

Cybersecurity, Privacy, and Generative AI in Schools



Cecelia Irvine-So
Practice Leader



Penny Liberogiannis
Senior Lawyer



Agenda

1

The emerging role and key risks of AI in schools

2

Case studies on recent hacking and breaches in schools

3

Existing legal framework

4

Updates to the Privacy Act and what schools need to do

5

Policy options for schools to respond to emerging AI

6

Interactive exercise: Simulated data breach and table discussions



“It’s a
deep fake”



World North America AI

AI-generated clip used to frame high school principal with racist remarks, police say

Ben Finley

April 26, 2024 – 6.25am

Saved

Share

A A A

0 Leave a comment



Listen to this article
4 min

A high school athletic director in Maryland has been charged with using [artificial intelligence](#) to impersonate a principal on an audio recording that



The emerging and potential role of AI in schools

Customised learning



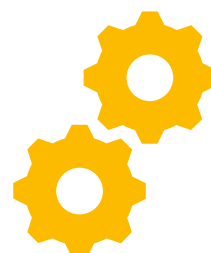
Tutoring or teaching assistants



Feedback



Automation of administrative tasks



Automation of marking



Generative AI





What is Generative AI

Generative AI:

- is a subset of AI.
- focuses specifically on creating content (text, images, music, media)
- ChatGPT
- Large Language Models (LLMs) are designed to understand and generate human-like text

Traditional AI:

- is about analysis and decision-making
- recognises patterns
- makes predictions



Identifying key risks

Privacy and data security risks

Volume of data
Unauthorised disclosure
Target for cyber-attacks
Child safety risks

Academic integrity

What is considered cheating
Rethinking assessments



Accuracy, reliability and suitability

Creativity and critical thinking
The black box problem
Learned skills



Discrimination

Biases learned by the system
Outcomes reflect the input
Input is the internet
LLMs can have gender bias

