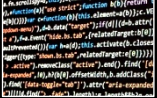
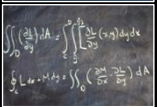
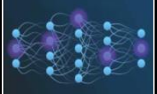


Jean-Jacques ESSOME BELL

Garder le **Contrôle** de l'**I**ntelligence **A**rtificielle



Afin que la machine ne domine jamais l'homme



Douala-Cameroun
Septembre 2024

Introduction

Dans un monde en constante évolution, l'Intelligence Artificielle s'impose comme une force omniprésente, transformant nos vies de manière profonde et souvent inattendue. Si l'IA offre des opportunités sans précédent en nous rendant la vie facile dans de nombreux domaines, elle pose également des défis majeurs et soulève des questions cruciales sur notre capacité à en garder le contrôle. En tant qu'outil puissant, l'IA peut enrichir notre quotidien, mais elle peut aussi, si nous ne sommes pas vigilants, devenir une source d'angoisse et de dépendance ; et cela du fait de ses limites, ainsi que de ses risques et de ses dangers.

Cet opusculé présente de nombreux conseils pratiques et des astuces pour naviguer dans cet univers complexe ; l'objectif étant de répondre à un certain nombre de questions.

- Quels sont les risques, dangers et limites de l'IA dont nous devons être conscients ? Et quels sont ceux qui concernent particulièrement le continent africain ?
- Comment pouvons-nous tirer parti des avancées technologiques, tout en préservant notre autonomie et notre liberté de choix ?
- Comment éviter que nos enfants ne soient en quelque sorte dominés par l'IA ?

En résumé, il s'agit d'explorer les attitudes et réflexes à développer pour que l'Intelligence Artificielle soit sous notre contrôle.



Vous trouverez à l'intérieur...

1



p. 07

Des bienfaits aux limites, risques et dangers de l'IA

2



p. 17

Attitudes et méthodes pour garder le contrôle de l'IA

3



p. 25

Le cas particulier des enfants dans le contrôle de l'IA

4



p. 33

Conseils concernant quelques limites et risques spécifiques

Quelques citations de prise de conscience...

« Le premier risque se présente dès lors que l'IA vous connaît mieux que vous ne vous connaissez vous-même. »

Kai-Fu Lee

« L'IA peut être utilisée à des fins malveillantes, comme la manipulation de l'opinion publique ou le cyberterrorisme. »

Satya Nadella

« Nous avons besoin d'une conversation sur le contrôle de l'IA, car elle transforme notre façon de penser, de communiquer et d'interagir avec le monde. »

Sherry Turkle

« Dans un monde où les enfants sont exposés à des algorithmes dès leur jeune âge, il est essentiel d'enseigner la pensée critique pour naviguer dans cette réalité. »

Katherine Hayles



1.

Des bienfaits aux limites, risques et dangers de l'IA

Des bienfaits aux limites et dangers de l'IA



Malgré ses nombreux avantages et bienfaits, l'IA possède un certain nombre de limites, mais également des risques et des dangers qu'il importe de connaître, et même d'anticiper afin de toujours en garder le contrôle.

1. Les bienfaits de l'Intelligence Artificielle

“L'Intelligence Artificielle a une finalité plutôt simple: nous rendre la vie facile...”



Gain de temps

La machine réalise des tâches plus rapidement que l'homme.

Exemples: résumé de longs documents, traduction de textes, rédactions diverses (articles de blog, CV, etc).



IA



Gain de précision

La machine effectue des tâches avec plus de précision et moins d'erreurs que l'homme.

Exemples: saisie automatique des données, robots conversationnels, etc.



Gain de qualité

La machine effectue des tâches trop complexes pour l'homme.

Exemples: ciblage des clients, robots intelligents dans les usines, etc.



Gain d'argent

L'IA permet de faire des économies et de générer des revenus.

Exemples: opportunités commerciales et financières, initiatives CLVM, optimisation de la main d'oeuvre, etc.

1. Les bienfaits de l'Intelligence Artificielle

Le but ultime de l'**Intelligence Artificielle (IA)** est de nous rendre la vie facile, en faisant exécuter à des machines des tâches qui nécessitent normalement une certaine forme d'intelligence.

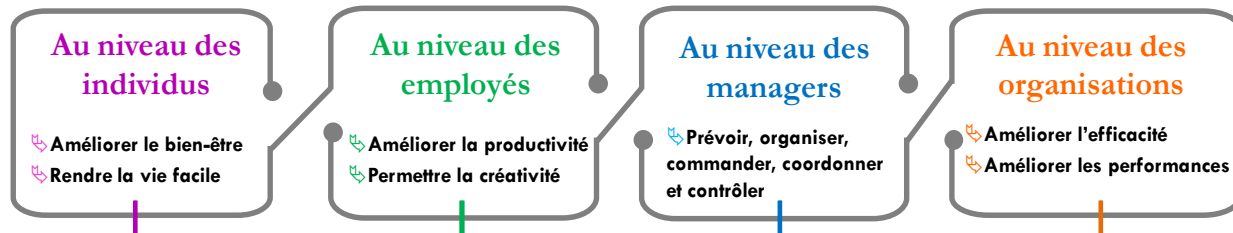
L'IA possède ainsi une multitude d'applications et de cas d'usage, et ses bienfaits apparaissent au niveau de:

► **l'amélioration du bien-être et de la qualité de vie des individus** (santé et forme physique, tâches domestiques, loisirs et divertissements, éducation et culture, économie familiale, communication et relations sociales);

► **l'amélioration de la productivité des employés** en les débarrassant de certaines tâches (génération de textes, analyse des données, etc), afin qu'ils puissent se concentrer sur la créativité;

► **l'assistance aux dirigeants des entreprises** dans leurs activités de management (planification, organisation, commandement, coordination et contrôle);

► **l'amélioration de l'efficacité et des performances des organisations** (gain de temps, amélioration de la qualité, résolution de problèmes complexes, amélioration des performances).



