Centre for the Fourth Industrial Revolution

WORLD ECONOMIC FORUM

Responsible Limits on Facial Recognition

Use Case: Flow Management

Part II

Pilot phase: Self-assessment, the audit management system and certification

WHITE PAPER

DECEMBER 2020

Cover: Getty Images/da-kuk

Inside: Getty Images/Prostock-Studio; Getty Images/Izisek; Getty Images/KENGKAT; Getty Images/Wonry; Getty Images/DNY59; Getty Images/Imaginima; Getty Images/MangoStar_Studio

Contents

3 Foreword

4 Preface

5 Introduction

7 Methodology

9 1. Test of the assessment questionnaire by Narita International Airport

10 1.1 General framework and objective
10 1.2 Case study: One ID programme at Narita International Airport, Japan

13 2. An audit framework to validate compliance with the principles for action

14 2.1 General framework and objective
15 2.2 Structure of the audit framework
17 2.3 Extract from the audit framework

18 3. A certification scheme to ensure the responsible use of facial recognition technology for flow management

19 3.1 General framework and objective
20 3.2 Certification process

21 4. From principles to certification: A journey to build accountability

22 4.1 An organization already offers a facial recognition system and wishes to get certified
24 4.2 An organization intends to develop a facial recognition system
26 4.3 Consequences of major non-compliance

27 5. Conclusion

29 Glossary

30 Contributors

32 Appendices

32 Appendix A: Answers from Tokyo-Narita International Airport to the assessment questionnaire
39 Appendix B: Audit framework

51 Endnotes

Foreword

The first global initiative to build trust and transparency for the use of facial recognition

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

Hideharu Miyamoto,
Senior Executive Officer, Narita International Airport Corporation, Japan

Julien Nizri,
Managing Director, AFNOR Certification, France

Toshifumi Yoshizaki,
Senior Vice-President, NEC Corporation, Japan

The need for seamless and contactless technology to accurately identify customers, employees and vendors has never been so critical. Along these lines, last year the World Economic Forum launched an initiative to build a governance framework for the responsible use of facial recognition technology and remote biometrics in the context of improving the airline passenger boarding experience. While last year these technologies were nice-to-have, now in the midst of an unprecedented global COVID-19 pandemic, remote biometrics have become a must-have.

Although the airport industry has used digital technologies for decades, artificial intelligence is fundamentally transforming the passenger experience, reducing waiting times and improving overall convenience. But this progress does not come without trade-offs and risks. Facial recognition is under scrutiny, with concerns over risks of bias, potential discrimination and personal data exposure. To proactively address these issues, we have combined our efforts to build a global definition of what represents the responsible use of facial recognition for flow management use cases and a governance framework to operationalize this definition.

Following the methodology proposed in this White Paper and with a view to ensuring transparency, Tokyo-Narita International Airport Corporation and NEC Corporation have joined the Forum initiative to test the assessment questionnaire presented in this pilot project's first White Paper entitled "A Framework for Responsible Limits on Facial Recognition – Use Case: Flow Management". Adopting a trailblazing approach, they decided to publicly share their results to demonstrate to industry players and policy-makers how transparency and passenger trust could be achieved.

This White Paper aims to advance the conversation on certification and third-party audits for the responsible deployment of facial recognition technology. In this regard, AFNOR Certification has played a critical role by designing an audit framework for this initiative's pilot phase that is now ready to be tested by organizations volunteering to validate the third-party certification scheme.

The World Economic Forum encourages organizations to join this initiative to test and adopt this policy framework and globally engage in the trusted use of facial recognition systems.

Preface

In Part II, the focus is on the last step – introducing the audit framework and certification scheme co-designed for industry actors.

In April 2019, the World Economic Forum Centre for the Fourth Industrial Revolution launched the Responsible Limits on Facial Recognition project. It seeks to address the need for a set of concrete guidelines to ensure the trustworthy and safe use of this technology through the design of a robust governance framework. To achieve this goal, the Forum is spearheading a multistakeholder, evidence-based policy project with France and Japan, which have recently joined the initiative, as anchor partners. The working group was initially composed of industry representatives considering procuring facial recognition systems (Groupe ADP and SNCF), technology providers (Amazon Web Services, IDEMIA, IN Groupe and Microsoft), policy-makers (members of the French Parliament), academics, civil society organizations and AFNOR Certification.

During the scoping phase, this working group decided to adopt a use-case-based approach because risks associated with facial recognition systems are highly contextual. Indeed, false positive and false negative results may lead to very different outcomes whether a given system is used to accelerate a boarding process or track a person of interest. Therefore, by focusing on a real-world application, a specific system in operation and the groups of stakeholders potentially impacted by that system (e.g. airline passengers), the opportunity to co-design a governance framework that effectively mitigates its associated risks is greater.

After careful consideration, the working group members decided to focus on "flow management" (using facial features as a means to access a service) primarily because this use case is likely to develop in the coming years. For instance, the organizers of the Tokyo Olympic Games announced the use of facial recognition to manage athlete and staff access to stadia and Olympic facilities.[1] Also, airports and airline companies have started using this technology.[2]

To design a balanced and actionable governance framework, the working group developed a method comprising four main steps: 1) define what constitutes the responsible use of facial recognition technology (FRT) by drafting a set of **principles for action**; 2) design a set of **best practices** to support the application of these principles; 3) assess, through an **assessment questionnaire**, if organizations comply with them; and 4) validate compliance with the principles for action through an **audit framework** and a **certification scheme**.

To design the latter, a partnership was signed with AFNOR Certification, whose expertise in auditing and certification is recognized internationally. Considering the initiative's primary concern to mitigate risks that may cause degradation or interruption of the service offered to passengers when airports deploy facial recognition systems, AFNOR Certification suggested designing a quality management system audit, rather than an audit of facial recognition algorithms, for two main reasons. First, airport companies are responsible for the quality of service they provide to their customers and need guidance on how to improve that quality. Building on solid foundations by following ISO 9000 quality management standards will support them in this process. Second, auditing facial recognition algorithms directly, though crucial, raises a set of fundamental challenges for a certification body. For instance, what is the acceptable threshold of performance for systems in operation, especially when this performance is dynamic and in constant evolution? How can legal and ethical considerations (e.g. fairness) be translated into quantitative requirements that can be assessed? How can the lack of explainability of AI systems be addressed and interpretable decisions be produced? These complex and open questions require further research. Yet, carrying out an audit of the human organization and rigour of the procedures guarantees a level of responsible use of FRT for flow management.

Interested in this practical approach, Japanese stakeholders seeking to pursue their efforts to ensure the responsible use of FRT at airports decided to join our initiative. To this end, the Japanese Government and NEC Corporation nominated two fellows to work at the World Economic Forum Centre for the Fourth Industrial Revolution Japan, where they play an integral role in shaping this initiative. The Government of Japan and NEC are highly valuable partners for two main reasons. First, FRT is being deployed in Japanese airports and thus offers an insightful use case. Indeed, last year, Narita International Airport – Japan's leading airport with over 41 million passengers per year[3] – announced plans to start using this technology in 2020 to streamline security checks and the boarding process through the One ID[4] service provided by NEC. More recently, the Japanese Government released guidelines for biometric data protection to regulate this use.[5] Second, Narita International Airport agreed to self-assess its facial recognition system using the assessment questionnaire presented in this project's first White Paper.[6] Thus, this collaboration represents a good opportunity for the airport to promote transparency and accountability for its local and global passengers.

In the first White Paper, published in February 2020, the first three steps of this method were presented in detail. In the present paper, which constitutes Part II, the focus is on the last step – introducing the audit framework and certification scheme co-designed for industry actors. Narita International Airport's answers to the assessment questionnaire are also presented as an example of rigorous self-assessment.

Introduction

The need for a balanced governance framework for facial recognition technology has never been so critical.

“ The overarching objective of this World Economic Forum initiative is to establish a comprehensive governance framework to ensure the responsible use of FRT.

Over the last few years, rapid technological advances, due mainly to progress in machine learning and sensors, have fuelled the development of FRT. This has enabled its trajectory from research to adoption in industry. Indeed, this technology has now expanded into various areas of public and private life, including in banking, retail, transportation, law enforcement and even healthcare.

The development of FRT creates considerable opportunities for socially beneficial uses, mostly through enhanced authentication and identification processes, including unlocking a phone, boarding a plane and accessing public services online. But it may also undermine civil liberties or lead to discriminatory outcomes. For example, in the United States, entertainment venues[7] have used this technology on consumers without advance notice or consent, data breaches[8] on remote biometrics are regularly reported and powerful solutions are being built through controversial data scraping practices.[9] Recently, the technology led to the wrongful arrest and custody of an innocent African American.[10]

These controversies have led to intensified policy activity. In the United States, various municipalities have banned the use of FRT by city agencies, including in California (San Francisco,[11] Oakland[12]) and five cities in Massachusetts (Boston,[13] Brookline,[14] Cambridge,[15] Northampton[16] and Somerville[17]), while Portland, Oregon has banned both public and private use of the technology in public spaces.[18] At the state level, Washington was the first state to pass legislation[19] to put guardrails on government use of FRT. Democratic lawmakers have proposed federal legislation[20] to permanently ban law enforcement agencies from using it. This initiative complements the list of policy proposals being discussed by various stakeholders, which includes different sets of principles,[21] a moratorium[22] and the creation of a new federal office[23] – a dedicated regulator for FRT.

Further, large technology companies have also formulated positions on this topic. Microsoft[24] has pledged to stop selling FRT to law enforcement agencies until a federal regulation is in place. Amazon Web Services (AWS)[25] has implemented a one-year moratorium on police use of its platform *Rekognition*, while IBM has announced that it will no longer offer, develop or research FRT.[26]

A European Commission White Paper published in February 2020 on the governance of artificial intelligence[27] actively discusses the possibility of introducing additional requirements to limit the deployment of FRT. In January, in the draft version of the paper, the European Commission reportedly considered introducing a five-year moratorium[28] on facial recognition use in public spaces to create time to develop appropriate regulation, but did not mention this option in the published version.

Unsurprisingly, most of this policy activity is related to government and law enforcement agencies' use of this technology because in these domains the risk of misuse and the surveillance of suppressed groups are high. Certainly, considering the sensitivity of biometric data, the use of facial recognition is intrinsically risky. Indeed, boarding a plane using facial recognition, or accessing a stadium, or using face-based advertising in retail also involves risks (e.g. privacy violations, inequalities in access to services because of performance gaps between different demographics, etc.). Therefore, the identification and effective mitigation of risks across use cases are needed. The overarching objective of this World Economic Forum initiative is to establish a comprehensive governance framework to ensure the responsible use of FRT, starting with one use case scenario: flow management.

The previously published White Paper, “A Framework for Responsible Limits on Facial Recognition – Use Case: Flow Management” had two main sections: a presentation of the pilot-based approach to policy-making and its application to the flow management use case through a method structured in four steps. The current White Paper is structured in four sections. First, the Narita International Airport use case is presented, detailing the airport's test of the assessment questionnaire conducted in collaboration with NEC. The answers appear in Appendix A. Second, the audit framework co-designed with AFNOR Certification is introduced, detailing its function, ambition, various uses and structure. The full framework is available in Appendix B. Third, the certification scheme is described, explaining its purpose, benefits and process. Finally, the step-by-step journey from the principles to the issuance of the certificate that an industry player should complete to demonstrate its trustworthy use of FRT for flow management applications is laid out. In the last section, each step includes the activities to run.

The two White Papers combined should empower organizations in their journey towards the responsible use of FRT. They present the documentation organizations should review (principles for action, best practices, assessment questionnaire and audit framework) and the scheme they need to complete to obtain certification.

These papers also serve a greater purpose: to advance the conversation about the regulation of FRT at the local, regional and international levels by providing a hands-on method for the mitigation of risks applicable to various use cases. But as this conversation and the pilot are still in progress, the World Economic Forum encourages organizations involved or willing to take part in this discussion to join in this journey.

Methodology

A pilot-based approach replicable in other FRT use cases

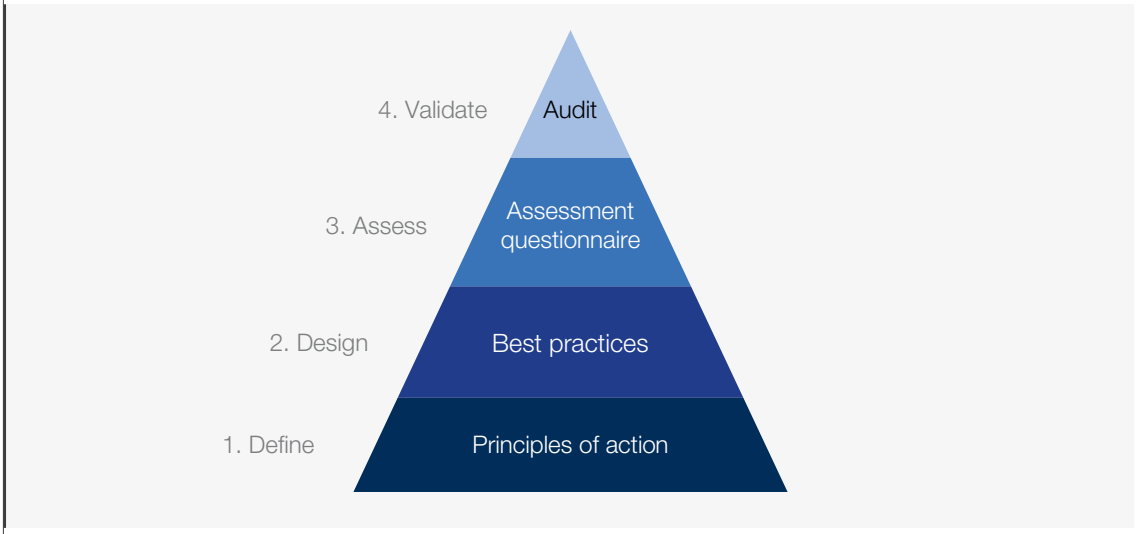
A four-step approach

The first White Paper, "A Framework for Responsible Limits on Facial Recognition – Use Case: Flow Management" introduced a method based on four main steps (Figure 1) to build a robust governance framework able to ensure the responsible use of facial recognition:

1. **Define** what constitutes the responsible use of facial recognition technology by drafting a set of principles for action. The first objective of the working group, composed of public figures, companies that design and procure facial recognition systems, regulatory bodies, academics and representatives of civil society, was to establish a shared definition, organized around 11 principles
2. **Design** a set of methodologies, tailored by use cases, to support product teams in the development of systems "responsible by design"
3. **Assess** to what extent the system designed is responsible through an assessment questionnaire that describes for each use case what rules should be respected to comply with the principles for action
4. **Validate** compliance with the principles for action through the design of an audit framework by a trusted third party.

