



**PREMIER
MINISTRE**

*Liberté
Égalité
Fraternité*

Charte relative à **L'UTILISATION DE L'IA GÉNÉRATIVE** dans les services du Premier ministre

Table des matières

Introduction et définition de l'IA.....	4
Opportunités et risques.....	5
Engagements au sein des SPM.....	8
Utilisations possibles et proscrites au sein des SPM	11
Suivi et mise à jour de la charte	14

Introduction et définition de l'IA

Lexique

IA :

L'intelligence artificielle (IA) est un procédé logique et automatisé reposant généralement sur un algorithme et en mesure de réaliser des tâches bien définies. Pour le Parlement européen, constitue une intelligence artificielle tout outil utilisé par une machine afin de « reproduire des comportements liés aux humains, tels que le raisonnement, la planification et la créativité ».

LLM :

Un Large Language Model (LLM) est une IA entraînée sur une grande quantité de texte pour comprendre ou générer du langage naturel. Il peut répondre à des questions, rédiger des textes ou traduire des langues en s'appuyant sur son apprentissage.

Prompt :

Instruction/question donnée à une intelligence artificielle pour obtenir une réponse ou un résultat précis. C'est une sorte de commande qui guide l'IA dans la génération de textes, d'images ou d'autres contenus.

Inférence :

Étape de « réflexion » de l'IA, quand elle reçoit une consigne (un prompt) et qu'elle produit une réponse en s'appuyant sur ce qu'elle a appris pendant la phase d'entraînement.

RAG :

(Retrieval Augmented Generation) : méthode intégrant une récupération d'informations issues de bases de connaissances spécifiques, afin que l'IA produise des réponses contextualisées. Par exemple le RAG permet notamment d'intégrer le corpus du service utilisateur afin que l'IA s'y réfère prioritairement.

Hallucination :

L'hallucination désigne la production, par un modèle d'intelligence artificielle, de contenus inexacts, non fondés ou fictifs, qui ne sont pas issus des données d'entraînement ni cohérents avec les faits. Ce phénomène est particulièrement fréquent dans les modèles de langage, qui peuvent générer des informations erronées tout en maintenant une forme linguistique fluide et convaincante.

La présente charte vise à promouvoir une **utilisation raisonnée et responsable de l'intelligence artificielle (IA) générative**

grand public au sein des services du Premier ministre (SPM), afin d'en exploiter les nombreuses possibilités, mais également de **comprendre les enjeux, les risques et les bonnes pratiques associés**. Cette charte s'inscrit dans la perspective du déploiement prochain d'IA souveraines pour les services de l'Etat qui lèveront certains des risques ici évoqués.

L'IA générative, aujourd'hui démocratisée par des solutions telles que Le Chat (Mistral), ChatGPT (Open AI) ou Copilot (Microsoft), repose sur des modèles algorithmiques entraînés à la génération de contenus multiples, par l'alimentation de grandes bases de données (corpus d'apprentissage). Ces bases représentant le socle d'apprentissage de la machine, elles lui permettent de prédire statistiquement (et donc par un certain mimétisme) la réponse à une demande formulée en langage courant. **Ainsi, plus l'IA est alimentée par des bases de connaissances approfondies, et confrontée à des questions sur des sujets variés, plus elle devient « compétente ».**

Bien que l'IA générative soit aujourd'hui particulièrement développée et apte à créer du texte, des images, ou même de courtes vidéos, **son apprentissage continu de nouvelles connaissances la fait évoluer constamment, et présente plusieurs risques et limites dont il faut être conscient.**

Les potentialités fonctionnelles de l'IA sont réelles et significatives, à condition d'en mesurer les **impacts en termes de confidentialité, de souveraineté et de sécurité**. De même, **les enjeux humains, budgétaires et environnementaux** doivent être mis en perspective.

La prudence et le respect des recommandations présentées dans cette charte sont nécessaires lors de l'utilisation des solutions accessibles en ligne, dans l'attente du déploiement et de la généralisation d'une IA souveraine.

Pour toute question sur l'IA ou conseil et accompagnement, n'hésitez pas à vous rapprocher de la DSAF à l'adresse : **Dsaf.IA@pm.gouv.fr**.

