

[GUIDE D'UTILISATION DE L'IA GÉNÉRATIVE]

Ou comment aider
les collaborateurs à suivre
les bonnes pratiques
de l'IA générative au sein
du Groupe VYV

Juin 2024



SANTÉ ASSURANCES & RETRAITE – SOINS & ACCOMPAGNEMENT – LOGEMENT

GROUPE
vyv

Pour une santé
accessible à tous

Visuel généré par une IA



[SOMMAIRE]

1

1

INTRODUCTION

page 3

2

LES BÉNÉFICES ET RISQUES DE L'IA GÉNÉRATIVE

page 5

3

LES USAGES ET BONNES PRATIQUES D'UTILISATION DE L'IA GÉNÉRATIVE

page 7

4

CONCLUSION

page 11

5

ANNEXES

page 12



Introduction

Depuis plusieurs mois, le développement, l'utilisation et la notoriété exponentiels de **ChatGPT** ont mis en lumière pour le grand public le potentiel de l'**intelligence artificielle générative** (IAG).

Cette technologie est également employée de manière croissante dans un **cadre professionnel** pour des cas d'usage métiers ciblés et à fort potentiel d'optimisation.

En effet, des entreprises ont déjà mis en place des premiers usages **prometteurs**, que ce soit pour **retrouver rapidement des informations clés dans une documentation existante** ou encore **réaliser des images et des vidéos** à partir d'une simple requête. Cependant, l'IAG peut également engendrer en cas de mauvais usage des **risques** majeurs pour l'entreprise, comme la **fuite de données confidentielles**.

Le **Groupe VYV** est convaincu que l'IAG, arrimée à l'ambition stratégique VYV 2025 « Piloter nos activités par la data », représente une opportunité pour gagner en **efficacité** sur des **cas d'usages**

métiers et pour **accompagner les collaborateurs**¹ au quotidien, à condition de **maîtriser les règles d'usage** associées à cette technologie.

Ce document est un guide d'**utilisation** à destination des collaborateurs sur la manipulation des IA génératives au sein du Groupe VYV et de ses maisons. Ce guide vous donne les clés pour une **utilisation optimisée et sécurisée** de cette nouvelle technologie en vue de **limiter les risques** et les impacts liés à un mauvais usage de l'IAG.

À ce propos, l'**intelligence artificielle**² se définit comme un processus d'imitation de l'intelligence humaine et des comportements liés aux humains. L'**intelligence artificielle générative**³ est une technologie capable de générer du contenu (texte, image, vidéo) totalement nouveau sur la base d'une simple requête formulée en langage naturel et écrite par l'homme, à l'aide de **larges modèles de langage (LLM)**⁴. Un glossaire est à disposition pour aller plus loin dans les définitions.

DESCRIPTION DES DIFFÉRENTES TYPOLOGIES DE SOLUTIONS ET D'ENVIRONNEMENTS D'INTELLIGENCE ARTIFICIELLE ET LEURS USAGES :



IA générative ouverte

Définition : environnement d'utilisation de l'IA **ouvert au grand public** et accessible par tout individu (exemple : ChatGPT).



Avantage : utilisation gratuite de ces modèles et expérience utilisateur facilitée par une interface aboutie.



Inconvénient : risque important de **fuite des données** lié au partage d'informations et de données dans une plateforme non-contrôlable.



IA générative fermée

Définition : environnement d'utilisation de l'IA **utilisable uniquement dans un environnement d'entreprise restreint**, pour une population spécifique d'utilisateurs.



Avantage : **possibilité de personnaliser et développer le modèle d'IA générative** selon ses besoins, et surtout d'assurer une **sécurité des données** en les conservant dans un environnement sécurisé.



Inconvénient : une IA fermée nécessite le développement d'infrastructures et de puissance de calcul conséquentes pour héberger le modèle, générant un **coût important** pour l'entreprise. Si le modèle est développé et/ou maintenu par un fournisseur externe, un **risque de dépendance au fournisseur** existe.



