Sécuriser l’adoption de l’IA générative avec le Zero Trust :

Sécuriser l’utilisation des applications publiques d’IA générative

LIVRE BLANC

Sommaire

Introduction

L'IA générative transforme le fonctionnement des gouvernements. Elle leur permet d'améliorer leur productivité, de simplifier leurs processus et de mieux servir leurs administrés. Toutefois, pour exploiter ce potentiel tout en limitant les risques, les agences doivent appliquer les principes du Zero Trust. Ce paradigme garantit qu'aucune entité (humaine ou machine) n'est fiable par défaut. Il impose une visibilité continue et des vérifications rigoureuses à chaque interaction.

Ce livre blanc est le premier de la série « Sécuriser l'adoption de l'IA générative avec le Zero Trust », une stratégie complète conçue pour aider les agences gouvernementales à exploiter l'IA générative en toute sécurité. La série comprend trois phases :

- La phase 1, décrite dans ce document, traite de la sécurisation des applications publiques d'IA générative. Elle aborde notamment les risques de fuite de données et l'usage non autorisé ou non validé de l'IA (l'IA fantôme).

- La phase 2 portera sur l'adoption, en toute sécurité, d'outils d'IA agentique afin de renforcer la productivité des employés.

- La phase 3 abordera le déploiement de systèmes d'IA générative pour les services aux citoyens, garantissant la protection des données et des infrastructures gouvernementales.

Chaque phase repose sur une approche proactive et multicouche, qui concilie innovation avec gouvernance et sécurité.

Sécuriser l’utilisation de l’IA générative publique

Synthèse

Les gouvernements prennent de plus en plus conscience du potentiel transformateur de l’IA générative, tant pour leurs opérations que pour les services aux citoyens. Cette technologie permet des gains de productivité importants et fait évoluer les services publics à travers de nombreux cas d’usage. Cela inclut notamment analyser l’opinion publique, fournir des chatbots d’assistance aux citoyens et aux équipes IT, faciliter la traduction, automatiser des processus internes comme la rédaction de fiches de poste, la synthèse de réunions ou la diffusion d’annonces publiques.

Les premiers utilisateurs dans l’administration observent déjà une amélioration de l’expérience et de la satisfaction des employés. L’émergence de grands modèles de langage (LLM) publics, comme ChatGPT, a stimulé l’expérimentation dans tout le secteur public et les administrations cherchent à comprendre et à tirer parti de ces nouvelles capacités. Cet intérêt généralisé souligne les opportunités d’améliorer les performances et les services grâce à l’intégration de ces outils d’IA avancés.

Mais l’intégration de l’IA générative, en particulier via des LLM publics ou des modèles tiers, soulève de sérieux défis de sécurité. L’utilisation non autorisée de ces outils (IA fantôme) peut exposer des données sensibles de citoyens, des informations professionnelles ou de la propriété intellectuelle. Le risque s’accroît encore dans les flux de travail intégrant la génération augmentée par récupération (Retrieval Augmented Generation ou RAG), le Model Content Protocol (MCP) ou des agents d’IA. Des acteurs malveillants ou soutenus par des États pourraient exploiter ces failles à des fins d’espionnage, de sabotage ou pour perturber des infrastructures critiques. L’IA générative expose aussi une très large surface d’attaque. Les mesures de sécurité classiques, souvent limitées à des contrôles binaires ou dépourvues de visibilité globale sur les environnements, ne suffisent pas à la gérer de manière efficace.

Pour en tirer pleinement parti, les agences doivent adopter une approche Zero Trust qui allie haut niveau de sécurité, visibilité totale et simplicité pour les utilisateurs. Les étapes qui suivent décrivent comment exploiter l’IA générative tout en réduisant les risques de fuite de données et en évitant une surcharge pour les équipes de sécurité.

1. Établir des cadres et des politiques de gouvernance de l’IA
2. Intégrer étroitement l’expérience utilisateur et la formation
3. Choisir la bonne architecture et privilégier la sécurité
4. Déployer la protection des données dès le départ
5. Adopter une approche multicouche de la protection

Examinons ces étapes plus en détail.

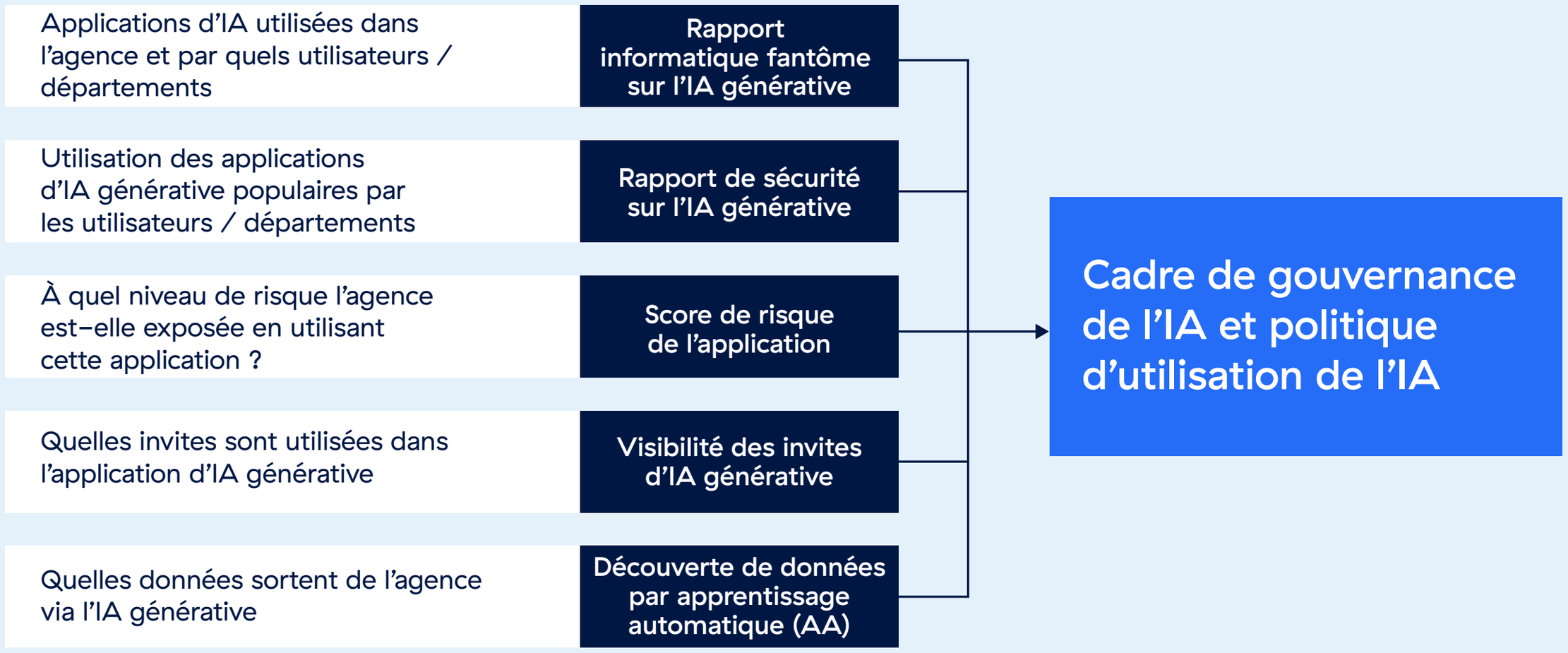
1. Établir des cadres et des politiques de gouvernance de l'IA

Pour tirer pleinement parti de l'IA générative, les agences doivent mettre en place des mesures de sécurité robustes, capables de traiter les risques sans brider la productivité des utilisateurs. Cette section explique comment appliquer une approche Zero Trust aux applications d'IA générative tout en veillant à ce que les contrôles de sécurité n'altèrent pas l'expérience des utilisateurs.

Élaborer des cadres et des politiques de gouvernance de l'IA est essentiel pour assurer une adoption sécurisée de ces technologies au sein des agences publiques. Cela passe souvent par la création d'un groupe de travail ou d'un organe de gouvernance chargé de définir et de mettre en œuvre les politiques. L'Alabama GenAI Task Force et son approche collaborative et interfonctionnelle illustre parfaitement cette démarche. Les agences doivent également s'appuyer sur des cadres Zero Trust reconnus, comme le modèle de maturité Zero Trust de la CISA ou la norme NIST 800-207. Elles peuvent les compléter par des cadres de sécurité propres à l'IA, tels que le cadre de gestion des risques de l'IA (AI Risk Management Framework ou AI RMF) du NIST, qui met l'accent sur la gouvernance, la cartographie, la mesure et la gestion, ou encore le TRISM de Gartner. En combinant un groupe de travail dédié et ces cadres éprouvés, les agences peuvent accélérer l'intégration sécurisée de l'IA générative dans l'ensemble de leurs services.

Pour accompagner cette démarche, Zscaler fournit des analyses qui permettent de suivre l'usage de l'IA dans les environnements de l'agence, d'évaluer les risques liés aux applications d'IA générative et de détecter les fuites de données. Grâce aux rapports de Zscaler, les agences accèdent à des données essentielles sur l'utilisation en cours des outils d'IA générative.

