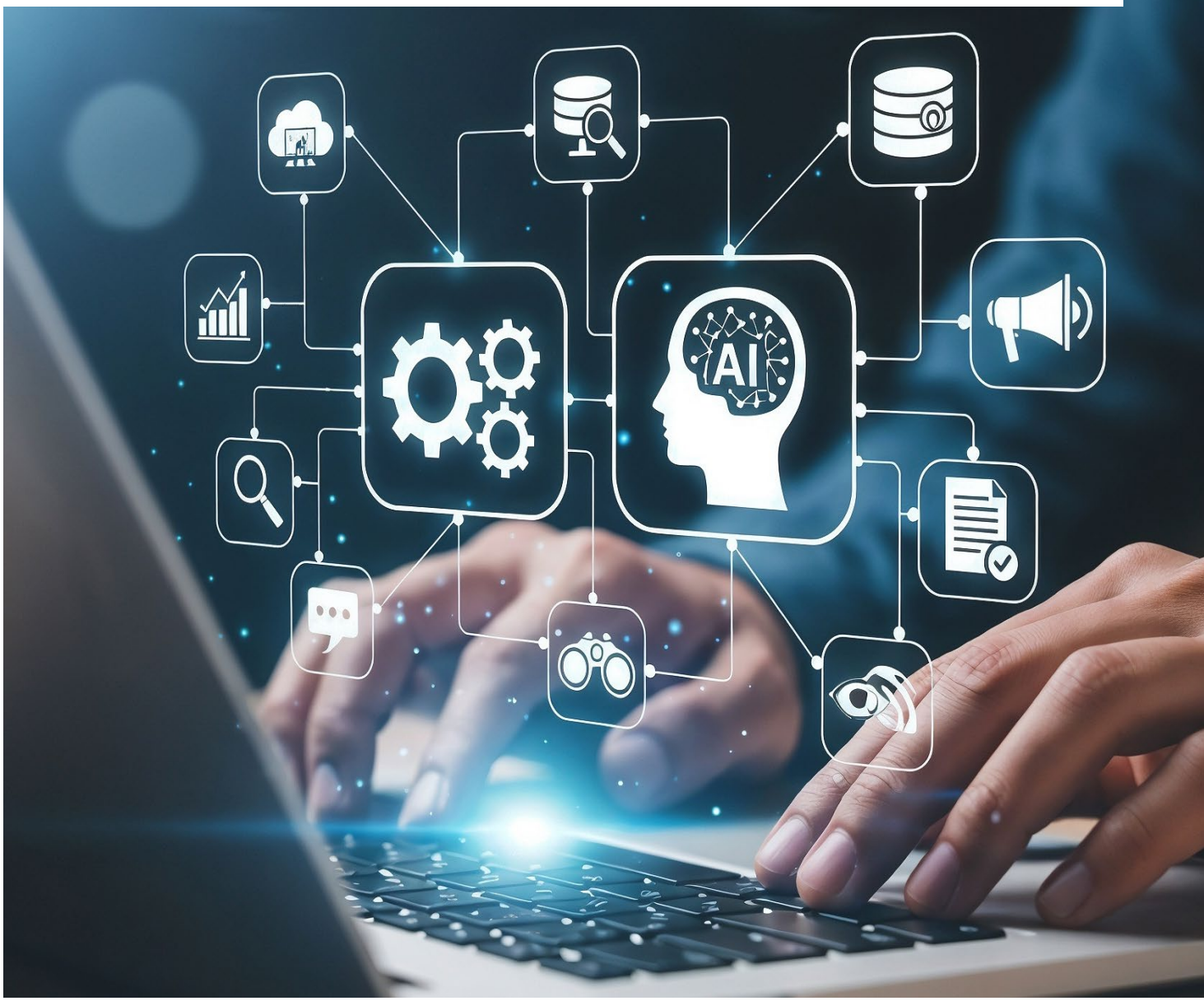
MINISTÈRE
DE L'INTÉRIEUR

Liberté
Égalité
Fraternité

FLASH INGÉRENCE ÉCONOMIQUE DGSI #117

Décembre 2025

RISQUES ASSOCIÉS À L'USAGE DE L'INTELLIGENCE ARTIFICIELLE DANS LE MONDE PROFESSIONNEL



Ce « flash » évoque des actions d'ingérence économique dont des sociétés françaises sont régulièrement victimes.

