

Human Technology  
Institute

REPORT

A background graphic consisting of numerous vertical blue lines of varying heights and thicknesses, creating a digital or data-like pattern.

# Facial recognition technology

Towards a model law

The Human Technology Institute (HTI) is building a future that applies human values to new technology. HTI embodies UTS's strategic vision to be a leading public university of technology, recognised for its global impact specifically in the responsible development, use and regulation of technology.

**Authors:**

Professor Nicholas Davis, Lauren Perry and Professor Edward Santow.

**Acknowledgements:**

- UTS staff: Professor Verity Firth, Jemima Back, Nella Soeterboek.
- Expert Reference Group members: Duncan Anderson, Distinguished Professor Fang Chen, Ivana Jurko, Kavita Kewal, Katie Kinsey, Owen Larter, Dr Monique Mann, Scott McDougall, Kieran Pender, Amanda Robinson, Roger Taylor.
- Dr Niels Wouters and the Paper Giant team.
- The Essential Research team: Peter Lewis, Hannah Barnett, Gavin White and Alissa Clement.
- Additional support and input from: Sue Glueck, Lee Hickin, Belinda Dennett, Kate Seward, Commissioner Angelene Falk, Sarah Ghali, Malcolm Crompton, Professor Mark Andrejevic, Alan Kirkland and Kate Bower.
- This project was undertaken by the Human Technology Institute, with funding from UTS and support from the UTS Centre for Social Justice & Inclusion. UTS acknowledges its generous donors including Microsoft, which provided a donation to the UTS Technology for Social Good program to advance work on responsible technology.

**Citation:**

Davis, N., Perry, L. & Santow, E. (2022) *Facial Recognition Technology: Towards a model law*, Human Technology Institute, The University of Technology Sydney.

© Human Technology Institute, The University of Technology Sydney.



Human Technology  
Institute

# Facial recognition technology

Towards a model law

**Acknowledgement of Country**

UTS acknowledges the Gadigal people of the Eora Nation, the Boorooberongal people of the Dharug Nation, the Bidiagal people and the Gamaygal people upon whose ancestral lands our university stands. We would also like to pay respect to the Elders both past and present, acknowledging them as the traditional custodians of knowledge for these lands.

<b>Foreword</b>	<b>5</b>
<b>1. Executive summary</b>	<b>6</b>
<b>2. Why do we need a model law for facial recognition technology?</b>	<b>12</b>
2.1. Purpose of the Model Law	12
2.2. The significance of ‘face data’	14
2.3. Defining facial recognition technology & its common functionalities	15
2.3.1. Facial verification	15
2.3.2. Facial identification	15
2.3.3. Facial analysis	16
2.3.4. A note on face detection	16
2.4. Glossary of key terms used in this report	17
<b>3. Our approach &amp; methodology</b>	<b>20</b>
3.1. Our overarching approach	21
3.2. Our methodology	22
3.3. Expert Reference Group	24
<b>4. The benefits &amp; risks of facial recognition technology</b>	<b>26</b>
4.1. Benefits of FRT	27
4.2. Technical & operational problems associated with FRT	28
4.3. Human rights risks	31
4.3.1. The right to privacy	31
4.3.2. Other human rights	32
<b>5. Facial recognition policy in the Australian &amp; global context</b>	<b>34</b>
5.1. A comparative law analysis	35
5.2. Australian law applicable to FRT	37
5.2.1. Privacy law	37
5.2.2. Other relevant Australian laws	39
5.3. Voluntary action by FRT Developers	39
<b>6. Model Law: overview</b>	<b>40</b>
6.1. To whom does the Model Law apply?	41
6.2. Outline of the Model Law	42
<b>7. Model Law: human rights risk assessment</b>	<b>44</b>
7.1. Human rights vulnerabilities and overall human rights risk	45
7.2. Factors relevant to the human rights risk assessment	46
7.2.1. Factor 1: spatial context	46
7.2.2. Factor 2: functionality of the FRT Application	48
7.2.3. Factor 3: performance of the FRT Application	49
7.2.4. Factor 4: the FRT Application’s role in decision making	50
7.2.5. Factor 5: prior, free and informed consent	52

7.3. Assessing the overall human rights risk level	55
7.4. Determining whether human rights limitation is justified	56
7.4.1. Which human rights are being restricted?	56
7.4.2. Is the human rights restriction legally justified?	56
7.4.3. Importance of 'reasonable, necessary and proportionate' criterion	56
<b>8. Model Law: the Facial Recognition Impact Assessment process</b>	<b>58</b>
8.1. FRIA Step 1 – use & risk assessment declaration	60
8.2. FRIA Step 2 – risk management declaration	60
8.3. Registration, publication & updating obligations	62
8.3.1. When FRT Deployers do not need to complete & register their own FRIA	62
8.3.2. FRIA updating obligations	62
<b>9. Model Law: risk-based legal requirements</b>	<b>65</b>
9.1. Mapping risk levels to obligations	65
9.2. Base-level legal requirements applying to all uses of FRT	65
9.2.1. Requirement to complete a FRIA	65
9.2.2. New legal requirements under the Model Law	66
9.2.3. Creation of a technical standard for FRT	67
9.3. Elevated risk: additional legal requirements	70
9.4. High-risk FRT Applications	71
9.4.1. Prohibition of high-risk FRT Applications subject to limited exceptions	71
9.4.2. Regulator authorisation	71
9.4.3. Law enforcement & national security	72
9.4.4. Genuine research	75
<b>10. Independent review &amp; dispute resolution</b>	<b>76</b>
<b>11. Implementation of the Model Law</b>	<b>78</b>
11.1. Primary legislation	79
11.2. Assigning & resourcing the regulator	80
11.3. Ensuring the law is accessible, clear & effective	81
11.4. A harmonised approach across all Australian jurisdictions	82
11.5. An Australian Government taskforce on facial recognition	83
<b>Appendix 1: Methodology &amp; consultation</b>	<b>84</b>
<b>Appendix 2: Summary of the Model Law's legal requirements</b>	<b>87</b>
<b>Appendix 3: Facial Recognition Impact Assessment Template</b>	<b>88</b>



# Foreword

## A need for reform

We are experiencing an extraordinary rise in the development and use of facial recognition technology (FRT) – in Australia and around the world. Yet, our laws were never drafted with this reality in mind.

As a result, Australian law does not effectively regulate FRT: our law does not reliably uphold human rights, nor does it incentivise positive innovation. Every liberal democracy around the world is facing a similar problem.

Facial recognition technology is being woven into the fabric of our personal, professional and communal lives. Increasingly, FRT applications are inside the devices that are used by, and on, Australians.

Many of us will have experienced FRT unlocking a smartphone, organising photos of friends and family, in home security systems, at passport control, and in monitoring and surveillance by employers and law enforcement. This list is rapidly expanding. While FRT is primarily used to identify an individual or to verify that they are who they claim to be, it is increasingly being used to assess characteristics, such as a person's age, gender or even emotions – albeit with widely-variable accuracy.

Well-designed, thoughtfully-implemented FRT offers convenience and efficiency, particularly in identifying people at scale. The technology can even enhance human rights: FRT is widely used by people who are blind or have low vision, and it can be used to locate missing people and identify victims of crimes.

However, this technology also threatens our human rights. Most obviously, FRT's reliance on sensitive personal information intrudes on the right to privacy. As FRT is deployed more widely, the risk of mass surveillance increases.

Particular human rights risks arise when FRT is used to make high-stakes decisions. For example, if an individual is wrongly identified as a criminal suspect, they could be unlawfully arrested and detained. Where errors caused by FRT disproportionately affect particular groups in our community – including women and people of colour – this can threaten the right to equality or non-discrimination.

There is a growing consensus – from leading voices in civil society, the private sector, government and academic experts – that change is needed. This report aims to respond to that need in two ways.

First, the report explains how current Australian law applies to the development and use of FRT. Drawing on leading research, it sets out the gaps in Australian law, especially where those gaps expose threats to Australians' human rights.

Second, this report proposes reform. It outlines a model law to regulate the development and use of FRT, as this affects people in Australia. The model law adopts a risk-based approach grounded in international human rights law, connecting Australian law with that of other jurisdictions. The model law fosters innovation by enabling the responsible use of FRT, while also protecting against the risks posed to human rights.

Australia needs a dedicated facial recognition law. This report urges the Federal Attorney-General to lead this pressing and important reform process.

**Professor Nicholas Davis**  
**Lauren Perry**  
**Professor Edward Santow**

*September 2022*

Part 1.

# Executive summary



# What is this report intended to achieve?

There is growing community concern about the rise of facial recognition

technology (FRT). As with other jurisdictions around the world, Australian law does not provide the legal guardrails necessary to ensure that FRT is developed and deployed in ways that uphold basic human rights.

This report proposes reform. It provides an outline of a model law for FRT (the Model Law). The Model Law aims to foster innovation and enable the responsible use of FRT, while protecting against the risks posed to human rights.

This report recognises that FRT can be used consistently with international human rights law, and indeed in ways that achieve public and other benefits. However, FRT necessarily also engages, and often limits or restricts, a range of human rights. As a result, the use of FRT can – and has been proven to – cause harm.

The Model Law is intended to be applied to any individual or organisation that develops, distributes, or deploys FRT in Australia. It covers use of FRT by both government and private sector organisations.

The precise human rights impact of FRT turns on how the technology is developed, deployed and regulated. Therefore, the Model Law proposed in this report focuses on how FRT is used in practice, adopting a risk-based approach grounded in international human rights law. While the report has been written primarily by reference to Australian law, the reform principles set out in this report are applicable to other, comparable jurisdictions.

---

*Australian law does not provide the legal guardrails necessary to ensure that FRT is developed and deployed in ways that uphold basic human rights.*

---

# Why is reform needed?

There is rapid, almost exponential, growth in the development and deployment of FRT and other remote biometric technologies. These technologies can identify and extract a wealth of sensitive personal information about an individual, often without the individual's knowledge, let alone consent. Australian law, like the laws of most jurisdictions around the world, was not developed with the prospect of widespread use of FRT in mind. In particular, our law was not drafted to address the challenges posed by FRT to human rights such as the right to privacy, freedom of assembly and association, freedom of expression and of movement.

Many civil society organisations, government and inter-governmental bodies and independent experts have sounded the alarm about dangers associated with current and predicted uses of FRT – including the inadequacy of existing law to protect communities and individuals from having their human rights restricted. Several leading trans-national technology companies have expressed concern that existing laws do not protect against harmful use of FRT. This has prompted a number of companies to voluntarily limit their own use of FRT, including in the products and services they sell. However, many other companies have not tempered their use of FRT.

In Australia and other similar jurisdictions, several existing laws apply to the development and use of FRT. For example, Australian privacy law includes several provisions dealing with the handling of biometric information. Yet, on the whole, these existing laws are inadequate in addressing many of the risks associated with FRT.

Some jurisdictions have responded to the rise of FRT by prohibiting certain uses of FRT. Most famously, in 2019, the city of San Francisco issued a legal moratorium that prohibits many uses of FRT by the San Francisco Police Department. While this sort of moratorium may be useful in addressing a very specific risk, it is a limited and blunt instrument, which can leave many uses of FRT unregulated. In addition, if a moratorium were introduced to prohibit all development and use of FRT (something that no major jurisdiction has done), it would preclude uses of the technology that have a demonstrable public benefit.

Against this backdrop, a small but growing number of jurisdictions have begun to explore a more nuanced approach to regulating FRT. Especially in the United States and Europe, risk-based laws have been proposed to enable beneficial forms or applications of FRT, while restricting or prohibiting harmful uses of FRT. This report has been drafted to apply the lessons from those reform processes to create a nuanced, risk-based, FRT-focused Model Law.

# What is facial recognition technology?

Facial recognition technology is defined in this report as any computer system or device with embedded functionality that uses data drawn from human faces to verify an individual's identity, identify an individual and/or analyse characteristics about an individual.

This report focuses on FRT, which is a specific form of biometric technology that has some unusual, if not unique, characteristics. In considering broader reform in this area, the authors urge that the reform principles set out in this report be adapted to apply also to other forms of remote biometric technology, including those based on an individual's voice, gait, ear, iris, body odour and other biometric data.

# How does the Model Law work?

The Model Law sets out a risk-based approach to FRT, grounded in human rights. Under the Model Law, anyone who develops or deploys an FRT Application must first assess the level of human rights risk that would apply to their particular FRT Application. In assessing this risk, it will be necessary to consider a range of factors including:

- how the FRT application functions
- where and how it is deployed (for example, the spatial context)
- the performance or accuracy of the application, and
- the effect of any decisions made in reliance on the FRT application's outputs.
- whether affected individuals can provide free and informed consent.

Drawing on these factors, the Model Law provides for a structured way of assessing the human rights risk of each specific FRT Application through a 'Facial Recognition Impact Assessment' (FRIA). FRT Developers and Deployers must complete this FRIA process, and assign a risk rating to the relevant FRT Application: base-level, elevated or high risk. That assessment can be challenged by members of the public and the regulator.

To address this human rights risk, the Model Law contains a cumulative set of legal requirements, limitations and prohibitions that apply based on this risk assessment. The Model Law imposes stricter legal constraints, and prohibitions, as the level of risk for any particular FRT Application increases.

Some of the Model Law's requirements are procedural – for example, FRIAs must be registered with the regulator and made publicly available to ensure transparency of operation and use. Other requirements are substantive – for example, the Model Law applies and extends existing privacy law obligations to FRT Applications. In addition, the Model Law provides for the creation of a new FRT technical standard that would have the force of law.

The Model Law prohibits the development and use of high-risk FRT Applications, subject to three exceptions: where the regulator provides specific authorisation; in genuine research; and in the context of law enforcement and national security agencies, where the Model Law provides for specific legal rules, including a 'face warrant' scheme.

Finally, the report recommends that a suitable regulator be legally empowered and resourced to oversee the development and use of FRT in Australia. The Office of the Australian Information Commissioner (OAIC) would be the most obvious candidate to regulate the development and use of FRT in the federal jurisdiction, with a harmonised approach in respect of the state and territory jurisdictions.

# Next steps for urgent reform

There is an emerging consensus across diverse stakeholder groups that reform in this area is both urgent and important.

This report calls on Australia's Federal Attorney-General to lead the reform process by taking four key steps:

1. The Attorney-General should introduce a bill into the Australian Parliament, based on the FRT Model Law set out in this report. This bill would apply to FRT within the regulatory purview of the Australian Government.
2. The Attorney-General should assign regulatory responsibility to the Office of the Australian Information Commissioner, or another suitable regulator, empowering that body to take a central role in the creation of an FRT technical standard, and in providing advice for FRT Developers, Deployers and affected individuals. The Australian Government should provide appropriate resourcing to the FRT regulator to fulfil these new functions.
3. The Attorney-General should initiate a process with his state and territory counterparts to ensure that the law on FRT is harmonised across all Australian jurisdictions. This process should ensure the law is consistent and easy to understand for FRT Developers, Deployers and affected individuals regardless of where one is located in Australia.
4. The Attorney-General should work with other relevant federal ministers to establish an Australian Government taskforce on FRT. The taskforce would have two functions. First, it would work with all relevant Federal Government departments and agencies, such as the Australian Federal Police, to ensure their development and use of FRT accords with legal and ethical standards. Second, it would lead Australia's international engagement on FRT, so that Australia can have a positive influence on the development of international standards and other assurance mechanisms for FRT, and to ensure that Australia's legal approach to FRT is consistent with international law and international best practice.

---

*Facial recognition reform is urgent and important. Australia's Federal Attorney-General should lead this reform process.*

---

Part 2.

Why do we need  
a model law for  
facial recognition  
technology?

This report presents the outline of a Model Law for FRT, setting out the key elements for reform in order to assist in drafting a bill of parliament.<sup>1</sup> Parts 2 to 5 of this report provide important context for the Model Law. Parts 6 to 10 of this report present the outline of the Model Law designed to regulate the development and use of FRT.

## 2.1. Purpose of the Model Law

The Model Law's purpose is to restrict, and in some cases prohibit, the development and deployment of FRT that risks human rights harm, while enabling FRT that is developed and used in ways consistent with human rights and Australia's liberal democratic values. The Model Law is intended to apply to FRT Developers and Deployers, where either the development or deployment takes place in Australia.

To achieve this, the Model Law and related consequential amendments to other legislation have the following goals:

- **uphold human rights** – Australian law should provide that human rights are protected in the development and use of FRT
- **apply a risk-based approach** – there should be a clear, straightforward legal framework that allows each FRT Application to be classified according to its level of risk, with legal restrictions and prohibitions on FRT calibrated to an FRT Application's relative level of risk to human rights
- **support for compliance** – the law should support FRT Deployers and Developers to meet their obligations
- **transparency in the use of FRT Applications** – the regulator and affected individuals should be able to understand how FRT Applications are deployed in Australia
- **effective oversight and regulation** – an appropriately-resourced regulator should be empowered to oversee the operation of the Model Law
- **accountability and redress** – where an affected individual considers the Model Law has been breached, they should be able to seek redress in a way that is simple and cost effective
- **jurisdictional compatibility** – the law should apply consistently across jurisdictions, including Australia's federal government, states and territories, and refer to global standards to support international interoperability.

1. This report does not itself contain a bill; rather it is the *outline* of a model law on FRT.

## 2.2. The significance of ‘face data’

Our face contains important information about us, which is unique to each individual. Information derived from one or more images of an individual’s face is biometric data.

Biometric data is at the heart of technology that enables ‘automated recognition of individuals based on their biological and behavioural characteristics’.<sup>2</sup> Generally, a person’s face is visible to the world – in this sense, our face is public – and images of a person’s face can be captured remotely. For this reason, FRT is often categorised as a form of ‘remote biometric technology’.<sup>3</sup>

The human face, and the information that can be derived from it, is special. The face is simultaneously deeply personal and publicly visible. We recognise one another by reference to each other’s faces, and our assessment of another person’s face is important in how we relate to and empathise with that person.

Faces also hold a concentration of important, sensitive information about people – data that can reveal or store information about someone’s gender, age, ethnicity, health conditions, emotional state, and behaviour. Face data can be captured in both static forms (from a single point in time) and dynamic forms (from moving, context-specific footage), and captured remotely by a wide variety of widely available devices. This differs to other biometric technologies like fingerprints and iris scans.<sup>4</sup>

Biometric templates and biometric information used for the purpose of automated verification or identification is ‘sensitive information’ under the *Privacy Act 1988* (Cth). This triggers the most stringent protections under that Act. However, even if an FRT Application might be deliberately designed *not* to identify any specific individual,

the information gathered from faces is still highly personal. For example, some FRT Applications use facial analysis to gather sensitive information about an individual’s emotional state without identifying the individual.

Face data is now very easy to obtain. Face data can be captured from almost any modern digital camera – including cameras embedded in smartphones and other such devices – and can be readily available through digital photos obtained from public online sources, such as on social media platforms like Facebook, LinkedIn, and Twitter. As a result, there are now many ways by which FRT can be performed on a person – an affected individual – without their knowledge, let alone their consent.

It is almost certain that, if you are reading this report, your face data exists in one or more FRT databases. For example, the Australian Government maintains such a database in respect of all Australian citizens who have a ‘biometric passport’. Similarly, national governments in countries such as the United States, the United Kingdom and China maintain FRT databases that include individuals of all nationalities who enter at many of those countries’ respective borders. More controversially, several private companies, such as Clearview AI, have created FRT databases using publicly accessible face data that they have obtained online, from social media platforms and search engines.<sup>5</sup>

In short, the special nature of the human face means that systems using ‘face data’ render us vulnerable to our rights being restricted or violated.

---

2. International Standards Organization, ‘ISO - ISO/IEC 2382-37:2022 – Information Technology – Vocabulary – Part 37: Biometrics’ <<https://www.iso.org/standard/73514.html>>.

3. Remote biometric technologies are AI systems created ‘for the purpose of identifying natural persons at a distance through the comparison of a person’s biometric data with the biometric data contained in a reference database, and without prior knowledge of the user of the AI system whether the person will be present and can be identified’: Explanatory Memorandum, Proposal for a Regulation of the European Parliament and of the Council: Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts 2021 (European Commission) 3.

4. Ada Lovelace Institute, *Countermeasures: The Need for New Legislation to Govern Biometric Technologies* in the UK (Report, June 2022) 1.

5. ‘Company Overview’, *Clearview AI* (Web Page, 2022) <<https://www.clearview.ai/overview>>.



## 2.3. Defining facial recognition technology & its common functionalities

This report defines FRT broadly, as any computer system or device with embedded functionality that uses biometric data drawn from human faces to verify someone's identity, identify a particular individual and/or analyse characteristics about a person. Within this definition, it is useful to distinguish between four different core *functionalities* of FRT.

### 2.3.1. Facial verification

Facial verification – also known as ‘one-to-one face matching’ – is a form of FRT used to verify an individual's identity. Put simply, an FRT Application that undertakes facial verification can be used to determine whether an individual is who they claim to be.

Facial verification works by using image-capture devices (such as cameras), algorithms and stored data to match face data from an input data source to a pre-existing, validated record. Facial verification is commonly used in smartphones, tablets and other such devices as an alternative to a password, to ‘unlock’ the device using an individual's face data. Facial verification is also increasingly common at national borders. For example, so-called ‘eGates’ at border control in Australia and other countries use facial verification to verify the identity of travellers by comparing face data derived from a photograph taken by a digital camera at the immigration control area to a digital record in, or connected to, the individual's passport.

### 2.3.2. Facial identification

Facial identification – also known as ‘one-to-many matching’ – compares unique data from your face with other records. Whereas *facial verification* compares new face data with a single, previously stored and validated entry, *facial identification* compares a captured image to a larger number of potential records, and searches for a match. An FRT Application that engages in facial identification can be used to answer the question, ‘who is this person?’