Each of these steps represents an additional level of commitment from industry actors to the trustworthy use of FRT.

FIGURE 1: **The four steps to ensure the responsible design and use of facial recognition technology for flow management use cases**



Source: World Economic Forum, "A Framework for Responsible Limits on Facial Recognition – Use Case: Flow Management", White Paper, February 2020

Tested in a policy pilot before deployment

In accordance with the experimental approach, each element of this governance framework (principles for action, best practices, assessment questionnaire and audit framework) will be tested and reviewed based on the practical findings of the policy pilot. If the results are satisfactory, the certification scheme will be deployed in collaboration with partnering certification bodies, starting with AFNOR Certification, which has played a key role in this initiative.

A replicable and scalable method

The four-step methodology could be replicated in other facial recognition use cases. Indeed, any responsible deployment strategy should start by establishing a clear definition of what constitutes responsible use in a given domain. This could be achieved by drafting a set of principles through a multistakeholder approach. The definition can then be operationalized in product development, given appropriate design requirements or best practices. Finally, it can be tested through a tailored assessment questionnaire and validated through the design of an audit framework. Therefore, interested stakeholders will mostly just need to determine what items are required to adapt these tools to their context and industry domain.

1 Test of the assessment questionnaire by Narita International Airport

The first publicly shared self-assessment on FRT by an organization

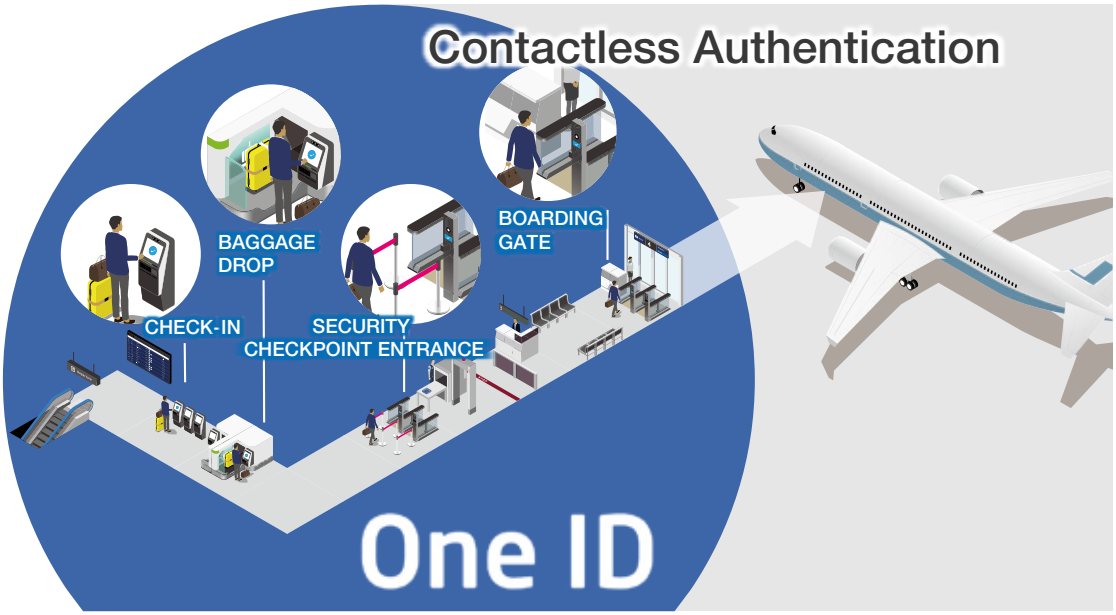
1.1 General framework and objective

The assessment questionnaire serves as a self-assessment document that details the requirements organizations must respect to ensure compliance with the principles for action. As such, the questionnaire can be used individually to perform an internal assessment of existing processes, as Narita International Airport chose to do (see below), or as a tool to measure readiness for the certification process (see section 4).

To build the assessment questionnaire, each principle for action was divided into a set of specific questions to be addressed by a cross-departmental task force appointed by management. In this regard, it is similar to the audit framework but is less detailed.

1.2 Case study: One ID programme at Narita International Airport, Japan

Presentation of the case study



The One ID programme will accelerate the passenger journey from check-in, baggage drop, security checkpoint entrance to the boarding gate

Source: Provided by NEC Corporation

As one of the world's leading airport companies, Narita International Airport aims to provide the best passenger experience possible, especially important in view of Japan's position as one of the world's most popular travel destinations. To achieve this goal, it decided to introduce a "One ID" programme using FRT to accelerate the "check-in to boarding" process. The project started in 2016 with a clear vision: for passengers to complete the entire process at "walking pace".

The system was designed to work using passengers' face data recorded at initial touchpoints, such as at self-service kiosks. This data is linked to their passports that include embedded integrated circuit (IC) chips and boarding pass information. Once the process is completed, travellers can go through the entire check-in to boarding process without presenting their passports or boarding passes. They will be able to pass through the security checkpoint entrance and boarding procedures at a walking pace. This smooth process saves a great deal of passenger time. It also represents a valuable tool to prevent the spread of disease, decreasing the risk of spreading viruses that cause disastrous effects such as those generated worldwide by the COVID-19 pandemic.

The following section presents the One ID programme for industry actors potentially interested in deploying similar solutions.

Key dimensions of the facial recognition system deployed

User experience

It usually takes approximately 15 minutes for an airline company to complete the boarding process for all passengers at the gate. To align with this target, the experiment was designed to complete the boarding process for 250 passengers spread across three lanes within 15 minutes. When Narita International Airport implemented the system requirements, it followed the guidelines of the Immigration Services Agency of the Ministry of Justice[29] and took into consideration the results of the US National Institute of Standards and Technology (NIST)'s[30] evaluation[31] of automated face recognition algorithms to achieve the highest level of performance.

A holistic approach was required to achieve the successful implementation of the facial recognition system to ensure that passengers go through the check-in to boarding process at a walking pace. More was needed than simply considering the processing capabilities of the facial recognition system. Thus the decision was taken for NEC, the technology solution provider, in collaboration with other vendors, to test the system's availability, reliability and performance at each touchpoint (during check-in, at the baggage drop, at the security checkpoint entrance and during boarding).

Bias mitigation

Aware of bias issues related to FRT, Narita International Airport took clear steps to mitigate their potential adverse effects, including by selecting technology provider NEC Corporation, a world leader in this field. Indeed, the NIST conducted a robust study[32] in 2019 on the effects of race, age and sex on the facial recognition software, using four large data sets. It evaluated 189 software algorithms, using federal government data sets that contained roughly 18 million images of 8.49 million people. The study revealed significant biases in facial recognition software for people of colour and women. NEC's proposed solution was found to be among the least biased across age, gender and ethnicity, while achieving a high level of accuracy.[33] In addition, Narita International Airport committed to complying with the Universal Design Basic Plan[34] of the Japanese Ministry of Land, Infrastructure, Transport and Tourism (MLIT) and the Tokyo 2020 Accessibility Guidelines[35] for the Tokyo Olympics and Paralympics, which both aim to create a more inclusive society for people with disabilities.

Further, NEC was primarily in charge of designing the facial recognition software, running verification tests and adjusting its parameters. It also worked in collaboration with airline companies on the installation of the devices, the lighting environment, operational tests and training scenarios to ensure the best deployment possible.

Data protection

Facial recognition is one of the most sensitive biometric technologies available. As such, its deployment requires a great deal of care and consideration regarding its potential impact on the privacy of the 40 million passengers who go through Narita International Airport every year. The Japanese Government is fully aware of this challenge. To ensure the responsible deployment of facial recognition by airports, the MLIT appointed a Personal Data Management Study Group, a multistakeholder task force comprising representatives of the International Air Transport Association (IATA), the Personal Information Protection Commission, legal experts and consumer associations. The working group conducted a year-long study that led to the publication of multidimensional guidelines for the management of personal data collected through the One ID programme at airports.[36] This document includes a list of processes airports should follow to address the privacy risks associated with that service and ensure a high level of data protection.

Narita International Airport strictly applies all the guidelines to ensure the responsible deployment of the One ID programme. It has paid particular attention to the processes related to major risks, such as data breaches. Indeed, since cyberattacks are increasingly more sophisticated, preventing them has become increasingly difficult, even for governments or global companies. To mitigate this risk, members of the Personal Data Management Study Group recommend deleting biometric data collected at check-ins within 24 hours. In addition, Narita International Airport runs regular cyber tests to assess the robustness of its servers.

Reflection on the self-assessment phase

Narita International Airport's journey to ensure the responsible use of FRT for flow management started when it began to consider deploying this technology in 2016. At that time, no official guidelines presenting the processes to implement to achieve this goal existed. Therefore, initially the airport conducted an internal reflection, but it is encouraged today to view the progress made in this domain over the past four years, first in Japan and currently internationally through the World Economic Forum initiative.

The multistakeholder process initiated by the MLIT through the Personal Data Management Study Group and the publication of the guidelines has proved useful. It also aided in the preparation of the self-assessment phase because many of the items listed in the questionnaire, such as those related to the right to information and consent, had already been addressed in the guidelines. Japan's lead on this topic was confirmed, and when the answers to the questionnaire were reviewed (see Appendix A), additional progress was observed, leading to the prospect of achieving an additional level of trust from local and global passengers through the initiative.

Reaching this milestone took time, however. While Narita International Airport joined the Forum's Responsible Limits on Facial Recognition initiative in early February 2020, the self-assessment process only began in early May and took six weeks to complete, longer than initially planned. Two issues caused the delay. First, the assessment questionnaire is comprehensive, covering all aspects of system management: data governance (e.g. data security and usability), performance and accuracy, and user experience (UX). Therefore, answering all the questions required the involvement of various departments both within Narita International Airport and NEC Corporation. Second, the Japanese Government's declaration of a state of emergency in response to the COVID-19 pandemic disrupted the initial timeline.

Narita International Airport hopes that its contribution will help to improve the assessment questionnaire and process, as envisaged in the iterative process adopted by the project community. For instance, it recommends strengthening the section related to data security and encouraging industry actors to follow its approach to delete biometric data within 24 hours after enrollment. It also suggests there is less need to run an external audit as suggested by the project methodology, as the processes introduced to ensure the responsible use of FRT are sufficiently robust. Indeed, the One ID programme was designed and implemented through a multistakeholder process that involved government officials, Narita International Airport, airline representatives, various vendors and legal experts. In the Airport's view, as a global initiative, it is essential to allow project partners to choose their enforcement mechanisms according to their organizational culture.

2

An audit framework to validate compliance with the principles for action

A quality management system audit to ensure the responsible use of FRT and validate risk-mitigation processes

This section details the work undertaken with AFNOR Certification to draft an audit framework to attest the compliance of organizations deploying FRT for flow management with the principles for action. It presents the general framework and its objective, the way it can be used and its structure, and provides an example of how it has been completed (the full audit framework is presented in Appendix B). An updated version of the principles for action are also provided, based on observations and feedback from the test conducted by Narita International Airport using the assessment questionnaire.

2.1 General framework and objective

Function of an audit framework
An audit framework's function is to serve as a reference document that details the requirements and processes of an audit for a defined scope. As such, it can be used to provide guidance on best practices, to conduct internal audits, to help formulate the needs that providers must meet during the development of a new project, and to enable participation in a voluntary or mandatory certification process. Usually, the stakeholders who acknowledge the need for an audit framework also determine how they want to use it. Within the scope of this pilot project, a multistakeholder community drafted an audit framework to use as a tool to validate compliance with the principles for action, which defines responsible use of FRT for flow management applications.

Designed for flow management applications
This audit framework was designed exclusively for flow management applications – that is, situations in which individuals' facial features are captured for use in accessing a service, such as boarding a plane or entering a concert hall. It is worth noting that the designed system offers an opt-in type solution in which users are offered the option to use this service for a perceived benefit. As such, it significantly differs from use cases in which facial recognition is deployed without the knowledge or consent of citizens or consumers. In this application, the audit framework is meant to address a set of issues specifically related to this type of use case and to validate compliance with the principles for action. It deals with concerns related to the governance of biometric data (e.g. consent, privacy), the performance of the facial recognition system across different demographics (e.g. identification and mitigation of biases, a defined performance threshold) and the empowerment of end users through the UX of the system (e.g. information display, right of access, availability of an alternative option). As such, it aims to be the first comprehensive framework for the responsible use of facial recognition for flow management applications. It is not intended for other use cases (e.g. person of interest tracking based on a warrant or terrorism risk, personalized shopping in retailing, identification of rare disease) and their associated risks.