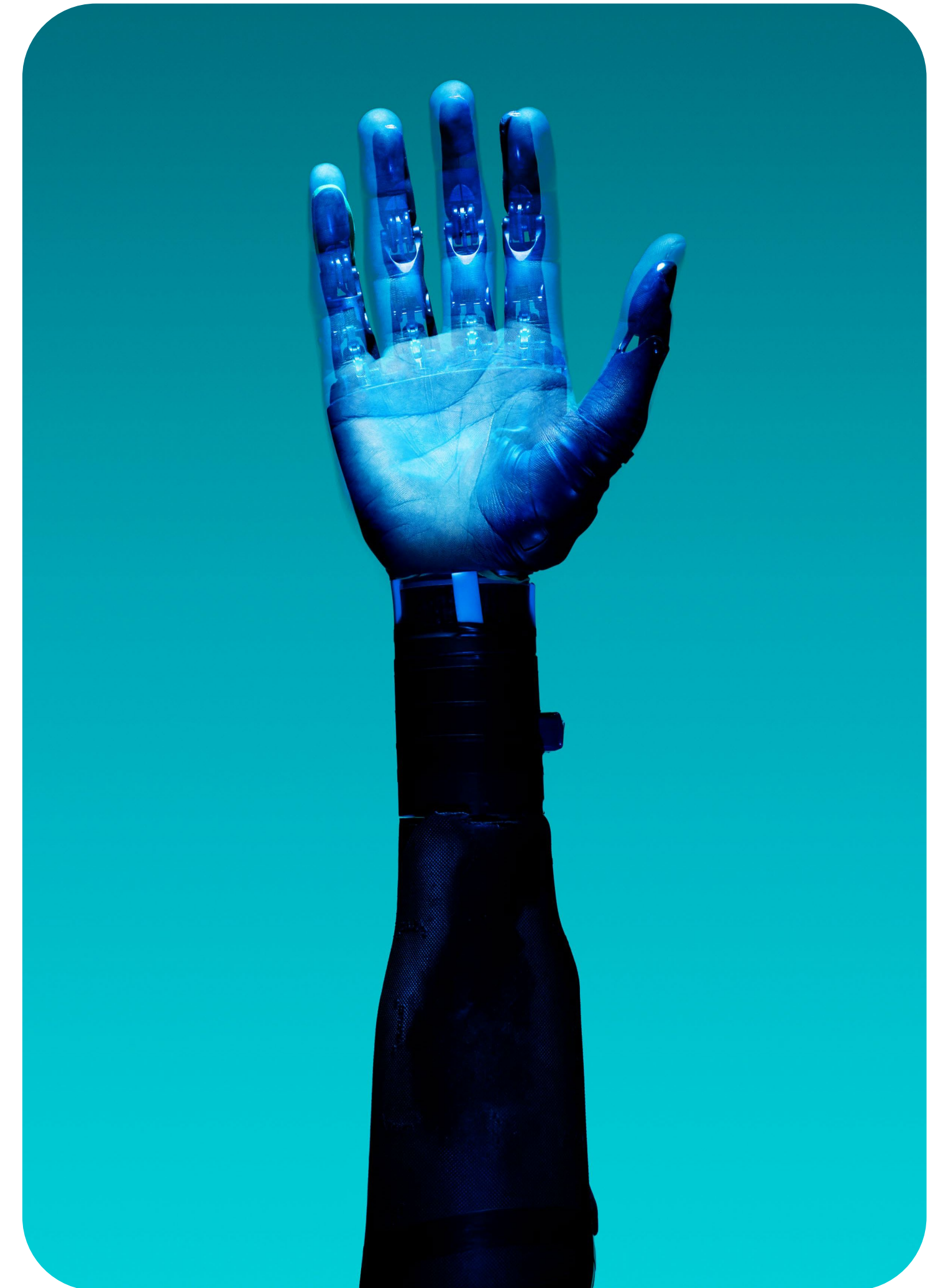
Privacy risks from AI in schools

Risks:

- Unintentional data exposure
- Inaccurate or biased AI models
- Lack of data minimisation
- Legal and regulatory compliance issues
- Loss of trust (think Robodebt)

Recommendations:

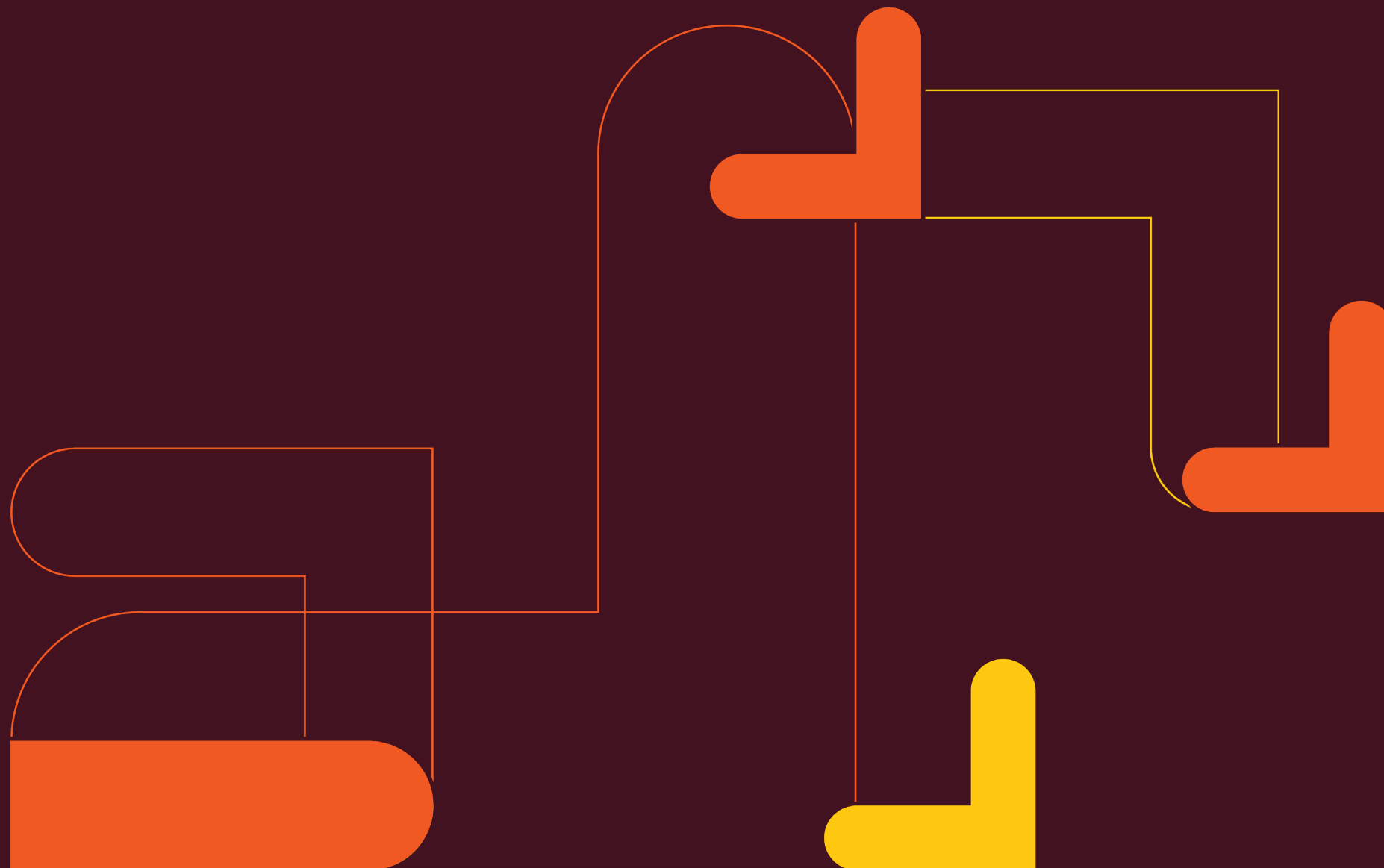
- Develop data literacy in staff and students
- Implement data handling guidelines
- Conduct Privacy Impact Assessments (PIAs)
- Promote accountability and privacy awareness