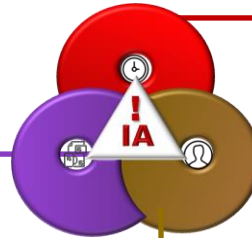
2. Les limites, risques et dangers de l'IA



“Mais l'IA est également une technologie qui comporte une kyrielle de limites et de dangers potentiels...”

Qualité et fiabilité des résultats

- 👉 Biais des données/algorithmes
- 👉 Hallucinations et erreurs
- 👉 Difficulté à assimiler le contexte et les abstractions
- 👉 Opacité de certains algorithmes et résultats
- 👉 Faible créativité et originalité
- 👉 L'effondrement du modèle
- 👉 Limites de la pensée catégorielle et de la rigidité des frontières entre les classes
- 👉 Conséquences des pannes du système



Problèmes d'ordre éthique

- 👉 Faux contenus et deepfakes
- 👉 Utilisation des données personnelles
- 👉 Violation de la vie privée
- 👉 Piratage informatique
- 👉 Exploitation malveillante de l'effet “boîte noire”

Impact social et organisationnel

- 👉 Impact négatif sur les emplois
- 👉 Adaptation organisationnelle des entreprises
- 👉 Grosse consommation d'énergie

2. Les limites, risques et dangers de l'IA

Focus sur quelques limites de l'intelligence artificielle

Les biais de données

L'IA étant essentiellement basée sur les données, les **biais de données** se produisent lorsque les données utilisées pour former un modèle d'IA ne représentent pas fidèlement la réalité. Cela peut se produire pour plusieurs raisons, comme un échantillonnage non représentatif ou des données historiques biaisées.

● Exemple 1: Demande de prêt bancaire

Si un modèle d'IA pour évaluer la solvabilité des clients utilise principalement des données de clients urbains, il pourrait ne pas bien évaluer les emprunteurs ruraux, qui ont des profils financiers différents.

● Exemple 3: Accès aux services de santé

Si un algorithme d'IA est utilisé pour prédire les besoins en soins de santé, il pourrait sous-estimer les besoins des populations rurales si les données de santé proviennent principalement de centres urbains.



● Exemple 2: Reconnaissance faciale

Si un système de reconnaissance faciale est formé principalement sur des images de personnes d'Europe ou d'Amérique du Nord, il aura du mal à identifier des personnes d'ethnies africaines en raison d'un manque de diversité dans les données d'entraînement.

2. Les limites, risques et dangers de l'IA

Focus sur quelques limites de l'intelligence artificielle

Les hallucinations

Les **hallucinations** en IA se produisent lorsque les modèles d'intelligence artificielle génèrent des informations incorrectes ou inventées qui semblent crédibles. Cela peut inclure des faits erronés, des réponses incohérentes ou même des images qui n'existent pas.

Exemple 1: Agents conversationnels

Un *chatbot* pourrait répondre à une question sur l'histoire d'un pays africain en fournissant des informations inexactes ou inventées. Par exemple, il pourrait affirmer qu'un événement historique a eu lieu dans un pays alors qu'il s'est réellement produit ailleurs.

Exemple 3: Systèmes de recommandation

Un système qui recommande des produits ou des services peut suggérer des articles qui n'existent pas, ou donner des descriptions inexactes de produits, ce qui peut frustrer les utilisateurs.

Exemple 2: Génération de contenu

Lorsqu'un modèle d'IA génère des articles ou des rapports, il pourrait inventer des statistiques ou des études qui n'ont jamais été menées. Par exemple, un article sur la santé publique en Afrique pourrait inclure des données fictives sur la prévalence d'une maladie.

2. Les limites, risques et dangers de l'IA

Focus sur quelques limites de l'intelligence artificielle

L'assimilation du contexte ou des abstractions

Les limites de l'IA en matière d'**assimilation du contexte** désignent la difficulté qu'ont les modèles d'intelligence artificielle à comprendre le contexte d'une conversation ou d'une situation. De plus, leur capacité à saisir des **concepts abstraits**, qui nécessitent une compréhension nuancée, est souvent limitée.

● Exemple 1: Assimilation des abstractions

Dans l'exemple ci-contre concernant un entretien de recrutement de chauffeurs pour une entreprise au Cameroun, la compréhension de l'examineur (à gauche) vis-à-vis de la réponse d'un candidat pourtant expérimenté (à droite) sera différente selon que l'examineur vient d'un pays européen ou alors d'un pays d'Afrique Centrale.



