INTERPOL

unicri
United Nations
Interregional Crime and Justice
Research Institute

TOWARDS RESPONSIBLE AI INNOVATION

SECOND INTERPOL-UNICRI REPORT ON ARTIFICIAL INTELLIGENCE FOR LAW ENFORCEMENT

INTERPOL

unicri
United Nations
Interregional Crime and Justice
Research Institute

TOWARDS RESPONSIBLE AI INNOVATION

SECOND INTERPOL-UNICRI REPORT ON ARTIFICIAL INTELLIGENCE FOR LAW ENFORCEMENT

Disclaimer

The opinions, findings, conclusions and recommendations expressed herein do not necessarily reflect the views of INTERPOL, UNICRI, or any other the national, regional or international entity involved.

The designation employed and material presented in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contents of this publication may be quoted or reproduced, provided that the source of information is acknowledged. UNICRI and INTERPOL would like to receive a copy of the document in which this publication is used or quoted.

Acknowledgements

This report is a joint production of UNICRI and INTERPOL. It has been prepared by Mr. Odhran McCarthy and Ms. Maria Eira (UNICRI) with the support of Ms. Sophie van de Meulengraaf, Mr. Jake Kelley (UNICRI) and Mr. Kevin Joel Anthony (INTERPOL) and under the guidance of Mr. Irakli Beridze (UNICRI) and Mr. Steffen Ousdal (INTERPOL).

INTERPOL and UNICRI would like to express their appreciation for the participation of experts and officials of the second Global Meeting on Artificial Intelligence for Law Enforcement, held in Singapore in July 2019, and to acknowledge the contribution of all those who supported the organization of the meeting and the review of this report.

Copyright

© United Nations Interregional Crime and Justice Research Institute (UNICRI), 2020

Viale Maestri del Lavoro,10, 10127 Torino – Italy

Tel: +39 011-6537 111 / Fax: +39 011-6313 368

Website: www.unicri.it

E-mail: unicri.publicinfo@un.org

© The International Criminal Police Organization (INTERPOL), 2020

200, Quai Charles de Gaulle, 69006 Lyon – France

Tel: +33 4 72 44 70 00 / Fax: +33 4 72 44 71 63

Website: www.interpol.int

E-mail: edgci-ic@interpol.int

FOREWORD

Crime is not stagnant. It is dynamic, ever-evolving and ever-adapting to new realities, becoming increasingly more complex and generating a plethora of new challenges at point of inflection. Over the past two centuries, we have seen emerging types of crimes coming to the fore and traditional crimes taking on different forms or an entirely new scope altogether. So long as crime and criminality maintain this dynamic nature, law enforcement agencies must be prepared to keep pace and flexibly adapt to growing trends and developments in order to ensure the safety and security of our global community.

Artificial intelligence (AI) may be the ace up our sleeve to do just this. The ability of AI to alter the very nature of policing and enhance efficiency and effectiveness to, for instance, identify persons of interest in crowded spaces; forecast and predict violence; automatically sort, tag and classify large police operational data such as evidence or harmful materials; and even monitor for drivers of radicalization, is just beginning to be seen. Much more is on the horizon.

We have strived to shape this forum, giving it meaning and purpose, and positioning it to grow into a global platform for cooperation and collaboration amongst law enforcement on AI

This report on AI for law enforcement is the most recent product of the collaboration on AI between the Innovation Centre of the International Criminal Police Organization (INTERPOL) and the United Nations Interregional Crime and Justice Research Institute's (UNICRI) Centre for AI and Robotics. Together we have created a unique forum for law enforcement to discuss advancements in AI, as well as the impacts of using this technology to fight crime. Since we began our work in early 2018, we have strived to shape this forum, giving it meaning and purpose, and positioning it to grow into a global platform for cooperation and collaboration amongst law enforcement on AI. We had the honour and privilege to hold the second INTERPOL-UNICRI Global Meeting on AI for Law Enforcement at the 2019 edition of INTERPOL World – the world's foremost forum for the exploration of innovation for law enforcement – in Singapore this past July and are already making preparations for the third annual INTERPOL–UNICRI Global Meeting on AI for Law Enforcement in The Hague, the Netherlands in November 2020. The increasing interest and attention these meetings are receiving is both a reward for INTERPOL and UNICRI and reveals the growing relevance of AI for the criminal justice community.

Lawfulness, social acceptance, trustworthiness, responsibility and ethics are important concepts that readers will, with good reason, find regularly repeated throughout this report. Indeed, while there is great potential in AI, the use of this technology by law enforcement also raises very real and serious human rights concerns that can be extremely damaging and undermine the trust communities place in law

enforcement. Human rights, civil liberties and even the fundamental principles of law upon which our criminal justice system is based may be unacceptably exposed, or even irreparably compromised, if we do not navigate this route with extreme caution.

The turmoil created by the emergence of the novel SARS-CoV-2 coronavirus in late 2019 and the ongoing COVID-19 pandemic caused by the virus has served to underscore the importance of this. As the global law enforcement community finds itself thrust into the middle of an unparalleled situation, playing a critical role in halting the spread of the virus, preserving public safety and social order and tackling the rapidly changing face of crime, new technologies such as AI will be powerful resource. Yet, even in times of crisis, we must strive to uphold these fundamental principles and rights and ensure respect for the rule of law.

For these reasons, we feel the need to underscore clearly and from the very outset that nothing in this report should be perceived as an endorsement by either INTERPOL or UNICRI of any specific AI use case for law enforcement at this stage. In fact, we take great solace in the progressive, open and earnest discussions we had at the second Global Meeting this past summer on precisely this critical duality of AI that needs to be further understood and developed. We welcome the identification of the need for additional guidance and support by the law enforcement community to facilitate its adoption of AI and, in doing so, to avoid, not only the 'possible' but the 'inevitable', pitfalls of its use.

Through the close cooperation between INTERPOL and UNICRI, we hope to ultimately contribute to filling this gap, by supporting the identification of current and potential use cases and by providing guidance for the development, deployment and use of AI systems in law enforcement in both a lawful and trustworthy manner. This support and guidance starts with this report, which features key insights from discussions at the second Global Meeting, as complemented by further expert analysis and recent developments, and will continue with the third Global Meeting in The Hague.

AI is here to stay. The question we must therefore grapple with is not if law enforcement should use AI, rather it is precisely in what ways can or should law enforcement use AI and how it does so in the most responsible and appropriate manner.

We hope that we will begin to answer some of these questions in this report.

Anita Hazenberg

Director,

INTERPOL's Innovation Centre

Irakli Beridze

Head, UNICRI Centre

for AI and Robotics

EXECUTIVE SUMMARY

AI can be a powerful tool, enabling law enforcement to realise game-changing potential, enhancing its effectiveness and augmenting existing capacities in the fight against all forms of crime. It is also a double-edged-sword, which must be wielded carefully to avoid infringing fundamental human rights, such as the right to privacy, equality and non-discrimination, and undermining principles of law, such as the presumption of innocence, privilege against self-incrimination and proof beyond a reasonable doubt.

Under the auspices of INTERPOL's Innovation Centre and UNICRI, through its Centre for Artificial Intelligence and Robotics, the second Global Meeting on AI for Law Enforcement took place in Singapore in July 2019. This report captures the presentations delivered and discussions held over the course of the meeting and complements it with further analysis and insights on recent developments of relevance regarding the use of AI for law enforcement, including current trends in AI domains and AI regulations. It goes beyond being merely a report of proceedings and instead seeks to serve as a practical reference to law enforcement agencies that intend to design, develop or deploy AI systems in a responsible and effective manner. It seeks to support law enforcement to better conceptualise the application of AI and further deepen its understanding of the concepts of responsible use of AI and proposes a responsible path forward for law enforcement.

Chapter one describes the general landscape of AI use in law enforcement, noting in particular the growing interest of the law enforcement community and underscoring the need for collaboration between the stakeholders, so as to learn from each others success and failures and be able to better mitigate any harmful impact stemming from the use of this technology. The chapter concludes by highlighting that law enforcement must also remain mindful of other threats related to advancements in AI, in particular the malicious use by criminals and terrorist groups.

In chapter two, four main AI domains considered of relevance for law enforcement by INTERPOL and UNICRI are described, specifically, audio processing, visual processing, resource optimization and natural language processing. Some possible applications in each of these domains are described, along with some of the practical and technical challenges for law enforcement to consider when exploring these applications.

To further support law enforcement in conceptualising potential practical applications of AI and to foster the exchange of experiences and lessons learned, chapter three provides an overview of a selection of AI use cases that are being designed, developed or piloted in Australia, Germany, Japan and Norway by national and local law enforcement agencies. These use cases were presented by the law enforcement agencies in question during the second Global Meeting.

Acknowledging the reality that AI is in fact increasingly being integrated into law enforcement, chapter four dives into the most critical aspect: the responsible use of AI by law enforcement. The chapter begins by presenting the general principles that law enforcement should endeavour to adhere to, namely the respect for human rights, democracy, justice and rule of law, as well as the related requirements of fairness, accountability, transparency and explainability that should be adopted in order for law enforcement to meet these principles. Specific legal challenges, considerations on the importance of social acceptance by the public and recent ethical frameworks and statements of principles being developed in Europe, by the Organisation for Economic Co-operation and Development and others, are also described.

In chapter five, the journey towards realizing the responsible use of AI for law enforcement commences, with specific policing needs and recommended actions that were presented and discussed by representatives of the law enforcement community being identified.

Building upon these needs and recommended actions identified, chapter six presents a proposal by INTERPOL and UNICRI for the development of an operationally oriented toolkit to support and guide law enforcement in the design, development and deployment of AI in a responsible manner. The possible objective, structure, target audience and key points of this toolkit are described.

The report concludes with chapter seven, which defines a selection of points of action for the international law enforcement community, policy-makers and intergovernmental organizations to consider in order to support the development of the toolkit and to further develop and promote the concept of responsible AI for law enforcement.

TABLE OF CONTENTS

1. THE CONTINUED EXPANSION OF AI FOR LAW ENFORCEMENT

It is increasingly evident that AI and the related ecosystem of new and emerging technologies, including everything from the Internet of Things to quantum computing, can have a profound impact on society. This impact was discussedin detail, from a law enforcement perspective, at the first INTERPOL-UNICRI Global Meeting on AI for Law Enforcement in 2018 and presented in a report issued in 2019 – *AI for Law Enforcement*.[1]

The second INTERPOL-UNICRI Global Meeting on AI for Law Enforcement, was held at the INTERPOL World 2019 – an event attended by more than 6,000 persons that provided a platform for law enforcement agencies to engage with partners from business and academia. This meeting gave rise to productive discussions and demonstrated that considerable progress, innovation, research and development (R&D), has and continues to be made on various fronts with respect to the use of AI by law enforcement. The law enforcement community has begun to formulate a heightened sense of maturity in terms of its knowledge and understanding of the intricacies of the application and integration of AI and there are a growing number of AI applications being explored by law enforcement to address specific needs or requirements – so-called 'use cases'. Some of these use cases will be documented in more detail below. Although law enforcement is still largely just piloting AI-based tools, techniques and approaches, the true extent of the possibilities of AI are increasingly becoming attainable in the near to mid-term future.

