

AIPOI MEMBERS USE ONLY.
Please do not hand this out to
members of the public

AiPol

Australasian Institute of Policing



Journal of the Australasian Institute of Policing Inc.

Volume 17 Number 3 • 2025



ROBOCOP COMING SOON

AUSTRALIAN POLICE CAUTIOUS

Fov:240.0x240.0
W:938 L:591
XGY OPER-0.35

Our commitment is to
enhance and maintain
police real estate so you
can do your best work.





WE'RE RECRUITING

SECURITY | TECHNOLOGY |
MANAGEMENT | PROPERTY SERVICES

[SECURECORP.COM.AU/EN/NOW-RECRUITING](https://securecorp.com.au/en/now-recruiting)

OVER

25

YEARS'
EXPERIENCE

MORE THAN

2,500

EMPLOYEES

BASED IN

225

SERVICE
LOCATIONS

PERMANENT
OPPORTUNITIES

SUPPORTIVE TEAM

EXCELLENT PAY

TRAINING PROVIDED



Scan for more
information
and to apply

 **SECURECORP**



‘831
Lives lost
in the line
of duty,’

National *Police* Remembrance Day



29 September 2025 

National Police Remembrance Day is a solemn occasion for all Australians to pause and reflect on the lives of police officers who have made the ultimate sacrifice.

We also **remember** those whose **lives were lost** due to illness, injury, or other tragic circumstances, acknowledging the **lasting impact** of their service and sacrifice. •

“Their service. Our gratitude. *N*ever forgotten.”

29th of September 2025



 **TASER 10**

A new era in less-lethal technology

With an effective range of 13.7 metres and a magazine featuring 10 individually-targeted probes, TASER 10 gives operators more time and distance to stop a threat without taking a life.

AU.AXON.COM/TASER-10

 **AXON**

Thank you for your sacrifice & your service.



**The Hon
Anthony Albanese MP**



**The Hon
Richard Marles MP**



**The Hon
Jim Chalmers MP**



**The Hon
Pat Conroy MP**



**The Hon
Kristy McBain MP**



**The Hon
Matt Thistlethwaite MP**



Kara Cook MP



Senator Dorinda Cox



Matt Gregg MP



Dan Repacholi MP



Tracey Roberts MP



Federal Labor proudly supports police officers, who serve our community with courage and compassion.

We remember those police officers who have made the ultimate sacrifice and we convey our greatest sympathies to their fellow officers, families and loved ones.



**The Hon
Tony Burke MP**



**The Hon
Jason Clare MP**



**The Hon
Tanya Plibersek MP**



**The Hon
Dr Andrew Charlton MP**



Julie-Ann Campbell MP



Claire Clutterham MP



Sam Lim MP



Fiona Phillips MP



Dr Gordon Reid MP



Susan Templeman MP



Anne Urquhart MP

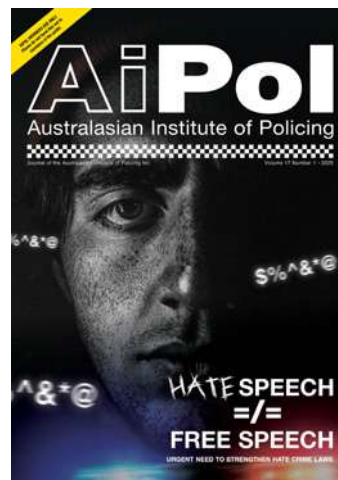


Vol. 17, No. 3
September 2025

Published by the Australasian Institute of Policing Inc.

A0050444D ABN: 78 937 405 524

ISSN: 1837-7009



Visit www.aipol.org to view previous editions
and to subscribe to receive future editions.

Contributions

Articles on issues of professional interest are sought from Australasian police officers and police academics. Articles are to be electronically provided to the Editor, aipoljournal@aipol.org. Articles are to conform to normal academic conventions. Where an article has previously been prepared during the course of employment, whether with a police service or otherwise, the contributor will be responsible for obtaining permission from that employer to submit the article for publication to Australasian Policing.

Contributors are expected to adhere to the Journal's publishing guidelines. These guidelines are available in this Journal. All papers are peer-reviewed.

Disclaimer

While every effort is made to check for accuracy, the Publishers or Editors cannot be held responsible for the content, errors or omissions inadvertently published in articles and advertisements in Australasian Policing. Views expressed by contributors are not necessarily those of AiPol, the Editors or the Publisher. No responsibility for loss occasioned to any person acting, or refraining from acting, as a result of material in this publication can be accepted.

Copyright

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or be stored in any retrieval system of any nature, without written permission of the copyright holder and the Publisher, application for which in the first instance should be made to the Publisher for AiPol.

TABLE OF CONTENTS

Editorial	9
President's Foreword	11
Governance, oversight and transparency of Artificial Intelligence within policing	16
Australia New Zealand Police Artificial Intelligence Principles - a Primer	21
Accountable AI Official Appointment	27
The tech whizzes working around the clock to catch criminals in NSW	28
Artificial intelligence and child sexual abuse: A rapid evidence assessment	33
'It's beyond human scale': AFP defends use of artificial intelligence to search seized phones and emails	43
Australian federal police using AI to analyse data obtained under surveillance warrants	44
Artificial intelligence and policing: it's a matter of trust	46
AI and policing: what a Queensland case study tells us	48
Effective, Explainable and Ethical: AI for Law Enforcement and Community Safety	50





Studying the Intersection of Brain Science and Criminal Behaviour

What if you could peer inside the criminal mind? What drives antisocial behaviour? Can an understanding of brain function change our approach to justice?

These and other compelling questions underpin the University of New England's groundbreaking Master of Neuroscience and Criminology degree – the first of its kind in Australia.

This innovative program bridges disciplines to explore new ways of understanding crime and the justice system. It was developed by experts in their fields to support professionals refreshing their careers in law enforcement, criminal justice and policy.

The curriculum will allow students to develop their understanding of neurobiological factors influencing criminal behaviour, and they are encouraged to apply that knowledge in evaluation of real-world criminal justice scenarios. With practical instruction in subjects like forensics, crime prevention, and the neurobiology of behaviour, students will gain the skills to examine complex social issues from an interdisciplinary perspective.

Students also gain access to UNE's Centre for Rural Criminology, a world-first hub of collaborative international research, which seeks to understand all aspects of rural crime in order to help build safe and resilient communities.



For those working in fields like criminal justice, law enforcement, mental health, or policy development, UNE's Master of Neuroscience and Criminology builds career-expanding expertise. It equips graduates to make meaningful contributions to

conversations about criminal responsibility, rehabilitation efficacy, and the neurological factors influencing human behaviour in justice contexts.

Short course offerings that deliver a Graduate Certificate or Graduate Diploma provide practical qualifications for those seeking to upskill mid-career. The full Master's degree offers research experience and the opportunity to develop specialised knowledge and practical expertise in the neurobiology of behaviour as it relates to criminological practice.

Working professionals undertaking the course are fully supported by UNE's flexible approach to study. The program is delivered through sophisticated online learning environments with 24/7 tutor support. UNE has consistently earned five-star ratings for Overall Experience and Student Support from The Good Universities Guide.

Apply now for UNE's Master of Neuroscience and Criminology:
www.une.edu.au/study/courses/master-of-neuroscience-and-criminology

une
University of
New England

Postgraduate **Neuroscience & Criminology**

The only postgraduate degrees taught by both Neuroscientists and Criminologists



Master of Neuroscience and Criminology



Graduate Diploma in Neuroscience and Criminology



Graduate Certificate in Neuroscience and Criminology

These online courses have a strong focus on neuroscience, and are taught by both neuroscientists and criminologists, allowing you to future fit your career in this rapidly developing field.

This combination of Neuroscience and Criminology will ensure that you learn how the brain works, what can go wrong in the brain, how to understand the brain's role in behaviour, and how this insight can support policing and inform legal decision-making.

- Freedom to choose specialisations in Criminology and Neuroscience
- Master's degree includes a research project of your choosing
- Completely online
- Limited number of Government subsidised places available

CRICOS 00003G | TESSA PRIV2054

Apply now to start October 2025 or February 2026.

For further information, contact Dr Kirstan Vessey:

kvessey@une.edu.au



une
University of
New England



Message from the Prime Minister: **Police Remembrance Day**

Police Remembrance Day is an important day for Australia. It provides a powerful moment in which we can acknowledge and reflect on the courage and commitment that our police officers exhibit in their service.

Today, my heart and mind are with the fellow police officers, family and friends of those that we have lost. Those who put themselves in harms way for the better of the Australian people.

Policing demands the highest standards of duty, diligence and morality. It can nurture and create people with extraordinary courage and selflessness. People that live and act not for themselves, but for the wellbeing of their communities.

Thank you for your service.

Your sincerely,

Authorised by A. Albanese, ALP, 334A Marrickville Rd, Marrickville NSW 2204



Linfox Logistics is Asia Pacific's largest privately-owned logistics company with operations employing more than 24,000 people and spanning 12 countries. Everyday our multinational team delivers food, medicines and resources across those 12 countries.

www.linfox.com • 03 8340 1000

Follow us:   

Proudly supporting AIPOL

Editorial

DR AMANDA DAVIES

Editor, Senior Researcher at the Charles Sturt University



“The reality is that not only are police organisations engaging with the possibilities of AI for improving operational responsibilities, on the other side of the law, the criminal community is similarly active.”

Welcome to the second edition devoted to the adoption of AI in policing, this edition focusing on the Australian domain. As we go to press there is a constant flow of discourse on the use of AI in policing, its benefits, disadvantages and challenges. Chief amongst which is the risk factor, cyber hacking is a genuine threat as the world has experienced and there appears to be no discrimination – attacking small, large, government and private entities.

Nevertheless, studies are signposting that AI is emerging as a valuable tool for policing, both for operational and administrative taskings. In Australia for example, AI-enhanced investigative tools and digital records systems have demonstrated the benefits of AI in transforming evidence handling and threat assessment procedures, enabling faster case resolution and more accurate risk forecasting.

As with many modern initiatives in policing, including the adoption of body worn cameras, a key challenge is balancing citizen privacy concerns with upscaling policing capabilities and mitigating increasing criminal activities. As discussed in the ANZPAA paper in this edition and the article Effective, Explainable and Ethical: AI for Law Enforcement and Community Safety by Wilson et al., (2020), the current and potential benefits of AI for policing engender adoption of this approach to policing whilst in parallel establishing sound governance, policies and procedures for its use.

The body of literature that is building associated with the use of AI in policing, particularly case studies (such as the article AI and policing: what a Queensland case study tells

us) highlights not only is the adoption of AI here to stay, it is having significant positive influence on community safety. Here also there is the cautionary warning that engagement in ethical, transparent adoption of AI tools is imperative. The article referencing Dr Teagan, Artificial intelligence and policing in Australia, offers a well balanced and compelling discussion, explaining in depth the nuanced benefits of AI and the manner in which it is able to analyse large data sources whilst also explaining why there is a critical need to ensure such use is strategically governed and secured.

Whilst across a diverse array of disciplines and professions AI is being adopted to enhance efficiency, effectiveness and innovation as with all 'new' trends in digital technology, there must be attention to monitoring the benefits and ensuring algorithmic bias and data privacy breaches are mitigated. The reality is that not only are police organisations engaging with the possibilities of AI for improving operational responsibilities, on the other side of the law, the criminal community is similarly active. As we move forward it will be critical that policing continues to develop the use of AI to tackle criminal activities and potentially increase the number of officers deployed to operational duties vs caught behind a desk with administrative tasks that could be assigned to AI tools. Digitalisation and particularly AI offers a 'brave new world' that has the potential to increase community safety and wellbeing, enhance officer safety and contribute to UN sustainable goals for a safe, modern and thriving community.

Nat's Coffee Shop



Delicious Coffee • Cakes • Pastries

📍 Ascot Vale Station

Opening Hours:

Mon–Fri: 6:45am – 1:00pm

Sat–Sun: 7:45am – 2:00pm

Also Available:

- ✓ Reusable Cups (8oz, 12oz, 16oz)
- ✓ Fresh Coffee Beans (250g & 1kg)
- ✓ Coffee Gift Cards



Get in touch!

☎ 0425 121 253

Follow us on Instagram

📷 @natscoffeeshop_

President's Foreword

JONATHAN HUNT-SHARMAN

President, Committee of management, Australasian Institute of Policing



It's a Brave New AI Policing World

A cautious approach adopted by Australian Police

As mentioned in our last edition, the integration of Artificial Intelligence (AI) into policing represents a significant shift in law enforcement practices globally. Countries such as the United States, Singapore, Thailand and the European Union have been at the forefront of integrating AI into law enforcement.



AI technologies are rapidly being adopted to enhance policing capabilities, improve public safety, and tackle the growing complexity of crime in the digital age. We are seeing AI increasingly being used in policing for data analysis, crime prediction, detection and prevention offering potential benefits like improved efficiency and targeted resource allocation.

As AI has continued to advance, we have also seen police forces throughout the world adopting what was once science fiction. China, Russia, Thailand, United Arab Emirates, Indonesia, Singapore and the USA now have stepped into the world of robo-cops patrolling streets and malls.

There are now many international examples of AI in policing and it is growing at an exponential rate and so are the mistakes! There are now many legal challenges occurring overseas to the use of AI in policing. As can be rightly expected, there are concerns about bias, privacy, and accountability. So where does that leave policing in Australia? How does Australian police adapt to this technology whilst not having over-reach into the freedoms that Australians expect? How can the *Independent Office of Constable*, Sir Robert Peel's legacy to policing in a democracy, be protected from AI over-reach? These are fundamental questions which I am pleased to say, Australian police are attempting to address through a considered, collaborative and measured approach.

AI within Australian policing is somewhat in its infancy. Australian governments and federal, state and territory police agencies are all very conscious of ensuring there is no over-reach using AI. They are conscious of the need for careful consideration and oversight, especially with technologies like facial recognition and predictive policing algorithms. Police jurisdictions are working in cooperation with academics to set a safe path through the AI minefield. They are working together to ensure appropriate accountability and transparency measures are in place to reduce judicial and societal concerns and to ensure national standards are adopted across jurisdictions.

There is no specific legislation regulating AI in Australia however the Australian government is currently working on a Bill to go before the Australian Parliament. Currently there are overarching policies and regulatory

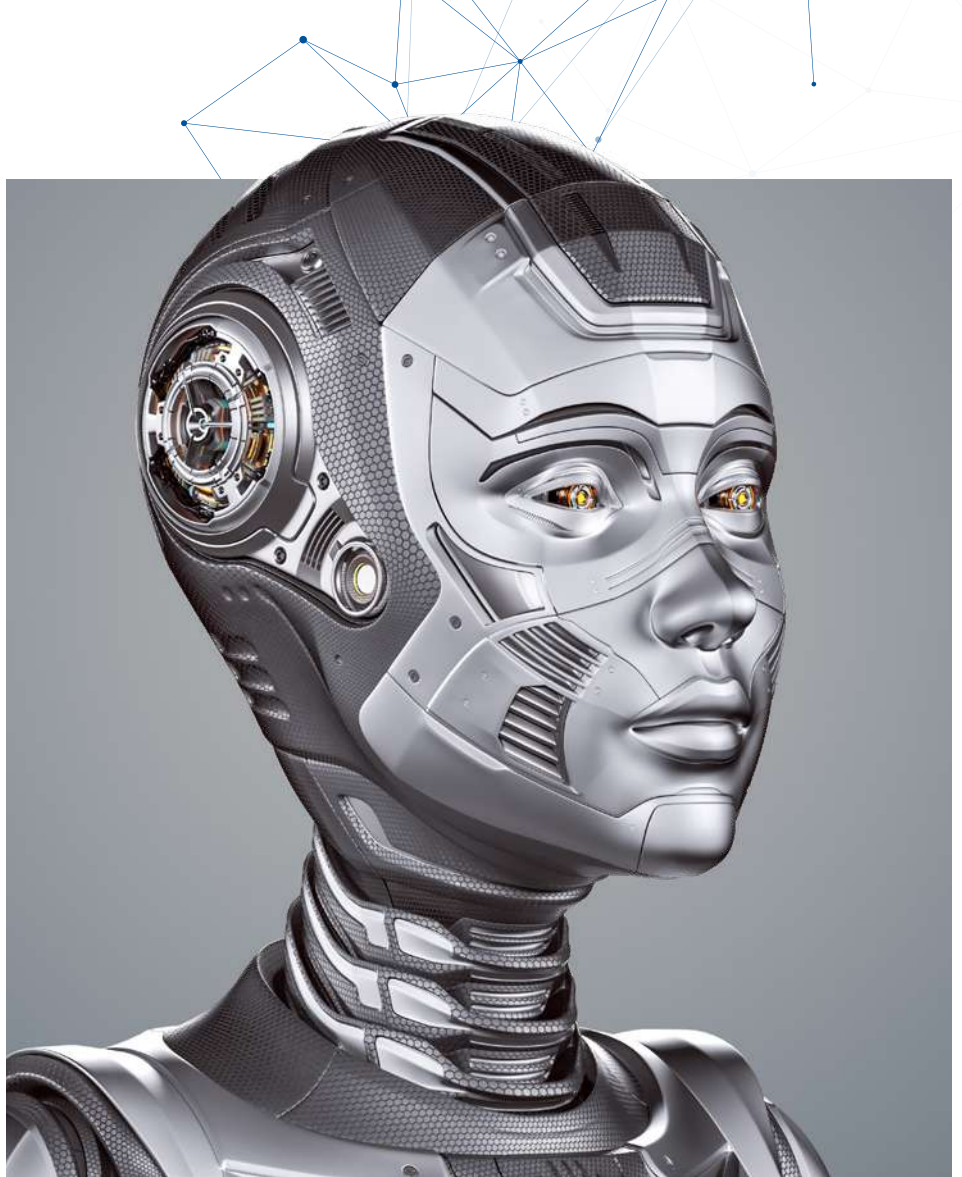
frameworks, that must be considered. These include Australia's *AI Action Plan*; *the Digital Economy Strategy*; (Australian-Government 2022), the *CSIRO AI road map* (CSIRO 2019); and the *AI Ethics Framework* (Australian-Government 2024). These overarching documents have been developed to ensure AI technologies are safe, secure and are trusted by the public. Among these policies, AI Ethics Framework is particularly significant. The framework aims to facilitate safer, more reliable, and accurate AI outcomes, to minimise the risk of negative impacts on those affected by AI, and to enable governments and businesses to adhere to ethical standards during the design and deployment of AI. These principles play a crucial role in enhancing public trust and transparency in AI usage, thereby multiplying its positive impact (Australian- Government 2024).

Similarly, ANZPAA (Australia New Zealand Policing Advisory Agency), an organisation to promote ethical and responsible use of AI in policing across Australia and New Zealand, developed nine principles including:

- **transparency;**
- **explainability;**
- **accountability;**
- **human oversight;**
- **fairness;**
- **skills and knowledge;**
- **proportionality and justifiability;**
- **reliability; and**
- **privacy and security.**

In 2024, the Australian Government published an interim response plan (Australian Government Dept of Industry and Resources 2024) which outlines three critical steps: **testing, transparency, and accountability.**

- **Testing:** Mandates that AI products are safe both before and after deployment;
- **Transparency:** Involves providing clear information about model design, data, and AI usage, including labelling and watermarking AI-generated content; and
- **Accountability:** Ensures that AI developers are properly trained, have knowledge of Australian laws and regulations, and that organisations have clear



accountability standards in place. The Australia New Zealand Policing Advisory Agency (ANZPAA) on behalf of police agencies in Australia and New Zealand, has been focusing of four (4) key areas:

- **Developing Ethical Frameworks:** Establishing clear ethical guidelines and standards for the use of AI in policing.
- **Ensuring Transparency and Accountability:** Insisting that AI algorithms are more transparent and establishing clear accountability mechanisms for their use.
- **Training and Education:** Developing Educational modules for Police officers on how to use AI tools responsibly and ethically, and how to interpret the data they generate; and
- **Collaboration and Dialogue:** Encouraging open dialogue and collaboration between law enforcement agencies, technology developers, academics and the community to address the ethical and social implications of AI in policing.

In Australia, the integration of AI into investigative processes is fundamentally changing the way Australian law enforcement agencies handle and analyse vast quantities of evidential data, enabling them to enhance efficiency and improve prosecution outcomes. AI algorithms are being used by Australian police to predict crime hotspots and identify potential offenders. AI-powered facial recognition systems are being used to identify suspects and monitor public spaces. AI is being used to analyse body camera footage to identify potential misconduct and improve officer accountability and AI is analysing voluminous documentary, real, and forensic evidence during police investigations and for Court preparation. Importantly all this is being done under the national regulatory framework and under the ANZPAA principles for policing.