● Exemple 2: Nuances émotionnelles

Les modèles d'IA peuvent avoir du mal à détecter des émotions subtiles dans les textes ou la voix.

Par exemple, une conversation où quelqu'un exprime de l'ironie ou du sarcasme peut être mal interprétée, affectant la qualité de l'interaction.

2. Les limites, risques et dangers de l'IA

Focus sur quelques problèmes d'ordre éthique liés à l'IA

Les faux contenus et les deepfakes

Les **faux contenus** désignent les informations fabriquées et en circulation, qui trompent les gens en leur faisant croire qu'elles sont réelles.

Les **deepfakes**, en particulier, sont des vidéos ou des audios générés par l'IA qui imitent de manière réaliste les voix et les visages des personnes.

● Exemple 1: Réseaux sociaux

Sur les réseaux sociaux (Facebook, X, etc), des faux contenus peuvent circuler rapidement, par exemple, des images ou des vidéos montrant des événements qui n'ont jamais eu lieu. Cela peut provoquer des paniques, des violences ou des tensions communautaires.

● Exemple 2: Utilisation malveillante des médias

Des deepfakes peuvent être créés pour déformer l'image de personnes, par exemple, en les plaçant dans des situations compromettantes. Cela peut nuire à leur réputation et à leur bien-être.

● Exemple 3: Manipulation des élections

Avant une élection, des deepfakes pourraient être utilisés pour créer de fausses vidéos de candidats africains faisant des déclarations mensongères. Cela pourrait influencer l'opinion publique et indirectement fausser les élections.

NB: Le mot « **deepfake** » est la combinaison de deux concepts : « **deep learning** » et « **fake** ». En termes simples, il fait référence à de **faux contenus** multimédias (vidéos, images, audios) générés à l'aide de l'apprentissage en profondeur.

2. Les limites, risques et dangers de l'IA

Focus sur quelques problèmes d'ordre éthique liés à l'IA

Les données personnelles et la vie privée

L'utilisation des données personnelles fait référence à la collecte, au stockage et à l'analyse d'informations sur les individus. Les dangers liés à ces pratiques incluent la violation de la vie privée, le vol d'identité et la manipulation des utilisateurs.

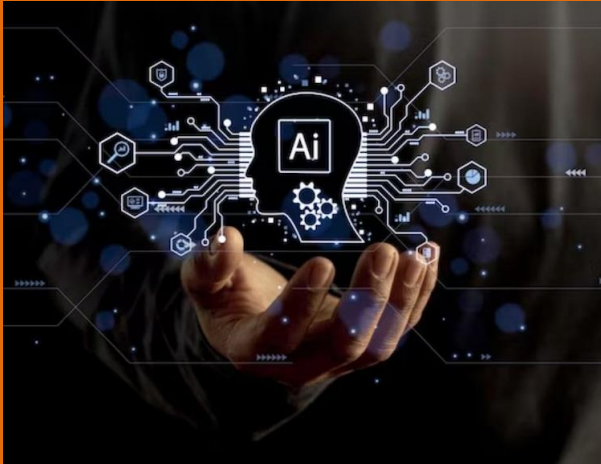
● Exemple 1: Vol d'identité et fraude financière

Si des informations personnelles telles que des données bancaires ne sont pas bien protégées, cela peut entraîner le vol d'identité, où des criminels utilisent ces informations pour accéder à des comptes et effectuer des transactions frauduleuses.



● Exemple 2: Surveillance et contrôle dans les régimes autoritaires

Dans certains pays africains, la collecte de données personnelles par le gouvernement peut être utilisée pour surveiller les citoyens, réprimer les dissidents ou contrôler la population, limitant ainsi la liberté d'expression.



2. Attitudes et solutions pour mieux contrôler l'IA

Attitudes et solutions pour mieux contrôler l'IA



Pour mieux naviguer dans un monde de plus en plus influencé par l'intelligence artificielle, tout en préservant son autonomie et en évitant de se laisser dominer par la technologie/machine, un individu peut adopter plusieurs attitudes et solutions. Celles-ci sont d'une grande diversité, et demandent une pratique régulière.

Attitudes et solutions pour mieux contrôler l'IA

1. Education et sensibilisation



Se former sur le fonctionnement de l'IA, ses capacités, ses implications et ses limites afin de pouvoir prendre des décisions éclairées.

Cela permet de déceler ses limites et ses biais, et d'identifier les applications utiles et celles qui peuvent être problématiques.

Illustration:

Participer à des cours en ligne sur l'IA, lire des livres accessibles sur le sujet, prendre part à des webinaires et conférences.

🔗 **Cours en ligne:** des plateformes comme Microsoft AI, Coursera, edX ou DS-IA Academy proposent des formations sur les bases de l'intelligence artificielle.

🔗 **Webinaires et conférences:** Participer à des événements comme le "AI Expo" ou "Web Summit" pour écouter des experts parler des avancées en IA.