Ayant vocation à illustrer la diversité des situations auxquelles les entreprises sont susceptibles d'être confrontées, il est mis à votre disposition pour vous accompagner dans la diffusion d'une culture de sécurité interne.

Il est également disponible sur le site internet : **www.dgsi.interieur.gouv.fr**

Par mesure de discrétion, le récit ne comporte aucune mention permettant d'identifier les entreprises visées.

Pour toute question relative à ce « flash » ou si vous souhaitez nous contacter, merci de vous adresser à :

securite-economique@interieur.gouv.fr

RISQUES ASSOCIÉS À L'USAGE DE L'INTELLIGENCE ARTIFICIELLE DANS LE MONDE PROFESSIONNEL

L'émergence et la démocratisation de l'intelligence artificielle (IA) au quotidien, notamment dans le monde de l'entreprise, peut amener à un changement radical de la manière de travailler. Les gains de productivité très importants permis par l'IA en font un outil de plus en plus incontournable. Sous-domaine de l'intelligence artificielle, l'IA générative représente une avancée considérable pour les entreprises, offrant des perspectives inédites en matière d'automatisation, d'optimisation et d'innovation.

Ce développement rapide de l'IA suscite à la fois fascination et inquiétudes. Ces craintes peuvent découler d'une méconnaissance des mécanismes de l'IA, de doutes concernant ses conséquences sur l'emploi, ou encore de préoccupations éthiques liées à la confidentialité et à l'automatisation de la prise de décision.

Dans le cadre de sa mission de sécurité économique, la DGSI accompagne les sociétés et les établissements de recherche dans la prise en compte des risques d'ingérences étrangères liés à ces outils novateurs. Ce « flash ingérence » évoque le cas de trois entreprises françaises ayant été confrontées à des dérives ou des comportements à risque associés ou permis par l'usage de l'IA.

1 DES SALARIÉS D'UNE ENTREPRISE FRANÇAISE STRATÉGIQUE ONT UTILISÉ UN OUTIL D'IA GÉNÉRATIVE POUR TRADUIRE DES DOCUMENTS CONFIDENTIELS

Le directeur des services informatiques d'une entreprise multinationale a appris que certains salariés avaient utilisé un outil d'IA générative « grand public » développé par une société étrangère pour la traduction de documents confidentiels. Après avoir interrogé les équipes, il est apparu que les salariés utilisaient régulièrement cet outil dans le cadre de leurs activités, sans l'aval de leur hiérarchie.

Des consignes ont été rapidement passées afin que les salariés privilégient l'usage d'une solution payante d'IA générative acquise par la société. L'équipe dirigeante a également mis en place un groupe de travail afin de définir une doctrine d'utilisation de l'IA en interne.

Commentaires

Le versement d'informations internes aux entreprises dans un outil d'IA générative, particulièrement si elles revêtent un caractère sensible, constitue un risque important de réutilisation de ces informations pour les entreprises concernées. Les versions grand public des principaux outils d'IA générative, gratuites et standards, utilisent souvent les données entrées par l'utilisateur pour entraîner leurs modèles.

La politique de confidentialité de certains outils d'IA générative impose le stockage des données d'utilisateurs dans des serveurs situés à l'étranger, parfois sans obtenir le consentement clair et explicite

des utilisateurs. Ce stockage implique que ces données puissent être soumises à des lois étrangères à portée extraterritoriale et puissent entraîner la responsabilité de l'entreprise. Par ailleurs, la question de la propriété des données issues de l'IA doit également être étudiée au regard des conditions d'utilisation de l'outil.

La connexion de certaines applications d'IA générative à des outils externes (interface logicielle de type API, plugins) augmente les vulnérabilités. Ces outils sont souvent moins sécurisés, ce qui peut entrainer des fuites de données d'utilisateurs, voire favoriser des attaques cybernétiques.