Technological advances are undeniably moving fast and perhaps even more rapidly than law enforcement can adapt and keep pace. If law enforcement is truly to capitalize on AI, it cannot do it alone. The need for a concerted effort for collaboration on AI for law enforcement was underscored at the second Global Meeting as one of the key points for law enforcement to seize the opportunities that AI presents.

Building cooperation on AI with stakeholders throughout the public sector, industry, academia, as well as related security entities, intelligence agencies, counter-terrorism bodies and so on, is an essential next step. Each stakeholder can both make a significant contribution and benefit from law enforcement's own unique experiences. Building upon each other's successes and learning from failures, considerable progress can be made in establishing the methods and practices for applying and integrating AI. Competition is not necessary, constructive. In the case of some of the major initiatives within the private sector, competition may not even be possible.

Indeed, the goal of law enforcement agencies should not be to create parallel AI R&D programmes, but rather to increase and promote knowledge- and technology-sharing, relationship-building and cross-stakeholder collaboration, in order to advance the work it requires on technical applications. Such

1 UNICRI & INTERPOL. (2019). *Artificial Intelligence and Robotics for Law Enforcement*. Retrieved from http://www.unicri.it/news/files/ARTIFICIAL_INTELLIGENCE_ROBOTICS_LAW%20ENFORCEMENT_WEB.pdf

a positive collaborative approach has been exhibited by the private sector as of late, wherein it is increasingly encouraged that research results be published and shared openly. One good practice discussed during the second Global Meeting was the recruitment of students from universities by the Netherlands National Police to carry out specific R&D activities within its premises and on behalf of the police.

The law enforcement community has begun to formulate a heightened sense of maturity in terms of its knowledge and understanding of the intricacies of the application and integration of AI

Nonetheless, interaction with entities outside law enforcement is not always easy. It may bring to light privacy or security concerns when the information shared is personal, harmful, sensitive or concerns ongoing investigations. Furthermore, practical challenges arise when different approaches, practices or techniques are used by collaborating stakeholders in R&D. During the second Global Meeting, some of the practical challenges of collaborating with academia were highlighted alongside the benefits. While there is no quick fix to this, one solution proposed for overcoming the challenge of the practical differences in multi-stakeholder collaboration would be to standardize systems, which would facilitate technology development and indirectly contribute to boosting cooperation between sectors.

At the same time, while the technological revolution is dramatically changing how people, the media, governments, businesses and policy-makers operate, there are two perspectives for law enforcement to bear in mind: the beneficial use *and* the malicious use. During his opening address at the 73rd General Assembly of the United Nations in 2018, the United Nations Secretary-General, António Guterres, identified the risks associated with the advances in technology as one of two epochal challenges for humanity, flagging the possibilities of mass unemployment and the malicious use of these technologies.[2] The latter in particular requires the continued attention of law enforcement.

Whether it is global positioning services (GPS), the mobile phone, the Internet, drones or cryptocurrencies, criminals have long been early adopters of technology. In this past year, the possibility of the use of AI for malicious purposes – as well as the abuse by criminal groups of existing AI systems used by

2 Guterres, A. (2018, Sept 25). *Address to the General Assembly*. New York, UN Headquarters. Retrieved from https://www.un.org/sg/en/content/sg/speeches/2018-09-25/address-73rd-general-assembly

business or public bodies – has become ever more real, with substantial and worrisome developments taking place.

According to the antivirus software developer AVG, phishing was the most common delivery method for malware in 2019 and is responsible for 32% of data breaches and 78% of cyberespionage incidents.[3] These cyberattacks can be used for fraud, extortion and espionage and can have damaging consequences. At the same time, AVG also identified a rise in AI-powered malware in 2019, indicating that cybercriminals are already exploiting automated tools to evade detection and conceiving new attacks to thwart the AI models, using the same smart automation technology security companies use to avoid malware. Another growing threat is the malicious use of small unmanned aeral vehicles (UAV) – commonly known as drones. Although their use is principally through manual remote operation and does not fully leverage AI, the advancing capabilities of AI may soon lead to greater autonomy. To date, drones have been used with improvised explosive devices (IEDs) in attacks by non-State actors in conflict zones, such as Syria, Iraq and Ukraine,[4] and to disrupt operations at airports, such as Gatwick Airport in the United Kingdom in December 2018.[5]

Perhaps the biggest eye-opener, in terms of the malicious use of AI, concerned advances in programmatically generated fake videos and images, or so-called 'deepfakes'. In 2019, the world witnessed the first noted criminal use of the technology in the United Kingdom.[6] Using AI-based software, the voice of a CEO of an energy company was successfully imitated and used to deceive an executive into transferring substantial sums of money into a private account. The criminal potential of deepfakes is enormous, as is the possibility to create social and political upheaval. In this regard, law enforcement must be prepared not only to leverage AI for good, but also to combat such current threats and to anticipate possible future ones.

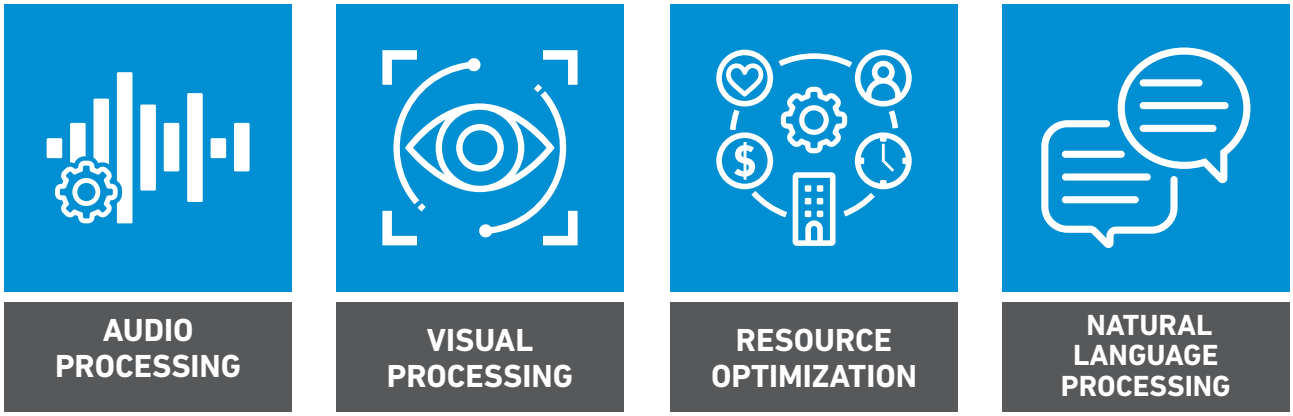
3 AVG (2019, Dec 16). *20:20 Vision — 5 Threats to Watch Out for This Year*. Retrieved from https://www.avg.com/en/signal/online-threats-in-2020

4 United Nations Security Council. Counter-terrorism committee executive directorate. (2019). *Greater efforts needed to address the potential risks posed by terrorist use of unmanned aircraft systems.* CTED Trends Alert. Retrieved from https://www.un.org/sc/ctc/wp-content/uploads/2019/05/CTED-UAS-Trends-Alert-Final_17_May_2019.pdf

5 BBC News (2018, Dec 3). *Gatwick Airport: Drones ground flights.* Retrieved from https://www.bbc.com/news/uk-england-sussex-46623754

6 Damiani, J. (2019, Sep 3). *A Voice Deepfake Was Used To Scam A CEO Out Of $243,000.* Forbes. Retrieved from https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/#af5474822416

2.
AI TECHNOLOGY DOMAINS



Horizon scanning is a valuable exercise. It can help law enforcement ensure it stays abreast of the most recent technological developments and identify those that are best-suited to contribute to its work, fill capability gaps or augment existing capabilities. Equally, it is a valuable exercise for policy- and decision-makers in the broader criminal justice community to, from a legal and ethical perspective, prepare frameworks for the eventual integration of such technologies into law enforcement.

In line with this, and following the first Global Meeting, INTERPOL and UNICRI carried out a preliminary horizon scanning of the AI field and identified four major domains involving AI-based technologies, which could be of most immediate relevance for law enforcement. These AI technology domains are:

For each of these four AI technology domains, specific use cases have been identified to illustrate their relevance and application for law enforcement. What follows is a brief overview of each of these four AI technology domains and the identified use cases.

It should be noted that these use cases are not future possibilities, rather they are present-day realities which are being developed and tested by academic and scientific researchers and, in some cases, by law enforcement and security agencies.

2.1 AUDIO PROCESSING

Audio processing is the manipulation of the characteristics of an audio signal to, for instance, enhance audio, separate sources, create entire new sounds or compress, store or transmit data. Capitalizing on advancements in scientific fields such as linguistics and anatomy, AI promises to open up the full potential of audio processing for law enforcement, allowing it to carry out AI-powered voice profiling.

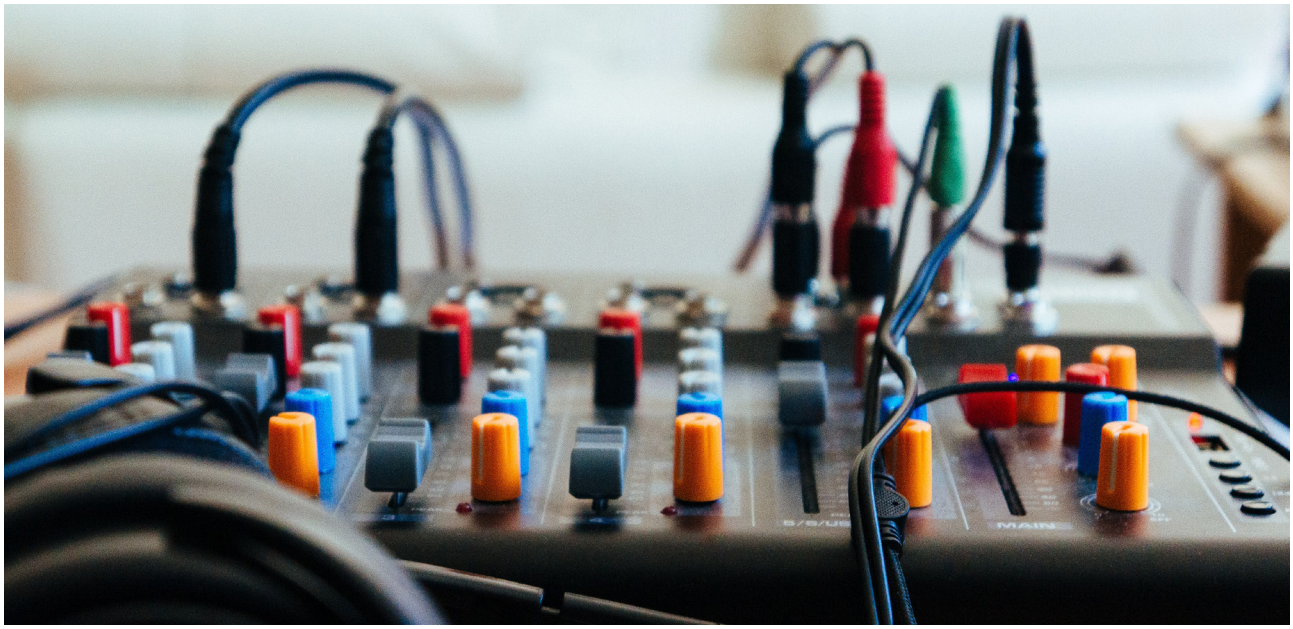
Profiling is a common technique used to extrapolate relevant information from an individual's psychological and behavioural characteristics. In law enforcement, it is often used as an investigative tool to support the identification of possible suspects or link seemingly disparate cases that may have been committed by a single offender. In addition to psychological and behavioural characteristics, the human voice is a valuable medium, carrying with it considerable personal information that, if appropriately deciphered, can support investigations.

