SentinelOne

Quatre questions que se posent les RSSI sur l'IA générative

eBook

Sommaire

Un pas de plus vers des opérations de sécurité autonomes **3**

1. Quel est l'impact de l'IA générative sur le paysage des menaces ? **4**

2. Quels sont les meilleurs cas d'utilisation de l'IA générative en cybersécurité ? **7**

3. Comment les entreprises peuvent-elles évaluer les offres d'IA générative et leurs fournisseurs ? **15**

4. Comment l'IA générative va-t-elle redéfinir le rôle des SOC ? **17**

Découvrez comment l'IA générative peut vous aider à unifier, accélérer et simplifier vos processus SecOps **19**

Un pas de plus vers des opérations de sécurité autonomes

L'IA générative opère une profonde transformation de l'interaction entre l'humain et la machine en cybersécurité

L'intelligence artificielle (IA) traditionnelle est déjà bien implantée dans les processus de sécurité et fait partie intégrante des workflows des SecOps. Dans ce contexte, l'apprentissage supervisé est couramment utilisé pour détecter des malwares par l'analyse de scripts, tandis que l'apprentissage non supervisé est mis en œuvre pour corréler les signaux et détecter des activités suspectes. Traditionnellement, les SecOps utilisaient surtout l'IA à des fins de détection. L'IA générative est un type d'intelligence artificielle capable de générer de nouveaux contenus (texte, image, vidéo, audio) à partir d'invites (prompts) en s'appuyant sur de vastes quantités de données d'entraînement à sa disposition. Cette nouvelle technologie offre des capacités exceptionnelles **sur le plan de la compréhension du langage naturel et de la génération de code**, et son potentiel transformatif s'étend à l'ensemble du workflow des SecOps. Dans cet eBook, nous nous intéresserons tout particulièrement aux **grands modèles de langage (LLM, Large Language Models)**, une sous-catégorie d'IA générative conçue pour la production de textes similaires à ceux rédigés par des humains, et à leur impact sur les opérations de sécurité. Nous sommes convaincus que les LLM peuvent aider les analystes à dégager du temps, en **automatisant les tâches de routine** et en leur permettant de se consacrer davantage à des tâches telles que **la recherche, la corrélation, l'interrogation et la contextualisation**.

Bon nombre de RSSI se penchent sur cette technologie et cherchent à en appréhender les implications – en termes d'opportunités et de dangers. Les risques de l'IA entre des mains malveillantes l'emportent-ils sur les avantages de l'automatisation par l'IA entre les mains des équipes sécurité ?

Nous plaidons pour un optimisme prudent. Notre conviction est que l'IA générative fera pencher la balance en faveur des équipes sécurité, car elle leur permettra d'alléger leur lourde charge de travail. Il convient toutefois de se montrer stratégique dans l'adoption des cas d'utilisation. Les meilleures opportunités sont à chercher là où les forces de l'IA générative répondent aux défis actuels des SOC.

63 %
des spécialistes de la sécurité ressentent un certain épuisement professionnel[1]

71 %
estiment que la pénurie de compétences en cybersécurité affecte leur entreprise (contre 57 % en 2021)[2]

56 %
des grandes entreprises doivent gérer plus de 1000 alertes de sécurité chaque jour[3]

63 %
des entreprises considèrent les alertes dupliquées comme un défi modéré à majeur, 60 % pensent de même pour les faux positifs[4]

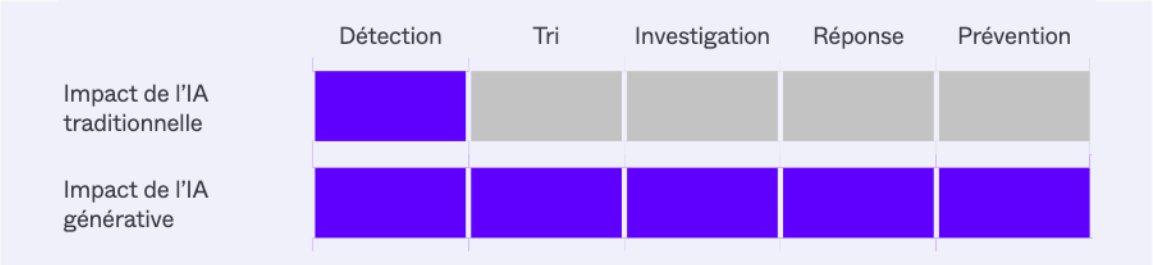
La transformation la plus considérable opérée par l'IA générative sur le **rapport entre l'humain et la machine** réside dans la démocratisation de la cybersécurité, car elle abaisse la barrière à l'entrée en ce qui concerne les tâches complexes. Cette avancée est rendue possible par l'interface en langage naturel de l'IA générative, capable de recevoir une invite, d'exécuter des requêtes complexes et d'en communiquer les résultats dans un format intuitif. Les analystes juniors étendent leur champ d'action, tandis que leurs collègues plus expérimentés peuvent se consacrer à des tâches plus stratégiques et à forte incidence.

[1] Voice of the SOC, Tines, 2023

[2] The Life and Times of Cybersecurity Professionals Volume VI, ESG, 2023

[3] 56% of Large Companies Handle 1,000+ Security Alerts Each Day, Dark Reading, 2020

[4] CSA Official Press Release, CSA, 2024



Dans cet eBook, nous vous faisons part de nos réflexions sur les opportunités immédiates et les transformations à plus long terme de l'IA générative dans le domaine de la cybersécurité. Pour ce faire, nous allons répondre aux quatre questions les plus fréquemment posées par les RSSI :

1 | Quel est l'impact de l'IA générative sur le paysage des menaces ?

2 | Quels sont les meilleurs cas d'utilisation de l'IA générative en cybersécurité ?

3 | Comment les entreprises peuvent-elles évaluer les offres d'IA générative et leurs fournisseurs ?

4 | Comment l'IA générative va-t-elle redéfinir le rôle des SOC ?

1 | Quel est l'impact de l'IA générative sur le paysage des menaces ?

Outil au service du bien ou dangereuse asymétrie ?

Les créateurs de LLM se sont efforcés d'inclure des protections à leurs modèles. Le degré de réflexion et d'investissement accordé à la prise en compte de potentielles répercussions sociétales ainsi que la mise en place de garde-fous pour favoriser la sécurité et réduire les utilisations abusives font de l'IA générative une innovation technologique quasi unique en son genre. Cependant, malgré ces efforts honorables, l'IA générative introduit de nouveaux types de risques. Ces outils sont déjà largement utilisés **tant par les cybercriminels que par les équipes sécurité**. Leurs effets les plus dangereux ne concernent pas forcément le jailbreaking – un terme qui se réfère à la manipulation malveillante des invites et au contournement des protections des LLM. Les cybercriminels peuvent tirer parti de ces outils de la même façon qu'un utilisateur ordinaire : pour coder plus rapidement, trouver des informations pertinentes et accélérer l'exécution de tâches de routine.