Case studies

These examples demonstrate real-life recent hacking incidents and data breaches in the education sector.





Case study 1 – online bullying and the Oyston case principle

Oyston Case Principle

- Ms Oyston was a Year 9 student who had brought a claim against her college in negligence alleging that she had suffered psychological harm from bullying from other students.
- The Court of Appeal agreed with the trial judge that the school had not taken sufficient steps to stop the bullying.
- The school had bullying policies in place, but critically, it had failed to implement its policy on innumerable occasions when the bullying incidents were reported.

Scenario

- A student at your school approaches you and informs you that they are being bullied at school.
- What do you do?
- Does what you do change if the bullying is online or in-person?



Case study 2 – student hacks into school via parent’s staff profile and changes results

Student hacks into school via parent’s profile and changes results

- Student accesses school software to view other student records (pastoral and assessment) which are private and confidential and the subject of a critical incident investigation.
- School considers expulsion and undergoes an in-depth investigation.

Scenario

- A student at your school hacks into their parent’s profile and changes their results.
- What do you do?



Case study 3 – image- based abuse; AI used to paste student faces onto explicit images

Image-based abuse; student faces pasted onto explicit images

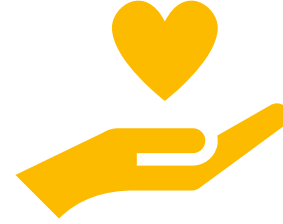
- ‘Deepfake’ images are becoming increasingly more prevalent as AI usage becomes more widespread.
- In June 2024, 50 students at Bacchus Marsh Grammar had their images taken from social media and pasted onto explicit images. In January 2025, a Sydney High School student was referred to police for allegedly creating deepfake pornographic images of female students.
- Students have also used images of teachers in explicit material.

Scenario

- A student at your school has used AI to paste other students’ faces onto explicit images.
- What do you do?



Existing laws and regulation affecting AI in schools



Duty of care



Academic integrity



Privacy (including privacy reforms)



Surveillance laws



Discrimination laws



Confidentiality, copyright and intellectual property



Duty of care

Reasonably foreseeable harm

What harm is reasonably foreseeable will depend on how you use AI in you school.

