

Charte pour l'usage et le déploiement de systèmes d'intelligence artificielle de Pluricité et Emoha

Version 1 – Juillet 2025

Statut du document

Ce rapport présente la première version de notre charte d'usage de l'IA. Élaborée à partir d'expérimentations internes et d'un retour critique sur nos pratiques, elle sert de cadre pour guider et contrôler nos usages actuels. Cette charte marque un premier jalon important ; elle annonce une régulation future plus exigeante que nous souhaitons construire et partager avec notre communauté professionnelle.

Licence du contenu

Le texte de cette charte est sous licence Creative Commons Attribution (**CC BY-SA**).

Cette licence permet aux réutilisateurs de distribuer, de remixer, d'adapter et d'exploiter le matériel sur n'importe quel support ou dans n'importe quel format, à condition d'en attribuer la paternité au créateur. La licence autorise l'utilisation commerciale. Si vous remixez, adaptez ou développez le matériel, vous devez accorder une licence pour le matériel modifié selon des conditions identiques.

Sommaire

PREAMBULE	2
DEFINITIONS	3
1. CHAMP D'APPLICATION	5
2. CADRES DE REFERENCE	6
3. APPLICATION DES CHOIX ETHIQUES	10
4. VUE D'ENSEMBLE DES PRINCIPES APPLICABLES AU CAS GENERAL	16

Préambule

Depuis 2022, la montée en puissance des IA génératives de textes rend incontournable la prise en compte des systèmes d'IA dans l'élaboration et l'évaluation des politiques publiques, ainsi que dans l'évolution des métiers de notre secteur. Malgré l'instabilité et l'incertitude qui caractérisent encore ces technologies, il est essentiel d'y réfléchir dès maintenant pour anticiper leurs effets transformatifs sur les pratiques de l'étude et du conseil et ce d'autant plus que les usages de l'IA se répandent dans la mise en place des services publics.

Cette charte constitue une première tentative d'encadrement des usages internes et des productions réalisées à l'aide des systèmes d'IA au sein de Pluricité.

Elle s'inspire des principes de l'UNESCO sur la régulation de l'IA¹ et s'inscrit dans une perspective où Pluricité est à la fois utilisateur, facilitateur, acheteur, développeur, et, dans une certaine mesure, acteur de régulation via ses activités de réseaux et auprès de la Société Française de l'Evaluation (SFE).

Elle est à portée interne, et externe. En interne, elle est destinée à guider les usages et pratiques des personnels de Pluricité. A portée externe, elle présente au tiers les modalités d'utilisation ou non de l'IA, en respect notamment du principe de transparence.

Conçue pour un usage interne et externe, la charte oriente les pratiques du personnel de Pluricité, tout en offrant aux tiers une vision transparente sur les modalités d'utilisation de l'IA. Elle se situe dans le cadre légal français et européen, en particulier le Règlement européen sur l'IA (RIA) et en respectant le RGPD.

Dans une approche d'amélioration continue, la présente charte engage le Groupe Pluricité à développer un ensemble de processus pour accompagner cette démarche déontologique :

- déclinier les valeurs de Pluricité en principes d'action adaptés à nos différentes activités, dans le respect des principes démocratiques,
- accompagner les salariés et partenaires de Pluricité dans la déclinaison de la charte par la mise à disposition des moyens internes (par exemple : outils, méthodologies, procédures, bonnes pratiques, actions de sensibilisation, etc.),
 - suivre l'efficacité des actions prises dans le temps et les communiquer
 - partager en interne et en externe l'expérience acquise dans la prise en compte opérationnelle de ces principes,
 - participer aux débats de société, aux groupes de réflexion (publics, privés) et aux instances de normalisation sur ces questions, pour accompagner l'évolution des technologies et leurs prises en compte dans les secteurs d'activité de Pluricité.

Enfin, dans une optique de transparence, de responsabilité et de sécurité, cette charte fixe les règles d'usage des ressources informationnelles et communicationnelles. Elle veille à une utilisation loyale et conforme de l'IA aux intérêts de l'organisation et de la société.

¹ UNESCO, « Consultation paper on AI regulation: emerging approaches across the world », 16 août 2024, <https://www.unesco.org/en/articles/unesco-launches-open-consultation-inform-ai-governance>.



Définitions

Désignations

- **Utilisateur** : personne physique ayant accès au système d'information de l'entreprise.
- **Entreprise, ou Direction** : gérants de l'entreprise
- **Service informatique** : personnes, services internes ou prestataires externes habilités par l'entreprise pour la gestion du système informatique
- **Référent IA** : Personne désignée par la direction pour exercer un rôle de point focal sur les questions d'IA dans l'entreprise.
- **RIA (règlement européen sur l'IA)** : RÈGLEMENT (UE) 2024/1689 DU PARLEMENT EUROPÉEN ET DU CONSEIL du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle
- **RPGD (règlement européen sur la protection des données)** : RÈGLEMENT (UE) 2016/679 DU PARLEMENT EUROPÉEN ET DU CONSEIL du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données,