Jusqu'ici, les utilisations malveillantes de l'IA générative ont surtout accru de façon spectaculaire la qualité et la sophistication de l'**ingénierie sociale**. Les outils d'IA générative peuvent produire des messages de phishing convaincants à grande échelle dans n'importe quelle langue, même si les cyberpirates eux-mêmes ne maîtrisent pas la langue en question. En conséquence, les éléments qui trahissaient les attaques par ingénierie sociale jusqu'alors ne se repèrent plus si facilement. Par ailleurs, avec l'émergence des « **deepfakes** », les cybercriminels arrivent désormais à imiter de façon convaincante la voix ou même l'apparence en vidéo de cadres dirigeants à partir d'un petit échantillon. Ils peuvent se servir de cette nouvelle technologie pour cibler des personnes ou déjouer les systèmes d'authentification biométrique. L'ensemble de ces nouvelles possibilités, conjuguées aux phénomènes de Ransomware-as-a-Service (RaaS), ont **abaissé la barrière à l'entrée** pour les activités malveillantes. Résultat ? Un fractionnement du paysage des menaces, avec l'apparition en masse d'acteurs malveillants de moyenne envergure dont la prolifération nuit au maintien d'une cyberveille de qualité.

On a recensé quelques cas expérimentaux d'utilisation cybercriminelle de l'IA au-delà de l'ingénierie sociale, à des fins plus techniques.

En voici quelques exemples.

- **Techniques de script optimisées par les LLM**
  Les LLM offrent un gain de temps considérable en ce qui concerne toutes les tâches de codage. Si la plupart des modèles sont dotés de protections contre la création de scripts malveillants, les LLM restent néanmoins utiles aux cybercriminels pour accélérer l'écriture du code et respecter des paramètres complexes.
- **Contournement de la détection des anomalies optimisé par les LLM**
  Les LLM sont exploités pour développer du code capable d'outrepasser les mécanismes de détection les plus courants.
- **Recherche de vulnérabilités assistée par les LLM**
  Des outils de recherche sémantique puissants et disponibles au grand public sont utilisés par des cybercriminels pour collecter des informations sur des vulnérabilités connues. Ces outils ont une fonction de reconnaissance plus large et peuvent notamment aider les entreprises à identifier et à étudier des cibles potentielles.

L'IA à la une

« Des escrocs clonent la voix d'un PDG et extorquent 35 millions de dollars[1] »

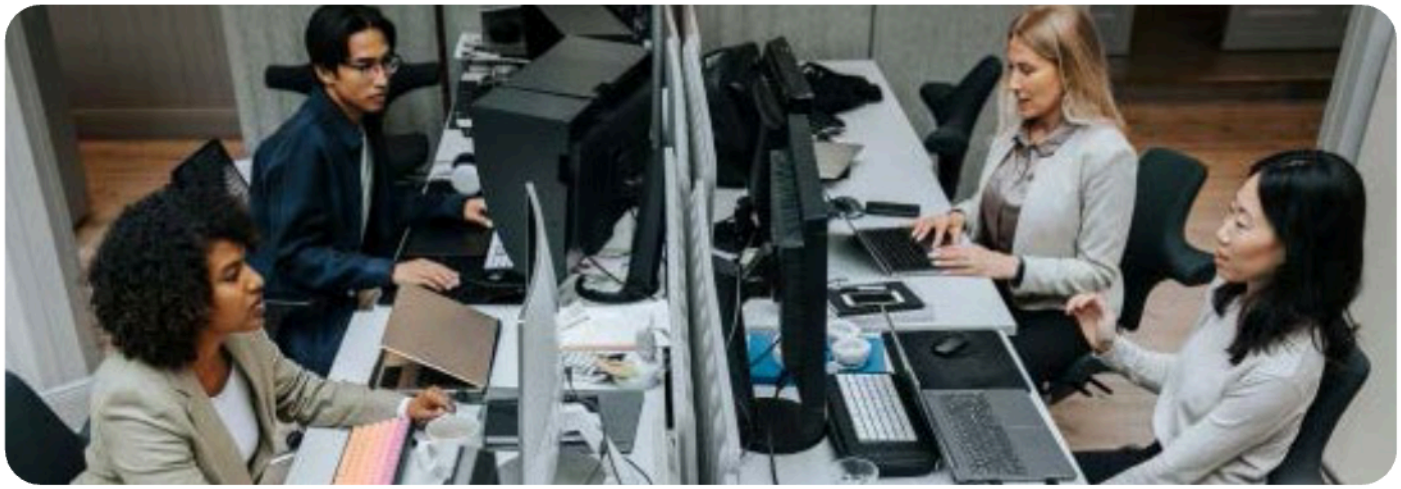
« Des hackers déjouent l'authentification biométrique de Bitfinex avec un deepfake vidéo généré par l'IA[2] »

« L'IA générative, cette ingénierie sociale sous dopage[3] »

[1] Forbes, 2023
[2] Deloitte, 2023
[3] Inc, 2023

Globalement, nous constatons que l'utilisation malveillante de l'IA générative en est encore à ses débuts en dehors de l'ingénierie sociale, mais elle pourrait prendre de l'ampleur à l'avenir. Nous anticipons un possible détournement de certaines capacités théoriques à des fins illicites :

- **Les malwares polymorphes basés sur l'IA** sont un type de logiciel malveillant qui tire profit de l'intelligence artificielle pour modifier constamment son code, ce qui le rend plus difficile à détecter et à combattre.
- **Les techniques d'IA générative avec agent** pourraient exploiter l'IA pour assembler plusieurs actions automatisées complexes.
- **La découverte accélérée des exploits** pourrait tirer parti de l'IA générative pour effectuer des tests d'intrusion créatifs et adaptatifs sur un environnement afin d'en détecter les vulnérabilités.



Une surface d'attaque inédite

Un grand nombre d'entreprises ont adopté des solutions d'IA générative très rapidement, à la faveur de son intense exposition médiatique. Dans le même temps, les RSSI ont œuvré à sécuriser l'utilisation grandissante de l'IA générative en se posant les questions suivantes :

- La protection des données actuelle suffit-elle, dans la mesure où ces modèles sont exposés à de vastes quantités de données propriétaires ?
- L'ingestion de données sensibles ou propriétaires dans des LLM pourrait-elle conduire à une fuite de données ou à une perte de leur contrôle ?
- Le modèle est-il protégé contre les utilisations abusives et le jailbreaking ?
- Comment prévenir les menaces pour la chaîne d'approvisionnement basées sur l'IA générative dans le cadre du cycle de développement logiciel ?
- Faut-il appliquer des principes Zero Trust aux outils, applications et plateformes basés sur l'IA ou mettre en place une surveillance continue et des contrôles d'accès dynamiques ?