There are some exciting developments in AI in policing in a number of Australian jurisdictions. For example:-

- The Australian Federal Police (AFP) has been at the forefront of utilising AI to enhance their investigative capabilities. The AFP routinely handles ultra large investigations - often involving 40TB or more of data (telephone intercepts, listening devices, CCTV, emails, sms, etc. For major fraud investigations, international drug syndicate



investigations, international child exploitation investigations the AFP's AI tools help sift through voluminous criminal intelligence and evidentiary material, structuring datasets, flagging anomalies, summarising imagery, and to analyse data obtained from telecommunications and surveillance warrants. etc. The AFP also collaborates with Microsoft using Azure AI tools to detect deepfake content and generate de-identified text summaries before presenting material to investigators. The AFP is also using AI to analyse transactional data to detect anomaly patterns in the context of combating financial crimes, where the rapid identification of suspicious transactions can prevent further illegal activities. AFP Operation Ironside involved distributing ANOM encrypted devices to members of international organised criminal syndicates that were covertly being monitored by the FBI and the AFP. This operation led to voluminous criminal intelligence and evidentiary material, often in

foreign languages, that was analysed through the use of AI. To understand the scale of this operation, in Australia Operation Ironside charged 383 alleged offenders with 2430 offences. More than 6.2 tonnes of illicit drugs and \$55.6 million in cash have been seized. globally, and excluding Australian statistics, more than 700 alleged offenders have been charged and 65 tonnes of illicit drugs seized. Investigations and charges are continuing whilst data obtained is still being analysed. For all AFP investigations using AI, an oversight committee ensures human supervision of AI-based decisions;

- The NSW Police operate a dedicated Facial Recognition Unit that cross references CCTV, body camera, and social media images with mugshot databases to generate investigative leads. Usage is governed by strict protocols - they cannot access licensing or passport images without formal requests, and images are manually reviewed before leads are passed on to investigators;

- The WA Police force deploy a cloud-based AI platform developed by Modis, powered by Microsoft Azure Cognitive services as a pilot program. This platform is designed to process and analyse large volumes of potential evidential data, including emails, text messages, social media posts, CCTV footage, and videos/photos. The AI's ability to handle such diverse data sources allows for the rapid identification of critical insights that might otherwise go unnoticed. For example, in one investigation, the AI system identified 18 new points of interest within 24 hours, information that was previously unknown to the investigators; and
- The Queensland (Qld) Police have been trialing the deployment of an AI with the aim of identifying and flagging high-risk domestic violence offenders.

AiPOL believes that the most fundamental protection for Australians when dealing with AI Policing is not the government policies, regulatory frameworks or even legislation, but is indeed the *Independent Office of Constable*.

In Australia, we are fortunate that the independent *Office of Constable* protects the rights of all Australian citizens.

This independence is even more important as we move to increase AI use in policing.

All Australians should be comforted that the *Independent Office of Constable* is the bedrock of modern day policing in Australia. Every Constable is an independent legal entity, the public's guarantee of impartiality. Officers of the Crown operate independent of undue influence, interference and with personal responsibility and accountability.

For readers who are not familiar with the *Independent Office of Constable*, it is important to understand that this is a constitutional and legal requirement. It means that individual police officers are personally responsible-legally and ethically, for the actions they take while performing their duties. Unlike other professions, a Constable can't simply say, 'I was just following orders' if those actions are unlawful, abusive or negligent. They must always ensure that their conduct is lawful, justified and proportionate. If a Constable abuses their powers or acts unlawfully, they can be criminally prosecuted and/or sued in civil court for damages and/or

;disciplined or dismissed This principle is vital because it ensures accountability in the exercise of police powers, protects individual rights; and ensures public trust in policing as there are serious consequences for officers for inappropriate or negligent behaviour.

This provides our protection against AI over-reach as the *Independent Office of Constable* requires the individual police officer/s involved to have clear oversight and understanding of the use of and outcomes from AI in policing. This requires mandatory human involvement for implementing and operating AI techniques within the field of law enforcement. Human intervention and human oversight are the two important factors that will ensure this fundamental principle is maintained when using AI in policing. The flow on effect being that the Australian public will embrace and support the appropriate use of AI in policing.

AiPOL will continue to influence decision makers in this important evolution of AI in policing. The policing profession must ensure that policing with AI is compatible with the public's expectation of the *Independent Office of Constable*.

Advertisement

NATIONAL POLICE REMEMBRANCE DAY 2025




We thank you for your service, today and every day.



Michelle

**MICHELLE ROWLAND MP
ATTORNEY-GENERAL
FEDERAL MEMBER FOR
GREENWAY**

Authorised by Michelle Rowland MP,
Australian Labor Party, Blacktown NSW 2148




NATIONAL POLICE REMEMBRANCE DAY

29th September 2025 

WITH HONOUR
THEY SERVED

DIVINE FAMILY FUNERALS

 divinefamilyfunerals.com.au

 (02) 9037 3900

Understanding Every Memory, Supporting Every Step

At Divine Family Funerals, we are more than just a funeral service; we are a team of compassionate individuals dedicated to honouring the lives of those who have passed. With over 20 years of combined experience, our funeral directors bring a wealth of knowledge, empathy, and understanding to every service we provide.

FUNERAL DIRECTORS



Jim Georges
Managing Director



Nick Dimitriadis
Managing Director

Honouring Every Journey with Dignity

Divinity Cremation offers dignified and affordable cremation services across Sydney. Whether you need a simple, direct cremation or a more traditional farewell, our compassionate team ensures each service is carried out with care, respect, and attention to detail.

We also provide low-cost options without compromising on quality or support.

Direct cremation package begin at
\$1,650 Incl GST

To explore our full range of services, please visit
www.divinitycremation.com.au

MEET OUR *Funeral Directors*

NICK & MARINA DIMITRIADIS

With deep compassion and years of experience, Nick and Marina guide families through one of life's hardest moments. Their care is personal, their support unwavering, and their commitment to each family truly heartfelt.



Divinity Cremations
— Memories Are Eternal —

**AUSTRALIA
NEW ZEALAND
POLICING AGENCY**

anzpaa.org.au

Governance, oversight and transparency of Artificial Intelligence within policing

Recent years has seen the expanded use of Artificial Intelligence (AI) within the criminal justice system, as data driven technology and algorithms play an increasing role in shaping law enforcement and criminal justice decisions across the globe.

Idea in Brief

- While police have an obligation to consider and implement new technologies which may advance their functionality and improve operational effectiveness, there must also be commitments to oversight and governance.
- Without sufficient safeguards, oversight and evaluation, the adoption and use of AI technology and capabilities by police may have serious implications on individual and societal rights to privacy, fairness and public perceptions of police legitimacy.



Considerations for Policing

In going forward with the procurement and implementation of emerging AI technology, police may wish to consider the following:

- Developing their own sets of principles to govern the response use of technology: Having police-led and developed standards in place can support proper consideration and assessment of the implications of technology within police agencies. ANZPAA is currently undertaking work to develop cross-jurisdictional AI principles for Australian and New Zealand police.
- Increasing transparency around the deployment of AI technology: Police can help to alleviate perceptions that AI technologies may lead to unfair or discriminatory outcomes through prioritising transparency around their processes for selecting and deploying AI. Developing technology tools 'in house' or in partnership with tech vendors: Participating in the development of AI technology can ensure capabilities are aligned with policing standards and obligations.

Introduction

Artificial Intelligence (AI) technologies, such as facial recognition technology (FRT) and predictive algorithms are increasingly playing a pivotal role in the criminal justice system.

When deployed, these technologies have the potential to improve the efficiency and functionality of policing and find solutions to complex problems. Police can utilise these technologies to support the detection and prevention of crime and generate new insights for policing. For example, AI technologies have provided significant assistance in detecting and prosecuting child abuse image trade and financial crime matters. However, as AI's deployment within the justice system has become more widespread, growing research has emerged to challenge its use. AI's critics have noted its capacity to perpetuate historical biases and the potential for serious implications on human rights and civil liberties.

This brief considers the governance and oversight of police's use of new and emerging technology. It examines lessons learned from international and



domestic examples and the potential repercussions if proper oversight and safeguards are not considered. The brief also identifies several opportunities for policing to develop their capability to avoid such repercussions.

Case Study

The Los Angeles Police Department (LAPD) recently terminated their use of a predictive policing program called PredPol, which aimed to identify when and where future crimes would occur based on past data. PredPol's algorithm was publicly criticised for reinforcing harmful patterns and creating a 'feedback loop'. Communities with a higher police presence will naturally have higher arrest rates, leading to datasets that appear to reflect higher crime rates, but which really reflect greater police attention. The use of these datasets may fuel the over-policing of Black and minority communities.

Algorithm bias

Biased algorithms within AI have made headlines across many industries in recent years, including within policing. Police agencies around the world increasingly use predictive algorithms based on historical crime data to assess offender risk levels and probability of further crime. A growing number of academics and researchers caution that these technologies may exacerbate discrimination, highlighting concerns relating to the dangers of human bias becoming embedded in the data that feeds AI algorithms and decision making.

Other studies have found that AI algorithms used to predict recidivism may demonstrate racial bias. Additionally, research on FRT has found that the technology is consistently less accurate on subjects of certain ethnicities and genders, in some cases leading to wrongful convictions. This is primarily due to a lack of diversity in the datasets that are used to train the technologies. Certain researchers have gone so far as to claim that FRT is inherently biased, while others agree that the use of larger datasets with better training methodologies should lead to greater accuracy.

Transparency and accountability concerns

There is often a shortage of information about how crime prediction and data-driven technologies are used within policing. Some critics have argued that the scepticism surrounding the use of this technology has less to do with the technology than a lack of transparency from the agencies administering it. In Los Angeles, details about the LAPD's use of predictive policing programs only emerged after years of campaigning from civil activists who demanded transparency around program operations.

Similarly, developments in technology used by police agencies have not always been accompanied by adequate safeguards, particularly if the technology is acquired commercially. While there may be great enthusiasm for the potential use of emerging technology by police and in the justice system, there does not appear to be a corresponding commitment to thorough evaluation and oversight processes.

Technology acquired from big data companies may not be subject to the review and oversight that police agencies require as part of their governance and accountability structures. For example, an audit of Operation LASER used by the LAPD found that the program used criteria that was inconsistent and imprecise.

In the UK, in a review into the use of new technologies in the justice system, the House of Lords outlined serious concerns around the lack of minimum scientific or ethical standards in place for AI tools before their adoption and use in the criminal justice sphere.

Considerations for policing

Developing principles to govern the responsible use of technology

In 2021, the European Union (EU) developed a draft set of Rules for the development, placement on the market and use of AI systems. Considering the EU's significant global influence, if these rules are adopted, there are likely to be significant effects on the development of new technologies and commercial strategies, even outside the EU.

To guide the use of emerging technology within their jurisdictions, police may wish to create their own principles. ANZPAA will be developing a set of crossjurisdictional Principles



**DEVELOPING AI
TOOLS IN-HOUSE MAY
HELP AVOID RISKS
ASSOCIATED WITH 'OFF
THE SHELF' SOLUTIONS.**

to guide the adoption and use of AI within policing. Having a robust set of police led and developed standards in place can support proper consideration and assessment of the implications of technology prior to implementation.

To increase transparency and oversight, police may also wish to work with experts in this space while developing principles and establish a mechanism to seek independent and specialist advice relating to their technology goals.

Case Study: New Zealand Police

New Zealand Police recently developed an emergent technology program in a public commitment to using technology safely and responsibly. An expert, independent panel was created to provide advice and oversight from an ethical and policy perspective. The panel acts as a reference group for proposed applications of new and emerging technology in policing. New Zealand Police have committed to making the panel's advice public wherever possible.

Supported by advice from the expert panel, a policy on trialling

or adopting new technology was developed. This policy governs approvals for all new technology-based capabilities, or new uses of existing technology. Noting key community concerns around FRT, with the use of this technology becoming more widespread, New Zealand Police also commissioned an external review of the use of FRT in policing. The review provided a detailed assessment of the opportunities and risks surrounding the use of FRT in New Zealand communities. The review made 10 recommendations, all of which were accepted by New Zealand Police.

Increasing transparency around deployment of technology

With transparency and accountability around police's use of technology key community concerns, police may wish to make transparency around the deployment of AI a priority when going forward. Police can pre-emptively provide reassurance to the public and their communities through transparency about a technology's purpose, benefits, data collection and storage methods, and safeguards in place. This may help to demonstrate responsible use of technology and inspire public trust and confidence.

This aligns with a procedural justice approach and may help to support public perceptions that police are exercising their authority in the deployment of AI technology lawfully. Should police aim to be clear and open about their processes for selecting and deploying AI technologies, this may help to address any perceptions that AI technologies lead to unfair or discriminatory outcomes (e.g., deployment is non-consensual or there is a lack of transparency).

Developing tools in-house or in partnership with tech vendors

At ANZPAA's recent Policing Forum on Artificial Intelligence, it was suggested that police may need to develop AI tools

'in house' wherever possible, in order to avoid some of the risks associated with 'off the shelf' solutions.

For example, the intellectual property protections of certain commercial products may sometimes prevent users from obtaining information on the technology being used and the data which it depends on. This makes it difficult to assess the data used to train the algorithm. Another risk relates to the data that underpins externally sourced technology, which may have been developed overseas within a different policing context. As such, applying this technology in an Australian and New Zealand context may mean that communities are not reflected appropriately, and that the technology may not operate fairly.

However, police may not always have the resources to develop technology 'in house'. Alternatively, police may seek to develop their internal capability to work with technology vendors to create police-appropriate technology solutions and capabilities. In doing so, police can work directly with vendors to involve themselves in technology development processes to ensure that its capabilities and functions are aligned with policing standards and obligations.

Collaborating on the development of technology or building it 'in house' may

avoid some of the complexities surrounding commercial engagement. If police are involved in the development processes, there is likely to be greater clarity and understanding as to how data was gathered and prepared, providing greater assurance that the model used is fair and of appropriate complexity.

Case Study: Clearview AI

Controversial facial recognition company Clearview AI was used by hundreds of police agencies around the world to help solve shoplifting, identity theft, credit card fraud, murder and child exploitation cases. Since 2020, the company has faced multiple lawsuits and has been accused of violating numerous privacy and data protection laws around the world, including not having a lawful reason to collect personal information and a failure to have mechanisms in place to stop data being held indefinitely.

Many police staff and agencies who used Clearview later admitted to having only a limited knowledge of how the program worked.

Australia New Zealand Police Artificial Intelligence Principles

PURPOSE

The Principles guide the ethical and responsible use of artificial intelligence (AI) by Australian and New Zealand Police and promote cross-jurisdictional consistency. The Principles reflect Police's commitment to community safety, harm minimisation and maintaining community confidence in the adoption and deployment of AI systems.

CONTEXT

There is no universally accepted definition of AI. For the purpose of these Principles, AI is defined according to commonly used definitions such as those published by the Australian Government Department of Industry, Science, Energy and Resources¹ and the AI Forum of New Zealand².

The Principles exist within the context of established legal, human rights and privacy obligations and reflect organisational commitments to building trust with First Nations people and communities. The Principles recognise that the technological and social landscape will continue to evolve, requiring ongoing review of police practices and commitments to using AI ethically and responsibly.

Transparency

Police organisations should ensure clear and understandable information about the use of AI systems is made publicly available to the greatest extent possible without undermining policing objectives.

Human Oversight

Police organisations should ensure that AI is only used to inform decision-making, rather than to independently make decisions or determine outcomes. There should be appropriate human oversight and control at all stages of the development, deployment and operation of the AI system. This includes oversight of decisions involving human discretion.

Proportionality and Justifiability

Police organisations should use AI systems in a reasonable, necessary, proportionate and lawful manner and respect human rights. In determining whether to use AI, police should consider all available policing options. On balance, the benefits to community safety should outweigh any potential negative impacts from the use of AI.

Explainability

Police organisations should ensure AI systems are able to be appropriately described in a meaningful and accessible manner so their use can be understood and challenged.

Fairness

Police organisations should design and/or use AI systems in a way that respects equality, fairness and human rights. AI systems should not be used to unjustly harm, exclude, disempower or discriminate against individuals, groups or communities. Potential harms and biases should be identified via risk assessments and appropriately managed.

Reliability

Police organisations should continuously monitor, test and develop AI systems, and ensure they are derived from relevant and contemporary data. This helps to ensure optimal functionality and that AI systems continue to meet their intended purpose.

Accountability

Police organisations should employ appropriate layers of governance and engagement at all stages to ensure they retain primary accountability for the AI system and the decision-making it informs. Police organisations should remain accountable for use of AI systems obtained through external vendors.

Skills and Knowledge

Police organisations should ensure members have appropriate training, skills and knowledge to develop, deploy and operate AI systems. This includes understanding the capabilities, limitations and risks associated with the AI system. The level of skills and knowledge required is determined by the use and application of the AI system and should remain contemporary.

Privacy and Security

Police organisations should ensure privacy and security are at the forefront of the design and use of AI systems. This includes compliance with relevant privacy, data collection, data sharing, data access, security and records management requirements and legal obligations.

FOOTNOTES:

- Defined as "a collection of interrelated technologies that can be used to solve problems autonomously and perform tasks to achieve defined objectives. In some cases, it can do this without explicit guidance from a human being (Hajkowicz et al. 2019:15). AI is more than just the mathematical algorithms that enable a computer to learn from text, images or sounds. It is the ability for a computational system to sense its environment, learn, predict and take independent action to control virtual or physical infrastructure."
- Defined as "advanced digital technologies that enable machines to reproduce or surpass abilities that would require intelligence if humans were to perform them."



More than a building, more than a workspace.
Welcome to Grosvenor Place.

**EXPERIENCE
MORE**

Reward your business with ultra-modern workspaces and stand-out amenities. Grosvenor Place's large floorplates are among the largest on offer in the Sydney CBD, and allow you to create an efficient workplace specific to your business needs.

For leasing
opportunities

Frank Sassine
0408 487 854



Daniel Kernaghan
0432 201 664



CIRCA

HERITAGE & LIFESTYLE

OUR STORY BEGINS WITH YOUR STORY

Founded as a direct response to an unmet demand for professional, knowledgeable advice within a unique sector of the real estate landscape, we join you in the mutual admiration of our architectural history and unique Australian lifestyle.

If you have any property enquiries, whether it be related to heritage, lifestyle, farm or commercial property, or wish to obtain a confidential market appraisal, please get in touch to discuss how we can assist you.

Sue Gratton
Director
0407 599 559

@circaheritageandlifestyle

CIRCAHERITAGEANDLIFESTYLE.COM.AU

Australia New Zealand Police Artificial Intelligence Principles - a Primer

Recent years has seen the expanded use of Artificial Intelligence (AI) within the criminal justice system, as data driven technology and algorithms play an increasing role in shaping law enforcement and criminal justice decisions across the globe.

May 17, 2024

**LEADING SNR CONSTABLE
JANIS DALINS, PHD**

AiLECS Lab

**ASSOCIATE PROFESSOR
CAMPBELL WILSON**

AiLECS Lab

Background

Australia New Zealand Policing Advisory Agency (ANZPAA) released its Policing Artificial Intelligence Principles in July 2023. Originally informed through background research by ANZPAA officers, the design was finalised over a two day focus group consisting of representatives from each Australian and New Zealand policing organisation. Given its academic and policing membership, AiLECS was in the unique position to contribute to their design throughout the entire process.

In addition to this primer, further discussion of the principles is included within an episode of ANZPAA's *Police Horizons* podcast¹

The principles themselves are intended to be as simple as possible, with the decision actively made to make them fit to a single page, in readily accessible, non-technical language. Critically, the method of implementation is not dictated, reflecting each jurisdiction's autonomy.

The Principles Themselves

The Principles² consist of nine individual concepts:

- Transparency
- Human Oversight
- Proportionality and Justifiability
- Explainability
- Fairness
- Reliability
- Accountability
- Skills and Knowledge
- Privacy and Security

The full text for each item is not included for brevity. The reader is advised to read this document in parallel with the principles themselves.

What is Artificial Intelligence? Should I care?

The principles do not attempt to define AI, instead quoting a previous definition from the (Australian) Department of Industry, Science, Energy and Resources.

In fact, these principles readily apply to any law enforcement task undergoing automation.

Takeout: Any data driven project should consider these principles. Don't get distracted by people calling it "AI" or otherwise.

An explanatory scenario

The principles document is intentionally abstract, primarily for reasons of brevity and clarity. For the purposes of this document, we will use a scenario of where 'AI' has 'gone wrong' in a law enforcement context.

Note: Our summary of this case is based upon media reportage of a lawsuit related to this incident. We do not claim first hand knowledge, nor claim that all elements are fact - only that they are alleged.