Facial identification and facial verification have only become widely accessible in the last decade, as machine learning algorithms, computer processing power and large datasets containing images of faces have become more available. Facial identification is being used increasingly by law enforcement, to identify an unknown criminal suspect or victim of a crime by comparing their faces to stored images of convicted criminals or faces that appear in other image databases such as driver licence photos. Some facial identification systems offer the ability to use any photo as input, and to find matches across billions of facial images gathered from social media and other internet sources.<sup>6</sup>

---

*An FRT Application that engages in facial identification can be used to answer the question, ‘who is this person?’*

---

6. ‘Clearview AI Releases 2.0 Version of Industry Leading Facial Recognition Platform for Law Enforcement’, *Clearview AI* (Web Page, 2019) <<https://www.clearview.ai/clearview-ai-releases-2-version-of-industry-leading-facial-recognition-platform-for-law-enforce>>.

### 2.3.3. Facial analysis

Facial analysis encompasses a wide variety of techniques that automatically draw inferences about the characteristics of an individual from their face data. What links these techniques is that they all rely on machine learning to identify correlations between certain facial features and movements, and certain human characteristics, emotions or behaviours.

Facial analysis can be divided into a number of sub-categories, including:

- *demographic* facial analysis looks to ascertain the age, sex or ethnicity of a person from face data
- *health information* facial analysis attempts to determine the health or disease status of a person from their face data
- *behavioural* facial analysis uses face data to identify information such as where someone is looking or what they are wearing
- *emotion* facial analysis draws inferences about a person's emotional state from their expression
- *intention* facial analysis tries to predict what a person wants from their face data.

It should be emphasised that facial analysis is, in general, highly controversial. Many types of facial analysis are unproven, and some rest on dangerous assumptions. For instance, one form of facial analysis claimed the ability to predict an individual's sexual orientation through facial analysis<sup>7</sup> – a process that resembles discredited theories such as phrenology. Notwithstanding this controversy, and the limited hard research to demonstrate the accuracy of facial analysis applications, this type of FRT is growing in popularity.

### 2.3.4. A note on face detection

'Face detection' is an automated process that seeks to detect when a human face is in a particular area, but it does not seek to identify any individual, nor provide other personal information about the individual.

Face detection, without other forms of FRT functionality, is similar to other forms of object recognition, and therefore sits outside the scope of this Model Law. Examples of face detection include digital cameras that detect faces or eyes to automatically use them as a point of focus, as well as features in mobile applications (including Instagram, SnapChat and others), which detect faces in order to overlay alternative visual features on screen.

---

*Many types of facial analysis are unproven, and some rest on dangerous assumptions.*

---

7. Yilun Wang and Michal Kosinski, 'Deep Neural Networks Are More Accurate than Humans at Detecting Sexual Orientation from Facial Images' (2018) 114(2) *Journal of Personality and Social Psychology* 246. The accuracy of the outcomes of this research has been challenged by several other researchers – see, e.g., Dawei Wang, 'Presentation in Self-Posted Facial Images Can Expose Sexual Orientation: Implications for Research and Privacy' (2022) 122(5) *Journal of personality and social psychology* 806.

## 2.4. Glossary of key terms used in this report

This table defines a number of key terms used in this report.


Term	Definition in this report
affected individual	An affected individual is an individual whose face data is used by an FRT Application.
automation	Automation refers to a computational system applying algorithms or other rules to particular fact scenarios with limited or no human involvement. A decision-making system may be wholly automated, in which case it produces decisions without human involvement. A system may be partially automated, meaning that it produces inferences, predictions or recommendations, which a human will use to make a final decision.
biometric data	Biometric data is information which pertains to the physical, physiological or behavioural characteristics of a person which can enable the unique identification of that person. Gait, fingerprints and images of the face are examples of biometric data.
biometric technology	Biometric technologies are any programs or systems which use biometric data to derive, assess and/or analyse information about people. Facial recognition technologies are a sub-type of biometric technology, as they can be used to verify, identity or analyse people through face data.
captured data	Captured data is face data collected and used as an input for a specific FRT use case.
a decision with legal or similarly significant effects	<p>The legal concept of a decision that produces ‘legal ... or similarly significant’ effects was introduced in Article 22 of the European Union’s <i>General Data Protection Regulation</i> (GDPR). This report applies this legal concept consistently with how it is generally interpreted in European jurisdictions:</p> <ul style="list-style-type: none"> <li>▪ a ‘decision with a legal effect’ is one that affects an individual’s legal status or legally recognised rights</li> <li>▪ A decision with a ‘similarly significant’ effect is one that has a significant impact on an individual’s life opportunities, behaviour or wellbeing.</li> </ul>
face data	Face data is data or information drawn from a human face in a way that can be used in an FRT algorithm, application or system.
facial analysis	Facial analysis is a functionality of FRT that attempts to draw inferences about the characteristics of an individual – including demographic features, health information, behaviour, emotional state, and intentions – from face data.
facial identification	Facial identification – also known as ‘one-to-many’ or ‘many-to-many’ face matching – is a functionality of FRT that compares captured face data to a set of reference data in order to search for matches and thereby identify one or more individuals.

<b>Term</b>	<b>Definition in this report</b>
facial recognition technology (FRT)	FRT is technology that uses face data to verify an individual's identity, identify an individual and/or analyse characteristics about an individual.
facial verification	Facial verification – also known as ‘one-to-one face matching’ – is a functionality of FRT that compares captured face data to a single reference image to assess if there is a match between the two, primarily for the purposes of verifying identity.
FRT Algorithm	An FRT Algorithm is software that uses face data to assist in performing an FRT function.
FRT Application	An FRT Application is a product or service that performs an FRT function (identity verification, identification and/or facial analysis).
FRT Deployer	An FRT Deployer is a person (including an individual, corporation or other organisation) that uses or deploys an FRT Application on one or more affected individuals.
FRT Developer	An FRT Developer is a person that creates an FRT Application (see definition above). Any person that uses or deploys an FRT Application on affected individuals is an ‘FRT Deployer’. An FRT Developer typically sells or provides an FRT Application to be used by an FRT Deployer. A person may be both an FRT Developer and an FRT Deployer if the person both develops the FRT Application and uses it on affected individuals.
FRT System	An FRT System embeds one or more FRT Applications as part of a larger decision-making process. An example of an FRT System is an online payment system that requires an affected individual to verify their identity using an FRT Application as one of several steps in making a payment.
individual	An individual is a natural person or human. An individual cannot be an organisation, corporation or other non-human entity.
person	A person is a legal person. A person includes a corporation, other type of organisation or an individual.
reference data	Reference data is a subset of ‘captured data’ (see definition above). Reference data is placed in a database for the purposes of future facial verification or identification.
remote biometric identification	Remote biometric identification involves the use of biometric techniques such as analysis of fingerprints, face data, irises, vein patterns, voices or ears to gather information on people ‘at a distance, in a public space and in a continuous or ongoing manner by checking them against data stored in a database.’ <sup>8</sup>

8. Directorate-General for Communications Networks European Commission, *White Paper on Artificial Intelligence – A European Approach to Excellence and Trust* (Website No COM/2020/65, Publications Office of the European Union, 19 February 2020) 18 <<http://op.europa.eu/en/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1>>.



**Part 3.**



**Our approach &  
methodology**

## 3.1. Our overarching approach

Facial recognition technology can be developed and used in ways that are consistent with or even promote human rights. However, the opposite is also true; poor design and deployment of FRT can threaten human rights. The Model Law thus tailors the legal requirements applicable to an FRT Application to the relative level of human rights risk posed by the FRT Application.

This is known as a risk-based approach. Risk-based approaches to regulation can help manage uncertainty across systems as a whole, rather than dealing with individual cases of harm after they have occurred. The risk-based approach adopted in the Model Law will impose greater restrictions on use cases where harm to human rights is more likely and more serious, and it will incentivise FRT Applications where this risk is lower.

The Model Law defines risk by reference to international human rights law. This report adopts human rights as the Model Law's normative foundation, because international human rights law applies throughout Australia and almost universally throughout the world.

International human rights law prioritises the protection of individuals and the broader society, while enabling the public and private sectors to advance a diverse range of other interests, ranging from the protection of national security to engaging in commerce. In other words, this framework provides a mechanism to reconcile human rights and other legitimate interests, even where they may be in tension.

For example, FRT Applications necessarily intrude on the right to privacy, because of the way they use personal information. Under international law, privacy is not an absolute human right. Provided that the intrusion on privacy is reasonable, necessary and proportionate to the pursuit of a legitimate aim, it will be permissible under international human rights law. The Model Law proposed in this report seeks to embody this approach.

---

*The Model Law tailors the legal requirements applicable to an FRT Application to the relative level of human rights risk posed by the FRT Application.*

---

## 3.2. Our methodology

This project was prompted in large part by the many civil society, academic, industry and government bodies that have, in recent years, expressed concern about the rise of FRT, its impact on human rights, and the inadequacy of existing laws in Australia and other comparable jurisdictions to strike an appropriate balance. Particularly influential on this report's approach have been:

- the Australian Human Rights Commission's 2021 *Human Rights & Technology* report<sup>9</sup>
- the landmark bipartisan report by Australia's Parliamentary Joint Committee on Intelligence and Security, which expressed deep concern for the inadequacy of privacy and other protections in the then Australian Government's proposed legal framework for FRT and other biometric technology<sup>10</sup>
- IBM's decision to stop selling general purpose facial recognition and analysis software products in June 2020<sup>11</sup>
- decisions by Amazon<sup>12</sup> and Microsoft<sup>13</sup> to ban police use of their facial recognition products in 2020<sup>14</sup>
- research led by Professor Mark Andrejevic, including in the 2020 White Paper, *Australian Attitudes to Facial Recognition: a National Survey*<sup>15</sup>
- Meta's decision to shut down its Facebook Face Recognition system in November 2021<sup>16</sup>
- the work by Professor Theodore Christakis et al on the use of FRT in public contexts<sup>17</sup>
- the Office of the Australian Information Commissioner's 2021 determinations regarding Clearview AI, and especially its determination regarding the Australian Federal Police's use of FRT provided by Clearview AI<sup>18</sup>
- the investigation by Australian consumer advocacy group CHOICE into the use of FRT by a number of major retailers<sup>19</sup>
- Microsoft's introduction in 2022 of customer eligibility requirements for the purchase and use of FRT Applications, and the company's withdrawal of facial analysis products that analyse emotional states and identity attributes<sup>20</sup>
- Ada Lovelace Institute's 2022 report, *Countermeasures: the need for new legislation to govern biometric technologies in the UK*.<sup>21</sup>

9. See Australian Human Rights Commission, *Human Rights and Technology Final Report*, (Report, March 2021) Ch 9.

10. Australian Parliament, Parliamentary Joint Committee on Intelligence and Security, *Advisory Report on the Identity-Matching Services Bill 2019 and the Australian Passports Amendment (Identity-Matching Services) Bill 2019* (Parliamentary Committee Report, October 2019).

11. Arvind Krishna, 'IBM CEO's Letter to Congress on Racial Justice Reform', *IBM Policy* (Web Page, 10 December 2019) <<https://www.ibm.com/policy/facial-recognition-sunset-racial-justice-reforms/>>.

12. Jeffrey Dastin, 'Amazon Extends Moratorium on Police Use of Facial Recognition Software', *Reuters* (online, 18 May 2021) <<https://www.reuters.com/technology/exclusive-amazon-extends-moratorium-police-use-facial-recognition-software-2021-05-18/>>.

13. Jay Greene, 'Microsoft Won't Sell Police Its Facial-Recognition Technology, Following Similar Moves by Amazon and IBM', *Washington Post* (online, 11 June 2020) <<https://www.washingtonpost.com/technology/2020/06/11/microsoft-facial-recognition/>>.

14. Jeffrey Dastin, 'Amazon extends moratorium on police use of facial recognition software' (2021), *Reuters* (online 19 May 2021) <<https://www.reuters.com/technology/exclusive-amazon-extends-moratorium-police-use-facial-recognition-software-2021-05-18/>>.

15. Mark Andrejevic, Robbie Fordyce, Luzhou Li and Verity Trott, *Australian Attitudes to Facial Recognition: A National Survey* (White Paper no.1, Automated Society Working Group School of Media, Film, and Journalism Monash University, Monash University, May 2020).

16. Jerome Pesenti, 'An Update On Our Use of Face Recognition', *Meta - Facebook* (Web Page, 2 November 2021) <<https://about.fb.com/news/2021/11/update-on-use-of-face-recognition/>>.

17. See for example Theodore Christakis, Karine Bannelier, Claude Castelluccia, and Daniel Le Metayer, *Mapping the Use of Facial Recognition in Public Spaces in Europe - Part 3*, (Report of the AI- Regulation Chair (AI-Regulation.Com) MIAI, 23 May 2022).

18. Office of the Australian Information Commissioner, Government of Australia, 'AFP Ordered to Strengthen Privacy Governance', *News and Media* (Web Page, 16 December 2021) <<https://www.oaic.gov.au/updates/news-and-media/afp-ordered-to-strengthen-privacy-governance>>.

19. Jarni Blakkarly, 'Kmart, Bunnings and The Good Guys Using Facial Recognition Technology in Stores', *CHOICE* (Web Page, 2 August 2022) <<https://www.choice.com.au/consumers-and-data/data-collection-and-use/how-your-data-is-used/articles/kmart-bunnings-and-the-good-guys-using-facial-recognition-technology-in-stores>>.

20. Sarah Bird, 'Responsible AI Investments and Safeguards for Facial Recognition', *Microsoft Azure Artificial Intelligence Blog* (Web Page, 21 June 2022) <<https://azure.microsoft.com/en-us/blog/responsible-ai-investments-and-safeguards-for-facial-recognition/>>.

21. Ada Lovelace Institute, *Countermeasures: The Need for New Legislation to Govern Biometric Technologies in the UK* (Report, June 2022).

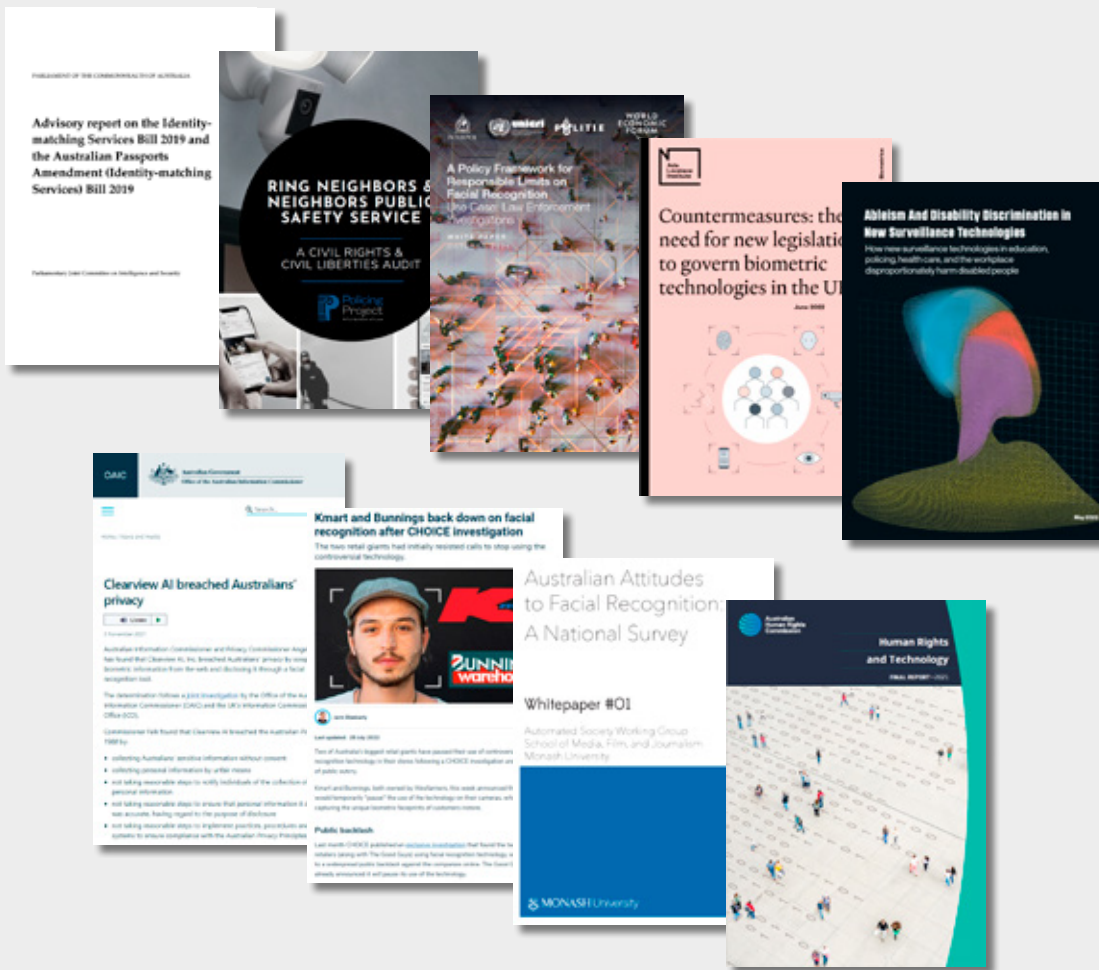


Many of these bodies themselves undertook extensive public consultation prior to expressing their concern for FRT. For example, one of the authors of this report, Professor Edward Santow, in his former role as Australia's Human Rights Commissioner, led an extensive public consultation process on human rights and technology, which underpinned the Commission's concern regarding the regulation of FRT.

In addition to drawing on the work referred to above, the authors have also solicited input from affected individuals, subject-matter experts, and stakeholders from government, industry and civil society. The report relies especially on advice from the project's Expert Reference Group.

This project also commissioned qualitative research from Paper Giant and Essential Research. That qualitative research involved the creation of an FRT simulation tool, which enabled focus group participants, drawn from diverse demographic groups within Australia, to experience a variety of different types of FRT. The participants then took part in structured group interviews convened and run by Essential Research. We summarise some of the key insights from that qualitative research in Appendix 1.

### Leading reports on facial recognition reform



### 3.3. Expert Reference Group

In February 2022, the authors of this report invited 11 national and international experts to form the project's Expert Reference Group (ERG).<sup>22</sup> The expertise of the ERG members spanned the commercial application of emerging technology, government policy, law and regulation, academic research and human rights. The ERG provided expert advice on the operation and outcomes of the FRT Project.

The ERG provided input via individual meetings with project team members and in three formal meetings of the full ERG, where the key issues for discussion were: the report structure and qualitative research; the drafting of this report;

and options to give effect to the report's proposed reform. While the ERG members provided invaluable advice to the project, the ERG was not asked to endorse the report itself. The co-authors – Nicholas Davis, Lauren Perry and Edward Santow – are the sole authors of this report.

A more detailed overview of the project's methodology is set out in Appendix 1.

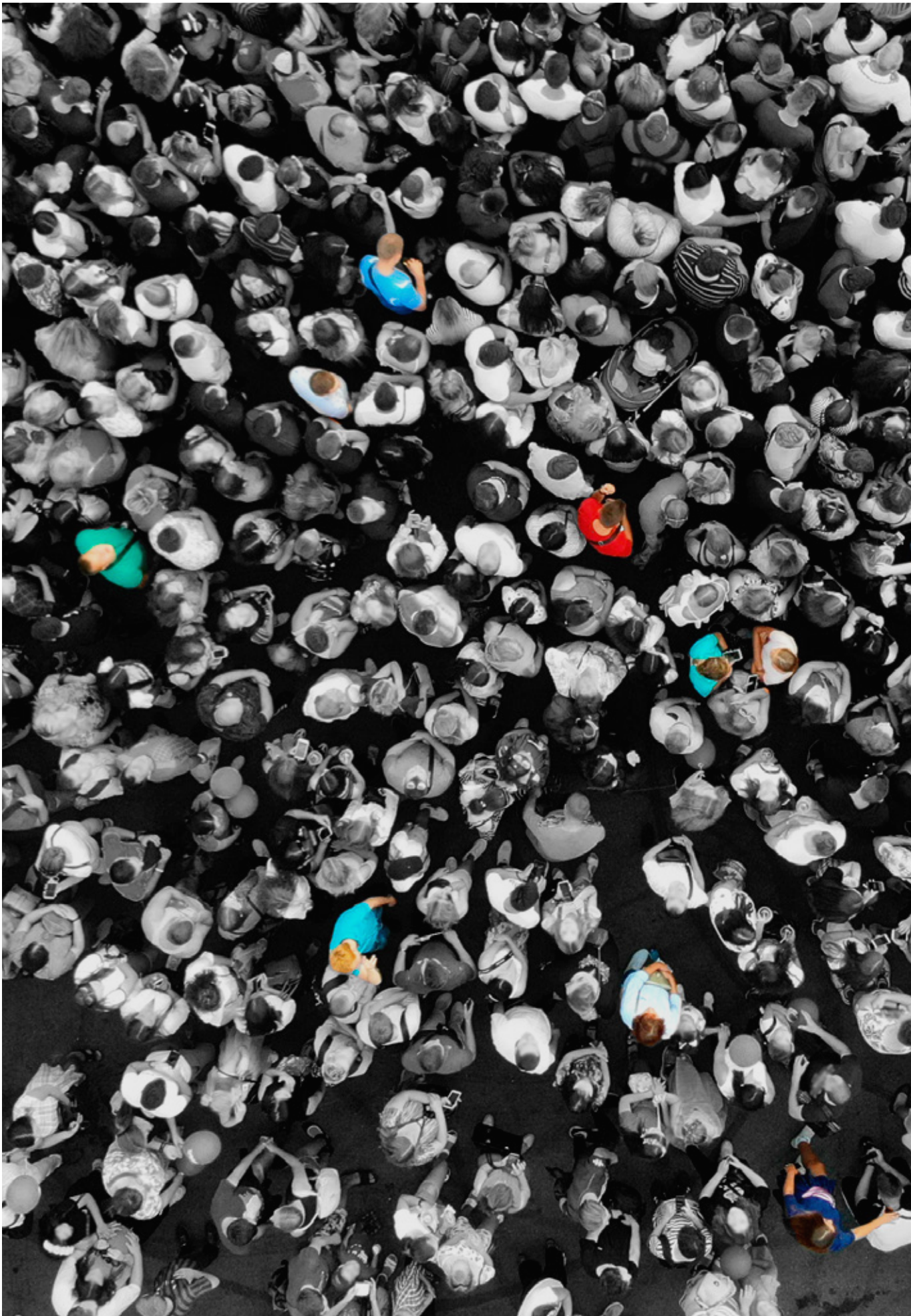
---

*The expertise of the ERG members spanned the commercial application of emerging technology, government policy, law and regulation, academic research and human rights.*

---

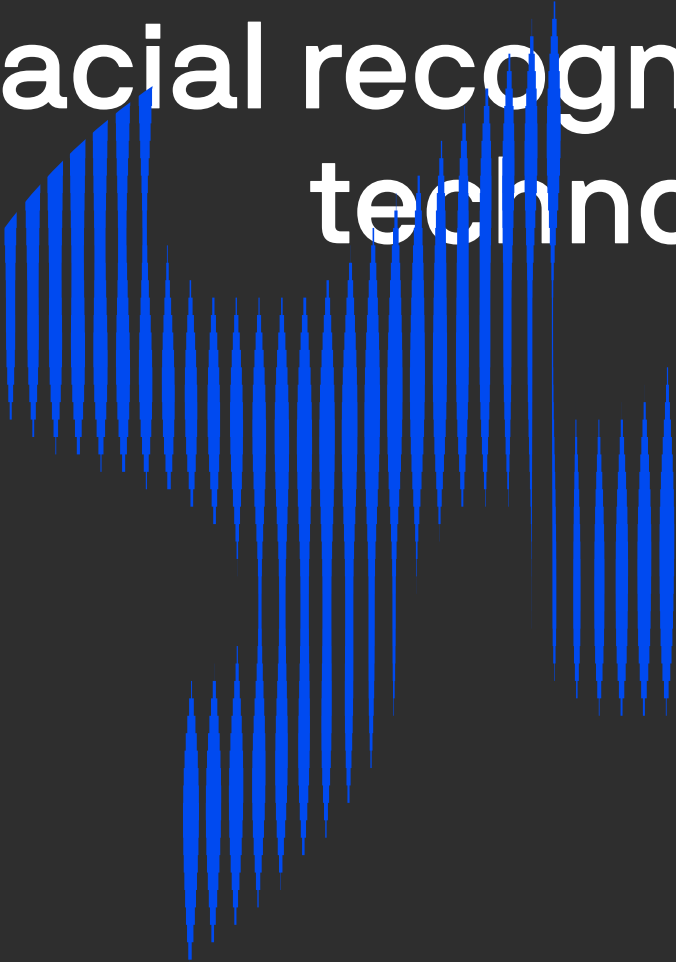
---

<sup>22</sup> The members of the Expert Reference Group are listed in Appendix 1.



