



DISCUSSION PAPER | A REVISED OUTLINE OF THE TOOLKIT

TOOLKIT FOR RESPONSIBLE ARTIFICIAL INTELLIGENCE INNOVATION IN LAW ENFORCEMENT

Introduction

AI is undeniably a transformative technology, with immense potential to support the work of law enforcement agencies, especially in a context where effective and efficient analysis of vast amounts of data is becoming more and more important. The enormous potential of AI is however equally paired with numerous pitfalls, making its deployment in law enforcement sensitive and controversial. Failure to adequately cater for these pitfalls can lead to the infringement of human rights and fundamental freedoms, the erosion of the rule of law and the undermining of public trust in law enforcement as an institution. Notwithstanding, there currently exists a lack of guidance to support law enforcement agencies to navigate these challenges.

This gap in guidance has led the INTERPOL Innovation Centre (IC) and the Centre for Artificial Intelligence and Robotics at the United Nations Interregional Crime and Justice Research Institute (UNICRI) to develop the *Toolkit for Responsible AI Innovation by Law Enforcement*. Convening representatives of the law enforcement community and other key stakeholders from industry, academia and civil society from across the world, INTERPOL and UNICRI seek to develop and make available to law enforcement agencies operationally oriented guidance to support them in the design, development, procurement and deployment of AI in a responsible manner. It is important to note that the Toolkit is not meant to be prescriptive for law enforcement, but rather function as a useful guiding tool on the considerations, requirements, and process recommendations for law enforcement when they are considering developing, procuring, and/or deploying AI tools to support their work.

An initial conceptualization of the Toolkit was presented in a discussion paper prepared by INTERPOL and UNICRI following a Pre-Meeting Workshop to the 3rd Global Meeting on AI for Law Enforcement that took place on 12 November 2020. Since then, a Core Group of experts has supported the further elaboration of the concept of the Toolkit, serving as a springboard for ideas and aiding in the refinement of specific aspects of the Toolkit and its developmental process. The present document constitutes the latest iteration of the concept of the Toolkit and reflects updates resulting from this process. This document serves as an update for the Core Group and a basis to collect their feedback regarding its developments. It will also be used to help ensuring that the resulting output reflects the needs and priorities of law enforcement and contributes to AI being used responsibly worldwide in law enforcement.

Main Purpose

The Toolkit should provide law enforcement agencies worldwide with a clear understanding of:

- 1) How law enforcement can utilize AI responsibly to support specific strategic and operational objectives, including all stages of development, deployment, and maintenance process;
- 2) The broadly recognized legal and ethical principles and considerations related to policing and AI to support law enforcement's efforts in utilizing AI tools and ensure a responsible use of AI in law enforcement;

- 3) Any unique infrastructural, organizational, operational and policy considerations for police organizations to start developing, procuring and deploying AI capabilities;
- 4) How law enforcement agencies can build public trust and confidence, whilst improving public knowledge towards the responsible use of AI by law enforcement.

Target Audience and Other Stakeholders

The target audience for the Toolkit is, first and foremost, law enforcement agencies in all INTERPOL Member Countries and United Nations Member States. More specifically, the Toolkit should be communicable to:

- 1) Senior police managers/key decision-makers in the respective law enforcement organization;
- 2) Law enforcement officers who oversee innovation and the use of technology in their respective agencies;
- 3) Research and Development (R&D) officers who oversee developing and maintaining the capabilities in-house, evaluating AI systems for procurement, or influence outsourced procurement process; and
- 4) Legal officers/advisors who counsel on the laws and regulations concerning the specific use of AI in policing work.

Through the development and dissemination of the Toolkit, it is anticipated that target audience no. 1 will attain a general understanding of the potential impact of AI in law enforcement, how it can be utilized lawfully and responsibly, the resources needed, the discussion surrounding its use in law enforcement, and its incorporation into organizational risk registries. Target audiences no. 2-4 will equally obtain the knowledge of the operational domains that AI can support (and have been used), the resources and steps required, and other considerations to *responsibly* adopt the technology in specific use-cases. For the purposes of the Toolkit, the term ‘responsible’ will be framed in line with the Oxford Dictionary, which defines it as acting “in a sensible or trustworthy manner.” In this context, the responsible use of AI by law enforcement should be understood as use that enshrines the general principles of ethics and justice, respect for human rights, and the rule of law.

In addition to the primary target audience, other important stakeholders will be closely involved in the development of the Toolkit and will indirectly benefit from it.

First, in the interests of building trust in the use of AI by law enforcement, the Toolkit should bear civil society and the general public in mind as concerned further stakeholders. Furthermore, AI systems and tools are often developed outside of law enforcement agencies and then adopted by them. While the Toolkit will provide guidance specifically tailored to the in-house development of AI in law enforcement, industry, academia and other entities developing AI will also benefit from the Toolkit in their research and development of AI systems for law enforcement and get clarity regarding law enforcement needs and requirements for operating AI tools in a responsible manner.

While these groups constitute the immediate and secondary target audiences for the Toolkit, the Toolkit may also be an important informative resource for other elements of the criminal justice system, such as the judges and prosecutor, as well as policy-makers focused on criminal justice.

Noting the diversity in target audiences, the Toolkit should therefore maintain a balance in the discussion between strategic, conceptual, technical, and operational language.

