

Version	V4
Date	Février 2026
Statut	Version de travail / Pour adoption SFE

Proposition de principes directeurs de la SFE pour l'usage de l'intelligence artificielle en évaluation.

Cadre général : Application de la charte SFE aux usages d'intelligence artificielle en évaluation

Depuis sa création, **la société française de l'évaluation (SFE) fédère l'ensemble des acteurs impliqués dans les pratiques évaluatives**. Tout en reconnaissant la diversité de leurs approches, la SFE promeut les démarches les plus à même de faire de l'évaluation des politiques publiques une composante essentielle de la vie démocratique et du fonctionnement des organisations publiques.

Adoptée en 2003 et actualisée en 2006, **la Charte de la SFE constitue le socle éthique et déontologique de cette communauté**. Ses sept principes fondateurs – Indépendance, Transparence, Pluralité, Respect des personnes, Compétence, Responsabilité et Utilité – guident les pratiques évaluatives et conservent toute leur pertinence à l'ère du numérique. Les principes directeurs pour l'IA déclinent la Charte SFE, dans un ordre de présentation spécifique aux enjeux liés à l'IA sans modifier la hiérarchie fondatrice des principes de la Charte SFE.

Le développement rapide des technologies d'IA, notamment générative, ouvre de nouvelles possibilités mais comporte des risques (biais, opacité, dépendance technologique). Leur usage suppose une vigilance accrue en matière d'équité, de responsabilité et de justice sociale.

Face à ces évolutions, **ces principes directeurs accompagnent la mise en œuvre éthiques des technologies d'intelligence artificielle en évaluation**. Ils précisent les exigences applicables vers lesquelles tendre pour un usage conforme aux principes généraux de la charte éthique de la SFE, en sus du respect des obligations contractuelles, des chartes internes des organisations et des cadres juridiques applicables (notamment RGPD, règlement européen pour l'IA, recommandations de la CNIL, propriété intellectuelle, ...).

Ces principes directeurs **s'adressent à toutes les personnes et institutions participant au processus d'évaluation** des politiques publiques, désignés sous l'expression "acteurs de l'évaluation".

L'adhésion aux principes directeurs vaut engagement à les mettre en œuvre. Ces principes s'appliquent à toutes les structures, quels que soient leur taille ou leur niveau de maturité. Lorsque certaines exigences ne peuvent être pleinement satisfaites à court terme, la structure s'engage à en reconnaître explicitement les limites et à définir une trajectoire de mise en conformité adaptée à son contexte.

1. Principe de Responsabilité

La responsabilité des productions évaluatives demeure pleinement humaine, même lorsque des outils d'intelligence artificielle ont été mobilisés.

Les personnes et institutions participant au processus d'évaluation à titre professionnel assument la qualité, l'intégrité et la pertinence des livrables, des analyses et du jugement porté sur le dispositif, la politique ou le programme évalué, quelle que soit la contribution des outils d'IA dans leur élaboration. Elles s'assurent que les usages des outils d'IA ne sont pas de nature à porter préjudice à l'action publique, au domaine de l'évaluation des politiques publiques et à celles et ceux qui en font leur métier. Cette responsabilité inclut la garantie que les usages des outils d'IA n'engendrent pas d'inégalités ou de traitements discriminatoires dans l'évaluation.

Pour cela, les acteurs de l'évaluation s'engagent dans une démarche de mise en œuvre des actions suivantes :

- Identifier dans son organisation les enjeux spécifiques liés aux outils d'IA utilisés, qu'il s'agisse d'impacts sur les pratiques, la qualité des évaluations ou les conditions de travail des professionnels.
- Superviser rigoureusement chaque production issue d'un outil d'IA, en s'appuyant sur les meilleures pratiques de contrôle pour chaque cas d'usage (par ex. : relectures systématiques, vérifications croisées avec d'autres sources, contrôles aléatoires par sondage, évaluation régulière des modèles utilisés, etc.).
- Sécuriser les traitements en choisissant des outils conformes aux standards de cybersécurité (ex. : chiffrement des données, fiabilité des infrastructures, garanties de sécurité, etc.).
- Participer activement aux échanges collectifs visant à faire évoluer la présente version des principes directeurs, en partageant les retours d'expérience et les difficultés rencontrées dans l'usage des outils d'IA.

2. Principe de Respect des personnes

L'évaluation respecte les droits, la dignité et la sécurité des individus. Les personnes et institutions participant au processus d'évaluation ont la responsabilité de protéger les données personnelles et prévenir toute forme de discrimination algorithmique.