Tested in a policy pilot before deployment
In line with the experimental approach, this audit framework will be tested and reviewed based on the practical findings of the policy pilot, in collaboration with AFNOR Certification and volunteering airports. If the results are satisfactory, the audit framework will be deployed in collaboration with partnering certification bodies, starting with AFNOR Certification, which has played a key role in the design of this initiative.

A quality management system audit
While drafting this audit framework, the working group decided to focus on mitigating risks that may cause degradation or an interruption of the service offered to the end users. For instance, when engaging in the process of boarding passengers onto a plane using captured facial features, what processes should be introduced to ensure that the passengers have equal access to this service regardless of their demographics or any physical condition that may impact the performance of FRT? If the system is dysfunctional, what reasonable alternatives are available to make sure someone does not miss their flight? Therefore, this framework is designed for an audit of the management of facial recognition systems, not their algorithms. The working group made this decision mostly because users of the technology (e.g. transportation companies, event organizations) are responsible for the quality of service they provide to their customers. In that regard, the principles for action serve as a set of requirements that describe how a high-quality facial recognition system should be designed and operated, while the audit framework details what processes should be implemented to ensure the effective delivery of that service to end users.

Built from a European perspective
The first version of the audit framework was drafted through a multistakeholder process, similar to the one followed to draft the assessment questionnaire. Careful attention was paid to the EU General Data Protection Regulation (GDPR). Also considered were the "Ethics Guidelines for Trustworthy Artificial Intelligence" of the European Commission's High-Level Expert Group on Artificial Intelligence, a vital document that paves the way for the ethical use of AI technologies across the EU. This means that Data Protection Officers can leverage this framework, in combination with legal aid, to check their compliance with EU data protection authorities when they process EU citizen data or operate within a country that has obtained an adequacy agreement under the GDPR with the EU, like Japan.[37] Organizations using FRT for flow management

applications that are not based within jurisdictions where the GDPR applies are also encouraged to use this audit framework to improve the responsible management of their system for the satisfaction of their end users. In fact, many jurisdictions across the world are now considering data protection laws, with the GDPR as a global compass. In this perspective, the work of the initiative presented in this paper may help to integrate FRT in future data-protection laws.

How to use the audit framework

As mentioned, the multistakeholder task force decided to use the audit framework as a tool to validate compliance with the principles for action, which defines what is responsible use of FRT for flow management applications. However, this validation can take different and complementary forms. Four ways in which organizations can use the audit framework have been identified:

1. **Guide of best practices**. An organization can use this audit framework as a blueprint to design and deploy its facial recognition system responsibly. In this case, it would integrate into the set of specifications those that relevant internal teams and external providers must respect during the project development.
2. **Self-assessment**. An organization that is about to deploy a facial recognition system or has already done so can conduct a self-assessment using the audit framework, similar to the assessment questionnaire. But as the audit framework is more comprehensive, the self-validation process would be more rigorous.
3. **Certification**. A trusted third party, ideally an accredited certification body, can assess the robustness of the processes implemented by an organization willing to comply with the principles.
4. **Regulation**. Policy-makers can also pass legislation that would require industry actors using facial recognition technology for flow management applications to be audited. In this case, it would become a statutory audit.

The third use of the audit framework, certification, was the option selected for this project. It is explored in the next sections of this paper. Although AFNOR Certification is the first certification body to participate in this initiative, others will be invited to join this project to build a network able to deliver this certificate globally, once the audit framework is tested and validated.

2.2 Structure of the audit framework

> The current version includes 10 principles, although they may continue to evolve based on the final results of the policy pilot.

To build the audit framework, the principles for action were transformed into a set of requirements that must be validated during the audit (described below). The requirements were listed by criteria and classified into three distinct types.

Principles for action

The first version of the principles for action was presented in the first White Paper published in February 2020. The principles were co-drafted in a multistakeholder process and define what constitutes the responsible use of facial recognition for flow management applications. Initially, 11 principles were identified but, during the testing phase, they were reviewed and updated to ensure effective implementation, completeness and relevance. As a result, the current version includes 10 principles, although they may continue to evolve based on the final results of the policy pilot.

1. **Proportional use of facial recognition systems**
   Facial recognition systems should be highly tailored according to the intended use. Organizations using facial recognition systems should take reasonable steps to assess the capabilities and limitations of the systems they intend to use and ensure that their systems are appropriate for purpose.
2. **Risk assessment**
   Organizations creating facial recognition platforms or using facial recognition as part of a service or system should conduct a comprehensive risk assessment of their systems, including the impact on privacy, potential for errors, susceptibility to unfair bias, vulnerability to hacking and cyberattacks, lack of transparency in the decision-making process and potential for civil and human rights infringements.
3. **Bias and discrimination**
   Organizations using facial recognition systems should take appropriate steps to ensure that all unfair bias or outcomes (i.e. not being recognized by FRT and consequently being subjected to a poorer quality of service) can be detected, identified and mitigated to the greatest extent possible. While acknowledging that the complete removal of bias represents one of the biggest challenges in AI research, organizations must assign appropriate resources to the implementation of tools and processes that minimize bias or unfair outcomes.
4. **Privacy by design**
   Organizations using facial recognition systems should design systems to support privacy, including privacy considerations in system requirements and carrying through privacy support in the design, development and testing of technology as well as in supporting business practices and ongoing system maintenance.

5. **Performance**
   Organizations creating facial recognition platforms or using facial recognition as part of a service or system should follow the standards for evaluating the accuracy and performance of their systems at the design (lab test) and deployment (field test) stages. Performance assessments should be auditable by competent third-party organizations and their reports made available to users of the systems.

6. **Right to information**
   Processes should be put in place to inform end users who have questions and/or need information on the use of facial recognition systems. End users should have access to their personal biometric data upon request.

7. **Consent**
   Individuals should provide informed, free, unambiguous, explicit and affirmative consent for the use of facial recognition systems. Thus, no unique biometric identifier should be created and maintained without explicit consent. Any time data subjects enrol for a new service powered by FRT, they should express clear consent with regard to the length of data retention and the terms of the data storage.

8. **Information display**
   When used in public spaces, clear signage should be deployed to ensure obvious communication with end users on the use of the facial recognition technology. Areas where facial recognition systems are used should always be delimited and indicated to individuals. A visual sign should also inform individuals when the system is in operation.

9. **Right of access to vulnerable groups**
   Facial recognition should not exclude anyone and should always be accessible to and usable by all groups of people, including elderly people and people with disabilities. It is recognized that there may be some instances, such as with infants and children, in which an exception to this principle is appropriate and an alternative to facial identification should be offered.

10. **Alternative option and human presence**
    A manual review (human overseeing) should be conducted for any use that could result in a consequential decision, such as causing a civil rights infringement. In the case of a fully automated system, a fallback system with a human presence should always be in place to address exceptions and unexpected errors, and for possible remediation purposes. A reasonable alternative to the use of facial recognition systems should always be in place.

Three types of requirements

Compliance with the audit requirements is assessed at three key stages: at the process design stage, implementation stage and operational stage:

– *Requirements related to the processes introduced in the design of a facial recognition system*. The purpose of these requirements is to assess the various processes implemented and the resources allocated in the design stage. The goal is to ensure that the design **guarantees the responsible and trustworthy deployment and use of FRT**.

– *Requirements related to the implementation of these processes while a facial recognition system is in operation*. The purpose of these requirements is to validate compliance with established processes, their continued implementation and long-term existence once the system is deployed in a specific real-world circumstance. It is essential to validate the durability of the system and the absence of drifts in its use or from the initial objectives. These requirements will **help build confidence in the operation and management of the system by ensuring that they meet the expectations established during the design stage**.

– *Requirements related to the system's functioning*: The purpose of these requirements is to validate that the system operates in accordance with the principles for action. Some of these requirements are linked to those established during the design phase. They will make it possible to take a snapshot of the system's operation, validate the user experience, and carry out various tests to make sure the system operates in compliance with the principles for action. These requirements will validate the **responsible use of the system**.

2.3 Extract from the audit framework

The full audit framework is presented in Appendix B of this White Paper, but an extract follows to illustrate its structure:

- In the framework's left column, the requirement numbers are listed.
- In the middle column, detailed text describes each requirement.
- In the right column, the relevant type of requirement is indicated, as described above.

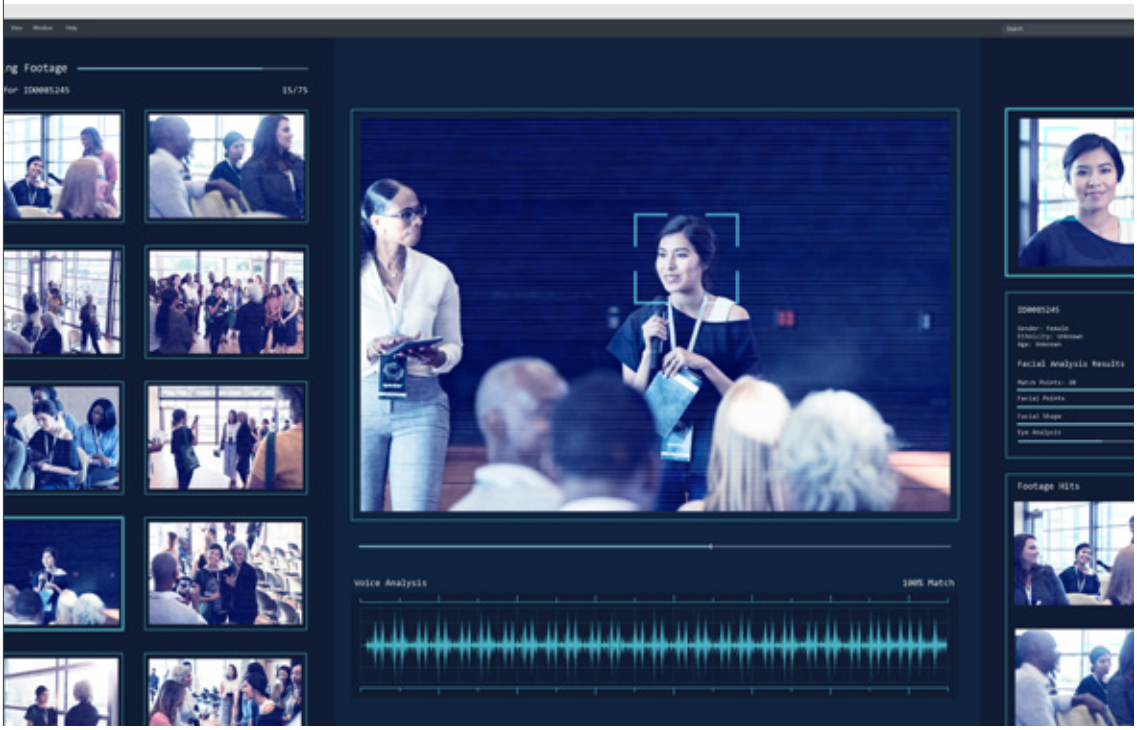
Proportional use of facial recognition systems
Requirement: Facial recognition systems should be highly tailored according to the intended use. Organizations using facial recognition systems should take reasonable steps to assess the capabilities and limitations of the systems they intend to use and ensure that their systems are appropriate for the intended purpose.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
1.1	Prior to any facial recognition project, the need leading to considering the use of a facial recognition system must be defined. Companies must describe the technical requirements to achieve the objectives assigned to their system and be able to guarantee that the system will only be used for its intended purpose.	✓		
1.2	The set of alternatives (excluding facial recognition) that fulfil the same need must be determined.	✓		
1.3	To fulfil the need, possible alternatives to the use of a facial recognition system must be identified. A documented process and methodology for analysing possible solutions must be set up. The objective is to assess the use of the facial recognition technology's relevance to its purpose and the resolution of the problem. To this end, companies must describe in detail the assessment and selection methodology, which must at least include: – A review of the identified advantages and disadvantages for every identified solution – A definition of the system's expected benefits for the various stakeholders (users, state, citizens, etc.) – A risk analysis covering false positive and false negative situations (in particular, the risks of violating civil rights) – A quantified assessment of the expected benefits – A comparative analysis of the different solutions – The conclusion that led to the preference for a facial recognition solution.	✓		
1.4	To validate the assumptions that led to the choice of facial recognition technology, companies must define the parameters to be respected to validate the relevance of its use (for example: expected false positive and false negative rates, expected performance).	✓		
1.5	These parameters must be checked in the use phase.		✓	
1.6	The facial recognition system was introduced to meet specific needs in the framework of particular uses. When used, the facial recognition system must be limited to the initially planned uses and validated for those uses.			✓

3

A certification scheme to ensure the responsible use of facial recognition technology for flow management

A certification scheme delivered by independent third parties to ensure trustworthy oversight.



