à l'adoption de l'IA est la perte de données, qu'elle soit accidentelle ou malveillante :

- L'exposition involontaire se produit lorsque des employés saisissent des données sensibles (par exemple, des informations confidentielles, des secrets commerciaux, des données réglementées) dans des modèles publics sans en mesurer les conséquences.
- Des initiés malveillants ou des attaquants externes peuvent exploiter les outils d'IA pour exfiltrer des données ou détourner les canaux de sortie du modèle.
- L'entraînement constitue également un risque : L'utilisation de données mal traitées pour l'entraînement des modèles peut conduire à des fuites d'informations confidentielles.

3 Gouvernance responsable de l'IA

À mesure que les systèmes d'IA prennent de l'ampleur, ils soulèvent des questions critiques de conformité et d'éthique qui touchent également à la sécurité :

- Les modèles d'IA peuvent involontairement encoder et propager des biais, ce qui entraîne une supervision réglementaire accrue ainsi qu'une atteinte à la réputation des organisations concernées.
- Le traitement inapproprié des données des employés ou des clients utilisées dans les flux de travail d'IA peut être contraire au RGPD, à l'HIPAA ou à d'autres lois sur la confidentialité des données.
- La substitution de la prise de décision humaine par le déploiement autonome de l'IA, en particulier dans les domaines à fort enjeu (par exemple, l'embauche, la sécurité, la finance), engendre des dilemmes éthiques et des lacunes en matière d'imputabilité.

Les fondements de la sécurité de l'IA

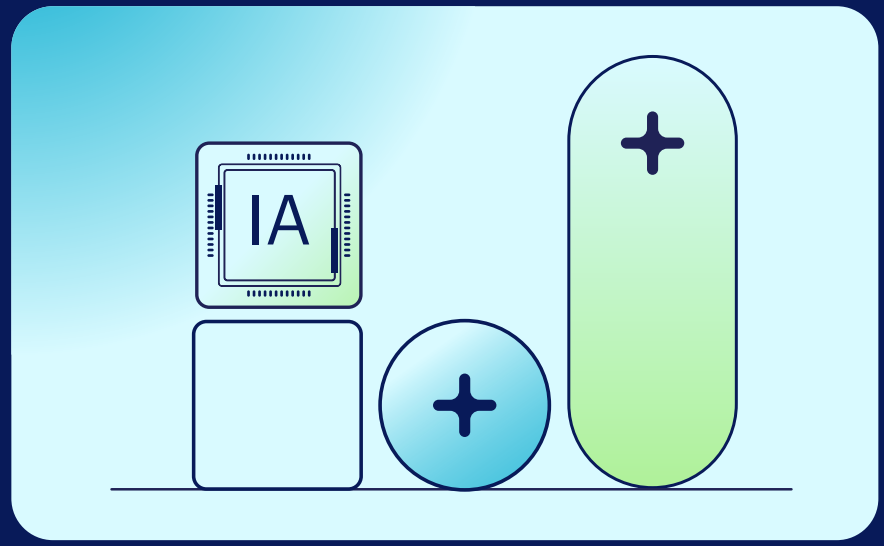
Le Zero Trust, un impératif

La sécurité de l'IA repose sur une approche zero trust, tout comme la sécurité du SaaS. Toutefois, elle présente des difficultés spécifiques liées à la manière dont les modèles d'IA traitent les données d'entrée et génèrent des données de sortie.

La sécurité de l'IA et du SaaS exige des contrôles d'accès stricts, une surveillance continue et une protection robuste des données pour atténuer les risques. Cependant, alors que la sécurité du SaaS se concentre principalement sur la protection des applications et des interactions avec les utilisateurs, la sécurité de l'IA doit également tenir compte de l'intégrité des données d'entraînement, de l'accès au modèle et du potentiel de manipulation adverse. Il est donc essentiel d'appliquer des politiques de sécurité adaptées au contexte et de détecter les menaces en temps réel pour prévenir les fuites de données, les accès non autorisés et l'exploitation des modèles d'IA.

Le zero trust doit bénéficier d'un cadre solide pour garantir la sécurité de l'IA : chaque demande doit être vérifiée, chaque flux de données doit être surveillé et l'accès doit être accordé sur la base d'une évaluation dynamique des risques plutôt que sur la base d'autorisations statiques. Cette approche nécessite une visibilité granulaire des mouvements de données et des contrôles de sécurité adaptatifs qui s'ajustent en fonction du contexte en temps réel.