🔗 **Lectures recommandées:** Lire des livres comme:

- « L'Intelligence Artificielle n'existe pas » de Luc Julia
- « L'Intelligence Artificielle pour les nuls » de John Paul Mueller
- « Petit livre de l'Intelligence Artificielle » de Anne-Sophie Baumann
- « L'Intelligence Artificielle pour et partout » de Jean-Jacques Essomé Bell

Attitudes et solutions pour mieux contrôler l'IA

2. Utilisation consciente de l'IA



Renforcer la protection des données personnelles.
Évaluer les outils d'IA que l'on utilise et comprendre leurs implications; ce qui permet de s'assurer qu'ils ne compromettent pas la vie privée ou le bien-être.
Limiter l'utilisation des technologies qui peuvent nuire à la vie privée ou à la santé mentale.

Illustration:

Limiter l'utilisation des réseaux sociaux qui exploitent les données personnelles. Lire les politiques de confidentialité des applications et choisir celles qui respectent des normes strictes. Protéger les données personnelles.

🔍 **Protection des données personnelles:** Utiliser des mots de passe forts (combinant lettres, chiffres et symboles). Procéder à une double vérification lorsque c'est possible en activant l'authentification à deux facteurs (2FA), afin d'ajouter une couche de protection supplémentaire.

🔍 **Paramètres de confidentialité:** Prendre le temps de configurer les paramètres de confidentialité sur les appareils et les applications pour limiter la collecte de données.

🔍 **Applications alternatives:** Choisir des applications comme Signal ou ProtonMail qui mettent l'accent sur la sécurité et la confidentialité des données.

Attitudes et solutions pour mieux contrôler l'IA

3. Collaboration entre humains et IA



Utiliser l'IA pour compléter les capacités humaines et améliorer notre efficacité.

Il s'agit d'utiliser les outils d'IA pour fournir des analyses et des recommandations sans que la machine ne remplace la prise de décision humaine.

Illustration:

Pour chaque domaine d'application ou métier, distinguer clairement ce qui est attendu de l'IA et ce qui doit rester la responsabilité de l'humain.

✎ **IA dans la finance:** un logiciel peut analyser des données financières et alerter les gestionnaires sur des anomalies; mais la décision d'agir revient toujours à l'homme.

✎ **IA dans la santé:** les médecins peuvent utiliser l'IA pour analyser des données médicales, mais la décision finale repose toujours sur leur expertise et leur jugement.

✎ **Utilisation d'outils d'assistance:** Utiliser des logiciels comme Grammarly pour améliorer ses compétences en écriture tout en gardant le contrôle créatif sur le texte.

✎ **IA dans l'éducation:** Les enseignants peuvent utiliser des plateformes d'IA pour personnaliser l'apprentissage des élèves tout en conservant une approche pédagogique humaine.

Attitudes et solutions pour mieux contrôler l'IA

4. Contrôle des résultats de l'IA



Utiliser des solutions d'intelligence artificielle qui offrent des explications claires sur leurs analyses et recommandations. Cela permet de comprendre comment l'IA arrive à certains résultats, et d'évaluer la pertinence de ces résultats.

Illustration:

🔍 **Transparence des algorithmes.** Choisir des solutions d'IA qui offrent des explications claires sur leurs analyses et recommandations. Cela permet de comprendre comment les décisions sont prises et d'évaluer leur pertinence.

🔍 **Vigilance face aux biais.** Être attentif aux biais dans les résultats, en vérifiant si certaines populations ou groupes (femmes, zones rurales, etc) sont défavorisés.

🔍 **Documentation sur la solution.** Lorsqu'elle est disponible, lire la documentation fournie avec la solution d'IA pour comprendre son fonctionnement, ses limites et ses capacités.

Attitudes et solutions pour mieux contrôler l'IA

5. Développement de la résilience



Cultiver une mentalité adaptable et résiliente face aux changements technologiques.

Être prêt à apprendre et à réévaluer ses choix.

Être adaptable face aux changements technologiques permet de mieux faire face aux défis.

Illustration:

Participer à des formations de développement personnel ou de gestion du stress pour améliorer sa capacité d'adaptation et de résilience.

Lire également des livres sur l'évolution des nouvelles technologies et les changements impliqués.

✗ **Formation en gestion du changement:** Participer à des séminaires sur la gestion du changement pour apprendre à s'adapter aux nouvelles technologies.

✗ **Pratiques de mindfulness (pleine conscience):** Intégrer des exercices de méditation ou de pleine conscience dans sa routine quotidienne pour améliorer la résilience mentale face aux perturbations technologiques.

✗ **Ouvrages sur les implications de la technologie:** Il existe des livres intéressants comme:

- « La guerre des intelligences » de Laurent Alexandre
- « La société de la déception » de Dominique Cardon
- « L'IA et le futur de l'humanité » de Maxime Dejean

Attitudes et solutions pour mieux contrôler l'IA

6. Développement d'autres compétences humaines



Développer des compétences qui ne peuvent pas être facilement reproduites par l'IA, telles que la créativité, la pensée critique et les compétences interpersonnelles (empathie notamment).

Illustration:

Prendre des cours de théâtre pour améliorer ses compétences en communication et en empathie ou participer à des ateliers de créativité pour stimuler l'innovation.

✂ **Formation en soft skills:** Suivre des cours en ligne sur des plateformes comme LinkedIn Learning qui se concentrent sur la résolution des conflits, la négociation, la créativité, etc..