Of the various ways available to validate compliance with the principles for action, the working group decided to focus on certification and to partner with AFNOR Certification. This trusted third party will be in charge of auditing volunteering industry actors by using the audit framework. Moving forward, when the certification scheme is tested and validated, the specific capabilities and competencies that a certifying body must have to conduct an assessment of FRT in accordance with the audit framework will be identified, and other certification bodies that have the requisite capabilities and competencies will be encouraged to adopt and run this certification scheme. This section presents the general framework of the certification scheme (its objectives, how it works, who is eligible, etc.) and the certification process in detail.

3.1 General framework and objective

Function of a certification scheme
The ultimate goals of a certification scheme are to: 1) ensure that the system or service being certified meets prespecified standards of quality (effectiveness, efficiency, safety and adherence to social values and norms); and 2) encourage ongoing improvements in quality. It achieves those goals through the performance of an independent assessment and the attainment of an objective judgement of a given system or product based on a defined set of requirements listed within an audit framework. In other words, its immediate goal is to rule on the level of compliance. As such, it is a robust and yet flexible signalling device for industry actors seeking to demonstrate the trustworthiness of their products or systems. Usually, applicant organizations that engage in such a process pursue various objectives: to improve their competitiveness, promote their best practices, achieve an additional level of trust with their customers and partners and/or comply with regulatory requirements.

A quality management certification
As already mentioned, a quality management system audit was co-designed, similar to the ISO 9000 family of standards. Consequently, the certification scheme focuses on the management of facial recognition systems for flow management applications and validates their compliance with the principles for action. Yet, different types of certification (e.g. products, services, professional skills, etc.) are beyond the scope of this initiative. Also, the certification scheme can take different forms depending on the objectives of the applicant organizations and the market demand. Presented below are three types of certification in related domains to illustrate how the voluntary certification scheme may evolve.

Benefits of being certified in related domains
– *Certification as a necessary step to access certain markets:* Service providers seeking to access new markets in the digital realm, in particular those that work with large organizations or public-sector bodies, must demonstrate their ability to secure their information systems by being ISO/IEC 27001 certified. Such certification allows client organizations to ensure the confidentiality, integrity and availability of the data that they have entrusted to their service providers through the implementation of internationally recognized data security processes.

– *Certification as a voluntary process permitted by a regulation:* Organizations that process European personal data must comply with the GDPR. As part of this obligation, they must make sure that their subcontractors to whom they entrust personal data are also compliant with this regulation. Subcontractors can apply for a GDPR certification as permitted by Article 42 of the GDPR to reassure the principal contractors and obtain a significant competitive advantage.

– *Certification cited in a regulation:* Any organization seeking to provide hosting services for personal health data on behalf of health professionals in France must obtain a Health Data Host (HDS) certificate. This sectoral legal obligation ensures that actors processing sensitive data, such as personal health data, implement technical and organizational measures to ensure their data protection.

Who should certify?
Once this audit framework is tested and validated by AFNOR Certification, the objective is to make it available to other certification bodies. Further, to ensure the independence and impartiality of the certification process, certification bodies should be operating to ISO/IEC 17021-1:2015. This standard contains "principles and requirements for the competence, consistency, and impartiality of bodies providing audit and certification of all types of management systems".[38] This point will be further detailed, including the identification of any specific capabilities and competencies a certifying body should have to conduct a meaningful assessment of FRT, when the certification scheme will be validated.

3.2 Certification process

Definition of the scope
The certification scheme is exclusively designed for flow management applications of FRT, operated by public- or private-sector organizations. As such, the audit framework explicitly states what aspects of the management of their facial recognition systems fall within the certification and those that are excluded.

Who should be certified?
Any organization that uses facial recognition for flow management is eligible for the certification scheme. It can apply either at the design stage when it starts building its system and is giving thought to the best way to manage it responsibly or once its system is in operation and it wants to improve the quality of its management. Either way, what is requested and assessed by the certification body is the effective compliance with the requirements of the audit framework.

Who should pay?
As this is a voluntary certification scheme for industry actors and public organizations seeking to establish quality management systems through a certificate, applicant organizations should bear the cost of the certification process.

Certification audit approach
The auditor commissioned by the certification body to perform the audit must assess, at the site, the effective implementation of the expected processes and compliance with the requirements of the audit framework in collaboration with the different departments involved in the certification process. This process involves interviews and discussions with employees. Also, evidence of compliance with the audit framework must be made available by the applicant organization for review by the auditors. In advance of any audit, the auditor should provide the applicant organization with clear guidelines as to which employees/functions are likely to be the subject of interviews and what evidence of compliance is expected to be collected and assessed.

Once the audit takes place, the auditors establish a series of observations, which AFNOR Certification advises to classify into five categories:

- **Major non-conformance**: Non-fulfilment of a requirement, calling into question the operation, efficiency or improvement of the facial recognition management system
  Major non-compliance must be the subject of corrective action and must be addressed before certification can be issued.

- **Minor non-conformance**: Failure to meet a specified requirement that does not in itself compromise the effectiveness or improvement of the facial recognition management system
  Minor non-compliance should be the subject of corrective action but does not by itself prevent the issuance of certification.

- **Sensitive point**: A latent risk of non-compliance
  Evidence of compliance with the requirements of the certification framework has been obtained, but the organization must modify its practices to eliminate this latent risk.

- **Strength**: Practice that exceeds the usual level of performance observed in response to the certification requirements

- **Note**: Observation about the compliance with the requirements of the audit framework

Once the audit is completed, a report that includes the auditor's findings is sent to the applicant organization. The applicant organization can then respond to any non-compliance issue identified by providing complementary documentation and the action plan it intends to implement.

Decision and issuance of the certificate
Based on the audit report and the auditor's recommendations, the certification body takes the decision to issue the certification and/or to require additional verifications (i.e. remote or on-site audit, etc.). The certificate is then issued for a year, subject to the completion of any corrective actions decided during the audit. The certificate explicitly states the aspects of the management of their facial recognition systems that fall within the scope of the certification. Once certified, organizations audited by AFNOR Certification are listed on its website.

4

From principles to certification: A journey to build accountability

A timeline for success, in which an external audit is one component of many key steps

This certification aims to provide a practical and workable tool for the continuous monitoring and improvement of the management of facial recognition systems for flow management applications. As such, the issuance of the certification for successful organizations represents a key milestone in their journey. Therefore, two phases can be distinguished:

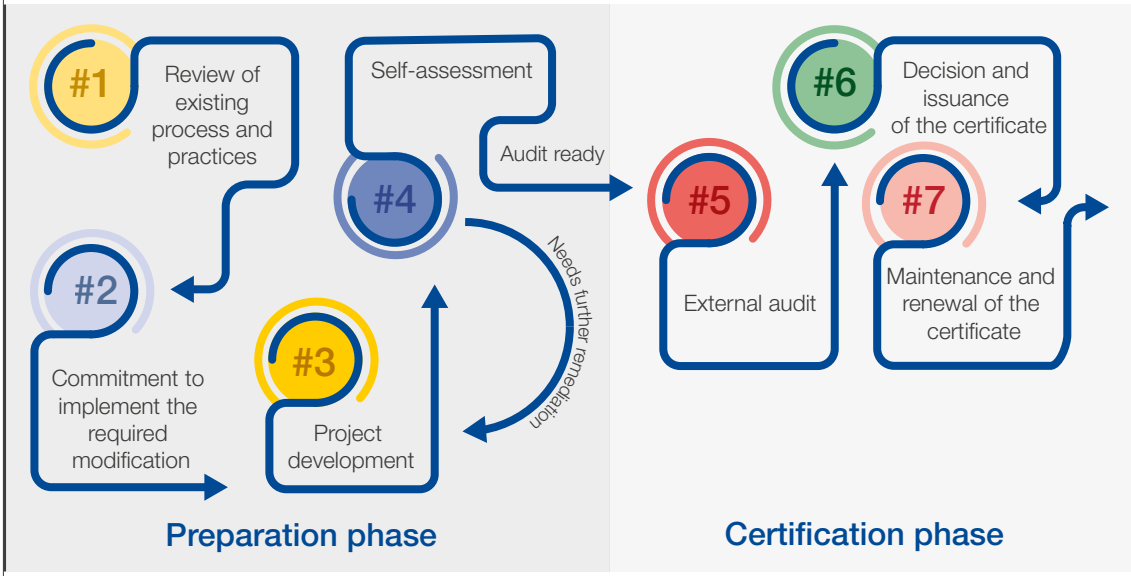
1. **The preparation phase.** Organizations that consider applying for the certification scheme should review the principles for action, implement the best practices and self-assess their processes using the assessment questionnaire, and take inventory of the subjects who may have to be interviewed and the evidence that will need to be collected and assessed as part of the certification process. These actions will create the conditions for a successful and minimally burdensome external audit.

2. **The certification phase.** Accredited bodies formally evaluate the processes implemented by candidate organizations against the requirements of the audit framework. The results of this evaluation determine the issuance of the certification.

To illustrate how such a journey would be completed, the steps that need to be taken and their associated activities are laid out in Figures 2 and 3. The **two possible scenarios** are: A) organizations already offer a facial recognition system and wish to get certified, or B) organizations intend to develop a facial recognition system.

4.1 An organization already offers a facial recognition system and wishes to get certified

FIGURE 2 **7 steps for the certification scheme, when the system is already in operation**



Source: World Economic Forum

– **#1 Review of existing processes and practices.** Any applicant organization should start with a thorough review of the management process of its facial recognition system. Then, it should assess the existing processes and practices against the elements of the governance framework (principles for action, best practices, assessment questionnaire or audit framework). This step allows organizations to validate their approach and check if meeting the requirements of the audit framework is feasible at this point.

– *Activities to run*:

  – A cross-departmental task force should be established.

  – This task force should run an inventory of existing processes and practices to assess if they are consistent with the elements of the governance framework.

  – The task force should take inventory of any data/evidence that must be generated, collected and assessed as part of the certification process and identify any gaps in the existing documentation or data archives.

- The task force should report the identified gaps to management, whether with regard to processes and practices or with regard to data and evidence used to demonstrate compliance.

- **#2 Commitment to implement the required modifications**. The management of the applicant organization validates internally the decision to go through the certification process and commits to allocating the required resources to address the identified gaps:
  - *Activities to run*:
    - The management of the applicant organization makes an internal commitment to adapt its facial recognition systems to comply with the requirements of the audit framework.
    - The management identifies the key stakeholders both internally (e.g. managing board directors, heads of units, members of the cross-departmental task force) and externally (e.g. national data protection authority).

- **#3 Project development.** The applicant organization implements the required modification to prepare the self-assessment process.
  - *Activities to run*:
    - The applicant organization uses the governance framework as a blueprint to modify its existing processes and practices.

- **#4 Self-assessment.** Once the gaps have been identified and filled, the applicant organization can perform a self-assessment to measure its readiness for the certification process. This self-assessment should be run by a team that was not involved in the modification of the problematic processes and practices to avoid self-validation (note that the results of the self-assessment may trigger further modifications). In practice, the applicant organization may need to go through steps 2 and 3 again to fine-tune the existing processes and practices.
  - *Activities to run*:
    - The applicant organization names a dedicated team to run a self-assessment or the internal audit.
    - The organization uses the assessment questionnaire and/or the audit framework to run the self-assessment.
    - The organization ascertains that all required evidence (data and documentation) to be used in support of compliance is up to date and can readily be retrieved.
    - The organization notifies employees who may be the subject of interviews that they may be interviewed as part of the audit (and make them aware of the importance of the audit and their candid participation in interviews as part of it).
    - The organization can compare itself with other organizations, such as Narita International Airport, that have conducted a self-assessment. (Organizations will be encouraged to publicly communicate the results of their assessment questionnaire to demonstrate how responsible use of FRT can be achieved.)
  - *Indicative timeline:*
    - The self-assessment should not exceed 1-2 days and should follow the conditions of the external audit. The time length to analyse potential gaps from the baseline and the implementation of corrective actions will depend on the findings of the self-assessment.

- **#5 External audit.** The external audit is run by the certification body. It takes place in two steps. First, the certification body reviews the processes that the applicant organization put in place to comply with the requirements of the audit framework by looking at how the system has been designed and implemented. Second, it assesses the effectiveness of those processes by auditing the system in operation at the site, at a date unknown to the applicant organization.
  - *Activities to run*:
    - The certification body will perform a documentary audit (review of the documentation related to the design and implementation of the system).
    - The certification body will conduct an audit of field activities (review of the relevant operational procedures, documentation and data).
  - *Indicative timeline:*
    - The audit should take two days. The timeline will depend on the volume of end users using the FRT service and thus may be longer depending on the project. In case of duplication of the system on several physical sites, a sampling audit methodology is put in place.