Avec les principes de zero trust en place, les entreprises peuvent adopter et faire évoluer en toute sécurité les technologies pilotées par l'IA sans compromettre la sécurité ou la conformité.



Conseil d'expert

La sécurité de l'IA repose sur une approche zero trust, tout comme la sécurité du SaaS. Toutefois, elle présente des difficultés spécifiques liées à la manière dont les modèles d'IA traitent les données d'entrée et génèrent des données de sortie.

Maîtriser la sécurité de l'IA

Six enjeux majeurs et leurs solutions

Enjeu n° 1 : manque de visibilité

Alors que les outils d'IA s'intègrent dans les flux de travail quotidiens, les entreprises sont confrontées à un problème de sécurité fondamental : elles ne peuvent pas sécuriser ce qu'elles ne peuvent pas voir.

Les employés accèdent à des applications approuvées et non approuvées à l'aide d'identifiants professionnels et personnels, ce qui brouille les frontières entre les utilisations approuvées et les autres. Cette prolifération incontrôlée augmente le risque de fuite de données, de perte de propriété intellectuelle et de violation de la conformité, en particulier lorsque des informations sensibles sont introduites dans des services d'IA non gérés ou fantômes.

La plupart des entreprises ne disposent pas de la visibilité granulaire nécessaire pour différencier les utilisations risquées et légitimes de l'IA. Les outils traditionnels ne parviennent pas à identifier les interactions spécifiques des modèles d'IA, à distinguer les comptes personnels des comptes d'entreprise et à fournir des informations en temps réel au niveau de l'utilisateur, de l'application ou de l'activité. Faute d'avoir une visibilité approfondie sur comment et où l'IA est utilisée, les équipes de sécurité ne sont pas en mesure de détecter les points d'exposition potentiels.

La solution de Netskope

Alors que les entreprises adoptent de plus en plus d'outils d'IA, il est crucial de conserver une visibilité et un contrôle sur leur utilisation. Netskope propose une solution complète qui permet de suivre les applications d'IA gérées et non gérées (fantômes), fournissant aux équipes de sécurité les informations dont elles ont besoin pour assurer une surveillance adéquate.

Les fonctionnalités clés incluent :

- **Connaissance approfondie des instances :** faites la distinction entre les applications d'IA personnelles et professionnelles (ChatGPT, Gemini, Copilot, etc.).
- **Tableau de bord de l'IA** : obtenez des informations approfondies sur les tendances d'utilisation de l'IA, les principales applications, la fréquence d'accès et les actions granulaires des utilisateurs, notamment les connexions, les publications, les chargements et les téléchargements.
- **Analyse du comportement des utilisateurs et des entités (UEBA) :** détectez les anomalies et les comportements à risque à l'aide de l'apprentissage automatique et identifiez les menaces telles que l'exfiltration de données, les risques d'initiés et les violations de politiques.
- **Visibilité fondamentale :** obtenez une meilleure visibilité sur votre écosystème d'IA, des utilisateurs aux applications, en passant par le trafic API et MCP. Netskope vous offre une vue unique sur l'utilisation, les ressources et les flux de données.

Cette visibilité globale permet aux équipes de sécurité d'agir rapidement et d'atténuer les risques liés à l'utilisation de l'IA dans l'entreprise.

Enjeu n° 2 : comprendre les risques liés aux applications d'IA

Les capacités de l'IA évoluent rapidement, tout comme le paysage des risques. Ce qui était autrefois une simple application SaaS peut désormais introduire discrètement des fonctions d'IA intégrées comme la génération de copies, les réponses intelligentes et les copilotes d'IA, sans en avertir les utilisateurs ou les équipes de sécurité. Cette tendance croissante entraîne un manque de transparence de plus en plus important sur les applications qui utilisent l'IA, la manière dont elles l'utilisent et les risques qu'elles introduisent pour l'entreprise.

Les équipes de sécurité doivent être en mesure d'évaluer dynamiquement les risques en fonction de la façon dont les fonctions d'IA sont intégrées, si elles conservent les données de l'entreprise ou les utilisent pour l'entraînement, et comment elles respectent les exigences de conformité. Sans ce niveau de connaissance, les entreprises risquent de s'exposer à des fuites de données, au vol de propriété intellectuelle, à des violations de la réglementation, voire à la manipulation de modèles d'IA. Alors que l'empreinte de l'IA dans le SaaS ne cesse de croître, la compréhension du risque applicatif n'est pas seulement une bonne pratique, c'est une nécessité pour toute entreprise qui cherche à adopter l'IA en toute sécurité.