Part 4.

# The benefits & risks of facial recognition technology



## 4.1. Benefits of FRT

Many use cases for FRT are consistent with human rights, and indeed some FRT Applications can be used in ways that advance human rights. Benefits associated with FRT include:

- **Convenience.** FRT Systems can offer convenience for affected individuals, and for the companies and governments that deploy them. Many people report that using facial verification to ‘unlock’ their smartphone or other similar device can be faster and more convenient than alternatives such as a passcode or another biometric authentication (such as fingerprint). As one participant noted in the focus groups conducted by Essential Research, swipe cards can be easily left at home making out-of-hours building access a challenge. This individual believes facial verification technology used to provide office access for employees would overcome this inconvenience.<sup>23</sup>
- **Security and safety.** Faces are a strong biometric tool, partly because faces are a unique identifier that cannot be accidentally misplaced by an individual. The fact that FRT does not require physical contact with a device also offers some public health benefits. For example, the use of FRT systems for contactless identification during the COVID-19 pandemic made some people feel safer and more secure.<sup>24</sup>
- **Efficiency.** FRT Systems tend to be more efficient than conventional, human-based systems, especially for verification or identification tasks involving large numbers of people. For example, in 2020, police in New Delhi stated that they used a facial recognition tool to identify almost 3000 missing children in four days.<sup>25</sup>

- **Healthcare and accessibility.** Facial recognition technology can be used to advance healthcare and improve accessibility for people with disability. For example, FRT may be used as a powerful diagnostic tool that can identify features of a genetic disorder that clinicians would otherwise miss, either because the relevant physical characteristic is so difficult to perceive visually, or because clinicians lack educational resources on how these symptoms present in diverse, non-European populations.<sup>26</sup> FRT has also been used to monitor newborns for quiet or non-expressed pain which can allow for timely medical intervention.<sup>27</sup> In addition, some FRT Applications are used by people who are blind or have low vision to identify the people and objects around them.<sup>28</sup>

Nevertheless, all FRT applications carry at least a base-level risk to human rights. Moreover, the claimed benefits of any particular FRT Application – including by reference to the metrics of convenience, security and safety, efficiency, and health and accessibility noted above – must be considered in the context of the application’s likely error rates as compared with other methods of identification, as well as any other human rights risks such as increasing surveillance and the restriction of the right to privacy.

23. Essential Research, *Facial Recognition Model Law Project: Findings from the qualitative research*, (Report commissioned by the University of Technology Sydney, May 2022) 17.

24. Pauline Norstrom and Anekanta Consulting, ‘Has Covid Increased Public Faith in Facial Recognition?’ (2021) 2021(11-12) *Biometric Technology Today* 5.

25. ‘Delhi Police Tells High Court It Requires More Information from Centre on Missing Children-India News, Firstpost’, *Firstpost* (23 April 2018) <<https://www.firstpost.com/india/delhi-police-tells-high-court-it-requires-more-information-from-centre-on-missing-children-4443161.html>>.

26. Kristina Grifantini, ‘Detecting Faces, Saving Lives’ (2020) 11(2) *IEEE pulse* 2, 3; Jeannine Mjoseh, ‘Facial Recognition Software Helps Diagnose Rare Genetic Disease’, National Human Genome Research Institute (Web Page, 23 March 2017) <<https://www.genome.gov/news/news-release/Facial-recognition-software-helps-diagnose-rare-genetic-disease>>.

27. Kristina Grifantini, ‘Detecting Faces, Saving Lives’ (2020) 11(2) *IEEE pulse* 2, 4.

28. KM Kramer, DS Hedin and DJ Rolkosky, ‘Smartphone Based Face Recognition Tool for the Blind’ (Conference Paper, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 11 November 2010) 4538.

## 4.2. Technical & operational problems associated with FRT

Problems with FRT can occur at multiple stages throughout the design, development and deployment of the technology. Errors and inaccuracies can arise from the technical operation of an FRT Application, including through problems related to training and reference data and the accuracy of an FRT Application's algorithm(s). These problems include:

- **Inaccuracy due to poor quality input data.** While the performance of many facial verification and facial identification algorithms has improved for some demographic groups in recent years, independent testing and benchmarking by bodies such as the US National Institute of Standards and Technology (NIST) demonstrate that FRT Applications commonly produce errors, particularly when low-quality photographs captured in real-world situations are used as input.<sup>29</sup>
- **Algorithm errors and failure rates.** FRT Applications can fail because the underlying algorithm makes mistakes when matching target and reference data or assessing characteristics in the case of facial analysis. Even when algorithms are tested using high-quality photographs for both captured and reference data, assessments of popular, commercially available algorithms find that they typically make mistakes about one to two per cent of the time in *facial verification*

and about three per cent of the time for *facial identification*. Academic models fail at about twice this rate and human assessors tend to be twice as bad again.<sup>30</sup> Furthermore, representations about the overall accuracy or reliability of various types of FRT should be caveated by the fact that within a particular FRT functionality (for example, facial verification or facial identification), there is wide variation in accuracy among different algorithms.<sup>31</sup> Real-world performance with algorithms other than those submitted for testing is almost certainly more variable.

- **Demographic variations in error rates.** Where and when FRT Applications fail can vary depending on the demographic characteristics of affected individuals. Analysis by NIST and others has shown large variation in error rates across demographic groups.<sup>32</sup> In general, computer-based algorithms make more errors at verification and identification when the relevant captured and reference data relate to dark-skinned people, women and people with disabilities.<sup>33</sup>

These problems are described in greater detail in Box 1 below.

29. See for example 'Face Recognition Vendor Test (FRVT)', NIST (Web Page, 30 November 2020) <<https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt>>.

30. Samuel Dooley et al., 'Comparing Human and Machine Bias in Face Recognition' (2021) arXiv:2110.08396v2, arXiv, Cornell University <<http://arxiv.org/abs/2110.08396>>.

31. NIST's most recent report on the performance of 293 one-to-many identification algorithms in controlled testing conditions found false negative error rates in test scenarios ranging from a few tenths of one percent to beyond fifty percent, leading NIST to conclude, 'This large accuracy range is consistent with the buyer-beware maxim.' Patrick Grother, Mei Ngan and Kayee Hanaoka, *Face Recognition Vendor Test (FRVT) Part 2: Identification* (No NIST IR 8271, National Institute of Standards and Technology, (Report, NIST IR 8271, 28 July 2022) <<https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8271.pdf>> ('Face Recognition Vendor Test (FRVT) Part 2').

32. In 2019, NIST analysis found a factor of 100 more false positive errors between individuals from different countries on identification tasks across high quality passport application photos. Patrick Grother, Mei Ngan and . Patrick Grother, Mei Ngan and Kayee Hanaoka, *Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*, National Institute of Standards and Technology (Report, NIST IR 8280, December 2019) <<https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf>> ('Face Recognition Vendor Test Part 3').

33. Samuel Dooley et al., 'Comparing Human and Machine Bias in Face Recognition' (2021) arXiv:2110.08396v2, arXiv, Cornell University <<http://arxiv.org/abs/2110.08396>>.

Errors and other problems in how an FRT Application is deployed or used by individuals and organisations can produce errors in outputs and outcomes. Common problems include:

- **Overuse.** Even where a particular FRT use case would not, of itself, be inconsistent with human rights, there is growing public concern about the sheer *amount* of FRT being used in Australia and other countries. This typically affects some demographic groups more than others. For example, research focused on the US city of Detroit showed high rates of police use of FRT in areas with high concentrations of African-American residents, as compared with white or Asian residents.<sup>34</sup>
- **User error.** ‘User error’ arises where an FRT Application is deployed incorrectly by an FRT Deployer, or where there are inadequate controls to check for accurate functioning of the application.
- **System error.** ‘System error’ refers to the situation where an FRT System – through poor design, operation or implementation – produces errors. A hypothetical example of a system error would be if an FRT System correctly identifies individuals using FRT but combines these identification outputs with incorrect information about the individuals’ income to produce decisions that incorrectly deny people social welfare.
- **Abuse.** This refers to deliberate misuse, such as covertly using FRT to track and monitor a current or former intimate partner; mischaracterising the outputs of an FRT Application to advance an FRT Deployer’s ulterior motive; or deliberately using FRT for an unlawful or otherwise illegitimate purpose, such as racial profiling.

---

*While the performance of many facial verification and identification algorithms has improved, FRT Applications commonly produce errors when the quality of photographs used in real-world situations is low.*

---

34. Detroit Community Technology Project, *A Critical Summary of Detroit’s Project Green Light and Its Greater Context* (Report, 9 June 2019).

## Box 1: the problem of technical inaccuracy in FRT Applications

Some one-to-many FRT Applications have been less accurate when identifying women, people of colour, young people, and people with a disability, for a range of reasons including a lack of diverse representation in algorithm training data sets.<sup>35</sup> Recent testing by the US's National Institute of Standards and Technology (NIST) shows that the highest performing algorithms have error rates as low as 0.1%, while historical analysis shows that the accuracy of facial identification is continually improving.<sup>36</sup> However, such tests are performed on relatively high-quality data sets in laboratory conditions.

Accuracy declines substantially in 'real world' use, with FRT error rates reaching over 20% when the same algorithms are presented with poorer quality images.<sup>37</sup> Furthermore, the rate of false positive errors remains higher among people of colour and women, underscoring concerns about unfairness or even unlawful discrimination for affected individuals in these groups when these FRT outputs are used to make decisions.<sup>38</sup>

Even with accuracy improving considerably, an accuracy rate of anything less than 100% may still result in many people being misidentified or unrecognised in contexts where FRT is widely adopted and deployed. In high-stakes use cases, the impacts of these errors can be harmful and irreversible.<sup>39</sup> While humans can also make errors when manually identifying people, a key difference is the scalability of errors which can occur en masse when FRT Systems are used.

Even apparently-small error rates can be problematic. Imagine if an FRT Application were used to scan the faces of travellers moving through a major airport such as Sydney over the course of a week. This would be approximately one million people.<sup>40</sup> Even if the Application had an error rate of only 1%, this could be expected to result in 10,000 individuals being the subject of an incorrect result.

Additionally, there are significant accuracy concerns related to the use of facial analysis across all demographics.<sup>41</sup> People with disability can be particularly prone to errors through facial analysis. For example, the use of some facial analysis applications in the remote proctoring of online exams or in monitoring students for aggressive behaviour has been found to incorrectly flag people with disabilities as suspicious or threatening because they may struggle to regulate their movements and body language.<sup>42</sup> Notably, in June 2022, Microsoft announced it would restrict the sale of its facial analysis products due to 'the lack of scientific consensus on the definition of 'emotions,' the challenges in how inferences generalize across use cases, regions, and demographics, and the heightened privacy concerns around this type of capability'.<sup>43</sup>

35. Joy Buolamwini and Timnit Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification' (Conference Paper, Conference on Fairness, Accountability and Transparency PMLR 81, 2018) 77; K. S. Krishnapriya et al, 'Characterizing the Variability in Face Recognition Accuracy Relative to Race' (Conference Paper, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019) 2278; Inioluwa Deborah Raji and Joy Buolamwini, 'Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products,' (Conference Paper, AAAI/ACM Conference on AI, Ethics, and Society, Association for Computing Machinery, 2019) 429.

36. Patrick Grother, Mei Ngan and Kayee Hanaoka, *Face Recognition Vendor Test (FRVT) Part 2: Identification* (No NIST IR 8271, National Institute of Standards and Technology, September 2019) 6.

37. Patrick Grother, Mei Ngan and Kayee Hanaoka, *Face Recognition Vendor Test (FRVT) Part 2: Identification* (No NIST IR 8271, National Institute of Standards and Technology, September 2019) 6.

38. Patrick Grother, Mei Ngan and Kayee Hanaoka, *Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*, National Institute of Standards and Technology (Report, NIST IR 8280, December 2019) 7-8.

39. Patrick Grother, Mei Ngan and Kayee Hanaoka, *Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*, National Institute of Standards and Technology (Report, NIST IR 8280, December 2019) 11.

40. Sydney Airport, *Market Summaries* (Web Page, 2022) <<https://www.sydneyairport.com.au/corporate/partner-with-us/aviation-opportunities/market-summaries>>.

41. Ada Lovelace Institute, *Countermeasures: The Need for New Legislation to Govern Biometric Technologies in the UK* (Report, June 2022) 25.

42. Lydia XZ Brown et al, *Ableism And Disability Discrimination in New Surveillance Technologies: How New Surveillance Technologies in Education, Policing, Health Care, and the Workplace Disproportionately Harm Disabled People* (Report, The Center for Democracy & Technology (CDT), May 2022) 8-9, 23.

43. Natasha Crampton, 'Microsoft's Framework for Building AI Systems Responsibly', *Microsoft on the Issues* (Blog Post, 21 June 2022) <<https://blogs.microsoft.com/on-the-issues/2022/06/21/microsofts-framework-for-building-ai-systems-responsibly/>>.



## 4.3. Human rights risks

A number of human rights can be engaged and limited by the development and use of FRT. Most obviously, the right to privacy is limited in almost all uses of FRT. Part 4 of the report summarises some common risks associated with FRT for privacy and other human rights.

### 4.3.1. The right to privacy

Face data – which can convey biometric information that reveals an individual's identity, racial or ethnic origin and health status – is sensitive information under the Privacy Act.<sup>44</sup> Whenever FRT is used without the consent of the affected individual or for purposes that are not lawful, it can breach the individual's privacy.

The increasing use of FRT in public and commercial places necessarily limits the privacy of affected individuals and the community more broadly. The cumulative increase in FRT corresponds with a cumulative intrusion on the right to privacy. Ultimately, this can result in mass surveillance. Not only does mass surveillance breach the right to privacy, it also can have a chilling effect on other rights, such as freedom of association and assembly, and freedom of expression and opinion.

For example, during the 2019 Hong Kong democracy protests, citizens covered their faces with masks<sup>45</sup> and destroyed 'smart lampposts' embedded with FRT surveillance capabilities<sup>46</sup> to protect their identities and their right to protest. Where FRT enables automated review of CCTV content, law enforcement can bypass the resource-constraints that previously prevented such intrusive, wide-scale monitoring, resulting in mass or constant surveillance.<sup>47</sup>

A related problem, known as 'lateral surveillance', occurs when individuals (as distinct from governments or corporations) use FRT in ways that limit the privacy and other rights of their peers, such as family, co-workers and other individuals with whom they come into contact. Lateral surveillance can be especially pernicious when used deliberately to cause harm, including to stalk, intimidate or harass. This is not a new problem,<sup>48</sup> but the growing availability of technology such as FRT that can enable lateral surveillance means that the problem is growing.<sup>49</sup>

---

*Not only does mass surveillance breach the right to privacy, it also can have a chilling effect on other rights, such as freedom of association and assembly, and freedom of expression and opinion.*

---

44. *Privacy Act 1988 (Cth)* s 6(1).

45. Paul Mozur, 'In Hong Kong Protests, Faces Become Weapons', *The New York Times* (Online, 26 July 2019) <<https://www.nytimes.com/2019/07/26/technology/hong-kong-protests-facial-recognition-surveillance.html>>.

46. Zak Doffman, 'Hong Kong Exposes Both Sides Of China's Relentless Facial Recognition Machine', *Forbes* (Online, 26 August 2019) <<https://www.forbes.com/sites/zakdoffman/2019/08/26/hong-kong-exposes-both-sides-of-chinas-relentless-facial-recognition-machine/>>.

47. Attorney-General's Department (Cth), *Report of the Comprehensive Review of the Legal Framework of the National Intelligence Community* (Report Vol 3, 4 December 2020) 190, 193.

48. See, eg, Mark Andrejevic, 'The Work of Watching One Another: Lateral Surveillance, Risk, and Governance' (2005) 2(4) *Surveillance & Society* 479.

49. See, eg, Thorin Klosowski, 'Facial Recognition Is Everywhere. Here's What We Can Do About It', *New York Times* (Online, 15 July 2020) <<https://www.nytimes.com/wirecutter/blog/how-facial-recognition-works/>>.

### 4.3.2. Other human rights

Facial recognition technology can limit a range of other human rights, especially:

- **The right to equality or non-discrimination.** When the accuracy of an FRT System or Application disadvantages certain demographic groups, this can lead to unlawful discrimination.<sup>50</sup> For example, in the *Bridges Case*, the Court of Appeal of England and Wales acknowledged the potential discriminatory impact of FRT and found that, by failing to investigate the possibility of discrimination, the South Wales Police had breached its statutory duty under the *Equality Act 2010* (UK).<sup>51</sup> Additionally, the Australian Human Rights Commission and the UN Committee on the Elimination of Racial Discrimination have warned that FRT can be used for a particular form of unlawful discrimination, known as ‘profiling’, where individuals are treated less favourably based on characteristics such as their race, colour, national or ethnic origin, or gender.<sup>52</sup>
- **The right not to be subject to arbitrary arrest or detention.** Especially when FRT is used by law enforcement to identify criminal suspects, false positive matches through the use of FRT Systems can lead to arbitrary arrest or detention. Examples of this phenomenon are starting to emerge overseas.<sup>53</sup>
- **The rights to equality before the law and to a fair trial.** Where the outputs of FRT Applications and Systems cannot be interrogated for accuracy, this can threaten the right to equality before the law and the right to a fair trial, when these outputs are used in legal proceedings.

---

*Facial recognition technology does not only affect the right to privacy. It can also restrict the right to equality and other human rights.*

---

50. See, e.g. Joy Buolamwini and Timnit Gebru, ‘Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification’ (Conference Paper, Conference on Fairness, Accountability and Transparency PMLR 81, 2018) 77; Inioluwa Deborah Raji and Joy Buolamwini, ‘Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products,’ (Conference Paper, AAAI/ACM Conference on AI, Ethics, and Society, Association for Computing Machinery, 2019) 429.

51. *R (on the application of Edward Bridges) v South Wales Police* [2020] EWCA Civ 1058, [210].

52. Australian Human Rights Commission, *Human Rights and Technology Final Report*, (Report March 2021) 115; UN Committee on the Elimination of Racial Discrimination, General Recommendation No 36: Preventing and Combating Racial Profiling by Law Enforcement Officials, CERD/C/GC/36 (24 November 2020) [35]-[36].

53. See, e.g., the ongoing legal proceedings filed by Robert Williams against the City of Detroit, Michigan after he was incorrectly matched with a criminal suspect by a facial recognition tool and arrested.



Part 5.

# Facial recognition policy in the Australian & global context



## 5.1. A comparative law analysis

Broadly speaking, there are three approaches to regulating FRT in Australia and other jurisdictions.

The first approach relies on privacy law and other existing laws, but does not deal explicitly or comprehensively with FRT. For example, the *Privacy Act 1988* (Cth) limits the collection and use of sensitive information, which includes facial images and other biometric data, but it does not expressly refer to FRT.<sup>54</sup> It leaves many aspects of FRT development and use unregulated.

