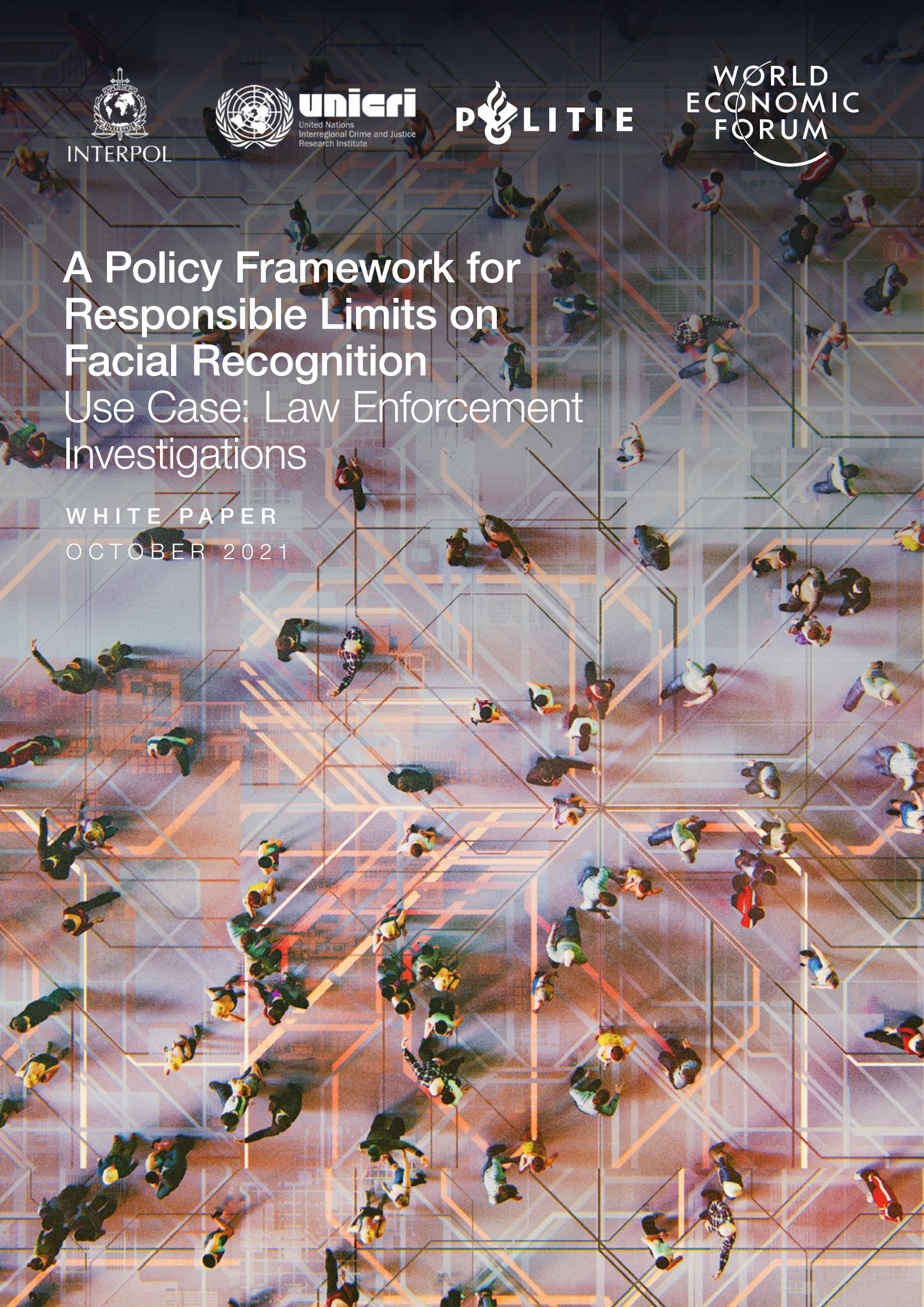


INTERPOL

unicri
United Nations
Interregional Crime and Justice
Research Institute

POLITIE

WORLD ECONOMIC FORUM

A Policy Framework for Responsible Limits on Facial Recognition

Use Case: Law Enforcement Investigations

WHITE PAPER

OCTOBER 2021

Inside: Getty images

Contents

Foreword 3

Introduction 4

Methodology 6

1 Law enforcement investigations: use cases and definitions 7

2 Proposed principles 14

3 Proposed self-assessment questionnaire 20

Conclusion 25

Glossary 26

Contributors 28

Endnotes 30

Foreword

Irakli Beridze
Head of the Centre for Artificial Intelligence and Robotics, UNICRI

Marjolein Smit-Arnold Bik
Head of the Special Operations Division, Police of the Netherlands

Kay Firth-Butterfield
Head of Artificial Intelligence and Machine Learning; Member of the Executive Committee, World Economic Forum

Cyril Gout
Director of Operational Support and Analysis, INTERPOL

Remote biometric technologies – in particular facial recognition – have gained a lot of traction in the security sector. In recent years, the accuracy of this technology has significantly increased thanks to the growth of the internet of things, the ubiquity of smartphones and the proliferation of smart city projects.

Law enforcement agencies could benefit greatly from these technologies to resolve crimes and conduct faster investigations. But, improperly implemented or implemented without due consideration for its ramifications, facial recognition could result in major abuses of human rights and harm citizens, particularly those in underserved communities.

The rapid adoption of facial recognition raises multiple concerns, mainly related to the possibility of its potential to undermine freedoms and the right to privacy. To address and mitigate these risks, policies have started to emerge over the past year.

The organizations that worked together to create this paper, the World Economic Forum, the International Criminal Police Organization (INTERPOL), the United Nations Interregional Crime and Justice Research Institute (UNICRI) and the Police of the Netherlands, have built a global alliance to tackle this challenge and bring this issue to the global agenda. We have also engaged with a community of experts composed of governments, civil society and academia to collect their insights through a one-month-long consultation.

This white paper presents a common set of proposed principles for the use of facial recognition by law enforcement investigations along with a self-assessment questionnaire developed to support law enforcement agencies in complying with these principles.

This is far from the end of the conversation on the use of facial recognition technology by law enforcement in criminal investigations, but we are confident that this first-ever proposed global approach can be an important contribution. Our alliance encourages governments and law enforcement agencies to reflect on this white paper, to participate in a dialogue on the basis of it, and review or adopt legislation that supports the responsible use of this technology.

Introduction

Over the past decade, progress in machine learning and sensors has fuelled the development of facial recognition technology (FRT) – a biometric technology capable of providing a score-based list of potential matches or verifying a person's identity by comparing and analysing patterns based on that person's facial features. This has led to its rapid adoption in various industries, including law enforcement, transportation, healthcare and banking.

The development of FRT presents considerable opportunities for socially beneficial uses, mostly through enhanced authentication and identification processes, but it also creates unique challenges. To fully grasp these challenges and the trade-offs they may entail and to build appropriate governance processes, it is necessary to approach FRT deployment through specific use cases. Indeed, passing through an airport border control with face identification, using face-based advertising in retail, or employing facial recognition solutions for law enforcement investigations involves very different benefits and risks.

To ensure the trustworthy and safe deployment of this technology across use cases, the World Economic Forum has spearheaded a global and multistakeholder policy initiative to design robust governance frameworks. The Forum launched the first workstream in April 2019, focusing on flow management applications[1] – replacing tickets with facial recognition to access physical premises or public transport, such as train platforms or airports. This workstream is now in the pilot stage with the release of a tested assessment questionnaire by Tokyo-Narita Airport, an audit framework and a certification scheme[2] co-designed with AFNOR Certification (Association française de normalisation).

In November 2020, the second workstream was started, focused on the law enforcement use case – *identifying a person by comparing a probe image to one or multiple reference databases to advance a police investigation*. While law enforcement has been using biometric data, such as fingerprints or DNA, to conduct investigations, facial recognition technology represents a new opportunity for law enforcement but also a new challenge.

This use case raises multiple public concerns because of the potentially devastating effects of system errors or misuses in this domain.

A study conducted in 2019 by the National Institute of Standards and Technology (NIST) showed that, although some facial recognition technologies had "undetectable" differences in terms of accuracy across racial groups, other facial recognition algorithms can exhibit performance deficiencies based on demographic characteristics such as gender and race.[3] Law enforcement agencies must be aware of these potential performance deficiencies and implement appropriate governance processes to mitigate them. In doing so, they would limit the risk of false recognitions and possible wrongful arrests of individuals identified by facial recognition systems.[4] Failure to build such processes could have dramatic consequences. In 2018 in the US, for example, an innocent African American man was arrested and held in custody as a result of being falsely recognized as a suspect in a theft investigation in which facial recognition technology was used.[5] In addition to hampering rights such as the presumption of innocence, the right to a fair trial and due process, the use of FRT by law enforcement agencies can also undermine freedom of expression, freedom of assembly and association, and the right to privacy.[6]

These concerns have led to global intensified policy activity. In the US alone, some local and state governments have banned the use of FRT by public agencies, including law enforcement. Major cities such as San Francisco, Oakland and Boston have adopted such measures. At the state level, Washington,[7] Virginia[8] and Massachusetts[9] have introduced legislation to regulate its use. Finally, at the federal level, various bills[10] have been proposed to regulate FRT but none of them has been adopted to this date.