Données fournies par Zscaler pour appuyer la création d'un cadre de gouvernance de l'IA et d'une politique d'usage



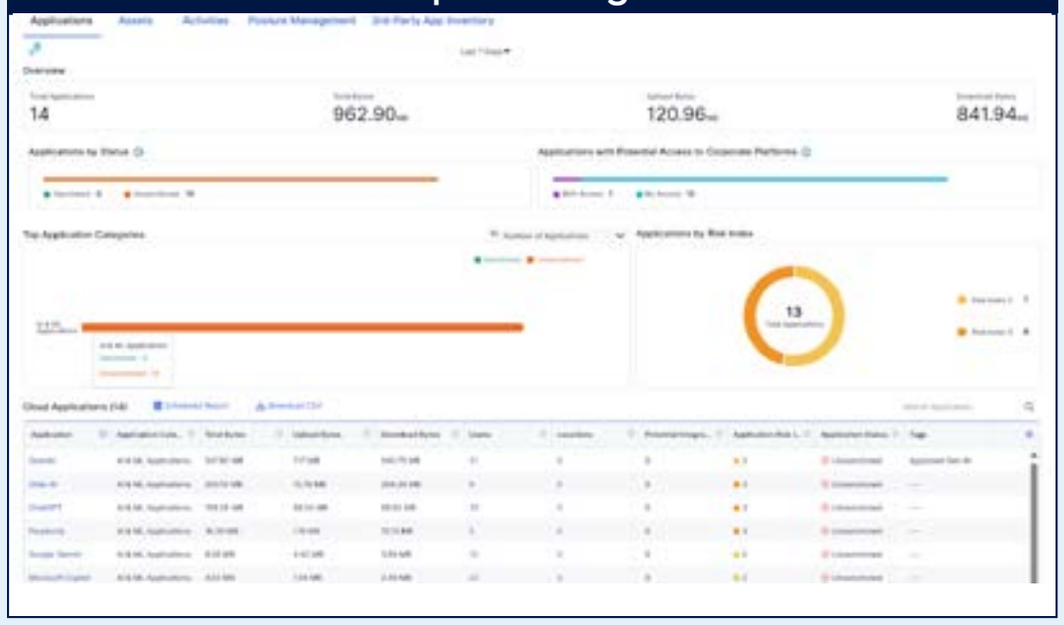
Comprendre l’utilisation actuelle de l’IA

Comprendre l’usage actuel de l’IA constitue une étape clé dans l’élaboration de cadres de gouvernance. En analysant les applications d’IA générative déployées, leurs usages et les risques associés, les agences peuvent déterminer les domaines où appliquer les politiques en priorité. Cette approche fondée sur les données garantit que le cadre reste pertinent, opérationnel et adapté aux besoins spécifiques de chaque agence.

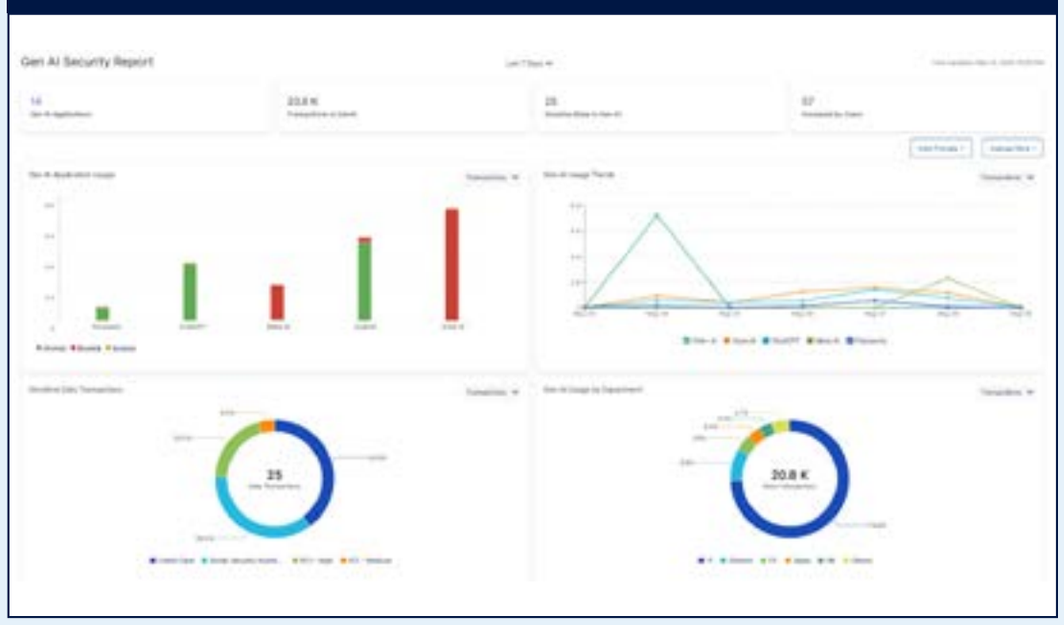
Zscaler fournit des rapports détaillés qui précisent les applications d’IA générative utilisées dans les agences et l’étendue de leur utilisation. Ceux-ci peuvent être affinés afin de spécifier les usages propres à certains départements ou sous-agences, et donner aux administrations une vision plus claire de leur usage de l’IA.

Données sur l’usage de l’IA fantôme

Détection des applications d’informatique fantôme pour l’IA générative

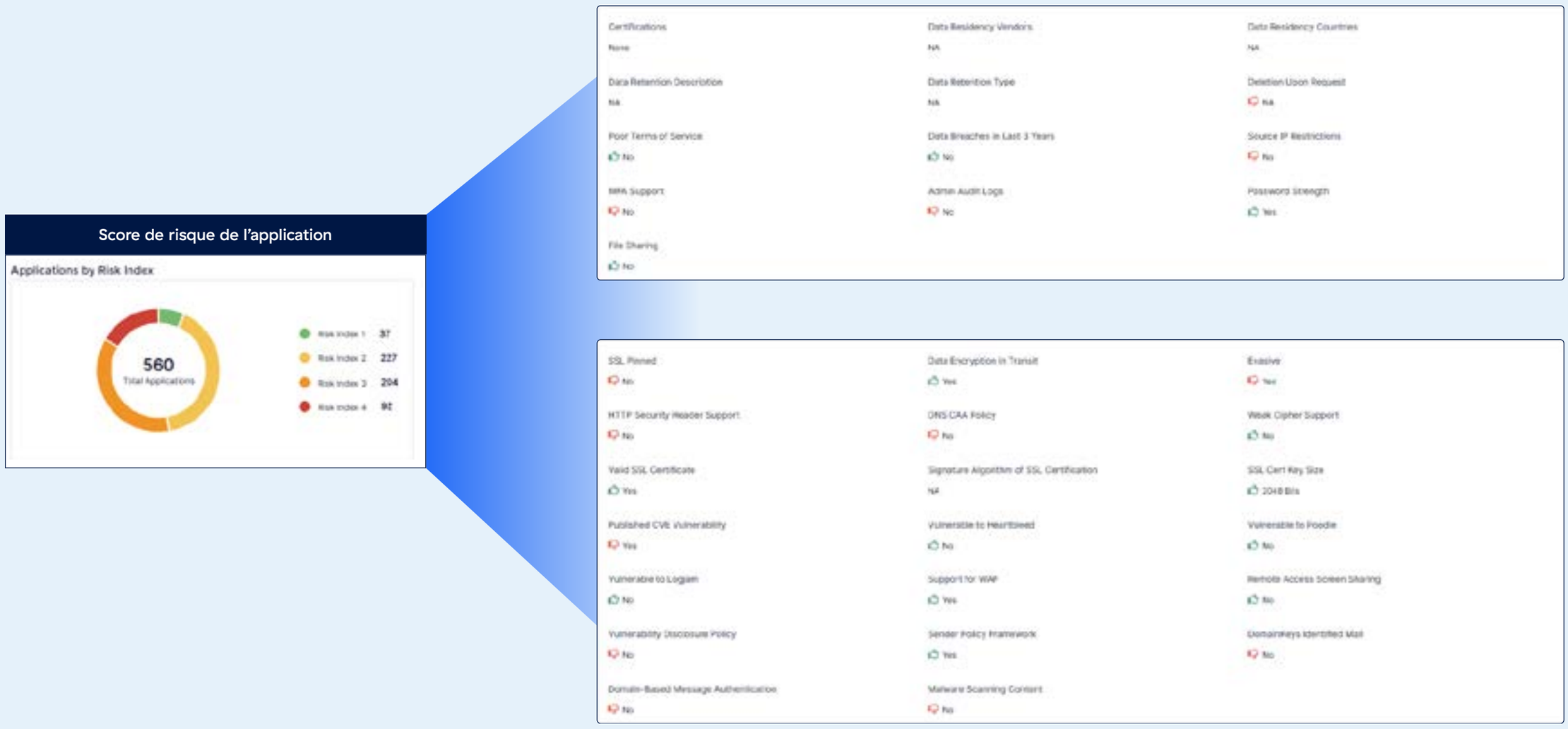


Rapport de sécurité sur l’IA générative



Les agences peuvent alors examiner plus en détail les risques liés à ces applications. L’équipe ThreatLabz de Zscaler, en collaboration avec des services externes de veille sur les menaces, évalue ces risques et leur attribue des scores agrégés de 1 à 5, afin de faciliter la prise de décision. Les agences peuvent également ajuster ces scores selon leurs priorités et leurs exigences propres. Les évaluations peuvent couvrir des facteurs essentiels comme les vulnérabilités de sécurité ou la conformité réglementaire. Cela permet aux responsables de concentrer leurs ressources sur les enjeux les plus critiques pour leur mission et leur sécurité. Le rapport ci-dessous illustre certains de ces facteurs de risque, par exemple, diverses vulnérabilités de sécurité ou le non-respect de la réglementation, qui aident les décideurs à définir leurs priorités.