The second approach involves issuing a moratorium that prohibits FRT in certain situations. For example, some jurisdictions have prohibited certain uses of FRT by law enforcement or the public sector. (These prohibitions remain subject to some, albeit narrow exceptions, so perhaps they would be best described as *limited* moratoria.) For example, the US State of Vermont has prohibited the use of FRT by law enforcement except where the technology is used to identify a victim or suspect of child sexual abuse who is already in police custody, using images that were already legally seized for that specific investigation.<sup>55</sup> On a municipal level, the City of San Francisco prohibits the use of FRT by any city officials or departments, including local law enforcement, in the absence of prior approval by the Board of Supervisors.<sup>56</sup>

The third approach regulates FRT directly. Jurisdictions that have passed this sort of FRT law include the US States of Washington and Illinois. The European Union's (EU) draft Artificial Intelligence Act presents another such example. Common features of this sort of law include:

- **Human oversight & review.** Under Washington State law, all decisions using FRT, which produce legal effects, must be subject to 'meaningful human review'.<sup>57</sup> Similarly, the EU draft Artificial Intelligence Act contains human oversight requirements aimed at minimising the risks to health, safety, and fundamental rights inherent in the operation of a high-risk AI system, or those that could arise from reasonably foreseeable misuse.<sup>58</sup>
- **Restriction of high-risk uses of FRT.** Washington State law prohibits high-risk, real-time FRT identification or tracking in the absence of 'exigent circumstances' and judicial authorisation.<sup>59</sup> That law also prohibits targeted use of FRT based on race, or political and religious affiliation, and requires independent testing of FRT software in its operational conditions to identify any unintended discriminatory results.<sup>60</sup> The EU draft Artificial Intelligence Act also prohibits the use of 'real-time' remote biometric identification systems in public spaces for law enforcement purposes unless it is targeted, or deployed to identify a victim or suspect of a serious offence or prevent an imminent threat to life or physical safety.<sup>61</sup> A deployment under one of these exception must balance the seriousness of the situation with the consequences of deployment on the rights and freedoms of individuals.<sup>62</sup> Any such deployment must also comply with 'necessary and proportionate safeguards and conditions in relation to use'.<sup>63</sup>

54. *Privacy Act 1988* (Cth), s 6(1) (definition of 'sensitive information').

55. Act of 7 October 2020, § 14 No 166 Vt Acts & Resolves (2020); Act of 4 May 2021, § 1 No 17 Vt Acts & Resolves (2021).

56. *Administrative Code* (San Francisco, California), ch. 19B, § 19B.2.

57. Wash Rev Code, § 43.386.030 (2020).

58. *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts 2021/0106/COD*, art. 14.

59. Wash Rev Code, § 43.386.080 (2020).

60. Wash Rev Code, § 43.386.080 (2020).

61. *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts 2021/0106/COD*, art. 5.

62. *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts 2021/0106/COD*, art. 5.

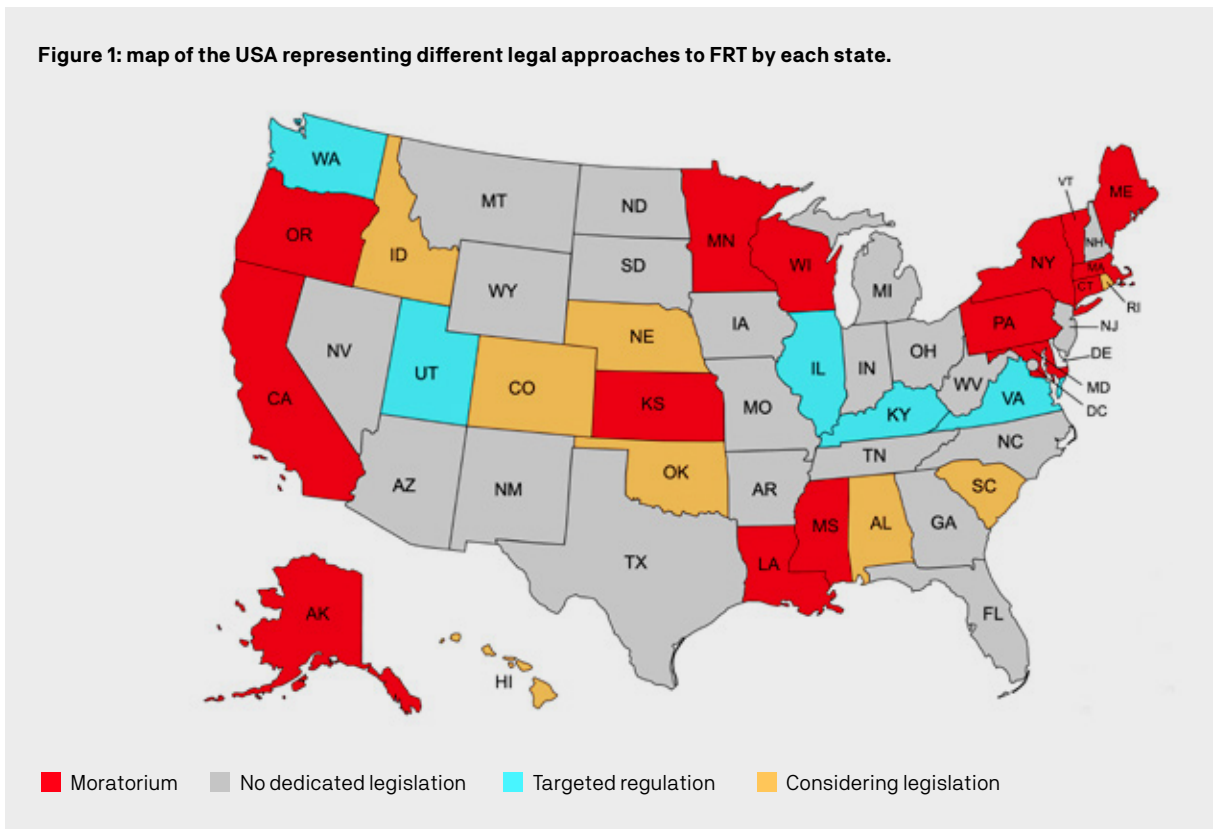
63. *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts 2021/0106/COD*, art. 5.

- Regulating private sector use.** The US State of Illinois has a law directed at private-sector use of biometric information, with strict consent requirements for the collection, disclosure, and trading of an individual’s biometric information.<sup>64</sup> This legislation was invoked by the American Civil Liberties Union against FRT Developer, Clearview AI, over the company’s failure to obtain consent before sharing biometric information. The case was settled in May 2022, with Clearview AI agreeing to stop selling its service to private companies in the United States, and to no longer work with Illinois public-sector bodies for five years.<sup>65</sup> In 2021, the Supreme People’s Court of China issued guidelines acknowledging the need for the private sector to secure informed consent prior to using an FRT Application.<sup>66</sup>

- Remedies and enforcement.** The Illinois law includes a cause of action for breaches of biometric technology restrictions in its law.<sup>67</sup> However, this is not true of all such laws. The Washington State branch of the American Civil Liberties Union strongly opposed that state’s law for failing to establish sufficient enforcement mechanisms.<sup>68</sup>

The movement for FRT law reform is growing, especially in North America. The map in Figure 1 summarises the legal approaches to FRT within the United States.

**Figure 1: map of the USA representing different legal approaches to FRT by each state.**



64. *Biometric Information Privacy Act*, 740 Ill Comp Stat 14/1 § 15 (2007).  
 65. For further information, see Drew Harwell, ‘Clearview AI to stop selling facial recognition tool to private firms’, *The Washington Post* (online, 9 May 2022) <<https://www.washingtonpost.com/technology/2022/05/09/clearview-illinois-court-settlement/>>.  
 66. Supreme People’s Court (People’s Republic of China), Legal Interpretation [2021] No. 15: Regarding the handling of personal information using facial recognition technology, Article 4, <<https://www.court.gov.cn/fabu-xiangqing-315851.html>>.  
 67. *Biometric Information Privacy Act*, 740 Ill Comp Stat 14/1 § 20 (2007).  
 68. Jennifer Lee, ‘We Need a Face Surveillance Moratorium, Not Weak Regulations: Concerns about SB 6280’, *ACLU Washington* (Web Page, 31 March 2020) <<https://www.aclu-wa.org/story/we-need-face-surveillance-moratorium-not-weak-regulations-concerns-about-sb-6280>>.

## 5.2. Australian law applicable to FRT

Australia does not currently have a law dedicated to regulating FRT. As summarised below, existing privacy, anti-discrimination, and state-level human-rights laws impose some limited regulation on the development and use of FRT.

### 5.2.1. Privacy law

The *Privacy Act 1988* (Cth), especially through the Australian Privacy Principles (APPs), regulates the collection, use and disclosure of personal information by government agencies and private sector organisations.<sup>69</sup> However, there are many exceptions to the requirements in the Privacy Act, and certain categories of organisation – including small business organisations, media organisations and political parties – are exempt from complying with the Act as a whole.

As previously noted, face data, as a type of biometric data, is ‘sensitive information’ and attracts additional protections beyond those applicable to personal information more generally.<sup>70</sup>

Individual consent is generally required for the collection of sensitive information, subject to some limited exceptions.<sup>71</sup> Notably, law enforcement bodies, such as the Australian Federal Police, may collect sensitive information without consent if the collection is ‘reasonably necessary for’ or ‘directly related to’ their functions or activities.<sup>72</sup> This is a broad exception. It adds to community concern that excessive use of FRT by law enforcement can increase the risk of mass surveillance and other human rights concerns.

The APPs also contain other requirements, such as to take reasonable steps to notify an individual whose personal information has or will be collected, limitations on using personal information for a purpose beyond that for which it was collected, and prohibitions on using or disclosing sensitive information about an individual for the purpose of direct marketing.<sup>73</sup>

Government agencies must conduct a privacy impact assessment for projects that will involve new or changed ways of handling personal information resulting in significant privacy impacts. It is likely that the use of facial recognition technology would fall within this definition.

To date, there has been some, albeit limited, regulatory action taken by reference to Australian privacy law in respect of FRT. Perhaps the most significant action related to the activities of the company, Clearview AI, and the Australian Federal Police (see the case study in Box 2). In addition, on 12 July 2022 the OAIC announced it had opened investigations into the personal information handling practices of two companies (Bunnings Group Limited and Kmart Australia Limited), following an investigation by consumer advocacy group CHOICE regarding these retailers’ use of FRT.<sup>74</sup>

69. *Privacy Act 1988* (Cth) s 6(1) (definition of ‘personal information’).

70. *Privacy Act 1988* (Cth) s 6(1) (definition of ‘sensitive information’).

71. *Privacy Act 1988* (Cth) sch 1, sub-cl 3.4.

72. *Privacy Act 1988* (Cth) sch 1, sub-cl 3.4(d)(ii).

73. *Privacy Act 1988* (Cth) sch 1, cls 5–7.

74. Office of the Australian Information Commissioner, ‘OAIC opens investigations into Bunnings and Kmart’ (Media Release, 12 July 2022).

## Box 2: Case study – police use of Clearview AI's face-matching service

Clearview AI, a company based in the United States, is an FRT Developer that offers its customers, generally law enforcement agencies in a range of countries, access to its FRT Application. For a period, the Australian Federal Police (AFP) was a customer of Clearview AI, using its FRT Application. In 2021, the Australian Information and Privacy Commissioner made two separate but related determinations, following her investigations into the actions of Clearview AI and the AFP.

The first investigation concerned the actions of Clearview AI, which had created a database of over three billion images of people, with many of these images obtained by the company 'scraping' images from social media and other websites. The Information Commissioner found that Clearview AI did not obtain the consent of affected individuals to create this database, and that in a number of respects in offering its service to the AFP, Clearview AI had failed to comply with its obligations under the Privacy Act.<sup>75</sup> The Information Commissioner ordered that Clearview AI take steps to cease its actions that had resulted in privacy breaches.

The second investigation concerned the AFP as a user of Clearview AI's FRT Application. The Information Commissioner found that, in the circumstances, the AFP was obliged to conduct a privacy impact assessment prior to its use of Clearview AI's service – something it had failed to do.<sup>76</sup> In addition, the Information Commissioner found that the AFP had taken inadequate steps to protect the privacy of Australians in contravention of APP 1.2.<sup>77</sup> The Information Commissioner ordered the AFP to cease using this service.

*Australia's privacy law provides limited protection against harmful FRT practices.*

75. *Commissioner Initiated Investigation into Clearview AI, Inc (Privacy)* [2021] AICmr 54 (14 October 2021), [76].

76. *Commissioner Initiated Investigation into the Australian Federal Police (Privacy)* [2021] AICmr 74, [76]. This failure to conduct a privacy impact assessment was found to amount to a breach of the Privacy (Australian Government Agencies – Governance) APP Code 2017.

77. *Commissioner Initiated Investigation into the Australian Federal Police (Privacy)* [2021] AICmr 74, [94].



### 5.2.2. Other relevant Australian laws

At the time of writing this report, there has been no significant litigation or regulatory action taken in respect of FRT outside of the privacy law context. Nevertheless, it is at least theoretically possible that a range of other Australian laws could be invoked in upholding human rights in the context of the development and use of FRT.

Those other laws include:

- **Anti-discrimination law.** Federal, state and territory legislation prohibit discrimination on a range of grounds including race, sex, disability, and age.<sup>78</sup> Where an FRT System unfairly disadvantages individuals by reference to any of these protected attributes, because the FRT System relies on an FRT Application that is disproportionately inaccurate in identifying people from those groups, it is at least arguable that using this system would contravene relevant anti-discrimination law.

It would have been beyond the Australian Information Commissioner's remit to make findings regarding anti-discrimination law in her determination regarding Clearview AI; however, she did refer to the disproportionate burden of FRT identification errors falling on certain groups, such as people of colour.<sup>79</sup>

- **Human rights legislation.** While Australia does not have a national human rights act or bill of rights, three Australian states and territories have enacted statutory human rights act. They contain a number of legal protections applicable primarily of bodies performing government functions in those particular states and territories. It is at least arguable that if one of those governments used FRT in a way that infringed human rights, such as the right to a fair trial, it could be found to contravene a state or territory human rights act.<sup>80</sup>

## 5.3. Voluntary action by FRT Developers

Some FRT Developers and Deployers – that is, organisations that sell and use FRT Applications – have voluntarily limited or ceased their commercial activities regarding FRT. Notably, in 2020, FRT Developers, Amazon and Microsoft, voluntarily ceased the sale of FRT Applications to law enforcement agencies in the United States, until appropriate federal regulation is introduced in that country.<sup>81</sup> IBM also withdrew entirely from the production of FRT Applications.<sup>82</sup> In 2021, Meta (the owner of Facebook) shut down its automated facial recognition feature, citing the absence of clear law that balances growing societal concern regarding facial recognition with its benefits.<sup>83</sup>

Other FRT vendors have responded to increasing criticism by publishing guidelines or principles that govern their design of the software. For example, Thales stated in 2021 that it would apply ethical principles, such as transparency, consent, security, and precision among other principles, in its design of FRT Applications.<sup>84</sup>

Many FRT Developers and Deployers have not taken such voluntary action, and there is as yet no industry-wide code of conduct governing the development and use of FRT by the private or public sectors.

78. The key federal legislation is: *Racial Discrimination Act 1975* (Cth); *Sex Discrimination Act 1984* (Cth); *Disability Discrimination Act 1992* (Cth); *Age Discrimination Act 2004* (Cth).

79. *Commissioner Initiated Investigation into Clearview AI, Inc (Privacy)* [2021] AICmr 54, [197], [207]–[217].

80. See *Charter of Human Rights and Responsibilities Act 2006* (Vic); *Human Rights Act 2004* (ACT); *Human Rights Act 2019* (Qld).

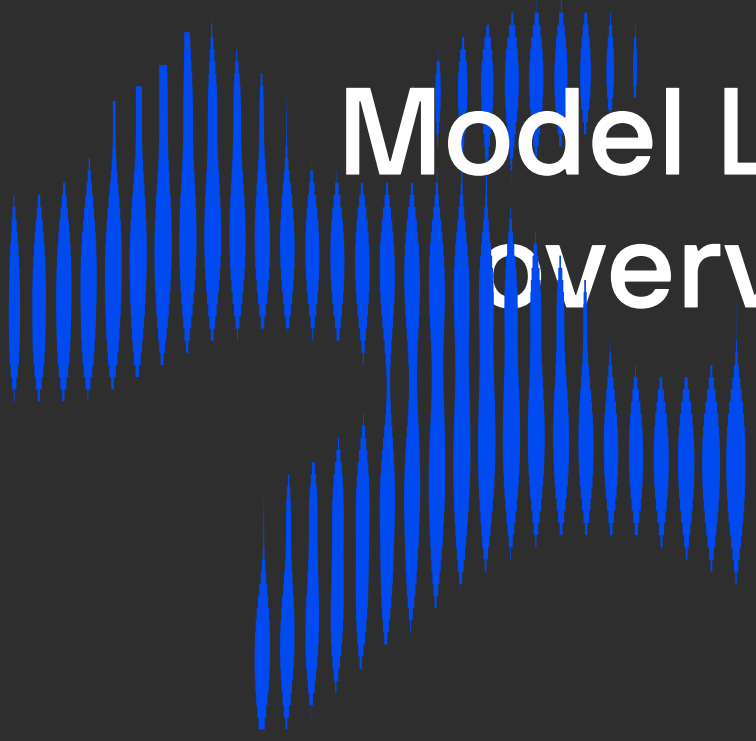
81. Jay Greene, 'Microsoft won't sell police its facial-recognition technology, following similar moves by Amazon and IBM', *The Washington Post* (Online, June 11 2020) <<https://www.washingtonpost.com/technology/2020/06/11/microsoft-facial-recognition/>>.

82. Jay Greene, 'Microsoft won't sell police its facial-recognition technology, following similar moves by Amazon and IBM', *The Washington Post* (online at June 11 2020) <<https://www.washingtonpost.com/technology/2020/06/11/microsoft-facial-recognition/>>.

83. Jerome Pesenti, 'An Update On Our Use of Face Recognition', *Meta Newsroom* (Blog Post, 2 November 2021) <<https://about.fb.com/news/2021/11/update-on-use-of-face-recognition/>>.

84. Thales Group, *Designing an ethical, socially accountable facial recognition system: A vision from Thales (Report 2021)* 10.

Part 6.



# Model Law: overview

Part 6 provides an overview of the proposed Model Law and gives some guidance in reading the remainder of the report.

## 6.1. To whom does the Model Law apply?

The Model Law imposes new legal obligations on persons developing and deploying or using FRT in Australia, and to those developing FRT outside Australia if their technology is marketed or offered for use in Australia.

A person is an *FRT Developer* if they: (a) develop an FRT application intended for active use by themselves or others in a substantial form (for example, beyond a mere sub-component or non-functional system), or (b) offer, market or distribute an FRT application for use in Australia.

Examples of FRT Developers include:

- an Australian startup offering access to a novel, pre-trained FRT algorithm via an application programming interface (API)
- a European-based technology firm selling access to a cloud-based suite of many different FRT Applications
- a Japanese company selling FRT verification software and hardware specifically for use at building entries in Australia.

A person is an *FRT Deployer* if they use or deploy an FRT Application on one or more affected individuals. Any natural person whose face data is captured or processed by an FRT Application is an affected individual.

Examples of FRT Deployers include:

- a person purchasing or licensing FRT software as part of a computer system that uses facial recognition to determine whether someone is authorised to enter a building.
- an individual using an application to categorise their digital photos by whose faces appear in each picture.

An FRT Developer that also directly deploys their FRT Application on affected individuals would be both an FRT Developer and an FRT Deployer.

---

*The Model Law imposes new legal obligations on persons developing and deploying or using FRT in Australia.*

---

## 6.2. Outline of the Model Law

Parts 7 to 10 of this report explain the operation of the proposed Model Law for FRT. Figure 2 provides a summary of the Model Law's key elements, as well as references to which Parts of the report these issues are elaborated upon.

**Figure 2: a summary of the key elements of the Model Law outlined in this report**

	Report Part
<b>Scope of the law</b>	Purpose of the Model Law ..... <b>2.1</b>
	The significance of face data ..... <b>2.2</b>
	Definitions of key terms ..... <b>2.4</b>
<b>Risk assessment approach</b>	Five factors relevant for assessing FRT Applications ..... <b>7.2</b>
	Risk assessment for FRT Application: base, elevated or high ..... <b>7.3</b>
<b>Justification for any human rights restriction(s)</b>	Specification of human rights restricted by FRT Application ..... <b>7.4.1</b>
	Need for legitimate aim for any rights restriction ..... <b>7.4.2</b>
	Test: Is any rights restriction reasonable, necessary and proportionate? ..... <b>7.4.3</b>
<b>Facial Recognition Impact Assessment (FRIA) process</b>	FRT Developers & Deployers must undertake a FRIA, subject to exceptions ..... <b>8</b>
	FRIA Step 1 - use declaration and risk assessment ..... <b>8.1</b>
	FRIA Step 2 - risk management declaration ..... <b>8.2</b>
	Must register completed FRIA with regulator and make available ..... <b>8.3</b>
<b>Legal requirements: base-level and elevated risk applications</b>	Legal requirements for base-level risk applications and above ..... <b>9.2</b>
	Creation of a technical standard for FRT ..... <b>9.2.3</b>
	Legal requirements for elevated risk applications ..... <b>9.3</b>
<b>Legal requirements for high risk FRT applications</b>	General prohibition on high-risk FRT Applications, subject to three exceptions ..... <b>9.4.1</b>
	Exception 1: regulator authorisation process ..... <b>9.4.2</b>
	Exception 2: special legal rules for law enforcement and national security ..... <b>9.4.3</b>
	Exception 3: genuine research uses within an established ethics and governance framework ..... <b>9.4.4</b>
<b>Independent review</b>	Regulator has power to conduct, of its own motion, audits and reviews of FRIAs ..... <b>10</b>
	Affected individuals have standing to make a complaint to regulator ..... <b>10</b>
	Judicial review available for decisions by regulator ..... <b>10</b>



Part 7.



# Model Law: human rights risk assessment

As described in Part 4, the use of FRT can, and generally will, limit human rights. How those rights are limited, and whether that limitation is justified, will depend on a range of factors, including the FRT Application's purpose, functionality, effect on individuals, and the context in which it is used. The Model Law requires these factors to be considered individually and in combination to produce an *overall human rights risk assessment*.