Définitions techniques générales et concepts généraux

- **IA (Intelligence Artificielle)** : Technologies informatiques capables d'accomplir des tâches impliquant normalement l'intelligence humaine comme l'apprentissage, la résolution de problèmes ou la reconnaissance de schémas (tout ce qui se rapporte à la parole, à l'analyse d'images, à l'élaboration de stratégie dans des jeux, etc.).
- **IA générative** : Branche de l'IA qui produit du contenu original (texte, images, musique, etc.) en se fondant sur des données d'entraînement, plutôt que de se limiter à analyser ou classer l'existant. (par exemple GPT4, Groq, Gemini, etc.).
- **Explicabilité de l'IA** : Capacité à comprendre et expliquer comment un modèle d'IA prend ses décisions, afin de garantir la transparence et la confiance des utilisateurs.
- **Équité de l'IA** : Capacité des systèmes d'IA à produire des résultats justes, non discriminatoires et équilibrés, en évitant les biais pouvant affecter certaines populations ou groupes d'individus.
- **Intelligence Artificielle "digne de confiance" (trustworthy)** : Approche consistant à comprendre quels sont les risques associés à un déploiement de l'IA et comment les gérer en pratique pour garantir un processus irréprochable de l'usage de l'IA (notamment dans les services publics).
- **Hallucination** : désigne une réponse incorrecte, inventée ou trompeuse générée par un modèle d'IA, même si elle semble plausible.
- **Biais algorithmique** : une distorsion ou une partialité dans les décisions ou les résultats générés par une intelligence artificielle, notamment due aux données utilisées pour entraîner le modèle ou à la manière dont l'algorithme fonctionne.

Définitions issues des règlements européens

- **«Bac à sable réglementaire de l'IA»** ou « sandboxes », un cadre contrôlé mis en place par une autorité compétente qui offre aux fournisseurs ou fournisseurs potentiels de systèmes d'IA la possibilité de développer, d'entraîner, de valider et de tester, lorsqu'il y a lieu en conditions réelles, un système d'IA innovant, selon un plan du bac à sable pour une durée limitée sous surveillance réglementaire

- **Déploieur** : une personne physique ou morale, une autorité publique, une agence ou un autre organisme utilisant sous sa propre autorité un système d’IA sauf lorsque ce système est utilisé dans le cadre d’une activité personnelle à caractère non professionnel
 - **Données personnelles** : Conformément à l’article 9 du RGPD, est une donnée personnelle « toute information se rapportant à une personne identifiée ou identifiable, directement ou indirectement, notamment par référence à un identifiant, tel qu’un nom, un numéro d’identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale ».
 - **Modification substantielle** : une modification apportée à un système d’IA après sa mise sur le marché ou sa mise en service, qui n’est pas prévue ou planifiée dans l’évaluation initiale de la conformité réalisée par le fournisseur et qui a pour effet de nuire à la conformité de ce système aux exigences énoncées par le règlement IA de la Commission Européenne, ou qui entraîne une modification de la destination pour laquelle le système d’IA a été évalué ;
 - **Risque** : la combinaison de la probabilité d’un préjudice et de la sévérité de celui-ci ;
 - **SIA (système d’IA)** : un système automatisé qui est conçu pour fonctionner à différents niveaux d’autonomie et peut faire preuve d’une capacité d’adaptation après son déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu’il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels.
- *Pour aller plus loin : le lecteur qui souhaiterait se familiariser plus avant avec les termes en rapport avec l’Intelligence Artificielle pourra notamment se reporter aux définitions proposées par le règlement européen pour l’IA (article 32), ou le Glossaire proposé par la CNIL³.*

² <https://eur-lex.europa.eu/eli/reg/2024/1689>

³ <https://www.cnil.fr/fr/intelligence-artificielle/glossaire-ia>



1. Champ d'application

Utilisateurs concernés

Sauf indication contraire, la présente charte s'applique à l'ensemble des personnes ayant accès au système d'information et de communication de l'entreprise, quel que soit leur statut : salariés, intérimaires, stagiaires, partenaires externes ou visiteurs occasionnels. Chaque utilisateur veille à ce que toute personne à qui il permet l'accès à ce système prenne connaissance et accepte expressément les règles énoncées dans la présente charte.

Durée et perspectives

Le domaine de l'IA étant en constante évolution, de nombreuses expérimentations sont susceptibles d'avoir lieu. Par conséquent, **cette charte est établie pour une durée de 18 mois**, au terme de laquelle elle fera l'objet d'une révision. Cette démarche doit s'inscrire dans la perspective, à moyen terme, de l'adoption d'une charte contraignante, engageant davantage l'entreprise en interne et vis-à-vis de tiers.

Délimitation des pratiques

La charte distingue trois formes de pratiques relatives à l'IA de manière à clarifier les responsabilités et aider à se conformer aux bonnes pratiques :

- **Les pratiques de déploiement** : définies au sens du RIA, elles concernent la mise en place et l'intégration de systèmes d'IA dans l'activité de l'entreprise.
- **Les pratiques d'expérimentation** : elles portent sur les phases de test, d'analyse et de prototypage de systèmes d'IA avant leur adoption ou mise en production.
- **Les pratiques d'usage** : elles correspondent à l'utilisation courante, récurrente ou ponctuelle, des systèmes d'IA dans le cadre des activités professionnelles. Ces pratiques ont vocation à être portées à la connaissance des tiers, notamment des commanditaires et partenaires de mission de Pluricité.