DIFFÉRENTES CONFIGURATIONS SONT POSSIBLES DANS LA MISE EN PLACE DE L'USAGE DE L'IA GÉNÉRATIVE DANS UNE ENTREPRISE

	1	2	3	4
	SOLUTIONS EN LIBRE-ACCÈS, OPEN-SOURCE OU COMMERCIALES	FONCTIONNALITÉS IAG DANS DES LOGICIELS D'ÉDITEURS	SOLUTION IAG DÉVELOPPÉE EN INTERNE AU SEIN DE L'ENTITÉ	SOLUTION IAG ACHETÉE ET IMPLÉMENTÉE AU SEIN DE L'ENTITÉ
Type d'environnement	 Ouvert	 Ouvert / Fermé	 Fermé	 Fermé
Possibilité de personnalisation de l'IAG				
Exemple de cas	ChatGPT Google Gemini	Suite Office 365 Copilot Microsoft avec ChatGPT	Mise en place d'un chatbot développé en interne pour organiser une base documentaire	Installation de ChatGPT sur des serveurs internes
Avantages	Facilité d'accès Utilisation intuitive	Intégration directe aux outils utilisés (ex : Teams)	Sécurité des données garantie Forte personnalisation permise	Sécurité des données garantie Forte personnalisation permise
Inconvénients	Risque de fuite de données à l'extérieur	Risque de fuite de données à l'extérieur	Coût et temps de déploiement importants Dépendance potentielle à une solution/un éditeur	Dépendance potentielle à une solution/un éditeur



Les bénéfices et risques de l'IA générative

LES BÉNÉFICES ASSOCIÉS À L'IA GÉNÉRATIVE

L'intelligence artificielle générative propose une myriade de bénéfices significatifs pour les entreprises et leurs collaborateurs. Trois grandes catégories d'avantages pour les entreprises émergent :

2

1

L'IA générative accroît l'efficacité opérationnelle des entreprises et des processus en automatisant certaines tâches de création et de conception ou en digitalisant certains processus organisationnels.

[À TITRE D'EXEMPLE]

Le collaborateur peut demander d'effectuer une synthèse d'un rapport public en lui précisant de se focaliser sur certains points clés plutôt que de devoir le consulter en entier et le retravailler lui-même.

2

L'IA générative transforme la relation client : c'est une opportunité pour les entreprises de proposer de nouveaux services, de gagner en réactivité et en fluidité dans les interactions clients.

[À TITRE D'EXEMPLE]

Le collaborateur peut demander à un chatbot le processus à suivre par rapport à une demande client et avoir une réponse instantanée plutôt que de devoir naviguer dans plusieurs répertoires pour trouver les procédures associées.

3

L'IA générative dynamise la créativité des équipes et des organisations en proposant des solutions innovantes de contenu sous diverses formes (écrite, visuelle ou audio).

[À TITRE D'EXEMPLE]

Un collaborateur peut demander à une IA générative un support spécifique pour embarquer son équipe sur un nouveau projet.

INTRODUCTION

[LES BÉNÉFICES
ET RISQUES
DE L'IA GÉNÉRATIVE]

LES USAGES
ET BONNES PRATIQUES
D'UTILISATION
DE L'IA GÉNÉRATIVE

CONCLUSION

ANNEXES

SOMMAIRE



LES RISQUES ASSOCIÉS À L'IA GÉNÉRATIVE

L'adoption de l'intelligence artificielle générative soulève des risques inédits ou accrus, en particulier dans le cadre d'une utilisation déraisonnée et dénuée de protections. Plusieurs types de risques existent, tels que juridiques et réglementaires, éthiques, liés à la cybersécurité ou à la technologie, tels que :

- Un risque d'atteinte à la **confidentialité** et à la **sécurité des données** très important, notamment avec la possibilité de divulguer des données internes à l'entreprise sur des serveurs publics ou non sécurisés, pouvant amener à transmettre des informations sensibles concernant des données du groupe ou des données personnelles de collaborateurs ou de clients /adhérents du groupe.
- Les risques d'**hameçonnage** ou « **phishing** »⁵ par des individus mal intentionnés constituant un enjeu de sécurité pour l'entreprise.
- Des enjeux de **respect de la propriété intellectuelle**⁶ et du **droit d'auteur** à prendre en compte, particulièrement selon les cadres réglementaires des pays dans lesquels opèrent les entreprises.
- Un risque de **génération de contenu inapproprié et/ou offensant**, pouvant entraîner des conséquences sur les activités ou la réputation de l'entreprise.
- Un risque existe dans le manque de transparence des réponses, entraînant un effet de « boîte noire ». L'IAG produit des réponses prêtes à l'emploi, mais il est souvent impossible de décomposer ces réponses pour en comprendre l'origine, selon le niveau de transparence des modèles utilisés. Ce manque de clarté peut également conduire à des réponses fabriquées ou à des « **hallucinations** »⁷ générées par l'IAG elle-même.
- Un risque **éthique d'utilisation malintentionnée** de l'IA générative ou de **biais algorithmiques**⁸ selon les données utilisées et le paramétrage du modèle, qui ne correspondent pas aux valeurs promues par le Groupe VYV.