✂ **Projets collaboratifs:** Participer à des groupes de travail sur des projets artistiques ou communautaires, où la créativité et la collaboration sont essentielles .



3. Le cas particulier des enfants: contrôler la dépendance à l'IA

Le cas particulier des enfants

Pour aider un enfant à ne pas devenir trop dépendant de l'intelligence artificielle, il est important d'adopter une approche équilibrée et proactive. L'objectif est de cultiver des compétences humaines essentielles et de favoriser des interactions authentiques, tout en utilisant l'IA de manière réfléchie et constructive.

En intégrant les 9 recommandations des pages qui suivent, les parents peuvent aider leurs enfants:

- à développer une relation saine avec la technologie, en les rendant plus autonomes et moins dépendants de l'intelligence artificielle;
- à développer un esprit critique et une créativité durable face à l'intelligence artificielle.

De cette façon , les enfants pourront, non seulement tirer parti des outils technologiques, mais aussi préserver leur humanité et leur originalité dans un monde de plus en plus automatisé.

Le cas particulier des enfants

1. Éduquez l'enfant sur l'IA



Initiation précoce

Expliquez ce qu'est l'IA de manière simple.

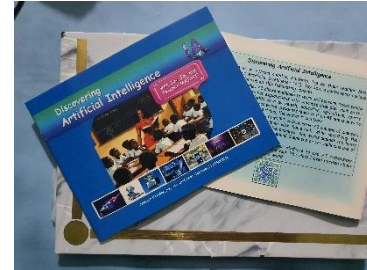
Utilisez des livres illustrés (bandes dessinées) ou des vidéos adaptées aux enfants pour rendre le concept accessible.

Exemple: « *A la découverte de l'Intelligence Artificielle* », de Jay-Jay Essomé Bell

Familiarisation avec les limites

Discutez des limites de l'IA.

Par exemple, « *l'IA ne peut pas remplacer les émotions humaines ni le jugement personnel* ».



2. Fixez des limites d'utilisation



Règles claires

Établissez des règles sur le temps d'écran.

Par exemple, « *pas plus d'une heure par jour* » ou « *pas d'utilisation pendant les repas* ».

Alternatives aux écrans

Proposez des alternatives, comme « *lire des livres* » ou « *jouer à l'extérieur* », pour remplacer le temps passé devant un écran.

Le cas particulier des enfants

3. Eduquez aux médias



Analyse critique

Enseignez à l'enfant à évaluer les informations qu'il trouve en ligne, à distinguer le vrai du faux et à comprendre les intentions derrière les contenus.

Création de contenu

Encouragez-le à créer ses propres histoires ou vidéos; ce qui peut renforcer ses compétences en narration et en expression.

4. Fournissez des modèles de comportement



Exemplarité

Montrez l'exemple en limitant vous-même votre usage des technologies et en privilégiant les interactions humaines.

Partage des expériences

Partagez des moments sans technologie et parlez des bienfaits d'une vie équilibrée.

Par exemple, un week-end ou un pique-nique sans téléphone.

Le cas particulier des enfants

5. Enseignez la vérification des faits



Utilisation d'outils de vérification

Familiarisez les enfants avec des sites comme Snopes, FactCheck.org ou des outils de vérification des faits pour vérifier les informations.

Exemples pratiques

Proposez des affirmations controversées et demandez aux enfants de rechercher leur véracité en utilisant ces outils.

Par exemple, organiser des jeux récompensés sur ce sujet.

6. Encadrez l'utilisation des réseaux sociaux



Responsabilité numérique

Discutez des implications de partager des informations sur les réseaux sociaux.

Montrez comment une information peut être déformée en ligne.

Exercices pratiques

Analysez ensemble des posts sur les réseaux sociaux pour identifier les éléments de désinformation ou de manipulation.

Par exemple, identifier les fake news sur l'actualité sportive.

Le cas particulier des enfants

7. Apprenez-lui à reconnaître les biais



Identification des biais

Expliquez-lui ce qu'est un biais et comment il peut influencer la présentation des informations. Discutez des biais politiques, culturels ou commerciaux.

Par exemple, les biais lors des sondages électoraux ou lors de l'évaluation des performances d'une équipe de football.

Comparaison de sources

Montrez-lui plusieurs articles sur le même sujet provenant de différentes sources pour discuter des différences de ton et de contenu.

8. Encouragez la créativité



Activités manuelles

Proposez-lui des activités artistiques comme la musique, le dessin, ou le bricolage, qui stimulent l'imagination sans technologie.

Jeux de rôle

Organisez des jeux de rôle ou des improvisations qui encouragent la créativité et les interactions sociales.

Le cas particulier des enfants

9. Encouragez les interactions humaines



Jeux en groupe

Favorisez les jeux de société ou les sports d'équipe qui nécessitent la coopération et le dialogue.

Temps en famille

Instaurer des soirées sans écrans pour discuter, jouer, ou cuisiner ensemble, renforçant ainsi les liens familiaux.







4. Conseils concernant quelques risques spécifiques

Conseils concernant quelques risques spécifiques

Les solutions d'Intelligence Artificielle Générative comme Chat-GPT, Gemini, etc, permettent de réaliser de nombreuses tâches (recherche d'informations, assistance à la rédaction, traductions, etc). Mais cela ne va pas sans certains risques dont il faut être conscient; ce qui implique l'adoption de certaines solutions.