La solution de Netskope

Netskope s'attaque à la complexité croissante des risques liés aux applications d'IA grâce à son Indice de confiance dans le cloud (CCI - Cloud Confidence Index), qui fournit des informations en temps réel et continuellement mises à jour sur plus de 85 000 applications cloud et SaaS. Grâce à des évaluations des risques dynamiques liés à l'IA, CCI aide les équipes de sécurité à garder une longueur d'avance sur les risques et à garantir la conformité.

Les fonctionnalités clés incluent :

- **Évaluation en temps réel des risques liés à l'IA :** identifiez les applications intégrant des capacités d'IA et comprenez les risques qui y sont associés.
- **Aperçu du traitement des données d'entreprise :** évaluez la façon dont les applications gèrent les données de l'entreprise, notamment en ce qui concerne la conservation des données, l'entraînement des modèles et le partage avec des tiers.
- **Suivi de la conformité :** restez en phase avec les exigences réglementaires telles que le RGPD, le SOC 2 et l'ISO 27001.
- **LLM et MCP sécurisés :** évaluez plus de 85 000 applications SaaS, y compris les applications d'IA et les fonctionnalités d'IA intégrées, ainsi que les serveurs MCP publics, en identifiant les attributs risqués, les types d'authentification et les versions de protocole.

Avec CCI, les équipes de sécurité peuvent faire face à la complexité des risques liés aux applications d'IA en toute confiance et garantir la sécurité et la conformité de leur entreprise.

Enjeu n° 3 : intégrité des modèles d'IA

Alors que les entreprises exploitent de plus en plus les outils d'IA générative (à la fois les modèles personnalisés et les applications d'entreprise comme Microsoft Copilot), garantir l'intégrité des données utilisées pour entraîner ces modèles devient une préoccupation de premier plan. Ces systèmes d'IA sont souvent entraînés sur de vastes jeux de données qui peuvent inclure des documents d'entreprise sensibles, des e-mails, des présentations, des feuilles de calcul et des informations commerciales exclusives.

Si des données sensibles ou propriétaires sont incorporées par inadvertance dans des jeux de données d'entraînement, elles peuvent être exposées non seulement par les résultats du modèle, mais aussi par de le biais de requêtes contradictoires, de fuites de données, et conduire à des violations potentielles de la conformité. À mesure que l'adoption de la GenAI se développe dans différents services, les équipes de sécurité ont de plus en plus de mal à contrôler la manière dont les données d'entraînement sont sourcées, validées et protégées.

+ Microsoft Copilot peut être entraîné sur le contenu de la suite Office d'un utilisateur, qu'il s'agisse de documents Word ou de feuilles de calcul Excel. Si des données confidentielles ou sensibles sont stockées à ces endroits et que les contrôles d'accès ne sont pas correctement configurés, il existe un risque que Copilot fasse apparaître des stratégies commerciales sensibles, des renseignements financiers ou des informations sur les clients dans ses réponses.

La solution de Netskope

Netskope One DSPM (gestion de la posture de sécurité des données) donne aux entreprises les moyens de surveiller et de protéger les données sensibles dans les environnements cloud et les référentiels de données. La détection et la classification des données critiques, telles que les dossiers financiers, les PII et la propriété intellectuelle, permettent à Netskope de s'assurer que ces informations ne sont pas utilisées pour entraîner des modèles d'IA sans autorisation appropriée.

Les fonctionnalités clés incluent :

- **Surveillance continue des environnements cloud :** détectez et classez les données sensibles en temps réel, afin d'empêcher toute utilisation non autorisée dans l'entraînement des modèles d'IA.
- **Visibilité sur l'accès et le partage des données :** obtenez des informations en temps réel sur la manière dont les données sont consultées et partagées dans le cloud afin de pouvoir prendre des mesures correctives immédiates, si nécessaire.
- **Conformité et prévention des fuites de données :** protégez les données sensibles pour garantir la conformité, empêcher les fuites de données et maintenir le contrôle sur la propriété intellectuelle.
- **Gestion performante de la posture de sécurité :** assurez une posture de données appropriée et découvrez, étiquetez et classez vos données structurées et non structurées.