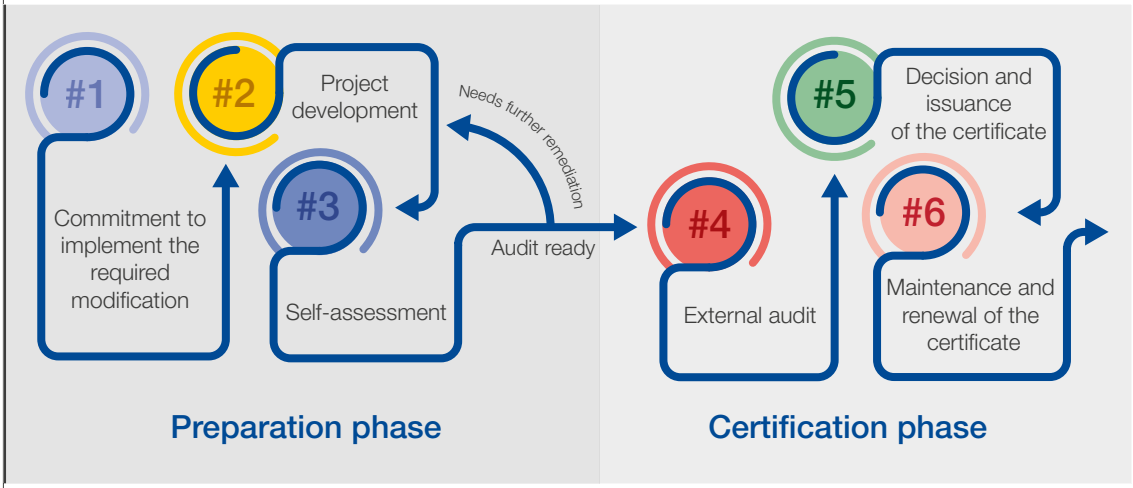
- **#6 Decision and issuance of the certificate.** Based on the results of the audit and the auditor's recommendations, the certification

body makes its decision. It can either issue the certification if all requirements are met or request additional corrective actions. In the second case, applicant organizations will have the time to implement corrective measures before the certification body makes its final decision.

- The certificate is valid for a period of three years and is subject to review based on an annual follow-up audit.
  - *Activities to run*:
    - The certification body will draft the audit report and publish its decision.
    - The certification body will issue the certificate if the requirements are met.
- **#7 Maintenance and renewal of the certificate.** An annual audit validates the maintenance of the requirements. In the event of non-compliance, the certification will be withdrawn.
  - *Activities to run*:
    - Annual audits comprise:
      - A documentary audit (review of the documentation related to the design and implementation of the system)
      - An audit of field activities (review of the operational procedures).
  - *Indicative timeline:*
    - The maintenance audit (monitoring) should take one day per year (the duration will also depend on the sampling rules applied initially).
    - At the end of three years, a certificate renewal audit is conducted. The renewal audit should take two days. The timeline will depend on the volume of end users using the FRT service and thus may be longer depending on the project.

4.2 An organization intends to develop a facial recognition system

FIGURE 3 **6 steps for the certification scheme, when the system is not yet in operation**



Source: World Economic Forum

- **#1 Commitment to comply with the governance framework.** The audit focuses on: 1) requirements related to the processes introduced in the design of a facial recognition system; 2) requirements related to the implementation of these processes while the system is in operation; and 3) requirements related to the functioning of the system. Therefore, organizations that review the policy framework while designing their facial recognition systems will have an advantage over those that manage a system already in operation. Indeed, they have the opportunity to build a system that is "responsible by design".
  - *Activities to run*:
    - The applicant organization uses the governance framework as a blueprint to design its facial recognition management system.

- The management of the applicant organization makes a commitment to design and implement a responsible facial recognition system in compliance with the elements of the governance framework (including ensuring that documentation and data will be generated and retained in a readily accessible location for use in an audit).
- The management will engage with a certification body accredited to issue this certificate and will establish a timeline for the audit.

- **#2 Project development.** The applicant organization prepares the self-assessment process.
  - *Activities to run*:
    - The management of the applicant organization makes an internal commitment to build its facial recognition system according to the requirements of the audit framework.
    - The management identifies the key stakeholders both internally (e.g. managing board directors, heads of units, members of the cross-departmental task force) and externally (e.g. national data protection authority).
    - The management includes the elements of the governance framework and requirements of the audit framework into the set of specifications for the self-assessment.

- **#3 Self-assessment.** Once the different task forces have identified and filled the identified gaps, the organization can perform a self-assessment to measure its readiness for the certification process. This self-assessment should be run by a team that was not involved in the gap analysis to avoid any self-validation (note that the results of the self-assessment may trigger further corrections). In practice, the appropriate stakeholders may need to fine-tune the existing processes and practices accordingly.
  - *Activities to run*:
    - The organization names a dedicated team to run a self-assessment or the audit (requirement 2.4 of the audit framework).
    - The organization uses the assessment questionnaire or the audit framework to run the self-assessment.
    - The organization can compare itself with other organizations, such as Narita International Airport, that have conducted a self-assessment.
  - *Indicative timeline:*
    - The self-assessment should not exceed 1-2 days and should follow the conditions of the external audit. The time length to analyse potential gaps from the baseline and the implementation of corrective actions will depend on the findings of the self-assessment.

- **#4 External audit.** The external audit is run by the certification body. It takes place in two steps. First, the certification body reviews the processes that the applicant organization has put in place to comply with the requirements of the audit framework by looking at how the system has been designed and implemented. Second, it assesses the effectiveness of those processes by auditing the system in operation at the site, at a date unknown to the applicant organization.
  - *Activities to run*:
    - The certification body will perform a documentary audit (review of the documentation related to the design and implementation of the system).
    - The certification body will conduct an audit of field activities (review of the operational procedures).
  - *Indicative timeline:*
    - The audit should take two days. The timeline will depend on the volume of end users using the FRT service and thus may be longer depending on the project. In case of duplication of the system on several physical sites, a sampling audit methodology is put in place.

- **#5 Decision and issuance of the certificate.** Based on the results of the audit and the auditor's recommendations, the certification body makes its decision. It can either issue the certification if all requirements are met or request additional corrective actions. In the second case, applicant organizations will have the time to implement corrective measures before the certification body makes its final decision.
  - *Activities to run*:
    - The certification body will draft the audit report and publish its decision.
    - The certification body will issue the certificate if the requirements are met.

- **#6 Maintenance and renewal of the certificate.** An annual audit validates the maintenance of the requirements. In the event of non-compliance, the certification will be withdrawn.
  - *Activities to run*:
    - Annual audits comprise:
      - A documentary audit (review of the documentation related to the design and implementation of the system)
      - An audit of field activities (review of the operational procedures).
  - *Indicative timeline:*
    - The maintenance audit (monitoring) should take one day per year (the duration will also depend on the sampling rules applied initially).
    - At the end of three years, a certificate renewal audit is conducted. The renewal audit should take two days. The timeline will depend on the volume of end users using the FRT service and thus may be longer depending on the project.

4.3 Consequences of major non-compliance

As the certification scheme is voluntary, in the event of minor non-compliance, companies have one month to make the required corrections. If the auditors identify major non-compliance, the certification is withdrawn. The main consequence is that companies must immediately stop communicating information about the certification and may be requested to publicly inform its consumers that it no longer holds the certification.

If policy-makers pass legislation that transforms certification into statutory law, major non-compliance would be treated in a much more consequential manner:

- The law can impose the immediate shutdown of the facial recognition system until the identified issues are fully addressed. The certification body suspends the certificate during this period for later reactivation if applicable.
- The law can authorize the use of the problematic facial recognition system for a defined period during which the company must address the issues identified. If the company resolves the issues, the certification body does not suspend the certificate. However, if the company fails, the system is shut down and the certificate is withdrawn.

The legislation must specify both the type of major non-compliance issues that require an immediate shutdown of the facial recognition system and the process that the company must follow to resume operations once they have been addressed.

5 Conclusion

A certification scheme is the appropriate regulatory response for flow management use cases.

If successful, this policy pilot will pave the way for the design of a standard for the responsible application of facial recognition systems.

Civil society organizations around the world are increasingly aware of both the opportunities and risks associated with FRT and are urging elected officials to act as the technology's influence over society rapidly increases. Certain policy-makers in the United States and European Union have heard this call and acknowledge the pressing need to create a robust governance framework. Yet consensus on the path forward is lacking.

This White Paper argues that a certification scheme is the appropriate regulatory response for flow management use cases. Entrusting a certification body like AFNOR Certification to assess compliance with the principles for action is an agile and robust means to ensure the trustworthy design and use of FRT for flow management applications.

Organizations willing to comply with this approach can undertake a rigorous multi-step process that starts with a governance framework review (principles for action, best practices, assessment questionnaire and audit framework), which can be used as a guide to design or improve an existing facial recognition system. The journey, however, does not end with the issuance of the certificate. Indeed, being trustworthy is an iterative and continuous assessment effort. Therefore, the main partners of this Responsible Limits on Facial Recognition project hope that, in the long run, certified organizations will develop an organizational culture that contributes to the identification and mitigation of ever-evolving risks, for the benefit of technology users, customers and society at large.

Governments have a key role to play in fostering this culture. Once the policy pilot is completed and has proven to be successful, policy-makers will be encouraged to take into consideration the proposals in this White Paper and pass legislation that makes this certification mandatory for industry actors that use FRT for flow management applications.

The next steps of this policy pilot are to test the audit framework and certification scheme with industry actors, assess their relevance and the amount of work they create for actors seeking certification, and review them based on the observed results. If successful, this policy pilot will pave the way for the design of a standard for the responsible application of facial recognition systems. Once the pilot project is completed, a multistakeholder coalition of actors committed to respecting and promoting this certification model will be formed.

Industry players, public actors, civil society representatives, certification bodies, policy-makers and academics are encouraged to join this journey and participate in an open and experimental approach to strengthen this certification model and ensure its impact.

This project's use-case-based approach could serve as a blueprint for stakeholders seeking to ensure the responsible use of FRT in other applications. As such, it carries important insights. Therefore, organizations interested in deploying this method in other use cases are invited to contact the Centre for the Fourth Industrial Revolution of the World Economic Forum.

Glossary

Accuracy of facial recognition: The accuracy of a facial recognition system is based on a combination of two conditions: 1) how often the system correctly identifies a person who is enrolled in the system; and 2) how often the system correctly finds no match for a person who is not enrolled. These two conditions, which are referred to as the "true" conditions, combine with two "false" conditions to describe all possible outcomes of a facial recognition system (see the definitions of true positive, true negative, false positive and false negative).

Algorithm: An algorithm is a series of instructions for performing a calculation or solving a problem, especially with a computer. Algorithms form the basis for everything a computer can do and are therefore a fundamental aspect of all AI systems.

Audit: The basic function of an audit framework is to serve as a reference document that details the requirements and processes of an audit for a defined scope.

Biometrics: Biometrics covers 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.

Certification: The basic function of a certification scheme is to perform an independent assessment to reach an objective judgement of a given system or product based on a defined set of requirements listed within an audit framework.

Computer vision: Computer vision is 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.

Enrolment: Enrolment is the process of enrolling images of individuals for template creation so they can be recognized. When a person is enrolled in a verification system used for authentication, their template is also associated with a primary identifier that will be used to determine which template to compare with the probe template.

Explainability: Explainability is a property of AI systems that can provide a form of explanation for how conclusions are reached to improve decision understanding and increase trust from operators and users of the systems.

Face detection: Detection finds human faces and answers the question, "Are there one or more human faces in this image?"

Face identification (or one-to-many): Face identification answers the question, "Can this unknown person be matched to an enrolled template?" Identification compares a probe template to all enrolment templates stored in a repository, so it is also called "one-to-many" matching. Candidate matches are returned based on how closely the probe template matches each of the enrolled templates.

Face verification (or one-to-one): Face verification answers 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 a customer ID), and facial recognition is used as a second factor to verify the person's identity. Verification is also called "one-to-one" matching because the probe template (one person) is compared only to the template stored for the (one) person associated with the identification presented.

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

False negative: A false negative is a test result that incorrectly indicates that the person in the probe image is not enrolled and they are not matched when they have been enrolled. Depending on the use case of facial recognition, the consequences of false negatives can vary greatly.

False positive: A false positive is a test result that incorrectly indicates that the person in the probe photo is enrolled in the system when they have not been enrolled. Depending on the use case of facial recognition, the consequences of false positives can vary greatly.

Probe image: A probe image is an image submitted to a facial recognition system to be compared to enrolled individuals. Probe images are also converted to probe templates. As with enrolment templates, high-quality images result in high-quality templates.

Template: Images of people are converted into templates, which are then used for facial recognition. Machine-interpretable features are extracted from one or more images of an individual to create that individual's template.

True negative: In a true negative, the person in the probe image is not enrolled and is not matched.

True positive: In a true positive, the person in the probe image is enrolled and is correctly matched.