In October 2023 Harvey Eugene Murphy Jr (MURPHY) was arrested for an alleged armed robbery of a Sunglass Hut

retail outlet in the Houston, Texas area in January 2002. He was incarcerated for several weeks before his alibi, being that he lived in California at the time of the alleged incident, was confirmed and charges dropped. In fact, his arrest occurred when he identified himself to the Texas Department of Motor Vehicles in order to renew his Texan driver licence upon his return to the State³. MURPHY alleges that during his incarceration, he was bashed and raped by other prisoners.

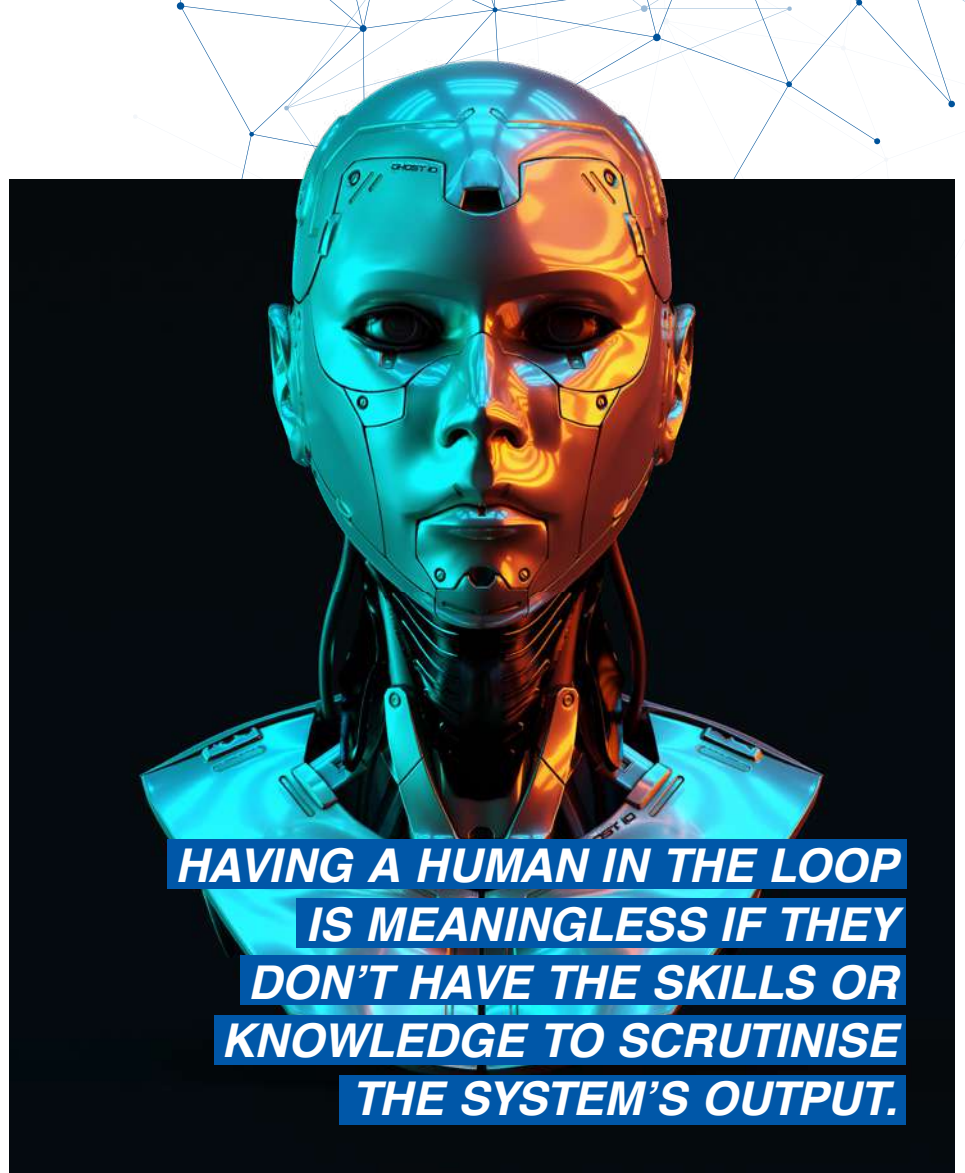
According to the lawsuit, the footage used as the basis for MURPHY's misidentification was of low quality, and may have been based upon a 1980s era mugshot⁴. It occurred after a Sunglass Hut employee shared the CCTV footage with Macy's (a separate retail store chain), who in turn jointly notified police of the identification. Given the circumstances of the case, MURPHY alleges that facial recognition software is the only reasonable explanation as to how he came to be involved in the matter. As at January 2024, no confirmation or otherwise of this aspect has been observed.

Transparency Transparency is a simple concept - be as open as possible without undermining your core mission, including victims' rights.

However, this is not just being open about using AI - go more broad:

- What are you using it for?
- How did you procure it?
- How did you test it?
- What data did you use to train/test it? Whose was it, and how did you procure it?
- How is all data involved in this system stored?
- Who has access to the data and the system itself?
- How are you monitoring performance?
- What will you do when things go wrong?
- Is there a right of appeal for affected parties?

There is a definitive lack of transparency alleged across our scenario (Section 3). Whereas a company/product name is *mentioned* in some reports, the basis for MURPHY's identification itself is not confirmed. We may not expect a specific product or tool to be publicly named, but it does posit the question as to whether



the investigating police, judicial officer issuing the arrest warrant, or the arresting officer were aware as to how MURPHY was identified. Would all of these parties have undertaken their duties in exactly the same way if they'd known?

The CCTV footage was shared between non-law enforcement parties, and the data used to train this unidentified system is completely unknown - though we note MURPHY's inference that a potentially 40 year old mugshot may have been accessible to the system. How much further did this information sharing go? Were customers aware?

Takeout: Be as open as possible about the entire project - not just the use of AI (or otherwise). Often it's the process itself as a whole that needs to be understood.

Human Oversight

Again, go further with this principle - don't only think about oversight of the algorithm. Think about oversight *throughout* the entire system. For our scenario:

- Was the training/test data gathered automatically? Is it accurate?
- Is the output accurate?
- What are the consequences for outputs? Are your users blindly following results, or are they applying suitable levels of scrutiny?

Our scenario, as reported, includes human oversight in the form of judicial review - presumably at the point of warrant issue, but definitely after arrest. However, this does not cover the vast bulk of the overall process. As mentioned previously, there is no indication as to whether any of the people in the loop understood the potential nature of the hypothetical facial recognition system, including the possibility for false positives. Given the completely opaque nature of the system, it is impossible for us to provide any indication of the oversight present during development and operation.

Takeout: The human needs to be in the loop throughout the entire project, not just outputs.

Skills and Knowledge

Having a human in the loop is completely meaningless if they don't have the appropriate skills and knowledge to perform their role. Not everyone needs to be a data scientist, in fact, that'll possibly be your project's smaller demographic. Developers, maintainers, management and users (both immediate and downstream) need to understand the aspects relevant to their roles.

For example, in our aforementioned scenario, this requires:

- Understanding of the facial recognition system's limitations, particularly as they apply to low quality imagery;
- Assuming the human in the loop validated their identification in some way (presumably by checking against a photo held by the system), how strong was their view of a match? Was it sufficient for its purpose in this case?
- Familiarity of the issuing judicial officer and the arresting police officer with automated facial recognition and its limitations? Did they even know it was used in this case?

You will note that many of the required aspects aren't technical, and they're not issues of mathematics. If anything, they're business, legal and process concerns.

Whilst not seeking to criticise the parties involved in the incident, we do ask - would the entire process, including the arrest and subsequent incarceration, have occurred differently if the entire chain of authorities involved had known of the use of automated facial recognition, its limitations, and the imagery's alleged low quality? If the answer is 'yes', was there really a human in the loop?

Takeout: Your human is not in the loop if they don't understand their position. Ensure skills commensurate to the role, plus knowledge of where the role fits in the wider system.

Proportionality and Justifiability

Now, this is where things become less clear-cut, and experience and judgement take the lead.

A 2020 report by Monash University⁵ gives us some hints regarding community expectations - when presented with the question "What are the main reasons for your distrust in the development and use of facial recognition technology in the federal government?", the strongest

responses related to invasions of privacy and being watched (25% and 28% respectively), yet issues around technical accuracy and reliability were regarded by a larger proportion as 'not very important' or 'unimportant'.

Switching over to application, the same study sees an interesting result - when ranked by strength of support for social use cases, every scenario with greater than 50% 'support' relates to policing. In fact, the only law enforcement scenario with less than 50% is "To identify people for minor offences", with 47%.

Justifications for the collection of data come down to the mission, **not** an issue with AI - rather, it comes down to justifications for the bulk collection, storage and access to sensitive data such as people's faces, the privacy implications, and the risks of misuse by both internal (i.e. trusted) users and external parties.

Our scenario raises several issues of proportionality - armed robbery is a serious offence, so if aligning with the aforementioned survey, the use of automated facial recognition in this context would most likely be seen as reasonable by most members of the general public. The capture and sharing of CCTV by private entities perhaps less so, particularly if used for offences such as shoplifting. What the survey does not consider, though, is the justifiability of the outcome based on the AI system, rather than use case. In our scenario, it is alleged the identification was used as the main (if not only) grounds for arrest and several weeks' incarceration, rather than merely a starting point for further police enquiries.

Takeout: Your choice (or otherwise) to use AI may not be the most relevant factor. Measure the risk/rewards of your mission itself, and remember to consider outputs beyond the first step, and across all possible paths.

Explainability

As with any explanation of technology, it is important to consider the relevant audience. Describing how an AI system works from a technical perspective is only one component of explaining its use. Many AI algorithms are resistant to simple explanations along the lines of "this is exactly how the output was generated" because of their inherent complexity. However, in many cases this is probably not the most important aspect of interest.

For instance, it may be as crucial to explain how the data that trained the AI was collected, how results from the AI were interpreted, how the AI integrates with other aspects (both human and automated) of an investigation and/or what safeguards are in place to govern its use.

If our scenario's nature is as reported, it is reasonable to say that beyond the process of incarceration and subsequent release, the overall explainability of how MURPHY came to be arrested in the first place is rather poor - in fact, the lack of explainability appears very much a key point of contention. One can only imagine that if his alibi was unable to be established, a subsequent trial would have focused almost exclusively on this aspect - not just how the hypothetical algorithm came to its conclusion, but what the overall system and process actually were.

Takeout: Consider for whom explanations are being made and avoid unnecessary complexity. Understand and explain how the overall system works, including the safeguards you installed to keep things reliable

Fairness

If a hypothetical facial recognition system used to identify and arrest wanted criminals was 99% accurate, is it fair? To a reasonable person, sure, but it comes down to why that 1% of mistakes is occurring. If it's due to some inherent variability in the system and seems effectively random, well then, the 1 person from every 100 being stopped for formal identification encounters a near one-off inconvenience. But what if that error is due to the system having a bias against something inherent to you? Then *you* become that 1%, and if your life involves regularly passing by a checkpoint related to our theoretical facial recognition system, you're going to get stopped each and every time until someone works out a guardrail to prevent such harassment.

In our scenario, do you regard it as 'fair' if even one person ends up being incarcerated for several weeks on the basis of an incorrect identification? Do you think the storing of involuntary data such as mugshots in private repositories in perpetuity is fair? Your answer may well be 'yes', and in your circumstances, most if not all may well agree, but you need to understand your system's further impacts in order to understand the question.

Takeout: Think beyond performance in percentage accuracy terms. Will your system somehow impact upon or harm people, even if working exactly as designed?

Accountability

Accountability is incorporated into all legislation related to policing, either explicitly or via case law. To our knowledge no legislation specifically placing burdens of responsibility onto technology (and away from humans) currently exists, nor is planned in the foreseeable future.

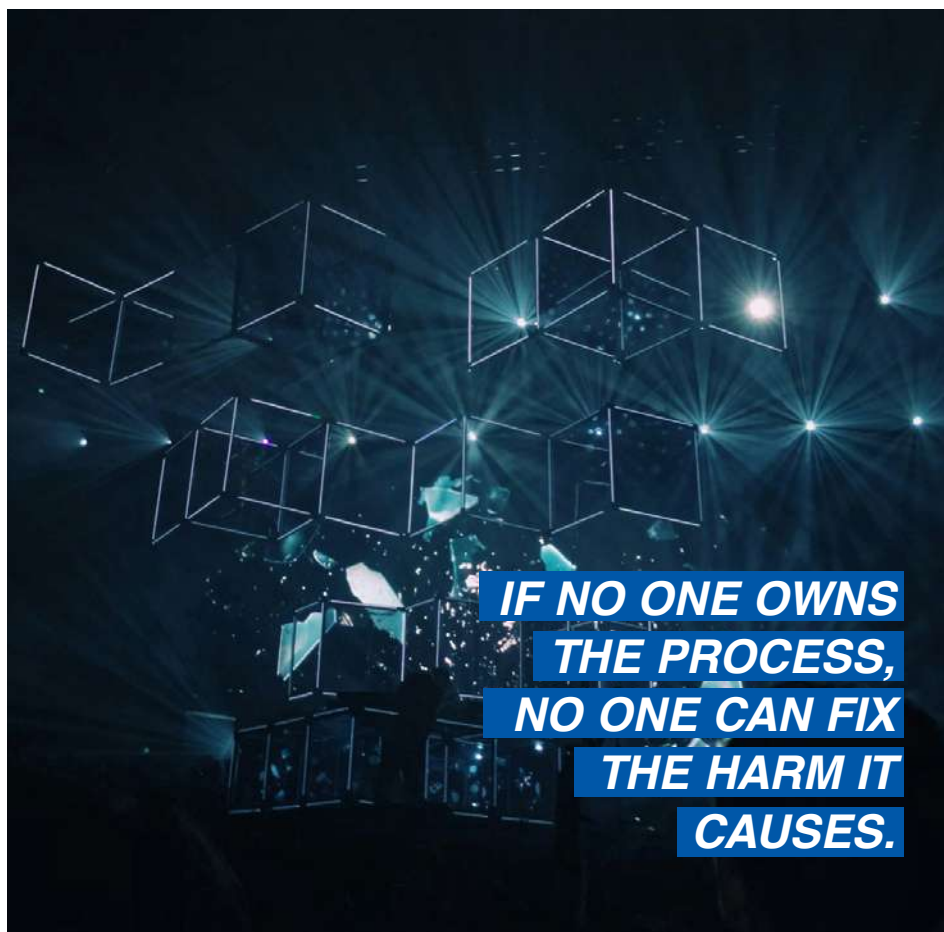
Persons affected by any policing activities have a right of appeal. Automation by any means does not remove this right, necessitating clear roles for any such appeal or internal review. This should not be limited to prosecution actions or other outputs. Ownership needs to be established throughout the project lifecycle, hence the need to ensure responsibility for aspects such as data collection, storage and deletion.

We would suggest most (if not all) of these roles already exist in most policing organisations, just not in this context. Our scenario shows that some rights of appeal exist, both in criminal court (the alibi resulting in MURPHY's release) and civil, through the lawsuit's existence. In this instance, the lawsuit is aimed against Macy's and Sunglass Hut rather than law enforcement, though one could imagine a similar suit being launched against police (and potentially the judiciary) if evidence of negligence or malpractice was established in their acceptance and use of the identification provided to them. The reputation costs of such a case establishing a lack of accountability and ownership over a process including incarceration could ultimately make any financial settlement pale into insignificance.

Takeout: Establish ownership over every project element, not just the technology and data, and do it early.

Privacy and Security

Data privacy and security are key tenets of policing, though have evolved of late. Issues around AI confidentiality, integrity and



accessibility remain unchanged. What has changed, however, is data is now used to train models (aka algorithms) to make the inferences and decisions we want to automate.

What does this mean? Algorithms adapt to data they're trained on. Being mathematics, every change can theoretically be reversed or at least analysed, without necessarily requiring full access to the model itself. As an example - private phone numbers from ChatGPT⁶. So two things - does the data still exist if it can theoretically be reconstructed, and could your model become a back door to your sensitive information?

Beyond our technical type attacks, our scenario raises multiple questions, such as:

- Under what authority were staff able to pass CCTV footage between their organisations?
- Where was the data used in the hypothetical facial recognition system

sourced and stored? Is it accurate, i.e. was the photo used to identify MURPHY actually of him?

- Is there a right to be forgotten? If, as alleged, the system held a nearly 40 year old mug shot photo of MURPHY, should it, particularly if it's a privately owned initiative?

If our system learned from matches (presumably as approved/rejected by users), what happens when there's a mistake? Is the model actually learning incorrect data, effectively making your quality assurance process an actual threat?

Takeouts:

- Your project will involve data throughout its lifecycle. Make sure you know where it's coming from, how you're treating it, who has access to it, how you've secured it, and how it's destroyed.
- Does your system adapt as it operates? If so, is your data actually deleted?

1 <https://open.spotify.com/episode/50gqOdXKYcoEcsyldahg8v?si=UXrgkEdeRYaLxMdijFXrJA>

2 <https://www.anzpa.org.au/resources/publications/australia-new-zealand-police-artificial-intelligence-principles>

3 <https://www.theguardian.com/technology/2024/jan/22/sunglass-hut-facial-recognition-wrongful-arrest-lawsuit>

4 <https://www.vice.com/en/article/man-jailed-raped-and-beaten-after-false-facial-recognition-match-dollar10m-lawsuit-alleges/>

5 Australian Attitudes to Facial Recognition: A National Survey https://www.monash.edu/__data/assets/pdf_file/0011/2211599/Facial-Recognition-Whitepaper-Monash,-ASWG.pdf

6 Scalable Extraction of Training Data from (Production) Language Models, <https://arxiv.org/abs/2311.17035>

McDonald's® Proudly Supporting NSW Police

FREE Medium McCafé® Coffee.

Use this ad as evidence
to receive offer.



VALID TO 31/12/2025 AT ALL McDONALD'S NSW RESTAURANTS

THIS VOUCHER MUST NOT BE DISTRIBUTED TO OR REDEEMED BY A CHILD UNDER 15 YEARS. Hand in this voucher when ordering to receive offer. Limit of one voucher per person per day. Not to be used in conjunction with or to discount any other offer or with a meal purchase. Only bona fide original vouchers will be accepted, no photocopies or images of vouchers are valid. Vouchers are void if tampered with in any way. McDonald's reserves the right in its absolute discretion to verify the validity of this voucher. Not redeemable via McDelivery®, mobile ordering or self-ordering kiosks in Restaurants. Extra charges apply for soy, syrups and other variations. M002856

Proudly supporting local Police

Remembrance Day is an occasion to honour those officers who have lost their lives while performing their duties. It is an important reminder of the continued dangers that our brave men and women in blue face each day in serving the community.



The Hon Sonia Hornery MP

State Member for Wallsend
Deputy Speaker

67 Nelson Street,
Wallsend NSW 2287
wallsend@parliament.nsw.gov.au
(02) 4950 0955

Authorised by Sonia Hornery MP. Funded using parliamentary entitlements.

JEZZINI

PROPERTY SERVICES

PROFESSIONAL FACILITIES MANAGEMENT
SERVICES GOING ABOVE AND BEYOND
FOR OUR COMMUNITY

INTEGRITY . QUALITY . COMMUNITY

A Company You Can Trust

For over 35-years, Jezzini Property Services has consistently delivered superior facilities management services to its clients.

It operates using an Integrated Management System of Quality, Safety and Environmental standards and prides itself in its impeccable, zero-complaints, perfect track record.

Jezzini operates across the following sectors:

Commercial / Office
Government / Local Councils
Hospitality / Entertainment
Industrial

Jezzini Property Services is passionate about its community and committed to maintaining its hard-earned reputation of extremely high-quality standards.

Mob: 0414 907 042
www.jezzini.com.au

*Proudly Supporting our
Police in the community*

Accountable AI Official Appointment

The AFP recognises the critical importance of enhancing technology governance and upholding public trust by ensuring the responsible adoption of new police capabilities.



February 26, 2025

AUSTRALIAN FEDERAL POLICE

afp.gov.au

The policy for the responsible use of AI in government, which took effect on 1 September 2024, sets mandatory requirements for departments and agencies to ensure the responsible use of AI in government.

The AFP is pleased to announce the appointment of the Manager Technology Strategy and Data as the accountable official under this policy. The key responsibilities of this appointment include:

- ensuring the implementation of the policy within the agency
- notifying the Digital Transformation Agency where the agency has identified a new high-risk use case via email

“THE AFP IS COMMITTED TO IMPLEMENTING THE AUSTRALIA NEW ZEALAND POLICING ADVISORY AGENCY’S AI PRINCIPLES.”

- serving as the contact point for whole-of-government AI coordination
- engaging in whole-of-government AI forums and processes
- keeping up to date with changing requirements as they evolve over time.

The AFP is committed to implementing the Australia New Zealand Policing Advisory Agency’s AI Principles and adopting a human-led approach to remain in step with public expectations and uphold responsible AI.