Sur un autre plan, la confidentialité des données en entreprise est un enjeu crucial qui est d'ailleurs régulièrement relevé par le public lors de certains séminaires ou de certaines séances de formation sur l'IA. Cet enjeu soulève plusieurs problématiques importantes qui concernent par exemple la protection des données, le respect de la réglementation, la sensibilisation des employés, etc.

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

Les solutions d'Intelligence Artificielle Générative comme Chat-GPT, Gemini, etc, et qui sont basées sur les grands modèles de langage, permettent de réaliser de nombreuses tâches, aussi bien dans le domaine personnel que dans le cadre professionnel:

- recherche d'informations;
- traduction de documents;
- assistance à la rédaction;
- génération d'images;
- synthèse de données;
- planification d'événements;
- création de divertissements, etc.

Les quatre premières tâches constituent les cas d'usage les plus courants, mais également ceux qui comportent le plus de risques. Il est donc important d'être conscient de ces derniers, et d'adopter certaines attitudes.

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

1

Le cas de la recherche d'informations

Il s'agit des requêtes en recherche d'informations sur un sujet donné. La recherche d'informations concerne donc la navigation dans Internet pour fournir des réponses précises, et plus globalement l'extraction de données pertinentes à partir de vastes ensembles d'informations.



Risques

- Production d'informations obsolètes. Par exemple, les connaissances de ChatGPT sont basées sur des données collectées jusqu'en octobre 2023.
- Erreurs et inexactitudes sur des sujets complexes ou spécialisés.
- Réponses biaisées ou incomplètes en raison des biais dans les données utilisées par l'algorithme d'IA.
- Absence (dans certains cas) de références claires (sources) sur les informations fournies.

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

✂ Dans le cas de la **recherche d'informations**, 4 solutions peuvent s'avérer particulièrement utiles.



La vérification des faits

Ne vous fiez pas à une seule source d'informations.

Il s'agit de croiser les informations fournies avec d'autres sources; par exemple en comparant les résultats de plusieurs solutions d'IA (ChatGPT, Gemini, Claude, etc); ou même en faisant une simple recherche sur Google.

Certains outils permettent de vérifier des faits en ligne: FactCheck.org, Snopes.com, Hoax-slayer.com, etc.



L'analyse critique

Évaluez les réponses.

Il s'agit de se poser des questions sur la logique et la pertinence des réponses fournies par l'IA.

Si une réponse semble vague ou incomplète, il faut rechercher des informations complémentaires.

Dans certains cas, vous pouvez également rechercher si des arguments ou des preuves sont fournis; ou même vérifier que le ton est neutre (surtout en ce qui concerne les informations sur des personnalités ou sur des sujets sensibles)

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT



L'utilisation pratique

Utilisez les réponses comme point de départ.

Il s'agit de garder à l'esprit que l'IA est juste une sorte d'aide pour générer des idées. Il faut donc adapter les réponses à votre contexte ou à vos besoins spécifiques, et donc approfondir ces réponses.



La compréhension des limites

Prenez conscience des biais et des dates.

Il s'agit de garder à l'esprit que les données utilisées par l'IA peuvent comporter des biais (par exemple la représentativité concernant les populations africaines), ce qui va se refléter sur les résultats produits.

Il faut également tenir compte des limites de date pour des sujets qui évoluent rapidement. Par exemple Chat-GPT4 se limite aux données produites avant octobre 2023

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

2 Le cas de la traduction de documents

Il s'agit de la conversion de texte d'une langue en une autre en tenant, compte du contexte et en respectant les idiomes et les nuances linguistiques.



Risques

- Production de traduction de traductions littérales qui ne tiennent pas compte des contextes culturels ou des idiomes.
- Mauvaise interprétation des textes ou des phrases ambigus en absence de contexte.
- Traduction incorrecte de certains termes appartenant à des domaines spécialisés.
- Problème de traduction des phrases complexes ou des structures grammaticales inhabituelles.
- Problèmes de traduction des langues avec des alphabets non latins ou des caractères spéciaux, notamment les langues africaines.

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

🔗 Dans le cas de la **traduction de documents**, il est recommandé d'opter pour l'une des 4 solutions qui suivent.



L'utilisation de solutions d'amélioration ou de vérification

Utilisez les solutions dites de plugin, des outils alternatifs ou des outils de vérification.

- **Utilisation des outils alternatifs:** D'autres solutions peuvent être utilisées pour comparer les traductions (par exemple *Google Translate*, *iTranslate*, etc).
- **Utilisation des plugins:** Certaines plateformes intégrant ChatGPT proposent des plugins de traduction dédiés (par exemple *DeepL*).
- **Utilisation des outils de vérification:** Certains outils aident à corriger les erreurs dans les textes traduits (par exemple *Grammarly*).



La facilitation par le contexte

Fournissez un contexte à l'IA.

Il s'agit de donner à l'IA (à travers le prompt) des informations contextuelles (comme le sujet ou le ton) pour l'aider à produire des traductions plus pertinentes.

Vous pouvez également fournir des exemples de phrases traduites pour guider l'IA sur le style souhaité.