Staff? Students? Learning? Writing assessments? Decisions? Debt recovery?

Risk mitigations

- Standards of acceptable use
- Training and education
- Technological barriers to access
- Locked down or local AI only (garbage in, garbage out)





Academic integrity

Use case of AI in writing and marking assessments?

How can we ensure quality control of performance of AI?

Unsuitable for research as hallucinates results.

Educating students about primary sources vs GenAI.

GenAI used to do the **creative** thinking, instead of students





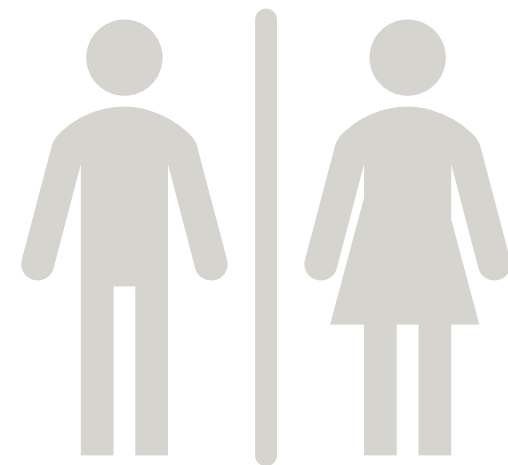
Discrimination

Sex Discrimination

Google translate has been found to have gender bias:

- Nurse = feminine
- Doctor = masculine

Facebook's ad algorithms excluding women from seeing jobs



Racial Discrimination

Facial recognition has less accuracy on people of colour.

Amazon's face surveillance technology called Rekognition.



Disability Discrimination

Accessibility limitations could lead to indirect discrimination.





Privacy

Inputting data may be considered a “**disclosure**”.

Are these disclosures permitted by the privacy principles?

APP 6



How will ChatGPT or any other generative AI system **use** your input data in the future?

APP 6

Downloading material may include **malware**.

What AI systems are staff using?

Could these open pathways for **malicious** actors?