By taking a holistic approach that balances legal, ethical, privacy, public perception, and operational risks and benefits, the AFP can enhance its operational effectiveness while maintaining police legitimacy and safeguarding public safety.

The AFP is currently working on a number of governance measures that will be released as part of its commitment to responsible AI and transparency.



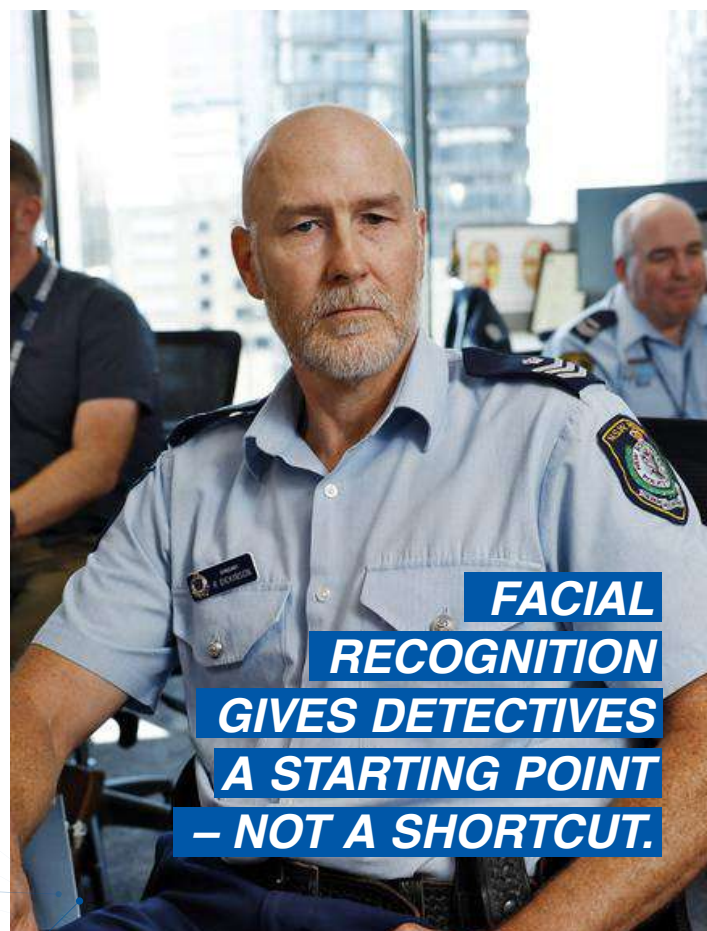
The tech whizzes working around the clock to catch criminals in NSW

Facial recognition is proving to be an invaluable tool for NSW Police, as a team of tech whizzes work around the clock to catch criminals across the state.

June 3, 2025

MARK MORRI

dailytelegraph.com.au



**FACIAL
RECOGNITION
GIVES DETECTIVES
A STARTING POINT
– NOT A SHORTCUT.**

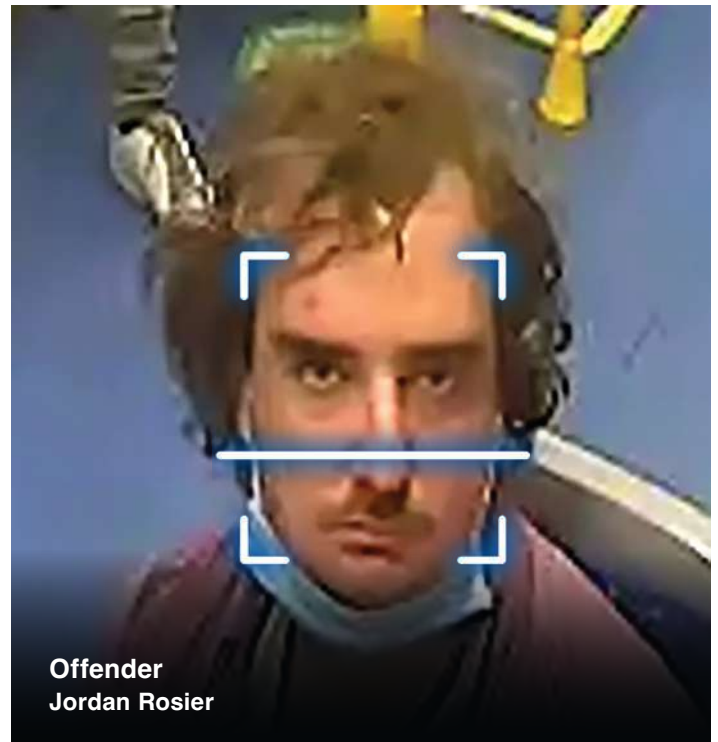


Offender
Keven Dixon

Offence

Sexually touching another person without consent

Charged with inappropriately touching a female victim, 22, on a train at Turramurra in February 2023. Facial recognition was used to identify Dixon and he was charged, convicted and received 24 months' imprisonment.



Offender
Jordan Rosier

Offence

Sexual touching

In 2021, Rosier got on light rail at Haymarket and touched two girls on their thighs. Facial recognition was used to identify him. He was charged, convicted and received 10 months' imprisonment.

Facial recognition software is being used by NSW Police to identify suspects in crimes ranging from violent protests to sexual offences and murders, The Daily Telegraph can reveal for the first time.

The team of tech whizzes –secretly works around the clock, monitoring images of wanted people captured on CCTV cameras, police body cams and social media before feeding them into a NSW Police database that cross-checks against people's mugshots to help identify them.

Detective Superintendent Damien Beaufils, head of the State Intelligence Covert Services branch, told the Telegraph that the Facial Recognition Unit had proved an "invaluable tool" in cracking some major crimes recently.

He said the system gave detectives a starting point in investigations, adding: "We don't have access to driver's licence images or passports -unless we make a request, which has to be justified and processed, and not instantaneous like on TV – it is not used to track people."

The Star and Crown casinos and Venues NSW pubs and clubs as well as Sydney Allianz Stadium, the SCG and Qudos Bank Arena all use real-time facial recognition software to identify barred fans.

In the case of the stadiums, the facial recognition tech can spot a person in their database as they are approaching the venue.

However, the NSW Police are not allowed to use the technology in the same way as these organisations do.

"We are very conscious to identify criminals and also –victims in a lawful manner using facial recognition and have

strict protocols and parameters which we are allowed to operate under," Supt Beaufils said.

The unit has had some recent successes, particularly in the aftermath of the Wakeley riots outside the Good Shepard Church last year.

Facial recognition was also used to help investigators in the alleged attempt to kill John Ibrahim at his Dover Heights home in July last year, leading to the arrest of two men allegedly involved in supplying cars.

They have not entered pleas and remain before the courts.

The team has had increasing success with sex offenders and other crimes on public transport.

"We have a terrific relationship with (public transport operators) and the quality of the CCTV they are able to provide us with has led to a significant number of arrests for those who have committed offences on trains and buses," Supt Beaufils said.

The Facial Recognition Unit's team members follow a strict process, which involves them scrutinising images themselves – as opposed to using computers or AI – before providing leads to help detectives in investigations. "The officers in this unit are highly trained," Supt Beaufils said.

Victims' advocate Howard Brown said it was "crazy" that police did not have access to legally obtained images such as driver's licences to help catch criminals.

"They are collected legally and with the person's consent, to have police limited to a database basically of known criminals is crazy," he said. "Cleanskins commit crimes.

We honour the sacrifice
of police officers
who have given
their lives
in the line of duty



September 29 2025



Janelle Saffin MP

MEMBER FOR LISMORE

lismore@parliament.nsw.gov.au
www.janellesaffin.com.au
02 6621 3624



Authorised by Janelle Saffin MP. Funded using Parliamentary entitlements.

NSW POLICE REMEMBRANCE DAY



On this National Police Remembrance Day, we honour and thank you for your efforts and commitment to justice. Thank you for standing in a blue line that inspires pride, respect, and service to our community, especially during challenging times.

I particularly pay tribute to the police officers in my electorate of Bankstown who put themselves in danger to protect and help our community stay safe. I especially acknowledge those who have made the ultimate sacrifice and offer my prayers to their families.



The Hon. Jihad Dib MP

Member for Bankstown
Minister for Customer Service and
Digital Government
Minister for Emergency Services
Minister for Youth Justice

Shop 21/1 Broadway Plaza
Punchbowl NSW 2196

P: 9759 5000 **F:** 9759 1945
E: Bankstown@parliament.nsw.gov.au

Authorised by Jihad Dib MP
Funded using Parliamentary entitlements.

NATIONAL POLICE REMEMBRANCE DAY



On **National Police Remembrance Day**, we pause to reflect on the ultimate sacrifice made by police officers in the line of duty. We honour the courage, commitment, and selfless dedication of the men and women who have died while serving our communities—protecting the safety and wellbeing of others. As a former police officer, this day holds deep personal significance. I stand with all those remembering a colleague, friend, or loved one lost in service. May we never forget the names etched on the National Police Memorial, and may their legacy live on through our enduring respect and gratitude. **Lest we forget.**



PHIL DONATO MP
STATE MEMBER FOR ORANGE

02 6362 5199 orange@parliament.nsw.gov.au 194a Lords Place Orange NSW 2800

Authorised by Phil Donato MP. Member for Orange. Funded using parliamentary entitlements.

**LUKE
GOSLING**
FEDERAL MP FOR SOLOMON
DARWIN + PALMERSTON



Thank you for your service.

Today and everyday.

(08) 8928 0180
 Luke.Gosling.MP@aph.gov.au
 3/266 Trower Road, Casuarina, NT 0810
 Luke Gosling MP



Authorised by L. Gosling, ALP, 3/266 Trower Road, Casuarina NT 0810

ADVERTISEMENT

Thank you for your service to our community



We salute your
dedication and
commitment to
public safety

Ed Husic MP

FEDERAL MEMBER FOR CHIFLEY

Office: Shop 41 Plumpton Marketplace,
Cnr Hyatts & Jersey Road, Plumpton NSW 2761

Email: contact@edhusic.com **Phone:** (02) 9625 4344

f [ehusic](#) **@** [@edhusicmp](#) **edhusic.com**

Authorised by Ed Husic MP, ALP, Shop 41 Plumpton Marketplace,
Cnr Hyatts & Jersey Road, Plumpton NSW 2761.



**" In honour of NSW Police Officers
whose lives were cut tragically short while
preforming their duty in protecting and
serving their community. "**

Robert Borsak & Mark Banasiak
Members of the NSW Legislative Council



Authorised by Robert Borsak MLC & Mark Banasiak MLC,
Parliament House Macquarie Street Sydney NSW 2000
Funded using Parliamentary entitlements

Dr Joe McGirr MP

INDEPENDENT MEMBER FOR WAGGA WAGGA



**Proudly supporting our
local NSW Police**

64 Baylis Street, Wagga Wagga

(02) 6921 1622

waggawagga@parliament.nsw.gov.au

www.joemcgirr.com.au

Authorised by Dr Joe McGirr MP, Member for Wagga Wagga, 64 Baylis Street, Wagga Wagga

Thank you for protecting our community

MELISSA McINTOSH MP

FEDERAL MEMBER FOR LINDSAY



Authorised by Melissa McIntosh, Liberal Party of Australia, 331 High Street, Penrith NSW 2750.

A: 331 High Street, Penrith NSW 2750
P: (02) 4722 0600 **f:** [MelissaMcIntoshMP](#)

E: melissa.mcintosh.mp@aph.gov.au

W: melissamcintosh.com.au



PROUDLY SUPPORTING OUR LOCAL POLICE AND
ROAD SAFETY

For bookings please call

Mob: 0476 649 616

Find us on Facebook

Email:

weighemall@outlook.com

For a full breakdown of our
services and pricing list please visit

www.weighemall.com.au

At Weigh 'Em All, we provide a reliable mobile vehicle weighing service based in Perth
and surrounding areas, catering to vehicles up to 8 tonnes.

OUR SERVICES

**- Mobile weighing service for: caravans, boats, horse floats, 4x4 touring vehicles
and race cars - Full weight report on completion. - Fully mobile, we come to you.**

Your safety is our priority. We ensure your setup is within legal weight limits
—right down to the kilo—for a safer, more comfortable towing experience.



THE NATIONALS
for Regional NSW

**On behalf of the Coffs Harbour
electorate I would like to
honour all members of
the NSW Police Force who
have lost their lives in the
line of duty.**

GURMESH SINGH MP

Member for Coffs Harbour

Shadow Minister for Emergency Services,
Regional Health and for the North Coast
Deputy Leader of The Nationals

coffsharbour@parliament.nsw.gov.au
(02) 6652 6500

Authorised by Gurmish Singh MP, 1/3 Park Avenue Coffs Harbour NSW 2450
Funded using Parliamentary Entitlements



"WHERE *caring* IS A DOING WORD"

"SOCIAL COMMUNITY HUB" - BULLABURRA

Offering a unique approach
to **disability support.**

Normalising the need for help.

- ✓ In-Home Support
- ✓ Community & Social Access
- ✓ Accessible Transport
- ✓ Day Program
- ✓ Respite & SIL
- ✓ Highly Trained & Skilled Staff
- ✓ Nursing

☎ 1800 00 5557

✉ office@live-life-learn.com.au

✉ PO Box 221, Blaxland NSW 2774

📍 Blue Mountains, NSW, Australia



Trends & issues in crime and criminal justice

Artificial intelligence and child sexual abuse: A rapid evidence assessment



**THE SCOPE AND SCALE OF
CHILD SEXUAL ABUSE—
PARTICULARLY ONLINE—
HAS BECOME TOO
GREAT FOR HUMAN-LED
APPROACHES ALONE.**

January 11, 2025

HEATHER WOLBERS, TIMOTHY CUBIT, MICHAEL JOHN CAHILL

Australian Institute of Criminology

Abstract

This study examined the intersection of artificial intelligence (AI) and child sexual abuse (CSA), employing a rapid evidence assessment of research on the uses of AI for the prevention and disruption of CSA, and the ways in

which AI is used in CSA offending. Research from January 2010 to March 2024 was reviewed, identifying 33 empirical studies.

All studies that met inclusion criteria examined AI for CSA prevention and disruption—specifically, how technology can be used to detect or

investigate child sexual abuse material or child sexual offenders. There were no studies examining the uses of AI in CSA offending.

This paper describes the state of current research at the intersection of AI and CSA, and provides a gap map to guide future research.

Introduction

In recent years, the development of artificial intelligence (AI) has rapidly increased, with availability of this technology significantly expanding. Development of AI technologies has extended to the field of child sexual abuse (CSA), with the scope and scale of the problem—particularly online—becoming too great for manual human-led approaches to manage effectively. However, the use of AI in this field has

extended beyond prevention. In early 2023, the US National Center for Missing and Exploited Children received reports of ‘fake’ child sexual abuse materials (CSAM) that offenders had produced with the assistance of generative AI tools (Murphy 2023). Similarly, Australia’s eSafety Commissioner has noted reports of children using AI to generate sexually explicit images of other children, suggesting that it was an indication of a more widespread issue (Long 2023).

This review considers the current state of research literature studying the use of AI in the field of CSA, focusing on studies investigating the use of AI for offending and the prevention and disruption of CSA.

Artificial intelligence for child sexual abuse offending

According to media reports, surveys and academic reviews, AI technologies are increasingly playing diverse roles in the

creation of CSAM, including fabrication (eg CSAM deepfakes), 'nudifying' pictures of clothed children, and manipulating images or videos to depict known or unknown children in sexually abusive scenarios (Milmo 2023; Okolie 2023). Existing CSAM has been used to train AI models, meaning offenders are using AI to produce new depictions of previously abused children. Further, reports indicate that offenders are using AI to alter photos from victims' social media and other online posts and using these altered images to sexually extort the victims (Garriss & DeMarco 2023). According to a survey of 1,040 people aged nine to 17 years in the United States, one in 10

(11%) minors said they knew of cases where their peers had used AI to create sexually explicit images of other minors (Thorn 2024b).

The Internet Watch Foundation found that, in a single month, 20,254 AI-generated images were posted to a CSAM forum on the darknet (Internet Watch Foundation 2023). Concerns have been raised that the ability to generate CSAM using AI could support an increase in CSAM consumption. Growth in AI-generated CSAM creates significant challenges for law enforcement, who work to detect and prevent the distribution of CSAM online. Ultimately, increases in the volume of CSAM online may influence the ability to investigate CSA, as AI-generated can be indistinguishable from real CSAM (Theil, Stroebe & Portnoff 2023). The malicious use of AI technologies for the production of CSAM is growing and is likely to continue to grow without multi-sector intervention (Theil, Stroebe & Portnoff 2023).

Artificial intelligence for the prevention and disruption of child sexual abuse

As identified in academic research, there are a diverse range of cyber strategies used to combat online CSA (Edwards et al. 2021; Singh & Nambiar 2024). As the field of AI continues to develop, so too does the development of CSA disruption strategies that use AI. For example, published research has shown that AI could be used to identify suspicious financial transactions procuring CSA (eg Cubitt, Napier & Brown 2021; Henseler & de Wolf 2019) or aid in law enforcement investigations by examining

CSAM (eg Brewer et al. 2023; Dalins et al. 2018; Westlake et al. 2022). AI technologies may ease the burden on law enforcement by reducing the risk of psychological harms among CSA investigators (Puentes et al. 2023), while increasing the capacity and timeliness of investigations. Further, AI technologies can have a much larger reach across online spaces than traditional methods of prevention and disruption.

While CSAM can be detected and removed across online spaces with hashes (ie unique digital fingerprints), this method is limited to known CSAM on platforms proactively using hashes, meaning there is limited efficacy and it cannot stop the upload and proliferation of undetected, new or edited content. AI has the potential to help address this challenge. For instance, Thorn has developed a machine learning tool to automatically detect, review and report CSAM at scale (Thorn 2024a). This tool is used to screen new content that gets uploaded to Flickr and other online platforms—a task too large for human moderation alone. Beyond detecting CSAM, AI has a range of potential uses for addressing CSA. These include conversation analysis, chatbots, honeypots and web crawlers, all of which show promise in combating CSA, albeit with some significant limitations (eg narrow scope or generalisability, privacy and legal concerns, and lack of robust evaluation; Singh & Nambiar 2024).

Study aims

AI is the ability of a computer system to simulate human intelligence, such as learning, problem solving, reasoning and decision making—all with some level of autonomy (High-Level Expert Group on Artificial Intelligence 2019). There are several domains of AI, including machine learning (in which an algorithm is trained to learn patterns in existing data), computer vision (interpreting visual information), natural language processing (understanding and generating human language), and generative AI (creating original content). We considered the domains of AI to develop search criteria and identify literature examining AI and its intersection with CSA.

This study aims to establish current AI capabilities in relation to CSA, with the intention of identifying key areas of progress and gaps where further

research is required. A rapid evidence assessment was conducted to address the following research questions:

- What are the uses of AI as a part of CSA offending and what are the areas of future risk?
- What are the uses of AI as a part of CSA prevention and disruption, and what are the areas of future potential?
- What are the key gaps in current research that should be addressed?