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

La vérification humaine

Procédez à la relecture et à la vérification de la traduction produite.

- Vous devez d'abord relire vous-même la traduction si vous avez quelques notions sur la langue de traduction.
- Vous pouvez également faire relire par un locuteur natif (ou toute autre personne compétente), afin de corriger les erreurs et d'améliorer la fluidité.

L'utilisation des glossaires

Utilisez des glossaires pour des termes techniques.

Il s'agit d'utiliser, de créer ou d'intégrer un glossaire pour des termes techniques ou spécifiques à un domaine (télécoms, santé, économie, ingénierie, etc) afin de garantir la cohérence avec la traduction.

La création d'un glossaire suppose que pour chaque terme technique vous fournissiez le mot en français (par exemple), la traduction en anglais (si c'est la langue cible), la définition et un exemple d'utilisation.

Terme en français	Traduction en anglais	Définition	Exemple d'utilisation
AVC	Stroke	Interruption de l'apport sanguin au cerveau	« Il a été hospitalisé après un AVC ».

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

3 Le cas de l'assistance à la rédaction

Il s'agit de la création et de l'amélioration de textes; ce qui inclut la génération de contenu de type rédaction d'articles de blog, production d'un curriculum vitae, rédaction d'une lettre de motivation, etc.



Risques

- Textes manquant d'originalité et de style unique, et pouvant paraître stéréotypés.
- Risque que le texte généré (article, cv, etc) soit similaire à d'autres textes, ou qu'il y ait un soupçon de plagiat.
- Curriculum vitae pouvant être trop généralistes, et manquant de contexte lié au secteur d'activité ou au poste.
- Risque d'erreurs grossières dans les CV (notamment les données techniques), jetant un discrédit sur la candidature.
- Existence d'outils de détection générés par l'IA, entraînant le rejet de certains documents.

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

✂ Dans le cas de l'**assistance à la rédaction**, diverses solutions peuvent se révéler efficaces.



La fourniture de contexte précis

Guidez la solution d'IA et soyez spécifique.

- Donnez des informations plus précises sur le sujet, le public cible et l'objectif du contenu (blog, cv, etc) afin d'améliorer les réponses.
- Partagez des exemples de ce que vous recherchez en termes de style, de ton, etc.



L'adaptation du document (CV, lettre de motivation)

Ajustez le document pour qu'il reflète votre personnalité.

- Modifiez le CV ou la lettre de motivation en mettant en avant des compétences et expériences pertinentes et en lien avec l'offre.
- Incluez des exemples concrets de vos réussites (augmentation des ventes, projets réussis, etc).
- Faites relire votre CV par des amis ou des professionnels du secteur pour obtenir des retours constructifs.

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

4 Le cas de la génération d'images

Il s'agit d'interpréter des instructions textuelles (ou des exemples existants) pour produire des images réalistes ou artistiques, souvent adaptées à des styles spécifiques.



Risques

- Désinformation et manipulation de l'opinion publique par la création de fausses représentations.
- Violation de la vie privée et des droits d'image par la création d'images réalistes de personnes sans leur consentement.
- Reproduction ou amplification des biais présents dans les données d'entraînement, entraînant des représentations stéréotypées ou inexacts de certains groupes.
- Qualité parfois mauvaise ou inappropriée des images générées, nuisant à la crédibilité des contenus visuels.

Conseils concernant quelques risques spécifiques

L'exploitation des résultats des grands modèles de langage comme ChatGPT

🦋 Il est important de pouvoir **reconnaître une image générée par l'IA**.



Quelques astuces pour détecter les images par IA

- **Cherchez les incohérences ou les déformations dans l'image.** Il peut s'agir d'éléments disproportionnés ou asymétriques, des yeux déformés, des ombres incohérentes avec l'éclairage, des objets qui flottent, etc.
- **Regardez attentivement les extrémités du corps humain (les mains et les pieds),** car il s'agit du genre de détail que les IA ont parfois du mal à reproduire de manière parfaite. Il vous faut donc compter le nombre de doigts et d'orteils, leur entrelacement, leur taille.
- **Observez le visage.** Les visages créés par l'IA peuvent parfois présenter un aspect plastifié, avec des reflets exagérés ou une absence de rides sur la peau, rendant le visage étrangement parfait.
- **Faites attention à l'arrière-plan de l'image.** Celui-ci peut en effet comporter des objets complètement flous, ou même des objets qui n'existent pas dans la réalité.
- **Examinez les éléments textuels et les symboles.** Les textes peuvent être incohérents dans leurs lettres ou dans leur disposition. Cela arrive très souvent dans les images où figurent des panneaux de signalisation routière.