APP 11

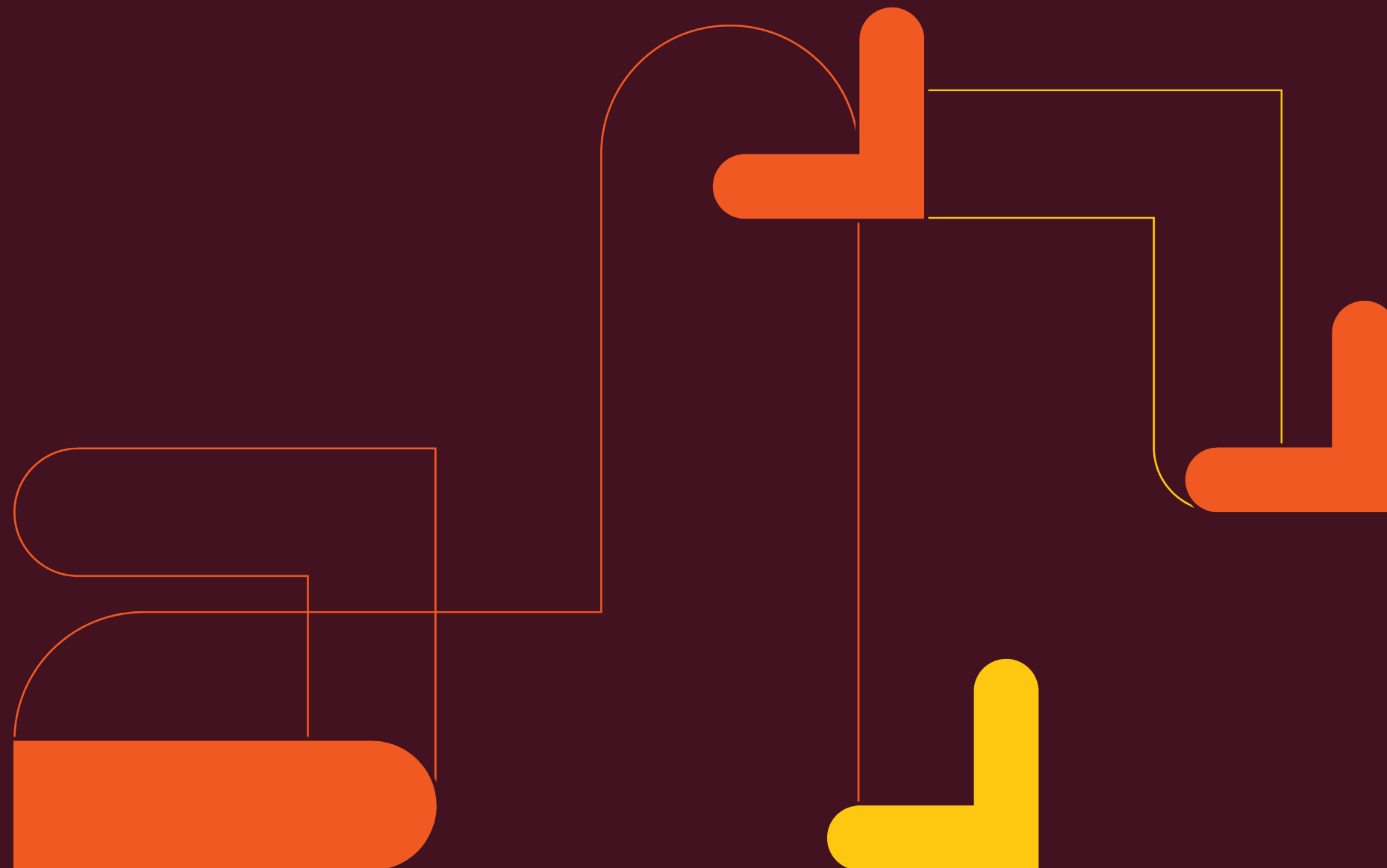
NDB Scheme





The black box problem

How AI makes decisions isn't clear.
We don't always know what led to the
decision being made.





Updates to the Privacy Act and what schools need to do

A statutory tort for serious invasions of privacy

Individuals may sue for serious invasions of privacy in cases where the individual entertained a reasonable expectation of privacy, though exceptions do exist

Development of a Children's Online Privacy Code

Information Commissioner will draft and implement a contemporary Children's Online Privacy Code within 2 years from the date of which the powers come into effect

Enhanced OAIC Powers

OAIC has been given ability to use investigation and monitoring powers, replacing its weaker regulating entry and inspection powers

Greater transparency of automated decision-making

Schools must clearly outline where a computer has access to student personal information

In 2025, schools should be reviewing their privacy policy and privacy management frameworks to respond to the incoming changes and ensure they are equipped to continue handling information responsibly.