Contributors

Lead authors

Ichirou Akimoto
Senior Manager, Global AI Business Development, NEC Corporation, Japan; Corporate Fellow, World Economic Forum

Yusuke Inoue
Japanese Government Fellow, World Economic Forum, Centre for the Fourth Industrial Revolution Japan

Sebastien Louradour
French Government Fellow, Artificial Intelligence and Machine Learning, World Economic Forum LLC

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

Jérémie Mella
Project Manager, AFNOR Certification, France

Project community

The World Economic Forum thanks the project community members for their insightful review and feedback:

Didier Baichère
Member of the National Assembly, France

Xavier Blondeau
Manager, Video Protection, SNCF, France

Vincent Bouatou
Director, Innovation Lab, IDEMIA, France

Pascal Briand
IT Manager, Passenger Experience Division, Groupe ADP, France

Laurent Dahmani
Deputy General Manager, AFNOR Certification, France

Jean-Luc Dugelay
Professor of Digital Security/Imaging Security, EURECOM, France

Marine Dunoguier
Director, Embedded Systems, Alcatraz, USA

Rosanna Fanni
AI Ethics Researcher, Belgium

Louis-Thomas Fernandes
LAF Expert, SNCF, France

Romain Galesne-Fontaine
Director, Communications and Institutional Relations, IN Groupe, France

Hervé Genty
Head, Data Retail Safety, SNCF, France

Brice Gilbert
Project Manager, Cybersecurity, AFNOR Certification, France

Ilana Golbin
Director, Emerging Technologies and Responsible AI, PwC, USA

Bruce Hedin
Principal Scientist, H5, USA

Matissa Hollister
Assistant Professor of Organizational Behaviour, McGill University, Canada; Academic Fellow, World Economic Forum

Luc Julia
Senior Vice-President and Chief Technical Officer, Samsung Strategy and Innovation Center, Samsung Electronics, Republic of Korea

Eva Kaili
Member of the European Parliament

Takayuki Kitagawa
Manager, Corporate Strategies Office, Narita International Airport Corporation, Japan

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

Gautier Martin
Project Manager, Airport Services and Product Development, Groupe ADP, France

Hidehisa Matsumoto
Senior Manager, Corporate Strategies Office, Narita International Airport Corporation, Japan

Franck Maurin
Product and Solutions Director, Passenger Facilitation and Border Control, IDEMIA, France

Cédric Mazière
Manager, Video Protection, SNCF, France

Michaël Mesure
Director, LAF, SNCF, France

Shaun Moore
Chief Executive Officer and Co-Founder, Trueface, USA

Anand Rao
Global Lead, Artificial Intelligence, PwC, USA

Arthur Ribemont
Project Manager, Privacy, AFNOR Certification, France

Mathieu Rondel
Director, Expertise and Operational Performance, Airport Operations Division, Groupe ADP, France

Stéphane Séjourné
Member of the European Parliament

Karen Silverman
Founder and Chief Executive Officer, Cantellus Group, USA

Emilia Tantar
Chief Data and Artificial Intelligence Officer, Black Swan LUX, Luxembourg

Isabelle Valverde
Head, Flow Operations, SNCF, France

Philippe Weiss
Head, Facial Recognition Project, SNCF, France

Tomohiro Yamane
Director, Aviation Innovation Policy Planning and Research, General Affairs Division, Civil Aviation Bureau, Ministry of Land, Infrastructure, Transport and Tourism, Japan

Independent observers

Thanks also go to the French Data Protection Authority (CNIL) for its role as independent observer, and specifically to:

Marie Duboys Fresney
Legal Counsel

Félicien Vallet
Privacy Technologist

Appendices

Appendix A: Answers from Tokyo-Narita International Airport to the assessment questionnaire

1. Proportional use of facial recognition systems

Assessment questions	Narita International Airport self-assessment responses
What are the alternatives to your facial recognition system? Why have you rejected them? What are the criteria used to determine the advantages and disadvantages of these alternatives?	Fingerprint recognition and iris recognition are considered as alternatives. As a result of the comparison, it was determined that facial recognition was superior in the following areas: 1. Convenient with no need for passengers to operate equipment 2. Its contact-free use turned out to be useful during the COVID-19 pandemic 3. A unique piece of information that no one else has and that can be collected without special manipulation 4. The concept of walking pace can be realized
How did you assess the appropriateness of your system with regard to its purpose?	It was evaluated based on the following three points: 1. High facial recognition accuracy 2. High end user (airline) requirement satisfaction 3. Processing performance such as walking pace
Describe the technical requirements for addressing the objectives of your system in a format understandable by the appropriate authorities.	The main points are as follows: 1. Facial recognition accuracy 2. Walking pace speed 3. Compatibility with existing airline systems in deployment (no need for massive modification) 4. Compliance with IATA standards
Have you carried out a risk analysis of the false positive and false negative situations (in particular, the risks of violating civil rights)?	We analysed the risk of misboarding and other factors with a third-party committee (Personal Information Protection Committee, MLIT: including lawyers, university professors and consumer groups).

2. Risk assessment

Assessment questions	Narita International Airport self-assessment responses
Have you rigorously assessed the risks related to the use of your system before (e.g. risk assessment framework) and during its operational functioning (e.g. audit framework) through the following dimensions?	
Privacy	This risk was evaluated by a third-party committee (Personal Information Protection Committee, MLIT: including lawyers, university professors and consumer groups).
Errors	We implemented system error monitoring and detection functions to enable system users (airlines, etc.) and system owners (Narita International Airport) to determine the errors.
Unfair bias	This risk was evaluated by a third-party committee (Personal Information Protection Committee, MLIT: including lawyers, university professors and consumer groups).
Hacking and cyberattacks	We conducted Security by Design, a system evaluation by a third party, and carried out penetration tests, etc.
Transparency in the decision-making process	The approval process was taken by the approving authority based on the contract. We also held a system specification coordination meeting with system users (airlines, etc.) to determine the specifications.
Human and civil rights infringement	This risk was evaluated by a third-party committee (Personal Information Protection Committee, MLIT: including lawyers, university professors and consumer groups).

3. Bias and discrimination

Assessment questions	Narita International Airport self-assessment responses
What are your definitions of unfair bias in your use case? Describe the metrics used to evaluate each of them.	We adopted a facial recognition system because the passport embedded integrated circuit (IC) contains facial information and has been used in other airports. We defined unfair bias as differences in performance (metrics: accuracy and certification time) among people of different races and with disabilities, such as wheelchair users. However, those whose passports do not have IC or those who cannot receive support due to country-specific reasons are not eligible.
What is your risk analysis framework? Describe the risks of unfair bias identified for your use case and the groups described by end-user characteristics for which you evaluated bias.	We used JIS Q 31000 based on ISO 31000 as a risk analysis framework for the entire system. As for unfair bias evaluation, we required vendors to submit the results of the public review by NIST, etc., as a verification item of the facial recognition system.
How are risks prioritized in this process? How are competing interests resolved?	The availability and security of the entire system were given top priority. Individual functions were reviewed based on the impact on the entire system.
Please describe the existing best practices for detection, identification and mitigation of unfair biases that were applied in this case.	We reviewed facial recognition vendors in NISTIR 8280 and selected a vendor with the best algorithms at the time of the choice.

Assessment questions	Narita International Airport self-assessment responses
What action plans have you put in place to mitigate the main risks identified? For each risk of unfair bias, what mitigation was identified and how were mitigations evaluated to ensure effectiveness?	As for unfair bias evaluation, we required vendors to submit the results of the public review by NIST, etc., as a verification item of the facial recognition system. Since NEC's system was the most accurate at the time of the choice, we decided to combine it with the operations.
What are the test cases and acceptance tests used for your facial recognition system?	We conducted acceptance tests as a system (including operational tests taking into account the local light environment) from a use case perspective. We required vendors to submit the results of the public review by NIST, etc., as a verification item of the facial recognition system itself.
What is the distribution of your training set and how well does it align with that of the end users of your system? If there are gaps, how did you evaluate the impact of the gaps and remediate them?	We designed use cases, tested (system acceptance) and created training scenarios with end-user airlines. We required vendors to submit the results of the public review by NIST, etc., as a verification item of the facial recognition system itself.
What kind of trade-offs are you facing in the deployment of your system? How do you address them?	We are facing trade-offs between the protection of personal information and usability, etc. Personal information protection is managed through a multistakeholder process.
If there were any gaps between the release criteria and actual performance, how were the gaps mitigated?	To reduce the accuracy deterioration of facial recognition due to sunlight, we took additional physical measures, such as shading films and curtains, and adjusted the software, such as the facial recognition parameter, to comply with the requirements.

4. Privacy by design

Assessment questions	Narita International Airport self-assessment responses
What processes (e.g. a task force) and resources (e.g. a charter of best practices) have you implemented to support the privacy of data subjects, for example, to avoid the over-collection of biometric data in relation to the purposes of use?	By working with the company's personal information protection office and legal counsel, we identified the legal issues and organized the direction of the solution and held discussions with the Civil Aviation Bureau and the Personal Information Protection Committee. Subsequently, the Civil Aviation Bureau established an expert panel (third-party committee) and introduced a multistakeholder process to evaluate the process from the viewpoint of the protection of personal information. A guidebook was prepared on the premise of domestic expansion of the One ID programme, while avoiding criticism.
Have you established a data protection officer position?	As for the system for the protection of personal information, internal regulations had already been established prior to the start of this project, and a responsible officer has been assigned in accordance with these regulations.
How do you promote a close collaboration for the product development phase of your facial recognition system, including product managers, a legal team, UX designers, data scientists and developers, to ensure a high level of data protection?	We conducted Security by Design, a system evaluation by a third party, and carried out penetration tests, etc.

5. Performance

Assessment questions	Narita International Airport self-assessment responses
For the lab and field tests, what existing standards (e.g. International Organization for Standardization [ISO], AFNOR Certification and European Committee for Standardization [CEN]) are you following to evaluate the accuracy and performance of your system? What criteria were used to choose the standards and norms that you follow?	The latest version of NEC products that complies with the biometrics standards of ISO/IEC JTC 1/SC 37 is used for the facial recognition system alone.
Have you submitted your facial recognition system to the National Institute of Standards and Technology (NIST) for evaluation?	We used the NIST evaluation (IR 8271) as a reference.
What process have you established to ensure the auditability of the performance results of your facial recognition system? What steps have been taken to allow a sufficient audit by a third party?	Based on performance requirements for walking pace, system tests (functional, non-functional tests: availability, reliability, performance and biometric tests) were conducted and verified.
What is the relevance of the performance tests conducted concerning the use case that has been considered?	Relevant We conducted an implementation test at the boarding gate, which required higher processing capacity than the conventional 250-passenger, two lanes for 15 minutes. We verified labour savings and reduced ground staff compared to the current operation.
How do you justify the chosen performance threshold that induces a theoretical rate of false positives and a measured rate of false negatives?	We comply with the specifications of facial recognition gates required by the Immigration Services Agency of the Ministry of Justice.

6. Right to information

Assessment questions	Narita International Airport self-assessment responses
What processes have been implemented to keep end users informed about the use of your system and their biometric data? Also, what processes, including the means for escalation and remedy, have been implemented when the system is believed to have caused harm? Best practices include but are not limited to providing for customer support and enquiries.	We comply with the “Guidebook on the handling of personal data in One ID services that utilize facial recognition technology at airports” by the Ministry of Land, Infrastructure, Transport and Tourism and we present it to users. – Email address: same as above – Phone number: same as above – Customer support FAQ: same as above – Customer support chatbot: not planned.

Assessment questions	Narita International Airport self-assessment responses
Could a data subject access, retrieve or ask to delete personal data (photo, video, biometric data linked to a person's identity, such as account event history, consent history, biometric data deletion history, shared information, history of use of biometric data) in a machine-readable format within a reasonable period (e.g. no more than 30 days)?	The data is deleted within 24 hours.
Have you established and publicly disclosed (e.g. on your website) the governance principles that guide the design and use of your system in a format that is intelligible to non-experts?	They will be published on the website, in leaflets, etc.
Have you established any process that enables individuals to access relevant information about the functioning of the system anonymously?	None.

7. Consent

Assessment questions	Narita International Airport self-assessment responses
Does the consent policy provide explicit and clear information to users, and more specifically?	It is clearly indicated on the display at the time of registration. The information is also available on the website and on posters placed in the terminal.
Is the consent page accessible at most after two clicks and is it easily visible on the "profile" page?	One-click is all it takes. To show the rules, scroll on the same page. Aborting an operation before the procedure is completed does not generate a token. The profile page is not displayed.
Is a summary of the main provisions available on this same page?	Yes.
Does this summary contain the following information?	
a description of all intended purposes	Yes.
the data retention period	Yes.
the data-sharing policy (including with which third parties this data will be shared)	Yes.
the means put in place to protect, secure and store data.	Yes.
Is this summary concise, comprehensible to non-experts and less than the equivalent of two A4-sized pages in length?	Yes.

Assessment questions	Narita International Airport self-assessment responses
Does the page for giving or not giving consent allow users to indicate it for each of the existing purposes?	No token is generated unless the user agrees to everything.
Are all these options available on the same page?	Yes.
Is the list of existing purposes up to date?	It is up to date.

8. Information display

Assessment questions	Narita International Airport self-assessment responses
What means have been put in place to inform individuals that they are entering an area in which a facial recognition system is being used? Are these means visible and explicit enough for the public? Is a user rights reminder display in place?	The facial recognition area and general area are clearly distinguished by the use of a special logo indicating facial recognition, etc. We comply with the "Guidebook on the handling of personal data in One ID services that utilize facial recognition technology at airports" by the Ministry of Land, Infrastructure, Transport and Tourism and we present it to users: posters, panels, etc., at the airport.
For premises access, flow management and/or enrolment in a public space, does the volume of the recording zone not exceed the capture space defined and identified by the users? How do you ensure that the capture space is understood by end users (please provide evidence based on evaluation/research/testing)?	We verify the facial recognition system at the time of deployment to ensure that it does not exceed the capture space.
Does a display of sufficient size relay the purpose of the facial recognition system? How do you ensure that the display is noticeable and legible (please provide evidence based on evaluation/research/testing)?	A special logo and colour scheme for facial recognition are used to increase visibility and make the display recognizable regardless of its size.

9. Right of access to vulnerable groups

Assessment questions	Narita International Airport self-assessment responses
Can you detail how the system has been designed and evaluated to support elderly people and people with disabilities (including visual and auditory)?	We comply with the "Guidebook on the handling of personal data in One ID services that utilize facial recognition technology at airports" by the Ministry of Land, Infrastructure, Transport and Tourism and we present it to users: posters, panels, etc. at the airport.
Is your facial recognition system accessible to everyone, including elderly people and people with disabilities?	Accessible. However, the subject is set to a height of between 130 cm and 190 cm so cameras can capture images.
What resources have you allocated to support elderly people and people with disabilities?	We assigned staff (airport and airline staff) to each touchpoint.

