AI GUIDELINE PRINCIPLES

AI TO PROTECT, ETHICS TO GUIDE

FOREWORD

Artificial intelligence today represents a major strategic level for adapting and modernizing public policies in the field of homeland security. Its integration into operational systems can enhance threat detection, analysis and anticipation capabilities, in a context marked by the growing complexity and scale of security risks.

However, the use of AI in this particularly sensitive area calls for greater ethical vigilance. We need to ensure that the technologies deployed are in line with the age-old values of the Gendarmerie type-forces and the respect individual and collective freedoms.

AI can be used in a wide range of homeland security applications, including automated detection of weak signals, massive data processing for investigations, predictive decision-making systems, and biometric recognition tools for identification purposes. These technologies offer significant operational gains, provided they are rigorously managed.

Artificial intelligence systems must meet a legitimate, clearly defined objective, and be subject to rigorous impact assessment, guaranteeing traceability of decisions and the guarantee of human control.

The fight against algorithmic bias is also a major challenge. We need to ensure the quality, representativeness and neutrality of the data sets used to learn the models, in order to avoid any direct or indirect discrimination.

Finally, the use of artificial intelligence by the Gendarmerie type-forces must be accompanied by ongoing dialogue with the supervisory authorities, scientific experts and civil society. This requirement for accountability conditions the social acceptability of these tools and reinforces their institutional legitimacy.

Thus, the development of AI in the field of homeland security is inconceivable without a solid, shared ethical framework. The State has a duty to set an example in this area: it is a question of public confidence, the protection of freedoms, and the long-term effectiveness of our security policies.

The four ethical principles — Non-Maleficence, Justice, Beneficence, and Autonomy — share a number of common control mechanisms.

These mechanisms form the foundation of ethical governance for AI in homeland security and are applied consistently across all principles:

- Ethical and Fundamental Rights Impact Assessment (EFRIA) : a structured, pre-deployment evaluation of ethical, legal, technical, and societal risks.
- Human oversight : ensuring meaningful human control over AI-supported decisions, particularly in high-risk or sensitive contexts.
- Audits : regular, independent evaluations to detect systemic risks, biases, or errors.
- Traceability : systematic documentation of data, parameters, and reasoning to enable accountability and verification.
- Explainability : ensuring that AI systems are understandable to both operators and those affected by their decisions.
- Training and awareness : equipping operators with the knowledge to use AI responsibly, without blind reliance on automated systems.

While these mechanisms are common, their emphasis and application differ depending on the principle:

- Under Non-Maleficence, they focus on preventing harm and ensuring proportionality.
- Under Justice, they emphasize fairness, data quality, and non-discrimination.
- Under Beneficence, they aim at maximizing societal benefits while preventing systemic risks.
- Under Autonomy, they safeguard human agency, informed consent, and institutional accountability.

This structure ensures both coherence (through unified safeguards) and specificity (through principle-driven adaptations), aligning with broader ethical standards.

AI GUIDELINE PRINCIPLES

For gendarmerie type-forces

Beneficence

Promoting Public Safety and Community Well-being.

Beneficence emphasizes that AI should support public safety without compromising human dignity.

Non-Maleficience

Preventing Harm and Protecting Civil Liberties.

Non-maleficience underscores the importance of preventing harm, particularly in the high-stakes context of law enforcement.

Autonomy

Using enforcement agencies must use AI in ways that respect individual agency, privacy, and legal rights, and provide mechanisms for public transparency and participation.

Justice an Fairness

Promoting Equality and Eliminating Discrimination.

Justice and fairness demand that AI in law enforcement be used in ways that uphold equality, protect vulnerable communities, and ensure that the rule of law applies equality to all.

The FIEP is a international association of gendarmerie and police forces with military status. Its aim is to strengthen institutional cooperation between its members, in order to better meet security challenges. It will enable the forces to maintain the highest technical standards, but also to develop capabilities and exchange values with other gendarmerie ans police forces with military status, with real prospects for a return to internal security and global stability.

NON MALEFICIENCE

Preventing Harm and Protecting Civil Liberties

CLASSIFICATION : Risk Prevention & Harm Minimization

DEFINITION :

AI systems must be designed and deployed to limit causing physical, psychological, legal, or reputational harm, particularly to vulnerable populations. It requires that no harm be done, directly or indirectly, to individuals or groups, including when they are the subject of investigation or surveillance. In this context, algorithms must avoid discriminatory biases, limit invasions of privacy and prevent errors that could lead to serious consequences, such as unjustified arrest or social stigmatization. Non-maleficence requires both technical safeguards (auditability, explainability) and legal protections. Effective human oversight is essential in sensitive applications to ensure decisions remain accountable and proportionate.