Le respect des personnes ne se limite pas à la conformité au RGPD et à l'évaluation des risques au sens du règlement IA. Il implique une vigilance active face aux biais algorithmiques, aux asymétries informationnelles et aux risques d'instrumentalisation ou de standardisation excessive des expressions individuelles.

Pour cela, les acteurs de l'évaluation s'engagent dans une démarche de mise en œuvre des actions suivantes :

- Vérifier les conditions d'usage et de conservation des données transmises aux outils d'IA en ligne. Une vigilance renforcée est requise pour les outils gratuits ou freemium,

dont les conditions d'usage peuvent inclure la réutilisation des données ou des transferts hors UE.

- Obtenir un consentement éclairé des personnes concernées, en précisant clairement les finalités et modalités d'utilisation des outils d'IA (ex. : « Vos réponses pourront être analysées par un outil numérique d'analyse automatisée, dans les conditions suivantes : [...] »).
- Privilégier les solutions locales ou souveraines, mieux encadrées sur le plan juridique et éthique, et du contrôle des transferts de données.

3. Principe d'Opportunité

Les outils d'IA ne sont utilisés que s'ils apportent une plus-value démontrée, dans un esprit de frugalité et de proportionnalité visant à un usage raisonné. Les personnes et institutions participant au processus d'évaluation à titre professionnel justifient l'usage des outils d'IA par des gains concrets (efficacité, exhaustivité, etc.), en tenant compte de ses impacts sociétaux.

Pour cela, les acteurs de l'évaluation s'engagent dans une démarche de mise en œuvre des actions suivantes :

- Éviter le recours à des outils d'IA lorsqu'une méthode plus « traditionnelle » pourrait être plus simple, robuste et moins consommatrice de ressources.
- Bien distinguer les usages à titre d'expérimentation sur les outils d'IA, de l'usage commun intégré au processus évaluatif.
- Maintenir des alternatives non-IA opérationnelles pour chaque étape clé du processus, afin de garantir la continuité de l'évaluation en cas de défaillance ou d'indisponibilité des outils d'IA.
- Documenter les gains attendus et obtenus grâce à l'usage des outils d'IA, qu'ils soient en termes de temps, de précision, de fiabilité ou de nouvelles perspectives d'analyse.
- Choisir uniquement des outils offrant des garanties reconnues de fiabilité, de neutralité et de maîtrise de leurs impacts sociétaux et environnementaux, en privilégiant les solutions transparentes et auditable.

4. Principe de Transparence

L'usage d'outils d'intelligence artificielle doit être **explicite** et **compréhensible** par toutes les parties prenantes. Les personnes et institutions participant au processus d'évaluation ont la responsabilité de rendre compte de manière claire et accessible du **rôle joué par ces outils dans le processus d'évaluation**, afin de renforcer la confiance et la lisibilité des résultats.

Pour cela, les acteurs de l'évaluation s'engagent dans une démarche de mise en œuvre des actions suivantes :

- Distinguer clairement les contenus et productions intégrant une contribution des outils d'IA de ceux réalisés exclusivement par les intervenants humains, en particulier pour les tâches liées au jugement évaluatif (référentiel, conclusions).
- Engager une discussion avec les parties prenantes sur le niveau de transparence attendue.
- Se mettre en capacité de mentionner dans les productions (cahiers des charges, référentiels, rapports, etc.) :
 - o Les outils utilisés (nom, version, fournisseur) ;
 - o La typologie des données d'entrée mobilisées et leur provenance ;

- Lorsque pertinent, les instructions (prompts) utilisées dans le cadre d'un traitement automatisé ;
 - Les vérifications et validations réalisées par des évaluateurs humains.
- Envisager et expérimenter l'usage de mentions ou pictogrammes standardisés pour signaler la nature des contributions des outils d'IA.

5. Principe de Distanciation

L'évaluation doit préserver son indépendance intellectuelle et opérationnelle, y compris vis-à-vis des fournisseurs d'outils d'IA. Les personnes et institutions participant au processus d'évaluation à titre professionnel exercent une posture critique et distanciée, en considérant systématiquement les biais propres aux technologies mobilisées ainsi que les choix (techniques, économiques, politiques ...) des acteurs qui les conçoivent et diffusent.

Pour cela, les acteurs de l'évaluation s'engagent dans une démarche de mise en œuvre des actions suivantes :

- Identifier et déclarer les biais des outils utilisés, en cherchant à comprendre leur fonctionnement et leurs données d'entraînement (ex. : surreprésentation de certains groupes dans les bases de données).
- Documenter les limites des outils (ex. : "Ce modèle ne prend pas en compte les contextes locaux spécifiques").