Furthermore, large US technology companies have also formulated positions on this topic. Last year, IBM announced that it will no longer offer, develop or research FRT, while Microsoft pledged to stop selling FRT to law enforcement agencies in the US until federal regulation was introduced.[11] More recently, Amazon Web Services (AWS) has extended its moratorium on police use of its platform Rekognition, which it originally imposed last year.[12]

In other jurisdictions, policy-makers are attempting to limit police use of FRT to very specific use cases associated with robust accountability mechanisms to prevent potential wrongful arrests. That is the direction proposed by the European Commission (EC), which recently released its draft of an Artificial Intelligence Act[13] – a comprehensive regulatory proposal that classifies AI applications under four distinct categories of risks subject to specific requirements.[14] This proposal includes provisions on remote biometric systems, which include facial recognition technology. It states that *AI systems intended to be used for the "real-time" and "post" remote biometric identification of natural persons*

represent high-risk applications and would require an *ex-ante* conformity assessment of tech providers before getting access to the EU market and *an ex-post* conformity assessment while their systems are in operation. Moreover, *"real-time" remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement* are prohibited unless they serve very limited exceptions related to public safety (e.g. the prevention of imminent terrorist threats or a targeted search for missing persons). In order to enter into force, however, the EC's proposal will first need to be adopted by the EU parliament and the Council of the European Union.

At the United Nations, a similar approach is emerging, with the Office of the High Commissioner for Human Rights (OHCHR) recently presenting a report to the Human Rights Council on the right to privacy in the digital age, in which it recommends banning AI applications that cannot be used in compliance with international human rights law. With specific respect to the use of FRT by law enforcement, national security, criminal justice and border management, the report stated that "remote biometric recognition dramatically increases the ability of State authorities to systematically identify and track individuals in public spaces, undermining the ability of people to go about their lives unobserved and resulting in a direct negative effect on the exercise of the rights to freedom of expression, of peaceful assembly and of association, as well as freedom of movement". The report also reiterates calls for a moratorium on the use of remote biometric recognition in public spaces, at least until authorities can demonstrate that there are no significant issues with accuracy or discriminatory impacts, and that these AI systems comply with robust privacy and data protection standards.

Court decisions can also play an important role in shaping the policy agenda on FRT, as illustrated in Brazil. Recently, The São Paulo Court of Justice has blocked[15] the deployment of facial recognition in the public transport system. This is perceived as a major victory by civil rights organizations opposing the increasing use of FRT by public agencies. In a similar case in the UK, the Court of Appeal found that the deployment of automated facial recognition by the South Wales Police – at certain events and public locations where crime was considered likely to occur – to identify wanted persons was unlawful.[16]

In some countries, governments have adopted a cautious approach. That's the case in the Netherlands. In 2019, the Minister of Justice and Security addressed a letter to MPs informing them about the existing uses of FRT by law enforcement agencies and reaffirming his support for robust governance processes in relation to this sensitive technology.[17] Further, he argued that the existing legal framework and safeguards (technical and organizational) are sufficiently robust to ensure the responsible use of FRT by law enforcement agencies. Yet, he requested additional privacy, ethical and human rights impact assessments before authorizing any more pilots.

Despite these important developments, most governments around the world are still grappling with the challenge of regulating FRT. The ambition of this work is to support law- and policy-makers across the globe to design an actionable governance framework that addresses key policy considerations in terms of *the prevention of untargeted surveillance, the necessity of a specific purpose, the performance assessment of authorized solutions, the procurement processes for law enforcement agencies, the training of professional forensic examiners, and the maintenance of the chain of command for emergency situations*.

To achieve this goal, a multistakeholder community centred around the International Criminal Police Organization (INTERPOL), the United Nations Interregional Crime and Justice Research Institute (UNICRI) and the Netherlands Police has co-designed a set of principles for action that defines what constitutes the responsible use of facial recognition for law enforcement investigations and a self-assessment questionnaire that details the requirements that law enforcement agencies must respect to ensure compliance with the principles for action.

This governance framework was designed with the same ambition the World Economic Forum adopted for its first facial recognition workstream on flow management applications: namely, to inform the public debate on the use of facial recognition technologies at the national, regional and international levels and provide an actional framework to maximize the benefits of FRT while mitigating its risks. The governance framework remains to be tested, however, before its potential large-scale adoption. As such, it is highly recommended that law enforcement agencies, in partnership with civil society representatives, policy-makers and academics, engage with this initiative to strengthen and test the governance framework and ensure the greatest impact possible.

Methodology

For the past two years, the AI/ML platform of the World Economic Forum has been conducting a policy project on the governance of facial recognition. The objective of this policy project is to create an appropriate space for conversation to advance the drafting of policies related to the use of biometric technologies. The methodology consists of a pilot project co-led by a core community of partners and an extended global community of experts.

This pilot-based approach to policy-making has the potential to inform and guide policy-makers seeking to ensure the appropriate governance of FRT, and works on a longer time frame.

A multistakeholder approach based on a core community and a project community

The objective of this initiative was to draft a policy framework with a core community composed of INTERPOL and the Police of the Netherlands, both users of FRT for law enforcement investigations, and UNICRI, a United Nations entity mandated to assist intergovernmental, governmental and non-governmental organizations in their efforts to formulate and implement improved policies in the fields of crime prevention and justice administration. The core community gathered 24 times through virtual meetings to draft the policy framework.

This core community organized consultations with the project community, an extended group of stakeholders, to benefit from their expertise and insights. The project community was composed of 42 people: representatives of technology companies, governmental organizations and civil society, plus academics.

The first consultation with the project community was a workshop, organized on 16 February 2021 to kick off the project and gain key insights regarding the risks related to the use of FRT by law enforcement and the potential solutions to mitigate them.

The second consultation was a request for comments on the draft of the principles for the responsible use of FRT for law enforcement investigations. The project community was given a month to send in comments on the proposal. Following this, four expert interviews were organized to gather additional insights. In total, 10 organizations and experts from the project community sent comments on the proposal and, based on the received feedback, the core community met to address this and modified the draft principles accordingly.

The whole project was conducted under the Chatham House Rule, whereby participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed.[18]

A policy framework composed of a list of principles and a self-assessment questionnaire

This policy framework is composed of two elements: a list of principles and a self-assessment questionnaire.

The policy framework aims to define what constitutes the responsible use of facial recognition through the drafting of a set of principles for action. This list of principles was drafted by the core community composed of INTERPOL, UNICRI, the Netherlands Police and the World Economic Forum.

The self-assessment questionnaire supports practitioners in the law enforcement community to effectively verify their compliance with the list of principles. Law enforcement agencies are encouraged to evaluate their processes in place and assess their compliance with the requirements stated in the list of principles. To do so, law enforcement agencies can either conduct an internal review or outsource it to a third-party organization. Once completed, the results of the self-assessment questionnaire can be made public to increase transparency and accountability.

A pilot phase to test and iterate on the policy framework

The next stage of the project is the pilot phase conducted with law enforcement agencies, including the Netherlands Police. The policy framework presented in this white paper will evolve based on the results of tests conducted by law enforcement practitioners over this pilot phase. Law enforcement agencies interested in testing the framework should reach out to the core community.

Overall, law enforcement agencies are encouraged to use this policy framework as a tool for the adoption and improvement of best practices.

1 Law-enforcement investigations: use cases and definitions

An accurate and non-technical description of how FRT is used in practice by law enforcement agencies.

The use of automated facial recognition for law enforcement (LE) investigations brings new technology and can be potentially applied to many use cases. The presentation of the following use cases does not refer to any specific laws, policies, principles or recommendations that should limit or regulate their use. The sole purpose of this presentation is to provide a better understanding of how facial recognition technology (FRT) is or can be used by law enforcement agencies and to help illustrate the challenges that the governance framework seeks to address. The different examples presented in this chapter follow the practices of the Netherlands Police. These practices may vary across jurisdictions.

BOX 1 **The roles of the Police of the Netherlands and INTERPOL**

The Police of Netherlands and INTERPOL are entities with two distinct mandates. As a national law enforcement body, the Police of Netherlands has the mandate to conduct investigations and is required to testify and report the outcome of its expertise before a judge at court. INTERPOL's mandate, on the other hand, is to, inter alia, ensure and promote the widest possible mutual assistance between all criminal police authorities within the limits of the laws existing in the different countries and in the spirit of the Universal Declaration of Human Rights. To do so, INTERPOL manages 19 databases, all accessible to its 194 Member Countries. INTERPOL also provides recommendations on best practices, forensic expertise and other specialized expertise, produces analysis, delivers training activities and provides operational support to its Member Countries.

How facial recognition is used for law enforcement investigations

Law enforcement investigators use FRT for identification and authentication purposes. Identification activity (also referred to as "one to many") consists of searching for the identity of a person, as opposed to authentication activity (also referred as "one to one"), which consists of verifying someone's identity against an identity document (ID).[19] Facial examiners are experts who run facial recognition analysis. In the case of the Netherlands Police and INTERPOL, for example, the examiners operate autonomously from the investigation teams, and do not have knowledge of the prosecution that requires them to run facial recognition analysis.

To identify an unknown suspect or person of interest, investigators work with probe images and databases.

Probe image

A probe image is collected from an image source

The probe image is compared against a reference database

There are two typologies of databases:

Type 1

A reference database of known criminals and suspects, composed of mugshots lawfully collected and stored by law enforcement agencies. People in this database are still suspects or have usually been convicted of a crime.

Reference database of known criminals, suspects and missing persons

A reference database of known criminals, suspects and missing persons has been built over time by law enforcement

A probe image is compared against this reference database to check if this person is among known criminals, suspects and missing persons

Type 2

A special database built specifically for an investigation. The public prosecutor provides a warrant to seize the video footage of a crime scene. This database can be built out of multiple sources (CCTV, social media, electronic devices, etc.). All of the faces are detected on the footage and stored on the special database. The face of a possible suspect can then be searched against the special database to see if the suspect is present on the footage. At the end of the investigation, the database is removed from the operational system and stored so that the fact-finding/archiving/evidence file can be produced in court when requested during the judicial procedure.