Avec Netskope One DSPM, les entreprises peuvent protéger de manière proactive leurs données sensibles, en veillant à ce que l'entraînement des modèles d'IA reste sécurisé, conforme et contrôlé.

Enjeu n° 4 : menaces ciblant les systèmes d'IA

Les cybercriminels font évoluer leurs tactiques pour exploiter les vulnérabilités spécifiques à l'IA. Ils utilisent l'injection de prompt, l'empoisonnement des données et les entrées adverses conçues pour fausser les résultats ou exfiltrer des données sensibles. En outre, les applications d'IA sont souvent intégrées à des systèmes d'entreprise plus vastes, ce qui en fait un point d'entrée potentiel pour les mouvements latéraux, l'élévation de privilèges ou le vol de données.

Qu'il s'agisse d'un acteur malveillant qui tente de manipuler les résultats d'un modèle d'IA, d'extraire des données d'entraînement ou d'exploiter des contrôles d'accès faibles autour des API d'IA, la surface d'attaque s'étend rapidement. Ce problème est aggravé par l'absence de cadres de sécurité uniformes pour protéger les systèmes d'IA, de sorte que de nombreuses entreprises ne sont pas préparées à se défendre contre de nouveaux vecteurs d'attaque. À mesure que l'adoption de l'IA augmente, les équipes de sécurité doivent détecter et atténuer de manière proactive les menaces qui ciblent spécifiquement les environnements d'IA, avant que les données sensibles, les opérations ou les processus de prise de décision ne puissent être compromis.

La solution de Netskope

Netskope s'attaque aux menaces croissantes ciblant les systèmes d'IA avec une approche de sécurité multicouche qui intègre une protection avancée contre les menaces, une visibilité approfondie et des défenses spécifiques à l'IA.

Les fonctionnalités clés incluent :

- **Défense unifiée de l'IA :** Netskope One AI Guardrails atténue les attaques sophistiquées, y compris les tentatives d'injection de prompt et de jailbreaking, grâce à une analyse approfondie et en temps réel de la totalité du trafic.
- **Protection contre les menaces avancées :** utilisez l'apprentissage automatique, le sandboxing et l'analyse heuristique pour détecter et bloquer les menaces connues et celles de type zero-day, y compris les logiciels malveillants cachés dans les fichiers soumis aux outils d'IA.
- **Red Teaming et évaluation des vulnérabilités :** automatisez les simulations adverses pour découvrir les vulnérabilités et vous assurer que vos modèles privés sont sécurisés, conformes et résilients face aux menaces avancées avec Netskope One Red Teaming.
- **Surveillance proactive de l'activité de l'IA :** détectez les menaces et les vulnérabilités émergentes grâce à la surveillance en temps réel des interactions avec l'IA, afin de garantir une stratégie de défense complète.

En combinant ces technologies, Netskope fournit une solution intégrée qui aide les entreprises à sécuriser leurs systèmes d'IA contre les cybermenaces sophistiquées et les vecteurs d'attaque en constante évolution.

Enjeu n° 5 : exposition des données

L'un des enjeux les plus pressants et les plus importants en matière de sécurité de l'IA est le risque d'exposition des données. Lorsque les employés de différents services adoptent des outils d'IA afin de booster leur productivité, ils peuvent, sans le savoir, communiquer, ou charger dans des modèles d'IA publics, des données sensibles telles que du code source, des dossiers clients, des documents financiers ou des droits de propriété intellectuelle. Une fois exposées, ces données peuvent être conservées, utilisées pour l'entraînement de modèles ou même faire l'objet d'une fuite, en fonction des politiques de confidentialité et des pratiques de traitement des données de l'application.

Contrairement aux canaux traditionnels de partage des données, les plateformes d'IA peuvent agir comme des boîtes noires, offrant peu de transparence sur la manière dont les données sont stockées, consultées ou utilisées. En l'absence de garde-fous, les entreprises s'exposent à de graves risques, allant des violations de la réglementation et du vol de propriété intellectuelle à l'atteinte à la réputation et au désavantage concurrentiel.

+ Netskope Threat Labs a constaté que le code source était exposé dans près de 50 % des cas de violation des politiques liés à l'IA. C'est dire la facilité avec laquelle les actifs critiques d'une entreprise peuvent être compromis par des actions apparemment anodines, comme le fait de coller un bout de code dans un chatbot d'IA pour le déboguer ou l'optimiser.