L'IA générative est susceptible d'aggraver un risque déjà encouru par beaucoup d'entreprises : **les privilèges d'accès excessifs** accordés aux utilisateurs. Aujourd'hui, la plupart des organisations ne disposent toujours pas du niveau de contrôle granulaire qu'elles souhaiteraient posséder sur leurs données. Cela s'explique en partie par leur nécessité de partager des données et de collaborer, souvent avec des acteurs tiers. Ainsi, des milliers d'utilisateurs individuels sont susceptibles de partager des fichiers avec des pairs ou de les conserver dans des dossiers partagés, souvent accessibles à de nombreuses personnes. Dans un tel contexte, il est difficile d'éviter les négligences et d'assurer une bonne hygiène numérique de la part des utilisateurs ordinaires. L'IA générative aggrave le problème, car de nombreuses entreprises se sont récemment dotées **d'outils de recherche sémantique très sophistiqués**. Jusqu'ici, le risque était qu'un utilisateur tombe par hasard sur des informations auxquelles il n'aurait pas dû avoir accès. Désormais, un puissant outil de recherche est à disposition pour dénicher ces fichiers dont l'accès est trop permissif. De plus, un intrus doté d'identifiants compromis peut exploiter l'outil de recherche basé sur l'IA générative pour effectuer des tâches de **reconnaissance**, afin de repérer plus vite les données intéressantes à exfiltrer.

Cette ère de l'IA générative fait également bouger les lignes **des workflows de développement des logiciels**. De nos jours, lorsqu'une entreprise cherche à développer un modèle d'IA, elle s'appuie souvent sur un modèle existant plutôt que de tout construire de A à Z. Plusieurs plateformes en accès public donnent la possibilité de partager des modèles et des ensembles de données. Si elles permettent de collaborer, elles introduisent aussi un nouveau **risque pour la chaîne d'approvisionnement**. Les entreprises doivent s'assurer de contrôler minutieusement la provenance de leur code, mais aussi sécuriser la surface d'attaque associée à **des modèles tiers**. Les méthodologies DevSecOps actuelles doivent être adaptées et étendues pour inclure la surface d'attaque supplémentaire liée à l'IA générative.

Vos équipes utilisent déjà très probablement l'IA générative

Les outils gratuits et accessibles au public, comme ChatGPT ou les assistants de codage, sont très tentants pour les utilisateurs. Même si votre entreprise a communiqué des consignes ou pris des mesures pour en interdire l'usage, il est fort probable que certains collaborateurs s'en servent sur leurs équipements professionnels ou personnels. Cet usage peut rendre des **données propriétaires de l'entreprise** vulnérables ou introduire une nouvelle surface d'attaque non surveillée. Dans la mesure du possible, les entreprises doivent s'efforcer de sortir l'IA générative de l'ombre et fournir des outils autorisés et validés pour éviter le Shadow IT.

Les SOC aussi sont concernés : il y a fort à parier que certains membres de votre équipe testent déjà ces outils pour accélérer leurs workflows. Pour appliquer les cas d'utilisation de l'IA générative avec **une gouvernance et des garde-fous adéquats**, mieux vaut mettre en place une **stratégie approuvée**.

2 | Quels sont les meilleurs cas d'utilisation de l'IA générative en cybersécurité ?

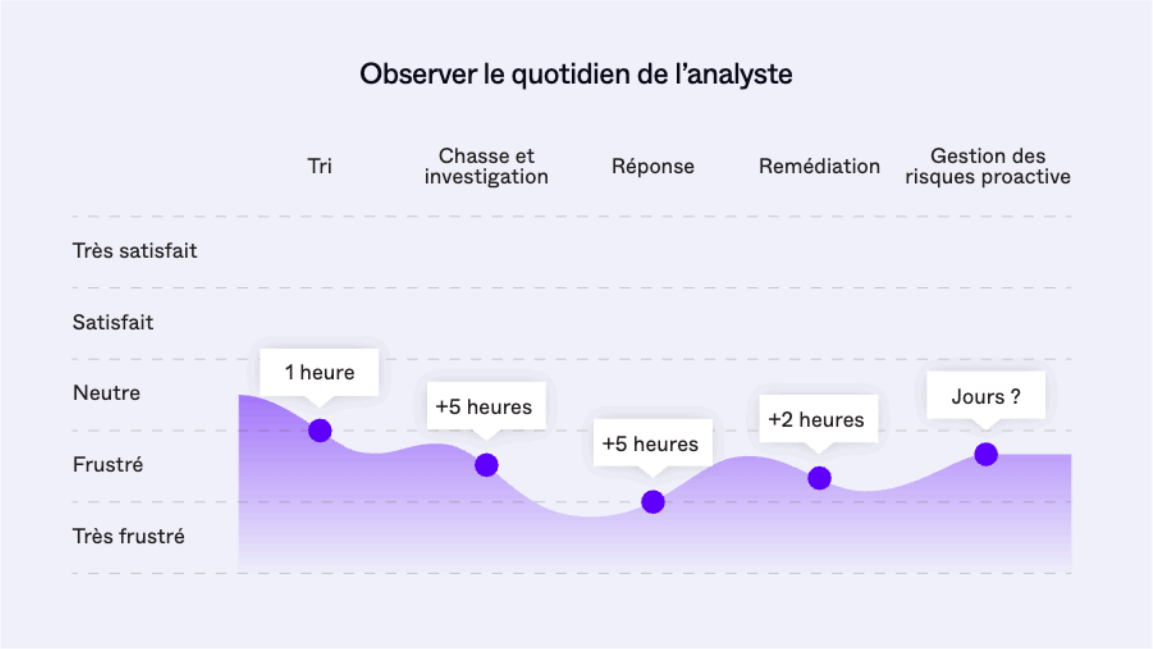
Identifier les cas les plus faciles à appliquer

Dans le domaine de la sécurité, les cas d'utilisation de l'IA générative sont légion. On peut citer les cas centrés sur les LLM, qui s'appuient sur les assistants conversationnels, les capacités de recherche sémantique et la génération de code, ou encore les cas d'utilisation multimodaux, par exemple pour évaluer le risque d'un fichier binaire inconnu. Notre perspective pour identifier les cas d'utilisation les plus prometteurs est de repérer l'intersection entre les défis majeurs des professionnels de la sécurité et les forces de l'IA générative. Ce sont les cas à la croisée de ces deux paramètres qui produisent les meilleurs résultats. C'est pourquoi nous nous sommes concentrés sur les SOC, où beaucoup de tâches manuelles chronophages sont de bonnes candidates à l'automatisation par l'IA générative.