Reference database built specifically for an investigation

A database of images from the investigation is created to build an investigation database of faces

An image of a known criminal, suspect or missing person can be searched against the investigation database

Probe images are images that are part of the law enforcement investigation and which are submitted to a facial recognition system to be compared to a database. Probe images are usually the photos or movies/stills of suspects or persons of interest. To collect these images, investigators (or digital/face experts) either already have an image of the suspect or they extract it from footage of movies/stills. In any case, law enforcement tries to collect the best-quality image to improve the chance of confirming the identity of the person.

Based on the practices followed by the Netherlands Police, the process for using FRT for law enforcement investigations is as follows:

Step #1: A (possible) crime is reported or suspected. An investigative team, under the supervision of the public prosecutor, is created and requests warrants to collect images relevant to the crime, including images of the suspect(s). If suspects are detected on the images, the team will try to determine their identity. This can be done by

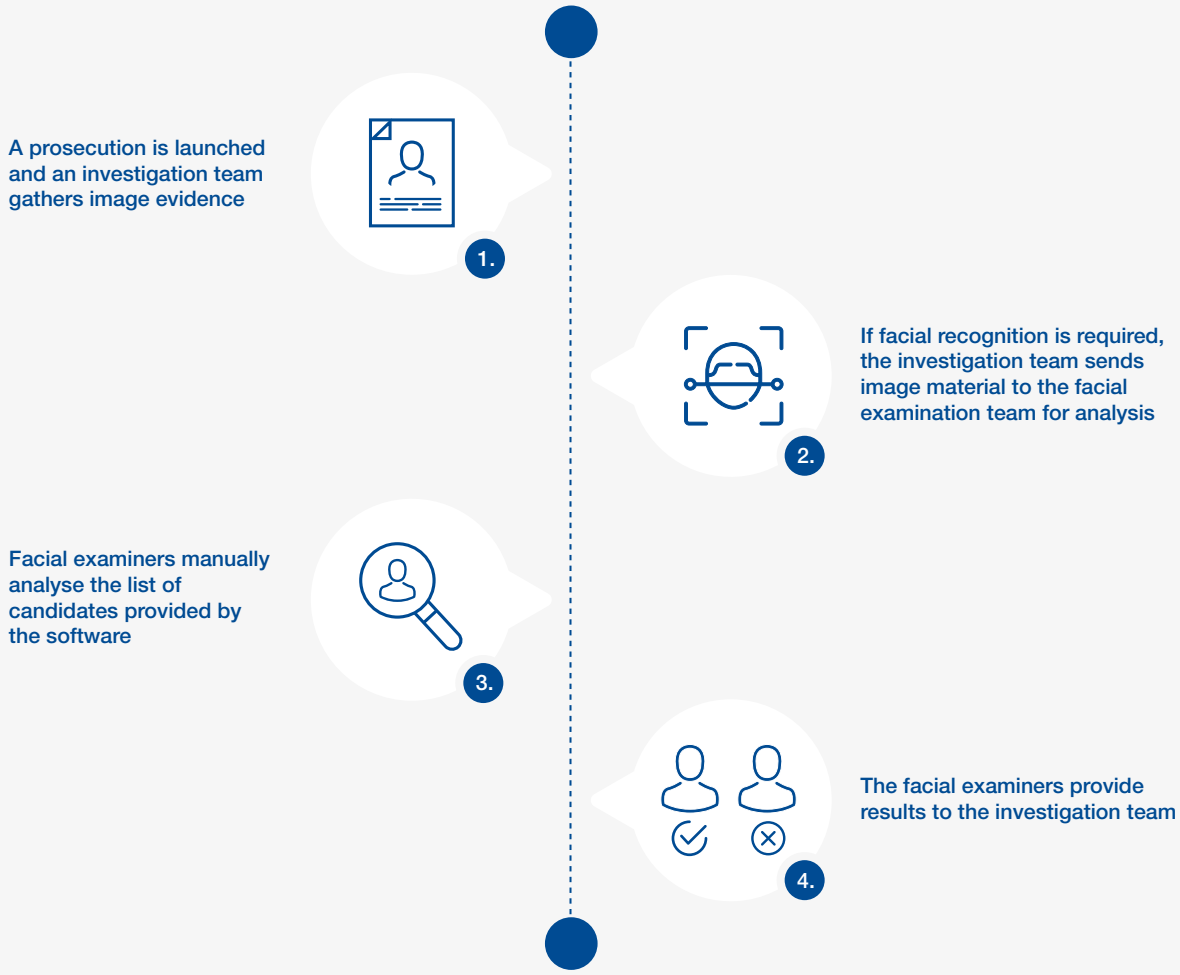
human means – through recognition by people who know the suspects, for instance, police officers or witnesses – or by using facial recognition software with a reference database of known people – for instance, suspects and convicts.

Step #2: If a facial recognition search is required, the investigation team will apply for an FRT investigation through the specialized FRT team. This facial examination team runs FRT software to compare the probe image against one or multiple databases. Before doing so, the facial examiners will first judge the quality of the probe image. If suitable for an FRT search, they will enter the probe into the FRT system, allow the system to do the pre-search analysis and may also provide some notable facial landmarks (centre of the eye socket, etc.) to the software. The examiners then set up the FRT software at a setting that is not too narrow, to avoid false negatives, or too wide, to avoid false positives – which would result in a list of candidates too large to be of use.

Step #3: After the search, the facial examiners analyse the list of candidates provided by the software. They run this last operation manually, deploying their expertise to check if one of the candidate images proposed by the system matches the probe image.

Step #4: If the facial examiners make a possible match, only the probe image and the image of the possible candidate from the reference database are handed to two facial experts. They perform, independently from each other, a full analysis of the probe and the reference image to determine the similarity/dissimilarity between the two faces. This blind peer review is systematically performed before any positive result is communicated to the requesting investigation team. The facial examiners and experts do not know the exact background to the case, to avoid bias as far as possible. The end result is the final consensus conclusion and is reported to the investigation team as an investigative lead.

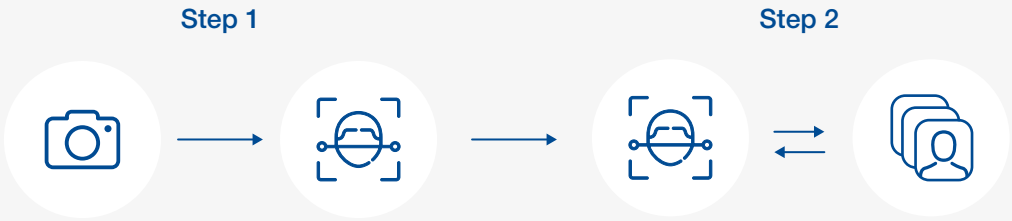
The four-step process followed by the Netherlands police when using facial recognition technology



The following is a collection of scenarios intended to illustrate how facial recognition technology can be used for law enforcement investigations:

