

**Charte sur l'usage
de
l'intelligence artificielle générative
au CEA**

PROJET

L'Intelligence artificielle (IA) générative fournit de nouveaux outils pour faciliter le travail des collaborateurs du CEA et leur permettre de dégager du temps pour se consacrer aux activités à plus forte valeur ajoutée.

Afin de tirer le meilleur parti de ces nouvelles possibilités, le CEA met en place une politique et des outils pour en favoriser l'usage avec une approche centrée sur l'humain où l'IA générative est positionnée comme une aide.

Bien que l'IA générative soit un levier de performance, elle présente également plusieurs risques et limites. La présente charte a pour objectif d'encadrer son usage afin de minimiser ces risques et de maîtriser ces limites.

Après en avoir pris connaissance et dans le respect de cette charte et de la charte d'utilisation des moyens informatiques et des services internet au CEA, les salariés travaillant au CEA peuvent utiliser à la fois les services intranet internes d'IA générative (tels que Maïa¹ ou Scott.id²) et pour des données publiques, les services internet (tels que LeChat, ChatGPT, Gemini, Claude, Perplexity).

Les directives et recommandations principales pour un bon usage des services d'IA internes et internet sont les suivantes :

1. Protégez les données sensibles :

- ✓ Le traitement des informations professionnelles non publiques doit être fait exclusivement sur les services internes. Dans le cas des informations classifiées au titre du secret de la défense, elles doivent être traitées uniquement sur l'outil interne DAM Scott.id.
- ✓ N'utilisez pas d'informations confidentielles et en diffusion restreinte ou limitée sans avoir vérifié les conditions d'emploi de ces outils.
- ✓ Sur les services internet ne saisissez que des informations publiques. Le transfert de données CEA est interdit.
- ✓ N'utilisez des données à caractère personnel que dans le strict respect de la réglementation relative à la protection des données personnelles.

2. Gardez la maîtrise des contenus générés :

- ✓ Respectez les principes d'intégrité scientifique et d'éthique.
- ✓ Corrigez attentivement les contenus générés en veillant à ne pas reproduire d'erreurs ou de plagiat.
- ✓ Mentionnez explicitement l'utilisation de l'IA générative et son rôle (par exemple «aide à la rédaction», «génération de schéma», «synthèse») dans vos productions.

Tableau comparatif des usages admis en fonction de la solution utilisée		
Niveau de sensibilité du contenu	Service interne CEA d'IA générative	Service internet d'IA générative
Public	OK	OK
Diffusion limitée*	OK	INTERDIT
Confidentiel (+ portée)**	SELON LES CONDITIONS D'EMPLOI***	INTERDIT
Diffusion restreinte	SELON LES CONDITIONS D'EMPLOI***	INTERDIT
Classifié	INTERDIT SUR LE RESEAU INTRA	INTERDIT
	SELON LES CONDITIONS D'EMPLOI A LA DAM (Scott.id)	

* informations librement partagées en interne CEA, mais dont la reproduction ou la diffusion par un acteur externe au CEA doit faire l'objet d'un accord du CEA.

** Comprend des données personnelles et sensibles au sens du RGPD.

*** Les conditions d'emploi des services précisent les niveaux de sensibilité des données que l'outil permet de traiter.

3. Adoptez une démarche sobre et raisonnée :

- ✓ Utilisez l'IA générative uniquement lorsque cela apporte une réelle valeur ajoutée.
- ✓ Privilégiez les outils internes, frugaux et adaptés aux besoins professionnels.

4. Renforcez vos compétences :

- ✓ Formez-vous régulièrement aux outils et bonnes pratiques liés à l'IA générative.
- ✓ Osez expérimenter ces nouveaux outils, avec l'appui des ambassadeurs Maïa.
- ✓ Partagez vos retours d'expérience avec les autres utilisateurs pour contribuer à un usage collectif efficient, éclairé et sécurisé.

Cette charte sera mise à jour afin de prendre en compte les évolutions législatives, technologiques, les nouveaux risques et opportunités en termes d'usage, et les retours d'expérience des utilisateurs.

1. Plateforme de services d'IA générative qui traite des données non classifiées
 2. Plateforme utilisée par la DAM, qui traite des données classifiées

Introduction

L'IA générative offre au CEA des opportunités majeures pour optimiser ses processus, stimuler l'innovation et renforcer sa compétitivité.

À partir d'instructions formulées par l'utilisateur (les prompts), l'IA générative, une fois entraînée, est en capacité notamment :

- ✓ De générer et améliorer du contenu de textes : correction grammaticale, orthographique, reformulation, traduction, rédaction d'articles, de synthèses ou de comptes-rendus.
- ✓ D'analyser et d'évaluer des documents : dégrossissage de données, analyse préliminaire de rapports, évaluation de candidatures ou de projets.
- ✓ De rédiger des appels à projets, présentation, ou notes internes.
- ✓ D'être une assistance technique : génération de code, aide à la programmation, ou analyse de données pour en extraire des tendances.
- ✓ De créer des contenus multimédia : schémas, graphiques, ou supports audiovisuels.