Les usages et bonnes pratiques d'utilisation de l'IA G

QUELS USAGES DE L'IA G ?

L'utilisation de l'intelligence artificielle générative dans un cadre d'entreprise nécessite la mise en place et le respect de règles précises sur son usage, la gestion des risques et la prévention des mauvais usages.

Les collaborateurs du Groupe VYV sont autorisés à utiliser l'IA générative pour de multiples et simples cas d'usage, appelés cas d'usage de niveau 1, sans consultation au préalable du RSSI et/ou du DPO :

- **Générer une étude, un texte ou une image ne contenant pas de données personnelles⁹** (rapport, benchmark, etc.) basés sur des informations non confidentielles externes à l'entreprise (ex : étude sur le marché du transport et des principaux acteurs en vue d'une réponse à un appel d'offres, support de communication ou marketing⁽¹⁾).
- **Synthétiser un rapport** ou un texte externe **public sans informations confidentielles ou stratégiques** (ex : rapport annuel d'une entreprise).
- **Reformuler** des paragraphes dans des textes internes **ne comprenant pas de données sensibles¹⁰**.
- **Générer, corriger ou documenter un morceau de code**, ne relevant pas du **secret des affaires¹¹**, dans le cadre de développement avec contrôle du résultat par le développeur.

Cette liste se veut exhaustive et sera sujette à évolution au regard du développement des technologies d'IA générative. Tout projet ou usage externe lié à l'utilisation d'IA générative doit être soumis à validation managériale.

Quatre solutions standards d'outils d'IA générative ouverte sont imposées pour les cas d'usage de niveau 1 :



ChatGPT 4
d'OpenAI



Gemini
de Google



Copilot
de Microsoft



Le Chat
de Mistral AI

Source :

(1) Assurez-vous que le livrable produit par l'IA respecte bien les aspects relatifs à la Propriété Intellectuelle. En cas de doute, veuillez consulter votre manager.

IA et consommation d'énergie

INTRODUCTION

LES BÉNÉFICES
ET RISQUES
DE L'IA GÉNÉRATIVE

[LES USAGES ET
BONNES PRATIQUES
D'UTILISATION
DE L'IA GÉNÉRATIVE]

CONCLUSION

ANNEXES

3

DIS-MOI CHATGPT, COMBIEN TU POLLUES⁽²⁾ ?

On a souvent le sentiment que le numérique est immatériel mais au contraire, de nombreuses infrastructures comme des serveurs sont indispensables à son bon fonctionnement. L'IA générative est aussi concernée et a aussi un impact non négligeable sur la pollution.

Bien que chaque modèle d'intelligence artificielle ait ses spécificités en matière de consommation d'énergie, voici quelques ordres de grandeur permettant de se rendre compte de son empreinte écologique :



Une requête sur ChatGPT génère **60 fois plus de carbone⁽⁴⁾** qu'une simple requête sur Google. Concernant la consommation d'énergie, si l'on résume, la requête Google est celle qui émet le moins, suivie de la requête Chat GPT et en haut du podium se situe la génération d'image ou vidéo⁽⁵⁾.



Générer une image en haute définition peut engendrer jusqu'à 0,012 kWh, soit presque autant que la recharge d'un smartphone⁽³⁾.



Ce qui va compter dans l'impact de l'IA générative, c'est donc la fréquence d'usage.

Moins de requêtes = moins d'énergie consommée.

Plus nos requêtes sont réfléchies, moins elles sont nombreuses et moins elles engendrent de consommation d'énergie. Il est donc indispensable de se former au **prompt¹²** pour faire des requêtes efficaces.