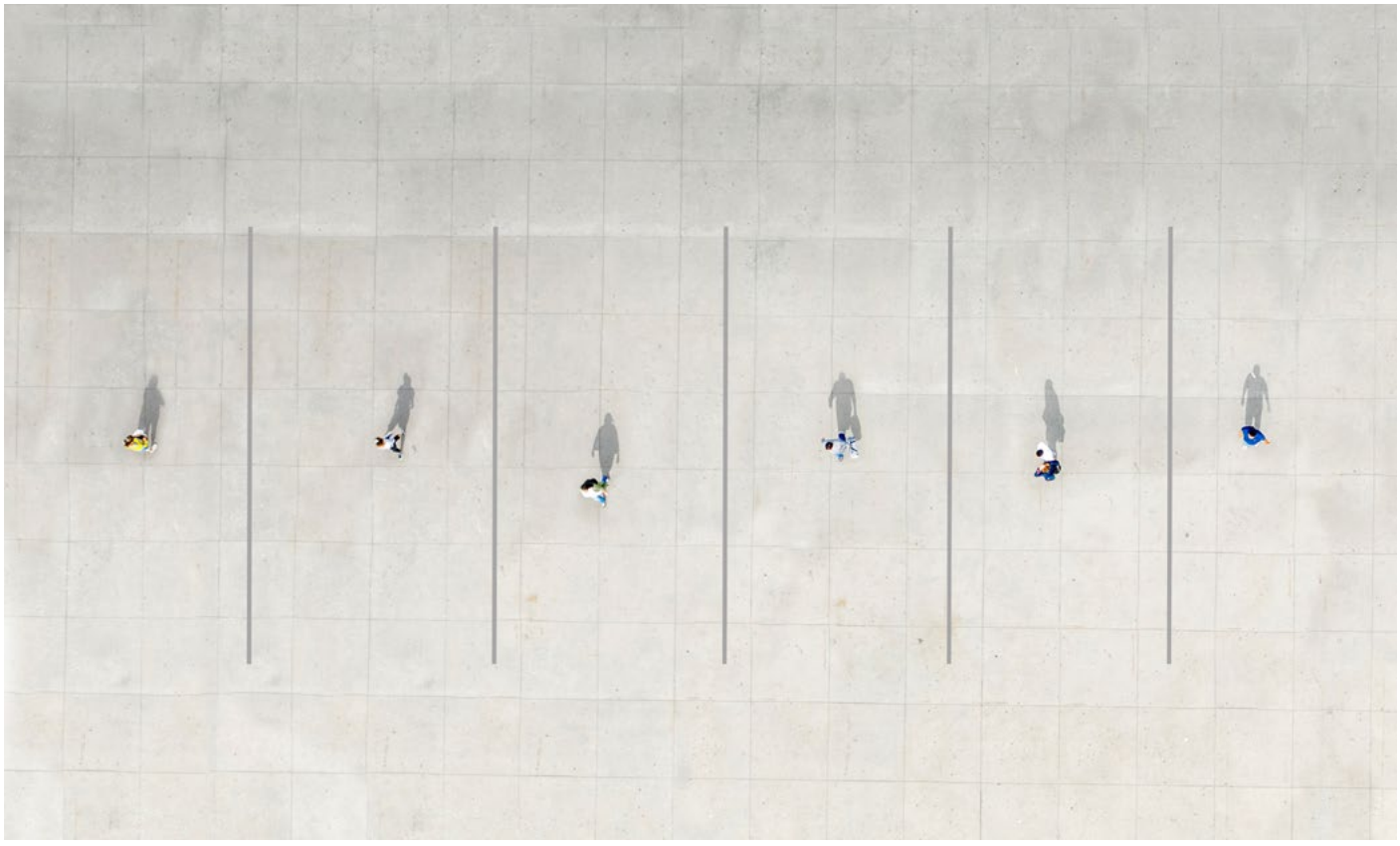
Identity checking at a border control

Border officers use identity controls to, *inter alia*, detect and possibly detain fugitives and wanted persons who are the subject of a valid INTERPOL Red Notice[20] and as such are recorded in the INTERPOL criminal database. Red Notices are published by INTERPOL at the request of member countries. The information published is also stored in the INTERPOL criminal database and made accessible to all member countries. While controlling the identity of people crossing a border, and upon assessment by the national border guard, border agents can make a request to INTERPOL to have the facial image of a controlled person compared to the facial images of criminal and missing persons stored in the INTERPOL Facial Recognition System (IFRS), the organization's facial image reference database. A photo of the controlled person is then converted into a probe image. In agreement with their national authorities, border officers send the probe image to their INTERPOL National Central Bureau (NCB) and to INTERPOL's headquarters for an urgent search in the IFRS. Then, INTERPOL facial examiners run the search in the IFRS. A list of potential candidate images is proposed by the system. Facial examiners analyse and manually compare the probe image with each candidate image and assess whether a lead emerges. If it is the case, a peer review is carried out by a second facial examiner and if the two agree on the positive conclusion, they subsequently inform the concerned INTERPOL National Central Bureau and border agents.



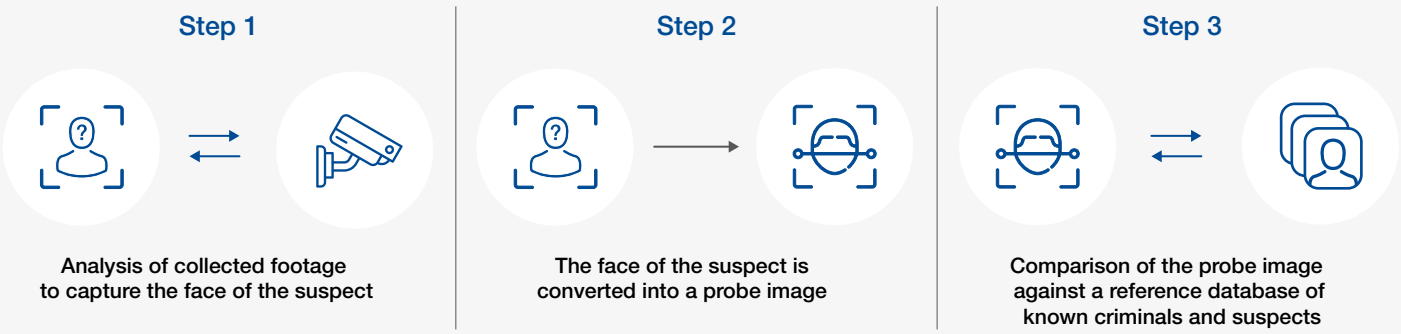
Collection of probe image

Comparison of the probe image against a reference database



Finding the identity of an ATM fraud criminal

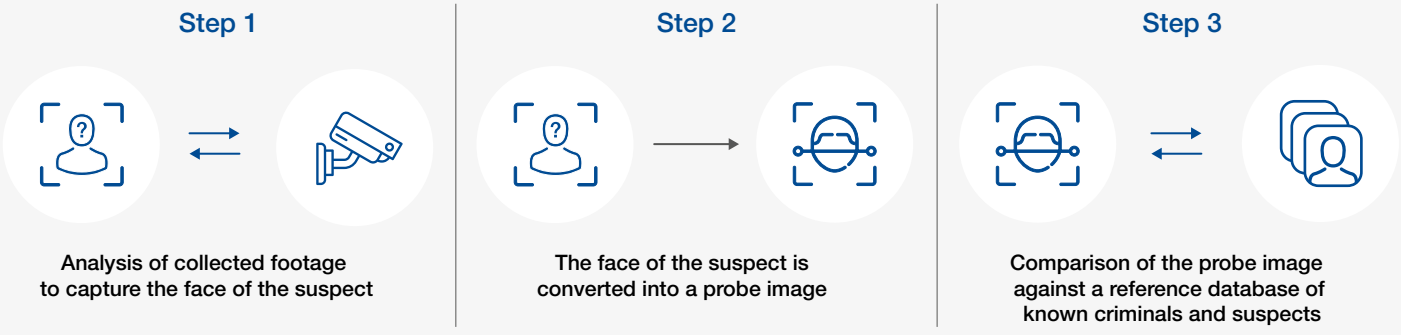
Fraudulently obtaining bank account data by usurping someone's identity allows a person to access a bank account and withdraw cash from an ATM machine. The video footage from the ATM machine enables investigators to collect a facial image of the offender. The quality of this image will vary depending on the exposure and whether the fraudster managed to hide his/her face. If the quality of the image is good enough, the photo collected will be compared against a database of known criminals using a facial recognition system. If facial examiners make a possible match, they follow the standard process described in Step #4 of the "How facial recognition is used for law enforcement investigations" section above.



The use cases in the next section highlight more specific examples of how facial recognition systems might be used.

Uncovering the identity of a rioter

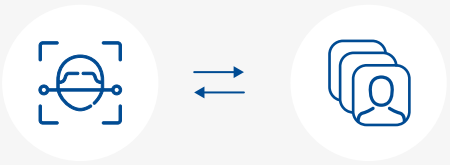
During a riot, a person attacks police officers and footage of the incident is collected from CCTV cameras. An investigation is launched and a warrant is provided to an investigation team to seize these images. The goal is to identify the assailant. To that end, the investigators, with the help of the police's digital experts, review the CCTV/video footage of the riot, looking for images of the wanted rioters. They collect images with the best angle, lighting and exposure possible to increase the quality of the image(s) and give the best chance of obtaining matches and identifying the rioters. If facial examiners make a possible match, they follow the standard process described in Step #4 of the "How facial recognition is used for law enforcement investigations" section above.



Looking for the identity of a museum thief

A piece of art has been stolen in a museum. A public prosecutor launches a criminal investigation. The investigation uncovers the identity of a potential thief and a warrant is given to collect video footage from the museum. Then, using a facial recognition tool, the investigators collect images of the faces of all visitors and staff who appear in the footage and build an investigation database from it. A list of candidate images is displayed by the system, reviewed and analysed to establish whether a serious potential match is detected that would confirm the involvement of the suspect. If facial examiners make a possible match, they follow the standard process described in Step #4 of the "How facial recognition is used for law enforcement investigations" section above.

Video footage from the museum is collected to build a database of faces that appear in it



The probe image of the suspect is compared against this reference database to check if this person shows up in the collected footage

Actively looking for a terrorist in public spaces

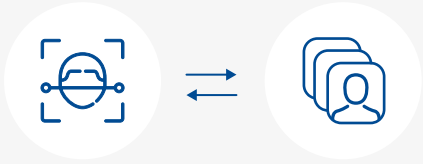
Note: the following example is a potential use case and has not yet been used by the Netherlands Police.

In the aftermath of a terrorist attack, where the terrorist remains at large, CCTV can be seized by law enforcement to collect a probe image of the fugitive terrorist. This probe image can then be distributed to all police patrols actively looking for the fugitive. In addition, the probe image can be compared in real time against other images of the suspect collected from separate CCTV footage or different image sources located in the terrorist's assumed vicinity. This real-time comparison may generate a potential lead that can be sent to police patrols, which can conduct a stop-and-check based on this alert.

Step 1

Collection of the probe image from video footage

Step 2

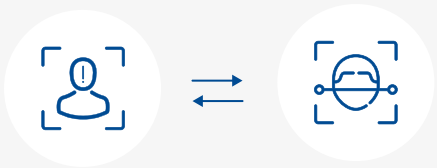


Comparison of the probe image with other images collected from separate CCTV footage (located in places where the fugitive is likely to be or where the fugitive was temporarily present)

Step 3

The video comparison leads to a potential match

Step 4



A stop and check by law enforcement patrols is done on the potential match

The following use cases provide examples of how facial recognition is currently used to conduct investigations in child abuse and missing persons cases. Considering the sensitivity of these investigations, the cases below present the different instances in which facial recognition can be used rather than providing a detailed presentation.