Elle permet donc d'automatiser les tâches répétitives comme la rédaction ou l'analyse de données, accélérant ainsi l'analyse en amont de la prise de décision et la production de contenus. En recherche, elle permet d'explorer de nouvelles hypothèses, d'accélérer la R&D et de favoriser la créativité pluridisciplinaire.

Stratégiquement, elle renforce l'autonomie technologique du CEA grâce à des outils internes sécurisés comme Maia ou Scott.ia sur les réseaux classifiés de la DAM, en réduisant la dépendance aux solutions internet.

Pour l'ensemble de ces raisons, le CEA souhaite promouvoir une utilisation responsable de l'IA générative par ses collaborateurs au quotidien.

Cette charte vise à les guider dans l'utilisation de ces nouveaux outils, disponibles en interne CEA ou sur internet, en établissant des directives et recommandations regroupées autour de quatre thèmes: la protection des données sensibles, la maîtrise des contenus générés par l'IA, une utilisation sobre et raisonnée de l'IA générative et le renforcement des compétences des collaborateurs.

Focus sur le Règlement (UE) 2024/1689 européen sur l'intelligence artificielle (AI Act)

- ✓ Le [règlement européen sur l'intelligence artificielle](#)³, adopté le 21 mai 2024, également désigné « AI Act », est le premier cadre législatif au monde qui encadre le développement, la mise sur le marché, la mise en service et l'utilisation des systèmes d'IA.
- ✓ Ce texte établit des règles harmonisées dans l'ensemble de l'Union européenne pour garantir que les systèmes d'IA respectent les droits fondamentaux, les valeurs européennes et les exigences de sécurité. Il s'appuie sur une approche basée sur les risques, catégorisant les systèmes d'IA selon leur impact potentiel, de risque minimal à inacceptable.
- ✓ L'AI Act s'applique notamment aux fournisseurs et utilisateurs de systèmes d'IA au sein de l'Union européenne.
- ✓ Ce règlement impose la transparence et le contrôle humain, entre autres principes.
- ✓ Ses principes sont pris en compte dans la présente charte.

Protection des données sensibles lors de l'utilisation de l'IA générative

Opportunités

L'utilisation responsable de l'IA générative permet d'exploiter les données de manière sécurisée pour générer de la valeur : accélération des analyses de documents (appels d'offres, candidatures ou projets scientifiques), aide à la rédaction de rapports, à la préparation d'un dossier de soumission d'un projet à un guichet institutionnel, élaboration de compte-rendu de réunions, traduction, etc.

L'IA générative peut également faciliter la collaboration entre les salariés en fournissant des outils puissants pour l'analyse de données et la génération de contenus. Avec ces outils, nous pouvons non seulement augmenter notre efficacité opérationnelle, mais aussi stimuler la recherche et le développement, voire améliorer la gestion de certaines infrastructures.

Risques

Sans mesures adaptées, l'utilisation de l'IA générative présente des risques liés à la diffusion de données sensibles lors de l'utilisation de services internet d'IA générative qui sont principalement de deux ordres :

- ✓ **Diffusion de données à caractère personnel** : le risque de non-respect du règlement général sur la protection des données à caractère personnel (RGPD), notamment quant aux principes de confidentialité, de consentement ou de droit d'accès et de droit à l'oubli est avéré en cas d'utilisation de données personnelles dans les requêtes soumises à un service internet d'IA générative.
- ✓ **Perte de confidentialité des données d'entrée et risques liés à la propriété intellectuelle et au secret professionnel** : les données confidentielles (sous forme par exemple de textes, photos, codes) peuvent être utilisées par le fournisseur du modèle d'IA générative et aboutir à une perte de contrôle des données, perte de propriété intellectuelle et atteinte au potentiel scientifique et technique ainsi qu'au secret professionnel, notamment dans le cadre des évaluations réalisées sur les dossiers de marché, de candidatures ou d'appels à projets.