Vérification par des outils spécifiques

- | | |
|----------------------|-------------------------|
| • AI or Not | • Deepware Scanner |
| • Deep Fake Detector | • Hugging Face Detector |

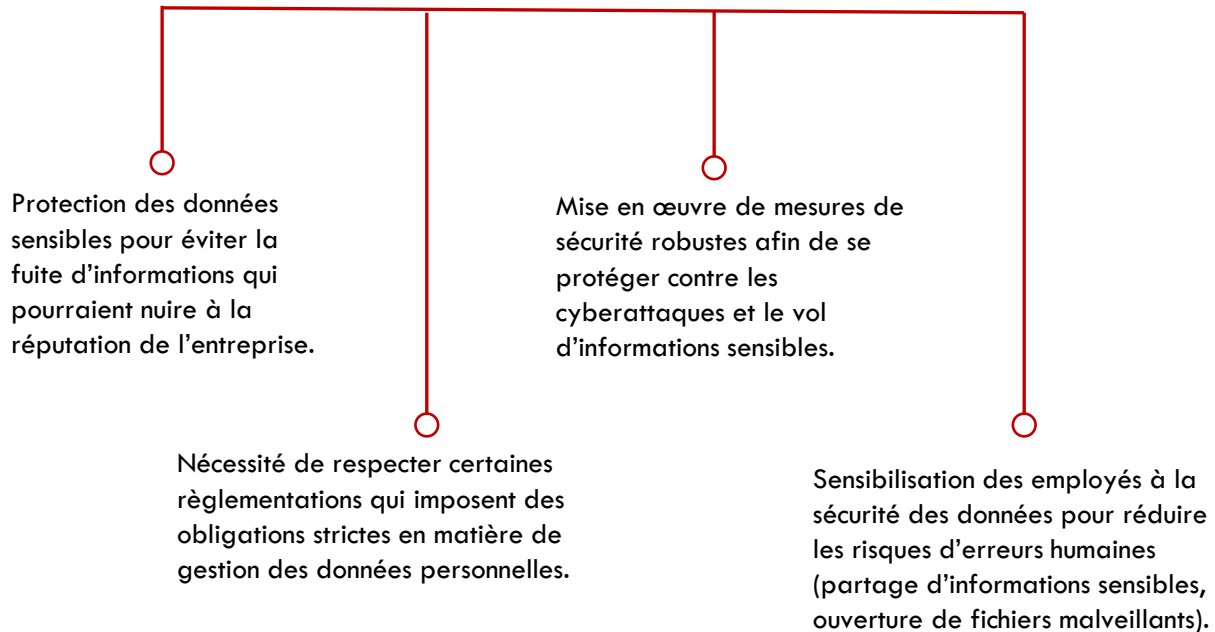
○ Conseils concernant quelques risques spécifiques



Conseils concernant quelques risques spécifiques

Le problème de la confidentialité des données dans les entreprises

🔒 La confidentialité des données en entreprise est un enjeu crucial qui soulève plusieurs problématiques importantes.



Conseils concernant quelques risques spécifiques

Le problème de la confidentialité des données dans les entreprises

Comment les entreprises gèrent-elles le problème de la confidentialité des données?

“Des Comités de Gouvernance sont créés pour gérer les problèmes liés à l'éthique et au respect de la vie privée en cas d'utilisation des données d'un client...”



Politiques de confidentialité

Rédaction de politiques de confidentialité qui expliquent comment les données sont collectées, utilisées et protégées.



Conformité réglementaire

Politiques de confidentialité mises en conformité avec la réglementation, Evaluation des risques, identification des vulnérabilités et mise en place de mesures correctives.



Gestion des accès

Définition de niveaux d'accès pour s'assurer que seuls les employés autorisés peuvent accéder aux informations sensibles.



Technologies de sécurité

Utilisation de techniques telles que:

- le chiffrement des données pour protéger les données sensibles;
- l'authentification multi-facteurs qui nécessite plusieurs vérifications.

Conseils concernant quelques risques spécifiques

Le problème de la confidentialité des données dans les entreprises

Comment les entreprises gèrent-elles le problème de la confidentialité des données?



Gestion des fournisseurs

Exigence de contrats stipulant les responsabilités des fournisseurs en matière de protection des données, et évaluation de la sécurité des pratiques de ceux-ci.



Formation et sensibilisation

Organisation de sessions de formation des employés sur les meilleures pratiques en matière de confidentialité des données et de sécurité informatique.



Anonymisation et minimisation des données

- Anonymisation ou pseudonimisation des données pour réduire les risques en cas de violation.
- Minimisation (collecte uniquement des données nécessaires et suppression en cas de besoin) afin de réduire le volume de données sensibles.

À l'ère de l'Intelligence Artificielle, le défi qui se pose à nous est immense : comment tirer le meilleur parti de ces technologies tout en préservant notre autonomie et notre humanité ?

Cet opusculé a permis de présenter quelques conseils pratiques et des astuces pour garder le contrôle sur l'IA, soulignant l'importance d'une approche consciente et réfléchie.

Il faut garder à l'esprit que l'IA, bien que puissante, n'est qu'un outil. Sa valeur réside dans la manière dont nous l'utilisons. En adoptant certains comportements, nous pouvons transformer l'IA en un allié précieux qui enrichit notre quotidien, tout en évitant les pièges de la dépendance et de la manipulation, ainsi que certaines limites de cette technologie.

D'une certaine manière, la clé réside dans l'éducation et la sensibilisation. En effet, en comprenant mieux le fonctionnement de l'IA, nous pouvons participer activement à son développement et à son intégration dans nos vies.

Et pour terminer, gardons à l'esprit que l'Intelligence Artificielle ne doit pas dominer nos vies, mais plutôt les enrichir et les améliorer, en nous permettant de réaliser notre plein potentiel.