An individual's unique voice print can be lawfully collected by law enforcement from wiretaps, radio-transmitted voices, wearable smart devices or other surveillance devices. AI voice profiling applications can make use of such voice prints after the samples are broken down into millisecond fragments. Examining this voice print and analyzing how sound vibrates off and through the body's unique physical structures, the system can deduce physical traits such as height, weight, facial structure, age and even make predictions regarding personality, physical health and mental well-being. Additionally, information about the physical environment surrounding a speaker can be deduced, for example: if the person is indoors or outdoors, the size of the room in which they are speaking, the material of the surrounding walls, if there are windows, what kind of equipment is being used nearby, and even the time of day on the basis of signatures left in the recording by fluctuations in the local electrical grid.

Using these insights, a construction of a person of interest's face or entire body can be generated, providing law enforcement with a realistic and actionable representation. It is possible for this technology to generate a prediction of the future physical characteristics of an individual in say five or ten years based solely on voice recordings.

While it is a promising application, voice profiling has not yet been widely explored by law enforcement and there is still some distance to go before it becomes viable in court. Nevertheless, in 2014, a voice profiling application developed by Carnegie Mellon University played an integral role in an investigation into hoax distress calls to the United States Coast Guard, leading to an arrest.[7] Such hoax distress calls are a federal crime in the United States and waste government resources and unnecessarily put the lives of coast guards at risk. In this case, the voice-based profiling application was used to create a profile of the caller, which assisted officials in eliminating false leads and, ultimately, expediting the identification of the subject and his location.

7 Singh, R. (2019). *Profiling Humans from their Voice*. Carnegie Mellon University, Pittsburgh, Springer.



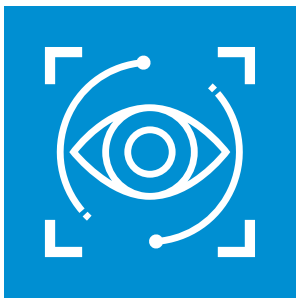
It is, however, pertinent for law enforcement to note that the technology does still face several practical challenges, including:

- Quality degradation that arises with noisy conditions, non-ideal acoustic environments and channels of insufficient quality;
- Overlapping human voices, such as multiple simultaneous conversations, multi-speaker conversations and conversations taking place in crowded spaces;
- Keeping pace with the increase in the number of new technologies that can be used to record or transmit audio, each of which produces different acoustic signatures;
- Voice manipulation or vocal artistry that can mask the true voice of a speaker;
- The emergence of deepfake voice applications that can generate credible fake voices; and,
- The decreasing trend of certain ranges of unique characteristics in voices brought on by factors such as globalization, technology, media and entertainment.

2.2 VISUAL PROCESSING

In a biological context, visual processing is the interpretation and understanding of visual information that allows us to identify what we see, to interpret size, shape, distances etc. From a technological perspective, visual processing, or computer vision, is the mimicry of the human visual system by a machine and it concerns the extraction, analysis and understanding of information from images.

The work of law enforcement has long been supported by visual information – be it pictures or videos of persons, vehicles or locations of interest. In fact, law enforcement was very much transformed with the advent of surveillance technology, in particular CCTV, which became widely available during the 1960s and 1970s. Surveillance technology has allowed law enforcement to quickly identify victims, perpetrators or other persons of interest and, in doing so, to solve crimes. Surveillance technologies have undergone considerable advancements over the years, including body-worn cameras (bodycams) and patrol drones, but it is with AI that perhaps the most impressive outcomes may materialize.



In the past decade, surveillance systems, accompanied by major advancements in machine learning applications, have created systems that can be trained to do the jobs of humans more efficiently and effectively. Early machine learning systems for surveillance were not advanced enough, as they had low performance, were time-consuming and required high-skilled engineers to re-train the model for each new deployment. Since 2012 however, deep learning has revolutionized the areas of image processing and object recognition. The integration of deep learning models into surveillance systems has allowed for tremendous progress in this field, specifically by improving detection and analysis of objects, human faces and bodies, leading to substantial drops in error rates of misidentification of persons of interest. These advanced systems can conduct face detection and recognition, as well as the recognition of facial expressions. They can also conduct human body detection, person identification, attribute recognition, human behaviour recognition, and body movement (gait) recognition. Concerning objects, they can conduct object tracking, vehicle identification and re-identification, license plate recognition and crime scene classification. AI-enabled visual processing systems can be used to identify abnormal behaviour and for both black- or white-listing persons to facilitate or limit entry into specific buildings or closed events, such as concerts and festivals. This can significantly augment the preparedness of law enforcement and security forces tasked protection of such events. At the same time, this domain continues to advance with other developments such as face search systems on the horizon that can use a snapshot of an individual to search through live camera systems in order to locate them.

In the past decade, surveillance systems, accompanied by major advancements in machine learning applications, have created systems that can be trained to do the jobs of humans more efficiently and effectively

These systems face a number of significant practical challenges however, including non-frontal subjects or covered subjects and the use of facial accessories, like glasses, jewellery and masks. Notwithstanding this, these systems will only get more advanced with the increasing amounts of data and the evolution of recognition methods, such as the facial-points identification method, which will enable law enforcement to overcome these challenges in its use of visual processing technologies.

2.3 RESOURCE OPTIMIZATION

In times characterised by more limited resources, an increasingly complex operating environment, and no significant decrease in global crime rates, law enforcement is increasingly being challenged to do more with less. Historically, law enforcement agencies have turned to technology for support in reducing inefficiencies within policing, improving operations and logistics and, ultimately, in striving for more efficient response systems that optimize the necessary amount of resources for a specific situation. Resource optimization – helping law enforcement do more with less – is a further domain in which AI can play a significant role.

Together with smart sensors, the Internet of Things, next generation telecommunication network (5G, Wi-Fi 6), and augmented reality, AI is key to realizing the concept of 'smart policing', supporting strategic planning and decision-making processes, increasing efficiency by enabling the improved allocation of officers, vehicles and equipment and diminishing emergency response times. Technological advancements and improved strategies of dynamic matching in resource supply and demand have already helped to decrease the response time for emergency calls. Further improving these dynamic systems and processes will, ideally, result in increased security. Although crime mapping has been pursued by law enforcement since the 1990s, the data and machine learning algorithms required for significant breakthroughs have just now become available. These tools allow for well-informed decisions in order to significantly cut down the emergency response times and strengthen the connection between alert, response and reaction.

Five main areas can be identified with regards to the use of AI for resource optimization:

Hot spot mapping – the collection of historic crime data from local law enforcement departments, combined with additional datasets, including weather predictions, police patrol history and criminological knowledge, to predict crime hot spots in a jurisdiction.

Deployment of resources on demand – the allocation of police resources (i.e., personnel, vehicles and equipment) based on the actual demand for the area in which crime hot spots have been identified.

Patrol route scheduling – the use of identified hot spots to optimize patrol routes and schedules.

Dispatch of resources for calls – dispatching the nearest available resources to respond to service calls based on their predicted response time.

Response route plotting – identifying the optimal route by factoring in distance and time and then deploying the resource(s) based on availability and optimal response time.

Resource optimization technology has seen a lot of advancements in recent years and has been and will continue to be factored into many operations of public management, including by law enforcement. However, the use of AI for resource optimization does require law enforcement to be prepared to answer a number of essential questions during the design process, which will determine the efficacy of the tool. In essence, when it comes to understanding and predicting optimal decisions, such a system needs to know what 'optimal' is and how to calculate it. Furthermore, it needs to know when, where and how incidents occur; how resources can be deployed; how well deployed resources will perform; how long cases take; how other variables, such as traffic, day or time of the week and weather affect incident patterns and responses and many more related questions.

Robust deployment optimization will also require plans that will work well with multiple incident scenarios and the overall objective must be to minimalize the failed incident response. Human controllers will need to carry out ongoing performance evaluation of these systems. This can be done by comparing deployments designed using generated incidents tested on actual data. Ultimately, evaluations will help improve the systems, since the machine learning model builds off past successes and failures.

2.4 NATURAL LANGUAGE PROCESSING

Natural Language Processing (NLP) – otherwise known as computational linguistics – is a field of AI that, in essence, enables machines to read, understand and derive meaning from human languages. It has proven useful in the extraction of information from large datasets, especially those containing unstructured data – data that is not or cannot be contained in a row-column format - like the text of an email. In light of this, NLP has found its way in daily life, such as in many applications that provide predictive or suggestive text and word or grammar checks.

In a law enforcement context, NLP offers considerable potential, especially in the review and classification of evidence. This can be done by extracting information and analyses from text-based sources, such as emails, online chats, written or typed documents or images thereof. In doing so, law enforcement can save time and resources in extracting relevant information from the data and converting it into actionable insights for digital forensics to aid investigations.

Detecting language through speech can also be carried out, which can be equally advantageous for law enforcement. Once language is detected in an audio clip, NLP applications can be used to run audio searches or to create automated transcripts of conversations, which may be useful, for example, for statement-taking machines. Once audio has been transcribed, the text can be processed for information extraction, namely by classifying the topic, sentiment, and intent of text or by clustering (i.e., grouping together similar and dissimilar groups). Recent research has increasingly focused on unsupervised learning as machine learning models are now able to understand the topic and context of the information just by using text, a fact that opens up the entire content of the world wide web for NLP applications.

NLP also has the potential to be useful in improving digital security. For instance, NLP-enabled filters can classify and analyze emails to block phishing attacks. In the era of fake news, it can also help to combat disinformation by determining if a source is accurate and trustworthy.