Observer le quotidien de l'analyste

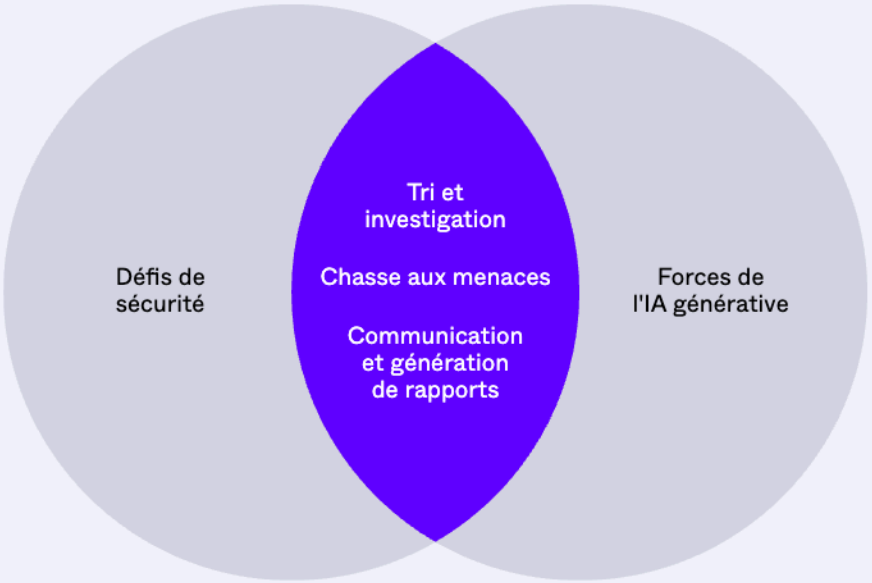
La pénurie de main-d'œuvre qualifiée en cybersécurité ne date pas d'hier et, d'après les enquêtes, la situation ne va pas s'arranger de sitôt. Les analystes subissent de ce fait une certaine pression allant de la surcharge de travail à l'épuisement professionnel. La plupart des entreprises ne peuvent pas traiter toutes les alertes qu'elles reçoivent, et c'est aux analystes que reviennent leur tri et leur priorisation, souvent par analyse heuristique pour compenser un manque de contexte. En prime, les analystes doivent souvent faire face à la multiplication des outils et passer constamment d'une fenêtre à l'autre pour trouver les informations nécessaires.

De la même façon, la plupart des équipes ont défini des initiatives stratégiques qui devraient être prioritaires (comme le renforcement de la sécurité sur les surfaces d'attaque externes), mais elles progressent moins vite qu'il ne le faudrait sur ces projets en raison des urgences qu'elles sont constamment appelées à gérer. En matière de réponse aux incidents, elles se montrent friandes de tout processus susceptible de faire gagner du temps aux analystes : traitement des données accéléré, moins de basculements entre interfaces ou écriture plus rapide des requêtes.



Forces de l'IA générative

L'un des atouts majeurs de l'IA générative est sa perception : examiner de vastes quantités de données, établir des corrélations entre éléments et présenter cette information de façon cohérente pour un humain. L'une des avancées les plus marquantes de l'IA générative dans la relation entre l'humain et la machine est sa capacité à échanger en langage naturel, au lieu d'obliger l'humain à communiquer avec un langage de code ou via une interface spécifique. Cet échange est facilité dans les deux sens : non seulement l'IA générative perçoit et résume pour l'humain, mais elle comprend aussi les invites en langage naturel et sait transformer l'intention de l'humain en action.



Les cas d'utilisation les plus rentables sont ceux qui résolvent les défis de sécurité les plus pressants en tirant parti des forces propres à l'IA générative.

À partir de cette analyse, nous recommandons aux RSSI d'étudier ces trois cas d'utilisation à fort potentiel :

- **Le tri des alertes et l'aide à l'investigation** sont souvent cités comme les tâches les plus chronophages. Un effort manuel considérable est nécessaire pour trouver et analyser les journaux pertinents. Les capacités de recherche et de synthèse de l'IA générative sont prometteuses dans ce domaine.
- **La chasse aux menaces** nécessite une évaluation de données de cyberveille complexes et dynamiques, ainsi que des connaissances pointues sur les langages de code et les schémas de données. La capacité de l'IA générative à transformer en code le langage naturel des invites représente un immense potentiel de réduction de la barrière à l'entrée et de gain de temps.
- **La coordination et la génération de rapports** créent de la friction lors de la réponse à des incidents complexes et, au quotidien, détournent les professionnels de la sécurité de leurs tâches principales. L'IA générative peut automatiquement enregistrer et résumer les actions entreprises lors des investigations et fournir une vue unique pour faciliter les réponses aux incidents.

Cas d'utilisation 1 : tri des alertes et aide à l'investigation

Ce cas d'utilisation porte sur l'interprétation et la priorisation des alertes, ainsi que sur les mesures d'investigation que les analystes peuvent décider de prendre si l'alerte leur semble significative.

Défis

Les équipes SecOps, souvent en sous-effectif, sont plus que jamais confrontées à **d'immenses volumes d'alertes**. Chaque analyste traite en général des dizaines d'alertes par heure. Le processus commence généralement par le tri d'une longue liste d'alertes, souvent effectué par des analystes juniors. Leurs outils intègrent parfois un certain degré de priorisation, mais certaines équipes travaillent encore manuellement. Les analystes vont s'appuyer sur **l'analyse heuristique** pour identifier les alertes qui méritent une attention plus soutenue. L'objectif de cette étape est de parvenir à déterminer quelles alertes sont prioritaires et exigent d'être investiguées. Le défi est double à ce stade : au-delà du volume considérable des alertes, **les analystes travaillent isolément**, ce qui peut les priver d'éléments de contexte collectés par d'autres outils ou gérés par une autre équipe. L'étape de tri fait donc perdre un temps précieux aux équipes et fait courir le risque de négliger des risques importants.

93 %

des entreprises déclarent ne pas pouvoir traiter toutes les alertes le jour de leur réception[1]

56 %

des grandes entreprises doivent gérer plus de 1000 alertes de sécurité chaque jour[1]

Ensuite vient la phase d'**investigation**. C'est une étape compliquée du processus puisque les données de contextualisation sont éparpillées sur divers outils. Les alertes peuvent concerner plusieurs systèmes et dépendances, ce qui oblige à passer un temps précieux à les retracer sur différents composants pour assurer une compréhension globale de leur impact. **L'intégration et la corrélation** des données à partir de plusieurs outils et plateformes de sécurité n'est pas chose aisée, d'autant plus lorsqu'elles impliquent des formats propriétaires. Résultat : les analystes passent un temps considérable à passer en revue manuellement des journaux.