Mitigation for people with disabilities, children, families and others for whom the system does not work or is undesirable may be to use an alternative option that has been tested to determine that it works.	We substitute with manual operation by staff (airport and airline staff).

10. Alternative option and human presence

Assessment questions	Narita International Airport self-assessment responses
Have you put in place a manual review process for situations in which the matching of a face and an identity document with a photo leads to a false negative, especially during the enrolment phase?	There is no manual review process. Full automation and elimination of false negatives is achieved by adding a cloud service that refers information to the airline host for verification of passenger information, passport information, etc. In the event of an error, a manual process by a person similar to the current operation is to be carried out.
For facial recognition systems, is the alternative option systematically implemented and:	Implemented: we substitute the current manual operation by staff.
operated by human agents? (Are these operators trained to handle exceptional situations?)	We substitute the current manual operation by staff.
reasonable; that is, it does not introduce disproportionately adverse consequences (e.g. doubling the time needed to go through the security check)?	It can be used as an equivalent system to the existing one, and can be handled in the same processing time as before.
Is there an alternative process for people who don't accept the use of their biometrics?	It can be used as an equivalent system to the existing one. Manual operations are prepared as an alternative process.

Appendix B: Audit framework

1. Proportional use of facial recognition systems

Requirement

Facial recognition systems should be highly tailored according to the intended use. Organizations using facial recognition systems should take reasonable steps to assess the capabilities and limitations of the systems they intend to use and ensure that their systems are appropriate for the intended purpose.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
1.1	Prior to any facial recognition project, the need leading to considering the use of a facial recognition system must be defined. Companies must describe the technical requirements to achieve the objectives assigned to their system and be able to guarantee that the system will only be used for its intended purpose.	✓		
1.2	The set of alternatives (excluding facial recognition) that fulfil the same need must be determined.	✓		
1.3	To fulfil the need, possible alternatives to the use of a facial recognition system must be identified. A documented process and methodology for analysing possible solutions must be set up. The objective is to assess the use of facial recognition technology's relevance to its purpose and the resolution of the problem. To this end, companies must describe in detail the assessment and selection methodology, which must at least include: – A review of the identified advantages and disadvantages for every identified solution – A definition of the system's benefits for the various stakeholders (users, state, citizens, etc.) – A risk analysis covering false positive and false negative situations (in particular, the risks of violating civil rights) – A quantified assessment of the expected benefits – A comparative analysis of the different solutions – The conclusion that led to the preference for a facial recognition solution	✓		
1.4	To validate the assumptions that led to the choice of facial recognition technology, companies must define the parameters to be respected to validate the relevance of its use (for example: expected false positive and false negative rates, expected performance).	✓		

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
1.5	These parameters must be checked in the use phase.		✓	
1.6	The facial recognition system was introduced to meet specific needs in the framework of particular uses. When used, the facial recognition system must be limited to the initially planned uses and validated for those uses.			✓

2. Risk assessment

Requirement

Organizations creating facial recognition platforms or using facial recognition as part of a service or system should conduct a comprehensive risk assessment of their systems, including the impact on privacy, potential for errors, susceptibility to unfair bias, vulnerability to hacking and cyberattacks, lack of transparency in the decision-making process and potential for civil and human rights infringements.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
2.1	A full risk assessment of the facial recognition system must be conducted. The analysis should take into account the following items: – Impact on privacy – Error potential – Susceptibility to bias – Vulnerability to cyberattacks (hacking, ransomware, etc.) – Lack of transparency in the documented decision-making process – Potential violation of civil rights The analysis should also rank the solutions implemented to mitigate the risks. (See an example of tools to conduct the risk assessment in the appendix.)	✓		
2.2	The risk analysis must include the implementation of a risk processing plan. The analysis should also rank the risks and the solutions implemented to mitigate them.	✓		
2.3	The actions resulting from the risk analysis and processing plan must be implemented and maintained. Indicators to assess their effectiveness and their maintenance in operational condition must be set up.		✓	✓

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
2.4	FRT users must set up systems making sure they deploy their projects in accordance with the principles for action. To do this, they must carry out one of the following actions: – A self-assessment based on the assessment questionnaire – An internal audit based on this standard. The resulting conclusions and deliverables must be accessible. In particular, they validate the deployment of the system by including the recommendations. Third-party auditors must have access to the methodology used. This approach must also be applied when the system is running to make sure the principles for action are complied with.	✓		

3. Bias and discrimination

Requirement

Organizations using facial recognition systems should take appropriate steps to ensure that all unfair bias or outcomes can be detected, identified and mitigated to the greatest extent possible. While acknowledging that the complete removal of bias represents one of the biggest challenges in AI research, organizations must assign appropriate resources to the implementation of tools and processes that minimize bias or unfair outcomes.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
3.1	A definition of bias within the scope of your facial recognition use must be provided. In particular, a review of biases should be carried out.	✓		
3.2	A description of the best practices that have been applied to your use case to detect, identify and mitigate bias must be provided.	✓		

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
3.3	Specifications must be drawn up for suppliers. The specifications should be drawn up on the basis of a documented risk assessment in order to take appropriate measures to guarantee that all risks (including biases) or unfair outcomes can be detected, identified and mitigated as far as possible. The assessment must at least include the following items, which are set out in the appendix: – Description of the identified risks of bias for your use case and the characteristics of the end-user groups that could be subject to these risks of bias – Definition of the system's end-user characteristics (for example, taking into account age groups, gender, ethnicity), grouped together by prioritizing those groups that require special attention because of the risks of bias they may be subject to – Consideration of the accessories and elements that may impact the algorithm: sunglasses, hats, beards, masks, etc.; the specifications must also take these situations into account – Implementation of parameters to assess each bias identified during the various stages of the use process; these parameters will, in particular, make it possible to rank the risks of bias – Analysis of each step in the use process (for example, by observing biases based on image capture and biases based on model performance) to identify and assess the associated risks of bias – A ranking of bias risks and the processing of diverging interests	✓		
3.4	Regarding the risks from using the system, it must be possible to define and document how the identified risks will be mitigated. Processes and resources to guarantee that potentially discriminatory outcomes are detected and mitigated in the best possible way when using the technology (see the example in the appendix) are needed.	✓	✓	
3.5	For each identified risk of discrimination, the assessment of the facial recognition system performance to mitigate this bias must be determined, with details of the parameters used and measurement systems (see the model in appendix B4). The implementation of indicators are necessary to assess and validate the effectiveness of strategies. These assessments should be carried out during the design phase and during the operation of the system in order to validate compliance with the indicators.	✓	✓	
3.6	The implementation of corrective and mitigating actions must take place during system operation when bias drifts compared to the objectives are identified.		✓	✓
3.7	Biometric system tests must be conducted along with the creation of an acceptance document to validate the algorithm.	✓	✓	

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
3.8	The distribution of your training data must be determined and items that are similar/different to those of the system users must be measured. If there are differences, the impacts must be assessed and reduced.	✓		
3.9	The trade-offs for your customers/users (for example, trade-offs between advantages and disadvantages produced by the technology) must be identified and described. A process for resolving arbitrations when diverging interests come to light must be set up.	✓	✓	
3.10	The processes and resources (see 3.4) to guarantee that potentially discriminatory situations are detected and mitigated during system use in the best possible way to reduce impacts on users must be implemented. Regarding the risks involved in the use of the system, actions to mitigate the risks must be implemented.		✓	✓
3.11	Criteria to determine that the system is ready for deployment and use must be defined (for example, system performance, discriminatory situations, etc.). During the use phase, compliance with these criteria must be guaranteed.		✓	

4. Privacy by design

Requirement

Organizations using facial recognition systems should design systems to support privacy, including privacy considerations in system requirements and carrying through privacy support in the design, development and testing of technology as well as in supporting business practices and ongoing system maintenance.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
4.1	Companies must comply with applicable standards and regulations covering the protection of personal data.		✓	
4.2	A documented process and resources must be set up to ensure the confidentiality of biometric data. The process must be deployed and maintained during the use of the system.	✓	✓	
4.3	The training of facial recognition product teams that natively respect privacy (including product managers, the legal team, UX designers, data scientists and developers) must be implemented to provide a high level of data protection.	✓		

5. Performance

Requirement

Organizations creating facial recognition platforms or using facial recognition as part of a service or system should follow the standards for evaluating the accuracy and performance of their systems at the design (lab test) and deployment (field test) stages. Performance assessments should be auditable by competent third-party organizations and their reports made available to users of the systems.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
5.1	Users of the technology must obtain guarantees from their suppliers that the construction of the specific or API-accessible database includes sufficiently equal samples of the subgroups that make up the end-user population and collect data accordingly. To do this, they provide their suppliers with the characteristics of the end users. – Suppliers must determine the criteria that led them to choose their assessment method and the standards that were used to choose the software. These items are part and parcel of the specifications for the selection of the system.	✓		
5.2	Suppliers of the technology must provide the elements that make it possible to validate the performance threshold expectations requested in the user's specifications.	✓		
5.3	It must be possible to demonstrate and validate that the chosen performance threshold (which induces a theoretical false positive rate and a measured false negative rate) is respected in the operation of the system.		✓	✓
5.4	The operational assessments and their reports must be auditable and can be consulted by independent third parties.		✓	
5.5	Processes must be implemented to make sure performance assessments can be audited. Steps must be taken to allow the sufficient auditing of these results by auditors.	✓	✓	

6. Right to information

Requirement

Processes should be put in place to inform end users who have questions and/or need information on the use of facial recognition systems. End users should have access to their personal biometric data upon request.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
6.1	A documented process to keep end users informed of the use of the system and the use of their biometric data must be implemented. The process must be able to include changes in the use of the system to inform users of them.	✓		

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
6.2	The system for informing end users about the use of the system and their biometric data must be durable and take account of system developments and the use of biometric data.		✓	
6.3	Users must have access to information on the use of their biometric data. Information on the use of biometric data must be up to date.			✓
6.4	Documented processes (for example, escalation and resolution procedures) to deal with prejudicial outcomes suffered by users must be implemented.	✓		
6.5	Users must be able to declare a prejudice. Best practices include, but are not limited to, making the following available: – Email address – Phone number – Customer support FAQ – Customer support chatbot			✓
6.6	The traceability and processing of cases of prejudicial outcomes communicated by users must take place.		✓	
6.7	Measures must make it possible for users to access in a legible format, retrieve and ask to delete personal data (photo, video and biometric data linked to a person's identity: account event history, consent history, biometric data deletion history, shared information, history of use of biometric data) within a reasonable time (for example, no more than 30 days).	✓		✓
6.8	Requests for access, recovery and erasure of personal data must be traced and implemented.		✓	
6.9	A process to allow individuals to anonymously access relevant information on the system's operation must be implemented.	✓		✓
6.10	Communication must be made to the general public (for example on the website) about the governance rules that guide the design and use of the system in a form that non-experts can understand.			✓
6.11	Relevant information about the operation of the system must be public and accessible.			✓

7. Consent

Requirement

Individuals should provide informed, free, unambiguous, explicit and affirmative consent for the use of facial recognition systems. Any time data subjects enrol for a new service powered by FRT, they should express clear consent with regard to the length of data retention and the terms of the data storage.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
7.1	The definition of the measures that will be implemented covering consent must be provided to guarantee that users can give informed, explicit and affirmative consent for the use of the facial recognition system.	✓		

7.2	The consent policy must be available online and provide users explicit and clear information, namely:
	– The consent page must be accessible at most after two clicks and must be easily visible on the "profile" page.
	– A summary of the main provisions is accessible on this same page. It must contain the following information:
	  – A description of all intended purposes
	  – The data retention period
	  – The data-sharing policy (including with which third parties this data will be shared)
	  – The means put in place to protect, secure and store data.
	– This summary must be concise, understandable by non-experts and less than the equivalent of two A4-sized pages in length.
				✓

7.3	The consent web page must make it possible to give or withdraw consent for each of the existing purposes. All of these options must be on the same page.			✓
7.4	On each subscription, users must clearly express their consent to the duration of the data retention.			✓
7.5	During a third-party audit, companies must be able to provide the elements needed to demonstrate that each user clearly expressed their consent.		✓	
7.6	Companies must make sure that consent provisions are durable and accessible to users.		✓	
7.7	If the facial recognition service evolves, the list of existing purposes must be kept up to date and explicitly communicated to end users. Provisions must be implemented to include changes.		✓	✓

8. Information display

Requirement

When used in public spaces, clear signage should be deployed to ensure obvious communication with end users on the use of the facial recognition technology. Areas where facial recognition systems are used should always be delimited and indicated to individuals. A visual sign should also inform individuals when the system is in operation.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
8.1	The design of the information system and communication on the use of the facial recognition system must comply with the requirement, taking into account: – Information to users (what is to be communicated and how, etc.) about the use of facial recognition and the area where the facial recognition system is used – A methodology for determining the area where the facial recognition system is to be implemented	✓		
8.2	All information systems must be in place, including: – Clear information for individuals entering an area where the facial recognition system is used. This resource must be sufficiently visible and explicit to individuals – An indicator (visual for example) must inform individuals when the system is in operation – A display listing user rights must be visible – A sufficiently large display used to remind users of the purpose of the facial recognition system must be present			✓
8.3	Measures must be taken to ensure that the capture area is clearly understood by users.		✓	
8.4	The information display must be permanent. Measures to ensure that this is the case must be implemented.		✓	

9. Right of access to vulnerable groups

Requirement

Facial recognition should not exclude anyone and should always be accessible to and usable by all groups of people, including elderly people and people with disabilities. It is recognized that there may be some instances, such as with infants and children, in which an exception to this principle is appropriate and an alternative to facial identification should be offered.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
9.1	A description of how the system was defined and assessed to not exclude anyone, including elderly and/or disabled people (in particular visual and auditory) must be provided. Whether the facial recognition system is accessible to the elderly and disabled must be indicated.	✓		
9.2	The planned facial recognition system perimeter must be effective in operation. An assessment should also be made as to whether the user perimeter is consistent or whether an alternative solution becomes relevant (management of encountered situations).		✓	✓
9.3	The definition of cases where it is accepted that an exception to this principle is appropriate and an alternative to facial recognition should be proposed must be provided. The resources allocated to support the elderly and disabled must be described.	✓		
9.4	A description of the alternative option for infants, children and their families must be provided and implemented.	✓		✓
9.5	Provisions to make sure the system does not exclude anyone must be deployed. Its effectiveness must be assessed and made sustainable (implementation of the alternative solution).		✓	✓

10. Alternative option and human presence

Requirement

A manual review (human overseeing) should be conducted for any use that could result in a consequential decision, such as causing a civil rights infringement. In the case of a fully automated system, a fallback system with a human presence should always be in place to address exceptions and unexpected errors, and for possible remediation purposes. A reasonable alternative to the use of facial recognition systems should always be in place.