In a law enforcement context, NLP offers considerable potential, especially in the review and classification of eviden

For these reasons, several law enforcement entities have taken note of and begun to work with NLP to process and transform language, and then extract information in order to search through documents, detect expressed sentiments, summarize texts, and translate between languages. NLP is, however, a very challenging task and considerable work is required in the development of an NLP application. Fortunately, through transfer learning – sharing parts of coding/scripts, model modules or libraries for algorithms – experts can help one another. For example, the Netherlands National Police has notably started open-source data projects in order to improve the success and results of NLP, as well as accessibility to use these technologies for good. Through transfer learning, experts will be able to collaborate in order to strengthen practices and models that can be utilized by law enforcement.

3. USE CASES FROM LAW ENFORCEMENT AGENCIES

It was recognized during the second Global Meeting that considerable development has taken place with respect to AI and law enforcement since the first Global Meeting in 2018, with many law enforcement agencies having substantially increased their interest in AI. The growth is reflected in the number of cases where States had adopted AI national strategies, action plans or related policy papers.[8] While not all of these policies specifically address law enforcement or crime prevention as a component of the national approach to AI, those of Germany,[9] Italy,[10] Lithuania,[11] the Netherlands,[12] the Republic of Korea[13] and the United Arab Emirates[14] notably do – albeit very briefly in most cases. Notwithstanding the brevity of these references, such acknowlegdements begin to lay strategic foundations for law enforcement agencies to explore the development of AI capacities.

This interest in AI for law enforcement was also seen in an increase of the number of units, centres and R&D laboratories focusing on AI that have been recently established by law enforcement agencies. For example, the Big Data Team in Germany's recently established Central Office for Information Technology in the Security Sector (ZITiS); the Artificial Intelligence for Law Enforcement of Community Safety (AiLECS) Lab established by the Australian Federal Police in collaboration with Monash University; and the Machine Learning and Big Data Team in Advanced Technology Planning within Japan's National Police Agency (NPA).

In line with this, several law enforcement agencies have begun exploring new concepts and applications through pilot projects and increased openness to collaboration and to addressing the inherent ethical, legal and social challenges that go with the use of AI in law enforcement. During the second Global Meeting, four law enforcement agencies were identified to provide specific updates on AI use cases and R&D being explored.

8 For a broad overview of State initiatives, refer to Campbell, T. (2019). *Artificial Intelligence: An Overview Of State Initiatives.* Retrieved from http://www.unicri.it/in_focus/files/Report_AI-An_Overview_of_State_Initiatives_FutureGrasp_7-23-19.pdf

9 Federal Government of Germany (2018). *Artificial Intelligence Strategy.* Retrieved from https://www.bundesregierung.de/resource/blob/997532/1550276/3f7d3c41c6e05695741273e78b8039f2/2018-11-15-ki-strategie-data.pdf?download=1

10 The Agency for Digital Italy (2018). *Artificial Intelligence at the Service of Citizens.* Retrieved from https://ia.italia.it/assets/whitepaper.pdf

11 Ministry of Economy and Innovation of the Republic of Lithuania (2019). *Lithuanian Artificial Intelligence Strategy: A Vision of the Future.* Retrieved from http://kurklt.lt/wp-content/uploads/2018/09/StrategyIndesignpdf.pdf

12 Ministry of Economic Affairs and Climate of the Netherlands (2019). *Strategic Action Plan for Artificial Intelligence.* Retrieved from https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/beleidsnotas/2019/10/08/strategisch-actieplan-voor-artificiele-intelligentie/Rapport+SAPAI.pdf

13 Government of the Republic of Korea (2017). *Mid- to Long-Term Master Plan in Preparation forthe Intelligent Information Society.* Retrieved from https://k-erc.eu/wp-content/uploads/2017/12/Master-Plan-for-the-intelligent-information-society.pdf

14 Government of the United Arab Emirates (2017). *UAE Strategy for Artificial Intelligence.* Retrieved from https://government.ae/en/about-the-uae/strategies-initiatives-and-awards/federal-governments-strategies-and-plans/uae-strategy-for-artificial-intelligence

3.1 NON-INTRUSIVE SURVEILLANCE SYSTEMS NORWAY

Videos and images collected by law enforcement through surveillance systems, such as closed-circuit television (CCTV) and body or car cameras, are often essential for law enforcement to prevent and investigate crimes and secure the prosecution of offenders within the court system. At the same time, however, images, videos and sounds contain information that can disclose the identity of individuals, the handling and use of which may present concerns regarding privacy.

In aiming for transparency, the Oslo Police District of Norway has been working with partners both within the police force and externally with industry and academia to explore the application of AI for the creation of heavily user-sensitive non-intrusive surveillance systems that can be employed in smart cities. The anonymization of videos and images collected using AI is at the core of this. More specifically, the images of people captured in surveillance footage are automatically anonymized by the AI system, by covering their face with a cartoon character or emoji. This enables the anonymous datasets to be handled and shared by the police and with police partners for use in a non-intrusive manner, for example, the anonymized data can be freely used for pattern recognition to identify acts such as vandalism, street fighting and movements that indicate intoxication. This project marks one of the first trials by law enforcement in which pattern recognition is combined with means of automated anonymization. This pilot is also notable in that it could even present a partial solution to addressing compliance with the General Data Protection Regulation of the European Union (EU), whereby law enforcement is required to be able to perform non-intrusive monitoring and evidence analysis with regards to privacy of the 'bystanders' or non-person of interest (non-POI).[15]

While non-intrusive surveillance presents opportunities for law enforcement and society to rethink surveillance, it nevertheless still requires careful consideration of privacy, data protection, de-anonymization and the practicalities of the use of anonymized data by law enforcement.

Images, videos and sounds contain information that can disclose the identity of individuals, the handling and use of which may present concerns regarding privacy

15 See *GDPR*, *inter alia*, Article 5(1)(c), which sets the 'data minimisation' principle. The full text of the GDPR is available at https://gdpr-info.eu

3.2 DATA AIRLOCK AND HARMFUL MATERIALS RECOGNITION AUSTRALIA

In Australia, the Australian Federal Police (AFP) has been pursuing several opportunities for the use of AI and machine learning to automate or assist with tasks such as tagging and organizing data. This includes the creation of a search engine system, akin to Google, for investigation data that will support AFP in making both sense and use of the large amounts of structured and unstructured data in its databases. Without the support of tools such as this, it may otherwise encounter difficulties in locating data points and making critical connections.

Another topic AFP is working on is the creation of a 'data airlock' system, which enables researchers to develop new algorithms without having access to the data. The data airlock is equipped with cryptography to provide an isolated and secure environment where researchers can put their algorithms and models in, execute them against the data, and extract the results of the research and analysis. Accordingly, data never leaves the data owner's environment, a feature especially relevant for organizations dealing with sensitive data. It is expected that the data airlock system will enable third parties to train, validate and test machine learning tools against real-world seized data, without requiring direct access to these materials. This could, for instance, help researchers to better understand and monitor the dark web.

A notable application of this system that is being explored by AFP focuses on the use of deep learning models to recognize, tag and cluster images and videos containing harmful material, such as child sexual abuse materials. Automated recognition of harmful materials, combined with the data airlock, will effectively protect law enforcement officers, investigators and researchers by diminishing their exposure to these materials.

The data airlock is equipped with cryptography to provide an isolated and secure environment where researchers can put their algorithms and models in, execute them against the data, and extract the results of the research and analysis

3.3 RECOMMENDER SYSTEM GERMANY

The Central Office for Information Technology in the Security Sector (ZITiS), within the German Federal Ministry of the Interior, has been working on a recommender system, similar to those used by platforms such as Amazon, Netflix and Spotify. Recommender systems work on the assumption that in large, diverse datasets, similar users select similar items. Based on this knowledge, and after an initial period of learning, items of potential interest are automatically recommended to the user.

For large law enforcement agency databases containing vast quantities of data from criminal investigations, finding relevant information is often like searching for a proverbial 'needle in the haystack'. ZITiS was specifically motivated to develop the recommender system to meet the exigencies of modern financial crime investigations that typically require the scrutinizing of vast quantities of financial data. Large cases of this kind could take years of work to search through and to find and cross-check relevant information for the case. With the increase of data and ever-growing databases, the frequency with which such large cases arise is likely to increase. The system itself is an active learning content-based tool, which means that the algorithm tries to interactively query valuable information for officers based on prior searches or similar user preferences.

While a promising application, there are some challenges concerning creating effective recommender systems for law enforcement. Primarily, it is essential that a recommender system with decision-making capacities is not biased. Bias will affect the accuracy and truthfulness of the information and could compromize an investigation. This will be a significant challenge for law enforcement to carefully navigate as recommender systems are in fact designed to encourage users in a specific direction, so as to avoid searching in multiple directions. For instance, recommender systems in the private sector, such as those used by Amazon, Netflix and Spotify, are designed to encourage users to 'buy more products', 'watch more films' or 'listen to more music'. Law enforcement will need to ensure that its recommender systems truly seek to provide valuable information for law enforcement and not merely encourage 'more arrests'. Another challenge is that ordinary recommender systems use limited data types, such as either audio or video exclusively, whereas with law enforcement these systems would have to work with a very broad range of types of data, such as emails, images, audio, video. Other potential challenges include ensuring accurate meta-data creation, the variety of content, the addition and contextualization of new content, explanation of recommendations, and the 'cold start' – starting from nowhere and having to link cases, individuals and evidence that may be relevant.

Finding relevant information in large law enforcement databases is often like searching for a proverbial 'needle in the haystack'

3.4 MAJOR EVENTS SCREENING, SURVEILLANCE AND BEYOND JAPAN

The Government of Japan has initiated cooperation with several tech companies as part of an AI development policy to maximize security and prevent crime or terrorism interfering with the success of major events. As the next host country for the Olympic Games, for instance – an event which is expected to bring an estimated 600,000 overseas visitors – enhancing security is a top priority for national authorities in Japan, including the National Police Agency (NPA). In April 2019, NPA established a new office of "Advanced Technology Planning" that seeks to leverage the beneficial use of advanced technologies, such as AI.

The Police Information Communication Research Centre of the National Police Academy is additionally exploring three pilot applications that could strengthen security surrounding major events. These include the use of AI to: identify the models of cars in surveillance footage, analyse suspicious financial transactions that may indicate the laundering of money, and help identify movements or actions that may be considered suspicious. The Prefecture Police in Tokyo is similarly developing AI-enabled tools in pilot form that focus on identifying areas of high crime risks, which can serve to support in determining optimal patrol routes or crime prevention techniques.

Law enforcement in other prefectures in Japan, such as Kanagawa, have also been working on hotspot statistics to inform predictive policing. Utilizing deep learning methods, prefecture police are developing a tool to detect, analyse and predict the location and time that crimes and accidents are likely to happen based on statistics and relevant data feeds. To make these predictions the tool takes into consideration factors such as time of the day, place, weather, geographical condition, urban mobility and various data feeds from past crimes and accidents in the area, as well as knowledge of criminology. Predictive policing supports police officers in detecting patterns of crimes and accidents; providing holistic overview in an active crime investigation; and, finally, reducing the possibility of crime by implementing a refined patrol route.