Risques associés à l’utilisation de l’IA fantôme

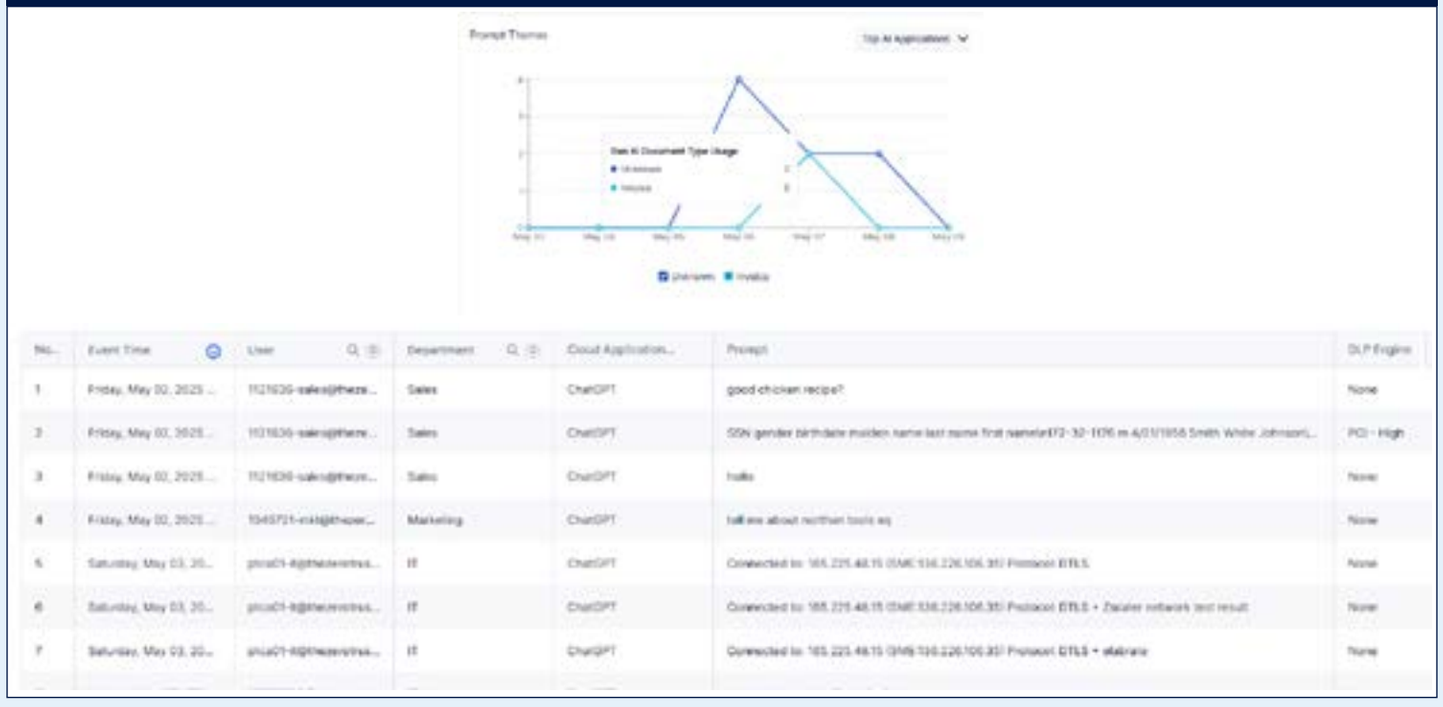


Analyse détaillée des interactions des utilisateurs avec les applications d’IA générative

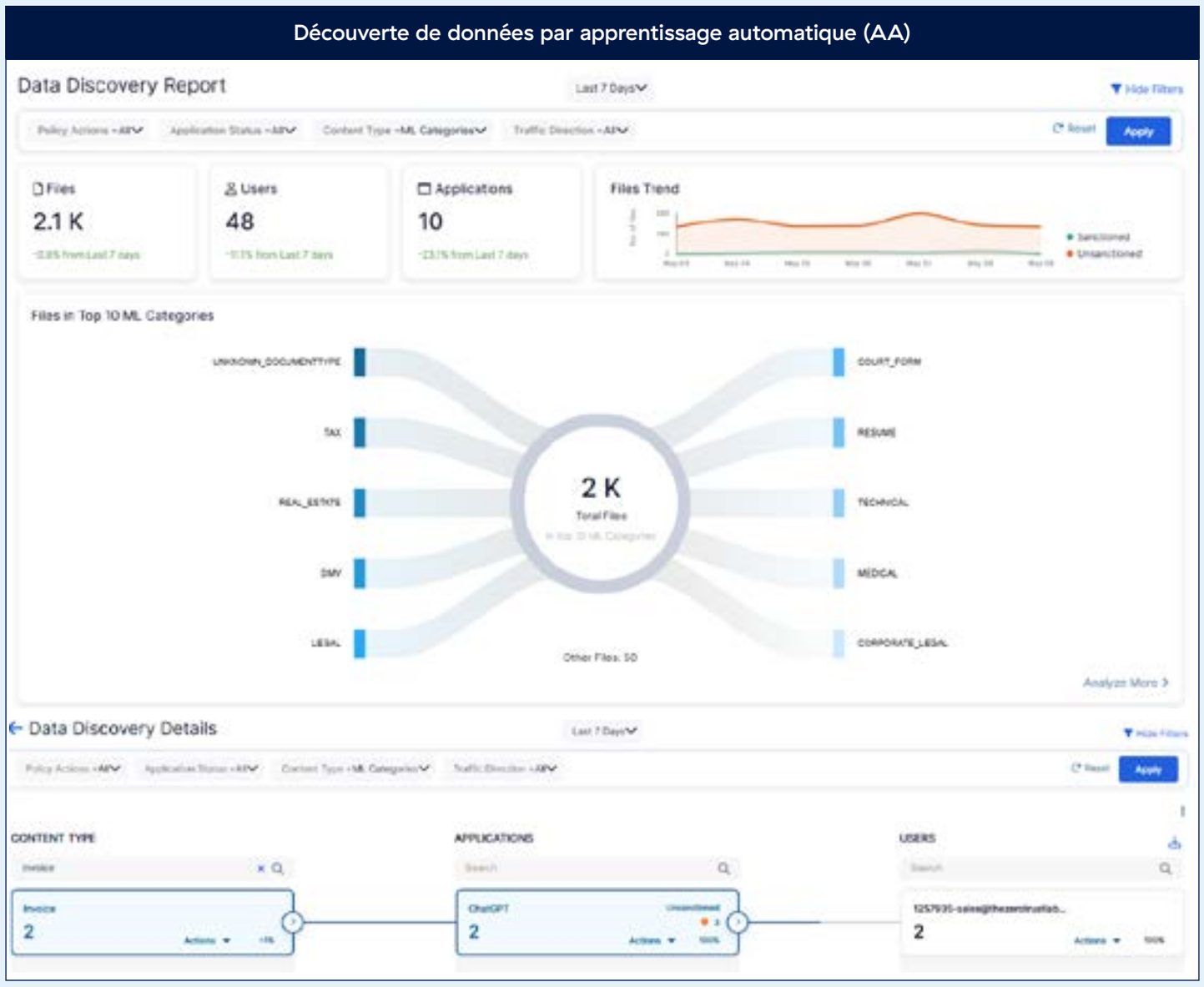
Zscaler va plus loin que l’analyse des applications, il fournit également une analyse fine de chaque transaction, de chaque requête et de chaque interaction utilisateur au sein des applications d’IA générative. Cela inclut des données détaillées sur ce que les utilisateurs partagent, non seulement via des transferts de fichiers, mais également via le clavier, le presse-papiers et d’autres canaux. Ces informations sont précieuses pour les agences, car elles permettent de mieux comprendre la nature des données partagées, d’affiner les politiques de sécurité et de garantir la conformité aux règles de gouvernance.
Ce niveau de visibilité est également indispensable pour les audits. Il peut être exporté sans difficulté vers le SIEM de l’agence afin d’assurer un suivi et une analyse complets.

Rapport sur les fuites de données en dehors de l’Agence

Visibilité des invites d’IA générative



Découverte de données par apprentissage automatique (AA)



Visibilité sur les données inconnues

Zscaler renforce encore cette visibilité en détectant les données qui fuitent via les applications d’IA générative, parfois à l’insu des agences. Grâce à ses fonctions optimisées par l’IA/AA, le rapport ML Discovery de Zscaler va au-delà des règles DLP classiques de simple surveillance. Il détecte et classe de manière proactive les données sensibles partagées avec les outils publics d’IA générative. Cela permet aux responsables des données et aux administrateurs de sécurité d’identifier les fuites non reconnues et les corriger avant qu’elles ne deviennent critiques.

Cette connaissance approfondie des données permet aux agences d'identifier en amont les informations à haut risque susceptibles d'être exposées à des LLM publics. Elle contribue également à établir ou affiner la propriété des informations sensibles, élaborer des politiques d'utilisation et appliquer des directives personnalisées afin de protéger les données stratégiques.

En croisant les informations sur les utilisateurs, les applications, les risques associés, les requêtes et les modèles de données, Zscaler contribue à élaborer des politiques et des procédures alignées sur les objectifs de l'administration. Ces informations orientent l'allocation des ressources et aident à définir les rôles et les responsabilités au sein du cadre de gouvernance Zero Trust. Elles permettent aux agences d'adopter une approche prospective qui concilie innovation et stratégie globale de gestion des risques.

2. Intégrer étroitement l'expérience utilisateur et la formation

L'expérience utilisateur et la formation jouent un rôle central dans l'adoption sécurisée de l'IA générative au sein des agences gouvernementales. Pour garantir une adoption fluide, il est essentiel que les mesures de sécurité et les programmes de formation maintiennent la productivité des utilisateurs tout en assurant un haut niveau de protection. Il convient donc d'éviter d'ajouter de nouveaux outils ou applications qui alourdiraient la charge d'apprentissage. De plus, les contrôles de sécurité doivent s'accompagner d'une sensibilisation continue des utilisateurs afin d'en maximiser l'efficacité. Les plateformes doivent s'intégrer naturellement aux flux de travail et aux canaux existants, tout en intégrant des mécanismes d'interaction et de rétroaction des utilisateurs. Cette approche aide les agences à s'aligner dès le départ sur des cadres tels que le AI Risk Management Framework (AI RMF) du NIST.

Fonctionnalités essentielles des plateformes qui soutiennent cette approche :

Accès fluide à l'IA générative

L'objectif principal des outils d'IA générative est de libérer les utilisateurs des tâches répétitives et de leur permettre de se consacrer à des missions où le discernement humain apporte une réelle valeur. Les mesures de sécurité appliquées à l'IA générative ne doivent donc pas perturber les flux de travail. Zscaler répond à cet enjeu en éliminant le besoin de logiciels supplémentaires ou de navigateurs gérés. Par exemple,

- **Agent unique Zscaler** Le même agent Zscaler qui sécurise l'accès aux applications publiques et privées gère aussi les contrôles liés à l'IA générative. Les utilisateurs bénéficient ainsi d'un accès fluide, sans ajout d'outils supplémentaires.

- **Accès sécurisé sans agent**
  Les utilisateurs peuvent accéder aux applications d'IA générative sécurisées via leur navigateur habituel et leur flux de travail existant (par exemple, une vignette dans le portail d'applications IDP), sans avoir besoin d'un agent.