Directives

Afin d'utiliser de façon appropriée et sûre les services d'IA générative, il est nécessaire de suivre certaines prescriptions :

- ✓ Sélectionnez l'outil d'IA générative en fonction de la sensibilité des données⁴ à traiter.
- ✓ Pour cela, veuillez-vous référer au tableau suivant et consulter les conditions d'emploi de l'outil concerné accessibles depuis l'outil :

Tableau comparatif des usages admis en fonction de la solution utilisée		
Niveau de sensibilité du contenu	Service interne CEA d'IA générative	Service internet d'IA générative
Public	OK	OK
Diffusion limitée*	OK	INTERDIT
Confidentiel (+ portée)**	SELON LES CONDITIONS D'EMPLOI***	INTERDIT
Diffusion restreinte	SELON LES CONDITIONS D'EMPLOI***	INTERDIT
Classifié	INTERDIT SUR LE RESEAU INTRA	INTERDIT
	SELON LES CONDITIONS D'EMPLOI A LA DAM (Scott.la)	

* Informations librement partagées en interne CEA, mais dont la reproduction ou la diffusion par un acteur externe au CEA doit faire l'objet d'un accord du CEA.

** Comprend des données personnelles et sensibles au sens du RGPD.

*** Les conditions d'emploi des services précisent les niveaux de sensibilité des données que l'outil permet de traiter.

- ✓ Préservez le besoin d'en connaître : assurez-vous de ne pas proposer ou rendre accessible à un tiers non autorisé des données ou informations sensibles.
 - ◇ Le traitement des informations professionnelles non publiques doit être fait exclusivement sur les services internes. Dans le cas des informations classifiées au titre du secret de la défense, elles doivent être traitées uniquement sur l'outil interne DAM Scott.ia.
 - ◇ Ne saisissez sur les services internet que les informations publiques, excluez en particulier de vos requêtes les informations confidentielles, en diffusion restreinte ou limitée, et plus généralement les informations internes au CEA. Le transfert de données CEA est interdit.
- ◇ Soyez vigilant aux informations publiques que vous manipulez sur les services internet car cumulées et croisées dans vos requêtes, elles peuvent constituer des informations sensibles.
- ◇ Sur les services internes, n'utilisez que des outils adaptés aux niveaux de classification de vos données ; vérifiez les conditions d'emplois de ces outils.
- ✓ Utilisez les services internes au CEA pour traiter des informations non publiques. A titre d'exemple, seuls les outils CEA (ex. Scott.ia, Maïa) sont autorisés pour les évaluations concernant les appels d'offres, les candidatures ou les projets scientifiques. Il est formellement interdit d'utiliser des plateformes internet pour ces activités.
- ✓ N'utilisez des données à caractère personnel que dans le strict respect de la réglementation relative à la protection des données à caractère personnel, notamment en vous assurant de l'existence d'une base légale (exemple : consentement, contrat, obligation légale) et, le cas échéant, en réalisant une analyse d'impact lorsque le traitement par IA générative présente des risques élevés.

Recommandations

Pour minimiser les risques et maximiser les avantages de l'utilisation de l'IA générative, voici quelques recommandations :

- ✓ Privilégiez les services IA générative proposés en interne pour vos usages professionnels.
- ✓ Privilégiez l'utilisation de solutions souveraines françaises ou européennes.
- ✓ Vérifiez les clauses générales d'utilisation (CGU) de chaque outil avant utilisation, sachant que la gratuité n'est pas suffisante pour garantir une autorisation du fournisseur pour un usage professionnel.

Opportunités

L'utilisation de l'IA générative ouvre de nouvelles perspectives pour la création et la gestion de contenus administratifs, techniques et scientifiques. Elle permet de produire rapidement des textes, des images, des vidéos et d'autres types de médias, ce qui peut considérablement accélérer les processus fonctionnels et de recherche et de développement.

Les textes produits bénéficient ainsi d'une correction grammaticale et orthographique, avec de meilleures tournures et ce dans différentes langues. En maîtrisant les contenus générés par l'IA générative, nous pouvons non seulement améliorer l'efficacité de nos travaux, mais aussi explorer de nouvelles idées et concepts innovants.

Limites de l'IA générative

Malgré ces atouts, nous sommes confrontés à certaines limitations dans l'usage de l'IA générative :

Phénomènes d'hallucination

L'hallucination en matière d'IA générative se réfère au phénomène de production de réponses crédibles mais factuellement incorrectes ou trompeuses, générées car les systèmes d'IA génératives fonctionnent sur des probabilités et des corrélations statistiques, non sur une compréhension réelle des faits. Cela entraîne des réponses plausibles mais pouvant être erronées.

créer et diffuser volontairement des contenus trompeurs ou manipulateurs, tels que de fausses nouvelles («fakenews»), des hypertrucages («deepfakes», «memes»), ou des informations erronées.

Biais dans les outils

Les modèles et les outils d'IA générative, en raison de leurs modes de conception et de fonctionnement, intègrent par nature de nombreux biais liés à la base d'entraînement, au fonctionnement de l'algorithme ou à l'intervention humaine lors des

Phénomènes de manipulation et de désinformation

Les outils d'IA générative peuvent être utilisés pour

Risques

Nous pouvons aussi être confrontés aux risques suivants :

Atteinte au droit d'auteur lors de la génération de contenu

Les modèles d'IA générative peuvent être entraînés sur des corpus de données contenant des éléments protégés par le droit d'auteur, ce qui peut conduire, à défaut de vérification, à la création de contenus dérivés violant les droits des créateurs originaux.