Utilizing deep learning methods, prefecture police are developing a tool to detect, analyse and predict the location and time that crimes and accidents are likely to happen based on statistics and relevant data feeds

4.
TAPPING INTO AI RESPONSIBLY

In the first INTERPOL-UNICRI report on AI for law enforcement, the legal and ethical perspectives of the use of AI by law enforcement were broadly introduced and a caveat issued that law enforcement must ensure responsible use of AI. Although the present report aims to be practical and operationally oriented, it is helpful to clarify the term 'responsible' from the outset in order to avoid ambiguous interpretations. The seminal 2019 white paper *AI and Ethics at the Police by Leiden University and TU Delft*[16] suggests that, from a legal perspective, to act responsibly means "to accept moral integrity and authenticity as ideals and to deploy reasonable effort toward achieving them."[17] Striving for moral integrity, in turn, implies "adhering to the values of freedom, equality, and solidarity."[18] For the purposes of this report, however, a more straightforward understanding will be adopted and the term 'responsible' will be framed in line with the Oxford Dictionary, which defines 'responsibly' as acting "in a sensible or trustworthy manner."[19] In this context, the responsible use of AI by law enforcement should be understood as use that enshrines the general principles of respect for human rights, democracy, justice and the rule of law.

To achieve these principles, law enforcement agencies must work to guarantee that the design and use of AI complies with the requirements of fairness, accountability, transparency and explainability (FATE). These requirements have emerged over recent years from a consensus within the AI community about what algorithms require in order to justify placing trust in them and to guarantee appropriate levels of safety. A brief explanation of each requirement follows.

Fairness implies that algorithmic decisions do not create a discriminatory or unjust impact on the end users. Automated decision should not be taken based on attributes, such as ethnicity, gender, sexual orientation, as they may lead to discrimination. At the same time, simply avoiding these attributes is not a solution as they may nevertheless be indirectly derived from other criteria. For instance, even when ´ethnicity' is not a criteria, if people of a specific ethnicity live in a certain area and 'address' is a criteria, the model may still make an unfair determination. Fairness requires that all AI systems are rigorously audited to show compliance with the right to non-discrimination and, in the event discrimination arises, measures to deliver the right to effective remedy must be put in place.

16 Zardiashvili, L., Bieger, J., Dechesne F. and Dignum, V. (2019). *AI Ethics for Law Enforcement: A Study into Requirements for Responsible Use of AI at the Dutch Police.* Delphi.

17 Dworkin, R. (2011). *Justice for Hedgehogs*. The Belknap Press, 111.

18 Ministry of Social Affairs and Employment. (2014). *Core Values of Dutch Society*. Pro Demos, House of Democracy and Constitution. Retrieved from https://www.prodemos.nl/wp-content/uploads/2016/04/ KERNWAARDEN-ENGELS-S73-623800.pdf

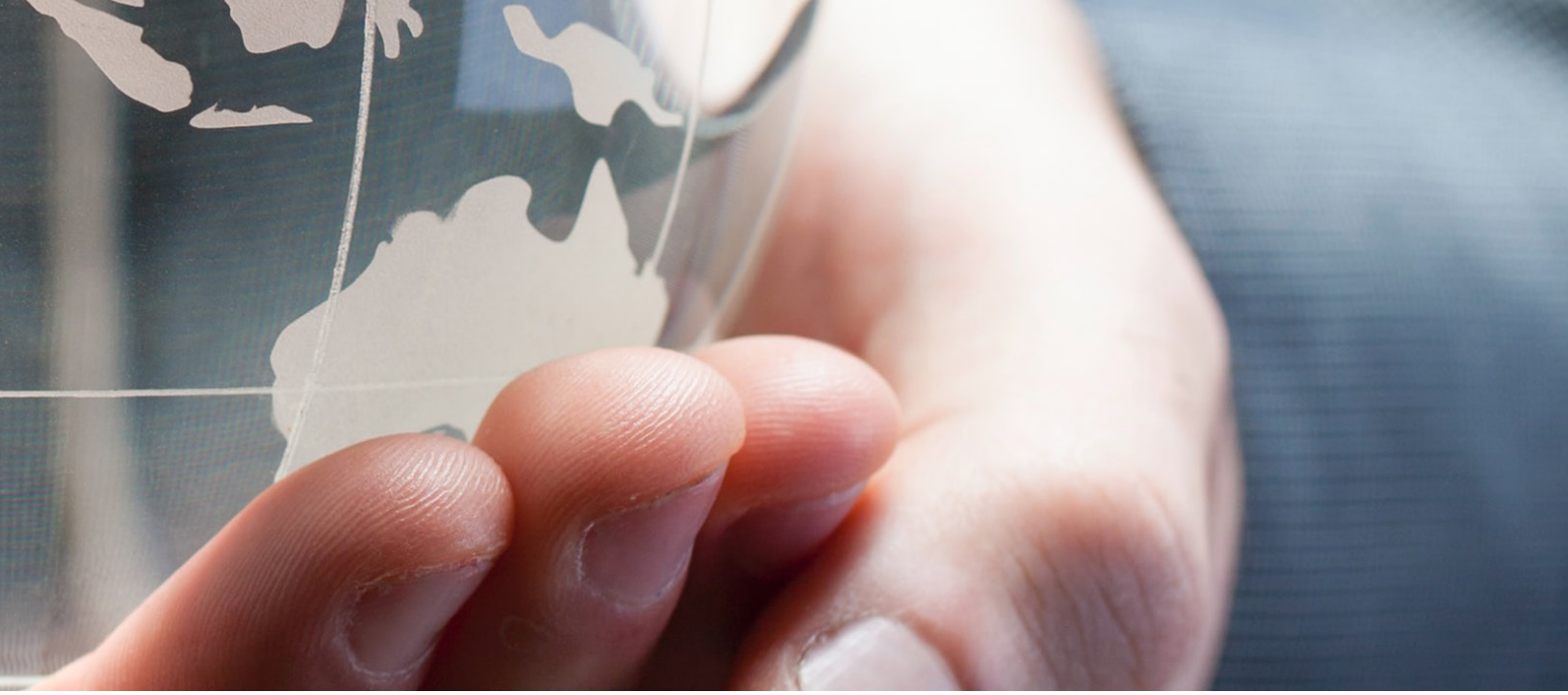
19 Responsibly. (2019). In Oxford Online Dictionary. Retrieved from https://www.lexico.com/en/definition/responsibly

Accountability can be understood as being responsible for an action taken and being able to provide a satisfactory justification for this action. The legal system is built on a fundamental assumption of human agents and so, replacing them with autonomous agents, such AI, throws this system into disarray. Given the difficulty in bringing an autonomous system before a court, the fundamental of who bears the responsibility for actions taken by or informed by an AI system must be asked. Is it the developer, manufacturer or end-user? Accountability requires that clear liability regulations and legal tools are elaborated to process cases with autonomous agents.

Transparency includes providing clear information about the human decisions taken at the time of the building of the model. This goes beyond providing complex 'terms of services' and instead includes matter such as: What is the goal of using AI in a specific context? Which decisions are fully automated? What is the machine learning model being employed? Which data is used? Which features in the dataset are being considered? Are any of sensitive individual attributes being considered? How is data privacy being respected? Transparency requires that questions such as these are clearly elaborated and that those who implement the AI system must be able to answer.

Explainability is closely associated with the requirement of transparency. It differs however in that explainability focuses on ensuring that algorithmic decisions can be understood by end-users in non-technical terms. This concerns the so-called 'black-box' problem. Deep learning systems are literally black boxes that combine and recombine attributes in many arbitrary ways. Once an input is provided, the internal behavior that leads the system to the output may not be clear. Explainability requires that end-users are able to interpret the information extracted from the black box and understand what elements used in the machine learning model were responsible for each specific outome. Unlike the other requirements of fairness, accountability and transparency, explainability is very much a technical challenge for developers and manufacturers. Several groups are however working to develop tools that can explain and present in understandable terms which features in the data were most important for the model, as well as the effect of each feature on any particular output.[20]

20 Kaur, H., Nori, H., Jenkins, S., Caruana, R., Wallach, H. and Wortman Vaughan J. (2020). *Interpreting interpretability: Understanding data scientists' use of interpretability tools for machine learning.* University of Michigan, Microsoft Research.



Beyond these four requirements, the concepts of safety and robustness are also often raised in terms of the general reliability of AI systems and their resilience to attacks and in terms of security. These additional requirements should equally be carefully taken into account by law enforcement for responsible AI. In order to ensure that AI systems are both safe and robust, two major practices need to be institutionalized when developing and deploying AI systems: first, periodical system integrity and updates by both internal engineers, police officers (being the users) and, if necessary, a trusted external partner; and, second, interoperability capability, which can be understood as ensuring that the system is 'easy to operate' by officers and 'feasible to be used with other future systems'.

If the use of AI by law enforcement is carried out in a manner contrary to these high-level principles and requirements, unethically and even illegally, the citizens that law enforcement is tasked to serve and protect are likely to exhibit adverse reactions and feel threatened by the law enforcement's use of this technology. These reactions may engender a resistance to and vocal criticism of the use of AI applications and other advanced technologies by law enforcement. To prevent this and ensure that that law enforcement can continue to tap into the positive potential of AI, public trust must be persevered.

Given the importance of this, a series of panel presentations were organized during the second Global Meeting to dive deeper into some of the legal and social aspects of the use of AI by law enforcement, as well as some of the developments in the ongoing ethical discourse. The sections that follow have been informed by these presentations.

4.1 LAWFULNESS

There are significant legal challenges for law enforcement in the field of AI. If law enforcement fails to overcome these challenges, the use of AI may infringe fundamental human rights, such as the right to privacy, equality and non-discrimination, as well as undermine principles of law, such as the presumption of innocence, privilege against self-incrimination and proof beyond a reasonable doubt. First and foremost, these concerns require that law enforcement ensures that its use of AI is in accordance with the law. Given that AI is fuelled by data this concerns, in particular, those laws concerning data privacy – which includes regulations on data collection – and data protection – which includes regulations on specific data retention, storage and processing of data. Furthermore, not only must law enforcement ensure conformity with the law at the time of use of any AI systems, but it must also ensure that this was the case in the context of its development.

There are two overarching legal instruments which regularly arise in discussion about the use of AI, not only by law enforcement, but also by other communities of end-users, namely: the EU General Data Protection Regulation (GDPR) and the EU Law Enforcement Directive 2016/680 (LED).[21] Although neither instrument was developed specifically for AI, as they concern data they end up being directly relevant for the development and implementation of AI. Moreover, given that GDPR is concerned with the personal information of citizens of the EU and is, in this regard, applicable to entities operating both within and outside the EU, GDPR in particular is an instrument of global relevance and merits close attention.