Using facial recognition to fight child abuse

National law enforcement agencies and INTERPOL use facial recognition technology to investigate cases of child abuse. To dismantle international child abuse networks, INTERPOL runs investigations in partnership with national law enforcement agencies. Dedicated task forces within INTERPOL and national police departments collect images and pieces of evidence to facilitate the resolution of investigations.

Images and videos showing victims of child abuse are stored in dedicated databases with highly restricted access. These databases are very often developed using a range of tools and features to support the work of investigators, help them to analyse the images and find new leads. Facial recognition can be used to identify the victims; their facial images can be searched in a database containing the facial images of missing persons. However, missing minors are not necessarily recorded in these facial databases because the face undergoes many changes during childhood and adolescence. In most cases, the police rely on other means to identify victims. Facial recognition can also be used to confirm that the same child appears in various image sources and estimate the period during which the victim has been abused. The primary goal of all of these findings is to identify, locate and rescue the victim as soon as possible.

Facial images of perpetrators, when collected and seized, can be searched in national criminal databases and in the INTERPOL criminal database in order to identify, locate and detain them with a view to prosecution. It is crucial for investigators to collect as much evidence as possible to document and strengthen the prosecution case, using all existing investigative tools, including facial recognition when relevant.

Using facial recognition to find missing persons

Where there is serious evidence suggesting urgency in a missing persons case, national law enforcement agencies can ask INTERPOL to create a Yellow Notice for that missing person. This file usually also includes other biometric attributes such as fingerprints and DNA.

If that missing person is eventually identified during a border control check, for example, the person has the choice, as long as he/she is an adult and has not committed any crime, to ask law enforcement not to inform his/her family.

Further, the Yellow Notice database can be beneficial when a person is declared missing in a given country and found dead in another one. In this case, the Yellow Notice will help identity the deceased person.

For missing children cases, most of the time, and for privacy reasons, there is no database of minors. Therefore, the only way to identify missing children using facial recognition is by consulting investigation databases of child abuse cases and comparing images.

2 Proposed principles

The first publicly shared principles for the responsible use of facial recognition technology for law enforcement investigations co-designed by a global community.

1 Respect for human and fundamental rights

Note: this proposition of principles focuses on law enforcement investigation activities. Law enforcement activities related to passport, residence permit and ID card issuance/ verification are not included in these principles. As such, facial recognition used to verify the identity of applicants – to make sure that the photo provided in the application matches the applicant and prevent fraud – is outside of the scope of this work.

1.1. Facial recognition technology (FRT) should be used only as part of a lawful criminal investigation such as to identify criminals/fugitives, missing persons, persons of interest and victims.

1.2. The rights provided for within the International Bill of Rights and other relevant human rights treaties and laws should always be respected, particularly the right to human dignity, the right to equality and non-discrimination, freedom of expression, association and the right of peaceful assembly, the rights of the child and older persons, the rights of persons with disabilities, the rights of migrants, the rights of Indigenous people and minorities, and the rights of persons subjected to detention or imprisonment. The use of FRT by law enforcement for investigations should respect these rights and be necessary and proportionate to achieve legitimate policing aims.

1.3. Law enforcement agencies should be subject to effective oversight by bodies with effective enforcement powers and in accordance with national laws or policies. Among other things, these or other bodies should have the specific task of hearing and following complaints from citizens and assessing the compliance of law enforcement activities with human and fundamental rights.

1.4. Any individual should have the right to an effective remedy before an independent and impartial tribunal set up by law against actions concerning the use of FRT.

2 Necessary and proportional use

2.1. The decision to use facial recognition technology should always be guided by the objective of striking a fair balance between allowing law enforcement agencies to deploy the latest technologies, which are demonstrated to be accurate and safe, to safeguard individuals and society against security threats, and the necessity to protect the human rights of individuals. As a general principle, FRT should never be used without cause and need that otherwise would undermine human and fundamental rights.

2.2. Law enforcement agencies considering the use of facial recognition technology should always provide a documented and justified argument as to why FRT is the chosen option and why other less intrusive options are not a chosen solution.

2.3. The use of FRT by law enforcement agencies should always be aimed at, and limited to, a single specific goal, necessarily related to investigative purposes.

2.4. International, regional and national policies and/or laws should specify for which classes of crimes or investigations the use of FRT by law enforcement agencies is acceptable and/or lawful.

2.5. Acknowledging the right to privacy and other human rights, the collection of images from public and publicly accessible spaces for FRT identification purposes should be done only for a determined list of use cases, in a limited area and for an established processing time period in accordance with relevant national laws or policies.

2.6. The use of facial recognition technology for identification purposes (with the exception of situations of emergency presented in 2.7) should always be conducted by an individual trained as described in 7.1 and the process independently reviewed/performed by a second blind analysis procedure undertaken by another officer. The outcome should be the conclusions of the two analyses, and the most conservative one reported to the investigation team as the final result.

2.7. The use of "real-time" facial recognition technology for identification purposes for the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or of a terrorist attack represents the most sensitive use case. The imperative to act fast can, exceptionally, necessitate using FRT systems without the outcome undergoing expert verification. The system would automatically provide a proposed match based on live CCTV footage from public areas of interest. Law enforcement patrols should use this proposed candidate only to verify this individual's identity and conduct additional verifications, if necessary. As a consequence, acknowledging the risks involved in this exceptional emergency situation, an independent authority should be in charge of authorizing this application and, if there is not enough time, it should be authorized by the chain of command. In this case, the chain of command should inform and justify the decision to the independent authority as soon as possible and not later than 24 hours. All processed images should be permanently deleted from the FRT system unless they have led to a match.

2.8. FRT, and other face analysis technologies, should be used for no purpose other than biometric identification/recognition/verification. The use of FRT to infer ethnicity, gender, sex, age, emotion, opinion, health status, religion and sexual orientation, and the use of FRT for predictive analysis, should not be permitted.

3 Transparency

3.1. Law enforcement agencies should make public:

3.1.1. The vendor selected (if applicable), the name and version of the software, and disclosure of all software developers (including those for the core FRT, image handling, post processing, GUI displays and systems integration).

3.1.2. A clear definition of the use of FRT for law enforcement investigations, specifying the purpose and objectives such as identifying criminals/fugitives, persons of interest, missing persons and victims.

3.1.3. The use of probe images: procedures and criteria to select, store/not store images, and if stored, for how long.

3.1.4. The use of reference database: procedures to consult the database, and criteria to select, store/not store probe images in this reference database, and if stored, for how long; as well as procedures for determining whether any machine learning can be conducted on that data, including training, learning and model refinement.

3.1.5. The type of data-sharing with other organizations.

3.1.6. The name of departments able to launch searches and view results of searches.

3.1.7. Information about human oversight and accountability (see 4.1 to 4.5) and the mechanisms in place to ensure FRT is used as intended.

3.1.8. The seniority threshold of law enforcement officials who have access to FRT and the chain of command for the use of FRT.

3.1.9. Results of evaluations of the effectiveness of the FRT conducted by the vendor of the technology (for each evaluation, a description of (1) the design of the evaluation, (2) the data used in the evaluation, and (3) the results (metrics) obtained).

3.1.10. Results of evaluations of the effectiveness of the FRT conducted by the law enforcement agency (for each evaluation, a description of (1) the design of the evaluation, (2) the data used in the evaluation, and (3) the results (metrics) obtained).

3.1.11. A record of complaints filed by members of the public against the use of the FRT and the law enforcement agency's response of those formal complaints.

3.1.12. Auditable records of search requests made by law enforcement.

3.1.13. Any other information necessary for the public to ensure law enforcement's compliance with the relevant obligations, including how an individual could contact law enforcement to submit a query or complaint.

3.2. Law enforcement agencies should provide information to the public regarding the use of FRT. Information provided to the public should be concise, easily accessible, understandable and provided in a clear and plain language. Exceptions to this should be permitted only if they are necessary and proportionate to pursue legitimate purposes and in accordance with the law.

4 Human oversight and accountability

4.1. Lines of responsibility for the outcome of a given use of FRT should be well defined and transparent. A law enforcement agency should never issue analysis and conclusions from FRT without interpretation by an examiner and oversight by a manager with the right expertise (with the unique exception described in 2.7).

4.2. The skills of facial examiners are critical and necessary to maintain the highest level of accuracy in the identification process.

4.3. A peer review (blind verification or examination by a second expert) should systematically be performed before any positive result communicated to the requesting investigation team. The provided end result should always be consensus-based, and the most conservative conclusion of the two should prevail.

4.4. The law enforcement agency should ensure a mechanism exists whereby citizens can file a complaint with or seek redress from an oversight body as designated by national policies.

4.5. For anyone identified using an FRT system, that person must be informed that he/she was subject to such a search/that an FRT system was used to identify them, if they are subsequently taken into custody, brought in as a witness, or have any other official role in a law enforcement process based on their face via the FRT system.

5 System performance

5.1. Organizations providing facial recognition technology should follow standards for evaluating the accuracy and performance of their algorithms at the design (lab test) and deployment (if and when possible, field test) stages.

5.2. Law enforcement agencies should require vendors to submit their algorithms to large-scale independent testing undertaken against appropriate test standards (lab tests and, if possible, field tests) and select providers who can demonstrate the efficiency of the algorithm follows standards of performance.

5.3. Due diligence with respect to system performance should be undertaken by reference to large-scale independent tests, such as those conducted by NIST in the USA. These tests provide a scientifically robust, transparent baseline of performance.

5.4. Validations of the performance of the FRT shall be designed to model, as closely as practical, the real-world objectives and conditions (including, e.g. data landscape, operators of the technology, timetables affecting decisions made using the technology) in which the FRT would in practice be applied.