- **Contrôles de sécurité flexibles** Zscaler ne se contente pas d’autoriser ou de bloquer l’utilisation de l’IA, il propose également une isolation du navigateur basée sur le cloud. Les utilisateurs qui accèdent à des applications d’IA générative sont redirigés vers cet environnement isolé tout en conservant l’ergonomie de leur navigateur habituel. Ce modèle permet d’appliquer des mesures de sécurité avancées — blocage du presse-papiers, interdiction d’impression ou de transferts de fichiers — sans perturber l’expérience utilisateur. Le tout est géré via une plateforme unifiée et un agent unique Zscaler afin de simplifier l’administration.

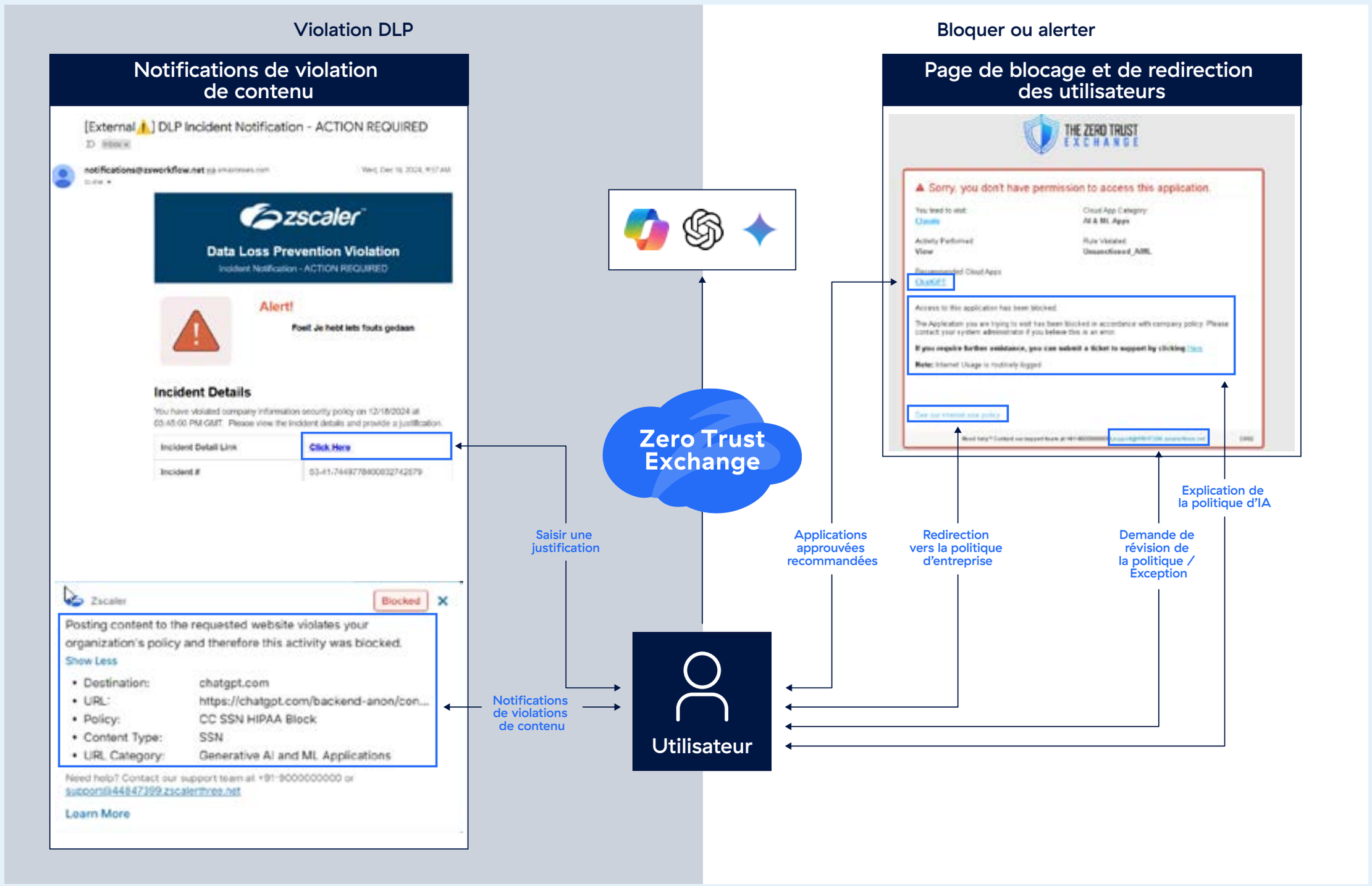
Ces contrôles peuvent être déployés avec un impact minimal sur l’infrastructure ou les postes existants. Les agences appliquent alors leurs politiques de sécurité tout en préservant une expérience utilisateur fluide et en réduisant la charge administrative.

Formation des utilisateurs et retours intégrés

Une sensibilisation continue à l’usage sécurisé de l’IA générative est essentielle, d’autant plus que la technologie évolue rapidement. La formation doit être régulière, continue et intégrée directement dans les outils et flux de travail des utilisateurs. Zscaler l’accompagne grâce à des notifications dynamiques qui alertent l’utilisateur lorsqu’une ressource est bloquée, isolée ou signalée pour contenu non conforme. Par exemple, si une application d’IA générative non approuvée est bloquée, Zscaler propose des équivalents approuvés afin d’orienter l’usage sans freiner la productivité. Si une violation de l’usage des données devait survenir, Zscaler s’intègre à des outils familiers tels que la messagerie électronique ou Slack, de sorte que les utilisateurs peuvent justifier leurs actions ou recevoir un retour adapté dans leur outil habituel.

En intégrant la formation des utilisateurs aux processus de sécurité, les agences posent une base de gouvernance solide pour l’usage de l’IA générative. Cette approche garantit non seulement que les utilisateurs savent comment interagir en toute sécurité avec la technologie, mais contribue également à créer un cadre évolutif pour gérer les incidents liés à l’IA générative et affiner les politiques d’utilisation dans l’ensemble de l’administration.

Formation et retour d’expérience des utilisateurs avec Zscaler



3. Privilégier la sécurité et choisir la bonne architecture

L'adoption galopante des outils d'IA élargit la surface d'attaque. Les administrations doivent donc s'appuyer sur une plateforme capable non seulement de les protéger contre les menaces optimisées par l'IA, mais aussi de sécuriser l'usage des applications d'IA générative, au sein d'une architecture Zero Trust et conforme aux exigences de la FedRAMP. Zscaler fournit une solution complète qui combine sécurité, simplicité et conformité pour les agences qui adoptent l'IA générative.

Ces agences peuvent déployer les fonctions de Zscaler dédiées à l'IA générative sans bouleverser leur infrastructure existante. Le trafic peut être dirigé vers la plateforme via l'application Zscaler, qui est compatible avec les principaux systèmes d'exploitation (Windows, Mac, iOS, Android et Linux), via un équipement réseau de site distant, ou simplement via un navigateur pour un déploiement sans agent. Une fois connectés, les utilisateurs sont authentifiés et autorisés avant de pouvoir accéder à toute ressource en ligne, y compris les applications d'IA générative. Agissant comme un proxy, la plateforme inspecte toutes les sessions TLS établies avec ces applications et offre aux agences une vision claire et un contrôle précis sur la manière dont les utilisateurs interagissent avec ces outils. Cette architecture offre aux agences une visibilité approfondie sur l'usage des applications d'IA générative et leur fournit des mécanismes de protection avancés de leurs données.

Inspection HTTPS et Websocket sécurisée avec Zscaler

Contrôle d'accès

- Filtrage d'URL
- Pare-feu cloud
- Contrôles granulaires des applications cloud
- Restrictions pour chaque tenant
- Sécurité DNS

Prévention des menaces

- Protection contre les malwares
- Cloud Sandbox
- Protection contre les menaces avancées (IPS)
- Isolation du navigateur cloud

Protection des données

- DLP inline
- EDM / IDM
- Contrôles de type de fichier
- Contrôles pour l'IA générative
- Isolation du navigateur cloud

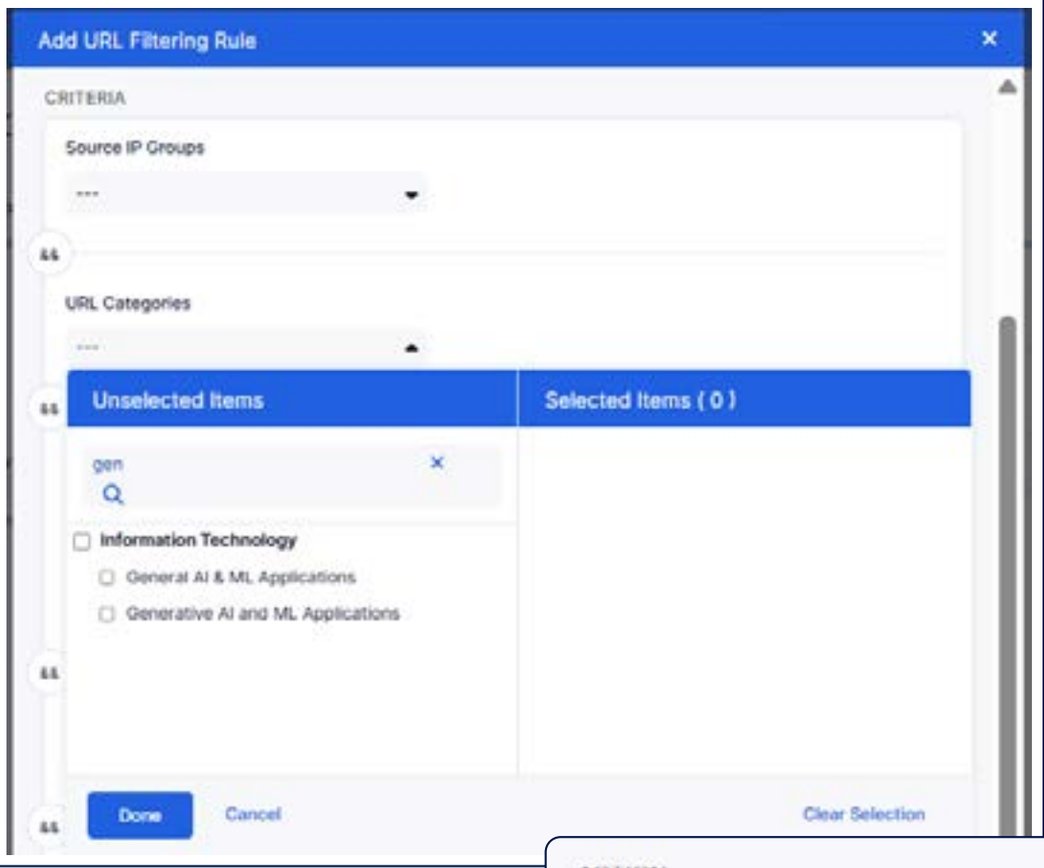
Automatiser la découverte et la gestion des applications d'IA générative

Grâce à l'inspection TLS, les agences accèdent à l'ensemble des capacités de Zscaler, notamment le contrôle granulaire des applications d'IA générative et d'apprentissage automatique. La catégorie Applications d'IA et AA gérée par l'équipe ThreatLabz, constitue un avantage essentiel. Cette catégorie englobe un large éventail d'applications d'IA, dont des outils populaires tels que ChatGPT, Gemini, MetiAI, Claude et bien d'autres.