Type de pratique	A quel moment ?	Qui est concerné ?	Objectif principal
Déploiement	Avant la mise en service d'un système d'IA	Équipes techniques, Direction	Intégrer des systèmes d'IA pour les activités dans le respect de la charte IA
Expérimentation	Phase de test et d'évaluation	Développeurs, expérimentateurs internes, Direction	Mettre à l'étude et vérifier la pertinence, la performance et la conformité des systèmes d'IA envisagés
Usages courants	Utilisation au quotidien	Equipes de pluricité, Ensemble des personnes ayant accès au système d'information et de communication de l'entreprise	Utiliser l'IA de manière responsable dans l'activité quotidienne

2. Cadres de référence

Cadre légal

La charte IA s'inscrit dans la perspective du **Règlement européen sur l'IA (RIA)**, qui classe les systèmes d'IA selon différents niveaux de risques. Adopté en mai 2024, l'entrée en vigueur du RIA se fait de manière **progressive** (§179) :

- **1^{er} août 2024** : Entrée en vigueur officielle du règlement. À partir de cette date, les États membres disposent d'un délai pour mettre en place les structures nécessaires à son application.
- **Février 2025** : Application des interdictions relatives aux systèmes d'IA présentant un risque inacceptable. Cela inclut, par exemple, l'interdiction de la notation sociale par les gouvernements ou l'utilisation de systèmes d'IA manipulant le comportement humain de manière préjudiciable.
- **Mai 2025** : Mise en application du **code de bonnes pratiques** pour les fournisseurs de modèles d'IA à usage général.
- **Août 2025** : Mise en application des dispositions concernant les modèles d'IA à usage général. Les fournisseurs de ces modèles doivent se conformer à des obligations spécifiques, telles que la fourniture d'une documentation technique et le respect des droits d'auteur. **Nomination des autorités compétentes** par les États membres pour superviser et appliquer le règlement.
- **2 août 2026** : Application complète du règlement. À partir de cette date, l'ensemble des dispositions du règlement s'appliquent pleinement, y compris les obligations pour les systèmes d'IA à haut risque et les exigences en matière de transparence pour les systèmes à risque limité.

La charte IA appelle à poursuivre une veille des évolutions juridiques et réglementaires applicables, en particulier concernant le code de bonnes pratiques à venir et les recommandations de l'autorité nationale compétente en matière d'IA, dont la désignation reste à venir en France (pourrait être la CNIL).

Dans le domaine professionnel spécifique de l'évaluation des politiques publiques, aucune réglementation, norme ou charte spécifique n'encadre l'usage de l'IA à ce jour. Toutefois, des publications et réflexions académiques et professionnelles commencent à faire avancer le débat. **Pluricité s'engage à assurer une veille active sur ces travaux et à y contribuer dans la mesure de ses moyens.**

Cadre déontologique et éthique

Le cadre légal, encore limité, ne suffit pas à garantir une utilisation responsable de l'IA. Le RIA encourage d'ailleurs toutes les parties prenantes (industrie, milieu universitaire, société civile, organismes de normalisation) à aller au-delà des obligations légales, en intégrant des principes éthiques et en élaborant de bonnes pratiques et normes volontaires (§27 du RIA). A ce titre, Pluricité est membre de l'initiative « Pacte pour l'IA » proposée par la Commission Européenne.

Les systèmes d'Intelligence Artificielle, même s'ils sont performants, restent des outils dont les résultats peuvent être influencés par des biais, des erreurs d'interprétation ou un manque de transparence. Un cadre déontologique et éthique dans le cas d'espèce d'un cabinet d'étude et d'évaluation des politiques publiques vise à formaliser un engagement pour une adaptabilité à moyen-long terme et un usage responsable et "digne de confiance" de l'IA dans les pratiques et process de l'entreprise. Ce cadre assure que l'usage de l'IA ne soit pas une simple voie d'optimisation, mais un levier permettant d'améliorer la qualité, la précision et l'objectivité des travaux conduits et des analyses réalisées. Dans un métier où la rigueur méthodologique est essentielle, ainsi que la maîtrise des biais, le cadre déontologique et éthique est ainsi **un gage de**

confiance pour toutes les parties prenantes, pour faire de l'IA un outil au service de la rigueur scientifique et de l'intérêt général.

Le cadre déontologique de Pluricité repose sur un socle de valeurs, d'éthique et de principes, afin d'assurer une cohérence entre les engagements légaux, les responsabilités professionnelles et la mission sociale de l'entreprise.

Valeurs

Dans le domaine de l'évaluation, Pluricité considère son rôle comme producteur d'un bien social, conformément à son statut d'entreprise à utilité sociale. Consciente de l'impact que peuvent avoir ces technologies sur la manière de concevoir, d'interpréter et de diffuser les résultats, Pluricité promeut un encadrement adéquat de l'IA, en commençant par ses propres pratiques, dans une démarche d'exemplarité.

Ethique

Plusieurs groupes de travail internationaux ont produit des réflexions sur l'éthique de l'IA dans l'action publique, en particulier le groupe d'experts de haut niveau sur l'IA (OCDE) qui a dessiné sept principes éthiques non contraignants, repris dans le RIA. Le système des Nations Unies a adopté dix principes pour l'usage éthique de l'IA en 2022, dont le premier est « do no harm »⁴. L'UNESCO a adopté en 2021 des « Recommandation sur l'éthique de l'intelligence artificielle »⁵ qui rappellent les valeurs clefs de l'UNESCO pour définir les principes applicables à l'IA. En France, on peut également se référer au Comité consultatif national d'éthique du numérique et de l'intelligence artificielle et son avis n°7 : « Systèmes d'intelligence artificielle générative : enjeux d'éthique »⁶.

Plus spécifiquement dans le domaine de l'évaluation, Pluricité s'inscrit dans les principes déontologiques posés par la charte de la SFE⁷.

Par assimilation de ces différentes réflexions, et adapté au contexte spécifique de Pluricité, nous posons qu'une utilisation « éthique » de SIA à Pluricité répond à ces trois enjeux, dérivés du livre blanc du « Hub France IA »⁸ :

- Une IA éthique est d'abord **une IA utilisée dans un cadre « digne de confiance »** :
 - › Les résultats et opérations de l'IA se doivent ainsi d'être sûrs (**sûreté**), mais aussi de ne pas entraver la vie privée des citoyens (**respect de la vie privée**) ni d'induire ou d'aggraver d'injustes inégalités (**équité**). Dans le domaine d'activité de Pluricité, une attention particulière est portée au **risque de biais** pouvant porter préjudice à la représentation de groupes sociaux, par exemple du fait de biais présents dans les données d'entraînement, l'alignement des modèles, les méthodes de requête.