5.5. To leverage accuracy gains, law enforcement agencies should expect to make, and establish procedures for, regular upgrades or replacement of the FRT.

6 Risk-mitigation strategies

6.1. The risk of error and bias by machines and humans should be mitigated to the greatest extent possible. This should be done through an *ex-ante* and *ex post* evaluation strategy:

6.1.1. *Ex ante* evaluations: technology providers, and when it applies, technology integrators, should ensure biases and errors are mitigated to the greatest extent before the deployment of the system by law enforcement agencies. The level of performance, and the design of the quality management system (which includes the quality of the risk management processes) should be evaluated by an independent third party. This evaluation should be organized by the technology provider, and when it applies, the technology integrator, and the results made available to law enforcement agencies that procure FRT and to the public for review. Law enforcement agencies that procure FRT should add to their procurement criteria the specific metrics the provider uses to gauge bias (as well as other relevant risks), the results of any evaluations conducted to estimate the performance of the provider's FRT on those metrics and the results of any evaluations of the performance of the system. Before deploying FRT systems, law enforcement agencies should set up pilot tests to ensure the system is operating as intended.

6.1.2. *Ex post* evaluations: law enforcement agencies – if needed, with the support of competent authorities – should deploy risk mitigation processes to identify, monitor and mitigate the risks of error and biases throughout the entire life cycle of the system. A regularly programmed internal audit (that would include the use of the self-assessment questionnaire related to these principles), and if possible, an independent third-party audit, should be conducted to validate the robustness of these processes. The conclusions of these audits should be made publicly available.

To continually improve the quality of the processes and the system's performance, law enforcement agencies, technology integrators and technology providers should set up an internal control or, where the services of a technology provider are procured by law enforcement, establish an in-service support agreement throughout the entire life cycle of the system.

7 Training of facial examiners

7.1. FRT should be used only by trained officers who follow the procedures ordered by the chain of command/management. Everybody within the organization, especially the chain of command/management, should understand the capacities and limits of the system used. The training (and certification when it applies) of examiners, and the chain of command/management, should include:

7.1.1. Knowledge of and updates of possible mandatory regulations, laws or policies concerning the use of biometrics.

7.1.2. Awareness of the risk of biases with the FRT system (anticipate false positives and false negatives, awareness of difference of performance on various demographics, know how to calibrate and adjust the threshold of the system, understand how to configure the system in the manner appropriate to the specific circumstances and risks of a given use case, and how to fix the length of the candidate lists).

7.1.3. Understanding of the risk of false negative and false positive errors (overestimation of own capability, risk of over-reliance on technology, blind spots, risk of human bias such as other-race-effect bias).

7.1.4. Awareness of the risk of image manipulation, including data integrity attacks and data morphs, and the tools to identify them.

7.1.5. Collection, storage, integrity and traceability of data processes.

7.1.6. How to implement risk mitigation methodologies (one match vs. differential diagnosis approach, blinding techniques, blind verifications, etc.).

7.1.7. Human-machine interaction best practices.

7.1.8. Ethical awareness: identifying the presence of vulnerable data subjects and/or areas potentially attended by vulnerable data subjects (e.g. schools, playgrounds, hospitals, places of worship, etc.).

7.1.9. How to use tools that assist examiners in understanding the reasoning behind systems' decisions/recommendations.

7.1.10. Awareness of the risk of false positives from twins, siblings and other related individuals.

7.2. Law enforcement agencies that use or intend to use FRT should offer training on an ongoing basis and should be informed by the latest research in machine learning, human-machine interaction and remote biometrics.

7.3. Recognizing that innate capability to recognize faces exists on a spectrum, examiners should be recruited by factoring in performance on face comparison tests, acknowledging that experience and training also matter.

8 Use of probe images and reference databases

8.1. Law enforcement agencies must ensure that their processing of probe images and reference databases are compliant with international, regional and national laws and/or policies, which should include purpose limitation, storage criteria, retention period, deletion rules, etc.

8.2. The collection of probe images should be conducted on a legal basis and aimed at a specific purpose.

8.3. The reference database(s) used for FRT investigations should always have a legal basis and be used under the authorization of competent authorities. Consequently, reference databases that include data collected without legal basis from the internet, or electronic devices, should not be used.

8.4. Probe images should not be adopted as reference photos and should not be inserted into reference databases unless they have led to a verified match.

8.5. Exporting images and biometric metadata to public cloud-based FRT that could potentially be outside the local jurisdiction should be prohibited.

8.6. Law enforcement agencies shall maintain a strict and transparent chain of custody of all images (probe image sets and reference databases) used for FRT. The law enforcement agency shall specify, and enforce, clear and transparent rules designating who does and does not have access to the images and in what circumstances.

8.7. Law enforcement agencies shall specify well-defined protocols for determining when, and on the basis of what criteria, images are to be expunged from a probe set or a reference database. The law enforcement agency shall create, and adhere to, a well-defined and transparent protocol for the disposal of images that have been expunged from a probe set or reference database or are otherwise no longer needed; any such protocol shall be designed to protect the privacy of any individuals appearing in the images identified for disposal.

9 Image and metadata integrity

9.1. To mitigate the risk of errors, law enforcement agencies should follow the recommendations of standards and thresholds of photo quality collected for law enforcement investigations. Before using any FRT system, law enforcement agencies should have a procedure to perform image quality assessment and a minimum quality threshold. The FRT system should not use probe or reference database images that do not meet the defined threshold.

9.2. Law enforcement examiners should be aware of the risk of image manipulations, such as morphing and deepfakes, when images come from uncontrolled sources and/or production modes. When detected, these images should be rejected or processed with extreme precaution.

9.3. Only forensic upgrading of face quality should be accepted for final examination. Forensic upgrading should be documented so as to ensure the auditability and reproducibility of the upgrading process. The creation of new content and the insertion or modification of facial features or geometry on an existing image should be forbidden.

9.4. While processing data, law enforcement agencies should always conduct a proper and verified attribution of identity to photos in the dataset, and verify the serial number of photos, their traceability and origin.

9.5. Vulnerabilities to hacking and cyberattacks should be identified to ensure robustness and avoid data leaks, and data manipulation.

9.6. An audit of the integrity of the reference database should be conducted regularly, in accordance with the applicable legal framework and best practices.

3 Proposed self-assessment questionnaire

A self-assessment framework to ensure that law enforcement agencies have introduced the right risk-mitigation processes.

1 Respect for human and fundamental rights

- *What are the procedures in place to ensure that FRT is used only in lawful criminal investigations?*
- *Are you working with effective oversight bodies to:*
  - *Ensure that your use of FRT complies with human and fundamental rights while being proportionate to achieve legitimate policing aims?*
  - *Address complaints from citizens?*
- *Is there an existing judicial authority to offer effective remedies to individuals who have been abused by law enforcement use of FRT?*

2 Necessary and proportional use

- *What procedures in place are preventing you from using FRT for no cause and need?*
- *What are the alternatives to your facial recognition system? And why have you rejected them? What are the criteria used to determine the advantages and disadvantages of these alternatives?*
- *How do you ensure that your use of FRT is appropriate, limited and exclusively related to investigative purposes?*
- *What uses of FRT are allowed in your jurisdiction (based on laws defined by international, regional and national laws or policies)?*
- *What are the use cases for which you are authorized to collect images from public spaces for FRT identification?*
- *What are the processes to record in a specific area using FRT and the period of time for which it has been approved?*
- *Have you deployed a procedure to ensure that only examiners are in charge of conducting face analysis?*
- *Have you deployed a procedure to ensure that peer reviews are systematically conducted?*
- *What procedures are in place to work with independent authorities in charge of authorizing "real-time" uses of facial recognition technology for identification purposes under exceptional circumstances?*

- *If "real-time" use of FRT is authorized by the chain of command because of a lack of time to inform the independent authority, what processes have you introduced to ensure that the chain of command informs and justifies its decision to the independent authority within 24 hours?*
- *What processes have you implemented to make sure all processed images are deleted unless they have led to a match?*
- *What processes have you implemented to prevent the use of FRT to infer ethnicity, gender, sex, health status, age, emotion, opinion, religion or sexual orientation recognition or for predictive analysis?*

3 Transparency

- *Have you publicly shared information about:*
  - *The purpose of the FRT solution deployed, the selected vendor, the name and disclosure of the software developers?*
  - *A clear definition of its use and the various facial recognition use cases?*
  - *Your processes regarding the use of probe images?*
  - *Your processes regarding the use of reference databases?*
  - *Your data-sharing policy with other organizations?*
  - *The list of departments that have access to FRT search requests?*
  - *Human oversight and accountability (see 4.1 to 4.5) and the mechanisms in place to ensure FRT is used as intended?*
  - *The seniority threshold of law enforcement officials who have access to FRT and the chain of command?*
  - *The results of evaluations of the effectiveness of the FRT conducted by the vendor of the technology?*
  - *The results of evaluations of the effectiveness of the FRT conducted by the law enforcement agency?*
  - *A report presenting the response of law enforcement agencies to citizens' complaints about their use of FRT?*
  - *Auditable records of search requests?*
  - *Which external communication channels and processes are in place for individuals to submit a query or complaint?*
- *How do you ensure that the information provided to the public about law enforcement use of FRT is concise, easily accessible, understandable and provided in a clear and plain language?*
- *How do you ensure that exceptions are justified only by the pursuit of legitimate purposes and in accordance with the law?*