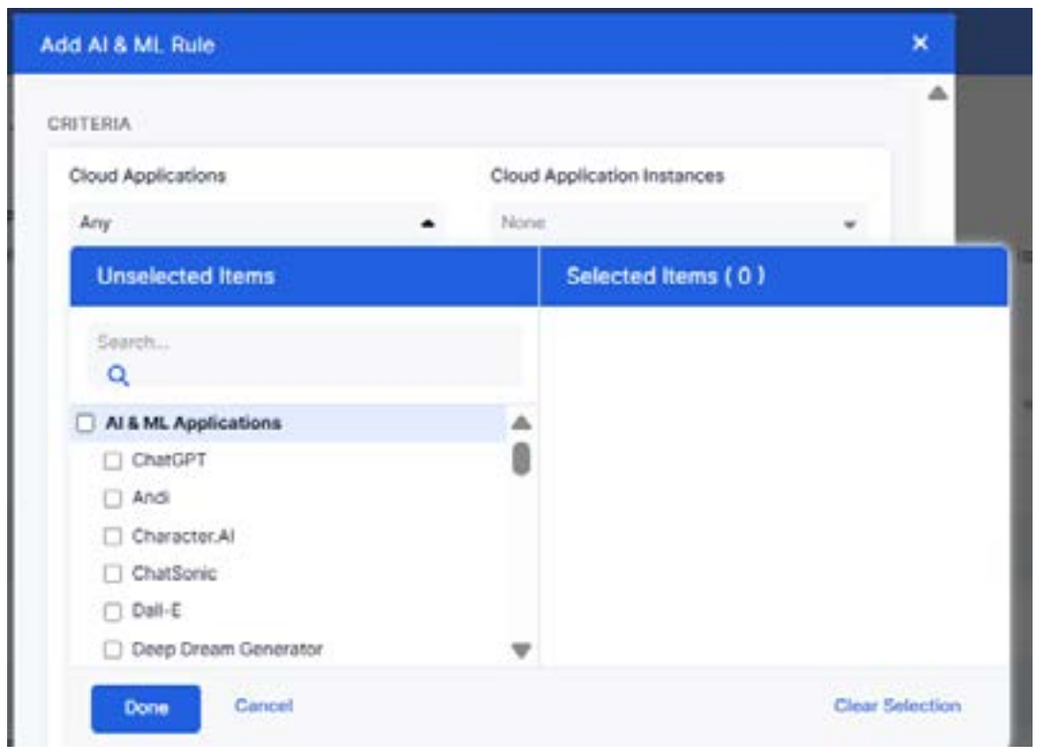
Cette classification permet d'appliquer des règles qui bloquent par défaut les applications non validées ou non reconnues, afin de réserver l'accès aux seuls outils approuvés. Les nouvelles applications sont automatiquement intégrées à ces catégories, ce qui évite aux agences d'avoir à les identifier et à mettre leurs listes à jour manuellement. Les agences gardent néanmoins la possibilité d'adapter cette classification en y ajoutant leurs propres domaines afin de mieux l'aligner sur leurs besoins spécifiques. Zscaler propose également des catégories dédiées telles que « Applications générales d'IA et d'AA » ou « Applications d'IA générative et d'AA ». Associées à la liste plus large « Applications d'IA cloud », elles assurent une couverture étendue qui réduit les risques de sécurité liés aux applications d'IA générative. Cette approche multicouche permet aux agences de gérer l'accès aux centaines d'applications développées et publiées chaque semaine.

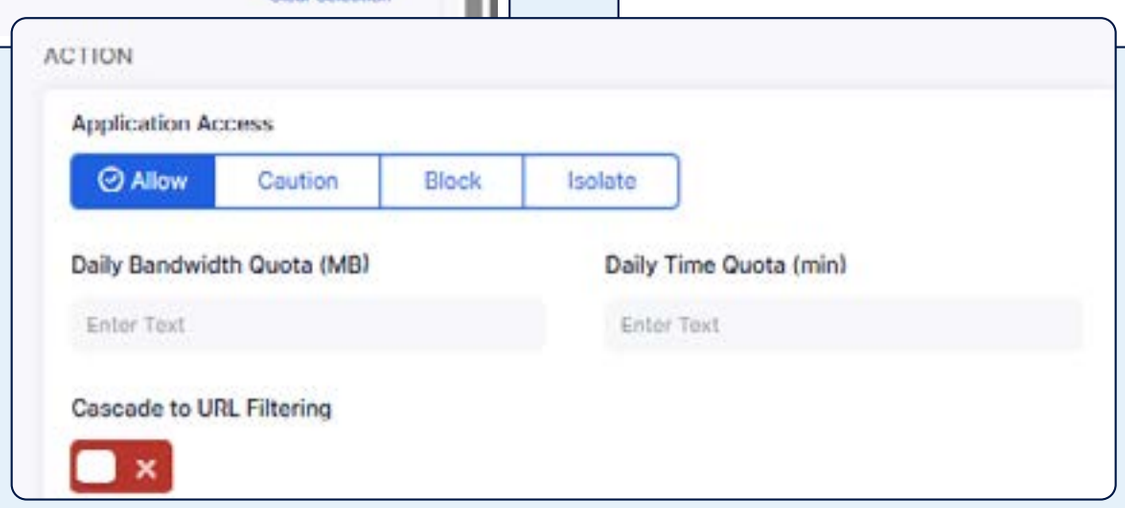
Sélection de catégories générales et d'applications d'IA spécifiques

Catégories d'URL pour Wide Net



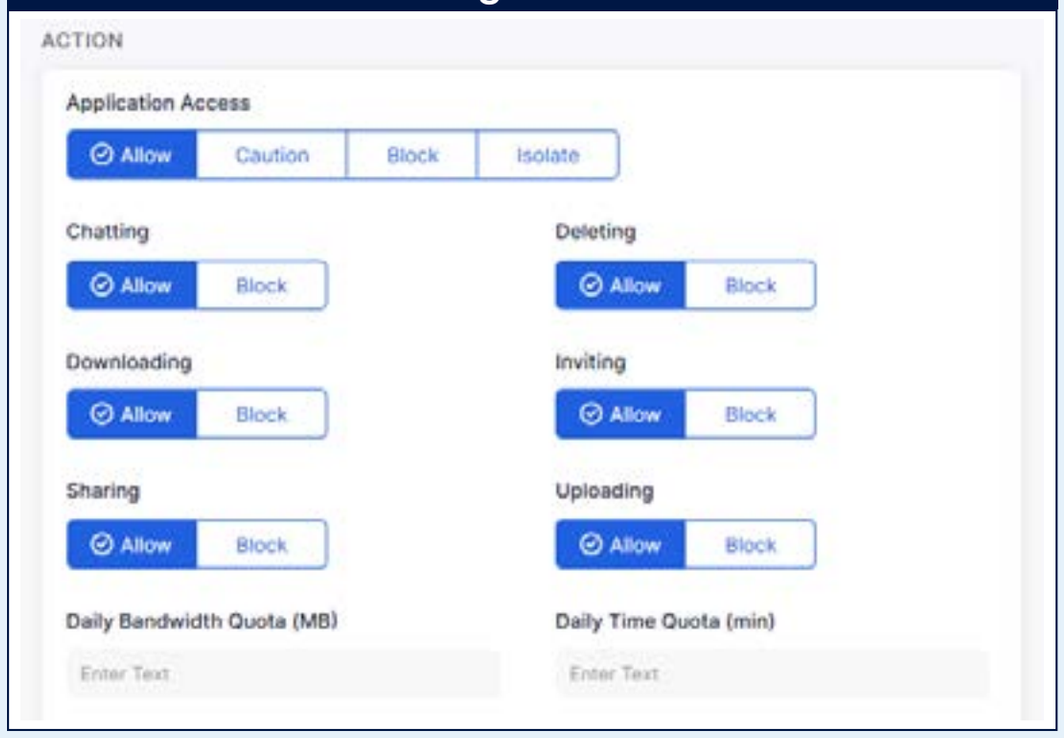
Contrôles granulaires des applications d'IA générative



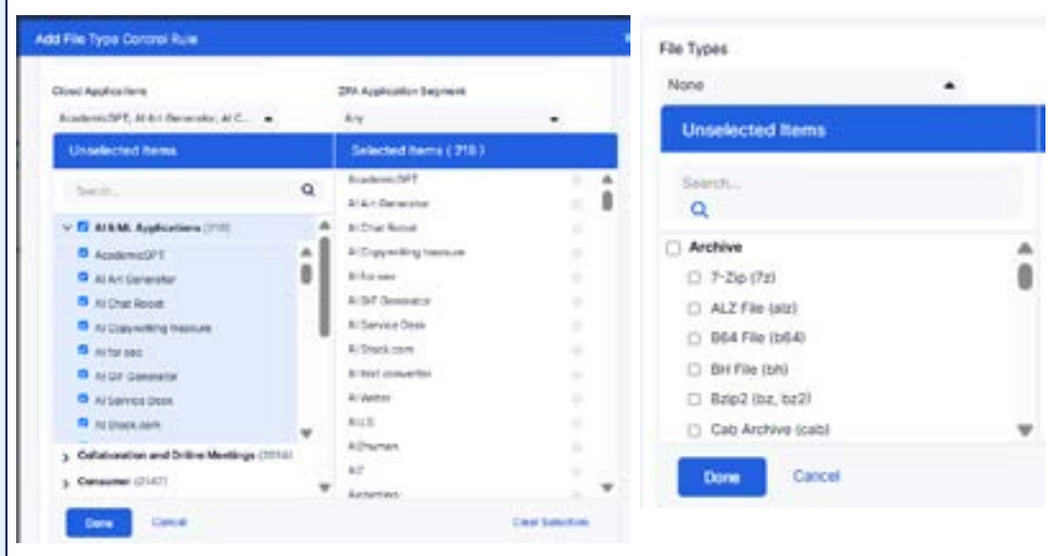


Contrôles granulaires pour les applications SaaS, Web et IA

Contrôles granulaires des applications d'IA générative



Contrôles de téléversement de fichiers pour les applications d'IA générative



Autoriser les applications validées via le contrôle de sécurité des applications SaaS