The Model Law requires that this human rights risk assessment should be undertaken through a Facial Recognition Impact Assessment (FRIA), which is described in detail in Part 8. The outcome of this risk assessment process determines the legal requirements that will apply to its development and use (see Part 9).

## 7.1. Human rights vulnerabilities & overall human rights risk

The Model Law sets out a non-exhaustive list of factors for assessing the overall human rights risk of an FRT Application. The Model Law requires that these factors be assessed individually to evaluate particular human rights 'vulnerabilities' associated with the FRT Application. Under the Model Law, any human rights vulnerability may be evaluated as 'moderate', 'significant', or 'extreme'.

The Model Law then requires these human rights vulnerabilities to be considered in combination. This results in an assessment of the overall human rights risk for the particular FRT Application. The Model Law sets out a three-level human rights risk scale: 'base-level', 'elevated' or 'high' risk.

Part 9 of this report sets out the Model Law's requirements for the use of FRT Applications assessed as posing a base-level or elevated risk, and the legal regime for high-risk FRT Applications.

---

*The Model Law requires that this human rights risk assessment should be undertaken through a Facial Recognition Impact Assessment (FRIA).*

---

## 7.2. Factors relevant to the human rights risk assessment

Every use or deployment of FRT will differ. The Model Law therefore provides an inclusive, rather than an exhaustive, list of the factors that are relevant in this risk assessment.

The factors that should be considered include:

1. the *spatial context* in which the FRT Application is expected to be used
2. the *functionality* of the FRT Application
3. the *performance* of the FRT Application
4. whether the FRT Application produces outputs that lead to a decision that has a *legal or similarly significant effect* for an individual or group and, if so, whether the decision *is wholly or partially automated*
5. whether affected individuals can provide *free and informed consent*, or withhold such consent, *prior* to the use of the FRT Application.

These factors, and the associated human rights vulnerabilities, may overlap. For example, both the spatial context and the functionality may create privacy concerns related to consent, which is also assessed in the fifth factor. Such overlap is valuable: it ensures that FRT Deployers and Developers consider potential rights restrictions from multiple perspectives.

Each of the five factors is addressed in turn below.

### 7.2.1. Factor 1: spatial context

‘Spatial context’ refers to the place or environment in which an FRT Application is deployed – such as in public spaces like a public street, or more controlled environments like a workplace. The spatial context is important in assessing any human rights impact of an FRT Application.

It is useful to distinguish between a number of categories of spatial context:

- **Open, publicly-accessible spaces** in Australia include public streets, supermarkets and other shops and public transport. Restricted, semi-public spaces tend to be controlled by a commercial organisation and include sports stadiums, clubs, pubs and cinemas.<sup>85</sup> In these spaces, an individual’s relative anonymity affords a reasonable expectation of privacy – one that allows the individual to participate freely in public life without the threat or reality of constant surveillance. The use of FRT in open, publicly-accessible spaces can therefore restrict civil and political rights, such as the rights to privacy, assembly, expression and freedom of association.
- **Restricted and closed spaces** include workplaces, schools and private clubs, where access is tightly controlled. The use of FRT by a person who exercises some control over this type of space can reduce the ability of individuals to enter, move and act freely.
- **Private spaces** include one’s own home or car. In these spaces, affected individuals generally have a high degree of control over the space and autonomy in how they behave in that space. However, private spaces can also render some people even more vulnerable to rights restrictions, including people in abusive relationships, employees such as carers operating in private homes, or people with disabilities. Hence, while the likely minimum human rights vulnerability for private spaces is moderate, FRT Deployers and Developers should consider the potential misuse of such systems to restrict people with less power in the home.
- **Virtual spaces** include immersive and semi-immersive online environments, such as social media platforms and some video games. In virtual spaces, vulnerabilities tend to be more contingent on other factors – such as decision effects – than in other spatial contexts.

<sup>85</sup> For example, as part of the gambling reform package introduced into South Australia in December 2020, an approved FRT System must be used in venues that operate 30 or more gaming machines. This reform lays out a set of specific legal requirements for FRT Developers and Deployers which govern the use of FRT in these restricted venues to identify barred patrons and self-excluded problem-gamblers. See South Australian Government Consumer and Business Services, ‘Facial Recognition Technology’ (Web Page, 2018) <<https://www.cbs.sa.gov.au/facial-recognition-technology>>.



**Table 1: FRT spatial context factors and vulnerabilities**

<b>Spatial context</b>	<b>Definition</b>	<b>Example</b>	<b>Likely minimum vulnerability</b>
Open, publicly-accessible spaces (public and commercial)	Spaces that are open more or less unconditionally to the public; there is no need for identification to enter.	Public: streets, parks, libraries, government service centres, etc. Commercial: shopping centres, supermarkets, etc.	Extreme
Restricted, semi-public spaces (primarily commercial)	Spaces accessible to members of the public, subject to conditions.	Stadiums, clubs and pubs, theatres.	Significant
Closed spaces	Spaces that are accessible only to a specific, limited number of people.	Workplaces, schools, private clubs, invitation-only events, GPs, some hospitals.	Moderate
Private spaces	Closed spaces not intended for commercial activity, which are controlled by individuals or private entities.	Private residence, car.	Moderate
Virtual spaces	Spaces that exist primarily online.	Video game, social media platform, government service portal.	Moderate: highly dependent on FRT purpose and effect.

*People have a reasonable expectation of privacy – even in open, public spaces.*

### 7.2.2. Factor 2: functionality of the FRT Application

As outlined in Part 2 of this report, FRT Applications can be divided into three broad categories based on their functionality, namely:

- **facial verification**, which compares a captured face to a known reference face, with the output being if there is a match between the two
- **facial identification**, which compares a captured face to a set of reference faces, with the output usually being a subset of reference faces that exceed a similarity threshold
- **facial analysis**, which analyses captured face data in an attempt to detect inherent or behavioural characteristics about the individual.

Where multiple FRT functionalities are applied in a specific operation, the FRT Developer or Deployer must consider the functionality that creates the highest level of vulnerability in assessing the system's risk.

There are inherent risks that generally attach to particular FRT functionalities. Facial verification applications usually can only determine whether an individual's face data matches a single, stored record. This means that facial verification applications can generally be used only to prove or confirm an individual's identity. By contrast, facial identification and facial analysis have a far greater range of potential use cases. This in turn engages a far greater range of human rights, often in more profound ways.

**Table 2: FRT functionality factors & vulnerabilities**

Functionality	Captured faces	Reference faces	Description	Likely minimum vulnerability
Facial verification	1	1	Compares a captured face against a reference held in a biometric token or on a remote database.	Moderate
Facial identification	1 or N	Many	Compares a captured face with a set of previously stored faces in a reference database, to search for a match.	Significant
Facial analysis	1 or N	N/A	Attempts to infer specific attributes of a person from face data, including demographic features, health information, emotions, intentions or behaviours.	Extreme

### 7.2.3. Factor 3: performance of the FRT Application

The third factor considers the performance of the FRT Application, by reference to the relative accuracy with which it can produce reliable results in line with its functionality.

As discussed in Part 4.2., there is wide variation in the accuracy of different FRT Algorithms. Even within a single functional type of FRT, such as facial verification or facial identification, this variation can be significant.

For example, NIST's most recent report on the performance of 293 one-to-many FRT Algorithms found false negative error rates in test scenarios ranging from a few tenths of one per cent to beyond fifty per cent, leading NIST to conclude, 'This large accuracy range is consistent with the buyer-beware maxim'.<sup>86</sup> Real-world performance with algorithms other than those submitted for testing is almost certain to be even more variable.

In addition, it should be observed that facial analysis in particular is highly experimental and therefore very concerning. Its objective accuracy is, at best, unproven. Many facial analysis FRT Applications appear to have very high rates of error or unverifiable outputs when they purport to identify demographic features, classify behaviour or diagnose health conditions. In the absence of very strict conditions, therefore, the use of facial analysis produces extreme risks for affected individuals.

The performance of an FRT Application refers primarily to the accuracy of its outputs, but also how effectively it maintains information security and so on. In practice, an FRT Application that performs poorly generally has high error rates – in other words, it fails to correctly verify, identify or analyse an affected individual at unacceptably-high rates. Such performance problems can affect all demographic groups equally, or they can disproportionately affect certain groups by reference to characteristics like skin colour, gender and age.

When an FRT Application performs poorly, this can result in an individual's human rights being limited or breached – especially where the FRT Application is used to make a decision about the individual.

Where an FRT Application performs poorly – by reference to measures such as accuracy, reliability, consistency and security – the vulnerability to human rights being limited generally increases.

In considering an FRT Application's performance, an FRT Developer or Deployer should take into account not only lab-based testing, but also the specific real-world conditions under which the FRT Application will be deployed, the skill and training of the individuals involved in deploying the Application, the suitability of reference or training data, and any other factors that may lead to errors. Furthermore, to address the risk of unlawful discrimination resulting from an FRT Application's performance, FRT Developers and Deployers should consider the distribution of performance or error rates across a range of affected individuals according to protected characteristics such as age, gender identity, disability, sex and race.

---

*Facial identification and facial analysis have a far greater range of potential use cases. This in turn engages a far greater range of human rights, often in more profound ways.*

---

<sup>86</sup> Patrick Grother, Mei Ngan and Kayee Hanaoka, Face Recognition Vendor Test (FRVT) Part 2: Identification (No NIST IR 8271, National Institute of Standards and Technology, September 2019) 36.

#### 7.2.4. Factor 4: the FRT Application's role in decision making

The fourth factor starts by asking whether the FRT Application produces one or more outputs that create, or materially contribute to, a decision. Where the answer to this question is 'yes' it will be necessary to consider whether the decision has a *legal or similarly significant effect* for an individual or group and whether the decision is *wholly or partially automated*.

This factor draws on Article 22 of the EU's GDPR, which states that an individual 'shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.'<sup>87</sup>

The focus here is on the impact of the use of the FRT Application on individuals. Where the impact is greater, the risk to human rights necessarily rises. Hence, as outlined, three questions need to be considered sequentially in respect of this factor.

##### **1. Does the FRT Application produce one or more outputs that create, or materially contribute to, a decision?**

Decisions tend to be directed towards individuals, and thus are generally more consequential – engaging human rights more directly. Hence, where the outputs of an FRT Application are used materially in a decision-making process, it will be necessary then to consider the other two questions.

For example, where an FRT Application is used in a video game or to give more convenient access to a personal device, it is not materially being used to produce a decision. Generally, this sort of scenario will result in a more minor impact on any affected individual, reducing the risk of their human rights being restricted to a significant degree.

##### **2. If the answer to Question 1 is 'yes', does the decision have a legal or similarly significant effect for an individual or group?**

Decisions that affect individuals' rights and interests also tend to be more consequential and thus engage human rights.

Under Article 22 of the GDPR, a 'legal effect' is something that affects a person's legal status or their legal rights. An FRT Application is used to make a decision that has a legal effect if the FRT Application was a material part of a decision-making process that determines an individual's legal rights. To take a hypothetical example: if a government agency uses an FRT Application to identify an individual, and if this identification determination is a material part of deciding whether the individual is entitled to a particular form of government social welfare, then the FRT Application is used in a decision that has a legal effect.

A 'similarly significant effect' is more difficult to define but, following the EU Data Protection Working Party guidance, it is clear that the threshold for such a decision is one that is 'more than trivial ... the decision must have the potential to significantly influence the circumstances, behaviour or choices of the individuals concerned. At its most extreme, the decision may lead to the exclusion or discrimination of individuals.'<sup>88</sup>

For example, imagine a department store uses an FRT Application to identify individuals on a list of people it has chosen to ban from entering the store. The effect of refusing entry to a particular individual based on an erroneous FRT Application identification may not affect the individual's legal rights, as they did not have an unconditional right of entry to the store, but it would be likely to have affected similarly significant rights of the individual.

87. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1, art 22.

88. Article 29 Data Protection Working Party, European Commission, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679* (2018) 21.

### 3. Was the decision wholly or partially automated?

Automation occurs when a computational system applies algorithms or other rules to a fact scenario in an automatic way. The use of automation in decision making can engage human rights and, in certain situations, the absence of a human in the decision-making process can result in human rights vulnerabilities.

A wholly automated decision-making system in this context produces decisions without any (or any meaningful) human influence over the ultimate outcome. For example, while a human may be involved in data entry or the verification of certain inputs, if human discretion is effectively absent from the system, it can be considered to be wholly automated.

A partially automated system maintains a ‘human in the loop’ with the power to confirm or override a decision with legal or similarly significant effect. Often, this will take the form of a system that produces inferences, predictions, or recommendations for human review: a human then uses this information to make a final decision.

Automation is increasingly being used in a wide range of business processes and by government, with a view to improving the efficiency, consistency and accuracy of routine operations and decisions. FRT Applications often play an essential part of such automated systems, providing a critical authorisation, verification, identification or analysis step which leads directly to an important decision.

There is nothing inherently problematic, in human rights terms, with automated decision making. However, there are several common problems associated with automation – and, by extension, the use of FRT within such automation – when the outcome is a decision that has a legal or similarly significant effect.

First, automation can result in discretion not being effectively applied to a decision. Humans typically are given the responsibility to exercise discretion; automation can make this difficult or impossible. As the Hon Justice Melissa Perry has observed in the context of government decision making, discretion can involve ‘complex and subtle questions’ that may be ‘beyond the capacity of an automated system determine’.<sup>89</sup>

In addition, many automated decision-making systems operate in ways that are difficult for FRT Developers and Deployers, let alone affected individuals, to understand or challenge. This can threaten accountability – a phenomenon sometimes referred to as ‘black box decision making’. That is, when an automated decision has a legal or similarly significant effect, a failure or inability to provide reasons for the decision can make it difficult or impossible to maintain accountability for a decision that is unlawful or incorrect. This can threaten the right to a fair hearing, and it can undermine the right to a remedy where a decision breaches human rights.

This is particularly true for government services where a decision involves the exercise of discretion.<sup>90</sup> As the Commonwealth Ombudsman has pointed out, community expectations of respectful treatment and fairness apply to automated systems, just as they do when a decision is being made manually, by a human being.<sup>91</sup>

89. Melissa Perry, ‘iDecide: Administrative Decision-Making in the Digital World’ (2017) 91(1) *Australian Law Journal* 29, 33.

90. Australian Human Rights Commission, *Human Rights and Technology Final Report*, (Report, March 2021) 56.

91. Commonwealth Ombudsman, *Automated Decision-making: Better practice guide* (Report, 2019) 2.

A third phenomenon associated with automation, and especially with wholly automated decision-making systems, is that decisions can be made at extraordinary pace and scale, with very low marginal cost. As a result, any errors that occur in wholly automated decision-making systems are also liable to be scaled-up to large groups of affected individuals, with the potential to create more widespread harm. While this problem is most acute in respect of wholly automated decisions, even when human review is present in an automated decision-making system, care must be taken to ensure that such review is meaningful.

When an FRT Application produces a significant data point that is used in an automated decision-making system, careful attention should be paid to such problems as the three described above. Where any such problems arise, the human rights vulnerability in respect of this factor would be higher.

Furthermore, even if an incorrect decision with legal or similarly significant effect made by an FRT System can be overturned or corrected ex-post, the effects of this decision on an individual in the interim (prior to rectification of the decision) can cause significant and irreversible harm. For example, if an FRT Application deployed by a government department fails to correctly verify the identity of an individual who is eligible for a welfare service, this individual may be severely disadvantaged by the delay in accessing relevant services while the decision is being reviewed and rectified.

If an FRT Application is used to create a decision with a legal or similarly significant effect, the likely minimum vulnerability would be significant, thereby leading to an assessment of elevated risk. If such a decision is partially or wholly automated, this vulnerability is likely to increase further.

### 7.2.5. Factor 5: prior, free & informed consent

Under Australian privacy law, it is generally necessary to obtain an individual's consent *before* collecting, using or disclosing, and storing their personal information, with stricter requirements where personal information is categorised as sensitive information. This requirement to obtain an individual's consent is not universal: there are many exceptions and exemptions where consent is not required under the *Privacy Act*, such as in law enforcement.

Nevertheless, where such a consent requirement exists, it should be 'free, prior and informed'. The individual should be able to exercise genuine autonomy, which involves their being properly informed of the proposed collection and use prior to the individual's personal information being collected, and the individual should be free to choose to consent or refuse to the proposed collection and use of their personal information.

The importance of consent in the specific context of FRT was emphasised in qualitative research commissioned for this project. The majority of research participants expressed a strong desire to be informed of the use of any FRT Application that affects them, and to be able to provide explicit consent before the collection of any facial data.<sup>92</sup>

However, it is well understood that compliance with such consent requirements in privacy law is erratic. There are many instances, especially in the online environment, where consent for collection and use of personal information is not, in substance, prior, free and informed.

That compliance with consent requirements in this general area of privacy law is problematic does not excuse a failure to obtain consent in respect of the use of biometric technology such as FRT. There have been many law reform proposals to improve practices regarding consent in the privacy law context.<sup>93</sup>

92. Essential Research, *Facial Recognition Model Law Project: Findings from the qualitative research*, (Report commissioned by the University of Technology Sydney, May 2022) 17.

93. See, eg, Australian Law Reform Commission, *For Your Information: Australian Privacy Law and Practice* (Report No 108 Vol 1, May 2008) Ch 19.

While this factor is important, consent should not be seen as the overriding protection in respect of FRT. The use of FRT engages a number of human rights beyond the right to privacy. Outside of the privacy context, purported consent generally does not justify a restriction of rights.

Even where consent is relevant in justifying a human rights limitation, not all individuals have the legal capacity to provide prior, free and informed consent in respect of the collection and use of their personal information. For example, because of an individual's age or cognitive capacity, they may be able to exercise limited or no autonomy in this area. Institutional contexts in which individuals are deprived of their liberty, such as in group care homes, immigration detention facilities and prisons may also restrict people's ability to give free consent. Where an individual lacks the legal capacity to consent, special care should be taken to avoid unnecessary use of FRT.

By way of illustration, there is an increasing push to use FRT in schools and other places where children are present. Either in the Model Law

itself, or in guidance issued by the OAIC, there should be a presumption against use of FRT where it is impractical or otherwise problematic to obtain free, prior and informed consent due to the fact that young, affected individuals have little or no option but to be exposed to FRT.

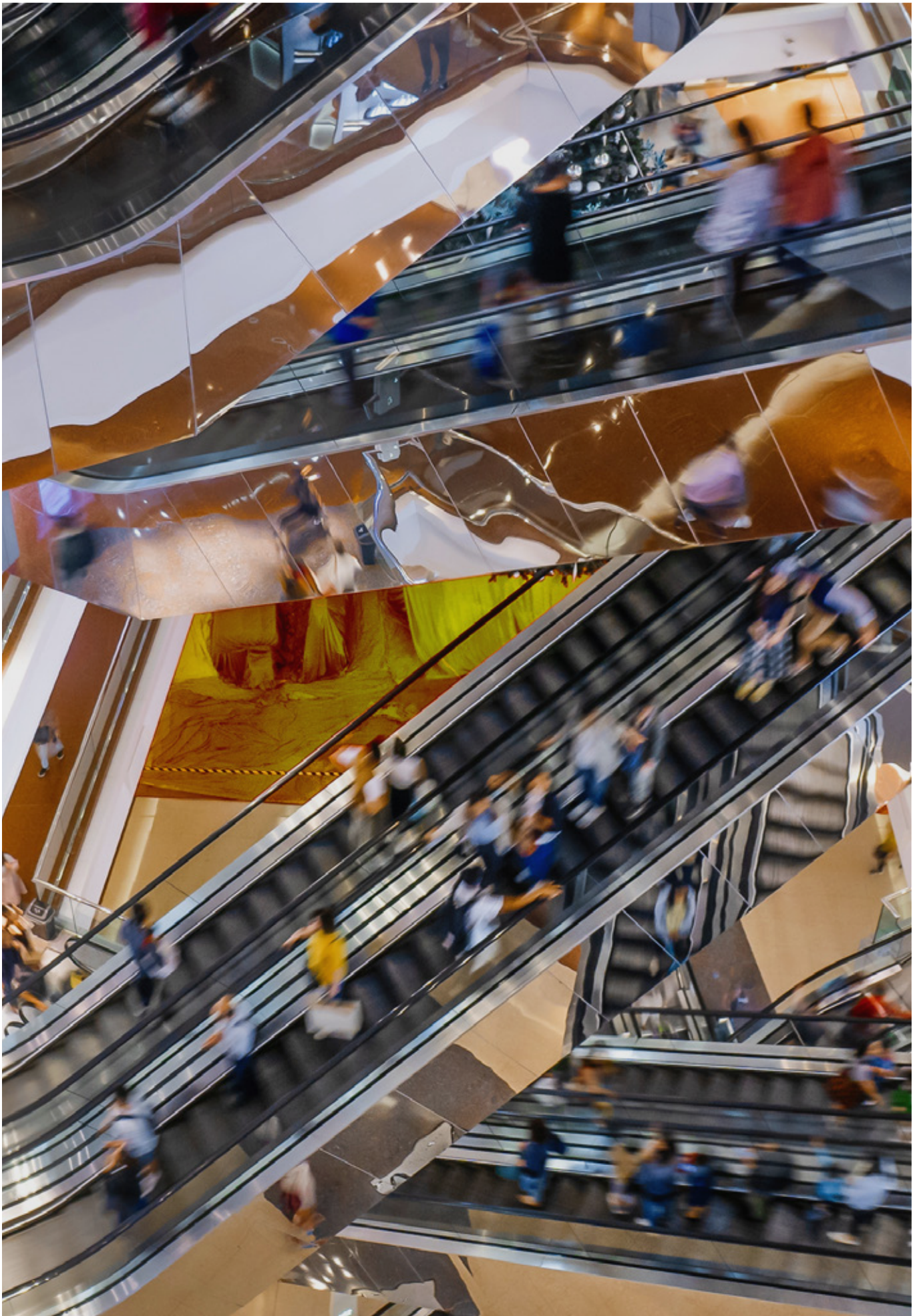
In its submission to the Privacy Act Review, the OAIC warns that obtaining consent is not a simple process. Expanding the circumstances and frequency of occasions in which consent is required can result in 'consent fatigue', undermining the quality of the consent received. Instead, the emphasis should be on defining consent as requiring a 'clear affirmative act that is freely given, specific, unambiguous and informed' and current in the sense that it lasts only as long as is reasonable.<sup>94</sup>

In view of the above, the Model Law does not treat consent as a threshold condition beyond which FRT deployments are lawful. Instead, consent is a necessary condition for ensuring that human rights are not breached by the use of FRT, but it is not the sole condition.