La solution de Netskope

Netskope offre une protection complète et contextuelle des données de l'entreprise, au repos et en mouvement. En combinant des évaluations des risques en temps réel, des contrôles en ligne et basés sur les API, etdes contrôles de posture, les politiques de sécurité unifiées de Netskope permettent une gouvernance précise des interactions à la fois des utilisateurs et des données dans l'ensemble de l'entreprise.

Les fonctionnalités clés incluent :

- **Prévention avancée des pertes de données (DLP) :** protégez les informations sensibles contre l'exfiltration grâce à des outils d'IA, que les utilisateurs soient au bureau, chez eux ou en déplacement.
- **Contrôle granulaire :** bloquez ou limitez les opérations à haut risque, telles que le chargement de code source ou de documents confidentiels.
- **Encadrement des utilisateurs en temps réel :** sensibilisez les utilisateurs aux violations des règles à l'aide de messages visuels, afin de réduire les récidives.
- **Inspectez chaque demande et chaque réponse :** identifiez et bloquez la diffusion de données brevetées ou protégées par des droits d'auteur dans les réponses d'IA pour empêcher que votre propriété intellectuelle ne soit utilisée pour produire les résultats des modèles génératifs.
- **Sécurisez le trafic API :** authentifiez et centralisez la gestion du trafic et l'inspection du contenu entre les applications privées et les LLM.

Grâce à ces fonctionnalités, Netskope assure une protection des données complète et adaptative qui évolue dans l'ensemble de l'environnement IA et cloud d'une entreprise.

Enjeu n° 6 : gouvernance, conformité et utilisation éthique

Alors que l'adoption de l'IA s'accélère, les entreprises sont confrontées à une pression croissante pour s'aligner sur les normes de gouvernance émergentes, les exigences réglementaires et les attentes éthiques, en particulier dans les secteurs très réglementés comme la finance, les soins de santé et les pouvoirs publics. Les pays du monde entier introduisent rapidement des cadres et des mandats spécifiques à l'IA, comme en témoignent la loi européenne sur l'IA, le cadre de gestion des risques de l'IA du NIST et les décrets américains sur la sécurité de l'IA. Ces règlements visent à garantir un développement et un déploiement responsables des systèmes d'IA, en imposant la transparence, la confidentialité des données, l'explicabilité et la non-discrimination.

Cependant, le respect de ces normes n'est pas chose facile. Les équipes chargées de la sécurité et de la conformité doivent comprendre comment l'IA est utilisée dans leur environnement. Elles doivent s'assurer que les données sensibles ne sont pas conservées de manière inappropriée ou qu'elles ne font pas l'objet d'un apprentissage. Elles doivent également prouver qu'elles respectent les directives légales et éthiques en constante évolution.

La solution de Netskope

Netskope assure la gouvernance de l'IA et la préparation à la conformité grâce à une visibilité approfondie, un contrôle des politiques et des informations en temps réel sur l'utilisation de l'IA dans l'entreprise.

Les fonctionnalités clés incluent :

- **Application granulaire des politiques :** contrôlez la manière dont les données sont partagées avec les outils d'IA, en veillant à ce que les données sensibles ou réglementées ne soient pas utilisées pour l'entraînement non autorisé de modèles tiers.
- **Contrôles de conformité en temps réel :** bloquez les chargements d'informations de santé protégées (PHI) vers des applications non conformes ou arrêtez le traitement des données financières dans les outils ne disposant pas des certifications adéquates.
- **Prise en charge des cadres réglementaires :** facilitez la mise en conformité avec des cadres comme la loi européenne sur l'IA ou le cadre de gestion des risques de l'IA du NIST.
- **Modération du contenu en temps réel :** filtrez et contrôlez automatiquement le contenu nuisible ou discriminatoire, y compris les discours de haine, les crimes, les armes et la violence.
- **Maîtrisez la gouvernance des données d'IA :** sécurisez l'ensemble du cycle de vie des données grâce à la détection et à la classification automatisées, ainsi qu'à un renforcement proactif avant le déploiement, afin de garantir la protection et la conformité de votre propriété intellectuelle.