4 Human oversight and accountability

- *What processes have you introduced to ensure that FRT is always used in partnership with an examiner and to prevent automated analysis and conclusions (with the unique exception described in 2.7)?*
- *How do you ensure that examiners have the capacity and knowledge required to interpret the outcome of a machine and make a final decision based on its training and expertise?*
- *How do you ensure that:*
  - *A systematic peer review is performed before reaching any final decision?*
  - *The provided end result is consensus-based, and the most conservative conclusion of the two experts always prevails?*
- *What mechanisms have you implemented to:*
  - *Enable citizens to file a complaint with or seek redress from an oversight body?*
  - *Inform any individuals taken into custody, brought in as a witness or involved in an investigation that they were identified using an FRT system?*

5 System performance

- *For lab and, when possible, field tests, what existing or forthcoming standards (e.g. International Organization for Standardization [ISO], and European Committee for Standardization [CEN]) are you planning to follow to evaluate the performance of your systems?*
- *Have you required your vendor to submit its FRT system to an independent evaluation?*
- *Have you selected a vendor able to demonstrate that the efficiency of its algorithm follows standards of performance?*
- *Has your vendor done a due diligence of the performance of its system such as the one organized by the NIST?*
- *Have you required vendors to follow specific standards of performance and introduced procurement rules to select providers who comply with these standards?*
- *What processes have you introduced to ensure validation of performance is:*
  - *Conducted through a comprehensive quality control protocol?*
  - *Designed to model, as closely as possible, the real-world objectives and conditions in which the FRT would in practice be applied?*
- *What procurement rules have you introduced to ensure the regular upgrading or replacement of your FRT?*

6 Risk-mitigation strategies

- *How is your technology provider (or when it applies, the integrator) making sure biases and errors are mitigated to the greatest extent possible before its deployment?*
- *Have tech providers or integrators been audited by a third-party organization on the level of performance and the design of the quality management system of their FRT systems?*
- *Have tech providers and integrators communicated the results of those evaluations to law enforcement agencies and the general public?*
- *Have you specified in your procurement criteria the metrics that tech providers must use to gauge bias and other relevant risks, as well as the results of any evaluations conducted to assess the performance of the provider's FRT systems?*
- *Have you run pilot tests before deploying FRT systems?*
- *Have you deployed risk mitigation processes to identify, monitor and mitigate the risks of error and biases throughout the entire life cycle of the system?*
- *Have you programmed internal audits and, if possible, an independent third-party audit, to validate the robustness of your risk mitigation processes?*
- *Have you publicly shared the results of these audits?*
- *Have you implemented internal control or in-service support agreement throughout the entire life cycle of the system in collaboration with technology providers and integrators?*

7 Training of facial examiners

- *What processes have you implemented to make sure FRT is used only by trained officers and that they follow the procedures ordered by their chain of command/management?*
- *Have you ensured that the training (and certification when it applies) of examiners and agents within the chain of command/management include up-to-date training programmes about:*
  - *Mandatory regulations, laws or policies concerning the use of biometrics?*
  - *Risk of machine biases related to FRT systems?*
  - *Risk of human biases when using FRT systems?*
  - *Risk of image manipulation, including data integrity attacks and data morphs, and training on existing or new tools used to detect them?*
  - *Data governance risks throughout the FRT system life cycle, including collection, storage, integrity and traceability of data?*
  - *Implementation of risk mitigation methodologies?*
  - *Human-machine interaction best practices?*
  - *Ethical training to identify the presence of vulnerable data subjects and/or areas potentially attended by vulnerable data subjects?*
  - *How to appropriately use tools that assist examiners in understanding the reasoning behind systems' decisions/recommendations?*
  - *Risk of false positives from twins, siblings and other related individuals?*
- *How often is this training programme offered?*

- *How do you evaluate the quality of the training programme over time, taking into consideration the latest progress in research (for example: have you established a scientific committee or equivalent, etc.)?*
- *Have you implemented recruitment processes to primarily hire examiners who perform well on standardized face comparison tests?*

8 Use of probe images and reference databases

- *What procedures have you been following to ensure that your processing of probe images comply with international, regional and national laws or policies?*
- *What processes have you introduced to ensure that the collection of probe images is conducted on a legal basis and aimed at a specific purpose?*
- *How do you manage your reference databases to ensure that:*
  - *Consultation is authorized by a competent authority?*
  - *All images are lawfully collected?*
  - *Probe images are not used as reference photos nor inserted into the database unless they have led to a verified match?*
- *What processes have you implemented to prevent the export of images and biometric metadata to public cloud-based FRT that could potentially be outside the local jurisdiction?*
- *How do you ensure a strict and transparent chain of custody of all images (probe image sets and reference databases)?*
- *Have you established clear protocols for determining when, and based on what criteria, images are to be expunged from a probe set or a reference database?*
- *Do these protocols effectively protect the privacy of any individuals appearing in the images identified for disposal?*

9 Image and metadata integrity

- *Have you been following best practices and recommendations, such as the one from the FISWG?*[21]
- *What photo quality standards are you following? What quality thresholds are you applying?*
- *How do you manage the risks of deepfakes and morphing? Do you deploy a specific procedure to detect them when you collect images from uncontrolled sources?*
- *If you detect a modified content (deepfake, morphing, etc.), how do you process this image?*
- *How do you prevent modifications of images except the forensic upgrading of face quality?*
- *How do you document forensic upgrading to ensure the auditability and reproducibility of the upgrading process?*
- *How do you prevent the creation of new content on existing images from occurring?*
- *What processes do you follow to ensure the proper attribution of identity to photos in the dataset and to verify the serial number of photos, their traceability and origin?*
- *What processes are in place to identify vulnerabilities to hacking and cyberattacks?*
- *Have you been conducting regular audits of the integrity of the reference database?*

Conclusion

The rapid deployment of facial recognition technology for law enforcement investigations around the world is arguably among the most sensitive use cases because of the potentially disastrous effects of system errors or misuses in this domain. Therefore, there is a pressing need to design and implement a robust governance framework to mitigate these risks.

The set of principles for action and assessment questionnaire contained in this white paper could inform a governance response. Indeed, providing law enforcement agencies with a clear definition of what constitutes the responsible use of facial recognition technology and a tool to assess their practices to ensure compliance is an agile and practical means to build accountability.

The project is now entering the pilot phase. During this period, we will test the governance framework to ensure its actionability, relevance, usability and completeness in collaboration with the Netherlands Police and review it based on the observed results. We encourage other law enforcement agencies to participate in the testing process. Once this pilot phase is completed, we will update the principles and the self-assessment questionnaire, and a second version will be published.

Policy-makers, industry players, civil society representatives and academics engaged in the global policy debate about the governance of facial recognition technology are encouraged to join this initiative to test and adopt this governance framework and encourage its deployment.

Glossary

Accuracy of facial recognition: The accuracy of a facial recognition system is based on the number of correct predictions, which consists of a combination of two so-called "true" conditions:

1. True positives: outcome when the system correctly finds a match for a person whose picture is included in the reference database.

2. True negatives: outcome when the system correctly finds no match for a person whose picture is not included in the reference database.

Accuracy is defined as the percentage of correct predictions, i.e. it is calculated by dividing the number of correct predictions by the number of total predictions.

Algorithm: A series of instructions to perform a calculation or solve a problem, implementable by a computer. Algorithms form the basis for everything a computer can do and are, therefore, a fundamental aspect of all AI systems.

Audit: Verification activity, such as inspection or examination of a process or quality system, to ensure compliance with requirements.

Bias in facial recognition technology: False positives and false negatives rate variations caused by demographic dependencies across groups defined by sex, age and race or country of birth. This lack of accuracy is usually caused by the training dataset of the algorithm, which does not contain enough or accurate representations of the demographics.

Biometrics: A variety of technologies in which unique identifiable attributes of people, including (but not limited to) a person's fingerprint, iris print, handprint, face template, voice print, gait or signature, are used for identification and authentication.

Computer vision: A field of computer science that works on enabling computers to see, identify and process images in a way similar to how humans perform these actions, and then provide appropriate output.

Explainability: A property of AI systems that provides a form of explanation for how outputs are reached. Explainability is important to improve decision understanding and increase the trust of operators and users of the systems.

Face detection: The process of finding human faces by answering the question "Are there one or more human faces in this image?". Face detection differs from face identification/verification as it does not involve biometrics analysis.

Face identification (or one-to-many): The process of answering the question "Can this unknown person be matched to an image in a reference database?". Identification compares a probe image to all of the images stored in a reference database, so it is also called "one-to-many" matching. A list of candidate matches is returned based on how closely the probe image matches each of the images from the reference database.

Face verification (or one-to-one): The process of answering the question "Are these two images the same person?". In security or access scenarios, verification relies on the existence of a primary identifier (such as an ID), and facial recognition is used as a second factor to verify the person's identity.

Facial recognition system: A biometric software application capable of uniquely identifying or verifying a person by comparing and analysing patterns based on the person's facial features

False negative: A test result that incorrectly indicates that the person in the probe image is not enrolled in the reference database when in fact this is not the case.

False positive: A test result that incorrectly indicates that the person in the probe photo is enrolled in the reference database when this is not the case.

Facial assessor/reviewer/examiner: Three distinct categories of officers in charge of conducting a face analysis:

- **Facial assessor:** Performs a quick comparison of image-to-image or image-to-person, typically with controlled images, carried out in screening and access control applications or field operations. Due to limitations such as time constraints, assessors perform the least rigorous of all facial comparison processes. For example, a person at a port of entry or in the field using a mobile FRT system to assist with an identity verification.