Adopted by both the European Parliament and the European Council in April 2016 and subsequently entering into force in May 2018, GDPR is a result of more than four years of discussion and negotiation and was designed to modernize laws that protect the personal information of individuals. It contains six core principles for the collection and processing of personal data: 1) Lawfulness, fairness and transparency; 2) Purpose limitation; 3) Data minimization; 4) Accuracy; 5) Storage limitation and 6) Integrity and confidentiality (security).[22]

The LED, also known as Police Directive, entered into force in May 2016 and aims to apply the rules governing personal data in GDPR to the activities of law enforcement. It has been heralded for its role in building "an area of freedom, security and justice with a high level of data protection, in accordance with the EU Charter of Fundamental Rights." Aiming at protecting individuals' personal data, while guaranteeing a high level of public security, the LED provides rights for data subjects, as well as obligations for "competent authorities" when processing data for "law enforcement purposes", i.e., prevention, investigation, detection, prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security.

As noted, neither GDPR nor the LED were adopted with AI in mind and, accordingly, some crucial provi-

21 See *EU Law Enforcement Directive 2016/680* available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016L0680

22 See *GDPR* Article 5(1), which sets out the six principles.

sions have yet to be tested in the context of AI in policing. In this regard, a number of possible situations regarding GDPR and the LED requirements have arisen and generated considerable debate.[23]

The scope of the restrictions on automated processing under GDPR and the LED is one such debated issue. Decisions based solely on automated processing, so-called automated decision-making (ADM) systems, which are increasingly used in predictive policing to analyze data to help predict either where crimes will occur or who will be involved in crime, raise practical concerns in terms of liability and accountability. The admissibility of ADMs before a court is equally contentious.

Although GDPR does not explicitly mention AI, it does reference the role of autonomous decision-making and, under Article 22, implies a right for explanation for automated decision-making, including profiling, which means that "controllers will need to design, develop and apply their algorithms in a transparent, predictable and verifiable manner". Interestingly, GDPR states that "the data subject shall have the right not to be subject to a decision based solely on automated processing", except if it is based on the data subject's explicit consent. In this context, adequate information should be provided to participants and/or generic or synthetic data should be used wherever possible.

The LED also specifically covers ADM systems under Article 11, which provides that ADM systems that produce an adverse legal effect on the data subject or significantly affects him or her, should be prohibited, unless authorized by EU or Member State law under which the controller is subject and which provides appropriate safeguards for the rights and freedoms of the data subject – at least the right to obtain human intervention on the part of the controller to express his or her point of view and to contest the decision. However, in accordance with the LED, profiling that results in discrimination against natural persons based on the processing of sensitive data, may not, under any circumstance, be authorized.

Thus, in order to enable effective legal protection, law enforcement agencies must be able to provide an explanation of an individual decision, and not just the logic behind it which can prove difficult in the case of ADM systems.

In light of the difficulties law enforcement may encounter in attempting to navigate these legal waters, legal experts should be involved in the processes of development and utilization of AI. This particularly the case where ADM systems are being considered for use in decision-making in order to ensure that no adverse effects concerning the data subject or any other individuals occur. Moreover, in light of some of the legal gaps, law enforcement agencies and other relevant national authorities may wish to consider developing specific and/or tailoring existing law enforcement regulations on the use of AI to frame and guide how it should lawfully act, catering for existing instruments of global interest, such as GDPR and the LED.

23 Hidvegi, F. Massé, E. (2018, Nov 1). *Mapping regulatory proposals for artificial Intelligence in Europe*. Retrieved from https://www.accessnow.org/cms/assets/uploads/2018/11/mapping_regulatory_proposals_for_AI_in_EU.pdf

4.2 SOCIAL ACCEPTANCE

Being a relatively new technology, people are often not fully aware of how AI really works and what it can and cannot do. This allows for worry, fear and concern to breed and, under the wrong circumstances, these feelings can hinder the integration and application of the technology. Arguably, nowhere is this more critical than when it comes to the use of AI in the public sector and, law enforcement in particular. Effective policing is very much predicated upon the trust of the community. Public safety can be jeopardized when communities lose trust in law enforcement. In this regard, social acceptance of the use of AI by the public is of paramount importance as law enforcement increasingly integrates AI. It is essential that law enforcement remains conscious of the need for this, as well as the importance of communication and information-sharing with key stakeholders and the general public.

A survey, conducted in 2018 by the Center for Higher Studies of France's Ministry of Interior, the University Jean-Moulin Lyon 3, the French National Police College and the French National Gendarmerie College, examined social acceptance by evaluating public trust toward the use of predictive policing by homeland security actors.[24] Of all AI applications being explored by law enforcement, predictive policing is regularly presented as the source of greatest concern. The project entailed a public survey of more than 2,000 individuals about public knowledge and opinion about predictive policing – both before and after they have received an explanation about the topic. Results showed that after a 150-words explanation, citizens were 28% more confident about law enforcement agencies using predictive policing – rising from 59% to 87%. More specifically, this included an increase in the number of respondents agreeing that predictive policing can be useful for preventing crimes after receiving the explanation, along with an increase in the number of respondents disagreeing that predictive policing is a threat for civil liberties. Furthermore, 59% of respondents indicated that they would even accept an identity-check based solely on predictive policing software.

While these results are indicative of the power of an explanation and better understanding, even if limited, the results should, however, be interpreted carefully. The social acceptance of predictive policing in this case may not necessarily be fully motivated by a newly acquired rational and objective under-

24 Piotrowicz, C. (2018). *Predictive Policing: European Law Enforcement Research Bulletin*, (4 SCE), 107-111. Retrieved from https://bulletin.cepol.europa.eu/index.php/bulletin/article/view/374

standing or appreciation of the nuances of the technology. It may instead be motivated by exceptional circumstances, such as the spate of terrorist incidents in France between 2015 and 2018, that preceded the survey, or misinformation regarding the technology. Concerning misinformation, it is notable that the study indicated that 44% of respondents justified the use of predictive policing to fight terrorism, while this is not per se the true purpose of the technology.

In order to communicate appropriately with the public, it is also essential to understand how best to reach stakeholders. In this regard, the survey also revealed that the public tends to prefer to get information about the use of AI by law enforcement from an independent authority or the government, rather than from academia or the private sector. On the other hand, law enforcement officers prefer to get professional training from an officer that has field-experience with the technology.

Communication and information-sharing are evidently critical ways of heightening social acceptance of the use of AI by law enforcement, but there are also several other ways which can be approached in parallel. These include the development of legislation or the implementation of the existing legislation; striking a balance between what amount of data is necessary and adequate for a useful output and what amount of details are excessive in relation to the purposes of processes – the so-called data minimization principle; data anonymization or the use of dummy or generic data wherever possible; seeking and obtaining explicit consent from participants or data subjects; and involving the public – particularly vulnerable groups – in the development and use of AI systems for law enforcement.

Finally, it is important that AI systems built for law enforcement should be developed and deployed with the mindset that mistakes cannot be prevented entirely, and that errors may inevitably occur. A risk assessment and mitigation scenario should, accordingly, be developed from the outset and discussed with the public.

4.3 THE ETHICS OF AI

AI ETHICS IN THE EUROPEAN UNION

In the context of the EU, there have been several noteworthy recent developments, through which the European Commission has sought to address these ethical, legal and social issues, and to ensure that AI systems remain human-centric and are, at all times, aimed at maximizing the benefits of this technology while preventing or minimizing its risks.

In April 2018, 24 Member States of the EU signed a *Declaration on Cooperation on Artificial Intelligence*, through which the need to develop an adequate legal and ethical framework and to cooperate to this end were identified[25]. Following this, in June 2018, the European Commission established a High-Level Expert Group on AI (AI-HLEG) – an independent group comprised of 52 eminent representatives from academia, industry and civil society. The group was tasked with elaborating recommendations on future-related policy development and on ethical, legal and societal issues related to AI, including socio-economic challenges. To this end, the AI-HLEG released its *Ethics Guidelines for Trustworthy AI* in April 2019[26]. While the European Commission is not the first player to release such guidelines, the Ethics Guidelines marks the first government-led initiative in this domain and is a significant step toward bringing not only the human-centric approach to AI to the global fore but also building international consensus around the notion of AI ethics.

According to the Ethics Guidelines, for AI to be considered trustworthy, it must, throughout its entire lifecycle, be: 1) Lawful, which entails complying with all applicable laws and regulations; 2) Ethical, which aims at ensuring alignment with ethical norms, and 3) Robust, both from a technical and social perspective. The development of trustworthy AI systems should be based upon established fundamental values, such as the respect for human dignity, democracy, justice and rule of law, while, at the same time, guaranteeing the freedom of the individual and citizens' rights in order to ensure equality and non-discrimination. On this basis, the Ethics Guidelines present four overarching ethical principles underlying the development, deployment and use of AI systems: respect for human autonomy; prevention of harm; and fairness and explicability.

25 European Union (2018). *Declaration on Cooperation on Artificial Intelligence.* Retrieved from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50951

26 High-Level Expert Group on AI (2019). *Ethics Guidelines for Trustworthy AI.* Retrieved from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419

Naturally, certain tension may arise between the above principles when attempting to put them into practice. Consider, as noted in the Ethics Guidelines, the use of AI for predictive policing. While this may help to reduce crime, it may also entail surveillance that encroaches upon individual liberty and privacy, thereby bringing the principle of prevention of harm and the principle of human autonomy into conflict, and necessitating deliberation on the most appropriate trade-off. Care must be taken to appropriately identify, evaluate, document and communicate these trade-offs and their solutions in a methodical manner.

A series of seven requirements for the development, deployment and use of AI systems in a trustworthy manner are further identified, namely: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity; non-discrimination and fairness; societal and environmental well-being; and accountability.

To implement those requirements, the Ethics Guidelines present both technical and non-technical methods. With respect to the former, trustworthy AI architectures should implement ethics and rule of law by design and privacy-by-design, as well as tests for validation of the system and quality of service indicators. Non-technical methods for implementing the requirements include: regulation; codes of conduct; standardization; certification; and participation in and efforts in terms of accountability, via, for instance, governance frameworks, education on and awareness-raising to foster an ethical mindset, stakeholder participation, social dialogue and the establishment of diverse and inclusive design teams. As a complement to these requirements, the Ethics Guidelines also contains an assessment list, or series of non-exhaustive questions intended to operationalize the key requirements and determine whether AI in any given use case can be considered trustworthy as per the Ethics Guidelines. In June 2019, the European Commission piloted the assessment list, inviting stakeholders to test it and provide feedback to be integrated into a revision of the assessment list. Notably, from a law enforcement perspective, the Netherlands National Police participated in the pilot phase.

In February 2020, the European Commission released a White Paper *On Artificial Intelligence – A European Approach to Excellence and Trust*, which signals the start of the process of the development of possible AI legislation in EU[27]. Notably, the White Paper builds on the ground work done by AI-HLEG in developing the Ethics Guidelines and again underscores that AI must be human-centric, ethical, sustainable and respects fundamental rights and values. It also suggests specific legal requirements such as, AI being trained on representative data, companies keeping detailed documentation on how the AI was developed and citizens being kept informed when they are interacting with an AI system.