Requirement n°	Description of the standard requirement	Process requirements related to the		
		design	implementation	system's functioning
10.1	Situations likely to lead to a decision with consequences such as the violation of civil rights (situations where the matching of a face with an identity document containing a photo results in a false negative, particularly during the enrolment phase) must be identified.	✓		

10.2	A manual review process must be implemented to avoid any situations prejudicial to users.	✓		✓
10.3	This process must be implemented and maintained in the operation of the system.		✓	✓
10.4	An alternative process that identifies who will use it (infants, children and their families, for example) must be implemented. This alternative must also take into account people who do not accept the use of their biometrics.	✓		
10.5	For facial recognition systems, the alternative option must be implemented and be: – Operated by human agents (these operators must be trained to handle exception situations) – Reasonable; namely, it does not result in disproportionate negative consequences (e.g. doubling the time required to pass the security screening)			✓
10.6	To guarantee that the alternative option does not lead to negative consequences, the following are required: – An analysis of the effectiveness of the measures and improvement of the system – Traceability of the manual review rate – Consideration of situations where a decision with consequences, such as the violation of civil rights, has been encountered		✓	

Appendix B1: Definition of risks

Risk n°	Identified risks	Risk description	Cause of the risk

Appendix B2: Risk analysis

Bias n°	Identified risks	End-user group characteristics	Step in the implementation process in which the risk will be encountered	Risk assessment parameters						Risk classification
-	-	-	-	Impact on users		Discrimination		Civil rights		Risk scoring
				Severity	Probability of occurrence	Severity	Probability of occurrence	Severity	Probability of occurrence	Potential severity level
N°1	*To be completed*	*To be completed*	*To be completed*	*To be completed*		*To be completed*		*To be completed*		Choose an item
				Choose an item	Choose an item	Choose an item	Choose an item	Choose an item	Choose an item	Choose an item

Definition of the criteria (indicators) that make it possible to identify the risk level.

Risk level:	Probability of occurrence:
– Very high = *To be defined*	– Very frequent = *To be defined*
– High = *To be defined*	– Frequent = *To be defined*
– Moderate = *To be defined*	– Moderate = *To be defined*
– Low = *To be defined*	– Low = *To be defined*

Appendix B3: Mitigation strategies

Risk n°	Identified risks	Risk mitigation strategy		Indicator to measure the performance of the strategy	Benefit of the mitigation strategy on the system
		Design	Implementation		

Appendix B4: Risk detection systems

Risk n°	Identified risks	Risk detection system	Measurement of indicators in operation to assess detection effectiveness		
		Implementation of indicators to assess the effectiveness of the strategies	I1:	I2:	I3:
		I1 I2 I3			

Endnotes

1. Shankland, Stephen, "Tokyo 2020 Olympics using facial recognition system from NEC, Intel", CNET, 1 October 2019, https://www.cnet.com/news/tokyo-2020-olympics-using-facial-recognition-system-from-nec-intel (accessed 1 October 2020).
2. McGinnis, Chris, "Facial recognition is coming to domestic air travel", SFGATE, 8 September 2020, https://www.sfgate.com/travel/article/Facial-recognition-domestic-flights-15550415.php (accessed 2 October 2020).
3. Opened in 1978, Narita International Airport (Airport code: NRT) offers flights to over 140 domestic and international destinations. It manages about 258,000 take-offs and landings a year. See the Narita International Airport website at https://www.naa.jp/jp for more information.
4. NEC, "NEC to provide facial recognition system for new 'One ID' check-in to boarding process at Narita Airport", Press release, 28 February 2019, https://www.nec.com/en/press/201902/global_20190228_01.html (accessed 2 October 2020).
5. Ministry of Land, Infrastructure, Transport and Tourism of Japan, "Guidebook on the handling of personal data in One ID services that utilize face recognition technology at airports", 13 March 2020, https://www.mlit.go.jp/report/press/kouku19_hh_000096.html (accessed 2 October 2020).
6. World Economic Forum, "A Framework for Responsible Limits on Facial Recognition – Use Case: Flow Management", White Paper, February 2020, https://www.weforum.org/whitepapers/a-framework-for-responsible-limits-on-facial-recognition-use-case-flow-management (accessed 1 October 2020).
7. Harwell, Drew and Geoffrey A. Fowler, "U.S. Customs and Border Protection Says Photos of Travelers Were Taken in a Data Breach", *The Washington Post*, 11 June 2019, https://www.washingtonpost.com/technology/2019/06/10/us-customs-border-protection-says-photos-travelers-into-out-country-were-recently-taken-data-breach (accessed 2 October 2020).
8. Ibid.
9. Manancourt, Vincent, "Controversial US facial recognition technology likely illegal, EU body says", Politico, 10 June 2020, https://www.politico.eu/article/clearview-ai-use-likely-illegal-says-eu-data-protection-watchdog (accessed 1 October 2020).
10. Burton-Harris, Victoria and Philip Mayor, "Wrongfully Arrested Because Face Recognition Can't Tell Black People Apart", American Civil Liberties Union, 24 June 2020, https://www.aclu.org/news/privacy-technology/wrongfully-arrested-because-face-recognition-cant-tell-black-people-apart (accessed 2 October 2020).
11. Conger, Kate, Richard Fausset and Serge F. Kovaleski, "San Francisco Bans Facial Recognition Technology", *The New York Times*, 14 May 2019, https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html (accessed 2 October 2020).
12. Ravani, Sarah, "Oakland bans use of facial recognition technology, citing bias concerns", *San Francisco Chronicle*, 17 July 2019, https://www.sfchronicle.com/bayarea/article/Oakland-bans-use-of-facial-recognition-14101253.php (accessed 2 October 2020).
13. DeCosta-Klipa, Nik, "Boston City Council unanimously passes ban on facial recognition technology", *Boston Globe*, 24 June 2020, https://www.boston.com/news/local-news/2020/06/24/boston-face-recognition-technology-ban (accessed 2 October 2020).
14. Ruckstuhl, Laney, "Brookline Passes Ban On Municipal Use Of Facial Recognition Tech", WBUR News, 12 December 2019, https://www.wbur.org/news/2019/12/12/brookline-facial-recognition-technology-ban (accessed 2 October 2020).
15. DeCosta-Klipa, Nik, "Cambridge becomes the largest Massachusetts city to ban facial recognition", *Boston Globe*, 14 January 2020, https://www.boston.com/news/local-news/2020/01/14/cambridge-facial-recognition (accessed 2 October 2020).
16. Cote, Jackson, "Northampton bans facial recognition technology, becoming third community in Massachusetts to do so", Mass Live, 27 February 2020 update, https://www.masslive.com/news/2019/12/northampton-bans-facial-recognition-technology-becoming-third-community-in-massachusetts-to-do-so.html (accessed 2 October 2020).
17. NBC Boston, "Boston Approves Ban on Facial Recognition Technology", 24 June 2020, https://www.nbcboston.com/news/local/boston-approves-ban-on-facial-recognition-technology/2148450 (accessed 2 October 2020).
18. Peters, Jay, "Portland passes strongest facial recognition ban in the US", *The Verge*, 9 September 2020, https://www.theverge.com/2020/9/9/21429960/portland-passes-strongest-facial-recognition-ban-us-public-private-technology (accessed 2 October 2020).
19. State of Washington, "Engrossed Substitute Senate Bill 6280", 66th Legislature, 2020 Regular Session, 12 March 2020, http://lawfilesext.leg.wa.gov/biennium/2019-20/Pdf/Bills/Senate%20Passed%20Legislature/6280-S.PL.pdf?q=20200331083729 (accessed 2 October 2020).
20. Congress.gov, "S.4084 - Facial Recognition and Biometric Technology Moratorium Act of 2020", 116th Congress (2019-2020), https://www.congress.gov/bill/116th-congress/senate-bill/4084/text?r=1&s=1 (accessed 2 October 2020).

21. Future of Privacy Forum, *Privacy Principles for Facial Recognition Technology in Commercial Applications*, September 2018, https://fpf.org/wp-content/uploads/2019/03/Final-Privacy-Principles-Edits-1.pdf (accessed 2 October 2020).

22. American Civil Liberties Union, "Coalition Letter Calling for a Federal Moratorium on Face Recognition", 3 June 2019, https://www.aclu.org/letter/coalition-letter-calling-federal-moratorium-face-recognition (accessed 2 October 2020).

23. Learned-Miller, Erik, Vicente Ordóñez, Jamie Morgenstern and Joy Buolamwini, *Facial Recognition Technologies in the Wild: A Call for a Federal Office*, Algorithmic Justice League, 29 May 2020, https://global-uploads.webflow.com/5e027ca188c99e3515b404b7/5ed1145952bc185203f3d009_FRTsFederalOfficeMay2020.pdf (accessed 2 October 2020).

24. Duffy, Clare, "Microsoft president calls for federal regulation of facial recognition technology", CNN Business, 18 June 2020, https://edition.cnn.com/2020/06/18/tech/brad-smith-microsoft-facial-recognition/index.html (accessed 2 October 2020).

25. The Amazon blog, "We are implementing a one-year moratorium on police use of Rekognition", 10 June 2020, https://blog.aboutamazon.com/policy/we-are-implementing-a-one-year-moratorium-on-police-use-of-rekognition (accessed 2 October 2020).

26. Peters, Jay, "IBM will no longer offer, develop, or research facial recognition technology", *The Verge*, 8 June 2020, https://www.theverge.com/2020/6/8/21284683/ibm-no-longer-general-purpose-facial-recognition-analysis-software (accessed 2 October 2020).

27. European Commission, "On Artificial Intelligence – A European approach to excellence and trust", COM(2020) 65 final, 19 February 2020, https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (accessed 2 October 2020).

28. Espinoza, Javier and Madhumita Murgia, "EU backs away from call for blanket ban on facial recognition tech", *Financial Times*, 11 February 2020, https://www.ft.com/content/ff798944-4cc6-11ea-95a0-43d18ec715f5 (accessed 2 October 2020).

29. Ministry of Justice of Japan, "Further Use of Facial Recognition Automated Gates (Notice)", http://www.moj.go.jp/ENGLISH/m_nyuukokukanri07_00016.html (accessed 6 October 2020).

30. The National Institute of Standards and Technology (NIST) is a physical sciences laboratory and a non-regulatory agency of the United States Department of Commerce. Among its various activities, it runs regular performance assessments of facial recognition solutions from private vendors, public organizations and academic institutions.

31. National institute of Standards and Technology (NIST), "Face Recognition Vendor Test (FRVT), Part 2: Identification", NISTIR 8271, September 2019, https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8271.pdf (accessed 6 October 2020).

32. National institute of Standards and Technology (NIST), "Face Recognition Vendor Test (FRVT), Part 3: Demographic Effects", NISTIR 8280, December 2019, https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf (accessed 6 October 2020).

33. Ibid.

34. PR Times, "'Narita Airport Universal Design Basic Plan' has been decided", Press release, 17 April 2018, https://prtimes.jp/main/html/rd/p/000000234.000004762.html (accessed 7 October 2020).

35. Tokyo 2020 Organising Committee, "The Tokyo 2020 Accessibility Guidelines", 24 March 2017, https://tokyo2020.org/en/organising-committee/accessibility (accessed 6 October 2020).

36. Ministry of Land, Infrastructure, Transport and Tourism of Japan, "Guidebook on the handling of personal data in One ID services that utilize face recognition technology at airports", op. cit.

37. European Commission, "EU Japan Adequacy Decision", Fact sheet, January 2019, https://ec.europa.eu/info/sites/info/files/research_and_innovation/law_and_regulations/documents/adequacy-japan-factsheet_en_2019_1.pdf (accessed 6 October 2020).

38. International Organization for Standardization (ISO), "ISO/IEC 17021-1:2015 Conformity assessment — Requirements for bodies providing audit and certification of management systems — Part 1: Requirements", https://www.iso.org/standard/61651.html (accessed 7 October 2020).

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