

PRINCIPES DIRECTEURS DE L'IA

*L'IA POUR PROTÉGER,
L'ÉTHIQUE POUR GUIDER*







AVANT-PROPOS

L'intelligence artificielle représente aujourd'hui un enjeu stratégique majeur pour l'adaptation et la modernisation des politiques publiques dans le domaine de la sécurité intérieure.

Son intégration dans les systèmes opérationnels peut améliorer les capacités de détection, d'analyse et d'anticipation des menaces, dans un contexte marqué par la complexité et l'ampleur croissantes des risques comme des crises sécuritaires.


Cependant, l'utilisation de l'IA dans ce domaine particulièrement sensible exige une grande vigilance éthique. Nous devons veiller à ce que les technologies déployées soient conformes aux valeurs séculaires des forces de type gendarmerie et au respect des libertés individuelles et collectives.

L'IA peut être utilisée au sein d'un large éventail d'applications de sécurité intérieure, notamment la détection automatisée de signaux faibles, le traitement massif de données à des fins d'enquête, les systèmes décisionnels prédictifs et les outils de reconnaissance biométrique à des fins d'identification. Ces technologies offrent des gains opérationnels significatifs, à condition d'être gérées de manière rigoureuse et dans un cadre réglementaire strict.

Les systèmes d'intelligence artificielle doivent répondre à un objectif légitime et clairement défini et être soumis à une évaluation d'impact rigoureuse, garantissant la traçabilité des décisions et le contrôle humain. La lutte contre les biais algorithmiques constitue également un défi majeur. Nous devons garantir la qualité, la représentativité et la neutralité des ensembles de données utilisés pour l'apprentissage des modèles, afin d'éviter toute discrimination directe ou indirecte.

Enfin, l'utilisation de l'intelligence artificielle par les forces de type gendarmerie doit s'accompagner d'un dialogue permanent avec les autorités de contrôle, les experts scientifiques et la société civile. Cette exigence de responsabilité conditionne l'acceptabilité sociale de ces outils et renforce leur légitimité institutionnelle.

Ainsi, le développement de l'IA dans le domaine de la sécurité intérieure est inconcevable sans un cadre éthique solide et partagé. L'État a le devoir de montrer l'exemple dans ce domaine : il en va de la confiance du public, de la protection des libertés et de l'efficacité à long terme de nos politiques de sécurité.



Les quatre principes éthiques - non-malfaisance, justice, bienfaisance et autonomie - partagent un certain nombre de mécanismes de contrôle communs. Ces mécanismes constituent le fondement de la gouvernance éthique de l'IA dans le domaine de la sécurité intérieure et sont appliqués de manière cohérente à tous les principes :

- Évaluation de l'impact sur les droits éthiques et fondamentaux (EFRIA) : évaluation structurée, préalable au déploiement, des risques éthiques, juridiques, techniques et sociétaux.
- Supervision humaine : garantir un contrôle humain significatif sur les décisions prises à l'aide de l'IA, en particulier dans les contextes à haut risque ou sensibles.
- Audits : évaluations régulières et indépendantes visant à détecter les risques systémiques, les biais ou les erreurs.
- Traçabilité : documentation systématique des données, des paramètres et du raisonnement afin de permettre la responsabilisation et la vérification.
- Explicabilité : garantir que les systèmes d'IA sont compréhensibles à la fois pour les opérateurs et pour les personnes concernées par leurs décisions.
- Formation et sensibilisation : doter les opérateurs des connaissances nécessaires pour utiliser l'IA de manière responsable, sans se fier aveuglément aux systèmes automatisés.

Bien que ces mécanismes soient courants, leur importance et leur application varient en fonction du principe :

- Dans le cadre du principe de non-malfaisance, ils visent à prévenir les préjudices et à garantir la proportionnalité.
- Dans le cadre du principe de justice, ils mettent l'accent sur l'équité, la qualité des données et la non-discrimination.
- Dans le cadre du principe de bienfaisance, ils visent à maximiser les avantages pour la société tout en prévenant les risques systémiques.
- Dans le cadre du principe d'autonomie, ils garantissent le libre arbitre humain, le consentement éclairé.
- Dans le cadre du principe d'autonomie, ils protègent l'action humaine, le consentement éclairé et la responsabilité institutionnelle.

Cette structure garantit à la fois la cohérence (grâce à des garanties unifiées) et la spécificité (grâce à des adaptations fondées sur des principes), en accord avec des normes éthiques plus larges.

PRINCIPES DIRECTEURS DE L'IA

POUR LES FORCES DE TYPE GENDARMERIE



Bienfaisance

Promouvoir la sécurité publique et le bien-être de la communauté.

La bienfaisance souligne que l'IA doit soutenir la sécurité publique sans compromettre la dignité humaine.