Although a European initiative, the Guidelines, and any subsequent efforts to further operationalize them in the form of EU legislation, are likely to reach beyond Europe. As the Guidelines themselves note, their aim is to foster reflection and discussion on an ethical framework for AI at a global level. In this regard, the relevance of this European approach to creating trustworthy AI systems at the service of humanity should be carefully noted.

OTHER AI ETHICS INITIATIVES

Discussions on the ethics of AI and the development of responsible AI in line with fundamental rights are not limited to Europe. Several public sector organizations, research institutions and private companies have issued statements of principles and guidelines or set up expert committees on AI to produce draft policy documents on how to approach AI. A recent study by Nature identified 84 documents containing ethical principles or guidelines for AI.[28] Notwithstanding the possibility of overlap, each of these documents is, in itself, a valuable instrument and constitutes a point of reference for law enforcement going forward with the responsible development and application of AI.

Perhaps most notably, the Organisation for Economic Co-operation and Development (OECD) established an expert group in May 2018 to elaborate principles on AI in society and subsequently adopted its *Principles on Artificial Intelligence* in May 2019 – the first set of intergovernmental policy guidelines on AI.[29] The Principles, which were adopted by 42 States, focus on promoting AI that is innovative and trustworthy and that respects human rights and democratic values. On a national level, several committees and expert groups have also been established to explore the ethical dimensions to AI, such as the Advisory Council on the Ethical Use of Artificial Intelligence and Data in Singapore and the Select Committee on Artificial Intelligence of the House of Lords of the United Kingdom.

Similar efforts are taking place in the private sector, especially among corporations who rely on AI for their business. Notably, major industry entities such as Google,[30] IBM[31] and Microsoft[32] have all established ethical principles upon which they will proceed to explore the application of AI. According to Google's principles, AI should be socially beneficial; fair by avoiding to create or reinforce bias; built and tested for safety; accountable to people and also should incorporate privacy by design principles; uphold high standards of scientific excellence; and be made available for uses that are in line with these principles. Google also defined that it will not pursue applications of AI such as "weapons or other technologies whose principal purpose or implementation is to cause or directly facilitate injury to people" or "technologies that gather or use information for surveillance violating internationally accepted norms". Deepmind, an AI research lab focused on deeplearning under the auspices of Google, has also created its own ethical advisory body and set out separate principles and good practices to develop its technology responsibly.

27 European Commission (2020). *White Paper: On Artificial Intelligence – A European Approach to Excellence and Trust.* Retrieved from https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

28 Jobin, A., Ienca, M. & Vayena, E. (2019). *The global landscape of AI ethics guidelines.* Nature Machine Intelligence 1, 389–399.

29 OECD. (2019). *Principles on Artificial Intelligence.* Retrieved from https://www.oecd.org/going-digital/ai/principles/

30 Google. *Artificial Intelligence at Google: Our Principles.* Retrieved from https://ai.google/principles/

31 IBM. (2019, Jan 17). *Coming soon: EU Ethics Guidelines for Artificial Intelligence*. Retrieved from https://www.ibm.com/blogs/policy/ai-ethics-guidelines/

32 Microsoft. (2020). *Our approach to responsible AI.* Retrieved from https://www.microsoft.com/en-us/ai/our-approach-to-ai

There has also been considerable development in terms of discussions on responsible AI within specific sectors. By way of example, in the financial sector, the Monetary Authority of Singapore (MAS) published its *Principles to Promote Fairness, Ethics, Accountability and Transparency in the Use of AI and Data Analytics in Singapore's Financial Sector* in November 2018 to provide guidance to firms offering financial products and services on the responsible use of AI and data analytics in order to strengthen internal governance around data management and use and, ultimately, to foster greater confidence and trust in the use of AI in this sector.[33] The Netherlands Central Bank (De Nederlandsche Bank) similarly released its own guidance document in July 2019, containing principles for the responsible use of AI in the financial sector to prevent any harmful effects for banks, their clients, or the credibility or reputation of the financial sector as a whole.[34] In the healthcare sector, there has also been movement in terms of responsible AI. For instance, the Royal Australian and New Zealand College of Radiologists proposed a framework in August 2019, containing nine ethical principles that are intended to guide all stakeholders involved in research or deployment of AI in medicine, including developers, health service executives and clinicians.[35] In the United States, the American Medical Association also advocates for a regulatory framework for the evolution of AI in health care since June 2018.[36]

33 Monetary Authority of Singapore. (2018). *Principles to Promote FEAT in the Use of AI and Data Analytics in Singapore's Financial Sector.* Retrieved from https://www.mas.gov.sg/~/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf

34 Van der Burgt, J. (2019). *General principles for the use of Artificial Intelligence in the financial sector.* De Nederlandsche Bank. Retrieved from https://www.dnb.nl/binaries/General%20principles%20for%20the%20use%20of%20Artificial%20Intelligence%20in%20the%20financial%20sector_tcm46-385055.pdf

35 The Royal Australian and New Zealand College of Radiologists. (2019). *Ethical Principles for Artificial Intelligence in Medicine*, Version 1. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=2ahUKEwjauKXQ1eLnAhVB3KQKHRzBDO0QFjABegQIBhAB&url=https%3A%2F%2Fwww.ranzcr.com%2Fdocuments%2F4952-ethical-principles-for-ai-in-medicine%2Ffile&usg=AOvVaw3IeDZFrKpdX0wvdOq6cdK6

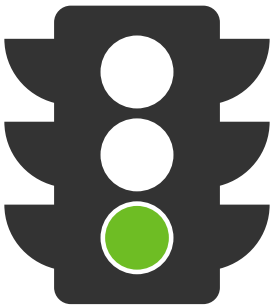
36 Crigger, E. & Khoury, C. (2019). *Making Policy on Augmented Intelligence in Health Care*. AMA Journal of Ethics. Retrieved from https://journalofethics.ama-assn.org/article/making-policy-augmented-intelligence-health-care/2019-02

5. POLICING NEEDS AND RECOMMENDED ACTIONS

Bearing in mind the preceding legal, social and ethical considerations, participants at the second Global Meeting were tasked with identifying some of the most pressing needs and recommended points of actions to be taken by policymakers and/or intergovernmental organizations to support law enforcement in the development, deployment and use of AI. A summary of the feedback collected follows.

Needs:

- Establish a common language and baseline between law enforcement, industry, academia and civil society stakeholders concerning the use of AI by law enforcement;
- Increase collaboration between law enforcement, industry, academia and civil society stakeholders;
- Share ideas regarding use cases and perspectives on the adoption of AI;
- Devise mechanisms for and collaborate on the evaluation of the use of AI in pilot projects;
- Standardize data, data collection, data protection measures;
- Develop common approaches for the anonymization of data; and,
- Set-up digital platforms for knowledge-sharing, project updates, R&D insights, and funding opportunities.



RECOMMENDED ACTIONS

Define a statement of principles on the use of AI in law enforcement that will guide law enforcement to ensure respect for human rights, democracy, justice and the rule of law and support it to prioritize the key requirements of fairness, accountability, transparency and explainability, as well as safety and robustness;

- Develop guidance for law enforcement on the implementation of new technology to support and encourage law enforcement agencies to explore and invest in new AI opportunities and to develop training in new AI applications and disseminate best practices;
- Create a knowledge-base with the law enforcement community on the requirements for the adoption of AI, such as what kinds of problems AI is capable of tackling, the current or inherent limitations and the resources (tools, data, expertise, computing power) required to implement AI solutions;
- Develop guidance for law enforcement on the admissibility of AI in court that assesses the impact and results of the specific use of AI in courts, while ensuring the respect for human rights and rule of law;
- Create an expert advisory committee that can provide guidance to law enforcement in terms of legislation and serve as a forum for discussing appropriate legislative models with legal experts and other key stakeholders;
- Identify an external global body to provide advisory support to law enforcement on ethical issues and to provide support in carrying out audits to check whether a system is responsible and complies with legal requirements;
- Foster a community and organize training courses and workshops to attract and connect different stakeholders from law enforcement, industry, academia, civil society and international bodies with the diverse backgrounds and essential perspectives to gather and synthesize views from cross-sections of society, in order to provide a balanced and facts-based picture of the opportunities and challenges of the use of AI and to highlight the application of AI to law enforcement and provide hands-on support.
- INTERPOL and UNICRI agreed to remain seized of these needs and actions and will seek to build upon them in forthcoming global meetings on AI for law enforcement.

6.
RESPONSIBLE AI INNOVATION TOOLKIT FOR LAW ENFORCEMENT

As has been noted, discussions on the responsible uses of AI are growing among States and throughout the private sector. At the time of drafting, more than 30 States have adopted national AI strategies or action plans since 2016, a large percentage of which highlight the importance of the ethical considerations to the use of AI. There have also been several notable developments in terms of statements of principles, such as the OECD's *Principles on Artificial Intelligence*, and proposed ethical frameworks, such as the European Union's *Ethics Guidelines for Trustworthy Artificial Intelligence*. The second Global Meeting set the stage for the law enforcement community to also take action, with participating representatives confirming the importance of responsible AI. They further clearly identified that law enforcement requires support and guidance to facilitate its adoption of AI and, in doing so, to avoid the many pitfalls.

Following the identification of law enforcement needs and requirements, and in response to growing global pressures, INTERPOL and UNICRI recognized that a toolkit for responsible AI innovation in law enforcement would be a valuable contribution in terms of providing support and guidance.

A 'toolkit' is considered to be the preferred format because as it would, departing from existing proposed approaches of 'guidelines', 'regulations' and 'frameworks,' seek to stimulate the positive potential of AI within the law enforcement community to develop, deploy and use AI systems, while providing recommendations to prevent any harmful effects. More specifically, it would help law enforcement to tap into AI in order to derive the most benefit from this technology in a lawful and trustworthy manner, rather than raising concerns regarding implementation, which would contribute further to creating conditions that undermine social acceptance.

The focus of the toolkit could include:

- A general explanation of AI, including a relevant working definition for law enforcement;
- Guidance on the use of AI for law enforcement, including the identification and compilation of major technology domains and possible use-cases;
- Considerations of examples or best practices of trustworthy, lawful and responsible use of AI in law enforcement – synthesis of important requirements, such as fairness, accountability, transparency and explainability, for consideration when a law enforcement agency intends to develop an AI-enabled project (in-house) or procure an AI-tool/system (external solutions) – and a series of recommended good practices that reflect the general principles and seek to build trust and social acceptance;

- Step-by-step recommendations for the development, implementation and maintenance of an AI-enhanced system in law enforcement, including a checklist of operational considerations to ensure that the system/operation is adherent to the aforementioned consideration.