Risques liés à l'éthique et à l'intégrité scientifique

Ces technologies permettent de créer des contenus textuels, visuels ou audio qui semblent authentiques mais qui sont en réalité générés par des algorithmes. Cela pose notamment des risques significatifs de fabrication, de falsification ou de manipulation de données, ainsi que de plagiat, compromettant ainsi l'intégrité scientifique des processus de recherche.

Directives

- ✓ Lorsque vous utilisez des services ou produits d'IA générative, qu'ils soient hébergés sur internet ou en interne CEA, veillez à l'exactitude et à la pertinence des données intégrées et exercez systématiquement un contrôle des contenus produits en toute responsabilité, en veillant à ne pas reproduire d'erreurs ou de plagiat.
- ✓ Soyez transparent et signalez explicitement l'utilisation de l'IA générative (outil, version, usage...) dans vos productions. En particulier, dans le processus de recherche, toute omission de ces mentions est considérée comme un manquement à l'intégrité scientifique ([Code de conduite européen pour l'intégrité scientifique, édition 23⁵](#)).

Utilisation sobre et raisonnée de l'IA générative

Opportunités

Adopter une démarche sobre et raisonnée dans l'utilisation de l'IA générative permet de combiner innovation et responsabilité environnementale. Cette approche, inscrite dans la stratégie de développement durable du CEA, s'accompagne d'une démarche proactive de mesure de notre empreinte environnementale liée à l'IA générative

(consommation d'énergie, d'eau, de ressources minérales et émissions carbone). Ces évaluations nous permettront d'affiner en continu nos pratiques pour les aligner sur nos engagements de réduction, tout en maintenant notre performance technologique.

Risques

L'IA générative présente un impact environnemental significatif, lié à la consommation énergétique des infrastructures numériques et à l'utilisation de ressources (eau, métaux rares) lors des phases d'entraînement et d'inférence. Une étude récente menée par Mistral AI en collaboration avec l'ADEME et Carbone 4⁶, révèle que :

- ✓ L'empreinte environnementale de l'entraînement de Mistral large 2a généré jusqu'à 20,4 ktCO₂e et a consommé 281 000 m³ d'eau sur 18 mois.
- ✓ Une simple requête de génération d'une

page de texte (400 tokens) via un assistant IA comme Le Chat produit 1,14 gCO₂e et consomme 45 mL d'eau, illustrant l'empreinte cumulative des usages quotidiens.

Recommandations

- ✓ Évaluez systématiquement si une solution plus sobre (moteur de recherche, outil interne léger) peut répondre à votre besoin.
- ✓ Choisissez le modèle adapté : privilégiez les versions légères et spécialisées plutôt que des modèles surdimensionnés.

Opportunités

Renforcer vos compétences en IA générative favorise une utilisation plus efficace et responsable de ces technologies. En acquérant des connaissances spécialisées et en maîtrisant les outils disponibles, vous pouvez non seulement améliorer votre productivité, mais aussi contribuer à des projets innovants.

Cette démarche favorise une meilleure compréhension des enjeux et des limites de l'IA générative, permettant ainsi de formuler des requêtes pertinentes et d'interpréter les résultats de manière critique. En outre, elle encourage la collaboration et le partage des connaissances, renforçant ainsi la culture de l'innovation au sein du CEA.

Risques

Une utilisation non maîtrisée de l'IA générative (mauvaise formulation des requêtes ou interprétation incorrecte des résultats) peut entraîner plusieurs risques, notamment :

- ✓ Les erreurs et biais.
- ✓ Le non-respect de dispositions légales ou réglementaires ou de principes d'intégrité scientifique.
- ✓ Une réduction de la pensée critique et créative.
- ✓ Une dépendance excessive.

Recommandations

- ✓ Formez-vous régulièrement aux outils et bonnes pratiques liés à l'IA générative : une formation sur l'IA générative est proposée par le CEA, permettant d'en appréhender les enjeux et les limites, d'accompagner et d'initier aux usages. De manière générale, suivez des formations adaptées pour maîtriser les outils, formuler des requêtes pertinentes et interpréter les résultats de manière critique. Une utilisation éclairée garantit efficacité et respect des dispositions légales et réglementaires et principes d'intégrité scientifique.
- ✓ Osez expérimenter ces nouveaux outils, avec l'appui des ambassadeurs Maïa et de votre hiérarchie.
- ✓ Partagez vos retours d'expérience avec les autres utilisateurs pour contribuer à un usage collectif efficient, éclairé et sécurisé.

Saclay le