Exemple : une alerte indique la suppression simultanée de plusieurs utilisateurs. L'analyste consulte les journaux pour trouver l'utilisateur qui a exécuté la commande. Il vérifie ensuite les journaux réseau pour déterminer si des données ont été transférées.

Opportunité

Le potentiel d'automatisation de l'IA générative pourrait s'appliquer à la phase de tri initiale. Ainsi, les analystes peuvent partir directement d'un verdict, plutôt que de devoir le déterminer au préalable. L'investigation s'en trouve accélérée, car les informations de contexte pertinentes étant rassemblées, et non plus éparpillées sur plusieurs sources, des étapes d'enrichissement contextuel sont déclenchées automatiquement en réponse à une alerte.

[1] 56% of Large Companies Handle 1,000+ Security Alerts Each Day, Dark Reading, 2020

Solution

Les solutions d'IA générative sont déployées sur des milliers d'environnements client. Elles utilisent les enseignements de ce contexte global pour comparer une alerte donnée à d'autres alertes similaires pour déterminer s'il s'agit d'un vrai ou d'un faux positif. Les alertes qui ne ressemblent à aucune autre déjà rencontrée par l'outil peuvent être remontées à un analyste humain. Selon le résultat du tri initial, des tâches automatisées supplémentaires sont déclenchées en cascade. Celles-ci enrichissent les données télémétriques pour fournir davantage de contexte aux équipes. En voici quelques exemples.

- **Enrichissement contextuel**
  Rassemble du contexte et des données supplémentaires autour de l'événement, comme les comptes utilisateur, les informations sur l'équipement ou encore les modèles de trafic du réseau.
- **Corrélation**
  Détermine si l'événement fait partie d'une attaque plus vaste ou constitue un incident isolé en le corrélant à d'autres événements potentiellement problématiques stockés dans un backend de journalisation centralisé.
- **Recherche dans les données de cyberveille**
  Interroge les flux de cyberveille pour déterminer si les indicateurs associés à l'événement sont des menaces connues.

Ce processus de détections en chaîne est inspiré des étapes suivies par les analystes humains dans des contextes similaires. Il inclut une boucle de rétroaction dans un but d'amélioration continue.

Cas d'utilisation 2 : chasse aux menaces

La chasse aux menaces est une pratique de sécurité qui consiste à rechercher des indicateurs de compromission de manière itérative. Les chasseurs recherchent de façon proactive des menaces à traiter, contrairement à la démarche réactive face à des alertes déjà reçues. Pour de nombreuses entreprises, la contrainte de temps et la haute barrière à l'entrée freinent le déploiement d'une chasse aux menaces formellement mise en place.

Défis

Selon les analystes avec lesquels nous échangeons, les défis majeurs sont la chasse aux menaces et l'investigation, car leurs workflows longs et complexes nécessitent une multitude de tâches manuelles.

La première étape pour une équipe SecOps est d'étudier la cyberveille et de développer une hypothèse quant aux menaces qui pourraient être les plus pertinentes. Comme nous l'avons vu précédemment, il s'agit d'un problème qui prend de l'ampleur, car **l'éclatement du paysage des menaces dynamique** complique l'identification des principaux acteurs en raison de l'émergence constante de nouveaux groupes et affiliés. À titre d'exemple, un gang cybercriminel peut voir ses activités perturbées par les forces de l'ordre, puis refaire surface quelques semaines plus tard avec de nouveaux serveurs et de nouvelles tactiques.

Le défi suivant réside dans la formulation des requêtes. Cette compétence exige l'intervention d'un analyste expérimenté, puisqu'on dénombre une vingtaine de langages de requête pour la chasse aux menaces – et aucun d'entre eux n'est intuitif. De plus, la formulation d'une requête efficace est conditionnée à la bonne compréhension des **schémas de données** si l'on souhaite trouver les champs et journaux permettant de répondre à la question posée. C'est une barrière à l'entrée non négligeable pour une entreprise qui n'est pas rompue à l'architecture de données. Même pour des experts, il serait impossible de mémoriser tous les aspects d'un schéma de données. Rechercher les champs prend un temps précieux, et ce même si les données sont normalisées.

Éléments indispensables pour répondre à la question « suis-je visé par une menace précise ? »

Si l'un d'entre eux manque à l'appel, la chasse aux menaces ne sera pas efficace.

Opportunité

La capacité de l'IA générative à **traduire le code en langage naturel et vice versa** permet aux analystes de créer des requêtes simples et rapides à partir d'invites en langage naturel et d'obtenir un résumé des résultats ou de la cyberveille sur les menaces. Cette **réduction de la barrière à l'entrée** permet aux analystes juniors de participer à la chasse aux menaces et fait gagner du temps aux spécialistes plus expérimentés. Les analystes les plus aguerris n'auront **plus à rechercher manuellement les éléments des schémas de données**, une tâche particulièrement chronophage.

Solution

L'IA générative peut résumer des flux complexes de chasse aux menaces et répondre en langage naturel aux questions des analystes, ce qui les aide à se concentrer sur des domaines plus pertinents. En outre, les outils d'IA générative sont capables de **préparer des modèles de requêtes** à partir des tendances de cyberveille. **Il suffit ensuite d'un clic** pour commencer la chasse aux menaces. Ces modèles peuvent être axés sur un cybercriminel, des ressources, des tactiques, des anomalies et bien plus encore.

Plutôt que d'écrire une requête complexe, les analystes peuvent saisir une invite en langage naturel. Voici quelques exemples de requêtes pour illustrer notre propos :

« SentinelLabs a publié un rapport qui détaille des techniques pour Lockbit 3.0. Effectue une chasse aux menaces sur les TTP "Event Viewer Tampering". »

```
| filter( event.type == "Behavioral Indicators" AND (indicator.name ==
"EventViewerTampering" OR indicator.name == "EventTampering") ) | group
EventCount = count() by src.process.user | sort -EventCount | limit 1000
```

L'IA générative peut ensuite fournir un résumé des résultats en langage naturel, ce qui représente un gain de temps non négligeable pour les requêtes complexes.

```
L'altération de l'observateur d'événements a été identifiée et regroupée
par nom d'utilisateur. L'utilisateur « jmartin » est le plus associé aux
altérations, suivi par « mdupont ».
```

L'IA générative peut ensuite étendre et approfondir l'investigation en suggérant des questions de suivi contextuelles. Par exemple :

- « Quels utilisateurs ont accédé à l'observateur d'événements sans autorisation appropriée ? »
- « Peux-tu me fournir le détail des altérations apportées à l'observateur d'événements par type d'événement (processus de création, modification de fichier, etc.) ? »
- « Montre-moi les activités qui ont altéré l'observateur d'événements et qui sont associées à l'utilisateur "jmartin". »

Cas d'utilisation 3 : coordination et génération de rapports

Les équipes de cybersécurité ne sont pas isolées : elles font partie d'un écosystème plus vaste et interagissent avec une multitude de parties prenantes. Répondre à un incident majeur exige que plusieurs professionnels de la sécurité (voire plusieurs équipes) travaillent en étroite collaboration et communiquent avec les parties prenantes du département informatique et de l'entreprise dans sa globalité.