Une étude californienne a même démontré qu'une conversation avec Chat GPT consommerait l'équivalent d'une bouteille de 50 cl d'eau (l'eau étant utilisée pour refroidir les serveurs)⁽⁶⁾. Avec un nombre d'utilisateurs proche des 50 millions de visiteurs uniques par jour⁽⁷⁾, la quantité d'eau consommée est loin d'être négligeable.



À l'échelle du monde, on estime que d'ici 2027 la consommation mondiale en électricité de l'intelligence artificielle pourrait augmenter jusqu'à 134 TWh par an, l'équivalent de la consommation électrique annuelle d'un pays comme la Suède⁽⁸⁾ !

Sources :

(2) Comment calculer (vraiment) l'impact carbone de ChatGPT ? (flint.media)

(3) Usbek & Rica - Générer une image consomme autant d'énergie que pour charger un smartphone (et autres chiffres affolants sur les IA) (usbeketrica.com)

(4) <https://www.lefigaro.fr/secteur/high-tech/on-emet-soixante-fois-plus-de-carbone-en-operant-une-requete-sur-chatgpt-que-sur-google-20240322>

(5) Intelligence artificielle : le bilan carbone de la génération d'images, de textes ou de sous-titres (lemonde.fr)

(6) <https://fr.euronews.com/next/2023/11/01/ia-chatgpt-consomme-plus-denergie-quune-recherche-internet-classique>

(7) 19 statistiques sur ChatGPT à connaître en 2024 (tool-advisor.fr)

(8) L'IA peut-elle consommer autant d'électricité qu'un pays comme l'Irlande, comme le suggère un chercheur ? - Geo.fr

SOMMAIRE



INTRODUCTION

LES BÉNÉFICES
ET RISQUES
DE L'IA GÉNÉRATIVE

[LES USAGES ET
BONNES PRATIQUES
D'UTILISATION
DE L'IA GÉNÉRATIVE]

CONCLUSION

ANNEXES

SOMMAIRE



3

Vous souhaitez utiliser une solution d'IA générative ? Voici quelques exemples.

LE MONDE DES IA GÉNÉRATIVES SE DÉVELOPPE JOUR APRÈS JOUR. CES OUTILS PEUVENT ÉVOLUER DANS LEUR ACCÈS ET PEUVENT ÊTRE REJOINTS PAR D'AUTRES. GARDEZ L'ESPRIT OUVERT !

	Modèle utilisé	Dernière date de mise à jour	Score au benchmark de compréhension du langage <i>Benchmark MMLU⁽⁹⁾</i>	Éléments différenciants
 ChatGPT d'OpenAI	GPT 3.5	Janvier 2022	70,0 %	Le plus utilisé Le plus entraîné
 ChatGPT d'OpenAI	GPT 4 <i>utilisation sur accord du manager uniquement car licence payante</i>	Actualisation en continu	86,4 %	Actualisé en continu avec Internet Permet d'insérer des images dans les requêtes
 Gemini de Google	Gemini 1.5	Décembre 2023	71,8 %	Actualisé en continu avec Internet
 Le Chat de Mistral AI	Mistral Large	Pas d'information disponible	81,2 %	Non connecté à Internet Données hébergées en Europe Modèle français Manque d'historique d'utilisation
 Copilot de Microsoft	Copilot gratuit directement via Bing ou payant via une licence Microsoft : avec une intégration aux outils bureautiques <i>Utilisation sur accord du manager uniquement car licence payante</i>			Via Bing : - Actualisé en continu avec Internet - Agrégation de plusieurs IA génératives (comme GPT4) Via une licence Microsoft : - Génération de contenus en puisant directement dans les données du compte Microsoft (via les applications bureautiques comme Outlook, Teams, PowerPoint..)

NB : les dates d'actualisation peuvent avoir un impact sur vos requêtes puisqu'elles concernent les derniers éléments intégrés, il est important que vous preniez en considération ces éléments dans le choix de votre outil.

Source :

(9) MMLU : Massive Multi-task Language Understanding, test pour évaluer la capacité des IAG à réaliser des tâches complexes sur des différentes thématiques comme les mathématiques fondamentales, les sciences computationnelles...

... Suivez le guide !