**Table 3: FRT consent factors & vulnerabilities**

Consent category	Definition	Examples	Likely minimum vulnerability
Consent is possible	All affected individuals have had the opportunity to provide, or withhold, prior, free and informed consent.	A bank offers FRT as an alternative to using a password to access a mobile banking application.	Moderate
Consent is problematic	Some affected individuals will not have the opportunity to provide, or withhold, prior, free and informed consent.	A supermarket uses FRT to scan customer faces and identify shoplifters. Signs are posted at the entrance to the store alerting customers to this fact.	Significant
Consent is highly challenging	A large proportion of affected individuals have little or no ability to provide, or withhold, prior, free and informed consent.	Surreptitious use of FRT on a public street, or use of FRT for a government service where affected individuals have no other alternative but to use the system, or an FRT Application deployed in schools with children under the age of 13 years old.	Extreme

94. Office of the Australian Information Commissioner, Government of Australia, 'Part 5: Notice and Consent', *OAIC Website* <<https://www.oaic.gov.au/privacy/the-privacy-act/review-of-the-privacy-act/privacy-act-review-issues-paper-submission/part-5>> ('Part 5')[5.37]-[5.40].





## 7.3. Assessing the overall human rights risk level

An FRT Developer or Deployer must consider each of the factors set out in Part 7.2. individually to evaluate specific human rights vulnerabilities posed by an FRT Application. It is then necessary to consider these factors, and their attendant human rights vulnerabilities, in combination to ascertain the *overall risk level* for the particular FRT Application. This risk level dictates the specific legal requirements that FRT Deployers and Developers must apply to comply with the Model Law, as detailed in Part 9.

As with the evaluation of human rights vulnerability in respect of each specific factor, the overall risk assessment is necessarily a *qualitative* process. On one hand, the assessment cannot be a mechanical process because it involves some exercise of judgment. On the other hand, the assessment is not simply a subjective analysis based on the views of the person or people undertaking the risk assessment. Properly undertaken, the Model Law's risk assessment process involves weighing up the relevant factors individually and in combination, by reference to hard data relating to the use, or likely use, of the relevant FRT Application.

In the event of a legal dispute relating to a particular FRT Application, the regulator or court would apply the same factors and risk framework embodied in the Model Law.

This Model Law provides a rebuttable presumption that the overall risk level of a proposed use should be matched to the highest level of vulnerability evident across each evaluated human rights factor. So, for example, if there are five factors considered in respect of a particular FRT Application, and four of these factors are evaluated as moderate and one is evaluated as extreme, then there should be a presumption that the overall risk level should be the highest available – namely, 'high risk'. However, as this is a rebuttable presumption, in some cases the overall risk rating in such a case would be lower based on the particular circumstances and the likely effect of implementing required and voluntary safeguards.

Consequently, and noting the rebuttable presumption referred to above, each risk assessment process would conclude with one of the following outcomes:

- If there are only *moderate* human rights vulnerabilities, the FRT Application likely would be *base-level risk*.
- If there is one or more *significant* human rights vulnerabilities, but no extreme vulnerabilities, the FRT Application likely would be *elevated risk*.
- If there is one or more *extreme* human rights vulnerabilities, the FRT Application likely would be *high risk*, and therefore prohibited unless an exception applies.

As explained in Part 9, high-risk FRT Applications should be prohibited unless special circumstances – such as limited law enforcement exceptions, research or assistive technology exceptions – apply.

## 7.4. Determining whether human rights limitation is justified

Even where an FRT Application is likely to limit human rights, its use may be permissible if the rights limitation is justified. In completing Step 1 of the FRIA (see Part 8.1.), it is necessary to consider whether there is such a justification. Under international human rights law, which applies to Australia and the vast majority of jurisdictions around the world, this process typically proceeds as follows.

### 7.4.1. Which human rights are being restricted?

It is first necessary to consider which human rights are likely to be restricted or limited by the FRT Application.

By its very nature, any FRT Application will use face data, which is sensitive information. It is almost certain to engage the right to privacy and, depending on the use context and a number of other factors, highly likely to restrict the right to privacy among other possible rights. In this situation, there is a second stage of the human rights risk analysis, as summarised below.

### 7.4.2. Is the human rights restriction legally justified?

Even if an FRT Application restricts human rights, this could still be justified if the following criteria are satisfied:

- **Does the FRT Application restrict only non-absolute rights?** Human rights are either absolute or non-absolute. Absolute rights, such as freedom from torture, can never be restricted or limited. Non-absolute rights can be restricted in certain circumstances, to accommodate other human rights and legitimate interests. The right to privacy is a non-absolute right and can be restricted in a number of situations.

- **Is any restriction of a non-absolute right in pursuit of a legitimate aim?** Non-absolute rights, such as the right to privacy, may be restricted where the restriction is needed to achieve another legitimate aim – such as to protect freedom of expression or uphold community safety.
- **Is any rights restriction reasonable, necessary and proportionate to achieving the legitimate aim?** This criterion is discussed in greater detail in Part 7.4.3. immediately below.

### 7.4.3. Importance of ‘reasonable, necessary and proportionate’ criterion

The third criterion above generally will be the most important in justifying a human rights restriction. It essentially asks whether there are better ways of achieving the aim that do not involve this level of human rights restriction.

The Model Law provides that, as part of the FRIA Process, FRT Developers and Deployers declare that they have considered the justification of any human rights limitation, and they reasonably believe their FRT Application complies with Australian law.

By way of illustration, imagine that a school proposes to use an FRT Application to mark the roll. This would necessarily restrict students’ right to privacy. The school might claim there is a legitimate need to restrict privacy in this way, because, for safety and other reasons, the school needs to know where students are at the start of each class. However, any assessment of reasonableness, necessity and proportionality would focus on the other options available to the school to mark the roll (and thus protect students), including the conventional way of calling out students’ names. Those other ways would not involve the collection and use of students’ biometric and other sensitive personal data. As a result, and depending on the specific circumstances involved, such a use of FRT could fail this ‘reasonable, necessary and proportionate’ criterion.

### Box 3: Case study – The *Bridges* case<sup>95</sup>

In the *Bridges v The Chief Constable of South Wales Police* case, the Court of Appeal of England and Wales considered a challenge to the use of an FRT Application (AFR Locate) by South Wales Police in the United Kingdom.<sup>96</sup>

AFR Locate was used to extract biometric data captured in a live camera feed and compared the captured data to headshot photographs on a police watchlist. If a match was detected, the tool alerted a police officer who decided what if any action to take (for example, arresting the individual).

Edward Bridges, a civil liberties campaigner, was scanned by AFR Locate in Cardiff in December 2017 and again while attending a protest in March 2018. Although Mr Bridges was not included on a watchlist, he contended that given his proximity to the cameras, his image would have been recorded by the AFR Locate tool.

Without making a factual finding on this issue, the Court acknowledged ‘scientific evidence that facial recognition software can be biased and create a greater risk of false identifications in the case of people from black, Asian and other minority ethnic (‘BAME’) backgrounds, and also in the case of women’.<sup>97</sup>

The Court found that the use of FRT can breach human rights to privacy and equality or non-discrimination, and that the police did not have lawful authority to use this tool. The Court identified two particular problems:

‘The first is what was called the ‘who question’ at the hearing before us [who can be put on a watchlist for surveillance using the AFR Locate tool]. The second is the ‘where question’ [location of the deployment]. In relation to both of those questions too much discretion is left to individual police officers.’<sup>98</sup>

The Court also found that the South Wales Police had failed to fulfil its positive duty to make enquiries regarding the potential discriminatory impact of the AFR Locate tool. Specifically, the police had ‘never sought to satisfy themselves, either directly or by way of independent verification, that the software program in this case does not have an acceptable bias on grounds of race or sex.’<sup>99</sup>

95. This case study is adapted from Australian Human Rights Commission, *Human Rights and Technology Final Report*, (Report, March 2021) 118.

96. *R (on the application of Edward Bridges) v The Chief Constable of South Wales Police* [2020] EWCA Civ 1058.

97. *R (on the application of Edward Bridges) v The Chief Constable of South Wales Police* [2020] EWCA Civ 1058, [164].

98. *R (on the application of Edward Bridges) v The Chief Constable of South Wales Police* [2020] EWCA Civ 1058, [91].

99. *R (on the application of Edward Bridges) v The Chief Constable of South Wales Police* [2020] EWCA Civ 1058, [199] – [201].

Part 8.

# Model Law: the Facial Recognition Impact Assessment process



As previously noted, the Model Law includes provision for FRT Developers and Deployers to undertake a rigorous, systemic impact assessment of their FRT Application before it is used in the real world. This process – known in the Model Law as a Facial Recognition Impact Assessment (FRIA) – is intended to achieve two aims:

1. First, it will assist relevant FRT Developers and Deployers to undertake the human rights risk assessment and management process required by the Model Law.
2. Second, the publication of FRIAs will provide some transparency to affected individuals, the regulator and FRT deployers about the operation of FRT applications.

Like other forms of impact assessment – such as privacy, human rights or algorithmic impact assessments – a FRIA involves the rigorous consideration of a number of specific matters in the process of designing, developing and then preparing for use of an FRT Application. The FRIA process is similar to the way in which the Australian privacy regulator can undertake or require that a government agency undertake a privacy impact assessment (PIA).<sup>100</sup> Where a person would be required under law to create a PIA in relation to the use of FRT, the FRIA would be a substitute for the PIA.

The FRIA process has two steps: a use declaration and risk assessment, and a risk management declaration. This part of the report sets out how this process works, including who must undertake a FRIA, when it must be registered with the regulator and made publicly available.

---

*The FRIA process is similar to the way in which the Australian privacy regulator can undertake or require that a government agency undertake a privacy impact assessment (PIA).<sup>100</sup>*

---

100. See *Privacy Act 1988* (Cth), ss 33C-33D. It is also common for such impact assessments to be voluntary. For instance, the World Economic Forum has designed an FRT self-assessment framework for law enforcement agencies wanting to employ these technologies. This assessment includes a series of questions under nine main areas for consideration. See World Economic Forum, *A Policy Framework for Responsible Limits on Facial Recognition Use Case: Law Enforcement Investigations* (Report, October 2021).

## 8.1. FRIA Step 1 – use and risk assessment declaration

Step 1 of a FRIA involves an FRT Developer or Deployer assessing the level of risk associated with their FRT Application, by reference to the factors specified in the Model Law (see Part 7.2.). This first step of the FRIA process is a use and risk assessment declaration.

All FRT Developers must complete FRIA Step 1 for base-level, elevated and high-risk FRT Applications. All FRT Deployers must complete Step 1 for elevated and high-risk FRT Applications. As explained in Part 8.3.1., FRT Deployers for base-level risk applications may be able to rely on a FRIA that has already been registered, leading to a simplified process.

Prior to any deployment of an FRT Application on affected individuals, Step 1 of the FRIA involves an FRT Developer or Deployer stating:

1. the use or range of uses to which the FRT Application will be deployed
2. the individuals or groups of individuals likely to be affected by the FRT Application
3. the factors that bear on the human rights risk assessment for the FRT Application
4. the overall human rights risk assessment of the FRT Application
5. they have considered the justification of any human rights limitation, and they reasonably believe their FRT Application complies with Australian law.

## 8.2. FRIA Step 2 – risk management declaration

After a person has completed Step 1 of a FRIA, all FRT Developers and those FRT Deployers not able to rely on a FRIA exception also must complete a second step in the FRIA process.<sup>101</sup> Step 2 in the FRIA process involves declaring the risk management steps taken by the relevant FRT Developer or Deployer.

Under the Model Law, Step 2 of the FRIA requires the relevant FRT Developer or Deployer to state:

- any specific human rights vulnerabilities or risks that were identified in Step 1 of the FRIA (see Parts 7.3. and 7.4.)
- any technical or operational limitations in the FRT Application that might affect the accuracy of its outputs generally, or in respect of particular

demographic groups (such as people of colour, women or people with physical disability), especially where this inaccuracy might result in unfairness or unlawful discrimination to affected individuals. This should include the results of system testing to identify performance gaps, including across demographic groups.

- measures taken to address problems referred to above
- any conditions of use, technical limitations or guidelines that must be followed in order for the registered risk assessment to be valid (this final requirement likely will be most relevant to FRT Developers marketing an FRT Application for use by FRT Deployers).

<sup>101</sup>. For more information on which FRIA steps apply to FRT Developers and Deployers, see Part 8.1. and Box 4.



## 8.3. Registration, publication & updating obligations

The Model Law generally requires that all completed parts of a FRIA – whether by an FRT Developer or FRT Deployer – be registered with the regulator and made available by the regulator in a searchable, online and public repository. This is a transparency measure, and the fact that the regulator maintains a register for FRIAs does not imply that the regulator endorses the FRT Application or its lawfulness.

In addition to registration with the regulator, in the interests of transparency for individuals dealing with government and private sector bodies that use FRT, and in line with Australian Privacy Principle 1.5, FRT Developers and Deployers also must take reasonable steps to make their registered FRIA available free of charge in an appropriate form.<sup>102</sup>

### 8.3.1. When FRT Deployers do not need to complete and register their own FRIA

The Model Law contains two exceptional situations where an FRT Deployer is not required to register a FRIA. The rationale for each of these exceptions is to avoid unnecessary regulatory burden for FRT Deployers and for the regulator. Given the rising use of FRT in Australia in a wide range of contexts, a high volume of FRIAs should be anticipated, and it would be unreasonably costly for the regulator to review and validate each registered FRIA.

1. The first exception applies where the FRT Deployer is a natural person, and all of the following conditions are met:
  - the individual is using or deploying an FRT Application for a non-commercial purpose
  - there is an existing registered FRIA that states the FRT Application presents no greater than a base-level risk
  - the individual uses the FRT Application within the relevant FRIA's use guidelines.

In this situation, the individual would not be required to complete their own FRIA for the FRT Application; they could simply rely on the existing, registered FRIA. There would also be no registration requirement in this situation.

While an individual would not be required to register a FRIA in these circumstances, they would be still required to comply with all other legal requirements in the Model Law regarding an FRT Application's use. This is important because of the risk of individuals deploying FRT Applications in ways that can cause harm, and even engaging in lateral surveillance as discussed in Part 4.3.1.

2. The second exception applies where the FRT Deployer wishes to use an FRT Application in accordance with another registered FRIA.

Given the rapidly-expanding market for FRT products and services, and the challenge for FRT Deployers to access detailed information about the functionality, training data, performance, and other attributes, it would be overly onerous for all Australian FRT Deployers of commercially-available FRT Applications to be required to individually complete a detailed FRIA. Hence, where both steps of a FRIA have been completed by an FRT Developer, an FRT Deployer may choose simply to be bound by the existing FRIA if all of the following conditions are met:

- the existing registered FRIA states the FRT Application presents no greater than a base-level risk
- the FRT Deployer agrees to use the FRT Application within the relevant FRIA's use guidelines
- the FRT Deployer registers with the regulator the fact they are using the FRT Application in accordance with the relevant FRIA.

The regulator would have the power to propose a simple mechanism by which an FRT Deployer would register such use and the specific FRIA on which they are relying.

### 8.3.2. FRIA updating obligations

FRIAs should be regularly reviewed and updated. A requirement to review and update the FRIA will be triggered if the FRT Developer or FRT Deployer makes changes to the design, purpose or context of the FRT Application that affect one or more of the vulnerability factors.

<sup>102</sup>. *Privacy Amendment (Enhancing Privacy Protection) Act 2012* (Cth) sch 1.



**Box 4: Which FRIA steps do I need to complete under the Model Law?****I'm an FRT Developer intending to market, sell or distribute an FRT Application in Australia:**

You must complete Steps 1 and 2 of the FRIA and register the FRIA with the regulator.

**I'm an FRT Deployer of an FRT Application in Australia for which a FRIA has already been completed and registered (a 'prior FRIA'):**

- If you are a natural person, intending to deploy a base-level risk FRT Application for a non-commercial purpose within the guidelines of the pre-existing, registered FRIA, you do not have to complete a FRIA at all.
- If you are a person intending to deploy a base-level risk FRT Application for a commercial purpose within the guidelines of the pre-existing, registered FRIA, you must register your intended use of the FRT Application with the regulator, indicating that you intend to rely on the prior FRIA.
- If the prior FRIA indicates the FRT Application creates an elevated risk, or if you wish to deploy the FRT Application outside the limitations outlined in the prior FRIA, you must complete and register Steps 1 and 2 of a new FRIA.
- If the prior FRIA has been assessed as high risk, you are prohibited from deploying the FRT Application unless you fall within the terms of an exception provided for in the Model Law.

**I'm an FRT Deployer of an FRT Application in Australia for which there is no prior FRIA:**

You must complete both steps of a new FRIA, register the FRIA and abide by the legal requirements flowing from your risk assessment of the FRT Application.

Part 9.

# Model Law: risk-based legal requirements



## 9.1. Mapping risk levels to obligations

Part 9 of this report describes the legal requirements and prohibitions applicable to FRT Applications. The specific legal requirements applicable to an FRT Application will depend on whether the Application has been assessed as base-level, elevated or high risk.

Under the Model Law, the legal requirements are cumulative, with each successive risk level incorporating the legal requirements of lower risk levels. Hence, use of a permissible high risk

FRT Application must comply with all the legal requirements relevant to both the base-level and elevated risk categories. Where legal requirements at different risk levels conflict or overlap, the requirement derived from the higher risk category will apply.

FRT Deployers and Developers each have legal responsibilities for the risk assessment, and they are each also responsible for compliance with the relevant legal requirements.<sup>103</sup>

## 9.2. Base-level legal requirements applying to all uses of FRT

The *base-level legal requirements* apply to any FRT Application assessed as presenting a base-level risk to human rights. As discussed in greater detail below, the base-level legal requirements are made up of:

1. the requirement to complete a FRIA
2. some new legal requirements set out in the Model Law, including some amendments to the *Privacy Act 1988* (Cth)
3. compliance with an FRT technical standard.

These requirements are discussed in turn below.

### 9.2.1. Requirement to complete a FRIA

Part 8 of this report describes when and how FRT Developers and Deployers are required to complete a FRIA in respect of an FRT Application. Some exceptions to the requirement to complete a FRIA are set out in Part 8.3.1.

Under the Model Law, where an FRT Developer or Deployer fails to comply with a legal requirement to complete or register a FRIA, they will be liable for a civil penalty. In addition, the regulator and courts will have the power to issue a mandatory injunction to prevent the ongoing use of an FRT Application until the person has completed and registered an appropriate FRIA with the regulator.

<sup>103</sup> The precise division of legal responsibilities, as between a developer and a user, will depend on the circumstances of the FRT application and how it is proposed to be used.

## 9.2.2. New legal requirements under the Model Law

### Privacy Act requirements would apply to FRT regardless of that Act's exceptions and exemptions

Under the Model Law, all FRT Applications used in Australia must comply with the Privacy Act, including the APPs, notwithstanding any otherwise applicable exceptions and exemptions in the Privacy Act.

This would have the practical effect of extending the operation of the Privacy Act and the APPs, because the Privacy Act currently contains a range of exemptions and exceptions that reduce or negate privacy protections for certain actors in certain situations. For example, small businesses, political parties and law enforcement agencies do not have to comply with the full set of legal obligations in the Privacy Act. In addition, the Privacy Act contains exceptions that reduce or negate certain privacy protections where, for example, an activity is required or authorised by another law.

In other words, regardless of whether a person would be able to avail themselves of an exemption or exception under the current Privacy Act, the Model Law provides that if the person is an FRT Developer or Deployer, they must comply with all of the general requirements in the Privacy Act in respect of the FRT Application. This would include anyone using or developing an FRT Application, such as universities, public schools, all businesses (including small business operators), media organisations, and political parties.

### Personal information derived from face data is 'sensitive personal information'

To recognise the special nature of face data, the Model Law clarifies that all personal information derived from an FRT Application or System, as well as face data specifically, should be treated as sensitive information under the Privacy Act, regardless of whether all such data is technically considered 'biometric' or 'health' information under the Privacy Act.

### Consent and notification

For the avoidance of doubt, the Model Law clarifies that any affected individual must have the opportunity to provide, or withhold, free and informed consent prior to the use of FRT on that individual. Applying the OAIC's preferred position, set out in its submission to the current AGD Privacy Act Review, consent should be a 'clear affirmative act that is freely given, specific, unambiguous and informed' and current such that it only lasts as long as is reasonable.<sup>104</sup>

This means that an affected individual must be given reasonably clear notification, at the time and place that FRT is used on the individual, that FRT is being used and for what purpose. This notification should be provided in a form that allows the individual to opt out, and it should contain information about how to obtain any applicable FRIA.

Take the following hypothetical example. If a small business embeds an FRT Application in its online shopping portal, with a view to allowing customers to log in and make payments with greater convenience, the small business would be an FRT Deployer. Consequently, any of its customers who choose to use the portal are affected individuals under the Model Law. As an FRT Deployer, the company would not be able to claim the current exemption in the Privacy Act for small businesses; instead it would be bound by the legal requirements in the Privacy Act, and must ensure that all face data and related personal information gained from the FRT Application is gathered, stored and managed in compliance with the APPs and other relevant provisions of the Privacy Act. This includes notifying individuals using the company's shopping portal that FRT is being used, its purpose, and the opportunity for the individual to decline the capture and processing of their face data.

In a rare set of circumstances, the Model Law provides that consent and notification are not required. For example, in some law enforcement and national security contexts, asking for consent or providing notification would compromise an investigation.