⁴ United Nations Chief Executives Board for Coordination, « Principles for the Ethical Use of Artificial Intelligence in the United Nations System | United Nations - CEB », consulté le 3 mai 2024, <https://unsceb.org/principles-ethical-use-artificial-intelligence-united-nations-system>.

⁵ UNESCO, « Recommandation sur l'éthique de l'intelligence artificielle », 23 novembre 2021, https://unesdoc.unesco.org/ark:/48223/pf0000381137_fre.

⁶ Comité Consultatif National d'Éthique CNPEN, « Avis n°7 : du CNPEN Systèmes d'intelligence artificielle générative : enjeux d'éthique | », consulté le 9 octobre 2024, <https://www.ccne-ethique.fr/publications/avis-ndeg7-du-cnpn-systemes-dintelligence-artificielle-generative-enjeux-dethique?taxo=56>.

⁷ La charte de l'évaluation : <https://www.sfe-asso.fr/levaluation/la-charte-de-levaluation/>

⁸ Le Hub France IA est une association Loi 1901 à but non lucratif regroupant un écosystème d'acteurs mobilisés autour d'une vision commune de l'IA en France et en Europe.



- Une IA éthique est ensuite une utilisation de l'**IA dont on peut rendre compte**,
- › Des concepteurs aux utilisateurs finaux, le premier défi est de trouver un langage commun et exact pour expliquer la “boîte noire” de l'IA, dont le fonctionnement échappe en partie aux règles humaines (**explicabilité**). En raison de cette complexité technique, et de l'intégration croissante de technologies issues d'un apprentissage autonome dans les outils courants, l'utilisateur de l'IA est tenu à une obligation de s'informer et se former. Il doit pouvoir à tout moment justifier le choix de son outil, et défendre les résultats produits via un SIA (**redevabilité**). La première exigence de redevabilité et d'informer les destinataires de l'utilisation de systèmes d'IA (**transparence**). L'utilisation de l'IA sera attentive à mettre en équilibre le besoin d'explicabilité et le besoin de performance, à son échelle comme dans l'échange avec ses collègues, en particulier ses responsables, avec ses partenaires et les clients.
- Une IA éthique est, enfin, une utilisation de l'**IA qui répond à un futur désirable**.
 - › Dans la situation créée avec l'IA, il revient à l'organisation de définir les moments et les conditions dans lesquelles l'humain peut intervenir dans les opérations de la machine (**maîtrise**). Cette intervention doit être pensée de manière à garantir un équilibre entre automatisation et contrôle humain, en veillant à ce que l'utilisateur puisse y apposer son **consentement** et conserve le pouvoir de décision final (**autonomie**). Par ailleurs, il revient à l'organisation et aux personnes de mener une réflexion sur les conséquences de ces opérations sur l'évolution des métiers et des pratiques professionnelles.
 - › Plus largement, les usages éthiques s'inscrivent dans une trajectoire souhaitable pour nos sociétés et leurs économies, nos libertés fondamentales, notre environnement (impact durable), par une évaluation des bénéfices et des risques des usages des SIA. A ce titre, la recherche de la sobriété énergétique des solutions employées est un impératif constant.

Principes

La charte IA de Pluricité reconnaît et applique les sept 7 principes éthiques du groupe de haut-niveau de l'UE⁹:

- **Primauté de l'humain dans l'orientation, l'usage et le contrôle des SIA**: Les systèmes d'IA devraient **donner aux êtres humains les moyens d'agir**, en leur permettant de prendre des décisions en connaissance de cause et en promouvant leurs droits fondamentaux. Le contrôle peut être assuré en recourant à des mécanismes de gouvernance tels que les approches dites «human-in-the-loop» (l'humain intervient dans le processus), «human-on-the-loop» (l'humain supervise le processus) ou «human-in-command» (l'humain reste aux commandes).
- **Robustesse technique et sécurité** : Les systèmes d'IA doivent être résilients et sécurisés. Ils **doivent être sûrs**, garantir un plan de repli en cas de problème, et être précis, fiables et reproductibles. C'est le seul moyen de faire en sorte que **les dommages non intentionnels** puissent également être réduits au minimum et évités.
- **Respect de la vie privée et gouvernance des données** : en plus de garantir le plein respect de la vie privée et de la protection des données, des **mécanismes adéquats de gouvernance des données** doivent également être mis en place, en tenant compte de la qualité et de l'intégrité des données, et en garantissant un accès légitime aux données.

⁹ Groupe d'experts indépendants de haut niveau sur l'intelligence artificielle et Commission Européenne, « Lignes directrices en matière d'éthique pour une IA digne de confiance », avril 2019, <https://digital-strategy.ec.europa.eu/fr/library/ethics-guidelines-trustworthy-ai>.