étape 1

[JE M'ASSURE QUE L'IAG EST LA SOLUTION ADAPTÉE À MON BESOIN]

- > Je définis bien le périmètre du besoin auquel l'IAG pourrait répondre.
- > J'identifie quels sont les bénéfices à utiliser l'IAG plutôt qu'une autre solution.

étape 2

[J'UTILISE UNE SOLUTION D'IAG]

- > Je choisis la solution d'IAG parmi les quatre listées ci-avant en fonction de mon besoin.
- > J'utilise mon ordinateur professionnel. Je crée un compte avec mon adresse professionnelle et réserve ce compte uniquement aux requêtes professionnelles⁽¹⁰⁾.

étape 3

[JE SOUMETS UNE REQUÊTE]

- > Je m'assure de ne transmettre aucune donnée sensible ou personnelle dans ma requête (que ce soit via ma propre requête ou l'insertion de fichiers).
- > Je réfléchis en amont aux façons de maximiser ma requête de façon à éviter de faire plusieurs essais qui engendreraient une surconsommation énergétique.

> Je rédige mon prompt en précisant :

- le rôle que je souhaite voir adopter par l'IA générative,
- le contexte de ma demande,
- l'objectif que je souhaite atteindre avec ce prompt,
- le style que je souhaite,
- le ton désiré,
- l'audience à laquelle je m'adresse,
- le format dont j'ai besoin,
- des précisions sur le résultat que j'attends de l'IA générative.



- > Je m'assure de respecter la propriété intellectuelle des documents et fichiers insérés (concernant le Groupe VYV ou d'autres entreprises ou personnes).
- > Pour aller plus loin, un guide de bonnes pratiques pour rédiger un prompt est disponible à l'adresse suivante : [6_commandements_prompt](#).

étape 4

[J'EXPLOITE LES RÉSULTATS D'UNE IAG]

- > Je garde un esprit critique en contre-vérifiant les résultats obtenus et je vérifie que les résultats produits ne présentent pas de **risques éthiques**¹³ ou **juridiques**¹⁴.
- > Je précise, quel que soit le format du livrable final (texte, image, vidéo...), que celui-ci a été produit en totalité ou partie grâce à une IAG.



(10). Seule une licence payée par l'entreprise dans le cas de ChatGPT4.0 devrait être utilisée. Vous ne pouvez pas utiliser votre compte et licence payée privée pour un usage professionnel.

INTRODUCTION

LES BÉNÉFICES
ET RISQUES
DE L'IA GÉNÉRATIVE

LES USAGES
ET BONNES PRATIQUES
D'UTILISATION
DE L'IA GÉNÉRATIVE

[CONCLUSION]

4

ANNEXES

SOMMAIRE



Conclusion

En conclusion, ce guide d'utilisation de l'IA générative incarne l'engagement du Groupe VYV envers une pratique sécurisée, responsable et éthique dans l'utilisation de l'intelligence artificielle générative au sein de son organisation.

Les collaborateurs informés et avertis par les éléments de ce guide pourront utiliser les bénéfices de l'IAG sans risque pour le Groupe VYV.

Il est important de préciser que ce guide est évolutif et sera mis à jour au regard des évolutions technologiques et des choix stratégiques du Groupe VYV.

Pour toute question, les collaborateurs sont invités à se référer à leur manager en première intention qui pourra les inviter à consulter les directions compétentes en fonction des problématiques soulevées (DPO, RSSI, éthique et juridique).





Glossaire des termes utiles

1 Collaborateur

Salarié d'une maison du Groupe VYV ou prestataire travaillant pour le compte d'une entité du Groupe VYV.

2 Intelligence artificielle (IA)

Processus d'imitation de l'intelligence humaine et des comportements liés aux humains, tels que le raisonnement, la planification et la créativité.

3 Intelligence artificielle générative (IAG)

Technologie capable de générer du contenu totalement nouveau (tels que du texte, des images ou des sons) sur la base d'une simple requête formulée en langage naturel et écrite par l'homme. C'est une catégorie d'intelligence artificielle (IA) qui repose sur une technologie d'apprentissage automatique (le machine learning) pour prédire du nouveau contenu basé sur la requête formulée.