In order to add the most values for law enforcement, it is useful to underscore the purpose, structure and level of abstraction of the document. First of all, as a toolkit, it should serve as a reference guide for law enforcement and, under no circumstances, should be considered compulsory or binding in nature. At all times, it should be practical and operational oriented and must avoid falling into conceptual discussions. There is no added value in the toolkit seeking to redefine well-established legal, ethical and social discussions surrounding the use of AI by law enforcement. Finally, the toolkit should seek to build upon work already done and avoid being just one more set of guidelines.

To ensure that the toolkit can be operationalized, it should furthermore have a clearly defined target audience and should specifically be communicable to, at least, the following target audiences that will play a central role:

- Senior police managers or key decision-makers in law enforcement agencies;
- Law enforcement officers responsible for innovation and the use of technology in their respective agencies;
- R&D officers tasked with developing AI capabilities in-house or that can influence outsourcing procurement processes; and
- Legal officers or advisors who provide counsel on the laws and regulation concerning specific use of AI in policing work.

In addition to these specific target audiences, the toolkit should cater for and be approachable by members of the general public in order to foster a sense of openness and transparency regarding the use of AI by law enforcement and, in doing so, build public trust.

Specific sections can be designed for each of the individual target audiences to provide tailor directives to frame and guide each target audience for their role in developing, deploying and using responsible AI systems. Taking the target audiences into consideration, the toolkit should seek to maintain a balance in the discussion between conceptual, operational and technical languages.

7. ROADMAP FOR ACTION

The development of a toolkit on responsible AI innovation for law enforcement is recommended as an essential prerequisite for law enforcement's continued exploration, application and integration of AI. By adopting a principle-based approach to building its AI capacities, law enforcement can lay the foundations for ensuring lawfulness and the public trust necessary for social acceptance.

As a follow-up of the second Global Meeting on AI for Law Enforcement, INTERPOL and UNICRI have carried out qualitative research and interviews with several global experts from different backgrounds – law enforcement agencies, industry, academia and other international bodies alike – to receive proper feedback about the feasibility of this work and how it should be framed in order to add value for its use by law enforcement agencies. The development of the toolkit, alongside the continued identification of law enforcement use cases, will be the focus of the third INTERPOL-UNICRI Global Meeting on AI for Law Enforcement in 2020.

While INTERPOL and UNICRI will lead the development of the toolkit, the approach adopted will be open, transparent and participatory in nature, particularly with respect to refinement of the principles and requirements that underpin the toolkit. In this regard, the present report constitutes an open call for engagement of interested parties for this process, in order to generate the most comprehensive output possible.

To further support this process, INTERPOL and UNICRI have also identified a series of related steps that could be taken by the international law enforcement community, policy-makers and intergovernmental organizations to support and feed into the process of securing responsible AI innovation for law enforcement. These include, but are not limited to:

- Continuing to build a forum and network of focal points within law enforcement at the national level for the purposes of sharing information on AI use cases, experiences and practices and facilitating the discussion on the responsible design, deployment and use of AI by law enforcement;
- Conducting comprehensive mapping of AI capabilities vendors, and levels of adoption across key law enforcement functions and with respect to key domains of crime;
- Developing a database on AI use cases for law enforcement and identifying commonalities in use cases;
- Coordinating and gathering experts to review and analyse existing policies, regulations, legislation, rules and procedures pertinent to the use of AI by law enforcement, and assessing readiness to adopt AI;
- Identifying needs and requirements for the development of an AI R&D programme;
- Organizing technical multi-stakeholder workshops, bringing together law enforcement, academia, industry and civil society organizations, focused on the development of guidelines and the exploration of further possible AI uses cases;
- Organizing public workshops to build public trust and social acceptance, inviting critical feedback from the public regarding the use of AI by law enforcement; and,
- Providing support for framing national standard operating procedures on the design, deployment and use of AI and associated technical materials for the delivery of national workshops in order to support the implementation of the toolkit.

ANNEX I
TERMINOLOGY

Understanding the opportunities and risks of AI may seem like an insurmountable challenge to a lay-person. Indeed, one of the reasons for creating misinterpretations about AI is the lack of understanding of this technology. A veil of confusion surrounds these subjects, which is due in part to their complex technical nature, but also to the jargon and buzzwords used in connection with AI.

In order to help law enforcement to pierce this veil and demystify some concepts to the issue of AI, some of the terms referred to in this report are described below. Examples are provided to facilitate the understanding of more complex concepts. These descriptions should not however be taken as definitions. It is beyond the scope of this report to provide a definition of these terms.

Algorithm: In computer science, an algorithm is a set of instructions that define a sequence of operations to perform a computation, in other words, it consists of the programming scripts behind any software.

Autonomous System: A system that can perform programmed tasks without the need of any human intervention. There are also semi-autonomous systems, which need human intervention at some point in its functioning.

Artificial Intelligence: A sub-field of computer science dedicated to the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, translation between languages, decision-making, and problem-solving. In contrast with other computer software, AI systems do not require explicit instructions from humans, but extract patterns and learn implicit rules from a considerable number of examples included in a database. AI applications are specialized at doing one particular task and in some cases, can even surpass human capabilities.

Big Data: Datasets that are too large or complex to be dealt with by traditional data-processing application software.

Computer Vision: An interdisciplinary scientific field that includes methods for acquiring, processing, analyzing and extracting information from digital images or videos.

Database: An organized collection of data, generally stored and accessed electronically from a computer system.

Deep Learning: A subfield of machine learning that uses artificial neural networks, algorithms inspired by the human brain, to learn from large amounts of data. Deep learning algorithms perform a task repeatedly, each time making minor modifications to its internal features to improve the outcome. The term 'deep learning' results from the several (deep) layers of the neural networks.

Deepfakes: A contraction of '*deep* learning and *fakes*'. A synthetic media in which images, audio or videos of people and events are generated or manipulated using generative neural network architectures. Deepfake algorithms leverage deep learning to generate visual and audio content that are difficult for humans to distinguish from authentic ones.

Facial Recognition: A system capable of identifying persons of interest from images or videos by comparing and analyzing patterns, shapes and proportions of their facial features and contours with faces within a database.

Internet of Things: A system of interrelated computing devices, including laptops, smartphones, sensors among others that transfer data over a network. A "smart home" is an example of the use of the Internet of Things, covering devices and appliances (such as lighting fixtures, thermostats or home security cameras) that can be controlled via devices associated with that ecosystem, such as smartphones.

Machine Learning: A subfield of AI that uses statistical techniques to give computer systems the ability to "learn" from data, i.e., progressively improve performance on a specific task. Machine learning algorithms do not require explicit programming instructions but rely on patterns and inference from an enormous number of examples, known as "training data". Once a mathematical model that correlates those examples is built, the machine learning algorithm is able to make predictions or decisions about new unseen examples, the "test data".

Malware: A contraction of '*mal*icious soft*ware*'. Malware is any piece of software intentionally designed to cause damage or steal data from a computer, server or computer network. Viruses, Trojans, spyware, and ransomware are examples of different kinds of malware.

Natural Language Processing: A subfield of computational linguistics and AI concerned with processing and analyzing large amounts of natural human language data. Tasks in natural language processing frequently involve speech recognition, natural language understanding, natural language generation and translation between languages.

Neural Networks: The base model of deep learning. Inspired by biological neurons, these algorithms use multiple layers of single units to progressively extract higher level features from the raw input. For example, if the input is an image, the first layers of the neural network may identify lines and curves, while the last layers may identify letters or faces.

Phishing: An email or electronic communications scam to trick people into downloading malicious software or to obtain sensitive information such as account credentials or financial information. Phishing attacks are not personalized and are usually sent to masses of people at the same time. Spear-phishing on the other hand specifically targets a victim.

Robotics: A branch of engineering that focuses on the development of robots - a machine capable of carrying out a complex series of actions that can be remotely operated or autonomous. Robotics encompasses the design, construction, operation, and application of robots, as well as computer systems for their control, sensory feedback, and information processing such as AI.

Speech Recognition: A computer software with the ability to convert audio speech to text information, i.e. it enables the recognition and translation of spoken language into text.

Supervised Learning: The most common sub-branch of machine learning, which consists on learning how to map an input to an output label, based on a number of examples of input-output pairs. For a certain number of input features (for example, fruit, red, round) a label is provided (for example, apple). Through the use of machine learning algorithms, the training dataset is used to build a mathematical model which enables predicting the classification of unlabelled data.

Unsupervised Learning: A machine learning task that finds patterns in data that have not been labelled, classified or categorized. It can be used for clustering analysis, which consists of grouping data elements with similar features.

ANNEX II
LIST OF ABBREVIATIONS

ADM	Automated Decision-Making
AFP	Australian Federal Police
AI	Artificial intelligence
AI-HLEG	High-Level Expert Group on AI
AiLECS	Artificial Intelligence for Law Enforcement of Community Safety
CCTV	Closed-Circuit Television
EU	European Union
FATE	Fairness, Accountability, Transparency and Explainability
GDPR	General Data Protection Regulation
GPS	Global Positioning Services
IC	INTERPOL's Innovation Centre
IEDs	Improvised Explosive Devices
IGCI	INTERPOL's Global Complex for Innovation
INTERPOL	International Criminal Police Organization
LED	Law Enforcement Directive 2016/680
MAS	Monetary Authority of Singapore
NLP	Natural Language Processing
non-POI	non-Person of Interest
NPA	Japan National Police Agency
OECD	Organisation for Economic Co-operation and Development
R&D	Research and Development
UAV	Unmanned Aerial Vehicles
UNICRI	United Nations Interregional Crime and Justice Research Institute
ZITiS	Central Office for Information Technology in the Security Sector

ABOUT INTERPOL

INTERPOL is the world's largest international police organization. Its role is to assist law enforcement agencies in its 194 Member Countries to combat all forms of transnational crime. INTERPOL works to help police across the world meet the growing challenges of crime in the 21st century by providing a high-tech infrastructure of technical and operational support. Its services include targeted training, expert investigative support, specialized databases and secure police communications channels.

Located in Singapore, within the INTERPOL Global Complex for Innovation (IGCI), INTERPOL's Innovation Centre (IC) works to create strategic partnerships with law enforcement, academia and the private industry on a global, regional and national level. These collaborations support INTERPOL in developing innovative solutions to policing threats and challenges.

ABOUT UNICRI

The United Nations Interregional Crime and Justice Research Institute was established in 1968. Within the broad scope of its mandate, the Institute contributes, through research, training, field activities and the collection, exchange and dissemination of information, to the formulation and implementation of improved policies in the field of crime prevention, justice and emerging security threats, due regard being paid to the integration of such policies within broader policies for socio-economic change and development, and to the protection of human rights.

In 2017, UNICRI opened its Centre for Artificial Intelligence and Robotics in The Hague, the Netherlands, with a view towards advancing understanding of artificial intelligence, robotics and related technologies vis-à-vis crime prevention, criminal justice, the rule of law and security. The Centre seeks to share knowledge and information on the potential beneficial applications of these technologies and to contribute to addressing any harmful effects and the malicious use.