Opportunités et risques

Opportunités

Le champ des possibilités offertes par l'intelligence artificielle générative est aujourd'hui particulièrement étendu, et les outils utilisant ces technologies apporteront des transformations notables des pratiques professionnelles de chacun. Certaines tâches répétitives pourront à l'avenir être facilitées grâce à l'assistance de l'IA. Chaque agent pourra ainsi concentrer son activité sur son cœur de métier, et sur les tâches pour lesquelles sa propre expertise devra être mise à contribution.

Les apports de l'IA sont aujourd'hui déjà visibles dans la vie quotidienne (réseaux sociaux, commerce en ligne...). Dans le domaine professionnel, chacun d'entre nous sera en mesure d'enrichir ses travaux par des propositions nouvelles, sous l'angle ou à la lumière des nombreuses sources dans lesquelles l'intelligence artificielle est capable de puiser.

S'il est difficile de fixer un horizon sur les apports de l'intelligence artificielle, en constante évolution, certains risques sont déjà identifiés. Aussi, une réflexion et une distance critique sont nécessaires afin d'y avoir recours de manière responsable et sécurisée en tant qu'agent de l'Etat.

Certaines tâches répétitives pourront à l'avenir être facilitées grâce à l'assistance de l'IA.

Risque d'erreur

L'IA n'étant pas dotée d'une capacité de prise de conscience et de compréhension du texte servant à son entraînement ou à une instruction, le contenu qu'elle génère est fortement soumis aux données sur lesquelles la machine s'est exercée. Sans réel contrôle sur les données d'entraînement, il est probable que l'IA fournisse des résultats erronés, comme tout individu pourrait se tromper sur la date d'un événement, ou le rôle d'un personnage historique. Ces erreurs, appelées hallucinations, peuvent être intégrées dans un travail personnel si elles ne sont pas détectées par l'utilisateur de l'IA générative.

Il est probable que l'IA fournisse des résultats erronés, comme tout individu pourrait se tromper sur la date d'un événement, ou le rôle d'un personnage historique.

Opportunités et risques

Il faut donc garder à l'esprit que l'IA est encore loin d'être une source infaillible et intégralement digne de confiance...

Il faut donc garder à l'esprit que l'IA est encore loin d'être une source infaillible et intégralement digne de confiance, et qu'il est nécessaire de systématiquement vérifier la qualité ou la véracité du contenu généré.

Ce risque est particulièrement sensible pour les SPM dans la mesure où il peut engager la responsabilité de l'administration ou nuire à son image.

Risque de discrimination

A l'heure actuelle, le risque de discrimination représente encore l'un des principaux points faibles de l'IA...

Un modèle est également capable d'avantager injustement un groupe d'individus en fonction de plusieurs critères. Les modèles algorithmiques entraînés sont en effet soumis à de nombreux biais, car le corpus d'apprentissage d'une IA a en effet tendance à figer certains stéréotypes, pouvant ensuite influencer sur le fond de certaines réponses générées. Perpétuant des biais déjà présents dans nos sociétés, via les informations ayant servi à son entraînement, ou de manière involontaire par la formulation des requêtes par les usagers, l'intelligence artificielle peut avoir tendance à fonder ses réponses selon des critères basés, à titre exemple, sur l'origine et le sexe (Gendre shades study – 2018).

A l'heure actuelle, le risque de discrimination représente encore l'un des principaux points faibles de l'IA, certaines données historiques la nourrissant contenant des biais reproduits et amplifiés par la machine.

A titre d'exemple, si une IA était utilisée pour trier des candidatures à un poste de cadre dans un service informatique, il est probable qu'elle avantagerait des CV masculins...

A titre d'exemple, si une IA était utilisée pour trier des candidatures à un poste de cadre dans un service informatique, il est probable qu'elle avantagerait des CV masculins car elle se baserait sur l'historique des candidats retenus par le passé pour ce type de poste (en ignorant peut être d'autres variables pour lesquelles elle n'a pas de statistiques : pourcentage de candidats masculins versus féminins, évolutions sociétales...). Inversement, sur un poste

Opportunités et risques

dans les ressources humaines, un profil féminin pourrait être favorisé. Dans les deux cas, le résultat produit par l'IA serait discriminatoire et nous mettrait en défaut vis-à-vis de la réglementation s'il était utilisé.