Outre le maintien d'une liste complète d'applications d'IA, Zscaler propose des contrôles précis sur la manière dont les utilisateurs exploitent l'IA générative. Ces contrôles sont incroyablement simples à appliquer, extrêmement puissants, et centralisés au sein d'une même plateforme. Le côté gauche de l'image illustre quelques exemples de contrôles granulaires, par exemple, une politique de sécurité appliquée à ChatGPT peut autoriser la messagerie tout en bloquant l'envoi de fichiers ou en limitant le partage des conversations. Les agences peuvent déployer ces règles à l'échelle d'un service entier ou les affiner au niveau de chaque utilisateur. Ces contrôles granulaires peuvent être encore affinés et restreindre les types de fichiers que les utilisateurs sont autorisés à transmettre aux applications d'IA générative, comme illustré dans l'exemple de droite. Ce filtrage permet aussi de bloquer le transfert de documents chiffrés.

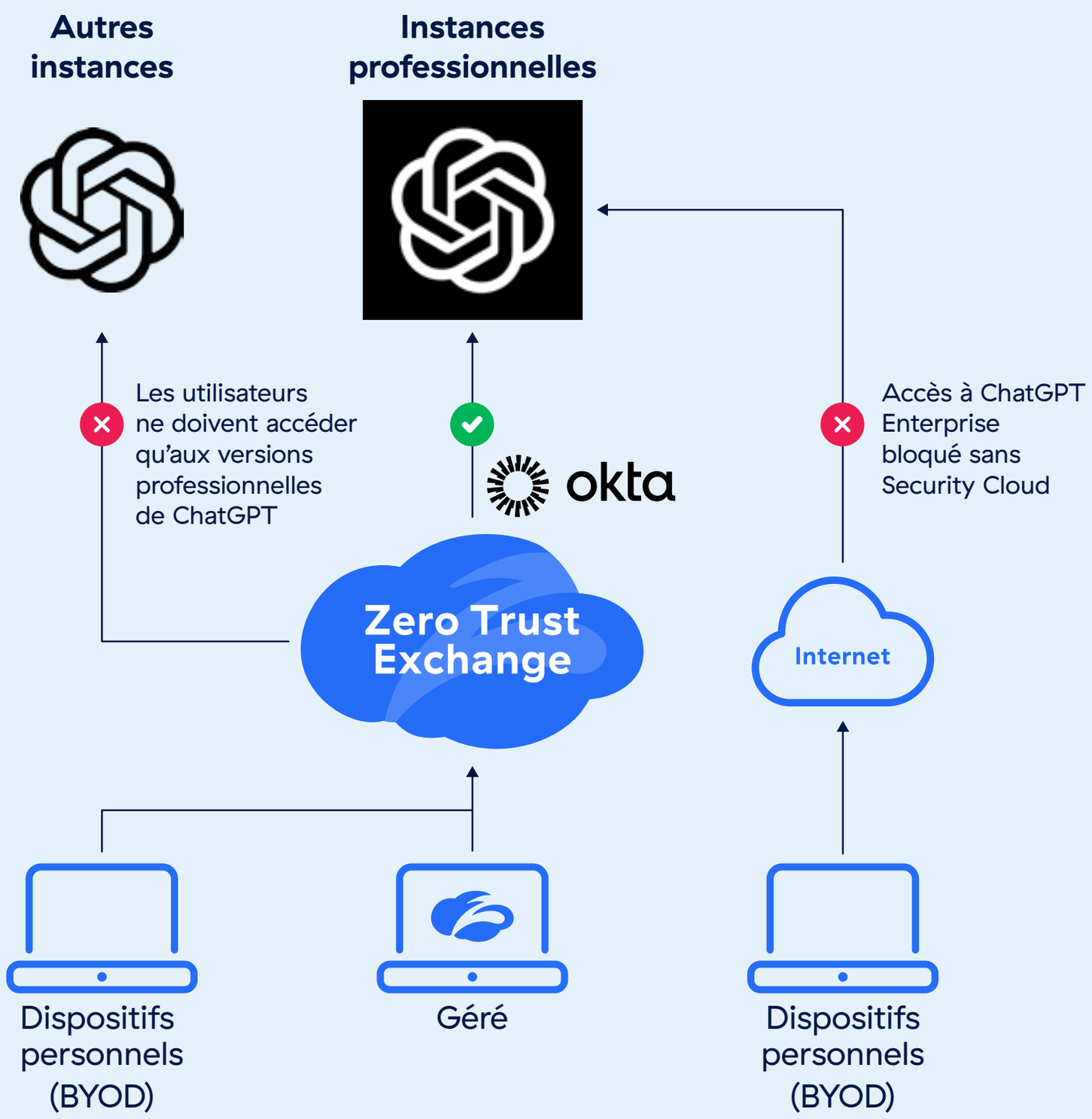
Restreindre l'accès aux versions professionnelles des applications d'IA générative

Les agences devraient privilégier les versions professionnelles des applications d'IA générative afin de renforcer la sécurité et garder le contrôle. Ces versions — par exemple ChatGPT Enterprise — donnent aux agences la maîtrise complète de leurs données et de leurs échanges, sans que les informations internes servent à entraîner les modèles. Elles sont conformes à la norme SOC2 et assurent le chiffrement des données, en transit comme au repos. Elles simplifient également la gestion des utilisateurs grâce à des fonctions telles que l'accès par équipe, la vérification de domaine, l'authentification unique (SSO) et des statistiques d'usage.

Les instances professionnelles doivent être associées au SSO pour maximiser la sécurité et offrir une meilleure visibilité et un plus grand contrôle de l'utilisation des applications. Avec le SSO, les agences peuvent définir des règles qui bloquent l'accès aux versions grand public des applications d'IA générative. Par exemple, le contrôle de locataire (tenancy control) de Zscaler pour ChatGPT veille à ce que seuls les environnements approuvés soient accessibles, les autres étant automatiquement bloqués. Les agences peuvent également appliquer des règles au niveau de la gestion des identités et des accès (IAM) à l'aide de listes autorisées (whitelists). Cela garantit que seules les versions professionnelles sont utilisées et que l'accès s'effectue dans des environnements sécurisés, comme le cloud de Zscaler. Les instances professionnelles peuvent en outre être rendues accessibles aux appareils non gérés ou personnels (BYOD) grâce à l'accès sans agent que propose Zscaler.

Une approche binaire du type « tout autoriser ou tout bloquer » ne suffit plus. Les agences doivent adopter une stratégie de sécurité multicouche, avec des contrôles précis adaptés aux différents usages des applications. Réunir toutes ces fonctions dans une plateforme unifiée permet non seulement de simplifier le déploiement, mais aussi d'appliquer plus facilement les principes fondamentaux du Zero Trust, à savoir l'accès sur base du moindre privilège, la surveillance continue et la protection globale de chaque interaction avec l'IA générative.

Contrôle d'accès aux versions validées des applications d'IA



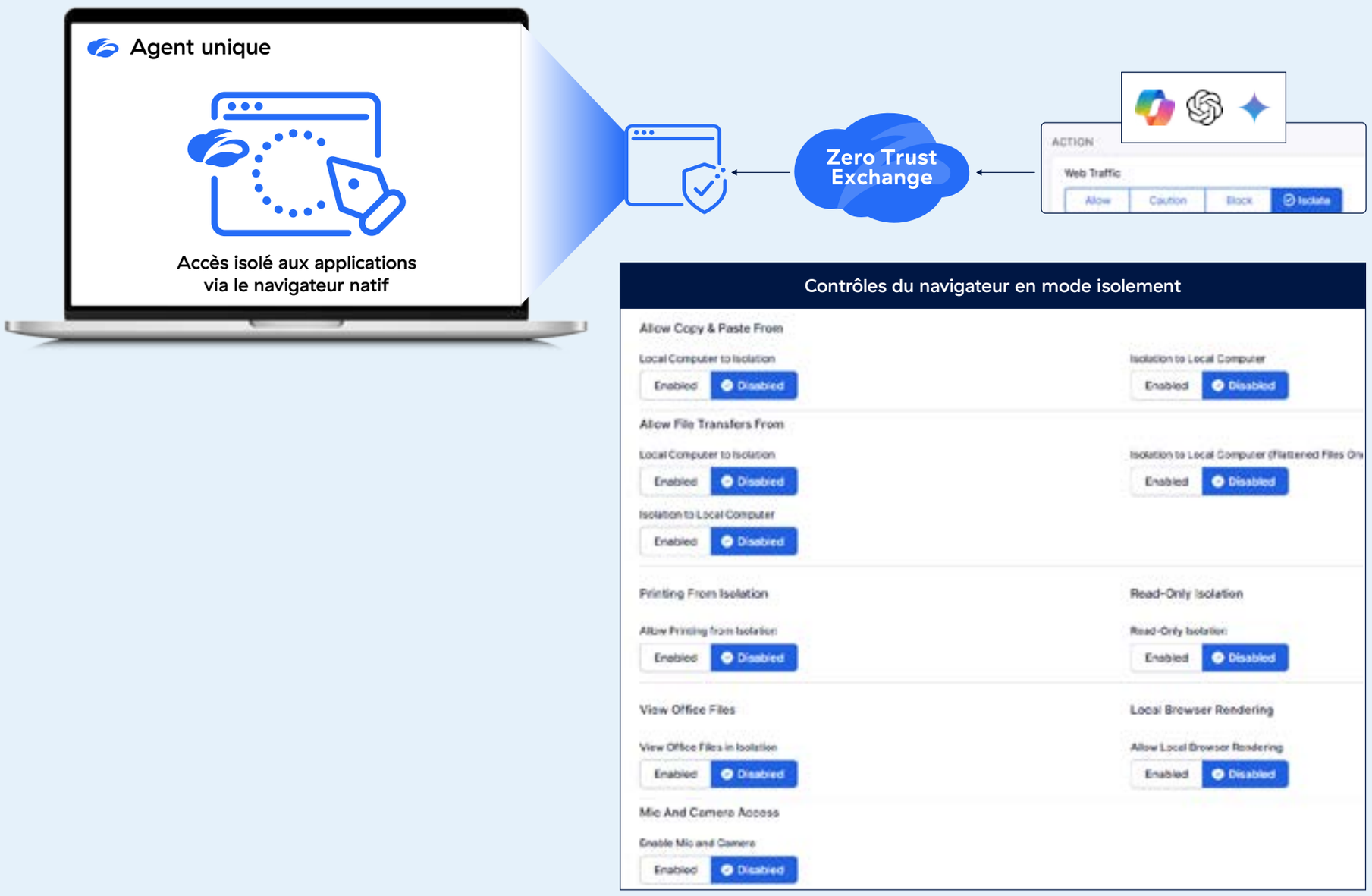
Réduire les risques liés aux applications d'IA générative non approuvées