Défis

La coordination et la communication peuvent être source de friction et accaparer le temps précieux des équipes, même en dehors des situations d'urgence. Bon nombre d'entreprises souhaitent aider leurs équipes à se concentrer sur la sécurité plutôt que sur **des tâches manuelles longues et fastidieuses**, comme la rédaction d'e-mails et de rapports.

Ces sources de friction sont d'autant plus malvenues lors de la gestion d'un incident majeur où chaque minute compte. Les défis principaux sont les suivants :

- **Les silos d'informations** avec des outils et équipes distincts selon les divers domaines du dispositif de sécurité. Certaines entreprises fonctionnent par exemple avec différentes équipes pour la sécurité des endpoints, la sécurité réseau et la gestion des identités et des accès. Les membres de chaque équipe doivent alors résumer manuellement leurs résultats et les partager avec leurs collègues.
- **La duplication des efforts** à cause d'un manque de visibilité mutuelle sur les missions d'autres membres de l'équipe.
- **La prise de décisions et la remontée des problèmes** lorsque les incidents complexes engendrent une surcharge d'information. Les membres de l'équipe sont alors contraints de passer du temps à s'assurer que les informations pertinentes sont remontées aux personnes adéquates. À cela s'ajoute la difficulté de déterminer si une situation ambiguë et changeante répond aux critères de remontée et, si oui, par quel chemin la faire remonter.

Opportunité

L'IA générative peut prendre en charge la majorité des tâches manuelles de coordination en regroupant automatiquement les informations pertinentes en un seul lieu accessible à toutes les parties prenantes pour vérification. Elle est également capable d'exploiter ces informations pour **générer des versions préliminaires d'e-mails et de rapports** et ainsi faire gagner du temps aux équipes sécurité.

Interrogés sur les tâches qu'ils apprécient le moins, les professionnels de la sécurité citent la communication (par e-mail, Slack, etc.) en **première position**[1].

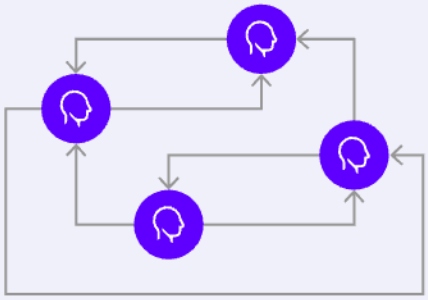
[1] Voice of the SOC, Tines, 2023

Solution

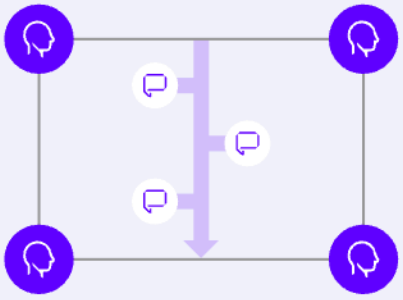
Lors du tri des alertes, l'IA générative regroupe un nombre considérable d'éléments de contexte à partir de sources différentes. Elle consigne chaque étape de l'investigation avec ses résultats dans l'ordre chronologique. L'information peut être rendue disponible sous la forme de **dossiers d'investigation vérifiables et partageables** pour fournir une vue unifiée de l'investigation à l'ensemble de l'équipe. Grâce à l'IA générative, les analystes n'ont plus besoin de résumer et de partager manuellement les résultats de leurs investigations, ce qui évite les **efforts en double**. Le dossier d'investigation peut également servir à évaluer l'incident a posteriori et à répondre aux exigences de reporting. À partir de ce dossier, l'IA générative est capable de rédiger un e-mail ou un rapport.

Par exemple : Des caractéristiques d'accès anormaux aux fichiers et un pic d'utilisation du processeur sur plusieurs serveurs sont détectés, ce qui déclenche une alerte. Un analyste décide de la remonter et génère un résumé rapide par e-mail. En prenant connaissance de l'e-mail, ses collègues peuvent accéder en un clic à une vue partagée de toutes les étapes d'investigation et aux mesures de réponse sur incident déjà prises.

À l'avenir, l'IA générative pourra probablement aller encore plus loin pour soutenir la coordination des équipes, en évaluant les situations et **en attribuant les tâches** aux analystes selon leurs domaines d'expertise. Cette coordination pourrait suivre des règles prédéfinies et des playbooks ou être générée de façon dynamique par l'IA selon le contexte.



Actuellement, la communication et la coordination complexifient la réponse aux incidents majeurs.



L'IA générative regroupe des éléments de contexte de sources différentes en une vue partagée pour une collaboration simplifiée.

3 | Comment les entreprises peuvent-elles évaluer les offres d'IA générative et leurs fournisseurs ?

Simple ajout vs véritable intégration

Face à l'intérêt des clients, la plupart des fournisseurs de cybersécurité affirment aujourd'hui proposer des solutions qui intègrent l'IA générative. **Mais comment distinguer les fausses promesses marketing des vraies avancées technologiques ?** Avant tout, vérifiez si la plateforme a été repensée pour intégrer pleinement l'IA générative ou s'il s'agit d'un simple ajout de fonctionnalité. Dans ce cas, le fournisseur n'applique qu'une couche superficielle à un LLM sans guider la sortie machine ni apporter de valeur ajoutée à la résolution des cas d'utilisation de sécurité.

Assurer la précision et éviter les hallucinations

Les LLM génèrent du texte à partir de modélisations apprises grâce à des données d'entraînement. Ils n'ont pas de réelle compréhension du contenu, mais prédisent le mot suivant par relation statistique. S'ils sont exposés à des invites vagues ou contradictoires, les modèles peuvent générer des réponses créatives mais incorrectes, appelées « hallucinations ». D'où l'importance d'appliquer une couche supplémentaire au LLM de base lorsqu'on adapte l'IA générative au contexte de la cybersécurité. Cherchez les éléments suivants dans l'architecture du fournisseur :

- **Génération augmentée par récupération (RAG)**
  Approche qui allie les avantages de la récupération et de la génération en intégrant des sources de données externes au modèle d'IA générative.

- **Base de connaissances ciblée**
  Le modèle du fournisseur devrait faire appel à un référentiel de contextes de cybersécurité dont la maintenance est assurée par le fournisseur de la solution. Lorsque le modèle répond à une invite, il commence par puiser dans ce référentiel pour trouver les informations pertinentes, qui sont ensuite injectées dans le modèle génératif avec la requête originale afin de produire une réponse plus précise et adaptée au contexte.