Method

Search strategy

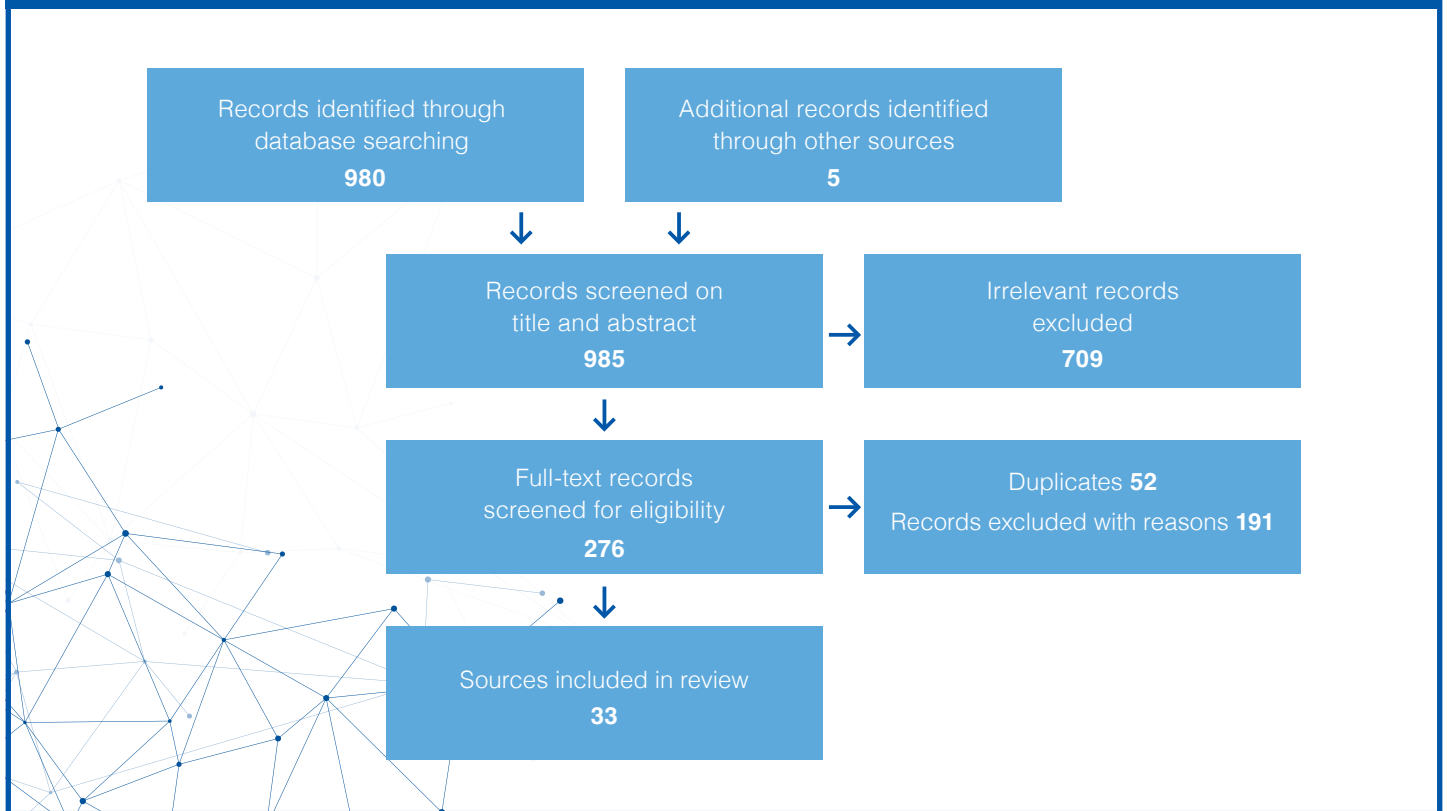
Rapid evidence assessments are accelerated systematic reviews of research undertaken

in a restrictive time frame. This study draws on research published in English between January 2010 and March 2024. Studies were excluded if they did not discuss AI in the context of CSA, include primary data (ie reviews or conceptual studies), explain the study's methodology in sufficient detail (eg if they did not detail the data source, the sample or data management for analysis), or have a direct application to CSA. We excluded studies examining the use of AI in medical settings to detect CSA, as a systematic review was recently published on this topic, which identified seven studies that examined the use of AI for predicting child abuse and neglect using medical or protective service data (Lupariello et al. 2023). Of note, our search yielded seven studies from medical settings, all of which were excluded for secondary reasons (ie they did not focus on detection or prevention of CSA).

The Australian Institute of Criminology's JV Barry Library searched 13 databases and relevant websites: the JV Barry Library catalogue, the Australian Institute of Criminology, EBSCO, ProQuest Criminal Justice, DeepDyve, arXiv.org, IEEE Xplore digital library, Office of the eSafety Commissioner, International Centre for Missing and Exploited Children, National Center for Missing and Exploited Children, Australian Centre to Counter Child Exploitation, Thorn and Google Scholar.

The search terms used combined keywords from three categories capturing AI and its relevant sub-domains, child sexual abuse, and specifying the group of concern to be individuals under 18 years of age:

Figure 1: Literature screening process



- Artificial intelligence ("machine learning" OR "artificial intelligence" OR algorithm* OR "deep learning" OR "unsupervised learning" OR "reinforcement learning" OR "generative artificial intelligence" OR "natural language processing" OR "computer vision" OR chatbot OR "image classification" OR "object detection" OR "augmented reality" OR "big data" OR "neural network");
- Child sexual abuse ("child sexual abuse" OR CSAM OR CSEM OR "child abuse material" OR "live streaming" OR "child exploitation" OR "child sexual abuse material" OR "child exploitation material" OR "image-based sexual abuse" OR "image-based abuse" OR "technology-facilitated sexual violence" OR online "sexual exploitation of children" OR "child pornography" OR "indecent images of children" OR grooming); and
- Child (child OR children OR "young person" OR "young people" OR adolescent OR teenage* OR youth OR minor OR "young adult").

Screening process

The rapid evidence search yielded 980 records. Five additional studies were identified through wider reading, resulting in identification of an initial 985 records. Titles and abstracts were screened to exclude irrelevant studies (n=709). This screening yielded 276 records, of which 52 were identified as duplicates and removed. The remaining 224 sources were assessed for eligibility against the selection criteria with full-text screening, and 191 were excluded because they did not meet the selection criteria. In total, the search yielded 33 sources providing primary information on the role of AI in relation to CSA (see Figure 1).

Limitations

Rapid evidence assessments do not provide the same exhaustive depth or detail as a full systematic review (Ganann, Ciliska & Thomas 2010). Databases that yielded a large number of hits were not searched in full. Rather, the first several hundred items returned by the search were screened, meaning that the most relevant sources were captured. However, search results were not screened exhaustively. Given that search results were presented in order

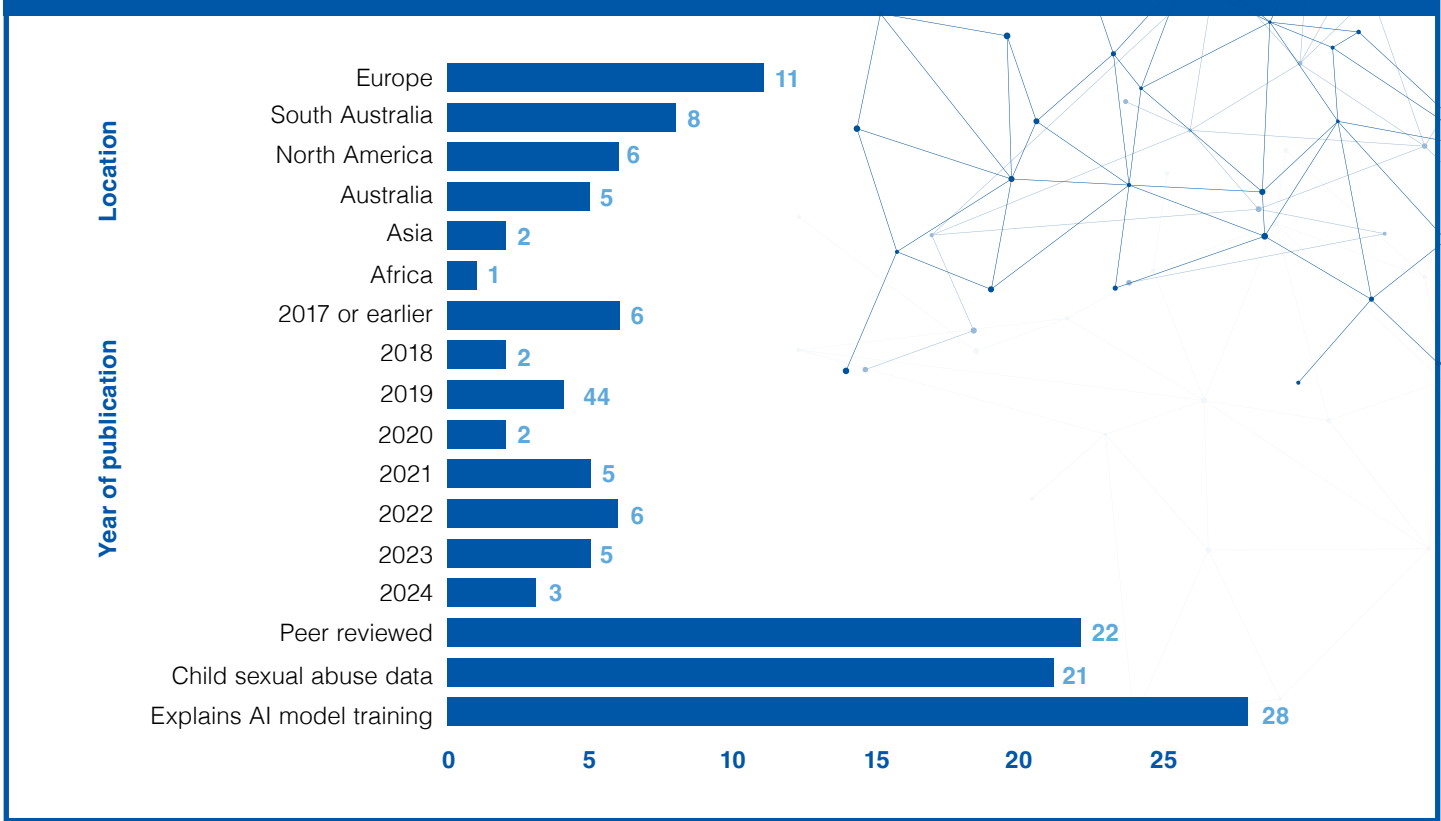
of relevance to the search terms used, the number of sources missed by using this methodology is likely limited.

Results

Study characteristics

Research at the intersection of AI and CSA was identified from Europe (n=11), South America (n=8), North America (n=6), Australia (n=5), Asia (n=2) and Africa (n=1). Identified research was published between 2011 to 2024, with a notable increase from 2020 onward. Two-thirds were peer reviewed (n=22). Non-peer reviewed sources included conference papers (n=9) or pre-prints (n=2). While research primarily relied on data from cases of CSA (ie CSAM files, or information on convicted CSA offenders), five studies used data on suspected rather than proven child sexual exploitation (eg reports to a hotline, risky online conversations). Five studies relied on interactions between suspected or known offenders and adults posing as children, and two relied on peripheral datasets using semi-nude non-sexual images of children and a database of faces for age estimation. The majority of included studies explained how their AI model was trained (n=28), with five using

Figure 2: Characteristics of included studies (n=33)



unsupervised modelling approaches. All identified research examined uses of AI for the prevention, disruption, detection or investigation of CSA, with no studies examining the uses of AI to perpetrate CSA. As reports of malicious use of AI to perpetrate CSA only began to emerge recently (Long 2023; Murphy 2023), it seems unlikely that enough time has passed for empirical research to have been produced on this subject matter. Quality of evidence regarding efficacy of artificial intelligence in the context of child sexual abuse

Quality of evidence regarding efficacy of artificial intelligence in the context of child sexual abuse

Among the 33 reviewed studies, 28 attempted to evaluate the discussed AI tool—typically considering accuracy, precision, recall or another metric specific to the intended goal. When a tool was evaluated, the employed metrics and study aims were diverse, meaning cross-study comparisons of efficacy were not possible. Additionally, the data or sample used for testing—particularly when the data were synthetically produced, or relied on a small or non-generalisable sample—raised questions about

how well the model would translate to a real-world setting. Only a minority of studies tested the tool in a real-world or near real-world setting (Brewer et al. 2023; Dalins et al. 2018; Guerra & Westlake 2021; Jin et al. 2024; Ngo et al. 2024; Peersman et al. 2016; Westlake et al. 2022). Due to these limitations, any findings regarding efficacy for each included study should be interpreted with some caution. For example, while a tool may demonstrate a high level of efficacy, this does not necessarily mean the tool would perform well if applied in the real world.

Artificial intelligence and child sexual abuse

Table 1 presents a summary of the AI tools discussed in published literature. These included tools used for detection, examination or investigation of CSAM or child sexual offenders online.

Artificial intelligence for detecting and investigating child sexual abuse material

The most common use of AI in the identified research was to detect CSAM. The primary intention of the AI tools studied was to reduce the burden of manual processing of CSAM by investigators, thereby mitigating mental

health impacts, while reducing the time needed to identify CSAM among very large datasets. Tools designed to detect CSAM tended to combine age identifiers (Dalins et al. 2018; Gangwar et al. 2021; Macedo, Costa & dos Santos 2018; Rondeau et al. 2022; Sae-Bae et al. 2014), with skin tone, nudity or pornography detectors (Dalins et al. 2018; Gangwar et al. 2021; Laranjeira et al. 2022; Macedo, Costa & dos Santos 2018; Oronowicz-Jaśkowiak et al. 2024; Peersman et al. 2016; Polastro & Eleuterio 2012; Rondeau et al. 2022; Sae-Bae et al. 2014).

Other models detected CSAM by analysing the words used to describe the picture (Peersman et al. 2016; Ulges & Stahl 2011) or language embedded in audio (Peersman et al. 2016). Three studies discussed tools that analyse file names or file paths to estimate the likelihood of them containing CSAM (AI-Nabki et al. 2023; Peersman et al. 2016; Pereira et al. 2021). Importantly, tools were frequently designed to be implemented in a specific setting, such as on peer-to-peer networking websites (eg Peersman et al. 2016) or on a personal computer or device (eg Polastro & Eleuterio 2012). Many of these studies used authentic CSA data

Table 1: Summary of how artificial intelligence is used in the context of child sexual abuse in the studies reviewed

Intention of the tool	Approach used
Detect CSAM	Detect CSAM on personal computers using file names and file paths (2, 22, 23)
	Detect CSAM using a combination of tools such as pornography detection, age estimation, skin tone/nudity identifier (8, 9, 16, 21, 22, 24, 25, 29, 32)
	Separate CSAM into discrete categories by type (8, 15)
Aid in the investigation of CSAM	Determine the age of children in CSAM (11)
	Match victims and offenders across CSAM videos using facial and voice recognition (4, 33)
	Identify patterns in the locations and folder or file naming practices of websites with CSAM (12)
Detect online child sexual offenders	Analyse the language used in online conversations to identify threats to children (1, 3, 5, 13, 17, 26, 30)
	Analyse conversations between children and offenders to identify whether offenders intend to contact offend or not (31)
	Distinguish real children from adults pretending to be children in chat rooms (17)
	Using a chatbot to interact with suspects and profile their interest in CSAM (18, 27)
	Scrape hashtags and images from tweets in real time to detect suspected human trafficking of minors (10)
Aid in understanding online child sexual offenders	Crawl the darknet to collect data on the behaviours of child sexual offenders who access and participate on dark websites (14)
	Analyse posts about CSAM and associated metadata to understand the characteristics, behaviours and motivations of CSAM creators (20)
	Understand the characteristics and typologies of offenders who live stream CSA (6, 7)
	Detect other files shared online by individuals who have shared known CSAM files (22)
Other	Analyse text-based reports of child sexual abuse (25)

Note: 1—Agarwal et al. 2022; 2—Al-Nabki et al. 2023; 3—Anderson et al. 2019; 4—Brewer et al. 2023; 5—Cardei & Rebedea 2017; 6—Cubitt, Napier & Brown 2021; 7—Cubitt, Napier & Brown 2023; 8—Dalins et al. 2018; 9—Gangwar et al. 2021; 10—Granizo et al. 2020; 11—Grubl & Lallie 2022; 12—Guerra & Westlake 2021; 13—Isaza et al. 2022; 14—Jin et al. 2024; 15—Laranjeira et al. 2022; 16—Macedo, Costa & dos Santos 2018; 17—Meyer 2015; 18—Murcia Triviño et al. 2019; 19—Ngejane et al. 2021; 20—Ngo et al. 2024; 21—Oronowicz-Jaśkowiak et al. 2024; 22—Peersman et al. 2016; 23—Pereira et al. 2021; 24—Polastro & Eleuterio 2012; 25—Puentes et al. 2023; 26—Razi et al. 2023; 27—Rodríguez et al. 2020; 28—Rondeau et al. 2022; 29—Sae-Bae et al. 2014; 30—Seedall, MacFarlane & Holmes 2019; 31—Seigfried-Spellar et al. 2019; 32—Ulges & Stahl 2011; 33—Westlake et al. 2022

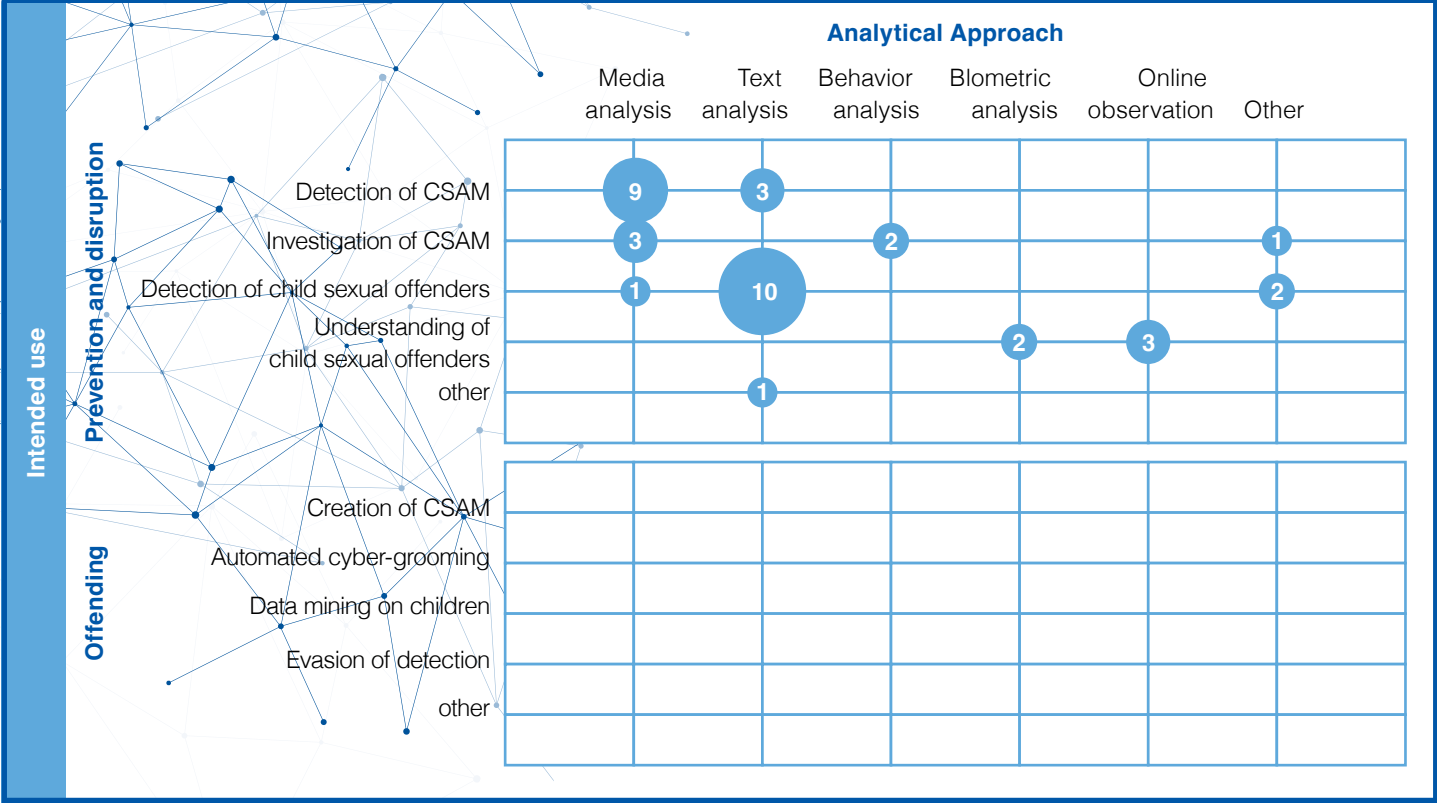
sources and reported that the tool of focus performed well at completing the intended task (Al-Nabki et al. 2023; Gangwar et al. 2021; Oronowicz-Jaśkowiak et al. 2024; Peersman et al. 2016; Pereira et al. 2021; Polastro & Eleuterio 2012). However, others showed limited efficacy (Dalins et al. 2018; Puentes et al. 2023; Ulges & Stahl 2011).

Notably, Peersman and colleagues (2016) described a toolkit that performed

multiple functions. While designed primarily to detect CSAM on peer-to-peer networks by analysing filenames, images and audio, the tool also flagged files shared by individuals who have shared known CSAM. This tool was evaluated using real-world CSA case data, demonstrating usefulness in law enforcement settings and considerable accuracy in detecting CSAM when combining filename and image classification.

Beyond the detection of CSAM, several studies discussed uses of AI to assist investigations in other ways. One study aimed to categorise the content of CSAM to aid with triaging (eg solo, non-penetrative, penetrative; Dalins et al. 2018), while another extracted features and labels of CSAM to describe the content without it ever being viewed (Laranjeira et al. 2022). Another important use was to extract and match the biometric features

Figure 3: Evidence and gap map for research on the uses of AI in the context of child sexual abuse



of victims and perpetrators shown in CSAM, allowing for a rapid detection of media associated with an investigation and the identification of links between files (Brewer et al. 2023; Westlake et al. 2022). Finally, a web crawler was designed to find patterns in the locations and CSAM naming conventions of websites with known CSAM (Guerra & Westlake 2021). Three of these studies tested the performance of the AI tool (Dalins et al. 2018; Laranjeira et al. 2022; Westlake et al. 2022), and just one demonstrated strong results. Specifically, Westlake and colleagues (2022) were able to identify and match victims and offenders across a test sample of authentic CSAM files with a high true match rate (between 93.8% and 98.8%) and a low false match rate (between 1.0% and 5.0%).

