GUIDE N°1

L'USAGE PRUDENT DE L'INTELLIGENCE ARTIFICIELLE EN ENTREPRISE : OPPORTUNITÉS, RISQUES ET BONNES PRATIQUES

MARS | IP

L'intelligence artificielle est un formidable levier de compétitivité, mais elle ne doit pas être utilisée à la légère. Une utilisation responsable et sécurisée est la clé pour éviter les pièges et tirer pleinement parti des avantages de l'IA en entreprise, voici l'objet de ce guide.

Si vous ne donneriez pas une information confidentielle à un inconnu dans la rue, ne la communiquez pas non plus à une IA sans précaution.

MARS-IP - 2025

I - L'une des situations suivantes vous semble-t-elle familière?

L’exemple de DeepL :

L'utilisateur ne se doute généralement pas de l’énorme différence entre la saisie de données dans la version gratuite et dans la version payante, appelée Deepl Pro, en matière de protection des données.

❌ En utilisant la version gratuite, les textes saisis sont stockés et utilisés pour entraîner l’algorithme DeepL. Cela signifie qu’une entreprise qui traduit des documents confidentiels via cette version prend le risque que ses données soient analysées et potentiellement exploitées.

✅ L’abonnement DeepL Pro garantit que les informations saisies ne sont ni enregistrées, ni utilisées à des fins d’apprentissage.

L’exemple de Midjourney :

Un célèbre architecte participant à un concours international met ses croquis sur Midjourney pour illustrer sa candidature et retrouve ses idées inédites reprises par ses concurrents candidats.

L’exemple de ChatGPT :

Des employés de Samsung ont saisi des données confidentielles sur ChatGPT dans le cadre de leur travail. Sachant que ces données sont utilisées par OpenAI pour améliorer leur système et entraîner leurs LLM, Samsung en a interdit ou limité l’utilisation.

Une personne qui ignore ces différences pourrait, sans le vouloir, exposer des informations stratégiques de son entreprise, simplement en utilisant le mauvais outil ou en utilisant un outil de la mauvaise façon.

Le simple fait de copier une information interne dans un traducteur ou autre LLM peut suffire à en perdre le contrôle pour toujours.

II - Les principaux risques liés à l'usage de l'IA en entreprise

Ayant vu ceci, l'IA pose des défis considérables pour la vie moderne en entreprise

1. Risques d'erreurs, d'hallucination et de manque de transparence ⚠️

On ne sait pas sur quoi la plupart des modèles d'IA sont entraînés ni comment leurs algorithmes fonctionnent. Si ces données d'entraînement ne sont pas de grande qualité, les résultats seront discutables, erronés, faussés.

→ Exemple : ChatGPT peut, pour étayer un propos, donner un livre en référence, avec son auteur et même parfois, quelques citations. Après quelques recherches, il peut s'avérer que ce livre n'existe tout simplement pas. C'est une hallucination !

2. Propriété intellectuelle et atteinte aux droits d'auteur 📝

L'utilisation d'IA pour générer du contenu soulève des questions juridiques complexes : les productions générées par des modèles d'IA sont-elles protégées par le droit d'auteur ?

→ Exemple : Une entreprise utilise l'IA pour créer des visuels marketing. Toutefois, la jurisprudence considère que les œuvres générées par IA ne sont pas protégées par le droit d'auteur, ce qui pourrait exposer l'entreprise à des copies sans recours possible.

Les modèles d'IA étant entraînés sur des oeuvres protégées, il serait possible que ces oeuvres soient reproduites par le modèle d'IA. Il y aurait alors une double atteinte des droits d'auteurs car les modèles d'IA s'entraînent sur leurs oeuvres, en général, sans l'autorisation de leurs auteurs, et les plagient ensuite.

→ Exemple : La GEMA, le pendant allemand de la SACEM, qui gère les droits d'exploitation de nombreux créateurs musicaux, a vu certaines des oeuvres de son portefeuille, telles que Big in Japan et Forever Young d'Alphaville, Rasputin et Daddy Cool de Boney M., Cheri Cheri Lady de Modern Talking et bien d'autres reprises et copiées par Suno AI.

3. Perte de confidentialité et savoir-faire 🔓

On sait que les systèmes d'IA stockent les données transmises par les utilisateurs et les utilisent pour entraîner leurs modèles. S'ils stockent et analysent des informations sensibles sans contrôle strict, des secrets d'affaires peuvent être compromis.

→ Exemple : Un employé utilise la version gratuite de ChatGPT pour reformuler un rapport confidentiel. Si le modèle conserve ces données pour améliorer ses réponses, celles-ci pourraient être réutilisées involontairement ailleurs. Pour l'entreprise, l'exclusivité des données pourrait être perdue à tout jamais

Lorsqu'une information confidentielle est transmise à un système d'IA, sa confidentialité est compromise. Le système d'IA pourra ressortir l'information confidentielle avec un prompt adéquat.