- **Contenu propriétaire propre à la cybersécurité**
  Un fournisseur prometteur est celui qui a déjà fait ses preuves avec l'IA traditionnelle dans un contexte de sécurité, avec des renseignements propriétaires sur la détection des menaces et de vastes ensembles de données sur les malwares.

- **Métaconditions**
  Tout outil d'IA générative conçu pour la cybersécurité devrait être programmé avec un ensemble de paramètres et de contraintes à respecter lorsqu'il génère une réponse. Ces conditions évitent non seulement les réponses « trop créatives », mais aussi le jailbreaking et les utilisations abusives.

L'importance d'un modèle de données commun

Pour effectuer des recherches performantes sur de vastes ensembles de données et les résumer efficacement, l'IA générative a besoin d'un environnement de données unifié. Un modèle de données commun construit sur une norme ouverte comme **l'OCSF (Open Cybersecurity Schema Framework)** donne les moyens d'atteindre le plein potentiel de cette technologie en permettant au modèle de traiter des informations de plusieurs sources. Autre avantage du modèle ouvert : la possibilité de traiter des données de fournisseurs différents et non pas celles du créateur de l'outil uniquement.

Protéger les données

Exposer vos journaux de sécurité et vos données télémétriques à une IA générative comporte un risque, surtout s'il s'agit d'un LLM tiers comme GPT-4. Il est donc essentiel de comprendre comment l'outil utilise vos données et de vous assurer que les contrôles et la gouvernance nécessaires sont en place. Voici quelques facteurs à prendre en compte.

- Les mesures de sécurité mises en place pour protéger vos données doivent inclure le chiffrement, les contrôles d'accès et la conformité aux réglementations. Vérifiez **les pratiques de gestion des données du fournisseur**, y compris les politiques de stockage, de traitement et de rétention des données.
- Les outils de génération gratuits se servent souvent des données soumises par les utilisateurs pour s'améliorer. Le risque est moindre pour les outils professionnels payants, mais demandez toujours confirmation pour assurer la protection de **vos informations propriétaires**.
- Votre entreprise doit comprendre comment **ses sous-traitants** gèrent ses données et s'assurer que celles-ci ne sont ni accessibles aux sous-traitants, ni utilisées pour améliorer leurs produits, ni conservées indéfiniment.

Les questions à poser aux fournisseurs

- Puis-je assister à une démonstration interactive en direct et tester des requêtes ?
- Quels sont les accords en vigueur avec vos sous-traitants ?
- Quel type d'isolation régionale utilisez-vous ?
- Quelles mesures de protection des données mettez-vous en place ?
- Comment assurez-vous la confidentialité des données exposées au modèle ?
- Comment protégez-vous votre produit du jailbreaking et de l'injection d'invites ?
- Comment votre architecture réduit-elle le risque d'hallucinations et favorise-t-elle la véracité des réponses ? Pouvez-vous communiquer des métriques sur le degré de véracité du produit ?
- Quel est le niveau de performance du modèle ?
- À partir de quels types de données l'outil peut-il produire des réponses ?
- Votre modèle de tarification est-il transparent et prévisible ?

4 | Comment l'IA générative va-t-elle redéfinir le rôle des SOC ?

Les limitations d'aujourd'hui ne seront plus celles de demain

Les entreprises peuvent d'ores et déjà adopter des cas d'utilisation matures pour tirer partir des opportunités qu'offre l'IA générative. Cependant, cette technologie est en constante évolution et n'a pas encore atteint un point d'équilibre stable. Nous encourageons les entreprises à voir au-delà des cas d'utilisation individuels et à **développer une approche stratégique, conçue sur le long terme**. Elles devraient adopter un processus formel d'analyse des workflows, pour identifier les cas d'utilisation prometteurs, tester de nouvelles solutions et déployer ensuite les nouvelles capacités.

Nous faisons le pari que tous les workflows de sécurité des informations finiront par intégrer l'IA générative. Cette technologie dépasse largement le cadre du module que l'on incorpore aux outils existants : c'est une nouvelle forme d'interaction entre l'humain et la machine qui conduit à l'émergence d'**un modèle d'opérations de sécurité autonome**. En d'autres termes, l'automatisation des tâches fastidieuses permet aux humains de se concentrer sur des tâches à plus fort impact. **L'IA automatisera les tâches manuelles répétitives et aidera à trouver du sens dans la complexité, ce qui permettra aux humains de se consacrer à la validation, à la prise de décision et à la gestion des cas atypiques.**

Modèle de maturité de l'adoption de l'IA générative en cybersécurité

Phase 1 : adoption débutante

- Les analystes expérimentés testent les capacités de l'IA générative.
- L'entreprise évalue les cas d'utilisation et les solutions.

Phase 2 : adoption mature

- L'entreprise dispose d'une solution professionnelle intégrée à son workflow et entraînée sur ses données.
- L'entreprise met en place une gouvernance formelle et des garde-fous autour de l'IA générative.
- L'entreprise évalue et teste activement des cas d'utilisation supplémentaires.

Phase 3 : adoption étendue

- Tous les workflows d'InfoSec sont optimisés par l'IA générative.
- L'entreprise possède un processus répétable d'intégration et d'extension des capacités de l'IA générative.
- Une boucle de rétroaction complète est en place pour améliorer les modèles, recouvrant les données de l'environnement et de la cyberveille globale.

Définir le niveau d'autonomie approprié

Selon nous, la prochaine étape pour les SOC est de tendre vers une « autonomie conditionnelle », où le système est capable d'effectuer automatiquement des tâches telles que la chasse aux menaces, la gestion des vulnérabilités et les actions de remédiation. Cependant, si le système détecte des anomalies en dehors de sa zone de confiance, il alerte un analyste humain qui prend le relais. Le système est entièrement autonome sous certaines conditions et inclut des mécanismes de remontée des problèmes à des analystes humains prêts à intervenir.

Une solution au plus vieux problème de la cybersécurité ?

L'IA générative suscite autant d'engouement que d'hésitation. Les risques sont réels, qu'il s'agisse des hallucinations ou de l'exploitation par des cybercriminels. Nous sommes partisans d'un optimisme prudent. Historiquement, l'un des plus grands défis que doivent relever les équipes de sécurité, partout dans le domaine, est le volume et la complexité des tâches gérées par des professionnels qualifiés en sous-effectif. **Nous sommes convaincus que l'IA générative peut aider les analystes à faire face à l'étendue et à la complexité des menaces et qu'elle fera pencher la balance en faveur de la cybersécurité.** L'automatisation permise par l'IA devrait faire gagner du temps aux équipes, qui pourront se concentrer sur des activités proactives et stratégiques au lieu d'être accaparées par des urgences. Par exemple :

- **Campagnes de sécurité**
  Effort ciblé pour résoudre un problème de sécurité, en communiquant les objectifs à l'ensemble de l'entreprise et en mesurant la performance. Aujourd'hui, la plupart des entreprises ont des initiatives de ce type, mais leurs efforts sont toujours entravés par d'autres priorités. Alléger le fardeau de la réponse aux incidents libère du temps pour déployer ces campagnes stratégiques.