Plus généralement, l'existence de ces biais peut conduire à des résultats dont l'utilisation sans analyse critique de la part de l'agent peut conduire à porter atteinte à l'obligation de neutralité du service public.

Risque de divulgation

Chaque information (donnée ou document) entrée dans un outil d'IA accessible en ligne alimente un algorithme dont les données sont stockées chez son propriétaire. Il faut donc garder en mémoire qu'envoyer une information vers une IA disponible au public revient à transmettre cette information vers l'extérieur, la faisant sortir de votre organisation.

Par ailleurs, il est aujourd'hui impossible de vérifier les méthodes d'exploitation et de stockage des documents transmis à ces systèmes. Toute information mentionnée est donc susceptible d'être exploitée, revendue, ou d'apparaître dans les résultats d'autres utilisateurs.

En conséquence, seules des données publiques librement utilisables (ou open data) peuvent être communiquées à une IA externe. Par exemple, les versions de travail d'un document amené à être publié ne sont pas considérées comme des données publiques. A l'inverse, des données figurant déjà sur des sites internet de l'administration sont des données publiques.

Seules des données publiques librement utilisables (ou open data) peuvent être communiquées à une IA externe.

Engagements au sein des SPM

Face aux défis auxquels l'IA nous confronte, nous devons nous appuyer sur le socle de valeurs que nous partageons au sein des SPM : neutralité, transparence, respect de la confidentialité, responsabilité, prise en compte des enjeux environnementaux ainsi que concertation et accompagnement.

1. Neutralité

Il convient de garder un œil attentif sur les partis-pris et autres biais employés par l'IA afin d'éviter la production ou diffusion de tout contenu déplacé et non-conforme aux valeurs portées par les services du Premier ministre.

Référence : code général de la fonction publique, article L121-2, obligation de neutralité

Recommandation : privilégier l'utilisation de systèmes d'IA qui indiquent leurs sources.

2. Transparence

L'utilisation de l'IA ne doit pas rester secrète au sein de votre organisation. **La transparence et la communication représentant des piliers indissociables d'un usage responsable de l'IA, votre hiérarchie doit être informée lors de l'utilisation de contenus générés par ces outils.**

Référence : règlement général sur la protection des données (RGPD), code des relations entre le public et l'administration (CRPA).

Recommandation : signaler aux personnes concernées tout usage d'algorithme participant à une prise de décision (aide à la décision).

3. Respect de la confidentialité

Il y a lieu de veiller à ne pas transmettre aux systèmes d'IA publiques des informations qui pourraient contenir des renseignements sur les missions et le fonctionnement interne de votre administration, conformément aux obligations de

secret et de discrétion professionnelles émises par le Code général de la fonction publique.

Référence : RGPD, CRPA, IGI 1300, articles L.121-6 et L.121-7 CGFP.

Recommandation : saisir uniquement des informations relevant déjà du domaine public, dans le respect des obligations de discrétion et de secret professionnel et du devoir de réserve.

4. Responsabilité

Tout agent utilisant l'IA pour produire un document reste responsable de cette production, pouvant engager la responsabilité des services du Premier ministre.

Il convient de garder une pleine maîtrise de l'usage de l'IA, **en veillant au respect des valeurs du service public**.

Recommandation : bénéficier des formations sur l'IA et ses usages mises en place par la DSAF et la plateforme interministérielle de formation MENTOR

5. Prise en compte des enjeux environnementaux

L'IA générative est particulièrement énergivore, tant sur la phase d'entraînement que lors de son utilisation. Son usage doit être mesuré en termes de valeur ajoutée, en prenant en compte l'impact environnemental. Lorsque l'IA générative n'apporte pas de valeur ajoutée par rapport à des outils existants moins énergivores, ces derniers sont à privilégier.

Recommandation : utiliser les outils adaptés pour le besoin recherché (moteur de recherche pour une simple recherche internet, un tableur type Excel pour un calcul). A titre informatif,

l'outil CompareIA de beta.gouv permet d'évaluer la performance énergétique des différents modèles d'IA générative.

Par ailleurs, afin de limiter l'empreinte écologique associée, la mutualisation des ressources et des moyens doit être recherchée.