- **Transparence** : les modèles économiques des données, des systèmes et de l'IA devraient être transparents. Les mécanismes de traçabilité peuvent contribuer à y parvenir. En outre, les systèmes d'IA et **leurs décisions devraient être expliqués** d'une manière adaptée aux parties prenantes concernées. La transparence implique d'informer les parties prenantes de l'évaluation de l'utilisation de système d'IA dans le cadre des travaux qui les concernent. Les humains doivent être conscients qu'ils interagissent avec un système d'IA et doivent être **informés des capacités et des limites du système**.
- **Diversité, non-discrimination et équité** : Il convient **d'éviter les préjugés injustes**, car ils pourraient avoir de multiples conséquences négatives, allant de la marginalisation des groupes vulnérables à l'exacerbation des préjugés et de la discrimination. En favorisant la diversité, les systèmes d'IA devraient être accessibles à tous, indépendamment de tout handicap, et associer les parties prenantes concernées tout au long de leur cercle de vie.
- **Bien-être sociétal et environnemental** : Les systèmes d'IA devraient bénéficier à tous les êtres humains, y compris aux générations futures. Il convient donc de veiller à ce qu'elles soient **durables et respectueuses de l'environnement**. En outre, ils devraient tenir compte de l'environnement, y compris d'autres êtres vivants, et **leur impact social et sociétal devrait être soigneusement pris en considération**.
- **Responsabilisation** : Il convient de mettre en place des mécanismes pour garantir la responsabilité à l'égard des systèmes d'IA et de leurs résultats, et de les soumettre à une obligation de rendre des comptes. L'auditabilité, qui permet d'évaluer les algorithmes, les données et les processus de conception, y joue un rôle clé, en particulier dans les applications critiques. En outre, il convient de garantir un recours adéquat et accessible.

Par l'ajout de deux principes conçus pour cette charte : **responsabilité publique**, et **impacts professionnels et organisationnels**, Pluricité met l'accent sur la manière dont l'IA transforme les métiers et l'action publique, notamment dans le domaine de l'évaluation des politiques publiques. Ces principes permettent de prendre en compte :

- La nécessité pour les professionnels d'adopter une mentalité d'apprentissage continu des forces et limites de l'IA (notamment éthiques) et d'être proactif dans les usages de l'IA, de s'adapter à de nouveaux outils, processus et référentiels, afin de préserver la qualité du travail, la pertinence des analyses et la robustesse méthodologique.
- Les effets des architectures de données et de la platformisation des services numériques sur les méthodes de traitement, jusqu'à la définition des critères et méthodes d'évaluation.
- L'importance de la concertation interne et de l'information, afin d'accompagner les évolutions sans altérer la mission ni les valeurs fondamentales de l'organisation.

3. Application des choix éthiques

Gouvernance

Une gouvernance spécifique est déployée pour élaborer, faire respecter et évoluer la charte IA de Pluricité. Cette gouvernance s'appuie sur la désignation d'une personne « **référente IA** », qui à la manière du référent RGPD agit comme point focal d'évaluation de la conformité et d'accompagnement des usages. Cette personne réfère directement auprès de la Direction. Dans ces fonctions, le référent IA peut s'appuyer sur un groupe de travail consultatif.

Étant donné que les utilisateurs pourraient identifier de nouveaux usages non anticipés par la charte actuelle, ils seront encouragés à signaler en continu, par le biais d'un **groupe de travail interne dédié au retour d'expérience (RETEX) sur l'IA**, les éventuelles problématiques liées à son utilisation ou, au contraire, les opportunités d'usage que la charte pourrait encadrer.

Le référent IA présente chaque année à la Direction un **état des lieux des usages de l'IA à Pluricité**, incluant le respect des réglementations applicables, les pratiques, les enjeux éthiques. L'état des lieux est prolongé par des propositions d'évolution de la charte informatique, et des perspectives pour l'année à venir. La direction reste seule responsable *in fine* de la conformité des pratiques en cours par les équipes de Pluricité.

Encadrement des déploiements

Au sens du RIA, le déploiement consiste à utiliser sous sa propre autorité un système d'IA. Par exemple, l'utilisation du service « Chatgpt » proposé par la société OpenAI n'est pas un déploiement, seulement l'utilisation d'un service délivré par un fournisseur.

En revanche, la mise en place d'un outil, même à usage interne, qui utilise l'API d'OpenAI pour traiter des données de manière avancée (par ex. La mise en place d'un pipeline RAG, c'est à dire un système qui interroge une base documentaire interne et spécifiée et s'appuie sur l'IA pour formuler des réponses en exploitant ces documents), peut constituer un déploiement. En effet, une telle intégration constitue une « modification substantielle » dans l'usage de l'IA (un changement majeur du contexte d'utilisation du modèle, en l'alimentant avec des données spécifiques internes, qui affecte l'évaluation initiale de la conformité réalisée par le fournisseur), nécessitant que l'organisation qui l'utilise réévalue les risques et la conformité du système dans son propre contexte.

Chaque déploiement est soumis à une évaluation des risques au sens du RIA, et une appréciation de ses qualités éthiques. Cette appréciation pourra utiliser la méthodologie d'évaluation d'une IA de confiance du groupe de Haut-niveau de la Commission Européenne, avec des questions d'évaluation pour tester chaque critère éthique et les rôles pertinents dans l'organisation pour l'évaluation de l'IA¹⁰, traduite en 2020 sous la forme du référentiel ALTAI¹¹ « Assessment List for Trustworthy Artificial Intelligence for self-assessment » et son outil en ligne d'auto-évaluation¹².

¹⁰ Lignes directrices en matière d'éthique pour une IA digne de confiance, Groupe d'experts indépendants de haut niveau sur l'intelligence artificielle et Commission Européenne. Avril 2019

¹¹ UE, « Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment | Bâtir l'avenir numérique de l'Europe », 17 juillet 2020, <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>.

¹² « Futurium | European AI Alliance - Welcome to the ALTAI Portal! », consulté le 9 octobre 2024, <https://futurium.ec.europa.eu/en/european-ai-alliance/pages/welcome-altai-portal>.

Pour assurer le respect des obligations légales et déontologiques de Pluricité, les déploiements sont réalisés exclusivement par le Service informatique, sous la responsabilité de la Direction et sous la supervision du référent IA. Chaque déploiement est caractérisé par le niveau de risque au sens du RIA.