6. Principe de Pluralité

L'évaluation repose sur des combinaisons d'outils et de méthodes mobilisant l'expertise humaine, éventuellement assistée par des outils automatisés ou recourant à des modèles d'apprentissage machine (IA). Aucune décision ou évaluation ne peut reposer uniquement sur des résultats générés par un traitement automatisé, sans validation ou interprétation par des experts humains qualifiés.

7. Principe de Compétence

L'usage des outils d'IA générative exige des compétences spécifiques, actualisées et exercées avec responsabilité. Les personnes et institutions qui y recourent doivent développer les compétences nécessaires pour être en mesure de les utiliser à bon escient, d'en interpréter correctement les résultats et d'en critiquer les limites.

Pour cela, les acteurs de l'évaluation s'engagent dans une démarche de mise en œuvre des actions suivantes :

- Se former régulièrement aux outils mobilisables en évaluation, tant sur le plan éthique et technique que sur la compréhension de leurs conséquences sociétales, notamment leur impact environnemental.
- Accorder une attention particulière à l'accompagnement des nouveaux professionnels (notamment juniors, stagiaires, alternants, ...) pour favoriser leur apprentissage et leur montée en compétences dans un cadre éthique et sécurisé.
- Préserver et renforcer les compétences cœur de métier (analyse qualitative, jugement professionnel), telles que définies dans le référentiel des compétences de la SFE, actualisé en 2025.
- Évaluer régulièrement les compétences en matière d'IA, par des dispositifs d'autoévaluation et de revues par les pairs.

8. Principe de Soutenabilité environnementale (nouveau principe)

Les outils d’IA générative amènent à étendre le principe de responsabilité au-delà du propos initial de la charte de la SFE, sur la question environnementale.

L’usage des outils d’IA doit tenir compte de son empreinte écologique et privilégier des solutions sobres, proportionnées aux besoins réels. Les personnes et institutions participant au processus d’évaluation à titre professionnel ont ainsi la responsabilité de réduire l’impact environnemental de leurs usages et de favoriser des pratiques numériques frugales et responsables.

Pour cela, les acteurs de l’évaluation s’engagent dans une démarche de mise en œuvre des actions suivantes :

- Privilégier les modèles légers, dits « frugaux », adaptés à l’usage visé, plutôt que des systèmes très puissants dont les capacités dépassent largement les besoins du projet.
- Limiter les usages énergivores, par exemple en évitant la multiplication de générations ou de calculs pour de simples variantes mineures, ou encore la répétition de mêmes prompts sur plusieurs outils, sans valeur ajoutée.
- Mesurer autant que possible l’impact environnemental de leurs pratiques, en s’informant sur la consommation énergétique des outils utilisés.

Avertissement en accord avec le principe n°5 de transparence :

Des outils d’intelligence artificielle ont été utilisés pour suggérer des idées, reformuler certains passages et vérifier la cohérence du texte. Leur usage est resté limité à un rôle d’appui technique : le contenu final résulte d’une réflexion collective et itérative, discutée par plusieurs instances de la SFE.

Outils utilisés	• ChatGPT – modèle o5 – OpenAI – licence Teams sans réutilisation des données. • Mistral Large – Mistral AI – accès plateforme/API sans réutilisation des données. • Gpt-oss – modèle open source sur machine local sans transmission de données.
Données transmises aux outils d’IA	• Charte éthique de la SFE (adoptée en 2003, actualisée en 2006). • Autres contenus originaux provenant exclusivement de la réflexion des membres du groupe de travail SFE.
Instructions (prompts) et nombre d’interactions	• Les instructions ont porté sur chaque principe indépendamment, à plusieurs étapes de la rédaction (idéation, reformulation, contrôle de cohérence).
Vérifications et validations humaines	• Tous les textes repris d’une production synthétique ont été relus plusieurs fois par plusieurs personnes différentes.
Consommation	• Estimation de 36 000 tokens utilisés (22000 sur gpt5 et 14 000 sur Mistral Large) • Soit, d’après les indicateurs de l’étude MISTRAL¹ d’aout 2025 : 91 grammes de Co² et 4 litres d’eau consommée. • Soit, d’après l’estimateur Compar:la du gouvernement français (Beta.gouv)² : 400 grammes de Co² pour 9 Watt/heure.

¹ Notre contribution pour la création d’un standard environnemental mondial pour l’IA | Mistral AI. URL : <https://mistral.ai/fr/news/our-contribution-to-a-global-environmental-standard-for-ai> [consulté le 11 septembre 2025].

² Beta Gouv. compar:IA, le comparateur d’IA conversationnelles. URL : <https://www.comparia.beta.gouv.fr/> [consulté le 11 septembre 2025].