Artificial intelligence for detecting and understanding child sexual abuse offenders

Two studies described a tool designed to initiate and hold conversations with potential online CSA offenders (ie a chatbot; Murcia Triviño et al. 2019; Rodríguez et al. 2020). This chatbot used generative and rule-based models to produce conversational posts and replies. Based on these interactions,

the tool then described each suspect's behaviour, classifying their disposition towards online child sexual offending as indifferent, interested or perverted. The efficacy of this classification was not numerically measured.

Several studies described AI methods designed to understand child sexual offenders through their online behaviours. A tool produced by Granizo and colleagues (2020) scraped posts from X (formerly Twitter) in real time to identify suspected cases of child trafficking. This tool demonstrated some efficacy at recognising the gender and age of individuals depicted in images by analysing their faces or torso.

Two studies discussed tools operating on the darknet. The first study discussed a web crawler that collected data on darknet ecosystems, detecting relevant content and providing information designed to reduce the anonymity of perpetrators, such as links to the surface web that may be used to trace darknet operators (Jin et al. 2024). Similarly, a tool developed by Ngo and colleagues (2024) processed CSAM discussions on the darknet and provided insights into the characteristics, behaviours and motivation of CSAM creators. Both tools performed well in identifying the content of interest

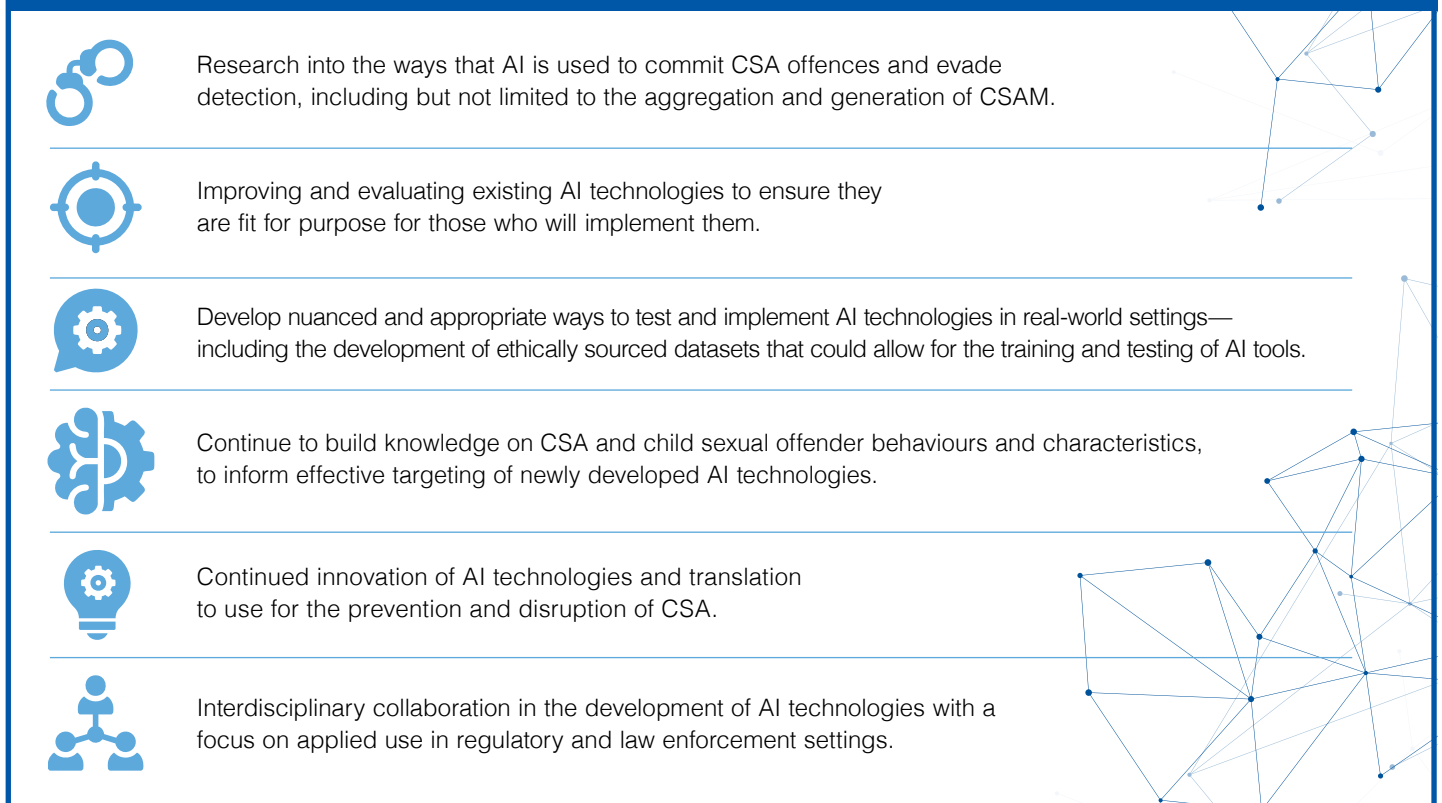
and were able to reveal meaningful information about online offending environments and offenders.

Peersman and colleagues (2016) described a model to detect other online files shared by those known to distribute CSAM online, while two studies used machine learning to analyse the characteristics and offending history of individuals who live streamed CSA (Cubitt, Napier & Brown 2023, 2021). In the 2021 study, the model had notable success in identifying individuals who would engage in prolific live streaming of CSA, successfully classifying more than 80 percent of cases (AUROC=0.85; Cubitt, Napier & Brown 2021).

Other uses of artificial intelligence in the context of child sexual abuse

One final tool used a large language model to identify the subject, degree of criminality, and level of impact relating to reports of CSA to a hotline (Puentes et al. 2023). Using this method, the authors aimed to automate the analysis of CSA reports, consequently expediting the process while reducing the exposure of analysts to potentially harmful content. The authors concluded that the approach was an appropriate starting point, but the efficacy was limited.

Figure 4: Summary of future research required on AI technologies in the context of child sexual abuse



Evidence and gap map

We constructed an evidence and gap map plotting the most common analytical domains of AI that were reported, and how they were used in the context of CSA (Figure 3). Importantly, these do not represent all possible AI capabilities and uses in the context of CSA. The analytical procedure and intended use categories broadly capture the most common AI capabilities and uses outlined in the research, but the figure does not critique the quality of evidence or efficacy of the AI technologies discussed. Dots on the graph show where the analytical approach and the intended uses intersect, with the size of the dot reflecting the number of studies. Intersections without dots indicate an absence of evidence, highlighting areas requiring further research.

The evidence and gap map highlighted a significant gap in evidence relating to how AI is used in the perpetration of child sexual offending, with no literature returned in this search. The two most developed areas of research focused on analysis of media (images, audio and videos) to detect CSAM, and text analysis to detect child sexual offenders.

Discussion

At the time of this review, research has examined AI tools used to detect CSAM or CSA offenders, to aid with investigations, and to improve our understanding of online environments where CSA is produced and shared. The two most common uses of AI were analysing images, audio or video to detect CSAM without requiring humans to view the content, and language processing to detect child sexual offenders through online conversations. Of note, several of the tools described undertook more than one task. For example, some tools were designed to both detect and categorise CSAM (Dalins et al. 2018), to detect CSAM by classifying media content and separately the text of file names (Peersman et al. 2016), or to consider both text and images simultaneously (Granizo et al. 2020).

The introduction of AI technologies in place of human decision-making offers important opportunities (Singh & Nambiar 2024). Benefits include automatically classifying CSAM images and improving the efficiency of detection or classification of images and videos when large volumes of data are obtained. These functions have the potential to

reduce the risk of psychological harm to investigators. The automated nature of these tools is particularly important given the demands placed on law enforcement by the recent dramatic growth in CSAM production and sharing. The rate of online sharing and viewing of CSAM is currently beyond manual human detection and intervention. Ultimately, this may mean that a human response alone is not an adequate solution to this increasingly AI driven problem, and opportunities to integrate AI technology into existing CSA prevention strategies should be explored.

Directions for research

The principal gap in research identified by this review was the use of AI for CSA offending. Further, the studies examined indicated that several of the AI technologies proposed for the prevention or disruption of CSA were not fit for purpose in their current form or did not have sufficient evidence to support their efficacy in a real-world setting. Figure 4 provides a summary of important areas for future research at the intersection of AI and CSA, informed by the evidence and gap map in Figure 3.

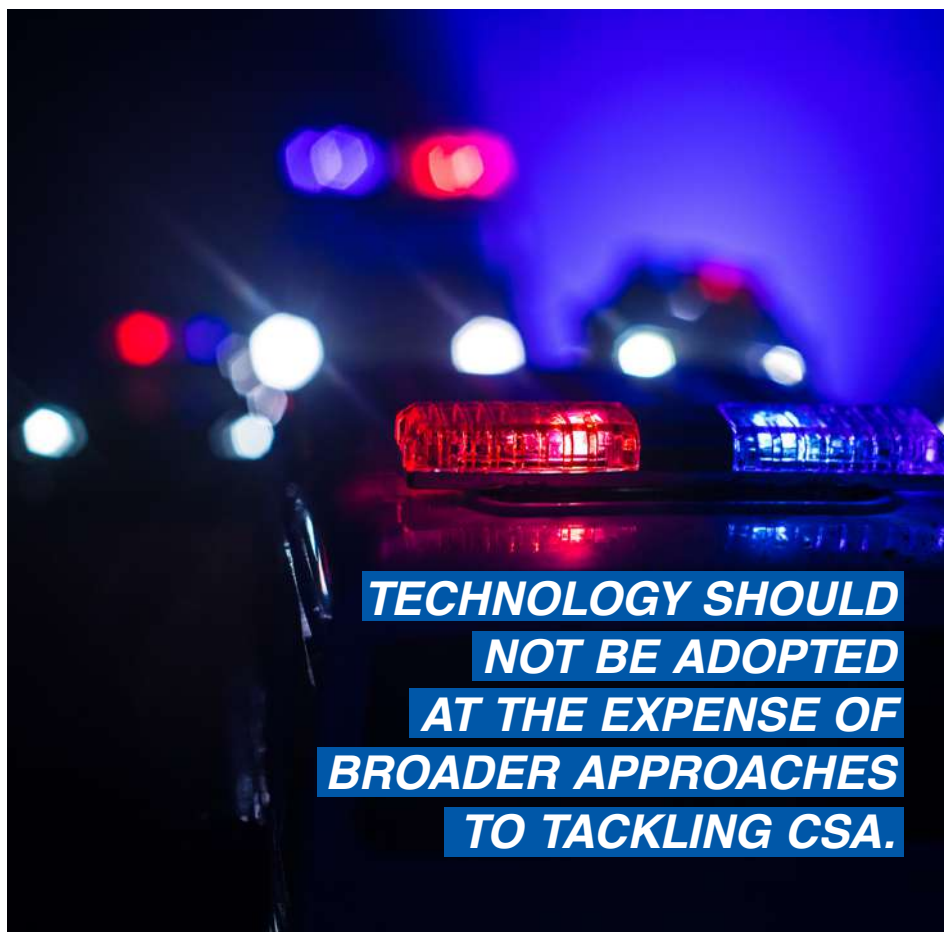
While there is evidence that AI is being used in the process of child sexual offending, particularly CSAM

offending, a trend that appears to be expanding (Internet Watch Foundation 2023), there were no identified studies investigating the nature and scope of AI use among offenders. Anecdotal evidence suggests AI is being used to generate deepfake CSAM or could automate cyber-grooming (Butler 2023; Internet Watch Foundation 2023). However, there is limited understanding in the research literature of the scope of this problem or emerging uses of AI for malicious purposes.

It is important to implement methods to deter the use of AI for illicit activities. Research should continue to explore the uses of AI among child sexual offenders to better understand the risks posed and ways to address these risks.

While there was a sizable amount of research on the development of AI technologies for CSA prevention and disruption, the research evaluating the efficacy of these models, or how they performed in real-world settings, was limited. It is important to note that tools should only be implemented after their performance has been evaluated. This evaluation should measure the tool's effectiveness (eg accuracy, reliability, specificity), as well as its application (ie whether potential users, such as law enforcement, can employ it and find it helpful). It may also be helpful to measure performance with different data sources or for different tasks, to clarify where models perform well and where they do not.

However, the adoption of AI technology alone for CSA prevention and disruption is unlikely to be a comprehensive solution for either CSA in general or AI produced CSAM. Tech-solutionism, in which technology is implemented as a standalone method of solving a given issue, is often criticised for oversimplifying complex problems, failing to address root causes, and leading to unforeseen negative consequences. Additionally, there are reasonable concerns regarding the use of AI prevention methods without supervision or validation by humans. If AI technology were to be found suitable for use, it should not be adopted at the expense of broader approaches to tackling CSA. Rather, it should be implemented as just one tactic among many to reduce the volume and impact of CSA. Detection and prevention approaches, when featuring AI, should continue to be transparent, should feature a degree of human supervision and should be considered part of a suite of complementary approaches to address CSA, rather than a standalone solution.



Developing and testing AI technologies in the context of CSA prevention may require the use of authentic CSA data sources that reflect real-world settings—for example, CSAM, offender chat logs and police case reports. Of course, there are important ethical and practical implications when building and accessing such data sources. It is appropriate that access to these data are tightly controlled; however, the difficulty accessing data is a significant barrier to assessing whether these AI tools may be useful and implementable. Progress in the field of AI for CSA prevention may therefore require consideration of how researchers can reliably evaluate the efficacy of their tool using appropriate datasets, under agreed upon conditions for access. Additionally, AI models could be advanced by improving knowledge of child sexual offender behaviour (Singh & Nambiar 2024).

AI technology will continue to develop rapidly. While caution should be exercised in implementing these models, innovative approaches to addressing CSA should be encouraged. Each of the directions suggested for future research would substantially benefit from interdisciplinary collaboration, particularly featuring the stakeholders who would ultimately use the technology.

Conclusion

The research literature has, to date, detailed a range of AI technologies developed with the aim of preventing or disrupting CSA—most commonly, those that detect CSAM or online child sexual offenders. The use of AI to address CSA is an emerging field, and while the evidence for the efficacy of AI technology in this context is limited, the field is moving and developing rapidly. Ultimately, with AI supported CSA offending becoming more widely reported, interest in AI approaches to prevent CSA is likely to grow. This review has emphasised the potential for AI technologies to identify, prevent and disrupt CSA. These technologies offer many advantages but must undergo strict evaluation and safety testing and adhere to ethical protocols before being considered for adoption to complement existing strategies.

Acknowledgements

This research was conducted as part of the National Office for Child Safety's Child Safety Research Agenda.



Thind Transport Group is one-stop solutions for providing logistics, warehousing, refrigeration, for the last 15 years within Australia.



WAREHOUSE SERVICE



REFRIGERATED TRANSPORT



CONTRACT CARTAGE



ROAD TRANSPORT SERVICE



SPECIALISED FREIGHT



CONTAINER SERVICE

Arrange a quote by visiting our website at www.thindtransportgroup.com.au

Call us on 0468 350 034 Or send us an email at info@thindtransportgroup.com.au

Proudly Supporting Our Police Members Locally



JD Interstate Transport can assist in transporting all sizes of freight, whether it's just one pallet or a B-double load, our experienced team ensures your freight is delivered safely and on time, and we offer excellent service at a competitive price.

We are renowned throughout the East coast of Australia for our professionalism and commitment to moving large quantities of freight between Brisbane, Sydney, Melbourne and Adelaide.

Phone us on (02) 4423 4366

www.dunntransport.com.au

Proudly Supporting Our Police Members

On Remembrance Day, we honour those who've served our nation with courage and sacrifice. A special thanks to Queensland Police for your dedication to keeping our communities safe.

I'm grateful for the trust you've placed in me by re-electing me to represent Queensland in the Senate.

Thank you for your service—and for your support.

Senator Malcolm Roberts

Electorate Office:

One Eagle - Waterfront Brisbane,
1 Eagle Street, Brisbane QLD 4000
GPO Box 228, Brisbane QLD 4300
senator.roberts@aph.gov.au

Tel: (07) 3221 9099



Malcolm Roberts
Senator for Queensland



Authorised by Senator M Roberts 1 Eagle Street, Brisbane QLD 4000



Thank you to our Police for service to our community. We commemorate those who have lost their lives in the line of duty.



Anoulack Chanthivong MP

Member for Macquarie Fields

Minister for Better Regulation and Fair Trading
Minister for Industry and Trade
Minister for Innovation, Science and Technology
Minister for Building / Minister for Corrections

Address: Unit 9 Centennial House,
12 Ingleburn Road, Ingleburn NSW 2565

Phone: 9618 2077
Website: connectwithanoulack.com
Email: macquariefields@parliament.nsw.gov.au

Authorised by Anoulack Chanthivong MP, Member for Macquarie Fields.
Funded using Parliamentary entitlements.



RUOK?™

A conversation could change a life.



Blacktown City Council is committed to working towards a community based on equality and respect to ensure that every person has the right to live a safe and meaningful life, free from all forms of violence.

Council applauds the outstanding police officers at Blacktown, Mount Druitt, and Quakers Hill Local Area Commands.

We are proud to work in partnership with our local police to help make Blacktown City a safe place to live, work and enjoy.

Thank you...

To the Newcastle Local Area Command
- our partners in making Newcastle a safe and liveable city for all members of our community.

newcastle.nsw.gov.au



“On Police Remembrance Day, we remember all those officers who have lost their lives serving their communities.”

Justine Elliot MP
Federal Member for Richmond

Authorised J. Elliot, ALP, 107 Minjungbal Drive, Tweed Heads South



Remembering NSW Police service and sacrifice

**- Police Remembrance Day
29 September 2025**

Dave Layzell MP
Member for Upper Hunter



Authorised by Dave Layzell MP, 94 John Street Singleton NSW 2330 - Funded using Parliamentary entitlements.

'It's beyond human scale': AFP defends use of artificial intelligence to search seized phones and emails

Australian federal police says it has 'no choice' due to the vast amount of data examined in investigations.

JOSH TAYLOR

[theguardian.com](https://www.theguardian.com)

The Australian federal police says it had "no choice" but to lean into using artificial intelligence and is increasingly using the technology to search seized phones and other devices, given the vast amount of data examined in investigations.

The AFP's manager for technology strategy and data, Benjamin Lamont, said investigations conducted by the agency involve an average of 40 terabytes' worth of data. This includes material from the 58,000 referrals a year it receives at its child exploitation centre, while a cyber incident is being reported every six minutes.

"So we have no choice but to lean into AI," he told a Microsoft AI conference in Sydney on Wednesday.

"It's beyond human scale, so we need to start to lean in heavily on AI, and we're using it across a number of areas."

Aside from being part of the federal government's trial of Copilot AI assistant technology, the AFP is using Microsoft's tools to develop its own custom AI for use within the agency, including undertaking work translating 6m emails that were all in Spanish, and examining 7,000 hours of video footage.

"Having ... a human sitting there going through 7,000 hours – it's just not possible. So AI is playing a heavy role in that," Lamont said.

One dataset the AFP is now working on is 10 petabytes (10,240TB), and an

**HAVING ... A HUMAN SITTING THERE
GOING THROUGH 7,000 HOURS –
IT'S JUST NOT POSSIBLE. SO AI IS
PLAYING A HEAVY ROLE IN THAT.**

individual phone seized could involve 1TB of data. Lamont said much of the work the AFP was seeking to use AI for was structuring of obtained files to make them easier for officers to process.

"When we do a warrant at someone's house now, there's drawers full of old mobile phones," Lamont said. "Now, how do we know that those mobile phones haven't been used in the commission of an offence? We have to go through them and then identify those components and see if there was ... any criminality in there."

The AFP is also developing AI to detect deepfake images and has been seeking to figure out how to quarantine, clean and analyse data obtained during investigations, through operating in a secure, fully disconnected environment.

The agency is also exploring whether generative AI could be used to create text summaries of images or videos before they are viewed by officers, to prevent them being unexpectedly exposed to graphic imagery. The AFP is also looking

at whether AI could modify such content by converting images to greyscale or removing audio.

The AFP has faced criticism over its use of the technology, most notably when its officers used Clearview AI, a facial recognition service built off photos taken from the internet.

Lamont said the AFP "haven't always got it right".

"We've had to strengthen our processes internally and I think this ... has been really key, because it's not just a set and forget," he said. "As technology evolves and as the processes evolve ... we have to continually look at how we're making sure that it's ethical and responsible, and so we've created a responsible technology committee within the organisation to assess emerging technology."

He said it was also important for the AFP to discuss its use of AI publicly and ensure that there was always a human in the loop making the decisions formed from AI use.

Australian federal police using AI to analyse data obtained under surveillance warrants



The AFP's use of AI has been limited so far but the agency hopes the technology will help police identify money laundering and potential fraud.