En combinant visibilité, intelligence de la conformité et application adaptative des politiques, Netskope permet aux entreprises d'adopter l'innovation en matière d'IA de manière responsable tout en répondant aux exigences éthiques et réglementaires actuelles et futures.

L’avenir de la sécurité de l’IA

Technologies émergentes et menaces

Alors que l’adoption de l’IA se développe et que de nouveaux cas d’utilisation, des copilotes aux agents d’IA personnalisés, se généralisent, le paysage des menaces évolue tout aussi rapidement. Tandis que la sécurité se concentre actuellement sur la protection des données et l’intégrité des modèles, deux nouveaux domaines de développement technologique sont en passe de soulever des difficultés encore plus grandes dans un avenir proche.

Tout d’abord, les systèmes d’IA agentiques, capables de prendre des décisions et d’agir avec un minimum de supervision humaine, sont en plein essor. Selon Gartner, d’ici 2028, au moins 15 % des décisions commerciales quotidiennes seront prises de manière autonome par l’IA agentique, contre pratiquement aucune aujourd’hui.[4] Cette évolution augmente considérablement la surface d’attaque, en particulier si les agents se voient accorder l’accès aux systèmes et aux données de l’entreprise via le MCP ou l’A2A (protocole agent à agent).

D’autre part, l’IA physique, telle qu’on l’observe dans les véhicules et les robots autonomes, gagne du terrain dans des secteurs comme la logistique, le transport et la fabrication. Ces systèmes introduisent des risques de sécurité dans le monde réel, où une IA compromise ou défaillante n’entraîne pas seulement une perte de données, mais aussi des dommages potentiels aux personnes et aux infrastructures.

À mesure que les capacités de l’IA se perfectionnent et s’intègrent plus profondément dans les opérations commerciales quotidiennes, les responsables de la sécurité doivent mettre en place une gouvernance stratégique et axée sur l’avenir.

Voici quelques éléments clés à prendre en compte pour garder une longueur d’avance :

- **Visibilité de l’utilisation de l’IA :** sachez quelles équipes construisent ou utilisent des modèles d’IA, qu’il s’agisse de l’informatique ouverte ou parallèle. Assurez une visibilité et une supervision centralisées sans étouffer l’innovation.
- **Fiabilité des données :** veillez à ce que les modèles soient entraînés sur des jeux de données sûrs, conformes et d’une grande intégrité. Des données de mauvaise qualité ou entachées d’erreur conduisent à des résultats inexacts, biaisés ou non fiables.
- **Autonomie et limites des risques :** à mesure que l’IA agentique gagne en capacité, définissez des garde-fous clairs pour l’autonomie. N’attendez pas pour mettre en place la gouvernance que les agents commencent à prendre des décisions à fort impact.
- **Gestion du cycle de vie des modèles :** traitez les modèles d’IA comme du code, avec un contrôle des versions, une analyse des vulnérabilités, des contrôles d’accès et des journaux d’audit.
- **Préparation culturelle :** la sécurité n’est pas seulement technique, elle est aussi comportementale. Sensibilisez les employés et les cadres aux risques de l’IA, à son utilisation en toute sécurité et à l’évolution du paysage réglementaire.

L’avenir de la sécurité de l’IA sera défini non seulement par la façon dont les entreprises se protègent contre les menaces d’aujourd’hui, mais aussi par la façon dont elles se préparent de manière réfléchie à ce qui va suivre.

[4] Gartner 2024 https://www.gartner.com/en/newsroom/press-releases/2024-10-21-gartner-identifies-the-top-10-strategic-technology-trends-for-2025

Prédiction

Le cabinet d’analyse Gartner prévoit que, d’ici 2028, au moins 15 % des décisions commerciales quotidiennes seront prises de manière autonome par l’IA agentique, contre 0 % en 2024.