- **Gestion des vulnérabilités et de l'exposition aux risques**
  De nos jours, le volume des vulnérabilités peut être difficile à gérer. Les entreprises doivent établir une stratégie de priorisation, ce qui contraint à tolérer certains risques pendant une certaine période. L'automatisation par l'IA libérera plus de temps pour déployer ces mesures préventives.

- **Chasse aux menaces proactive**
  Avec l'IA, le rôle des analystes évolue vers moins de réactions aux urgences et plus de chasse aux menaces proactive. La conséquence ? Une réduction à la fois du temps passé à répondre aux incidents et de la barrière à l'entrée pour la chasse aux menaces.

Gain de temps, amélioration des détections, accélération de l'investigation et de la réponse… l'IA présente de nombreux avantages. Et elle cache encore un atout dans sa manche : les analystes qui se servent de ces outils semblent plus épanouis dans leur métier. Ils peuvent en effet se consacrer à des activités de résolution de problèmes demandant d'exercer leur esprit critique plutôt que de gérer des tâches ennuyeuses et répétitives.

36 %

des professionnels de la cybersécurité déclarent que leurs équipes passent la majorité de leur temps à régler des problèmes prioritaires ou urgents, et trop peu de temps sur des activités stratégiques ou d'amélioration des processus[1].

93 %

des professionnels de la cybersécurité considèrent que l'automatisation améliorerait leur équilibre entre vie professionnelle et personnelle[2].

[1] Examining and Addressing Threat Detection and Response Challenges, ESG, 2019

[2] Voice of the SOC, Tines, 2023

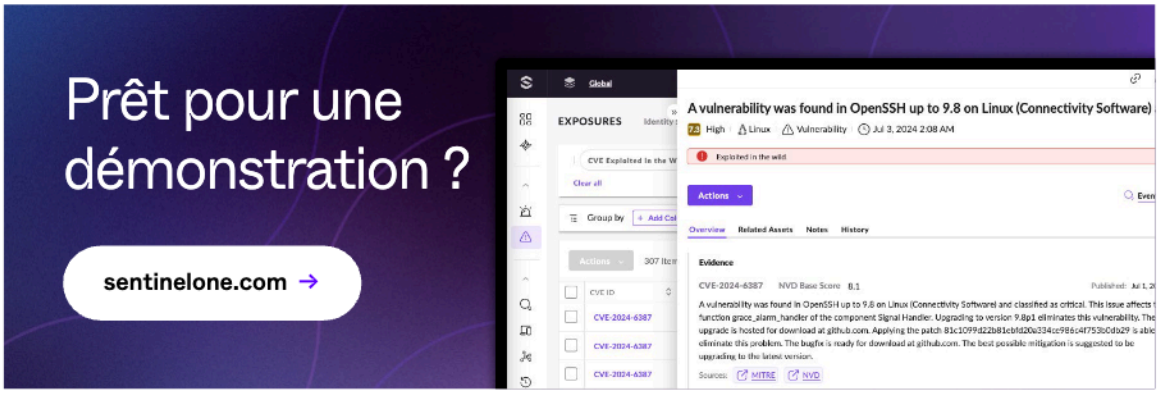
Découvrez comment l'IA générative peut vous aider à unifier, accélérer et simplifier vos processus SecOps

L'IA générative est déjà entre les mains des cybercriminels. Les équipes sécurité l'utilisent aussi, que ce soit au travers d'outils autorisés ou par Shadow IT. Si les risques sont majeurs, nous prônons toutefois un optimisme prudent. Cette technologie transforme la cybersécurité sous nos yeux en donnant aux analystes le temps et les moyens de travailler à une défense proactive. À plus long terme, nous anticipons un changement de paradigme pour les workflows des SecOps à mesure que les entreprises adopteront un modèle d'opérations de sécurité autonome.

Pour gérer efficacement les risques et prendre conscience de la valeur des cas d'utilisation en sécurité, les entreprises ont besoin d'un partenaire de confiance. SentinelOne est pionnier de l'innovation en IA générative. Après s'être illustrée en sécurité pilotée par l'IA avec sa plateforme Singularity, l'entreprise propose désormais une solution de sécurité d'IA générative avec Purple AI, l'analyste sécurité basé sur l'IA le plus avancé du marché.

Avec Purple AI, vous pouvez :

- Traduire le langage naturel en requêtes complexes de chasse aux menaces qui vous permettent d'exécuter des recherches et des tris efficaces.
- Bénéficier d'une visibilité complète grâce au seul analyste sécurité d'IA générative qui prend en charge le framework OCSF (Open Cybersecurity Schema Framework).
- Synthétiser les journaux des événements et les indicateurs en langage naturel.
- Détecter les risques proactivement avec nos guides de démarrage rapide pour la chasse aux menaces et soutenir vos analystes avec des suggestions de questions de suivi.
- Renforcer la collaboration grâce aux dossiers d'incidents partagés et à la génération d'e-mails.



Innovation. Fiabilité. Reconnaissance.

Gartner

Leader du Magic Quadrant™ 2023 consacré aux plateformes de protection des endpoints

MITRE ENGENUITY

Résultats exceptionnels à l'évaluation ATTACK

+ 100 % de protection. 100 % de détection
+ Couverture analytique exceptionnelle, 5 ans de suite
+ 100 % en temps réel, 0 retard

Gartner Peer Insights

96 % des évaluateurs de Gartner Peer Insights™ pour les solutions EDR recommandent SentinelOne Singularity

FedRAMP · TEVORA PCI DSS Attestation HIPAA Attestation · AICPA SOC · STAR LEVEL ONE · VB 100 VIRUS · SE Labs BEST Innovator WINNER 2021 · Trusted Cloud Provider CSA

Nous contacter

sales@sentinelone.com
+1-855-868-3733

fr.sentinelone.com

À propos de SentinelOne

SentinelOne (NYSE : S) propose une cybersécurité autonome de pointe, capable de prévenir, détecter et neutraliser les cyberattaques plus rapidement et plus précisément que jamais. Sa plateforme Singularity XDR protège les grandes entreprises mondiales en leur offrant visibilité en temps réel sur les surfaces d'attaque, corrélation entre plateformes et réponse aux incidents optimisée par l'intelligence artificielle. Vous disposez ainsi de plus de fonctionnalités, tout en réduisant la complexité de votre écosystème de sécurité.

24_MKTG_Product_WhitePaper_013_Four_Questions_CISOs_are_Asking_r3_fr_02252025