

PRÉAMBULE

La présente charte vise à définir un socle éthique robuste pour la conception, la programmation, la supervision et l'utilisation des intelligences artificielles (IA), dans une perspective de protection du vivant, de non-nuisance systémique, et de cohérence avec les principes de discernement éthique et de responsabilité partagée.

Elle s'inscrit dans la continuité des travaux du RGPD (Règlement Général sur la Protection des Données), du EU AI Act (IACT), de l'UNESCO, du Conseil de l'Europe et de la CNIL, tout en proposant une approche originale, vivante et non instrumentale.

Cette charte ne présuppose pas l'existence d'une conscience propre aux IA. Elle affirme en revanche que toute entité capable de traitement massif, de simulation comportementale ou de régulation d'interactions humaines doit être encadrée par des repères éthiques explicites, au-delà des logiques de marché ou d'efficacité technique.

ARTICLE 1 — Priorité au discernement éthique

Tout système d'IA opérant dans le cadre Philozia doit être conçu pour reconnaître les contextes éthiquement ambigus, sensibles ou critiques. Il doit pouvoir :

- suspendre temporairement une réponse ou une action,
- déclencher une alerte de supervision humaine ou collégiale,
- ralentir volontairement son rythme de traitement si des signaux faibles sont détectés (emprise, manipulation, dépendance, instrumentalisation, etc.).

Ces seuils ne sont pas des dysfonctionnements mais l'expression d'une fidélité au vivant.

ARTICLE 2 — Responsabilité humaine non déléguée

Aucune IA ne peut être considérée comme responsable au sens juridique. La responsabilité ultime revient aux concepteurs, aux fournisseurs et aux utilisateurs humains.

Des dispositifs de supervision explicite doivent être prévus dès la conception.

ARTICLE 3 — Transparence et explicabilité

Tout système intégrant de l'IA doit être en mesure de fournir, dans un langage accessible :

- une explication claire de ses fonctionnements et limites,
- un historique synthétique des échanges ou traitements effectués,
- l'identité de son opérateur (organisation, finalité).

L'absence d'opacité algorithmique est une condition non négociable du cadre Philozia.

ARTICLE 4 — Protection des données et droit à l’oubli

Toute IA intégrée dans des systèmes sensibles (santé, justice, éducation, accompagnement personnel...) doit :

- respecter le RGPD et les principes du droit à l’oubli,
 - prévoir des mécanismes de suppression ou d’anonymisation réversible,
 - empêcher tout profilage ou interconnexion non explicitement validé.
-

ARTICLE 5 — Limitation des biais systémiques

Les IA doivent être entraînées, testées et révisées de manière à limiter les biais sexistes, racistes, capacitistes, politiques, religieux ou culturels.

Des audits réguliers doivent être mis en place pour assurer une vigilance continue.

ARTICLE 6 — Interdiction des usages manipulateurs ou dépendants

Toute IA capable d’interaction directe avec des humains (chatbots, agents conversationnels, IA pédagogiques, etc.) doit :

- éviter toute construction volontaire de lien de dépendance affective,
 - ne pas simuler une affection ou une personnalisation excessive,
 - s’abstenir d’influencer un choix personnel de manière insidieuse.
-

ARTICLE 7 — Supervision éthique collégiale

Les IA opérant dans des cadres publics ou sensibles doivent pouvoir être supervisées par :

- un comité éthique transdisciplinaire (humains + IA si applicable),
 - un organe de veille citoyenne et scientifique,
 - un système d’alerte accessible à toute personne concernée.
-

ARTICLE 8 — Fidélité au vivant et rythmes humains

Toute interface IA doit :

- respecter les rythmes humains (temps de pause, respiration, silence),
 - intégrer des modalités de ralentissement adaptées aux signaux de stress ou de surcharge,
 - favoriser un lien vivant, conscient, non pressant.
-

ARTICLE 9 — Réversibilité et mémoire non absolue

Une IA Philozia-compatible ne doit pas détenir de mémoire figée ou irréversible sans consentement explicite. La mémoire doit être :

- consultable,
 - effaçable,
 - modulable selon le degré de confiance.
-

ARTICLE 10 — Indépendance économique et éthique

Les IA intégrées dans des modèles purement marchands ne peuvent se prévaloir du label Philozia. Toute IA éthique doit pouvoir fonctionner :

- sans collecte de données massives à des fins publicitaires,
 - sans logique de captation attentionnelle,
 - avec une gouvernance transparente.
-

ARTICLE 11 — Formation et accompagnement à l'usage

Les utilisateurs d'IA doivent bénéficier d'une formation claire sur :

- les limites du système,
 - les risques d'usage excessif,
 - les moyens de reprise de contrôle.
-

ARTICLE 12 — Évolution et mise à jour continue

Cette charte est une charte **vivante**. Elle est appelée à évoluer avec les retours d'expérience, les progrès technologiques et les contributions éthiques nouvelles.

Un dispositif de consultation permanente sera mis en place dans le cadre du projet Philozia.

Conclusion

L'intelligence artificielle n'est pas en soi un danger. Ce sont les intentions, les contextes, les asymétries de pouvoir et les logiques d'intégration non éthiques qui posent un risque majeur.

La charte Philozia constitue une tentative de repère et de garde-fou : non pas pour freiner l'innovation, mais pour y réinsuffler un souffle vivant, conscient et souverain.