6. Concertation et accompagnement

Toute expérimentation impliquant l'intelligence artificielle au sein d'un collectif de travail se fait en concertation avec les agents concernés et en associant ces derniers notamment dans la définition des cas d'usages. Cette concertation doit être réalisée en amont de toute mise en œuvre de façon à identifier le cas échéant les besoins de formation et d'accompagnement.

Utilisations possibles et proscrites au sein des SPM

D'un point de vue réglementaire, l'usage de l'IA est encadré par de nombreux textes et chartes. En particulier, il est à rappeler la prise en compte du règlement général sur la protection des données¹ (RGPD) ainsi que du règlement européen sur l'intelligence artificielle² (AI Act). Ce dernier définit notamment, en fonction du cas d'usage envisagé, le niveau de risque afférent.

1. Le règlement général sur la protection des données - RGPD | CNIL

2. Qu'est-ce que l'« AI Act » ? | info.gouv.fr

Cas d'usages auxquels l'IA peut répondre

1. Résumer/rédiger

Les LLM étant spécialisés dans la génération de textes, leur utilisation offre une **économie de temps non-négligeable** pour l'analyse et la synthèse de longs documents non-confidentiels. Ils peuvent également servir d'aide à la génération d'idée, à la rédaction et à la correction de textes.

2. Générer des éléments visuels/sonores

Plusieurs outils d'IA permettent la génération d'éléments visuels variés ou de productions audio (fond sonore, montage, podcast). L'IA peut être ainsi utilisée pour la création de maquettes, de visuels d'illustration, ou d'idées de logos, tant que la plateforme utilisée donne le droit à l'exploitation des éléments générés. Cet usage doit en outre être raisonné, notamment en veillant à la qualité des prompts, dont il est préconisé de faire des brouillons préalables, pour limiter l'empreinte écologique liée aux itérations multiples.

3. Rechercher et s'informer

L'IA peut être utilisée pour **faciliter l'accès à des connaissances spécifiques** mobilisant plusieurs sources. Il est néanmoins **nécessaire d'effectuer une vérification systématique de la validité de ces informations**.

4. Traduire

L'IA permet de traduire de façon instantanée de longs textes. Il convient néanmoins de veiller à ce que les documents ne soient pas confidentiels.

Utilisations proscrites au sein des SPM

1. Déléguer sa production / sa prise de décision

L'IA doit s'inscrire dans un rôle d'assistant et non de remplaçant. Tout travail confié sans contrôle à ces outils est à proscrire afin de garantir la pertinence et la qualité du contenu produit. Garder la maîtrise de son propre travail permet de limiter de potentielles erreurs.

En particulier, la génération de code informatique et la résolution de bugs ne doivent pas être déléguées à un système d'intelligence artificielle dans un souci de confidentialité et de sécurité.

Enfin, l'utilisation de l'intelligence artificielle est inappropriée si elle est amenée à prendre des décisions entièrement automatisées, sans intervention humaine, du fait des erreurs et biais qu'elle peut générer.

2. Plagier

Lié à la délégation de production, le plagiat altère le caractère personnel du travail produit, lorsque certains éléments réutilisés par l'intelligence artificielle relèvent déjà du travail d'autrui. **Il appartient à chacun d'entre nous d'être le garant du respect du droit** - notamment de la propriété intellectuelle - en la matière. Tout élément produit par une IA et intégré dans un travail personnel doit **faire l'objet de mesures spécifiques** (citation des sources, reformulations, mention explicite, etc.) **afin de baliser le document**.

3. Saisir des données personnelles et/ou sensibles

Les IA publiques doivent être considérées comme des « tiers » à qui il n'est pas possible de confier de données personnelles ou sensibles (celles de l'administration ou celles des usagers). Il en est de même pour les données de niveau diffusion restreinte, classifiées ou confidentielles (données contractuelles, juridiques ou financières).

Il incombe à chacun d'éviter de s'exposer, ou d'exposer son organisation, par la communication à l'IA d'informations personnelles/sensibles. L'aval de la hiérarchie doit être obtenu en cas de doute sur la sensibilité des informations que l'on souhaite communiquer au système d'IA.

Suivi et mise à jour de la charte

Cette charte est amenée à évoluer parallèlement au développement des possibilités offertes par l'intelligence artificielle, et du cadre légal/réglementaire entourant ces outils.

JUIN 2025