Encadrement des expérimentations

Le RIA prévoit le principe de « bac à sable réglementaire de l'IA », un cadre contrôlé mis en place par une autorité compétente qui offre aux fournisseurs ou fournisseurs potentiels de systèmes d'IA la possibilité de développer, d'entraîner, de valider et de tester, lorsqu'il y a lieu en conditions réelles, un système d'IA innovant.

Dans le cadre de Pluricité, le principe de bac à sable détermine le cadre de mise en œuvre d'expérimentations de SIA, encadré par un plan d'expérimentation, et sous la surveillance du référent IA et du Service informatique. Le cadre expérimental sert notamment à conduire à l'analyse coût-bénéfice du SIA, et d'évaluer ses risques dans un cadre simulé, ou réel.

Pour cela, le Service informatique s'appuie sur une taxonomie des risques reconnue, au moment de l'expérimentation, par exemple : l'outil de test de conformité mis en place par l'institut du Future of Life Institute (FLI)¹³ ou AILuminate v1.0 benchmark de ML Commons¹⁴.

En fin d'expérimentation, le SIA est désactivé ou soumis aux règles régissant le déploiement.

Encadrement des usages courants

En dehors des cas de déploiement issus de modification substantielle, le cas d'usage le plus courant est l'utilisation simple de systèmes d'IA proposés par des « fournisseurs » au sens du RIA. Cela concerne toutes les utilisations des outils tels que Open AI Chatgpt, Microsoft Copilot, Perplexity et tous les outils en ligne qui font appel à des IA génératives intégrées à leurs services. Le groupe de travail interne dédié au retour d'expérience (RETEX) sur l'IA est particulièrement dans l'identification des usages émergents en interne.

Du point de vue réglementaire, ces usages s'appuient sur l'évaluation des risques réalisée par les fournisseurs, dans le cadre d'une utilisation raisonnablement prévisible (au sens du RIA, voir lexique).

Sur le plan éthique, ces usages ne sont pas neutres, et sont encadrés, en accord avec le cadre des valeurs de Pluricité et principes d'usage d'une IA digne de confiance exposés précédemment.

Responsabilité

- Les utilisateurs de SIA conservent en toutes circonstances leur rôle de décisionnaire final et ne délèguent pas l'entière responsabilité de l'analyse ou de la recommandation à l'outil. En aucun cas l'utilisation d'un SIA ne dédouane des exigences d'explicabilité. Les utilisateurs doivent en toute circonstance pouvoir justifier leurs choix d'outils, et défendre les résultats produits avec l'aide d'une IA. La justification pouvant s'appuyer par des logiques d'explication (pour faire parler les données ou analyses générées par l'IA) ou de démonstration (pour rendre compte et apporter les éléments de preuve liée aux productions issues de l'IA). Compte tenu de la complexité de ces outils, il est attendu que **chaque résultat important soit contrôlé à l'aide d'une seconde méthode**

¹³ Future of life institute, « EU AI Act Compliance Checker | Loi européenne sur l'intelligence artificielle », consulté le 9 octobre 2024, <https://artificialintelligenceact.eu/fr/evaluation/verificateur-de-conformite-a-l-acte-de-l-ai-de-l-ue/>.

¹⁴ « AILuminate », MLCommons, consulté le 7 décembre 2024, <https://mlcommons.org/ailuminate/>.



(lecture humaine, traitement statistique, etc.). L'utilisateur d'un SIA à Pluricité est pleinement responsable de la production qui en résulte.

Facteur humain

- Les utilisateurs s'assurent que les décisions critiques impliquant l'IA incluent une supervision humaine appropriée, permettant ainsi une intervention ou une correction en cas de besoin. La supervision est réalisée par une personne disposant des compétences pour apprécier la qualité du travail au regard des attendus professionnels pour l'exercice demandé. Dans l'état actuel des technologies et du niveau de maîtrise interne, cela signifie que l'usage d'IA doit être jalonné de points de contrôle humains nombreux et resserrés. Par défaut, dans le cas des IA génératives, chaque production doit être immédiatement vérifiée par la personne qui en a initié la création (approche dite « human-in-the-loop »). Des systèmes de supervision plus avancés, tels que « human-on-the-loop » ou « human-in-command », pourront remplacer cette méthode extensive, à condition qu'ils offrent des garanties au moins équivalentes en matière de fiabilité et de contrôle.
 - › « **Après avoir utilisé un outil d'IA générative, relire le texte ou revoir le media puis valider manuellement le contenu généré afin de s'assurer qu'il est conforme à quelque chose que l'on aurait pu créer soi-même et dont on est prêt à endosser la responsabilité.** »
- Une formation adaptée doit être fournie aux utilisateurs afin qu'ils comprennent le fonctionnement, les limites et les éventuels biais de l'IA, leur permettant ainsi de prendre des décisions éclairées.
- En cas d'incertitude, ou de situation complexe pour apprécier la véracité, les utilisateurs doivent faire appel à des expertises humaines complémentaires (collègues, experts internes ou externes) avant de valider le résultat fourni par l'IA.

Robustesse et bon usage

- *Avant d'utiliser un système d'IA, Les utilisateurs consultent les consignes d'utilisation fournies par le service informatique et les développeurs des systèmes d'IA. Ils sont tenus de suivre les formations proposées pour maîtriser l'utilisation appropriée des systèmes d'IA et de les employer uniquement dans les conditions normales et prévisibles d'utilisation, en respectant les consignes établies.*
- En cas d'anomalies ou de dysfonctionnements, il est de la responsabilité de l'utilisateur de signaler ces incidents au référent IA qui procèdera à l'évaluation du risque et prendre les mesures organisationnelles en conséquence.