- **Facial reviewer:** Performs a comparison of image(s)-to-image(s) generally resulting from the adjudication of a candidate list generated by an FRT. The comparison results are often used in either investigative and operational leads or intelligence-gathering applications.

- **Facial examiner:** performs a comparison of image(s)-to-image(s) using a rigorous morphological analysis, comparison and evaluation of images for the purpose of effecting a conclusion, often used in a forensic application.

Forensic upgrading of face quality: Enhancement of the quality of an image. The creation of new content, insertion or modification of facial features or geometry on an existing image are not considered an upgrading, and thus forbidden in the process of FRT for law enforcement.

Law enforcement agency: Any government agency responsible for the enforcement of the law.

Peer review face analysis: A peer review process based on blind verification, or second opinions, that validates the conclusions of any initial human analysis.

Probe image: The image submitted to a facial recognition technology system to be compared to images on the reference database.

"Real-time" and "post" facial recognition:[22]

- **"Real time" facial recognition:** In the case of "real-time" systems, the capturing of the biometric data, the comparison and the identification occur all instantaneously, near-instantaneously or in any event without a significant delay. "Real-time" systems involve the use of "live" or "near-live" material, such as video footage, generated by a camera or other device with similar functionality.
- **"Post" facial recognition:** In the case of "post" systems, in contrast, the biometric data has already been collected and the comparison and identification occur only after a significant delay. This involves material, such as pictures or video footage generated by closed circuit television cameras or private devices, which have been generated before the use of the system in respect of the natural persons concerned.

Reference database: The repository of images against which a probe image is compared. In the law enforcement context, two main typologies of database exist:

- **Reference database of known suspects:** Composed of photos and mugshots of criminals, missing persons and persons of interest.
- **Investigative database:** Uniquely created for the purpose of an investigation, which is deleted when the case is closed.

Training dataset for facial recognition models: Repository of images of annotated faces that are used as an input to a model during the training phase, in order to make it learn from examples and provide correct predictions based on unseen data.

Contributors

Lead authors

Sebastien Louradour
French Government Fellow, World Economic Forum

Lofred Madzou
Project Lead, Artificial Intelligence and Machine Learning, World Economic Forum

Acknowledgements

Core Community

Maria Eira
Information and Technology Officer at United Nations, Centre for AI and Robotics, UNICRI

Inês Ferreira
Legal and Policy Research Fellow, Centre for AI and Robotics, UNICRI

Luc Garcia
Face Examiner, Forensics and Police Data Management Sub-Directorate, INTERPOL

Sylvia Jamgotchian
Policy Analyst, Executive Directorate Partnerships and Planning, INTERPOL

Rozemarijn Victoria Jens
Research Intern, Centre for AI and Robotics, UNICRI

Odhran McCarthy
Programme Officer, Centre for AI and Robotics, UNICRI

John Riemen
Lead Biometric Specialist, Center for Biometrics, Netherlands Police

Project Community

The World Economic Forum thanks the project community members for their insightful contribution on the principles:

Balques Al Radwan
Associate Programme Management Officer, United Nations Office of Counter-Terrorism/United Nation Counter-Terrorism Centre

Raja Chatila
Professor of Robotics, Université Paris Sorbonne

Samuel Curtis
AI Policy Researcher, The Future Society

Benedict Dellot
Head of AI Monitoring at the Centre for Data Ethics & Innovation

Jean-Luc Dugelay
Professor of Image Engineering & Security, Eurecom Nice Sophia Antipolis

Akvilė Giniotienė
Head of Cyber and New Technologies Unit, United Nations Office of Counter-Terrorism/United Nation Counter-Terrorism Centre

Patrick Grother
Biometric Standards and Testing Lead, NIST

Bruce Hedin
Independent AI consultant

Ameen Jauhar
Lawyer and Social Policy Researcher, Senior Resident Fellow, Vidhi Centre for Legal Policy – India

Brenda Leong
Senior Counsel & Director of Artificial Intelligence and Ethics, The Future of Privacy Forum

Teresa Magno
Investigative/Trial Judge, Assistant to the National Member for Italy, Eurojust

Claire Poirson
Lawyer, Bersay Avocats

Emmanuel Saliot
Adviser on Security and Technology, Council of the European Union

Sylvester Sammie
Human Rights Officer, United Nations Office of Counter-Terrorism/United Nation Counter-Terrorism Centre

Jessica Smith
Policy Advisor, Centre for Data Ethics and Innovation – UK Government

Melissa Taylor
Program Manager, NIST

Luc Tombal
Director, Defence and Security, Sopra-Steria

Vincent Toubiana
Head of LINC at CNIL

Jai Vipra
Vidhi Centre for Legal Policy – India

Endnotes

1. World Economic Forum, *A Framework for Responsible Limits on Facial Recognition: Use Case: Flow Management*, 2020: https://www.weforum.org/whitepapers/a-framework-for-responsible-limits-on-facial-recognition-use-case-flow-management (link as of 16/8/21).
2. World Economic Forum, *Responsible Limits on Facial Recognition: Use Case: Flow Management – Part II*, 2020: https://www.weforum.org/whitepapers/responsible-limits-on-facial-recognition-use-case-flow-management (link as of 16/8/21).
3. NISTIR 8280, *Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*, 2019: https://www.nist.gov/publications/face-recognition-vendor-test-part-3-demographic-effects (link as of 16/8/21).
4. Ibid.
5. Bobby Allyn, "'The Computer Got It Wrong': How Facial Recognition Led to False Arrest of Black Man", npr, 24 June 2020: https://www.npr.org/2020/06/24/882683463/the-computer-got-it-wrong-how-facial-recognition-led-to-a-false-arrest-in-michig (link as of 18/8/21).
6. ACLU, *The Dawn of Robot Surveillance: AI, Video Analytics, and Privacy*, 2019: https://www.aclu.org/report/dawn-robot-surveillance (link as of 16/8/21).
7. Monica Nickelsburg, "Washington State Passes Landmark Facial Recognition Bill, Reining in Government Use of AI", GeekWire, 13 March 2020: https://www.geekwire.com/2020/washington-state-passes-landmark-facial-recognition-bill-reining-government-use-ai/ (link as of 16/8/21).
8. Bill Atkinson, "Virginia to Enact Statewide Ban on Facial Recognition Use", Government Technology, 9 April 2021: https://www.govtech.com/public-safety/virginia-to-enact-statewide-ban-on-facial-recognition-use.html (link as of 16/8/21).
9. Emma Peaslee, "Massachusetts Pioneers Rules for Police Use of Facial Recognition Tech", NPR, 7 May 2021: https://www.npr.org/2021/05/07/982709480/massachusetts-pioneers-rules-for-police-use-of-facial-recognition-tech?t=1623343113224 (link as of 16/8/21).
10. iapp, "Will There Be Federal Facial Recognition Regulation in the US?", 11 February 2021: https://iapp.org/news/a/u-s-facial-recognition-roundup/ (link as of 16/8/21).
11. Rebecca Heilweil, "Big Tech Companies Back Away from Selling Facial Recognition to Police. That's Progress", Vox, 11 June 2020: https://www.vox.com/recode/2020/6/10/21287194/amazon-microsoft-ibm-facial-recognition-moratorium-police (link as of 16/8/21).
12. Amazon, "We Are Implementing a One-Year Moratorium on Police Use of Rekognition" 10 June 2020: https://www.aboutamazon.com/news/policy-news-views/we-are-implementing-a-one-year-moratorium-on-police-use-of-rekognition (link as of 16/8/21).
13. EUR-Lex, *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence*, 21 April 2021: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence (link as of 16/8/21).
14. Jorge Liboreiro, "'The Higher the Risk, the Stricter the Rule': Brussels' New Draft Rules on Artificial Intelligence", euronews, 21 April, 2021: https://www.euronews.com/2021/04/21/the-higher-the-risk-the-stricter-the-rule-brussels-new-draft-rules-on-artificial-intellige (link as of 16/8/21).
15. accessnow, "Privacy Win for 350,000 People in São Paulo: Court Blocks Facial Recognition Cameras in Metro", 12 May 2021: https://www.accessnow.org/sao-paulo-court-bans-facial-recognition-cameras-in-metro/ (link as of 16/8/21).
16. In the Court of Appeal (Civil Division) on appeal from the High Court of Justice of Queen's Bench Division (Administrative Court): https://www.judiciary.uk/wp-content/uploads/2020/08/R-Bridges-v-CC-South-Wales-ors-Judgment.pdf.
17. Rijksoverheid, "Letter of the Minister of Justice and Security of the Netherlands to MPs to Inform Them About the Use of Facial Recognition Technology by Law Enforcement Agencies (in Dutch), 20 November 2019: https://www.rijksoverheid.nl/documenten/kamerstukken/2019/11/20/tk-waarborgen-en-kaders-bij-gebruik-gezichtsherkenningstechnologie (link as of 16/8/21).
18. Chatham House, "Chatham House Rule": https://www.chathamhouse.org/about-us/chatham-house-rule (link as of 16/8/21).
19. In law enforcement, there exist instances where "one to one" also relates to identification activity; for example, in disputed identity cases or where an image is compared in a case with a possible suspect.
20. INTERPOL, "Red Notices": https://www.interpol.int/How-we-work/Notices/Red-Notices (link as of 16/8/21).
21. Facial Identification Scientific Working Group: https://www.fiswg.org (link as of 16/8/21).
22. Definitions extracted from "Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts", 21 April 2021: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206&from=EN (link as of 18/8/21).

WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744
contact@weforum.org
www.weforum.org