JOSH TAYLOR

theguardian.com

The Australian federal police has said it uses AI to analyse data obtained under telecommunications and surveillance warrants, as the agency promises full transparency over the use of the technology.

In a submission to the federal government's discussion paper on the responsible use of AI, the AFP said its use of the technology had been limited so far, including using AI to translate foreign materials into English.

But it noted that AI tools – including large language models (LLMs) – gave the AFP an opportunity to find useful information in large, lawfully collected datasets.

"By speeding the discovery task, members can make decisions earlier and execute the necessary actions accordingly," the AFP said.

The AFP indicated it could also potentially help analyse transactional data to identify irregular patterns like money laundering and potential fraud.

In 2021, the AFP came under fire from the privacy commissioner over employees

using the controversial Clearview AI facial recognition technology, which built its dataset from photos of people taken from social media without their permission. The AFP has ceased using the technology, but secretly met with Clearview AI after it claimed to have stopped using the technology.

The AFP said in its submission that it would be transparent and "proactively undertake due diligence into technologies before deployment", taking into account ethical considerations and robust governance and oversight.

"Policing is deeply connected to society and must reflect the values, norms and expectations of the community it serves and critically requires human oversight and accountability."

A spokesperson for the AFP confirmed that sensitive information obtained from warrants would be fed into LLMs or neural networks. But the agency said it ensures the data is protected, whether it is an in-house tool or when using a commercial product, so it would not feed into public datasets.

The lawfully collected data used could include data collected under a warrant, including telecommunications interception data and surveillance data.

The spokesperson also said all language translations are checked by a human.

"We can design regulation providing a platform for innovation while protecting Australians"

The inquiry, established by the Albanese government earlier this year, received 510 submissions from a wide variety of people and organisations. Submissions came from groups at the forefront of the technology, including Meta, Google, Amazon Web Services, Open AI, and Microsoft, and industries likely to be affected, including legal firms, healthcare organisations, business groups, banks, supermarkets, and film, music and television companies.

Many of the submissions raised concerns over AI, particularly large language models, being trained on their content without permission or payment.

ROY BUTLER^{MP}
INDEPENDENT MEMBER FOR BARWON

Thank you to all the
 Police Officers and their families
 for the sacrifices you make.
 You dedicate your lives to
 ensuring we live safely and
 for that I say thank you.

BROKEN HILL
 1/142 ARGENT STREET
 PH (08) 8087 3315

COBAR
 11 BARTON STREET
 PH (02) 6836 3722

NARRABRI
 60 MAITLAND STREET
 PH (02) 6792 1422

www.roybutler.com.au
 barwon@parliament.nsw.gov.au

Authorised by Roy Butler MP 60 Maitland Street Narrabri NSW 2390. Paid for using Parliamentary Entitlements July 2025

Arab Bank
 Australia
 Limited

Thank you.
 For your ongoing dedication
 to the community.

We Are
abal banking

arabbank.com.au @abalbanking

Arab Bank Australia Limited ABN 37 002 950 745 AFSL 234563.

*We remember and honour
 those who gave
 their lives in the
 line of duty.
 Thank you.*

**DUGALD
 SAUNDERS MP**

Member for the Dubbo electorate

Authorised by Dugald Saunders MP, 1/118 Tallmudge St Dubbo NSW 2830. Funded using Parliamentary entitlements.

**Remembering and
 honouring those who
 put their lives on the
 line to protect us.**

Edmond Atalla MP
 MEMBER FOR MOUNT DRUITT
 Parliamentary Secretary for Police and Counter-terrorism

Phone: (02) 9625 6770 **Email:** mountdruitt@parliament.nsw.gov.au
Office: Suite 201, Westfield Shoppingtown, Carlisle Avenue, Mt Druitt

Authorised by Edmond Atalla MP, Suite 201 Westfield Shoppingtown, Mount Druitt funded using parliamentary entitlements.

The Australian Recording Industry Association (ARIA) said AI that creates deepfakes or vocal clones without authorisation should be severely restricted.

"Such use of AI technology robs artists of control over their own voices and image, and can confuse and mislead fans who may be unaware they are not listening to genuine music created by their favourite artist. This can have a detrimental impact on an artist's career."

Getty Images said that AI developers needed to be transparent about the datasets their technology is built on to ensure that intellectual property and privacy rights are not being violated.

"One way to mandate transparency requirements is to require both private and public sector organisations to keep auditable records of all training datasets used including how the data was sourced," Getty Images said in its submission.

Free TV Australia said that content owners should be paid for the use of their

content, while also having the option to refuse AI tools access to their content.

Guardian Australia previously reported Google's submission to the inquiry called for flexible copyright laws to allow AI to be able to mine content from websites unless those websites opt out.

Since then, a number of news publications including the Guardian, have blocked ChatGPT owner OpenAI from mining their sites.

In a speech on Thursday, industry minister, Ed Husic, said most tech industry submissions called for using existing laws, while others had noted gaps in the legislation. He said the government wanted to get it right and would consider the responses over the next few months.

"I also know that we can design regulation providing a platform for innovation while protecting Australians – our communities and our national wellbeing," he said.

"The root of this debate isn't, should we regulate AI? It is, in what circumstances

should we expect people and organisations developing and using AI to have appropriate safeguards in place?"

The Law Council of Australia said the federal government should consider in the short-term regulating high-risk AI technology including biometric tech such as facial recognition and social scoring, as well as protecting from AI-generated fakes and scams.

The Shopping Centre Council of Australia raised concerns that its existing technology such as CCTV and "facial detection" screens might get caught up in any regulation of AI, while banks and financial institutions such as Visa, CBA and NAB said they had long deployed AI to detect potential fraud and argued that much of what they do is already covered by existing law.

The Business Council of Australia said the problem wasn't that there were no laws covering AI, just that there needed to be a better understanding of how existing laws apply before any new laws are made.

AUSTRALIAN STRATEGIC POLICY INSTITUTE

aspi.org.au

Artificial intelligence and policing: it's a matter of trust

From *Robocop* to *Minority Report*, the intersection between policing and artificial intelligence has long captured attention in the realm of high-concept science fiction. However, only over the past decade or so has academic research and government policy begun to focus on it.

Teagan Westendorf's ASPI report, *Artificial intelligence and policing in Australia*, is one recent example. Westendorf argues that Australian government policy and regulatory frameworks don't sufficiently capture the current limitations of AI technology, and that these limitations may 'compromise [the] principles of ethical, safe and explainable AI' in the context of policing.

My aim in this article is to expand on Westendorf's analysis of the potential challenges in policing's use of AI and offer some solutions.

Westendorf focuses primarily on a particular kind of policing use of AI, namely, statistical inferencing used to make (or inform) decisions—in other words, technology that falls broadly into the category of 'predictive policing'.

While predictive policing applications pose the thorniest ethical and legal questions and therefore warrant serious consideration, it's important to also highlight other applications of AI in policing. For example, AI can assist investigations by expediting the transcription of interviews and analysis of CCTV footage. Image-recognition algorithms can also help detect and process child-exploitation material, helping to limit human exposure. Drawing attention to these applications can help prevent the conversation from becoming too focused on a small but controversial set of uses. Such a focus could risk poisoning the well for the application of AI technology to the sometimes dull and difficult (but equally important) areas of day-to-day police work.

That said, Westendorf's main concerns are well reasoned and worth discussing. They can be summarised as being the problem of bias and the problem of transparency (and its corollary, explainability). Like all humans, police officers can have both conscious and unconscious biases that may influence decision-making and policing outcomes. Predictive policing algorithms often need to be trained on datasets capturing those outcomes. Yet, if algorithms are trained on historical datasets that include the results of biased decision-making, it can result in unintentional replication (and in some cases amplification) of the original biases. Efforts to ensure systems are free of bias can also be hampered by 'tech-washing', where AI outputs are portrayed (and perceived) as based solely on science and mathematics and therefore inherently free of bias.



Steve GEORGANAS MP
Federal Member for Adelaide

Proudly Supporting
Police

*Your Community
Voice*

161 Main North Road
NAILSWORTH SA 5083
Phone: 08 8269 2433
Email: Steve.Georganas.Office@aph.gov.au




LOGO DESIGN



BRANDING



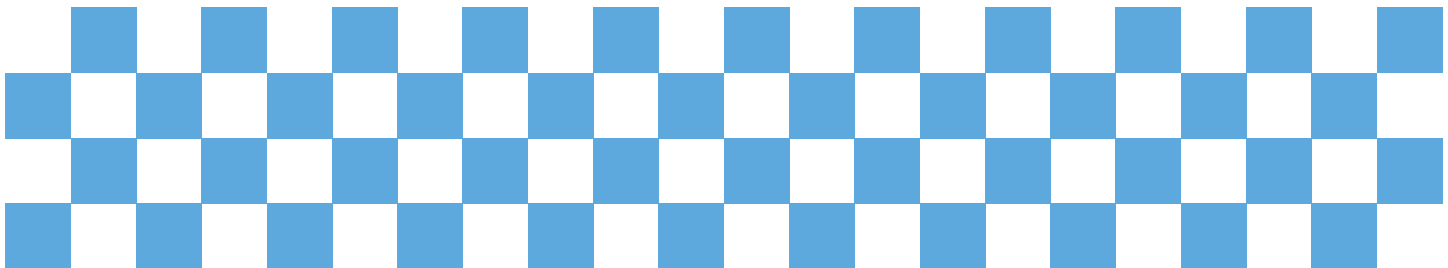
SOCIAL MEDIA



WEB DESIGN



info@officevanilla.com * @officevanilla

Related to these concerns is the problem of transparency and explainability. Some AI systems lack transparency because their algorithms are closed-source proprietary software. But it can be difficult to render even open-source algorithms explainable—particularly those used in machine learning—due to their complexity. After all, a key benefit of AI lies in its ability to analyse large datasets and detect relationships that are too subtle for the human mind to identify. Making models more comprehensible by simplifying them may require trade-offs in sensitivity, and therefore also in accuracy. Together these concerns are often referred to as the ‘AI black box’ (inputs and outputs are known, but not what goes on in the middle).

In short, a lack of transparency and explainability makes the detection of bias and discriminatory outputs more difficult. This is both an ethical concern and a legal one when justice systems require that charging decisions be understood by all parties to avoid discriminatory practices. Indeed, research suggests that when individuals trust the process

of decision-making, they are more likely to trust the outcomes in justice settings, even if those outcomes are unfavourable. Explainability and transparency can therefore be important considerations when seeking to enhance public accountability and trust in these systems.

As Westendorf points out, steps can be taken to mitigate bias, such as pre-emptively coding against foreseeable biases and involving human analysts in the processes of building and leveraging AI systems. With these sorts of safeguards in place (as well as deployment reviews and evaluations), use of AI may have the upshot of establishing built-in objectivity for policing decisions by reducing reliance on heuristics and other subjective decision-making practices. Over time, AI use may assist in debiasing policing outcomes.

While there’s no silver bullet for enhancing explainability, there are plenty of suggestions, particularly when it comes to developing AI solutions to enhance AI explainability. Transparency challenges generated by proprietary systems can also be alleviated when AI systems are owned by police and designed in house.

Yet the need for explainability is only one consideration for enhancing accountability and public trust in the use of AI systems by police, particularly when it comes to predictive policing. Recent research has found that people’s level of trust in the police (which is relatively high in Australia) correlates with their level of acceptance of changes in the tools and technology used by police. In another study, participants exposed to purportedly successful policing applications of AI technology were more likely to support wider police use of such technologies than those exposed to unsuccessful uses, or not exposed to examples of AI application at all. In fact, participants exposed to purportedly successful applications even judged the decision-making process involved to be trustworthy.

This suggests that focusing on broader public trust in policing will be vital in sustaining public trust and confidence in the use of AI in policing, regardless of the degree of algorithmic transparency and explainability. The goal of transparent and explainable AI shouldn’t neglect this broader context.

AI and policing: what a Queensland case study tells us

Policing agencies consider artificial intelligence a force multiplier because it can rapidly process more data than human brains and yield insights to help solve complex analytical problems.

Our limited understanding of how AI algorithms make decisions and produce their insights, however, presents a significant challenge to ethically and safely implementing AI policing solutions. The use of AI by Queensland police provides a valuable opportunity to study how we can mitigate the possible negative, ethical and operational effects of this problem.

The state is trialling AI as a risk-assessment tool to predict and prevent crime, in this case domestic violence. It screens data from police records to identify 'high risk of high harm' repeat offenders. Police then routinely pre-emptively 'knock on doors' to deter escalation to violence and, theoretically, lessen the likelihood of perpetrators reoffending.

Police say perpetrators have proved more likely to recognise their intervention outside a 'point of crisis' (domestic violence incident) and they believe this provides a 'turning point opportunity' for habitual offenders to deviate from a trajectory of repeated offending.

However, door-knocking for deterrence can have serious negative impacts, including the possibility of triggering further conflict within families experiencing repeated violence. This possible

antagonising effect on offenders in a 'precrime' strategy would have to be mitigated for this process to be ethical according to the federal framework, and for it to be effective in reducing domestic violence.

I raised this issue with the Queensland Police Service. They said the trial had demonstrably not driven further violence, and cited a 56% reduction in incidents in one cohort of high-risk, high-harm offenders with a possible victim cohort of 1,156 people.

These statistics are compelling evidence that the program could reduce offending by such perpetrators. It could also reduce deaths. The police say that 30% of domestic violence homicides in the state are carried out by offenders already known to them for domestic violence, and known offenders are significantly more likely to suicide.

The Queensland Police Service's aim is to prevent domestic violence, disrupt recidivist behaviour and 'arrest no one'.

How did the police deal with the limitations and potential pitfalls of AI?

First, the police own the AI and developed it in house, substantially increasing its transparency.

The barrier of commercial interests that prevent a company from sharing

the details of product development was removed, as in the case of AI used to make parole decisions by some US courts. Data scientists were employed to work closely with officers at all stages of the AI's development and deployment.

Owning the supply chain gives police as comprehensive an understanding as possible, given the technical limitations, of the processes by which the AI is developed, trained and then deployed and monitored by in-house data scientists. This includes understanding what human biases may have been coded into the AI, what mitigation strategies have been used, what AI biases might develop through its operation on live datasets, and what should be guarded against via monitoring once the AI is deployed into live datasets.

Police ownership also seems to have provided an opportunity for authentic policing knowledge and judgement to be included at the design stage, rather than as a retrofit after the proprietary development of the product.

Critically, police could ensure the AI was trained on state police data. That meant that, while it's not possible to avoid coding human bias into AI, they could be certain that the bias was from their own historical data and therefore



**Family-Built.
Ready to Move.
Made to Last.**

GM Portable Homes is Queensland and New South Wales' family run manufacturer of transportable buildings, offering modular, demountable, relocatable homes, mining accommodation, park cabins, granny flats, family accommodation, and commercial buildings.

0414 996 634 | gmporthablehomes@gmail.com
gmporthablehomes.com.au

Proudly Supporting Our Police In The Community



AVANT

www.avant.com.au ph: (02) 9675 4400




known and understood. Those training datasets serve as a historical resource from which police data scientists can glean information on the historical human bias of the police force and try to code safeguards against it into the AI.

Knowledge of AI decisions is similarly increased by the AI being owned and developed in house, because the data scientists using and monitoring it and the police officers employing have all the same information about it as those who developed and trained it.

To be eligible for assessment by AI, subjects must be already considered high risk and high harm through their previous interactions with police and have at least three domestic violence orders against them. This helps police know which of all the homes experiencing repeated family violence they should door-knock to deter further violence. Police don't have the resources to door-knock all homes with a record of domestic violence.

It's likely that using AI for high-risk criminal justice decisions will never be a good idea if we're striving for safe, fair, ethical and reliable AI use. But it can provide valuable insights and context to inform policing decisions.

Identifying at-risk victims is not the focus, despite the overall goal being to

prevent or reduce the rate and severity of violence endured by victims of habitual offenders.

So, the potential harm of over-policing subjects is arguably neutralised by the fact that police have already been interacting with them due to their repeated offending. This is not a risk-assessment of a general cross-section of a community, or even a cohort with a single record of violence.

But if AI decision-making were used in a higher-risk policing scenario, who would be accountable for incorrect decisions: police, computer scientists, policymakers or even the AI?

The eligibility criteria provide a key safeguard for limitations on transparency and explainability because if someone is incorrectly flagged for a door-knock, they are still a known, repeat perpetrator. If we accept the ethical and practical legitimacy (in terms of likelihood to achieve outcomes of harm reduction) of police pre-emptively door-knocking known offenders at all, then it can't be argued that incorrectly door-knocking someone at a slightly lower but still significant risk of triggering violence counts as over-policing or violating their right to privacy and equality. A net benefit logic applies.

Problems remain, though. Significant technological development is required to design comprehensively transparent and explainable AI.

Computer scientists tell us that it remains incredibly difficult, if it proves possible at all, to comprehensively understand how AI make decisions within live datasets as they develop more and more correlations that aren't visible to monitors, as in overfitting. We need to keep trying, and to hold AI to equal, or higher, ethical standards than human decision-making.

As for the net benefit argument, using AI as a solution could obfuscate both the root causes of a problem and the possible alternative, non-technical solutions. Can we, for example, better support victims in the family court to prevent them living in a perpetually violent home?

AI solutions are here to stay. Appropriate regulation in law enforcement scenarios is imperative to mitigate their significant potential impacts on justice outcomes and civil liberties. If Australia wants to ensure AI is safe, secure and reliable, we need an ethical framework that is compulsory and legally enforceable, not voluntary and aspirational.

Effective, Explainable and Ethical: AI for Law Enforcement and Community Safety



CAMPBELL WILSON

AiLECS Lab, Monash University

JANIS DALINS

AiLECS Lab, Australian Federal Police

GREGORY ROLAN

AiLECS Lab, Monash University

Abstract

We describe the Artificial Intelligence for Law Enforcement and Community Safety (AiLECS) research laboratory, a collaboration between the Australian Federal Police and Monash University. The laboratory was initially motivated by work towards countering online child exploitation material. It now offers a platform for further research and development in AI that will benefit policing and mitigating threats to community wellbeing more broadly. We outline the work the laboratory has undertaken, results to date, and discuss our agenda for scaling up its work into the future.

INTRODUCTION

Criminal activity is increasingly facilitated by technology; often characterised by the generation, distribution and/or monetisation of illegal material via computer networks. In particular, recent years have seen rampant growth in the production and online dissemination of harmful and offensive materials, such as child exploitation material (CEM) and violent media associated with online radicalisation. Furthermore, worrying trends are emerging in the algorithmic generation of realistic abusive material such as “deepfake” imagery. Many of these cybercrimes are organized and transnational. In the course of investigating and prosecuting such offenses, law enforcement agencies deal with significant challenges, both professional and personal. Analysis and classification of such material exposes police and judicial officers to significant psychological harm. This is exacerbated by the increasingly large volumes of data involved in such investigations.

Given their capacity to learn patterns from large datasets and make consequent predictions,

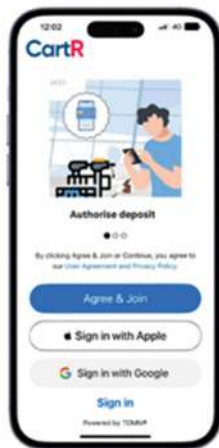
Artificial intelligence (AI)¹ technologies offer clear advantages for law enforcement in countering these threats while greatly reducing investigator exposure to harm. However, such technologies need to be developed and operationalised in accordance with appropriate frameworks for their legal, ethical, and explainable use, particularly given that preservation of community trust is vital for effective policing.

In this paper we describe a new research initiative: the AI for Law Enforcement and Community Safety (AiLECS) Laboratory [1], a collaboration between the Australian Federal Police (AFP) and Monash University. We envision the lab and its operations as a model and platform for AI research related to law enforcement on an international scale. Such a global orientation is necessary given the cross-jurisdictional nature of the problem domain and complexity of the research issues.

This is a high-stakes endeavour, with potential impact across several of the United Nations Sustainable Development Goals outlined in




Rent a Luggage Cart in Seconds -
Straight from Your Phone with CartR.



**CartR: Smarter Cart Rentals for Modern
Travellers**

Contact us at
support@cartr.com.au
www.cartr.com.au

Find us on
 



BURU REHAB

Improving Aboriginal social outcomes through mining rehabilitation



BURU REHAB PTY LTD

Mob: 0448 893 266
www.bururehab.com.au

We bring together the science, engineering,
expert earthworks operators and traditional
owner land management principles to deliver
world class mine rehabilitation outcomes for our
clients and for the traditional custodians of the land.