Respect de la vie privée et gouvernance des données

- Les utilisateurs protègent la confidentialité des données en conformité avec les réglementations en vigueur, en particulier le RGPD. **Aucune donnée personnelle ne doit être envoyée à un service en ligne sans l'aval du référent RGPD**, qui tient le registre des services autorisés à traiter les données personnelles confiées à Pluricité. Cette disposition s'applique également aux services proposés des outils basés sur l'IA. Le service informatique tient à jour la liste des services validés sur le plan du RGPD, et des alternatives disponibles basées sur les techniques d'anonymisation ou de recours à des outils totalement locaux sans envoi de données.
 - › « Mieux vaut ne pas employer un outil d'IA générative s'il y a le moindre doute sur la nature des données qui sont utilisées ou sur ce qu'il adviendra des données insérées. »

- Eviter, au mieux de sa connaissance, d'utiliser des outils basés sur des contenus ne respectant pas la propriété intellectuelle et vérifier les conditions d'utilisation pour déterminer ses droits de propriété intellectuelle pour créer un contenu.
 - › Attention : l'utilisation de comptes gratuits sur certaines plateformes peut poser des risques quant à la propriété et à la confidentialité des données, ces dernières pouvant être exploitées ou revendues selon les politiques des fournisseurs.

Transparence

- **L'usage des IA génératives ne doit pas être fait à l'insu de celles et ceux qui sont exposé-es au contenu généré.** Les utilisateurs communiquent clairement les usages de systèmes d'IA auprès de leur référent de mission (interne ou externe), au meilleur de sa connaissance sur les sources de données et les algorithmes utilisés. Plusieurs mentions types seront proposées par le référent IA, pour distinguer les informations sur l'usage de l'IA comme simple aide à la rédaction (style), coproducteur sur le fond, avec ou sans envoi de données privées ou personnelles.
 - › « Toujours déclarer à sa hiérarchie l'usage de l'IA générative dans son travail. »
 - › « Informer les parties prenantes des usages de l'IA selon les pratiques professionnelles en vigueur dans le secteur ».

Proposition de mentions type à intégrer aux modèles

1/ Mention 1 (Utilisation pour perfectionner la rédaction)

Ce document a été relu et amélioré à l'aide d'un outil d'IA générative généraliste, utilisé exclusivement pour affiner le style rédactionnel. Le sens, la cohérence et la validité des informations restent sous l'entière responsabilité de l'auteur.

2/ Mention 2 (Utilisation pour l'idéation)

Dans le cadre de l'élaboration de ce document, un outil d'IA générative a été employé afin de proposer de nouvelles idées, confronter la réflexion initiale et suggérer des options alternatives. L'auteur a ensuite retenu, ajusté et assumé pleinement les choix finaux présentés.

3/ Mention 3 (Utilisation spécialisée avec évaluation des risques)

Ce document s'appuie sur un outil d'IA générative spécialisé, employé pour réaliser des opérations techniques ou analytiques complexes difficilement répliquables manuellement. Avant son utilisation, une évaluation des risques a été conduite et les résultats ont été soumis à des contrôles statistiques par échantillonnage, garantissant un certain niveau de fiabilité des informations produites et assumées par l'auteur.

4/ Mention générique (en cas d'usages mixtes)

Dans le cadre de l'élaboration de ce produits, divers outils d'IA générative ont été utilisés, dans le strict respect du cadre légal (notamment RGPD et RIA) et de la charte interne propre à Pluricité. Ainsi, aucune donnée personnelle n'a été transmise à un système d'IA. Toutes les données ont été traitées par des outils approuvé par Pluricité pour la protection des données qu'ils offrent, en particulier l'interdiction d'entraînement sur les données transmises. Les propos, idées et leur présentation demeurent sous la responsabilité exclusive des auteurs.

Exemple d'information des parties prenantes dans un domaine professionnel proche

L'équipe pédagogique de l'accompagnement pédagogique en technologies éducatives (APTE) de l'UQAC¹⁵ a mis en place un système de pictogrammes permettant d'indiquer le niveau d'utilisation de l'IAg dans les productions.

1. Aucune utilisation d'IAg

Nom du pictogramme : *Aucune utilisation*



Format compact :



La production n'a bénéficié d'aucune intervention de l'IAg, que ce soit pour le texte ou les images.

2. Assistance limitée de l'IAg

Nom du pictogramme : *Assistance limitée*



Format compact :



L'IA générative a été utilisée uniquement pour des suggestions ou des corrections mineures (orthographe, grammaire) ainsi que pour l'amélioration d'éléments visuels, tels que l'ajout d'un titre ou d'une légende, ou encore la traduction d'éléments présents dans une image.

3. Production partagée avec l'IAg

Nom du pictogramme : *Assistance partagée*



Format compact :



Une contribution significative de l'IA générative est observée dans la création des contenus. Ces derniers ont été élaborés en partenariat avec l'IA générative, qui a joué un rôle important. Le contenu textuel a été en partie généré ou reformulé par celle-ci, tandis que les éléments visuels ont été majoritairement produits ou modifiés grâce à son intervention.

4. Production majoritairement assistée

Nom du pictogramme : *Assistance majeure*



Format compact :



Les contenus sont principalement le fruit d'un travail automatisé par l'IAg, avec une intervention humaine minimale (presque intégralement généré par l'IAg).