Non-malfaisance

Prévenir les préjudices et protéger les libertés civiles.

La non-malfaisance souligne l'importance de prévenir les préjudices, en particulier dans le contexte à haut risque de l'application de la loi.



Autonomie

Promouvoir l'égalité et éliminer la discrimination.

La justice et l'équité exigent que l'IA dans le domaine de l'application de la loi soit utilisée de manière à défendre l'égalité, à protéger les communautés vulnérables et à garantir que l'état de droit s'applique de manière équitable à tous.



Justice et équité

Promouvoir l'égalité et éliminer la discrimination.

La justice et l'équité exigent que l'IA dans le domaine de l'application de la loi soit utilisée de manière à défendre l'égalité, à protéger les communautés vulnérables et à garantir que l'état de droit s'applique de manière équitable à tous.

La FIEP est une association internationale regroupant des forces de gendarmerie et de police ayant un statut militaire. Son objectif est de renforcer la coopération institutionnelle entre ses membres, afin de mieux répondre aux défis en matière de sécurité. Elle permettra aux forces de maintenir les normes techniques les plus élevées, mais aussi de développer leurs capacités et d'échanger des valeurs avec d'autres forces de gendarmerie et de police ayant un statut militaire, avec de réelles perspectives de retour à la sécurité intérieure et à la stabilité mondiale.

NON-MALFAISANCE

Prévenir les préjudices et protéger les libertés civiles

CLASSIFICATION : Prévention des risques et minimisation des dommages

DÉFINITION :

Les systèmes d'IA doivent être conçus et déployés de manière à limiter les dommages physiques, psychologiques, juridiques ou réputationnels, en particulier pour les populations vulnérables. Cela implique de ne causer aucun préjudice, direct ou indirect, à des individus ou à des groupes, y compris lorsqu'ils font l'objet d'une enquête ou d'une surveillance. Dans ce contexte, les algorithmes doivent éviter les biais discriminatoires, limiter les atteintes à la vie privée et prévenir les erreurs pouvant entraîner des conséquences graves, telles qu'une arrestation injustifiée ou une stigmatisation sociale. Le principe de non-malfaisance nécessite à la fois des garanties techniques (vérifiabilité, explicabilité) et des protections juridiques. Une surveillance humaine efficace est essentielle dans les applications sensibles afin de garantir que les décisions restent responsables et proportionnées.

MÉCANISMES DE CONTRÔLE :

- **EFRIA** : évaluation complète avant déploiement des risques éthiques, juridiques, techniques et sociétaux..
- **Audits** : surveillance indépendante pour détecter les biais préjudiciables et les erreurs systémiques.
- **Traçabilité** : enregistrement du raisonnement pour permettre la responsabilité et la traçabilité des décisions.
- **Explicabilité** : systèmes compréhensibles pour les cas d'utilisation sensibles et à fort impact.
- **Contrôle humain** : obligatoire dans la prise de décision opérationnelle et juridique.
- **Conformité juridique** : respect de la réglementation des États membres en matière de droits fondamentaux.
- **Proportionnalité et limitation de la portée** : utilisation restreinte à des objectifs légitimes avec des périodes de conservation définies.
- **Formation et sensibilisation** : formation continue des forces de sécurité aux risques et aux limites des outils d'IA.

JUSTICE

Garantir l'équité et la qualité

CLASSIFICATION : Traitement équitable et non-discrimination

DÉFINITION :

Le principe de justice exige que les systèmes d'IA garantissent l'équité et l'égalité de traitement tout au long de leur cycle de vie. Sans garanties adéquates, l'IA peut reproduire ou amplifier les biais intégrés dans les données historiques, affectant de manière disproportionnée certains groupes spécifiques. La justice exige que les algorithmes ne fassent aucune discrimination portant préjudice, ni directe ni indirecte, fondée sur la race, le sexe, l'origine ou d'autres caractéristiques protégées. Cela implique un contrôle rigoureux de la qualité des données, la mise en place de mécanismes d'audit et d'évaluation de l'équité algorithmique, et une transparence suffisante pour permettre la traçabilité des décisions. L'objectif est de garantir que les outils d'IA, tout en améliorant l'efficacité des missions de sécurité, contribuent également à renforcer la légitimité des institutions et à préserver la cohésion sociale.

MÉCANISMES DE CONTRÔLE :

- **Contrôle de la qualité des données :** garantir des données d'entraînement représentatives, impartiales et équilibrées.
- **Évaluation de l'équité algorithmique :** mesurer et atténuer les disparités entre les variables sensibles.
- **EFRIA :** couvrir explicitement l'équité et la non-discrimination.
- **Audits externes et indépendants :** supervision par des experts juridiques, éthiques et scientifiques.
- **Transparence et traçabilité :** documenter les modèles, les critères et les résultats.
- **Formation :** sensibilisation à l'éthique de l'IA, aux biais et aux limites des outils afin d'éviter une utilisation aveugle.