104. Australian Government - Office of the Australian Information Commissioner, *Privacy Act Review: Submission by the Office of the Australian Information Commissioner* (Discussion Paper, 23 December 2021) 230.

### Providing an alternative without detriment

Central to the concept of consent being ‘free’ is that the individual involved should not be unreasonably disadvantaged if they opt to refuse or withhold their consent. As a general principle, individuals should not be required to consent to the use of an FRT Application in order to access goods, services or venues. Where an individual does not consent to the use of an FRT Application on them for this purpose, generally the FRT Deployer should provide an alternative way for the individual to access the relevant good, service or venue. This is sometimes known as providing ‘an alternative without detriment’ where an individual refuses to provide consent.

The Model Law applies this general principle, while acknowledging that this principle is not absolute. To this end, the Model Law provides that, unless the use of FRT is the sole practical means available to an FRT Deployer of deciding whether to grant individuals access to a good, service or venue, the FRT Deployer must offer individuals an alternative means of accessing the good, service or venue. This alternative means of access should be on terms that are substantially similar to access for individuals who consent to the use of FRT.

### 9.2.3. Creation of a technical standard for FRT

The Model Law provides for the creation of a technical standard for FRT Applications used in Australia (FRT Standard). The proposed FRT Standard would provide technical specifications to assist FRT Applications to be developed and deployed in ways that comply with the Model Law.

In drafting an Australian FRT Standard, it would be important to ensure that the Australian Standard draws from and is congruent with international standards, such as those developed by the International Standards Organization (ISO) and the Institute of Electrical and Electronics Engineers (IEEE). In addition, it will be important to consult with key stakeholders from industry, civil society and government, including Standards Australia.

As with other such technical standards – for example, standards dealing with access to premises or public transport under the *Disability Discrimination Act 1992* (Cth) – the proposed FRT Standard would be a legislative instrument made under the Model Law. As such, it would not set out any *new* legal requirements beyond those that exist in the Model Law itself, and any other relevant primary legislation. Instead, it would simply assist with compliance.

An advantage of creating an FRT Standard is that it can be promulgated by a responsible minister or statutory authority, and so it can be reviewed and updated more expeditiously than is ordinarily the case with primary legislation – something that is important given the speed with which research and development on FRT are advancing, as well as the need for the FRT Standard to remain congruent with relevant international standards.

---

*As a general principle, individuals should not be required to consent to the use of an FRT Application in order to access goods, services or venues.*

---

Another advantage of an FRT Standard is that it can be sensitive to, and aligned with, the risk level posed by each particular FRT Application. For example, the Model Law requires an FRT Developer or Deployer to take any steps that are reasonable in the circumstances to ensure the personal information it collects, uses and discloses is accurate, up-to-date and complete.<sup>105</sup> An FRT Standard could point to steps that typically would be considered 'reasonable' in this context. An FRT Standard could also outline stricter standards for FRT Applications that pose an elevated risk.

There is significant work currently underway, especially overseas, in respect of technical standards for artificial intelligence systems in general, and FRT and other biometric technologies specifically.<sup>106</sup> Bodies such as the ISO and IEEE are conducting important work in this area. An Australian FRT Standard should draw on, align with, and help shape those international processes, noting that Australian experts are already working on artificial intelligence technical standards relevant to an FRT Standard through the Standards Australia Committee IT-043.

For example, the ISO 30137 standard series applies to the use of remote biometrics in video surveillance systems including for real-time operation against watchlists and in post-event analysis of video data. The ISO standard establishes general principles for supporting operators of such video surveillance systems and emphasises the need to have suitably-trained staff, as well as effective governance arrangements.<sup>107</sup> The IEEE Inclusion and Application Standards for Automated Facial Analysis Technology working group is creating standards for facial analysis.<sup>108</sup> Meanwhile, ISO Committee SC42 has 26 standards under development that include the development of quality evaluation guidelines for AI systems,

the treatment of unwanted bias in classification and regression machine learning tasks, and AI system impact assessment.<sup>109</sup>

Any FRT Standard produced under the Model Law should cover issues including:

- **Security.** The loss or unauthorised disclosure of face data undermines authentication integrity of FRT Systems and is a severe breach of sensitive, personal data. Relevant technical standards for the protection of stored biometric data should be applied including the use of cryptographic techniques such as digital signatures and encryption. These provisions should also cover data handling and minimisation.
- **Access and audit logging.** FRT Deployers should be required to maintain detailed audit logs regarding face data collection, access to face data, and access to outputs. The standard should specify the time period for which logs must be kept and made available for auditing purposes.
- **Data quality.** It should establish a mechanism for assessing and assuring face image quality.
- **Explainability.** It should describe ways to provide FRT Developers and Deployers with explainability in relation to an FRT Application's outputs.
- **FRT Application performance and testing.** There should be a mechanism for assessing the performance of FRT Applications, at different overall risk levels, and by reference to different demographic groups.
- **Third-party activity.** FRT Deployers should be responsible for ensuring the security of any third-party services that interact with or gain access to face data or FRT Systems.

105. See for example, *Privacy Act 1988* (Cth) APPs 10.1 and 10.2.

106. ISO/IEC JTC 1/SC 37, an ISO standardisation subcommittee focused on biometrics, has more than 120 published standards (including amendments) in biometrics. These standards are enhanced by extensive work since 1990 by ISO/IEC JTC 1/SC 27 related to biometric data protections techniques, biometric security testing, and evaluation methodologies.

107. International Standards Organization, 'ISO/IEC 30137-1:2019: Information technology — Use of biometrics in video surveillance systems' <<https://www.iso.org/standard/64935.html>>.

108. Joanna Goodrich, 'Standards Working Group Takes on Facial Recognition: Chair of IEEE Standards Association working group explains what the organization is doing to help ensure the technology is used ethically', *IEEE Spectrum* (online, 17th September 2019) <<https://spectrum.ieee.org/standards-working-group-takes-on-facial-recognition>>.

109. International Standards Organization, 'ISO - ISO/IEC JTC 1/SC 42 - Artificial Intelligence' <<https://www.iso.org/committee/6794475/x/catalogue/p/0/u/1/w/0/d/0>>.



## 9.3. Elevated risk: additional legal requirements

The Model Law provides for *additional legal requirements*, beyond the base-level legal requirements discussed in Part 9.2., in respect of FRT Applications assessed as posing an elevated risk. As noted in Part 7.2.4., a common situation where an FRT Application will be deemed to be of *elevated risk* is when it is used to make a decision that has a legal or similarly significant effect.

Specifically, the Model Law includes the following additional legal requirements for any FRT Application assessed as elevated risk:

- An FRT Deployer intending to deploy an FRT Application which has been assessed as elevated risk may not rely on a prior FRIA, but must fully complete both Steps 1 and 2 of the FRIA process, and register a statement to this effect with the regulator (see Part 8).
- The FRT Developer and/or Deployer must provide for human review of any decisions with legal or similarly significant effect, where the FRT Application materially contributed to the decision. Affected individuals should be provided with clear information for how to obtain that review.
- The FRT Developer and/or Deployer must provide effective training to relevant staff in the lawful development and/or use of the FRT Application (including staff responsible for administering, operating, maintaining, updating, responding to queries or interpreting the output of FRT Systems).
- The FRT Deployer will have a duty of care in deploying the FRT Application lawfully on affected individuals. Any breach of this duty will be actionable by affected individuals, and can be the subject of a complaint to the regulator. On request by the regulator or a court, the FRT Deployer would be required to provide a report to the regulator detailing information not covered in the relevant FRIA related to:
  - a) the performance of the FRT Application
  - b) all requests for rectification of data
  - c) all divergences from standard operating procedures
  - d) all security breaches related to face data.

---

*When an FRT Application is used to inform a decision which has a legal or similarly significant effect, it will likely be deemed ‘elevated risk’.*

---



## 9.4. High-risk FRT Applications

### 9.4.1. Prohibition of high-risk FRT Applications subject to limited exceptions

The Model Law contains a general prohibition in respect of high-risk FRT Applications.<sup>110</sup>

As set out in greater detail in the remainder of Part 9.4., the Model Law provides some limited exceptions to this general prohibition, namely:

- 1. Regulator authorisation.** Where an FRT Developer or Deployer considers that use of its high-risk FRT Application is justified under international human rights law, it will be able to apply to the regulator for authorisation. If the regulator is satisfied that this is so, it could authorise the development and/or deployment of the FRT Application, subject to any conditions it sees fit (see Part 9.4.2.).
- 2. Law enforcement & national security.** The Model Law contains a special regime regulating the development and deployment of high-risk FRT Applications by Australian law enforcement and national security agencies, and organisations acting on their behalf (see Part 9.4.3.).
- 3. Genuine academic research.** High-risk FRT Applications are permitted for use in genuine academic research, where there are appropriate ethical and legal protections in place (see Part 9.4.4.).

### 9.4.2. Regulator authorisation

The Model Law grants the regulator power to authorise high-risk FRT Applications where the FRT Developer or Deployer demonstrates to the regulator's satisfaction that the relevant risks to human rights can be addressed or managed appropriately. The regulator's authorisation can be made subject to any special restrictions, which would have the legal status of enforceable undertakings.

One area where such an authorisation process may be valuable is in the area of health care and accessibility for people with disabilities, especially in the provision of professionally-supervised diagnostic, therapeutic or assistive technologies. For example, the regulator may authorise otherwise high-risk FRT Applications that have a clear benefit regarding accessibility, such as personal-use smart phone applications which use facial analysis to assist with emotion recognition for people who are blind or have a vision impairment.

The regulator should have the power to designate a specific certification process for assessing and granting such exceptions, including delegating certifications to qualified third-party providers. To ensure that such exceptions do not overwhelm the resources of the regulator, the regulator should have the ability to recover the cost of its authorisation process from relevant FRT Developers or Deployers.

<sup>110</sup>. That is, to use the language preferred by the OAIC, high-risk FRT Applications could be considered, in general, a 'prohibited practice' under the Privacy Act.

### 9.4.3. Law enforcement & national security

#### Background

Police and other security agencies have a special role in any liberal democracy to protect the community from criminal and other unlawful activity and to protect national security. As a result, these agencies are generally given extraordinary powers to restrict or even suspend rights (for example, through powers of arrest and detention). International human rights law recognises the importance of public order and national security, and the need for a range of human rights to accommodate these imperatives.

In this context, the use of FRT by law enforcement and national security agencies can be useful, especially in identifying people suspected of criminal activity, victims caught up in humanitarian disasters and for other similar purposes. However, the use of FRT for law enforcement and national security purposes is highly controversial.

Where FRT is misused, overused or used in error in a law enforcement context, the potential human rights risk can be extreme. For example, an error in identifying an individual with FRT can lead to the individual being unlawfully arrested, detained and having their rights to a fair trial restricted. Where those errors are unevenly distributed across the community, this can lead to unlawful discrimination and racial and other forms of profiling (see Part 4.2.). Moreover, the ever-increasing deployment of FRT in the community can push our society towards mass surveillance.

#### Overview of the Model Law approach for FRT in law enforcement & national security

In view of these considerations, this report seeks to strike an appropriate balance. The Model Law contains a special regime for the use of FRT in the law enforcement and national security context (the FRT in Enforcement Regime). This is set out in greater detail below but, in summary, the FRT in Enforcement Regime is as follows:

- The FRT Enforcement Regime applies to law enforcement and other government national security agencies or another entity operating on their behalf (collectively referred to as Enforcement Agencies). The responsible minister will be responsible for gazetting entities that are subject to the FRT in Enforcement Regime.
- FRT Developers and Deployers that are not Enforcement Agencies – such as companies wanting to use FRT to address the problem of loss prevention in their stores – do not fall within the scope of the FRT in Enforcement Regime, and so cannot benefit from the exception to the prohibition against high-risk FRT Applications that applies to Enforcement Agencies. Such other organisations may, however, seek to avail themselves of other exceptions in the Model Law (for example, the regulator authorisation regime discussed in Part 9.4.2.).
- Where an Enforcement Agency proposes to develop or use an FRT Application for law enforcement or security purposes, it must comply with the requirements applicable to base-level and elevated risk FRT Applications, and it must also comply with the special legal requirements set out in the FRT in Enforcement Regime (see below).
- The Model Law does not provide a general exemption allowing Enforcement Agencies to use high-risk FRT Applications without any restrictions. Instead, the FRT in Enforcement Regime provides a number of conditions for the use of high-risk FRT Applications in law enforcement, such as through the creation of a new ‘face warrant scheme’. Only where these conditions are met by an Enforcement Agency, will the Agency be permitted to use the relevant high-risk FRT Application.

## The FRT in Enforcement Regime

The remainder of this portion of the report (Part 9.4.3.) outlines the FRT in Enforcement Regime – that is, the Model Law’s special provisions that apply to Enforcement Agencies wishing to develop or deploy FRT Applications in a law enforcement or national security context.

### Regulator review

**A. Regulator review of FRIA.** Where an Enforcement Agency proposes to develop or use an FRT Application, its FRIA must be reviewed and authorised by the regulator prior to the FRT Application being used.<sup>111</sup>

**B. Performance measures & assurances.** The regulator must be satisfied that the FRT Application and any related FRT System are performing close to, or within a reasonable range of, the best-performing systems in the world. This should take into account the role of any Enforcement Agency staff or contractors involved in the process.

### Lawful data & use in court

**C. Lawfulness of face data.** For the avoidance of doubt, captured and reference face data, including personal data used to test algorithms, must be lawfully obtained. The regulator should also clarify any special circumstances, such as the purchase or use of public face data, including the circumstances under which Enforcement Agencies can use or rely on the results of FRT Systems in jurisdictions not covered by this Model Law.

**D. Evidentiary rules.** The outputs of an FRT Application cannot be used as the sole basis for ascertaining identity or attributing behaviour in criminal investigations or in criminal or other similar proceedings. Furthermore, any use of FRT in such proceedings must be revealed to the court, relevant affected individuals and other parties.

### Limitations on use

**E. Minimum seriousness threshold.** FRT Applications assessed as high risk should be authorised only in the context of the investigation of serious crimes (for example, crimes that attract a minimum custodial sentence of three years or more),<sup>112</sup> to identify deceased persons in coronial investigations or in limited circumstances during missing persons investigations, as determined by the regulator.

**F. Prohibition on facial analysis.** Enforcement Agencies are prohibited from using facial analysis functionalities, unless they receive specific authorisation from the regulator.

**G. Whistleblower protection.** Enforcement agencies are prohibited from using FRT Applications to identify whistleblowers or journalistic sources.

**H. Prohibition against wholly automated face identification systems.** Enforcement Agencies may not deploy wholly-automated FRT Applications or Systems (for example, with no human intervention in the decision-making process) for the purposes of facial identification.

### ‘Face warrant scheme’ for live or routine use of FRT in law enforcement and national security

**I. Face warrant scheme.** There will be a new ‘face warrant scheme’ whereby a judge, persona designata or similar independent authority has the power to consider applications by Enforcement Agencies to conduct live, repeated or routine use of FRT involving members of the public who are not suspected of having committed a crime.<sup>113</sup> Where a face warrant is granted, it should be time limited and for a specific purpose. Enforcement Agencies are prohibited from using FRT in live and routine circumstances without such a face warrant.

111. See discussion in Part 11.2. regarding the possible need for a specialised regulator for Enforcement Agencies’ use of FRT.

112. The regulator should carefully consider how to ensure that this seriousness threshold be applied so as to prevent uses of FRT under the enforcement regime exception unduly restricting rights. Custodial sentencing guidelines are an imperfect guide to the perceived seriousness of a crime, and could be used to justify the use of FRT in contexts where it is clearly inappropriate from a human rights perspective. An alternative approach is to use frameworks such as the Australian Bureau of Statistics National Offence Index.

113. In this context, ‘routine’ FRT refers to the repeated or systematic capture and analysis of face data of members of the public in ways that effectively replicate a live system. For example, tasking Enforcement Agency staff or officers to routinely take photos of participants in a public protest for immediate FRT processing, or the routine use of FRT at traffic stops, would create a de-facto ‘live’ system that operates by other means.

### *Structural protections*

**J. Separation of data handling, request processing & investigators.** To prevent misuse and enhance auditability, Enforcement Agencies must put in place structural barriers between staff using an FRT Application to identify individuals, and staff responsible for data handling and management.

**K. Benchmarking & operational testing.** Prior to the deployment of an FRT Application, rigorous testing of the Application and system as a whole must be undertaken and recorded. Specific attention should be given to ensure the Application performs equally across diverse demographics, including age, race, gender and disability.

**L. Human review.** Before any action is taken based on the use of an FRT Application to identify an individual using one-to-many facial identification, the Enforcement Agency should apply an automatic process of independent internal review, whereby an officer of the Enforcement Agency who is generally more senior than the initial officer independently reviews the output of the FRT Application.

**M. Standard operating procedures.** Enforcement Agencies must establish a clear set of publicly-available policies, processes and standard operating procedures for using and updating FRT Applications in accordance with human rights, including the FRT Deployer's plan to conduct periodic testing of any FRT System in operational conditions and address any performance gaps, such as across demographic groups. The results of this testing should be made publicly available. Enforcement Agencies must regularly review audit logs and other information to identify all deviations from standard operating procedures. Enforcement Agencies must,

on a regular basis, report to the competent regulatory body and make public a summary of all material or non-trivial deviations from standard operating procedures. Where a search, decision or investigation was made on an affected individual contrary to the standard operating procedure of the FRT Application, they must be promptly informed of this unless doing so would compromise an investigation.

### *Governance & accountability*

**N. Notice & consent.** Generally, the Model Law's provisions in respect of notice and consent apply in this enforcement context, except where the relevant Enforcement Agency's legitimate aim for deploying the FRT Application would be frustrated by providing for notice and/or consent for any affected individuals.

**O. Accountability, transparency, & reporting.** Accountability measures should be further enhanced to include ministerial oversight for the use of FRT by law enforcement agencies. Enforcement Agencies are required to register their FRIAs in a way that allows public scrutiny.

In addition to these legal requirements for individual Enforcement Agencies, the Model Law creates structural barriers to prevent the creation of public surveillance systems in Australia. For example, the Model Law includes a provision to make it unlawful for cities, municipal authorities, and private owners to link public CCTV systems directly to FRT Systems so as to create live, FRT-capable video surveillance systems.<sup>114</sup>

<sup>114</sup> For clarity, these restrictions on the use of FRT in CCTV systems are not intended to prevent law enforcement agencies from lawfully obtaining footage and applying FRT in a subsequent investigation. However, they would prevent, for example, a council from creating a live network of FRT-capable CCTV cameras surveilling public spaces.

#### 9.4.4. Genuine research

The Model Law provides an exception for academic researchers to conduct research involving FRT Applications, including those assessed as being high risk, provided that the research is safe and complies with the legal, professional and ethical rules that apply to genuine academic research in Australia. A researcher wishing to deploy a high-risk FRT Application still has a duty to complete and register a FRIA prior to undertaking FRT research.

For this exception to apply, two conditions must be met:

1. First, the research project should take place within the structure or under the formal supervision of a regulated research body, such as a public university. In this way, any use on individuals should be subject to prior approval by a relevant human research ethics committee, with a particular emphasis on ensuring prior, free and fully informed consent on the part of any affected individuals.
2. Second, the FRT application involved in the research, or any related FRT system, should not produce decisions that have a legal or similarly significant effect on affected individuals.

---

*Where FRT is misused, overused or used in error in a law enforcement context, the human rights risk can be extreme.*

---

**Part 10.**



**Independent  
review & dispute  
resolution**

The Model Law enables review of how FRT Developers and Deployers are complying with their legal obligations. Review can address errors or other legal problems with any registered FRIA, and it can also consider any alleged breach of the relevant law in respect of the use of FRT Applications.

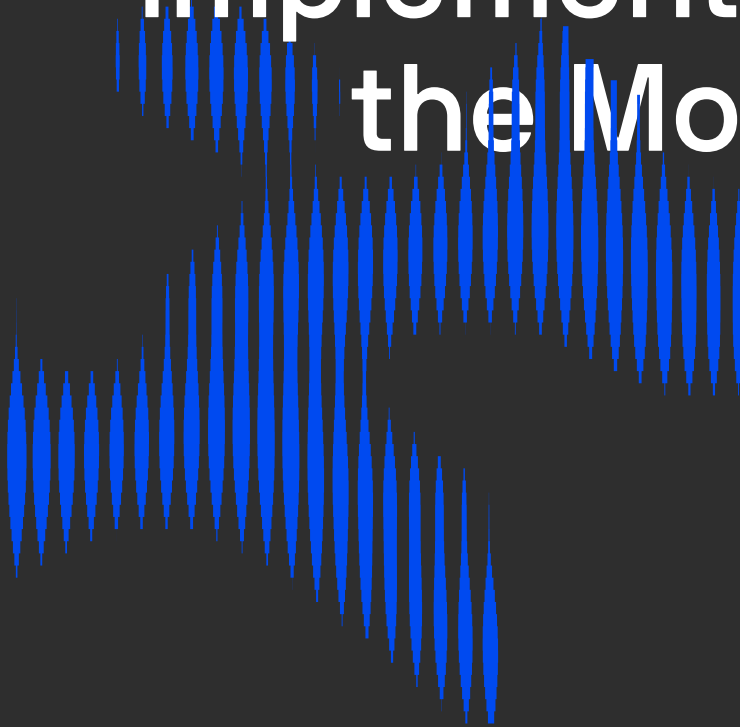
The Model Law grants review rights to the regulator to oversee the operation of this law, and also to affected individuals and some other related parties to make complaints for independent dispute resolution.

The Model Law grants the regulator a power to conduct, of its own motion, audits and reviews of FRIAs, including by reference to the relevant FRT Application's operation. If the regulator determines that the self-assessed risk level in Step 1 of a FRIA is incorrect, or that the statements in Step 2 of a FRIA should be amended, the regulator will be able to substitute its own view on these matters. If an FRT Developer or Deployer considers that the regulator made an error of law in this process, it will be able to seek judicial review of the regulator's decision.