Lorsque l'accès à des applications d'IA générative non approuvées (parce qu'elles n'ont pas de licence professionnelle ni de SSO) s'avère nécessaire, elles doivent être considérées comme à haut risque. Les données chargées dans ces applications peuvent être réutilisées pour entraîner les modèles, ce qui expose potentiellement des informations sensibles. Pour limiter ce risque, les agences doivent déployer des contrôles de sécurité supplémentaires afin de renforcer la supervision des échanges de données.

Zscaler propose une solution efficace : son navigateur Zero Trust. Cet outil permet un accès sécurisé aux applications d'IA générative non approuvées grâce à des contrôles avancés, tels que le blocage des transferts de fichiers, de l'impression ou de l'usage du presse-papiers. Il empêche aussi ces applications d'exécuter du code directement dans le navigateur de l'utilisateur : les interactions sont restituées sur des pages isolées. Cette approche protège contre le fingerprinting, le suivi par cookies tiers et d'autres vulnérabilités, tout en laissant les utilisateurs travailler avec le même navigateur déployé par l'agence.

Deux modes de déploiement sont possibles : via l'agent unifié de Zscaler, ou en mode sans agent. Sur les appareils gérés par l'agence, il est recommandé d'opter pour le déploiement avec agent, qui garantit que l'ensemble du trafic passe par la plateforme de contrôle de Zscaler. Lorsque l'installation d'un agent n'est pas possible, l'option sans agent de Zscaler constitue une alternative sécurisée, assurant un accès contrôlé aux applications d'IA générative sans compromis sur la sécurité.

Contrôles granulaires pour sécuriser les applications d'IA isolées tout en préservant l'expérience utilisateur



4. Déployer la protection des données dès le départ

L'absence de protection solide des données dès le début de l'adoption de l'IA générative peut entraîner des fuites de données, des violations des règles de confidentialité et une perte de confiance du public, qui compromettraient le succès même de ces outils. Le caractère conversationnel et convivial des applications publiques d'IA générative accroît le risque que des utilisateurs exposent involontairement des données gouvernementales sensibles. Des gestes aussi anodins que copier-coller des informations ou transférer un fichier, peuvent suffire à divulguer des détails confidentiels, par le contexte ou par des intégrations avec d'autres systèmes. C'est pourquoi la protection des données doit être au cœur de toute stratégie d'adoption de l'IA générative.

Zscaler aide les agences à relever ces défis grâce à ses capacités avancées de prévention de la perte de données (DLP). Conçue pour protéger les informations sensibles dès le départ, la solution DLP de Zscaler pour l'IA générative détecte et bloque le partage de données confidentielles — qu'il s'agisse d'une requête, d'un fichier envoyé ou d'un usage abusif — avant qu'elles n'atteignent les modèles d'IA générative publics. Cette approche proactive permet aux agences d'adopter l'IA générative tout en préservant la confidentialité et en maintenant leur conformité.

Accélérer l'adoption de la DLP

Déployer une stratégie de protection des données peut sembler complexe, surtout s'il faut concilier l'accès aux outils d'IA générative avec des garde-fous robustes. Zscaler simplifie ce défi avec une plateforme pensée pour les équipes réduites, qui facilite l'adoption de l'IA générative tout en intégrant des contrôles efficaces de protection des données. Cette approche permet d'étendre le cadre de sécurité à différents services et profils d'utilisateurs.

Les agences qui appliquent déjà des règles à d'autres destinations Internet peuvent les étendre sans difficulté aux applications d'IA générative. Qui plus est, Zscaler intègre directement les moteurs DLP et les dictionnaires déjà utilisés pour d'autres canaux dans les applications d'IA et d'AA, ce qui évite les doublons et accélère le déploiement. Pour celles qui partent de zéro, Zscaler fournit des dictionnaires prédéfinis, activables en quelques clics pour prévenir les fuites de données sensibles. Enfin, les documents ou ensembles de données connus peuvent être protégés grâce aux fonctions EDM/IDM, tandis que le balisage de Microsoft Information Protection (MIP) renforce la protection des données chiffrées ou classifiées contre toute exposition.

Pour affiner davantage les politiques, les capacités de découverte par l'AA (apprentissage automatique) de Zscaler identifient les informations sensibles et les fuites de données jusqu'alors inconnues dans les applications d'IA générative. Les agences peuvent ainsi faire évoluer en continu leur stratégie de protection. Elles peuvent affiner les dictionnaires existants ou créer des règles de détection personnalisées, par exemple avec des expressions régulières (regex) ou des mots-clés, afin d'adapter les contrôles à leurs besoins. Zscaler s'intègre également aux solutions de sauvegarde de données comme Rubrik, ce qui facilite l'identification et la protection des informations sensibles.

Accélérer la mise en œuvre de la DLP avec Zscaler

Jour O de la mise en œuvre

Données spécifiques à l'agence avec EDM et IDM

Dictionnaires prédéfinis à utiliser par les agences gouvernementales

- ABA Bank Routing Numbers
- Document financier d'entreprise
- Document juridique d'entreprise
- Document judiciaire
- Identifiants et secrets
- Cartes de crédit
- Informations sur les maladies
- Permis de conduire (US)
- Informations sur les médicaments
- États financiers
- Document d'immigration
- Document d'assurance
- Facture / Document de facturation
- Document juridique
- Document médical
- Informations médicales
- Document immobilier
- Numéros de sécurité sociale (US)
- Document fiscal
- Numéro d'identification fiscale (US)
- Documents du Department of Motor Vehicles (US)
- Informations sur les traitements

MIP / Étiquettes AP

Surveillance et visibilité continues

Identifier les fuites de données et les applications inconnues

Données issues des incidents

Saisie et retours des utilisateurs

Affiner et ajuster | Selon les besoins

Créer un dictionnaire personnalisé Regex / Mot-clé

Mots-clés simples ou composés avec gestion de proximité

Étendre EDM + IDM aux solutions de sauvegarde de données

L'application des règles en temps réel et une visibilité précise permettent aux équipes IT de sécuriser les données sensibles sans complexité supplémentaire ni supervision manuelle. Cette approche simplifiée facilitent l'adoption des outils d'IA générative. Les agences en retirent des avantages de productivité, tout en respectant la conformité et en maintenant la confiance du public, conformément au principe du Zero Trust : « Ne jamais faire confiance, toujours vérifier ».

Simplifier la gouvernance de la DLP

Le volume d'incidents que doivent traiter les équipes SOC et les responsables de données est l'un des principaux défis de la DLP, en particulier dans les grandes agences ou les services partagés. Selon les cas, le traitement de ces incidents peut aller d'un suivi des employés pour justification au renforcement de la formation, en passant par la gestion des exceptions ou la tenue d'un registre d'audit. Faute d'un système efficace, cette charge devient vite difficile à gérer.

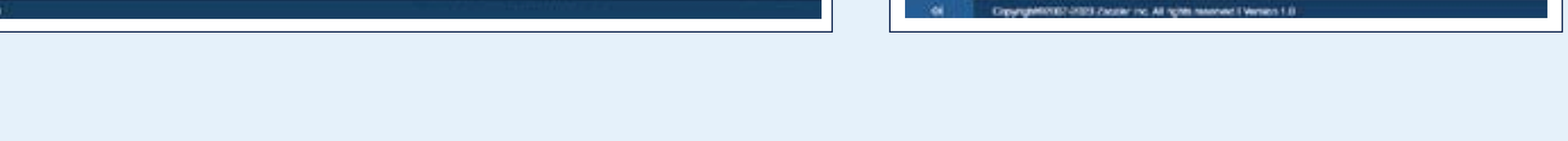
Workflow Automation simplifie ce processus grâce à une solution centralisée de gestion des incidents liés à l'IA générative. Il regroupe tous les incidents en un seul endroit, avec les métadonnées et les détails des actions ou données spécifiques qui ont déclenché la violation. Cette centralisation permet aux administrateurs d'examiner, de hiérarchiser et de corriger rapidement les incidents.

Workflow Automation peut regrouper les incidents par caractéristiques communes et leur attribuer un niveau de priorité. Ces groupes peuvent ensuite être attribués à des administrateurs spécifiques pour une résolution ciblée. L'automatisation joue un rôle essentiel, elle permet de créer des workflows qui notifient ou forment les utilisateurs impliqués dans un incident, exigent une justification, ou transmettent le dossier à un responsable ou à un propriétaire de données pour approbation. Ces workflows peuvent aussi déclencher directement des actions correctives, sans intervention manuelle.