BIENFAISANCE

Maximiser le bien-être social

CLASSIFICATION : Intérêt public et promotion de la sécurité

DÉFINITION :

Le principe de bienfaisance exige que l'IA génère des résultats positifs nets pour la société, en particulier dans le domaine de la sécurité intérieure. L'IA doit renforcer la capacité des institutions à protéger les citoyens tout en respectant les droits fondamentaux.

Les applications légitimes comprennent l'anticipation des menaces, l'optimisation de l'allocation des ressources ou l'accélération des enquêtes, à condition qu'elles servent l'intérêt général et évitent tout abus ou surveillance de masse. La bienfaisance repose également sur la transparence et le dialogue avec la société civile afin de maintenir la confiance dans les institutions souveraines.

MÉCANISMES DE CONTRÔLE :

- **Objectifs clairs et légitimes :** les projets d'IA doivent servir des objectifs définis, proportionnés et légaux.
- **Mesure de l'impact sociétal :** indicateurs permettant d'évaluer la sécurité, l'efficacité et l'optimisation des ressources.
- **Transparence sur les objectifs et les avantages :** communication publique des objectifs, des limites et des avantages.
- **EFRIA :** prise en compte des risques systémiques, notamment les effets dissuasifs disproportionnés ou la criminalisation excessive.
- **Amélioration continue :** adaptation des systèmes en fonction des retours d'expérience et des besoins opérationnels.
- **Répartition équitable des avantages :** garantir que les gains profitent à tous, sans accroître les inégalités.

AUTONOMIE

Respecter le contrôle et le consentement humains

CLASSIFICATION : Niveaux d'autonomie (autonomie totale/autonomie limitée / délégation d'autorité limitée) / Capacité de supervision humaine.

DÉFINITION :

Le principe d'autonomie fait référence à la préservation de la liberté — tant pour les opérateurs que pour les citoyens — dans l'utilisation des systèmes d'IA. Pour les opérateurs, l'autonomie signifie éviter un rôle passif : l'IA peut fournir des recommandations ou un soutien, mais ne doit jamais dicter ou remplacer le jugement humain. Chaque choix important doit rester sous le contrôle significatif de l'homme, garantissant ainsi la responsabilité et l'obligation de rendre des comptes. Pour les citoyens, l'autonomie est protégée en veillant à ce que l'utilisation de l'IA soit transparente et compréhensible, afin que les individus sachent quand la technologie est impliquée et puissent faire confiance au processus. L'autonomie nécessite donc des garanties qui préservent la liberté de choix des opérateurs et la clarté pour les citoyens. L'IA dans le domaine de la sécurité intérieure doit être considérée comme faisant partie d'un continuum technologique plus large : un outil d'assistance intégré aux opérations, mais jamais un instrument de contrôle absolu ou d'application automatique.

MÉCANISMES DE CONTRÔLE :

- **EFRIA** : évaluation des risques pour l'autonomie et l'action humaines.
- **Contrôle humain** : validation de toutes les décisions opérationnelles ou juridiques soutenues par l'IA par des agents humains.
- **Explicabilité** : fournir un raisonnement clair derrière les actions de l'IA.
- **Traçabilité** : documenter les processus automatisés en vue de la réalisation d'audits.
- **Transparence publique** : informer les citoyens des objectifs, de la portée et des limites du système.
- **Mécanismes de consentement** : permettre l'acceptation ou le rejet des interventions de l'IA lorsque cela est possible.
- **Formation** : s'assurer que les opérateurs comprennent les systèmes d'IA et évitent la délégation aveugle.

NOTRE ENGAGEMENT

ITALIAN ARMA DEI CARABINIERI

SPANISH GUARDIA CIVIL

ROYAL NETHERLANDS MARECHAUSSEE

MOROCCAN GENDARMERIE ROYALE

ROMANIAN JANDARMERIA

ARGENTINIAN GENDARMERIA NACIONAL

CARABINEROS DE CHILE

JORDANIAN PUBLIC SECURITY DIRECTORATE

TUNISIAN NATIONAL GUARD

QATARI LEKHWIYA FORCES

UKRAINIAN NATIONAL GUARD

PALESTINIAN NATIONAL SECURITY FORCES

BRAZILIAN NATIONAL COUNCIL OF GENERAL
COMMANDERS OF MILITARY POLICE

DJIBOUTIAN GENDARMERIE NATIONALE

KUWAITI NATIONAL GUARD

SENEGALESE GENDARMERIE NATIONALE

SAMMARINESE CORPO DELLA GENDARMERIA

MOLDOVAN GENERAL INSPECTORATE OF CARABINEERS

FRENCH PRESIDENCY GENDARMERIE NATIONALE