The second form of review could be initiated by an affected individual, or another party acting to support one or more affected individuals. In this case, they would lodge a complaint regarding an FRT Application with the regulator. Complaints could be made that some material part of the self-assessment in Step 1 of a FRIA, or the declaration in Step 2 of the FRIA, is incorrect according to the relevant law. A complainant could also claim that an FRT Developer or Deployer has breached a relevant law in using their FRT Application. In both of these scenarios, the regulator could resolve the dispute by issuing its own determination. The regulator's determination would be subject to judicial review for errors of law.

Part 11.

# Implementation of the Model Law





There is an urgent need to reform the law applicable to FRT. This report urges the Federal Attorney-General to lead this reform process, guided by three principles:

1. The aim of this reform should be to foster positive innovation for public benefit using FRT, while protecting against the risks of harm to human rights.
2. Australia's federal, state and territory governments should work cooperatively to ensure a harmonised legal approach that provides consistency of protection for all Australians, and a clear, straightforward compliance obligations for all FRT Developers and Deployers.
3. Australia should engage actively in leading international reform and standard-setting processes on FRT, such as through the International Standards Organization, with the aim of contributing positively to those processes, and also incorporating this work as appropriate in Australian law and policy, such as in the proposed Technical Standard.

The remainder of Part 11 of this report outlines the key steps necessary in achieving positive reform.

## 11.1. Primary legislation

This report contains an outline of a model law for FRT. To give this legal effect, it would be necessary to draft a bill, which the Attorney-General would need to introduce into the Australian Parliament. This bill would apply to FRT within the Australian Government's regulatory purview. That is, it could apply to all corporations operating in Australia and to federal government departments and agencies, including law enforcement and security bodies such as the Australian Federal Police.

The Attorney-General has three viable options in introducing this bill. The simplest option would be to amend the Privacy Act, by creating a new division of that Act dealing with FRT, and potentially other forms of remote biometric technology.

A related option would be to amend the Privacy Act through a larger package of reforms that take in the recommendations of the soon-to-be-completed Attorney-General's Department review into Australian privacy law.<sup>115</sup>

A third reform option would be to introduce a stand-alone bill. Such a bill would establish a legal framework that connects to, but is separate from, the Privacy Act. The Australian Government previously attempted such a bill (the Identity-matching Services Bill 2019), but that bill was the subject of widespread concern, and it lapsed with the previous parliament.

As noted above, each of these reform options is viable. Most crucially, reform is urgent and important, and so, the Attorney-General should pursue a legislative approach that ensures expeditious passage of this reform.

<sup>115</sup> On 12 December 2019, the then Attorney-General announced a review of the *Privacy Act 1988* 'to ensure privacy settings empower consumers, protect their data and best serve the Australian economy.' The review has published an Issues Paper (October 2020) and a Discussion Paper (October 2021). The review's final report is anticipated soon.

## 11.2. Assigning & resourcing the regulator

Key to the success of the proposed FRT Model Law would be to assign regulatory responsibilities to a body that has expertise in human rights, especially the right to privacy, and can work constructively with a wide range of stakeholders, including FRT Developers and Deployers in the public and private sectors. In addition to conventional regulatory functions associated with compliance and dispute resolution, the regulator also would take a central role in the creation of an FRT Standard, and in providing advice to FRT Developers, Deployers and affected individuals.

There are a number of viable options as regulator. The most obvious would be to make the Office of the Australian Information Commissioner (OAIC) the regulator. Other existing regulators, such as the Australian Human Rights Commission, could also be suitable alternatives. While there might be some advantage in creating a wholly new regulator for FRT and/or biometric technology more broadly, we do not consider this to be necessary to achieve the Model Law's objectives.

Where national security issues arise – for example, where the Australian Federal Police or another security service is an FRT Deployer or Developer – it will be necessary to provide for appropriate secrecy and clearance protections for the assigned regulator. The OAIC already has provision for addressing these concerns, especially in performing its current responsibilities under the *Freedom of Information Act 1982* (Cth). Alternatively, it may be appropriate for enforcement agencies to be regulated under the Model Law by their existing specialist oversight agencies.

Regardless of which body is selected to be the regulator for the FRT Model Law, the government will need to provide the necessary financial and other resources to enable the regulator to fulfil its remit. Given the breadth of new functions assigned to the regulator under the FRT Model Law, a rigorous process will be necessary to assess the one-off and recurring funds and other resources needed.

---

*Regardless of which body is selected to be the regulator for the FRT Model Law, the Government will need to provide the necessary financial and other resources to enable the regulator to fulfil its remit.*

---

### 11.3. Ensuring the law is accessible, clear & effective

FRT Developers, FRT Deployers and affected individuals will benefit from support in understanding their respective rights and duties under the Model Law. The regulator should play a central role in providing guidance and other support regarding:

- best practice approaches to completing the FRIA process
- how different FRT use cases can disproportionately affect particular groups, such as children, people with disability and people of colour
- templated examples of common FRT use cases, including illustrative risk assessments and suggested management or mitigation strategies to address human rights risks
- avenues of review and other advice for affected individuals concerned about an FRT Application, including advice on how to access and review a registered FRIA.

Undertaking the FRIA process, as well as compliance with the proposed FRT Technical Standard, will inevitably require some FRT Developers and Deployers to seek expert advice to ensure they are operating within the law. This is a feature, not a bug, of the Model Law, as it incentivises investment in review, assurance and governance of FRT Systems that will benefit all parties.

## 11.4. A harmonised approach across all Australian jurisdictions

Under Australia's constitutional arrangements, the federal, state and territory governments have overlapping responsibilities to uphold human rights, including the right to privacy, and each of these jurisdictions has passed its own legislation dealing with these issues. The Model Law, when implemented, should be consistent and easy to understand for FRT Developers, Deployers and affected individuals, regardless of where one is located in Australia.

It would be possible for federal legislation to govern most development and use of FRT – especially where the regulated entities are corporations. However, generally, federal, state and territory governments are responsible for regulating their own respective departments and agencies. There is a risk, therefore, that different legal rules would apply to the use of FRT by, say, the Australian Federal Police as compared with a state or territory police force. This would be an undesirable outcome, because it would fragment the human right protections applicable across Australia, and it would add unnecessary 'red tape' in complying with multiple federal, state and territory legal regimes.

For these reasons, the Attorney-General should initiate a process with state and territory counterparts to ensure that the law on FRT is harmonised across all Australian jurisdictions. This could be achieved in a number of ways. For example, the states and territories could refer to the Australian Parliament the power to legislate a comprehensive national law on this subject. Another option would be for the federal, state and territory governments to agree on a single FRT Model Law, with each jurisdiction committing to passing and maintaining laws that are uniform or at least consistent.

One specific issue will be whether there is a residual role for state and territory regulators. Again, those state and territory governments could simply agree with the federal government to assign full responsibility to a federal regulator. Another approach would be to establish a system for mutual recognition and consistency of approach between federal, state and territory regulators, especially in respect of FRT use by state and territory departments and agencies such as state police forces. If the OAIC were the federal regulator, then it could work with its state and territory privacy regulator counterparts through mechanisms such as Privacy Authorities Australia.

---

*The Australian Government should work with the state and territory governments to promote a consistent approach to the regulation of FRT in Australia.*

---

## 11.5. An Australian Government taskforce on facial recognition

The success of this reform will depend in part on the effective implementation of the policy underpinning the Model Law, especially given the complex technical issues that arise in respect of FRT.

To this end, the Attorney-General should work with other relevant federal ministers to establish an Australian Government taskforce on FRT. The proposed taskforce would have two functions.

First, it would work with all relevant Australian Government departments and agencies, such as the Australian Federal Police, to ensure their development and use of FRT accords with legal and ethical standards. This would involve the creation of training programs and policy material to support the achievement of the FRT Model Law's goals.

Second, the taskforce would lead Australia's international engagement on FRT, so that Australia can have a positive influence on the development of international standards and other assurance mechanisms for FRT, and to ensure that Australian law on FRT is consistent with international law and international best practice.

The taskforce should also advise the Government on ways to streamline the operation of Australian law in this area, given that many FRT Applications are developed wholly or partially in overseas jurisdictions. This might involve, for instance, mechanisms for mutual recognition of other impact assessments for FRT Applications, where those other assessments are conducted under comparable laws of other jurisdictions or under the auspices of the International Standards Organization, and they apply, in substance, the elements of the Model Law's FRIA process.

---

*There should be an Australian Government taskforce on FRT to apply international best practice in Australia.*

---

## Appendix 1: Methodology & consultation

### Expert Reference Group

As noted in Part 3 of this report, the Project formed an Expert Reference Group (ERG). The members of the ERG were:

- **Duncan Anderson**, Executive Director of Strategic Priorities and Identity, NSW Police Force
- **Professor Fang Chen**, Distinguished Professor and Executive Director, Data Science Institute, UTS
- **Ivana Jurko**, Co-founder and Lead - Evidence & Influence, Humanitech (from June 2022)
- **Kavita Kewal**, Assistant Secretary Identity and Biometrics Policy and Strategy Branch, Department of Home Affairs
- **Katie Kinsey**, Chief of Staff, Policing Project New York University Law
- **Owen Larter**, Director of Public Policy, Office of Responsible AI, Microsoft
- **Dr Monique Mann**, Senior Lecturer in Criminology, Deakin University
- **Scott McDougall**, Queensland Human Rights Commissioner
- **Kieran Pender**, Senior Lawyer, Human Rights Law Centre
- **Kate Pounder**, CEO, Technology Council of Australia
- **Amanda Robinson**, Director, Humanitech (until June 2022)
- **Roger Taylor**, Advisor to the Responsible AI Programme, Accenture

## The Human Rights and Technology Project

Much of this report builds on the extensive research and consultative work undertaken by the Australian Human Rights Commission (AHRC) during its three-year, national Human Rights and Technology Project. That project's 2021 Final Report was informed by 291 written submissions, 725 consultation participants and 2149 national survey participants.<sup>116</sup>

Specifically, Chapter 9 of the AHRC's report focuses on biometric surveillance, facial recognition and privacy. It explores the concerns raised by FRT, the views of Australians towards this technology, and presented three recommendations for further consideration. Of particular relevance to this report is the AHRC's finding that stakeholders 'urged that the regulatory approach to facial recognition should focus on risk', and its recommendation that 'all Australian governments work cooperatively to introduce legislation to regulate the use of facial recognition and other biometric technologies.'<sup>117</sup>

### Roundtable consultations and key informant interviews

The Project Team undertook a small number of roundtable consultations during the report drafting process to test hypotheses and proposals for the risk-based framework of the Model Law. This included seeking feedback on the 'vulnerability triggers' and minimum legal requirements for elevated and high-risk use cases of FRT. The sessions took place primarily in June 2022 and included senior representatives and subject specialists from across civil society, government (including police and law enforcement), and the technology industry.

In addition to formal and informal input from the Expert Reference Group, the Project Team also undertook several key informant interviews with experts from civil society, government, academia and industry.

### Qualitative research for this project

The report authors commissioned a piece of qualitative research to provide some grounding for this project. Essential Research was commissioned to explore community attitudes to FRT using a qualitative research methodology.

A critical component of that methodology involved the development and use of an FRT simulation tool, which was commissioned separately from the strategic design agency, Paper Giant. The tool simulated four hypothetical but plausible, real-world scenarios in which various forms of FRT could be used:

1. An organisation using FRT in place of a swipe card for people to enter a secure building, simulating one-to-one facial verification.
2. A bank using FRT as a form of identity confirmation for people applying for a home loan, again simulating one-to-one facial verification.
3. A border security agency using FRT at an airport to undertake an identity check, simulating one-to-many facial identification.
4. Police using FRT to monitor a public space and determine individuals' level of threat based on facial analysis of perceived aggressiveness.

116. Australian Human Rights Commission, 'About the Project', *Human Rights & Technology* (Web Page, 2021) <<https://tech.humanrights.gov.au/about-project>>.

117. Australian Human Rights Commission, *Human Rights and Technology Final Report*, (Report, March 2021) 117.

Essential Research used the tool as part of a series of six focus groups to collect qualitative data on the following key questions:

- What are participants' initial perceptions of facial recognition, prior to exposure to the FRT simulation tool?
- What are participants' perceptions of facial recognition after exposure to the FRT simulation tool?
- What rules or restrictions, if any, do participants think should be enforced around the use of FRT?

The focus groups were conducted live, via Zoom, with a total of 42 participants spread equally across age groups of 18 to 34 years, 35 to 54 years, and over 55 years. Within each age-group, two sessions were held based on participants self-selecting as either 'keen' or 'reluctant' users of technology. Conducting the focus groups online allowed for significant diversity across the participant cohort part, including through gender, socio-economic background, and geographical location – including representation across states and territories as well as between regional and metropolitan regions.

### **Summary of focus group findings**

The focus groups revealed that, prior to exposure to the FRT simulation tool, participants generally lacked deep awareness and understanding of FRT and how it is currently being used – especially by the private and government sectors. Generally, participants' perceptions of the technology were informed by reference to their own (often limited) experiences of using one-to-one facial verification to unlock smartphone or other such devices.

Many participants reported feeling that the proliferation of FRT is inevitable, and that there is little that individual Australians can do to influence this. Even so, many participants were concerned about the corresponding issues of privacy and data security which result from an increased adoption of FRT.

After participants completed the simulation, they were more readily able to identify both risks and opportunities presented by FRT. Notably, participants acknowledged the ease and convenience of using FRT to verify their identity in relatively low-risk settings where any possible errors in the technology could be easily addressed. By contrast, the use of facial analysis was seen as particularly concerning, and participants were more likely to challenge or question both the accuracy of this technology and the corresponding implications of a denial of access to a service or venue because of a decision made by facial analysis.

On the whole, most participants only considered the risks of FRT from a personal perspective, such as a breach of their own data security or how they might be inconvenienced by the technology, as opposed to considering the broader societal implications such as mass surveillance or discrimination.

Participants were asked whether there should be legal protections or restrictions on the use of FRT, and what these might be. Many participants responded that they wanted the law to:

- require notice and consent prior to being subject to FRT
- restrict the collection and use of FRT data to its original purpose only
- require that FRT not be the sole basis for decision making in elevated and high risk use cases
- require technical accuracy of FRT before it is rolled out more widely by industry and government.



## Appendix 2: Summary of the Model Law's legal requirements

**Table 4: Legal requirements under the Model Law**

Legal requirements under the Model Law	Base-level risk	Elevated risk	High risk
Complete & register FRIA Step 1 – Use and risk assessment declaration	#	✓	✓
Complete & register FRIA Step 2 – Risk management declaration	†	✓	✓
Comply with Privacy Act requirements regardless of that Act's exceptions and exemptions	✓	✓	✓
Treat 'face data' as sensitive information under Privacy Act	✓	✓	✓
Consent and notification requirements	✓	✓	✓
Comply with FRT Standard	✓	✓	✓
Provide for human review of any decision with legal or similarly- significant effect made using FRT Application	x	✓	✓
Provide effective training for relevant staff	x	✓	✓
Duty of care to deploy FRT Application lawfully	x	✓	✓
Auditing obligation, on request by regulator	x	✓	✓
General prohibition, subject to exceptions: <ul style="list-style-type: none"> <li>■ special authorisation by regulator</li> <li>■ special legal regime for law enforcement &amp; national security</li> <li>■ genuine academic research</li> </ul>	x	x	✓
Civil penalties for non-compliance	✓	✓	✓

### Legend

✓ = requirement applies to all FRT Developers and FRT Deployers.

x = requirement does not apply.

# = requirement applies to all FRT Developers. FRT Deployers may choose to rely on a prior FRIA.

† = requirement applies to all FRT Developers and some FRT Deployers.

## Appendix 3:

# Facial Recognition Impact Assessment Template

### FRIA Part One

Use declaration and risk assessment: for all developers and deployers except individuals using an FRT Application for non-commercial purposes (who are exempt from completing a FRIA).

#### Type of FRIA

1.1 Are you completing a FRIA as an FRT Deployer, as an FRT Developer or both?

---

1.2 If you are an FRT Deployer, do you intend to rely on a previously-completed base-level risk FRIA?

---

#### Use declaration

1.3 Where, when and how is the FRT Application intended to be used?

---

1.4 What is the FRT Application intended to achieve?

---

#### For FRT Deployers relying on a previously registered FRIA ('prior FRIA')

1.5 Which prior FRIA are you relying on?

---

1.6 Do you agree to be bound by the conditions of use detailed in the prior FRIA?

---

#### Risk assessment

1.7 Whose face data are likely to be captured, searched or analysed by the FRT Application?

---

1.8 In what 'spatial context(s)' do you or others intend to use the FRT Application?

---

1.9 What FRT 'functionalities' does the Application employ?

---

1.10 How accurate and reliable is the FRT Application?

---

1.11 How are face data and related outputs secured to protect unauthorised access?

---

1.12 Is use of the FRT Application designed to contribute to a decision with legal or similarly-significant effect?

---

1.13 Does the FRT Application contribute to a decision that is partially or wholly automated?

---

1.14 Can individuals easily give prior, free and informed consent to the use of the FRT Application?

---

1.15 Based on the vulnerability factors above, what is your overall risk assessment of the FRT Application?  
[base-level, elevated, high]

---

1.16 Have you considered the justification of any human rights limitations posed by the FRT Application?

---

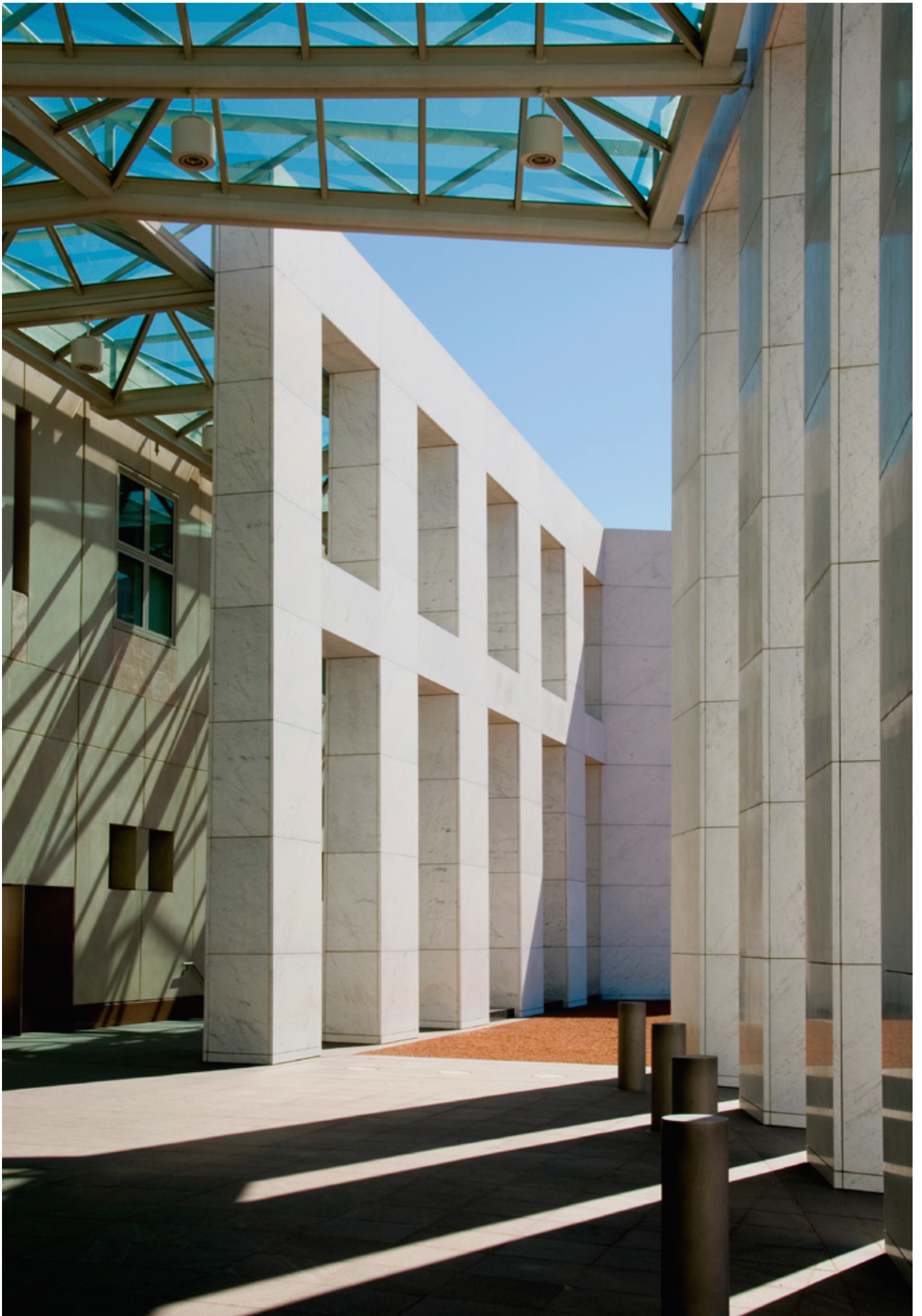
1.17 Based on your assessment of the FRT Application, do you believe that the FRT Application complies with Australian Law?

---

## FRIA Part Two

For all FRT Developers and FRT Deployers of 'elevated' or 'high risk' FRT Applications.

- 
- 2.1 How can people who have not consented to the use of the FRT Application still access your product or service without detriment?
- 
- 2.2 How will the performance and outputs of the FRT Application be assessed?
- 
- 2.3 How will human review and redress of any relevant decisions be conducted?
- 
- 2.4 How will face data and other related information be kept up to date, accurate and complete?
- 
- 2.4 How will any errors produced by the FRT Application be promptly identified, recorded, reported and rectified?
- 
- 2.5 What training will be available to support the responsible deployment of the FRT Application?
- 
- 2.6 What are the preconditions necessary to use the FRT Application safely and securely at the assessed risk level?
-







## For more information

Human Technology Institute  
[hti@uts.edu.au](mailto:hti@uts.edu.au)

University of Technology Sydney  
PO Box 123  
Broadway NSW 2007

[uts.edu.au](http://uts.edu.au)