2 UNE SOCIÉTÉ DÉLÈGUE ENTIÈREMENT L'ÉVALUATION DE SES PARTENAIRES COMMERCIAUX À UN OUTIL D'IA

Dans un contexte de croissance de ses activités sur des marchés étrangers, une société française a mis en place une procédure d'évaluation de ses partenaires commerciaux, ou *due diligence*, dans le but de réduire les risques juridiques, financiers et réputationnels liés à ces relations.

Pour ce faire, la société française a délégué cette tâche à un outil fondé sur de l'intelligence artificielle, créé par une entreprise étrangère, qui lui fournit un rapport d'évaluation sur son potentiel partenaire.

Par manque de temps et par méconnaissance des biais potentiels de l'outil, la société ne procède à aucune vérification complémentaire et oriente systématiquement ses décisions en fonction du retour fait par l'outil.

Commentaires

Afin d'utiliser des outils recourant à l'IA de manière éclairée et responsable, il est nécessaire de comprendre les différents biais auxquels les utilisateurs peuvent être confrontés :

- *une dépendance excessive aux outils d'IA peut diminuer la vigilance humaine ;*
- *les utilisateurs peuvent perdre le contrôle sur leurs données personnelles ou sur les décisions automatisées ;*
- *les IA peuvent reproduire ou amplifier des biais présents dans les données d'entrainement, pouvant mener à des décisions inéquitables ou discriminatoires ;*
- *les décisions prises par certains systèmes d'IA sont parfois qualifiées de « boîtes noires », difficiles à comprendre ou à remettre en cause par les utilisateurs. Cela complexifie la prise de décision éclairée et peut réduire la confiance ;*
- *les IA formulent leurs résultats sur la base de la réponse la plus probable à apporter statistiquement parlant et pas nécessairement la plus pertinente, ni exacte, dans le contexte de la question posée. L'IA va, par exemple, jusqu'à créer des évènements de toute pièce, on parle dans ce cas d'« hallucination ». Cela peut influencer les opinions, manipuler les comportements ou nuire à la réputation.*

3 UNE ENTREPRISE FRANÇAISE A ÉTÉ VICTIME D'UNE TENTATIVE D'ESCROQUERIE PAR HYPERTRUCAGE ASSOCIANT LE VISAGE ET LA VOIX DE SON DIRIGEANT GRÂCE À L'IA

Le responsable d'un site industriel d'un groupe français a reçu un appel en visio-conférence de la part d'une personne se présentant comme étant le dirigeant du groupe. Au premier abord, cet appel n'a pas suscité la curiosité du responsable, l'apparence physique et la voix de l'individu à l'écran correspondant bien à celles du dirigeant.

Néanmoins, l'individu usurpant l'apparence du dirigeant a rapidement demandé au responsable du site de procéder à un transfert de fonds dans le cadre d'un soi-disant projet d'acquisition du groupe.

Surpris par le caractère inhabituel de cette démarche, le responsable du site a mis un terme aux échanges et a alerté la direction de sa société par les canaux habituels. Il lui a été confirmé qu'il avait été victime d'une tentative d'escroquerie par hypertrucage (*deepfake*) associant le visage et la voix du dirigeant grâce à l'usage d'une IA.

Commentaires

L'IA peut être vecteur d'actes d'ingérence élaborés, aux méthodes inédites, commis par des acteurs malveillants. Les attaquants utilisent l'IA pour automatiser leurs attaques en exploitant notamment :

- *la génération automatique de contenus malveillants. Des messages frauduleux, des courriers électroniques de phishing ou de faux sites web sont créés par des IA capables d'imiter parfaitement le style humain, rendant la détection beaucoup plus difficile ;*
- *le* spear phishing *amélioré. Grâce à l'IA, les cybercriminels peuvent analyser rapidement les données publiques et privées d'une cible (réseaux sociaux, courriers électroniques, publications) pour personnaliser leurs attaques, augmentant ainsi les chances de réussite ;*
- *la création de* deepfakes. *Des contenus truqués (vidéos, messages, audios, images, documents, etc.), produits par l'IA, peuvent servir à manipuler l'opinion, commettre des fraudes ou du chantage ;*
- *les attaques par exemples contradictoires (*adversarial examples attack*). Des attaquants peuvent manipuler les données d'entrée d'un système d'IA (par exemple, en modifiant subtilement une image) pour tromper l'algorithme et provoquer des erreurs, voire des comportements dangereux, notamment dans des systèmes critiques, comme les véhicules autonomes ou les dispositifs médicaux ;*
- *l'empoisonnement des données. En injectant de fausses données dans le processus d'apprentissage, les hackeurs peuvent biaiser ou saboter un système d'IA.*