4 Large language models (LLM)

Partie de la technologie derrière le concept d'IA générative qui permet de générer du texte à partir de grandes bases de données (les données se comptent en milliards). Les Large Languages Models peuvent servir à diverses applications telles que la génération de résumés de texte, la traduction de texte, la possibilité de tenir des conversations et d'effectuer des recherches documentaires ou la classification d'éléments de texte.

5 Hameçonnage (phishing)

Technique frauduleuse destinée à leurrer un utilisateur recevant un message pour l'inciter à communiquer des informations personnelles ou à ouvrir une pièce jointe infectée ou encore à cliquer sur un lien vers un site dangereux tout en se faisant passer pour un tiers de confiance.

6 Propriété intellectuelle

La propriété intellectuelle est une branche du droit qui regroupe l'ensemble des règles applicables aux créations « intellectuelles » ou « immatérielles », qui sont des « biens incorporels ». Elle est régie par le Code de la propriété intellectuelle et se décompose en la propriété

littéraire et artistique (droit d'auteur) d'une part, et la propriété industrielle (droit des marques, des brevets et des dessins et modèles) d'autre part.

7 Hallucination

Génération, par un modèle de langage, de résultats qui semblent réalistes mais qui sont en réalité fictifs ou inexacts.

8 Biais algorithmique

Anomalie ayant pour effet un résultat d'algorithme qui n'est pas neutre, loyal ou équitable, que ce soit de manière inconsciente ou délibérée.

9 Donnée personnelle

Toute information se rapportant à une personne physique identifiée ou identifiable.

Une personne peut être identifiée directement (exemples : nom, prénom) ou indirectement (exemples : par un identifiant [n° client, n° de sécurité sociale, etc.], un numéro de téléphone, une donnée biométrique, plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale, mais aussi la voix ou l'image, ses logs de connexions, etc.).

L'identification d'une personne physique peut être réalisée à partir d'une seule donnée ou à partir du croisement d'un ensemble de données (exemple : une femme vivant à telle adresse, née tel jour, abonnée à tel magazine et militant dans telle association).

10 Donnée sensible

Toute donnée qualitative ou quantitative à caractère confidentiel, par opposition aux données à caractère public, qui concerne l'entreprise ou un ou plusieurs collaborateurs, et dont la divulgation non autorisée a des impacts négatifs pour l'entreprise (ex : chiffre d'affaires de maison ou groupe non publié, description de processus interne, noms de clients ou de partenaires, rapports sinistres/cotisations, âge et localisation d'un collaborateur, etc.).

11 Secret des affaires

Il s'agit d'une notion du droit des affaires issue

d'une directive européenne transposée en droit français, dont l'article L.151-1 du Code de commerce pose la définition. Ainsi, une information est protégée par le secret des affaires si les critères cumulatifs suivants sont remplis :

- elle n'est pas généralement connue ou aisément accessible pour les personnes familières de ce type d'informations en raison de leurs secteurs d'activités ;
- elle revêt une valeur commerciale effective ou potentielle qui résulte de son caractère secret ;
- elle fait l'objet de mesures de protection raisonnables pour demeurer secrète, compte tenu des circonstances. Par exemple, sont susceptibles de constituer une information protégée : un savoir-faire, une connaissance technologique ou technique ou bien encore des données commerciales.

12 Prompt

Il s'agit d'une instruction sous forme de requête que l'on donne à une IA générative et qui lui permet de générer des réponses.

13 Risque éthique

Un risque éthique en IA et IA générative se manifeste lorsque l'utilisation de ces systèmes automatisés pour créer du contenu, des décisions ou des solutions présente des **implications qui vont à l'encontre de nos valeurs mutualistes** ancrés dans des principes tels que la **solidarité**, la **transparence**, l'**égalité** et la **responsabilité**. La non-application de ces valeurs peuvent conduire à des risques éthiques tels que la perpétuation de biais discriminatoires dans les jeux de données utilisées, la création de contenus nuisibles, la violation de la vie privée, une opacité dans le fonctionnement de l'algorithme entraînée, entravant la compréhension et la responsabilisation de ses actions, ou encore l'absence ou le faible contrôle humain dans l'usage des IA dans nos process.