**We are not just an earthworks
company, we are a specialised mine
rehabilitation business.**

Proudly Supporting Our Police

Resolution A/70/L.1 [1] and adopted by the General Assembly in 2015. All member nations have committed to meeting these goals and their associated targets by 2030. We aim to contribute specifically to Goal 3 Good Health and Well-Being by reducing the incidence and impact of harmful material online, Goal 16 Peace, Justice and Strong Institutions by integrating academia, industry and law enforcement to enhance technological capacity building, and Goal 17 Partnerships for the Goals through establishing sustainable collaboration between researchers and practitioners in information sciences, policing, law, and ethics.

The paper is organised as follows. In section 2, we describe our initial drivers for research in this area, challenges, and related initiatives. In section 3 we discuss the infrastructure underlying the lab in the form of our innovative data airlock platform. Section 4 covers work done to date in AI algorithm development. In section 5, we discuss learnings from the partnership and outline our future plans. We conclude in section 6 with a call for greater collaboration in this domain.

MOTIVATIONS AND CHALLENGES

In July 2019, the AiLECS laboratory was officially launched by the AFP and Monash University in Melbourne, Australia. This collaborative initiative represents significant investment by both organizations in building expertise around AI and related technologies for enhancing investigative capabilities. Importantly, it also seeks to move towards limiting exposure of AFP members to traumatizing material. The primary initial motivation behind AiLECS was addressing online child exploitation, building on prior research work the lab investigators.

Motivator: The Scourge of Online Child Exploitation

The disturbing proliferation of online CEM is but one example of the need for urgent investigation of how AI approaches may assist investigators. In 2018 alone, the Australian Center for Countering Child Exploitation received over 18,000 individual reports of child exploitation, with each one potentially containing hundreds or thousands of abhorrent images and videos. As stated in the introduction, this is a crime that knows no national boundaries and many countries

face similar onslaughts. A recent study pointed towards the "exponential" growth of this material, stating that of the 23.4 million reports of CEM received since 1998 by the US National Center for Missing and Exploited Children (NCMEC), 9.6 million were received in 2017 alone [2].

Combating CEM places an increasingly traumatic drain on the human and digital forensic resources of law enforcement agencies. Development of AI tools to assist in this task is a natural focus of much research and development. When integrated with existing digital investigation workflows, AI techniques do promise much by reducing investigator workload, accelerating investigations, and improving the welfare of those tasked with viewing and labelling such material.

Despite the AI field being long-standing, there have been particularly strong advances in recent years. For example, large strides have been made in improving artificial neural-network based image classification algorithms, thus making semi-automatic triage and categorization of large

amounts of seized and previously unseen CEM images and videos more feasible [2], [3]. AI-based natural language processing techniques are also being investigated for assisting with classifying relevant evidentiary textual material [4].

AI in Law Enforcement: Challenges

In addition to countering CEM, AI techniques can be deployed to enhance many types of law enforcement activities. They are in a sense an evolution of data analysis capabilities that agencies already possess, and are suited to the increasingly large amounts of data that characterize many contemporary criminal investigations. However, aside from the challenge of research and development into improving the techniques themselves, their use in law enforcement poses a number of challenges:

- **Ethical Application:** Over the past few years there has been a growing discourse concerning the ethics of AI. Particularly in the law enforcement context, we need a clear and practical ethics basis upon which any AI intervention should proceed. Ethics frameworks (typically comprising statements of broad ethics principals) have been proposed by governments, private corporations, professional bodies, standards organisations, academia, and other individual and consortia stakeholders in AI [5]. Understandably, there is an emphasis in this discourse, on justice, fairness, and so on; and some of these can be be addressed through technological measures such as better data wrangling or research into algorithmic bias detection or fairness balancing.

AI systems based on machine learning algorithms are highly dependent on (possibly voluminous) training data. Human *biases* in the training data are easily reflected in the output of such systems. Training data must be collected in a way that ensures the data is representative, inclusive and accurate. In the law enforcement context, this is particularly important given the potential adverse effects of biased algorithms and consequent loss of community trust in police agencies.



From a broader perspective, our ethics stance needs to take into account the broader sociological implications of law enforcement work [6]. What are the purposes of law enforcement, and for AI interventions that seek to augment its activities? Along with justice and fairness, how is societal cohesion and interdependence best served? For example, the collection of large amounts of data in order to train algorithms in this context must also be mindful of individual rights to privacy. While lawful police activities may be the subject of certain exemptions under relevant legislation governing use of personal data, perception of undue surveillance is again a potential source of erosion of trust in agencies. Further, any degree of automation in law enforcement decision making potentially erodes human agency. How do we automate law enforcement work while retaining interpersonal discretion and *accountability* to those affected by the decision making?

- **Explainability:** Transparent decision making is particularly important in law enforcement. In an AI context, this means that explaining the output of traditionally opaque algorithms (e.g. deep neural networks) is a highly relevant area of research. Explainability of the algorithm itself is however not the entire picture—defending an AI enhanced policing approach in a court of law may require clear explanations of the provenance and collection methods of training data, how this data was curated and labelled prior to model training, how results were interpreted, and how predictions made by the model were applied in the investigation.
- **Data Provision and Access:** The use of evidence seized as part of real-world law-enforcement investigations as training data for machine learning algorithms must be carefully considered from legal, ethical and technical perspectives. Relevant legislation in individual jurisdictions will govern how this data can be processed.

From a purely technical perspective, more training data is better. However the transnational nature of technology facilitated crime poses challenges for the interchange of potentially sensitive data between countries. This is a challenge borne not just of legal restrictions on the export of such data, but also of security and logistical data management concerns.

Academic researchers outside of law enforcement agencies do not typically have legal access to data held by police. This ostensibly restricts their capability to develop, train, test, and compare machine learning models on such data. While partnerships between academia and law enforcement such as AiLECS do provide a platform for research collaboration, provision and processing of the training data itself remains fraught from a legal perspective. Additionally, in the case of highly distressing content such as CEM, it is important to be mindful of the psychological harm that may be inflicted on all of those dealing with the data or even with the concepts the data implies.

- **Data Labelling** The labelling of data for training supervised machine learning models in law enforcement poses some particular challenges. In addition to the resourcing challenges of labelling very large volumes of data, potentially across jurisdictions, there are considerations around effecting algorithmic bias through subjective labelling (alluded to earlier). Further, it may be the case that investigators are required to deal with evidential classification schemes that may not be as amenable to machine learning training (e.g. less objective, less distinct classes [3])

The AiLECS lab was created address these and other challenges of AI in law enforcement in mind, and to provide a scalable platform for research and development in this area. Efforts targeting algorithm development, sensitive data management, and national and international collaboration have been core to this mission, in addition to development of goals around ethical use.

Building on its initial motivating remit of countering CEM, the AiLECS lab has prioritised a number of areas of focus, in particular:

- Illicit image and video machine learning classification
- Image localisation (estimating image geolocation by content)
- Scalable near-duplicate image detection
- Natural language processing, in particular of short-text documents
- Explainable algorithms and auditable performance techniques
- The ethics of AI in law enforcement

A key objective of AiLECS is to open source as much of our work as is possible in order to rapidly disseminate research and bolster development in the area.

Related Initiatives

Police agencies around the world have investigated, and to varying degrees implemented, AI related technologies. This includes among other areas of application: facial recognition, optimised resource allocation, crime prediction, traffic policing and text/social media analyses.

Much of this has been in conjunction with commercial vendors, while direct collaborations between law enforcement and universities is often ad-hoc. Nevertheless, a number of examples of research collaboration between law enforcement and academia exist. Typically the applications of these initiatives are closely related to the local needs of the relevant jurisdiction. For example:

- The University of Cambridge and Durham Constabulary in the UK have developed a random-forest based reoffending risk model [7]
- Also in the UK, the Turing Institute has worked with the West Midlands Police to study the ethics of data analytics at scale for policing [8]
- The Netherlands Police Lab AI [9] is a collaboration between the Dutch Police, Utrecht University and the University of Amsterdam, and is a close (independently developed) analogue of AiLECS. The lab actively researches both the technical and social/ethical aspects of AI in law enforcement.

OPERATIONALISING AiLECS - Data Airlock Platform

In order to deal with the challenge of managing sensitive, distressing and legally restricted data, an innovative *Data Airlock* platform [3] is under

continued development within AiLECS. This infrastructure is intended to provide controlled and configurable access to large law enforcement datasets for researchers. Such a platform is necessary in order to scale up research in this area, particularly since international collaboration will be vital to further address the large scale technical challenges inherent in combating criminal network activity.

Specifically, the current data airlock platform comprises three zones within its underlying hardware, namely *public*, *sensitive* and *secure*. There is no reliance on any particular underlying hardware or cloud environment, with the platform being highly containerised. The three zones operate under different security and access models.

The raw sensitive data resides in the *secure* zone and runs models on the data in isolated virtual environments "airlocks". Layers of encryption and exfiltration controls are in place and configurable, and uploading of data to the secure zone is only possible with physical hardware access, with this zone isolated from the broader network.

The *sensitive* zone provides an environment for models submitted to the platform by researchers to be vetted by authorised personnel before they are allowed to execute on the sensitive data in their own airlock in the secure zone. This vetting can be either through manual review or via semi-automatic or automatic tools depending on the application.

The *public* zone is a web-enabled environment through which researchers can submit models and view the results of their execution, and any other allowable meta-data as determined by authorised personnel.

The current data airlock infrastructure, given the focus on deep-learning based illicit image classification is based on a combination of secure servers with an NVidia DGX-1 platform providing execution of models within the *secure* zone.

OPERATIONALISING AiLECS - Algorithm Development

The AiLECS lab represents a culmination and platform for extension of research into algorithmic approaches to law enforcement already undertaken, some of which is being operationalised by the AFP [10]. We describe below some of this work.

Monte-Carlo Filesystem Search

It is often the case that rapid triage of data on seized devices is required in the course of a police investigation. There may be a range of time constraints imposed - by law and/or operational resourcing. Additionally, the search may be undertaken in highly bandwidth constrained environments. Thus, techniques that can decrease the time taken to find potential criminal evidence through prioritised search algorithms are of particular value.

In [11] we describe the *Monte Carlo Filesystem Search* (MCFS), inspired by the success of algorithms that apply *Monte Carlo* methods to searching game trees. The algorithm was designed to be lightweight for in-the-field investigative environments and to explicitly incorporate investigative domain knowledge. As with game tree searches, the algorithm leverages file system tree exploration (new branch selection) and exploitation (expansion of visited branch) with the configurable incorporation of bias towards more probable locations of interest. Lightweight machine learning algorithms (e.g. Multinomial Naive Bayes), were used to train the scoring algorithms for filesystem node selection depending on the type of application.

MCFS was evaluated in a realistic investigative setting by training on real case data from police investigations, with speedups of around one third seen in typical file system search scenarios looking for known items of interest compared with uninformed file searching [11]. The extension to this work for web-search and image based similarity search is ongoing.

Stonefish Classifier

In [3], we introduce a classifier, "Stonefish" based on deep neural networks to test the capacity of such architectures to reliably identify CEM. Although other such classifiers have been proposed and tested for this task, see for example [12] and [13], a typical bottleneck is access to real world case data. In both of those works, such data was sourced from the Brazilian Federal Police under controlled conditions. In [13] this involved a sandbox approach whereby only feature vectors were exposed to researchers. This is similar in spirit to our data airlock approach, however we provide a more generalised platform to enable research scalability

and obviate the need for adhoc virtual environment construction. The Stonefish classifier architecture consists of three levels, namely:

- *Pornography detection*: The first level aims to identify pornographic content with high confidence. This is a mature area of research, so we adapted the pre-trained OpenNSFW [14] model for our environment and deployed it on our test corpus of data sourced from real-world contemporary AFP online child exploitation investigations (with appropriate legal and welfare controls in place [3]). This allowed us to partition the dataset to focus on those images most likely of interest to law enforcement and assess them for possible CEM.
- *Child detection*: We trained a deep neural network model to perform a binary classification task - to label images previously assessed as pornographic as either containing children or not. This is in itself a very difficult task. The estimation of age in images (generally containing faces) is a very active area of research (for a survey of approaches see [15]). In our case, we did not seek to estimate age beyond the very broad categorical assessment of child vs adult. In this layer, our approach utilised a deep convolutional neural network (CNN) pretrained on the ImageNet 1000 class, with the top stack of fully connected layers removed and retrained on images previously labelled according to whether or not they depicted children. Again, training data was provided by the AFP under carefully controlled conditions. While datasets containing illegal and distressing material must not be distributed, it is useful to note also at this point that there is a lack of available standard training sets of completely legal images of children. This is of course understandable given online safety and privacy and ethical concerns.
- *Standardised classification*: The third layer of the Stonefish classifier is tasked with assigning suspected CEM images to the appropriate standardised schema; in our case the CETS scheme, a *de facto* standard in use in Australia and various other jurisdictions [3].

This schema is used as the basis for determining the severity of offending in that it provides labels indicating the nature of the activity depicted. In this layer, a deep CNN architecture is also used to perform this multi-class classification, with training data sourced from real-world AFP investigations. We also tested the model on images downloaded during a random traversal of the Tor network [16]

Results of the Stonefish classifier showed that with this early iteration of the architecture, we could achieve overall accuracies of around 60-70% in identifying real world CEM. This indicates that significant further work is required and is ongoing. However results were certainly adequate for initial triage prior to more in depth examination, thus going some way towards reducing investigator burden.

Jurisdictionally Independent CEM Training Schemata

What became particularly striking during our initial work on classifying CEM was the inadequacy of schemas such as CETS for training machine learning algorithms. There is a lack of standardised terminology in legislation around child exploitation, and schemas such as CETS tend to be overly broad and vague. This is understandable from the point of view of the initial motivation of the schemas - which was to inform categories of sentencing. However this abstract nature is not suitable for machine learning training, where it is valuable to have as little ambiguity as possible (particularly so when large quantities of data and many human labellers are involved). We therefore proposed, also in [3], the *Majura* schema, which provides a manifestly objective and detailed set of categories for labelling CEM that not only covers broad types of activities but also other features such as participants, props, subtypes of activities etc. It is our hope that this schema, independent of jurisdictional constraints, will be able to, in addition to the data airlock platform, rapidly accelerate collaborative international development of AI techniques in this area.

DISCUSSION AND LEARNINGS

The AiLECS lab is scaling up its operations, building on the existing infrastructure and work we have already

described. We are broadening the nature of our research and actively seeking collaborators. In doing so, we reflect on our experience so far and on the way forward.

The University/Law Enforcement AI partnership

Universities have for many years been working closely with industry on applied research. A variety of models exist for these collaborations, depending to a large extent on the government funding models in place in various countries in addition to the overall strategic vision of the partners. The integration of academic research at scale with law enforcement agencies is relatively new, and in the specific area of AI research, provides great opportunity, albeit subject to a number of important considerations.

Law enforcement agencies are by their nature highly operational and agile, needing to respond rapidly to changing environments to counter threats to community safety. Resource allocation is crucial. It is important that performance indicators relating to collaborative research with universities are formulated such that results of research are as directly and easily applied in practice as possible. For example, modern policing relies heavily on ICT infrastructure tailored to operational requirements, including forensics and intelligence capabilities. From a technical point of view, AI systems developed through collaborative research should be constructed within a mindset of integration with existing systems and workflows.

However, this is not to say there is no place for fundamental longer term research under such arrangements. Indeed this is very much a value-add that is provided through such collaboration given the mission of university research. This forms part of the answer to the question as to why a law enforcement agency would want to partner with a university. In our view, this is further answered through considering the following:

- **Interdisciplinarity:** Collaborations such as AiLECS leverage the natural interdisciplinary capabilities of Universities. In the domain of AI, and in particular when applying AI in the high-stakes environment of policing, computer science is not the only discipline that will inform



research. Large research oriented academic institutions bring new perspectives by virtue of their broad topical remit and base of expertise. Naturally, from a technological point of view, computer science plays a crucial role. However, it is clear that to advance AI in law enforcement, research from the social sciences can provide valuable legal, ethical and criminological perspectives on the use of the technology. Further, research in bioinformatics, pharmacology, chemistry, ballistics, and a number of other fields directly relevant to law enforcement operations will likely see tighter integration with data driven AI techniques.

- **Capacity building:** Tighter collaboration between law enforcement and universities through research leads to a cross-pollination of expertise. Not only does this assist agencies to adapt to technological change, but promotes broad understanding

of the issues faced by police in the research and higher education sector. This in turn strengthens the community partnership on which law enforcement is best-based.

- **Non-commercial imperatives:** Commercial ICT vendors have been vital partners of law enforcement agencies, and have enhanced police work with a number of tools and systems. However, a mixture of commercial and non-commercial (e.g. government/university) partnerships may avoid risk of vendor lock-in and over-reliance on commercially driven products. This is potentially a particular concern with technologies such as AI which are currently hype-driven in some industry sectors.
- **Research culture:** Universities are typically research oriented and have highly developed infrastructure around the management of data, research student training and supervision and appropriate ethical oversight.

The Future: Effective, Explainable and Ethical AI

This is a pivotal time for the application of AI, with research efforts and increases in computational resources driving the technology forward apace. In order to harness AI for application in law enforcement, we believe platforms such as AiLECS will provide a crucial clearing house for research in the area. This can only take place however if the three tenets of *effective* (meeting mission objectives and enhancing capability), *explainable* (transparent and accountable) and *ethical* (respecting the rights, privacy and agency of humans) AI are adhered to.

- *Effective* Our AI development is underpinned by the further development of the data airlock infrastructure for the safe access to real data. We intend to open the data airlock for use by researchers anywhere in the world who wish to test their models against real world data they would not otherwise be able to access. We hope this will also build international collaboration that will further enhance the combating of large scale transnational criminal activity.
- *Explainable* It is important that the entire pipeline of AI application in law enforcement, from methods of data collection and curation, labelling, storage, cleaning, training through to model construction, operation and prediction is as transparent as possible. It is not unforeseeable that the future will see an increase in AI algorithms requiring defence in courts, should they be used in decision support. We are thus working on building frameworks against such transparency, in addition to investigating how explainable AI (XAI) techniques can be applied and improved.
- *Ethical* AI for law enforcement is a focus for AiLECS. We alluded earlier to this issue as a challenge for AI in law enforcement and, particularly, the danger in ethics frameworks as checklists of overly broad and potentially banal statements such as "do no harm", or "be fair". The more pressing challenge is the mapping of ethics against operational requirements of law enforcement agencies in real



**AI FOR LAW
ENFORCEMENT
MUST BE EFFECTIVE,
EXPLAINABLE
AND ETHICAL.**

environments and building this into tools, data pipelines, and juridical processes. We believe that, in the context of AI, building ethical understandings and implementations that are useful in practice is best achieved with law enforcement practitioners working closely with ethicists and researchers, in the context of actual case studies—a trajectory AiLECS will follow in further re-search in this domain.

CONCLUSION

We have introduced the AiLECS research laboratory and outlined our work and vision for collaborative research between academia and law enforcement. Squarely in the domain of "AI for good", the work of AiLECS will accelerate progress towards achieving the UN Sustainable Development Goals pertaining to good health and

wellbeing, peace, justice and strong institutions as well as those around partnerships for achieving the goals. Our data airlock infrastructure and open source goals are designed to rapidly operationalise research efforts and build international teams to assist in researching and developing ethical AI for community safety applications.

To this end, we actively seek collaboration on a number of fronts. Firstly, with AI researchers and law enforcement agencies, to enhance and scale-up these initiatives as we have described. Secondly, on a cross-disciplinary basis, seeking partnership with domain experts in law, criminology, social-sciences and so on, to ensure ongoing alignment and prioritisation in our problem domain. And finally with ethicists, to ensure the application of these technologies remains consistent with community expectations.



Dine Differently



The Warnbro Tavern is your go-to local spot, where friends and family gather for memorable meals and refreshing drinks. Whether you're popping in for a beer or a glass of wine or settling in for a feast in our restaurant, we're perfectly nestled in the heart of Warnbro, making us an ideal local venue for your evening.

Enjoy our live music and entertainment as you indulge in last drinks. At The Warnbro Tavern, we're your prime destination for great times and great tastes.

Reserve a table now at

www.warnbrotavern.com.au

or phone us on (08) 9593 1597

7 Hokin St Warnbro WA 6169



Proudly Supporting Our Police Members Locally



the ACADEMY NEWS

from Tasmanian Chinese Buddhist Academy of Australia

August 2025

EVENT RECAP

The annual Otlands Heritage & Bullock Festival in Tasmania, is held over two days and blends old-world charm with heritage experiences, local crafts, and good old country hospitality.



As part of the Southern Midlands community, the Academy participated in the parade and followed up with performances near the famous Callington Mill - showcasing the holy dragon and lion dances to the crowd.

more about us

Light of Tasmania
on WeChat



The World of Jin
Gang Dhyana



Academy
Website



Academy
Facebook Page



get in touch

facebook us
enquiry form on our website
email at contact@tascbaa.org