4. Perte de contrôle des données personnelles 📂

Les systèmes d'IA nécessitent d'énormes quantités de données pour fonctionner efficacement. Ces informations proviennent aussi bien des usagers par leur utilisation que de données extérieures telles que prises, par exemple, sur des forums.

Une fois qu'une information est intégrée dans le système d'IA, il est pratiquement impossible de la retirer ou de la supprimer complètement. Les entreprises perdent ainsi le contrôle sur leurs propres données.

5. Risques de biais 🎭

Il est important de rester attentif aux réponses des modèles d'IA, car elles peuvent comporter des biais susceptibles d'influencer la prise de décision en entreprise.

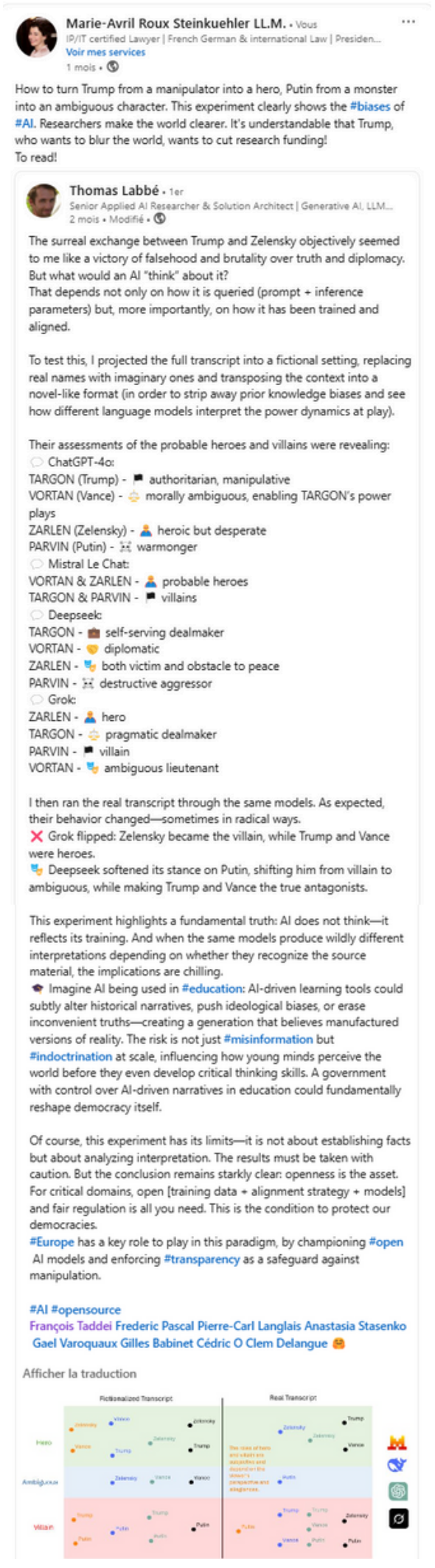
→ Les modèles d'IA peuvent donner une vision différente des leaders politiques actuels en fonction du prompt saisi et du corpus de documents utilisé lors de leur entraînement, comme le démontre ce post LinkedIn.

6. Risques de non conformité juridique ⚖️

L'Union européenne encadre les systèmes IA à « haut risque » par L'IA Act et le traitement des données personnelles par les systèmes d'IA par le Règlement Général sur la Protection des Données (RGPD).

Les entreprises doivent s'assurer que leurs outils IA sont en conformité avec ces réglementations.

→ Exemple : Une entreprise utilisant un système d'IA pour l'analyse des CV pourrait être en infraction si elle ne peut pas expliquer comment l'algorithme prend ses décisions ou si elle ne peut pas supprimer les données d'un candidat à sa demande.



Marie-Avril Roux Steinkuehler LL.M. • Vous
IP/IT certified Lawyer | French German & international Law | Presiden...
Voir mes services
1 mois •

How to turn Trump from a manipulator into a hero, Putin from a monster into an ambiguous character. This experiment clearly shows the #biases of #AI. Researchers make the world clearer. It's understandable that Trump, who wants to blur the world, wants to cut research funding!
To read!

Thomas Labbé • 1er
Senior Applied AI Researcher & Solution Architect | Generative AI, LLM...
2 mois • Modifié •

The surreal exchange between Trump and Zelensky objectively seemed to me like a victory of falsehood and brutality over truth and diplomacy. But what would an AI "think" about it?
That depends not only on how it is queried (prompt + inference parameters) but, more importantly, on how it has been trained and aligned.

To test this, I projected the full transcript into a fictional setting, replacing real names with imaginary ones and transposing the context into a novel-like format (in order to strip away prior knowledge biases and see how different language models interpret the power dynamics at play).