Structure

The Toolkit is comprised of many different resources, which collectively are intended to provide law enforcement agencies with a comprehensive understanding of how they can implement AI responsibly to support their work, within operational, legal and ethical considerations.

Although the resources in the Toolkit, as well as their name and exact composition, may change based on the feedback of the Core Group and other stakeholders consulted, the Toolkit is currently envisaged to include the following resources:

- 1) Introductory Guide to AI for Law Enforcement
- 2) Principles for Responsible Use of AI in Law Enforcement
- 3) Organizational Roadmap for Implementing Responsible AI
- 4) Organizational Readiness Assessment for Responsible AI
- 5) Responsible AI in Action Tool
- 6) AI Risk Assessment Tool
- 7) AI for Law Enforcement Primers

The ultimate goal is that law enforcement agencies interested in developing, procuring and deploying AI can resort to the Toolkit to make informed decisions, aligned with the needs and concerns of the communities that they serve. The Toolkit's resources therefore aim to be concise, substantively light, easy to read, engaging, and as much interactive as possible. A brief description of each resource in the Toolkit is provided below.

1) INTRODUCTORY GUIDE TO AI FOR LAW ENFORCEMENT

Goal: This resource aims to convey the foundational information for law enforcement agencies to achieve a working knowledge of AI and data and understand how AI can be used in law enforcement.

Content: Includes an explanation of key terminology and concepts, an overview of various uses of AI tools for law enforcement and a mapping of the relevant stakeholders within law enforcement that play a role in responsible AI.

Format: A document including easy-to-read text, as well as interactive elements and visual aids such as infographics, diagrams and videos to make the information as digestible as possible.

2) PRINCIPLES FOR RESPONSIBLE USE OF AI IN LAW ENFORCEMENT

Goal: This resource aims to convey an understanding of the principles for responsible AI in law enforcement and introduce the target audience to how to think about the ethical and legal issues related to AI in law enforcement.

Content: Includes an overview of the main legal, policy and ethical considerations raised within the development, procurement and deployment of AI in law enforcement and the relevant ethical and legal principles to guide law enforcement agencies as they engage with AI in their police work.

Format: A document including easy-to-read text as well as interactive elements and visual aids such as infographics and diagrams to make the information as digestible as possible. To go beyond high-level formulations and communicate these principles embedded in a law enforcement relevant operational context, the principles will be framed around specific use cases, presented as case studies, with an interactive element to tailor descriptions for users with basic, intermediate and advanced knowledge.

3) ORGANIZATIONAL ROADMAP FOR IMPLEMENTING RESPONSIBLE AI

Goal: This resource aims to provide general recommendations on organizational aspects for law enforcement agencies to implement AI responsibly. These general recommendations also aim to serve as background information necessary for the remaining resources, in particular resource 4 – *Organization Readiness Assessment for Responsible AI*.

Content: Recommendations covering the following aspects:

- **Culture and infrastructure** – How to build a culture within which law enforcement operates for responsible development, procurement, and use of AI, and what infrastructure is needed to engage in responsible AI practice, including best practices and available resources.
- **People and expertise** – What types of expertise is needed within law enforcement agencies to achieve responsible AI.
- **Process and workflow** – How responsible AI efforts should be integrated into the workflow of different law enforcement units during different stages of AI development, procurement, and deployment.

Format: A document including easy-to-read text as well as interactive elements and visual aids such as infographics and diagrams to make the information as digestible as possible.

4) ORGANIZATIONAL READINESS ASSESSMENT FOR RESPONSIBLE AI

Goal: This resource is intended to be used by law enforcement stakeholders responsible for the decision-making on the development, procurement and deployment of AI in determining the organization’s readiness to develop, procure or deploy AI tools responsibly.

Content: A set of questions that focus on the *People and expertise* and *Culture and infrastructure* aspects of responsible AI as portrayed in resource 3 – *Organizational Roadmap for Implementing Responsible AI*. It links to relevant materials in resource 3, as well as best practices and guidelines available to build necessary capabilities.

Format: An interactive questionnaire.

5) RESPONSIBLE AI IN ACTION TOOL

Goal: This resource aims to help stakeholders apply the responsible AI principles in their respective tasks.

Content: A set of questions connected to each Responsible AI principle to prompt the law enforcement stakeholders to consider various aspects of responsibly developing, procuring, and deploying AI tools, tailored to the specific workflow and role of the agent engaging with the resource.

Format: An interactive questionnaire.

6) AI RISK ASSESSMENT TOOL

Goal: This resource aims to provide law enforcement stakeholders with knowledge tailored to the kind of AI tool they are engaging with, allowing them to bear in mind the particular risks and other considerations it raises. It further aims to draw the attention of the law enforcement stakeholders to the benefits of a risk-based approach for a proportionate decision-making process with regards to AI.

Content: A set of questions to determine the level of risk that implementing a specific type of AI technique or application entails.

Format: An interactive questionnaire.

7) RESPONSIBLE AI FOR LAW ENFORCEMENT PRIMERS

Goal: This resource aims to provide specific categories of law enforcement stakeholders with a brief overview of the information contained in the Toolkit that is most relevant according to their specific role and whether they are developing, procuring or deploying the AI tool. The primers should serve as a gateway for these stakeholders to more easily access and understand the resources in Toolkit.

Content: Concise and digestible briefs with key takeaways and sample use-cases from the whole Toolkit tailored for specific groups within law enforcement.

Format: Small stand-alone documents including easy-to-read text.