◆ Encadrer l'usage de l'IA au sein de son entreprise

- **Définir les conditions d'usage de l'outil.**
  Comme pour tout outil informatique, le cadre d'emploi de l'IA doit être explicité dans la charte informatique. Cette dernière doit préciser les limites de l'utilisation de l'outil et notamment le niveau de sensibilité des informations qui peuvent être confiées à l'IA. Un document interne sous forme de guide pratique peut être envisagé en complément de la charte informatique afin de décrire ce que les collaborateurs sont autorisés à faire ou non avec un système d'IA générative.

- **Favoriser le recours à des IA génératives françaises.**
  Le recours à des solutions qui hébergent leurs données en France et qui respectent le règlement général sur la protection des données (RGPD) constitue un gage de sécurité et de souveraineté des données internes à l'entreprise.

- **Privilégier l'utilisation d'IA en local.**
  Une IA locale désigne un système d'intelligence artificielle qui fonctionne directement sur le système d'information (SI) de l'utilisateur, sans nécessiter de connexion constante à un serveur externe. Ce type d'IA est installé et exécuté en local ce qui signifie que toutes les données traitées restent sur le SI de l'utilisateur, assurant ainsi une plus grande confidentialité et un contrôle renforcé des données.

- **Former régulièrement ses équipes à l'usage de l'IA.**
  Il est crucial de former ses équipes à l'utilisation de l'IA en organisant des ateliers, aussi bien dans le but de démystifier l'IA, de rassurer les utilisateurs et de s'assurer de l'appropriation de l'outil que de développer une culture de la cybersécurité.

◆ Utiliser l'IA générative de manière raisonnée

- **Être transparent et signaler l'utilisation de l'IA générative à sa hiérarchie ou à son client.**
  En fonction des situations, prévenir du recours à l'IA, par exemple dans le cadre de la réalisation d'un livrable, permet au destinataire de mieux appréhender des erreurs éventuelles ou de mauvaises interprétations.

- **Ne pas soumettre de données personnelles (nom, téléphone, adresse, analyses médicales, photos personnelles, etc.) dans un outil d'IA.**
  Les IA génératives disponibles gratuitement sur Internet collectent les données soumises par l'utilisateur. Il est donc nécessaire d'anonymiser systématiquement les requêtes effectuées. Par mesure de précaution, il est préférable de ne pas utiliser ces outils pour des tâches sensibles, comme la rédaction de comptes-rendus de réunions confidentielles.

- **Faire preuve de vigilance face à la manipulation de l'information et les réponses biaisées.**
  Certains modèles d'IA générative sont conditionnés, dès leur phase d'apprentissage, pour éluder certaines questions ou thématiques, ou encore pour relayer des informations manipulées, voire de la propagande.

- **Vérifier l'exactitude et la pertinence des résultats obtenus en s'appuyant sur des experts qualifiés.**
  Il est fortement déconseillé de fonder des prises de décision uniquement sur les résultats fournis par l'IA, non vérifiés par des personnels qualifiés en interne. Lorsque la taille et les moyens de la société le permettent, la présence d'un data scientist au sein de l'entreprise permet d'apporter une expertise métier afin d'analyser le fonctionnement et la qualité de l'IA utilisée. En effet, les outils d'IA générative étant particulièrement simples d'utilisation, ils peuvent laisser penser à un utilisateur régulier qu'il développe une expertise alors que l'analyse d'un outil d'IA requiert d'importantes compétences techniques.

- **Avertir la DGSI de tout évènement suspect lié à l'utilisation de l'IA.**
  La DGSI se mobilise face aux nouvelles formes d'ingérences étrangères liées à l'IA. Tout évènement suspect peut lui être communiqué sur son adresse électronique dédiée aux sujets de protection économique :
  securite-economique@interieur.gouv.fr

MINISTÈRE
DE L'INTÉRIEUR

Liberté
Égalité
Fraternité