CONTROL MECHANISMS :

- **EFRIA :** comprehensive pre-deployment evaluation of ethical, legal, technical, and societal risks.
- **Audits :** independent monitoring to detect harmful biases and systemic errors.
- **Traceability :** recording of reasoning to enable accountability and reconstitution of decisions.
- **Explainability :** understandable systems for sensitive and high-impact use cases.
- **Human oversight :** mandatory in operational and legal decision-making.
- **Legal compliance :** adherence to member state regulation on fundamental rights.
- **Proportionality and scope limitation :** restricted use for legitimate objectives with defined retention periods.
- **Training and awareness :** ongoing education for security forces on risks and limits of AI tools.

JUSTICE

Ensure fairness and quality

CLASSIFICATION : Fair Treatment & Non-Discrimination

DEFINITION :

The principle of justice requires AI systems to guarantee fairness and equal treatment throughout their lifecycle. Without proper safeguards, AI can reproduce or amplify biases embedded in historical data, disproportionately affecting specific groups. Justice demands that algorithms neither directly nor indirectly discriminate on the basis of race, gender, origin, or other protected characteristics. This implies rigorous control of data quality, the implementation of mechanisms for auditing and evaluating algorithmic fairness, and sufficient transparency to enable decisions to be traced. The aim is to ensure that AI tools, while improving the effectiveness of security missions, also contribute to strengthening the legitimacy of institutions and preserving social cohesion.

CONTROL MECHANISMS :

- **Data quality control :** ensuring representative, unbiased, and balanced training data.
- **Algorithmic fairness evaluation :** measuring and mitigating disparities across sensitive variables.
- **EFRIA :** explicitly covering equity and non-discrimination.
- **External and independent audits :** oversight by legal, ethical, and scientific experts.
- **Transparency and traceability :** documenting models, criteria, and outcomes.
- **Training :** awareness on AI ethics, bias, and the limits of tools to prevent uncritical use.

BENEFICENCE

Maximize Societal Good

CLASSIFICATION : Public Interest & Safety Promotion

DEFINITION :

The principle of beneficence requires AI to generate net positive outcomes for society, especially in homeland security. AI should strengthen the capacity of institutions to protect citizens while respecting fundamental rights. Legitimate applications include anticipating threats, optimizing resource allocation, or accelerating investigations — provided they serve the general interest and avoid overreach or mass surveillance. Beneficence also relies on transparency and dialogue with civil society to maintain trust in sovereign institutions.

CONTROL MECHANISMS :

- **Clear and legitimate objectives :** AI projects must serve defined, proportionate, and lawful goals.
- **Measuring societal impact :** indicators to assess safety, efficiency, and resource optimization.
- **Transparency on objectives and benefits :** public communication of aims, limits, and advantages.
- **EFRIA :** addressing systemic risks, including disproportionate deterrent effects or over criminalization.
- **Continuous improvement :** adapting systems based on feedback and operational needs.
- **Equitable distribution of benefits :** ensuring gains benefit all, without increasing inequalities.

AUTONOMY

Respect human control and consent

CLASSIFICATION : Autonomy levels (full autonomy/ Limited autonomy/ Limited delegation of authority) / Capacity of human supervision.

DEFINITION :

The principle of autonomy refers to the preservation of freedom — for both operators and citizens — in the use of AI systems. For operators, autonomy means avoiding a passive role : AI may provide recommendations or support but must never dictate or replace human judgment. Every significant choice must remain under meaningful human oversight, ensuring responsibility and accountability. For citizens, autonomy is protected by ensuring that AI use is transparent and understandable, so that individuals know when technology is involved and can trust the process. Autonomy therefore requires safeguards that maintain freedom of choice for operators and clarity for citizens. AI in homeland security should be viewed as part of a broader technological continuum: a tool of assistance integrated into operations, but never an instrument of absolute control or automatic enforcement.

CONTROL MECHANISMS :

- **EFRIA :** assessing risks to human autonomy and agency.
- **Human oversight :** validation of all operational or legal AI-supported decisions by human agents.
- **Explainability :** providing clear reasoning behind AI actions.
- **Traceability :** documenting automated processes for auditability.
- **Public transparency :** informing citizens of system purposes, scope, and limits.
- **Consent mechanisms :** enabling acceptance or rejection of AI interventions where feasible.
- **Training :** ensuring operators understand AI systems and avoid blind delegation.

OUR COMMITMENT

ITALIAN ARMA DEI CARABINIERI	SPANISH GUARDIA CIVIL
PORTUGUESE GUARDA NACIONAL REPUBLICANA	TURKISH JANDARMA
ROYAL NETHERLANDS MARECHAUSSEE	MOROCCAN GENDARMERIE ROYALE
ROMANIAN JANDARMERIA	ARGENTINIAN GENDARMERIA NACIONAL
CHILEAN CARABINEROS	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	

Conception et réalisation graphique : COMCYBER-MI 250648 MB