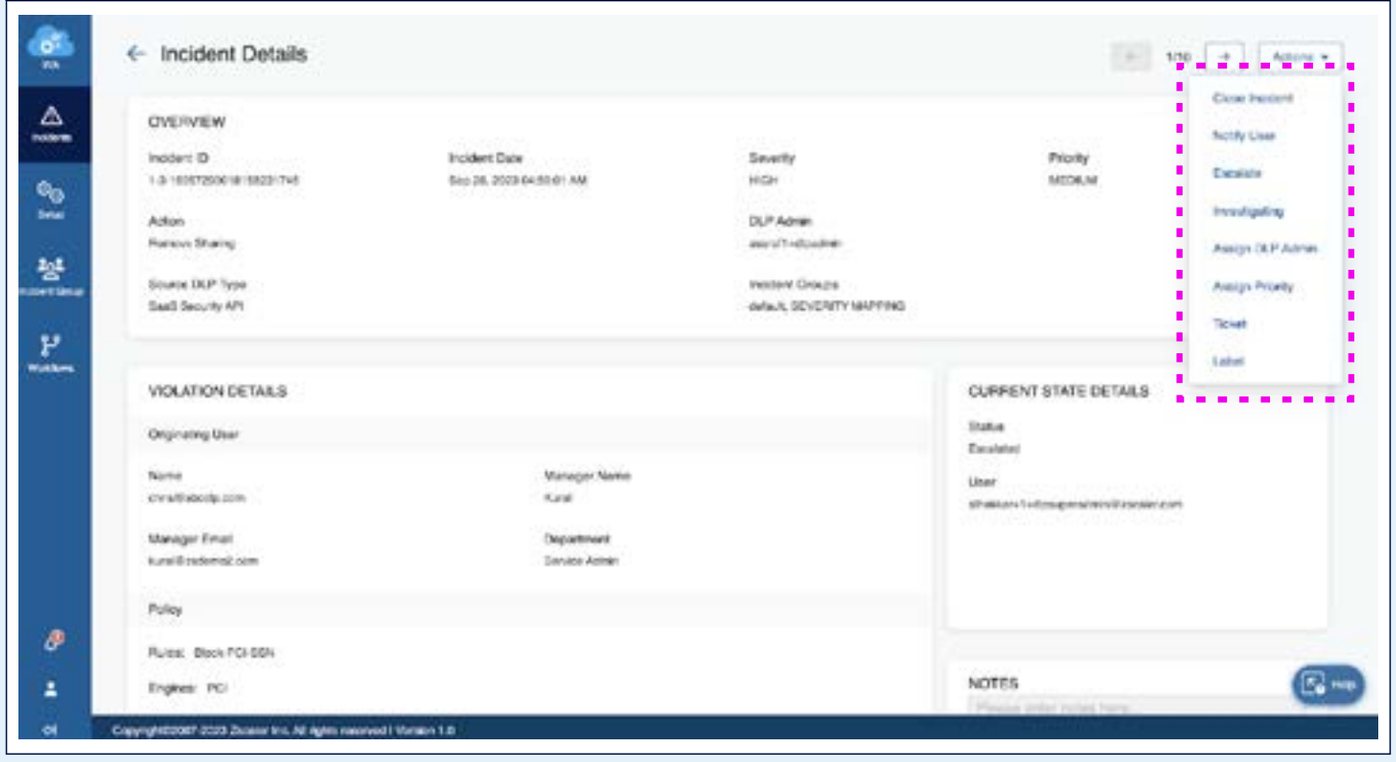
En s'appuyant sur Workflow Automation dans la DLP, les agences réduisent fortement les délais de résolution, allègent la charge des équipes SOC et obtiennent des informations exploitables sur les zones de risque. Ces informations servent ensuite à affiner les politiques de sécurité ou à améliorer les formations, afin que les utilisateurs puissent travailler en toute sécurité et de réduire le risque d'incidents futurs.

Optimiser la gestion des incidents grâce au case management et au coaching des utilisateurs

Workflow Automation
Justifier les incidents tout en formant les utilisateurs



Gérer, attribuer et faire remonter les incidents



Définir des routines pour les flux de travail

5. Rassembler l'ensemble et adopter une approche multicouche

Les administrations nationales et locales adoptent l'IA générative pour gagner en efficacité et améliorer leurs services, mais cette adoption doit se faire de manière sécurisée. Face aux milliers d'outils disponibles et à des risques tels que la fuite de données ou l'usage non approuvé, les agences doivent disposer d'une stratégie claire qui place la sécurité au premier plan, intègre les principes du Zero Trust et préserve la productivité. Une approche multicouche facilite cette démarche et réduit la pression sur les équipes informatiques : elle classe les applications par niveau de risque, applique des contrôles adaptés et automatise la gestion des incidents, Cette stratégie aide les agences à protéger leurs données sensibles, à fluidifier leurs opérations et offre aux utilisateurs un accès sûr aux applications d'IA générative, dans un cadre à la fois évolutif et maîtrisable.

Mettre en place des contrôles multicouche

Nous verrons dans cette section comment les agences peuvent sécuriser l'adoption de l'IA générative en combinant les différents volets dans une approche multicouche. Le marché compte déjà des milliers d'outils d'IA générative, et de nouveaux sont lancés chaque semaine. Sans stratégie claire, la gestion des règles et des incidents devient vite incontrôlable.

L’approche multicouche simplifie cette tâche : elle organise les accès et applique des contrôles adaptés au niveau de risque. Elle allège le travail des responsables de la sécurité, réduit fortement les risques de fuite de données et diminue le nombre d’incidents que doivent traiter les équipes informatiques et de sécurité. En adoptant cette approche structurée, les administrations peuvent exploiter l’IA générative de façon sûre et efficace, sans sacrifier leur performance opérationnelle.

Des outils tels que la détection de l’informatique fantôme, les rapports de détection de l’IA générative ou le suivi des requêtes fournissent une vision claire des usages. Ils servent de base pour ajuster les politiques et adapter les contrôles de sécurité à l’évolution des besoins. Ces informations constituent la base d’une approche pratique et multicouche de la gestion des applications d’IA générative.

Une méthode simple consiste à classer les applications en trois catégories : à haut risque, à risque moyen et à faible risque. Haut risque : ces applications doivent être bloquées pour éviter toute vulnérabilité. Risque moyen : elles peuvent rester accessibles, mais avec des contrôles de sécurité renforcés tels que l’isolation du navigateur et une protection plus stricte des données. Faible risque : elles sont directement utilisables, mais avec des restrictions ciblées sur certains contenus ou actions.

Approche multicouche pour sécuriser les applications d’IA

Cette structure permet aux agences d’adopter une approche Zero Trust de l’IA générative. Dans ce modèle, les applications inconnues, récemment publiées ou non approuvées sont bloquées par défaut. Les applications approuvées, mais pas encore officialisées, sont isolées via des couches de sécurité supplémentaires, tandis que les applications entièrement approuvées offrent une expérience utilisateur fluide, assortie de protections adaptées. Pour faciliter le déploiement et la gestion, les agences disposent d’outils tels que les étiquettes d’applications personnalisées ou les profils de risque. Les équipes sécurité peuvent établir des règles prédéfinies, qui s’appliquent automatiquement selon le niveau de risque attribué à chaque application. Il suffit d’étiqueter une application pour que les règles adaptées s’appliquent, réduisant la charge administrative tout en maintenant un contrôle strict.

Automatiser les workflows d’incidents

La gestion des incidents constitue également une considération critique Les agences doivent réduire au maximum le nombre de cas que les équipes SOC ou les administrateurs de données traitent manuellement. Les violations de faible ou moyenne gravité, par exemple, doivent être consignées à des fins d’audit et clôturées automatiquement, sans intervention manuelle lourde. S’agissant malgré tout de violations de politique, les utilisateurs doivent être avertis et invités à justifier leurs actions, une étape précieuse pour renforcer leur formation et encourager la responsabilisation.

Avec Zscaler, les politiques d’inspection de contenu appliquées à l’IA générative permettent de définir le niveau de gravité des violations, qui seront ensuite transmis aux outils d’automatisation du workflow. Les administrateurs peuvent ainsi concevoir des workflows adaptés à la gravité de chaque incident. Des attributs supplémentaires tels que la gravité et d’autres caractéristiques partagées permettent de regrouper les incidents et d’associer ces groupes à des workflows automatisés. Cette approche simplifie le traitement des incidents : chaque violation est gérée de manière appropriée, tout en allégeant considérablement la charge des équipes SOC.

Conclusions

Les agences gouvernementales doivent être à l'avant-garde de l'usage de l'IA générative afin de moderniser leurs opérations, accroître l'efficacité des employés et offrir de meilleurs services aux citoyens. Cependant, cette adoption doit reposer sur une architecture Zero Trust. En vérifiant, surveillant et contrôlant chaque utilisateur, appareil et interaction, quel que soit l'emplacement ou l'application, les agences peuvent sécuriser leurs initiatives en IA générative avec, au cœur de leur stratégie, une protection renforcée des données, une gouvernance claire et une expérience utilisateur simplifiée.

Zscaler permet aux agences publiques de tirer parti des gains de productivité offerts par l'IA générative grâce à son approche sécurisée et multicouche qui simplifie la gouvernance, fluidifie le déploiement et intègre la sécurité dans chaque interaction. En mettant en place des cadres de gouvernance de l'IA, en automatisant la découverte et la gestion des applications d'IA générative, en contrôlant l'usage des différentes instances et en intégrant dès le départ des capacités avancées de DLP, les agences réduisent considérablement leurs risques et peuvent faire évoluer leur stratégie d'adoption sans surcharger les équipes IT et sécurité.

Les responsables d'agences sont encouragés à adopter une approche stratégique et progressive face à l'évolution de l'IA générative. Première étape : sécuriser l'accès aux applications publiques d'IA générative. La suivante : recourir à l'IA agentique (document à paraître) pour renforcer la productivité sans compromis sur la sécurité. Enfin, nous verrons comment étendre les capacités de l'IA générative aux services destinés aux citoyens, en veillant à préserver la sécurité des systèmes à chaque étape. Avec Zscaler, les agences peuvent sereinement aborder ces étapes, accélérer l'innovation et maintenir les plus hauts standards de sécurité et de conformité.

Pour plus d'informations, contactez-nous ou votre équipe de compte pour organiser un atelier adapté à vos besions.

+1 408 533 0288 **Zscaler, Inc. (siège) • 120 Holger Way • San Jose, CA 95134** **zscaler.com/fr**

zscaler™

Zero Trust Everywhere