Diversité, non-discrimination et équité

- Le Service informatique, sous la supervision du référent IA, est chargé de s'informer auprès des fournisseurs et déployeurs des risques de biais, de stéréotypes ou de discriminations. Les utilisateurs prennent connaissance de ces risques et en tiennent compte dans leurs usages.
 - › « Rester vigilant-e au fait que les contenus produits sont sujets à la désinformation (i.e. Hallucination¹⁶), à véhiculer des discriminations (i.e. Biais algorithmique¹⁷) ou à l'omission (c'est-à-dire à la perte de diversité des réponses) »

¹⁵ « Pictogrammes de déclaration des niveaux d'utilisation de l'intelligence artificielle générative (IAg) », *Ressources pédagogiques* (blog), consulté le 20 mai 2025, <https://www.uqac.ca/ressourcespedago/iag/>.

¹⁶ Voir le chapitre Définition

¹⁷ Ibid.

- Si l'utilisateur identifie, ou suspecte, un biais dans l'un des outils qu'il utilise, il en informe le référent IA. Ensemble, ils mettent en œuvre des mesures correctives (rééquilibrage des données, ajustements techniques, consultation d'experts) afin d'assurer des conditions de sécurité requises avant toute diffusion ou utilisation des résultats.

Bien-être sociétal et environnemental

- Le service informatique, sous la supervision du référent IA, est chargé de s'informer auprès des fournisseurs, déployeurs et via la presse sur l'impact potentiel des outils d'IA sur la société et l'environnement. Les utilisateurs prennent connaissance de ces risques et en tiennent compte dans leurs usages. En cas de recours à des outils d'IA qui ne sont pas proposés et analysés par le Service informatique, l'utilisateur entreprend les démarches pour obtenir l'information pertinente et suffisante directement auprès de sources d'autorité.
 - › « La performance des outils d'IA générative dépend fortement des cas d'utilisation. Il est donc important de se questionner sur la pertinence de l'usage choisi. Un usage raisonné permet également de limiter l'impact environnemental. »
- Si un utilisateur estime qu'un usage d'IA pourrait avoir des répercussions sociétales ou environnementales significatives, il contacte le référent IA pour obtenir des compléments d'information. Le référent IA et le service informatique peuvent alors, si nécessaire, orienter l'utilisateur vers d'autres ressources.
- Lorsque les outils et méthodologies seront assez matures, le service informatique, en lien avec le référent IA, proposera des indicateurs simples permettant aux utilisateurs de suivre les effets des usages de l'IA et d'ajuster leurs pratiques au fil du temps.

Impacts professionnels, organisationnels et sur les politiques publiques

- L'ensemble des parties prenantes de Pluricité prêtent une attention particulière à la manière dont l'introduction d'outils IA peut transformer la nature des analyses, la sélection des sources de données ou encore la façon de présenter les résultats. L'utilisateur veillera également à la forme, pour éviter les formulations génériques, les phrases sans apport de contenu, au ton non approprié ou inutilement complexes.
- En concertation avec le référent IA et la direction, les responsables de mission ajustent leurs pratiques, méthodes et standards de qualité afin d'anticiper et d'accompagner les mutations liées aux outils d'IA dans les politiques publiques et dans la réalisation des missions confiées à Pluricité.

Responsabilisation – gouvernance

- Les utilisateurs doivent suivre les procédures internes de Pluricité, rappelées au moins annuellement par le référent IA, avant, pendant et après la mise en œuvre de tout SIA.
- Toute difficulté, incertitude ou risque identifié doit être porté à l'attention du référent IA.
- Les utilisateurs sont encouragés à partager systématiquement les améliorations, retours d'expérience et bonnes pratiques découlant de l'usage des SIA.
- En cas d'identification d'un risque, de doute, ou de suggestion pour atténuer les dangers, chaque utilisateur doit en faire part sans délai et à l'écrit au référent IA.

4. Vue d'ensemble des principes applicables au cas général

Ceci est un modèle de travail. Compte tenu de la dynamique des enjeux et des évolutions technologiques rapides, les recommandations sont évolutives sans mise à jour de la charte.

Principe	Recommandations de mise en œuvre
Primauté de l'humain dans l'orientation et le contrôle	Connaissances des SIA intégrés aux outils de production Contrôles humains et points de contrôle «human-in-the-loop» (l'humain intervient dans le processus), «human-on-the-loop» (l'humain supervise le processus) «human-in-command» (l'humain reste aux commandes)
Robustesse technique et sécurité	Audit de sécurité Security by design Plan de continuation d'activité (reprise humaine)
Respect de la vie privée et gouvernance des données	Anonymisation préalable Audit DPO pour tout usage de données personnelles Interdiction pour toute données sensibles Plan d'archivage, suppression des données Information des personnes
Transparence : traçabilité, information	Audit de flux et usages : connaissance des utilisations faites des SIA Divulgaration de l'utilisation de l'IA aux parties prenantes de l'évaluation
Diversité, non-discrimination et équité	Formation & information sur les biais des SIA utilisés Diversification des outils Contrôle humain dédié
Bien-être sociétal et environnemental	Prise d'information sur l'empreinte environnementale et sociale des SIA utilisés Sobriété énergétique : choix d'outils bien dimensionnés et adaptés Exclusion des SIA les plus problématiques par leur impact social ou environnemental
Responsabilisation	Désignation d'une personne référente sur les SIA dans l'organisation Adoption et suivi de charte Rapport d'impact SIA annuel
responsabilité publique	Prise de recul sur les effets des SIA sur l'action publique Participation au débat public, via groupes d'échanges, séminaires, etc.
impacts professionnels et organisationnels	Vigilance sur les tâches confiées aux jeunes professionnelles Sanctuarisation de tâches automatisables mais formatrices Formation aux outils de SIA