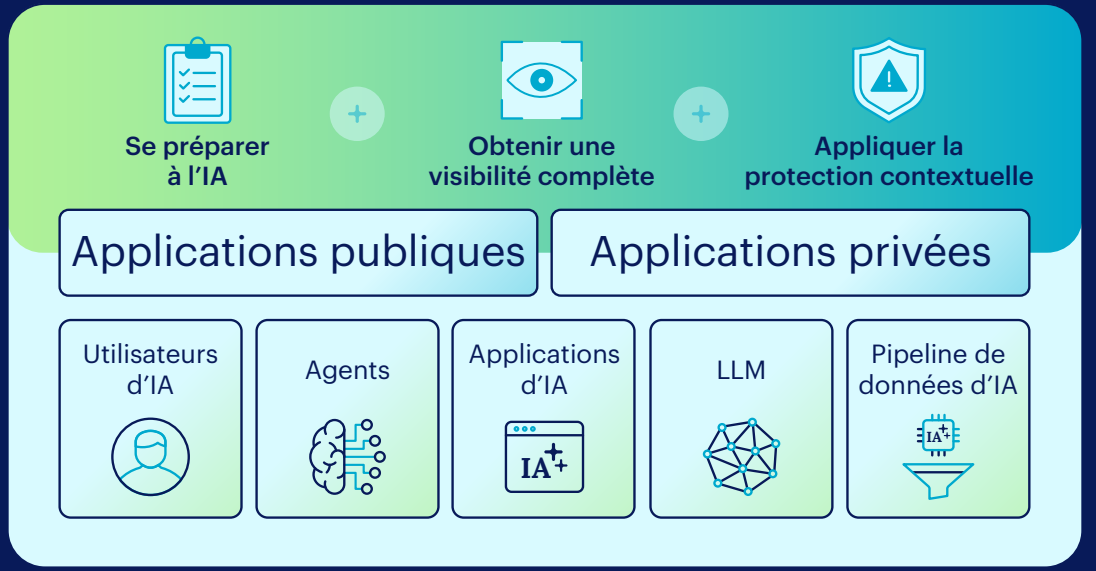
Conclusion

Sécurisez l'IA de bout en bout, partout, avec Netskope One

Alors que les entreprises s'empressent d'adopter l'IA, les responsables de la sécurité sont confrontés à une pression croissante pour protéger les données sensibles et garder une longueur d'avance sur les nouveaux risques ciblant leur écosystème d'IA. Le manque de visibilité sur l'utilisation de l'IA, l'exposition des données et les besoins de conformité sont autant de défis que les équipes de sécurité doivent relever pour pouvoir utiliser l'IA en toute sécurité au sein de l'entreprise :

- Manque de visibilité
- Comprendre les risques liés aux applications d'IA
- Intégrité des modèles d'IA
- Menaces ciblant les systèmes d'IA
- Exposition des données
- Gouvernance, conformité et utilisation éthique

Netskope One AI Security offre une solution unique pour gouverner votre écosystème d'IA et protéger vos données. Cet outil sécurise les utilisateurs et les agents automatisés dans les applications SaaS publiques, les outils d'IA privés et les flux de travail agentiques. En combinant haute performance et contrôles Zero Trust contextuels, Netskope permet aux organisations de tirer parti de l'IA en toute sécurité.



Recherche

Le cabinet d'analyse Forrester a constaté que Netskope permet de réduire de 80 % le risque d'une violation grave causée par une attaque externe, ce qui équivaut à une économie de 2 millions de dollars en coûts matériels annualisés liés à une violation.[5]

[5] Rapport de Forrester : The Total Economic Impact™ of Netskope SSE
https://www.netskope.com/resources/analyst-reports/forrester-the-total-economic-impact-of-netskope-sse

À propos de Netskope

Netskope (NASDAQ : NTSK), leader dans le domaine de la sécurité et des réseaux modernes à l'ère du cloud et de l'IA, répond aux besoins des équipes chargées de la sécurité et des réseaux en offrant un accès optimisé et une sécurité en temps réel, basée sur le contexte, pour l'écosystème de l'IA. Sa solution couvre en outre l'ensemble des composants : agents, applications, outils, grands modèles de langage (LLM), employés, appareils et données. Des milliers de clients, notamment plus de 30 entreprises du Fortune 100, font confiance à la plateforme Netskope One, son moteur Zero Trust et son puissant réseau NewEdge pour réduire les risques et obtenir une visibilité et un contrôle absolus sur les applications cloud, SaaS, web et privées. Cet outil permet d'assurer la sécurité tout en optimisant les performances, sans aucun compromis.

Vous souhaitez en savoir plus ?

Demander une démo

Ressources associées

- Sécuriser l'IA avec Netskope One
- Blog Maîtriser l'adoption de l'IA avec une sécurité de bout en bout, partout
- Netskope Threat Labs : Rapport sur les menaces liées à l'IA générative dans le cloud
- Sécuriser l'IA générative pour les nuls