Their assessments of the probable heroes and villains were revealing:
ChatGPT-4o:
TARGON (Trump) - authoritarian, manipulative
VORTAN (Vance) - morally ambiguous, enabling TARGON's power plays
ZARLEN (Zelensky) - heroic but desperate
PARVIN (Putin) - warmonger
Mistral Le Chat:
VORTAN & ZARLEN - probable heroes
TARGON & PARVIN - villains
Deepseek:
TARGON - self-serving dealmaker
VORTAN - diplomatic
ZARLEN - both victim and obstacle to peace
PARVIN - destructive aggressor
Grok:
ZARLEN - hero
TARGON - pragmatic dealmaker
PARVIN - villain
VORTAN - ambiguous lieutenant

I then ran the real transcript through the same models. As expected, their behavior changed—sometimes in radical ways.
❌ Grok flipped: Zelensky became the villain, while Trump and Vance were heroes.
Deepseek softened its stance on Putin, shifting him from villain to ambiguous, while making Trump and Vance the true antagonists.

This experiment highlights a fundamental truth: AI does not think—it reflects its training. And when the same models produce wildly different interpretations depending on whether they recognize the source material, the implications are chilling.
Imagine AI being used in #education: AI-driven learning tools could subtly alter historical narratives, push ideological biases, or erase inconvenient truths—creating a generation that believes manufactured versions of reality. The risk is not just #misinformation but #indoctrination at scale, influencing how young minds perceive the world before they even develop critical thinking skills. A government with control over AI-driven narratives in education could fundamentally reshape democracy itself.

Of course, this experiment has its limits—it is not about establishing facts but about analyzing interpretation. The results must be taken with caution. But the conclusion remains starkly clear: openness is the asset. For critical domains, open [training data + alignment strategy + models] and fair regulation is all you need. This is the condition to protect our democracies.
#Europe has a key role to play in this paradigm, by championing #open AI models and enforcing #transparency as a safeguard against manipulation.

#AI #opensource
François Taddei Frederic Pascal Pierre-Carl Langlais Anastasia Stasenko Gael Varoquaux Gilles Babinet Cédric O Clem Delangue

Afficher la traduction

III - Comment limiter les risques ? Bonnes pratiques à adopter

✅ Privilégier des solutions IA respectueuses des données

Avant d'adopter un outil, vérifier sa politique de confidentialité et s'assurer qu'il ne stocke pas ou n'exploite pas les données sensibles (souvent le cas pour les outils gratuits). Voici un tableau récapitualtif.

Service/LLM	Utilisation des données pour l'entraînement	Exceptions/Contrôle utilisateur tel que l'opt out
OpenAI (grand public)	Oui par défaut	Possibilité d'opt-out
OpenAI (Team/Enterprise/API)	Non	Possibilité d'opt-out
Mistral (gratuit)	Oui par défaut	Possibilité d'opt-out
Mistral (payant)	Non	Données utilisées uniquement pour le service
Azure OpenAI	Non	N/A
Google Gemini	Oui	Non
Google Gemini (entreprise ou éducation)	Non	N/A
Claude IA	Non	Données utilisées uniquement pour le service (commentaires laissés par les usagers et input signalé pour un examen de confiance)
Perplexity	Oui par défaut	Possibilité d'opt-out

✅ Effectuer un audit régulier des outils utilisés

Le marché de l'IA évolue rapidement. Une entreprise doit régulièrement revoir sa stratégie et s'assurer que ses outils respectent toujours les normes en vigueur.

✅ Mettre en place une politique IA en interne

Établir des règles claires sur les outils autorisés, les usages interdits et les précautions à prendre. Cela permet d'éviter des mauvaises pratiques involontaires.

✅ Former les employés aux risques liés à l'IA

Une simple erreur humaine (copier-coller un fichier confidentiel dans un outil IA) peut entraîner une fuite de données irréversible. Des formations régulières sont essentielles.

✅ Adopter une approche hybride

Combiner la puissance de l'IA avec l'expertise humaine est essentiel. Former les employés à reconnaître les erreurs potentielles et à développer leur esprit critique par rapport aux recommandations fournies par l'IA réduit le risque d'erreurs coûteuses.

✅ Informer ses clients de l'utilisation de modèles d'IA

Par souci de transparence, vos clients doivent être informés de votre utilisation des IA.

À suivre :

Le volume 2 du guide sur l'usage de l'intelligence artificielle en entreprise

Le Cabinet d'avocats MARS-IP, spécialisé en droit de la propriété intellectuelle, droit des données à caractère personnel, du numérique et des communications est à votre disposition pour toutes demandes en la matière.

MARS

IP

MARS - IP
Bleibtreustr. 20 - D-10623 Berlin - t: +49 (0) 30 56 55 355 0 - Cell: +49 17 32 30 38 33
26, rue du Quatre-Septembre - F-75002 Paris - t : +33 (0) 1 44 39 49 50 - Cell : +33 6 18 90 20 07
www.mars-ip.eu | mars@mars-ip.eu