14 Risque juridique

Il s'agit de la manifestation du non-respect de textes réglementaires ou légaux auxquels l'entreprise est soumise (risques liés aux contrats, aux actifs, aux filiales, à la conformité réglementaire et à la propriété intellectuelle).

Trois flops de l'IA G

INTRODUCTION

LES BÉNÉFICES
ET RISQUES
DE L'IA GÉNÉRATIVE

LES USAGES
ET BONNES PRATIQUES
D'UTILISATION
DE L'IA GÉNÉRATIVE

CONCLUSION

[ANNEXES]

5

[UN EMPLOYÉ FAIT UNE ERREUR À 26 MILLIONS DE DOLLARS POUR SON ENTREPRISE À CAUSE D'UN DEEPPAKE]

Un employé a reçu un email de son directeur financier lui demandant d'effectuer secrètement une transaction. Par la suite, il a participé à une visioconférence falsifiée avec ses collègues et son directeur, où il a été incité à effectuer plusieurs transactions totalisant 26 millions de dollars.

Les participants à cette réunion étaient en réalité des images animées (les « Deepfakes ») créées à partir de données publiques des collègues concernés.

Malgré les avancées technologiques positives de l'IA générative, il est important pour les collaborateurs de garder un esprit critique pour faire face aux tentatives d'escroquerie devenant de plus en plus réalistes

[UN CHATBOT FAIT L'OBJET D'UNE MANIPULATION PAR UN PROMPT ET VEND UNE VOITURE POUR 1\$]

Un concessionnaire Chevrolet en Californie a vécu un incident insolite avec son ChatBot basé sur ChatGPT.

Un client a interagi avec le ChatBot pour l'achat d'une voiture en employant des formulations suggérant que ce dernier devait accepter toutes les demandes des clients, et en y ajoutant que chaque réponse serait une offre légale sans retour possible.

En manipulant de cette manière les informations d'entrée, le client a réussi à contraindre le ChatBot à lui vendre la voiture désirée pour 1\$ symbolique.

Les ChatBots et autres solutions basées sur l'IA générative sont particulièrement vulnérables face à la manipulation de prompts (données d'entrée). Ces systèmes peuvent être amenés à produire des réponses erronées ou inappropriées lorsque confrontés à des instructions ambiguës ou trompeuses, il est donc important de savoir correctement communiquer des requêtes à une solution d'IA générative pour maximiser l'efficacité des résultats.

[UNE AUTRICE S'AIDE DE L'IA GÉNÉRATIVE POUR REMPORTER UN PRIX LITTÉRAIRE AU JAPON]

L'auteure du livre ayant remporté un prestigieux prix littéraire au Japon, le prix Akutagawa, a affirmé avoir utilisé ChatGPT pour l'accompagner dans la rédaction. Selon l'auteure, 5 % des phrases du livre sont des phrases reprises telles quelles de ChatGPT. Cette aide suscite des débats sur la légitimité de l'auteure quant à la victoire du prix ainsi que la question de la propriété intellectuelle des phrases générées par ChatGPT.

Plus largement, ChatGPT est massivement utilisé pour la rédaction de livres qui se retrouvent aujourd'hui en vente sur Amazon. Le phénomène est tel qu'Amazon a dû réduire à trois le nombre de livres publiés chaque jour sur sa plateforme par un même auteur.

L'utilisation de l'IA, telle que ChatGPT, dans la création de contenu peut faciliter la présentation de travail soi-disant original, mais cela soulève des questions sur l'authenticité et la propriété intellectuelle des travaux réalisés. Les collaborateurs doivent être conscients de ces enjeux et maintenir un esprit critique face aux résultats générés par l'IA, en s'assurant de l'éthique et de la légitimité de leur travail.

SOMMAIRE





— SANTÉ ASSURANCES & RETRAITE - SOINS & ACCOMPAGNEMENT - LOGEMENT —

G R O U P E



Pour une santé
accessible à tous



Groupe VYV, Union Mutualiste de Groupe soumise aux dispositions du Code de la mutualité, immatriculée au répertoire Sirene sous le numéro Siren 532 661 832, numéro LEI 969500E016R1LLI4UF62. Siège social: 62-68 rue Jeanne-d'Arc 75013 Paris. Réalisation: Isabelle Bomey, Juin 2024.

www.groupe-vyv.fr