Privacy reform – automated decision making

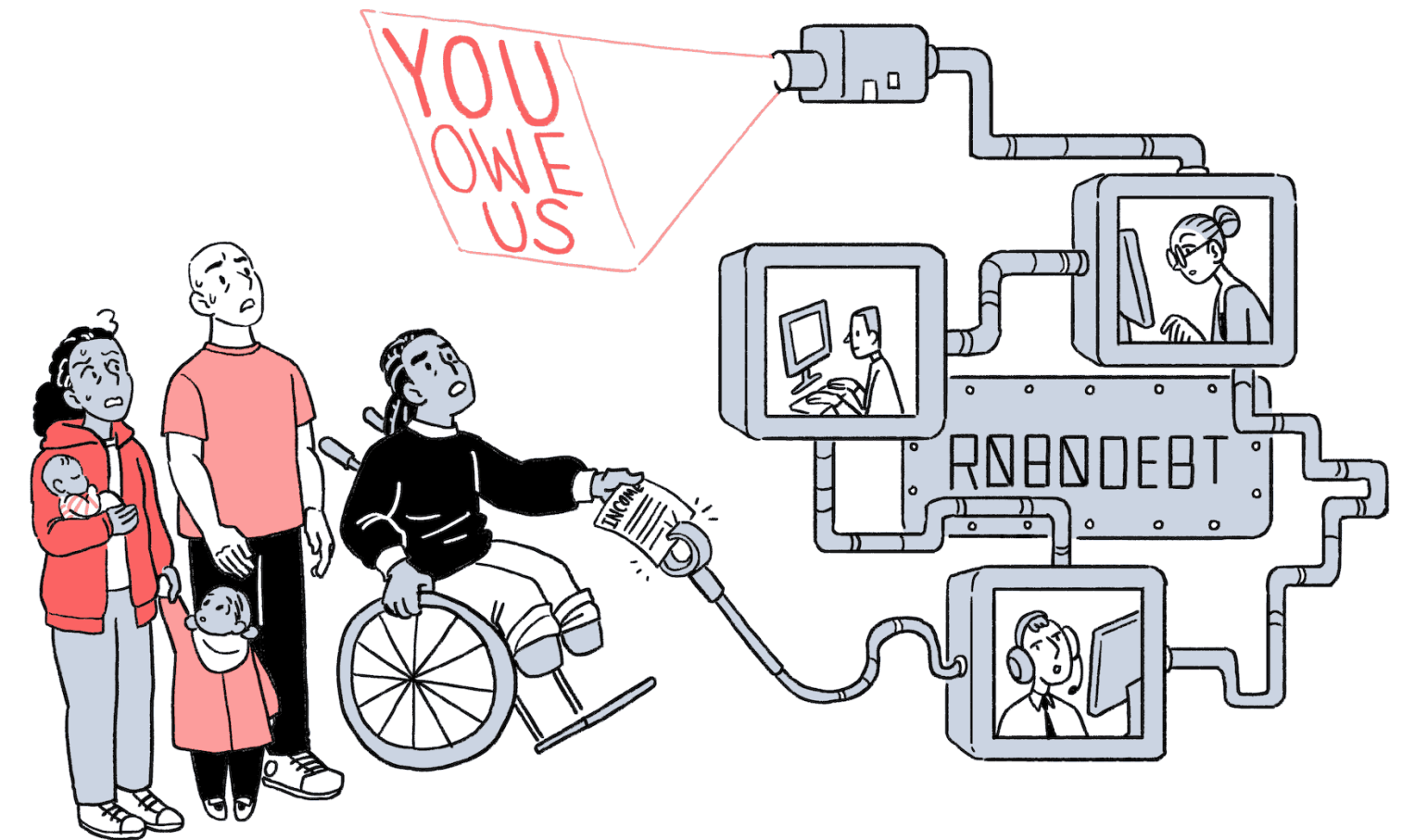
Automated decision making is ***distinct*** from AI. Robodebt was automated decision making, ***not*** AI.

Privacy Act Review

Acceptance, or acceptance in principle, of recommendations.

Enshrine a right to request meaningful information about how automated decisions are made.

This is in addition to the access, correction and complaint rights already existing, and a possible right to deletion.





Surveillance laws

Surveillance is **physical**.

Privacy is about **information** and data.

Surveillance of students during online learning or assessments to stop cheating.

Consequences of data collection and psychological impact of surveillance.





Confidentiality, copyright and intellectual property

Ownership of curriculum,
worksheets, assessments.

If teachers load worksheets into
GenAI, the IP could be used by
others also using that platform.

Where will the input data pop up in
the future?

Could the school infringe copyright
of others by using material taken
from GenAI?





Understanding the Terms of Service and contractual requirements

- Privacy requires open and transparent terms – plain English.
- If you can't understand it, don't sign it.
- What is the training data?
- What are the retention periods?
- Is there a breach notification clause?
- Is there are requirement for the other party to comply with the *Privacy Act 1988* (Cth)?



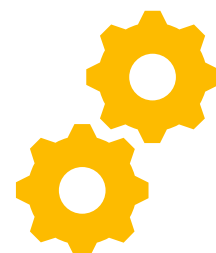


The AI Ethics Framework

1. AI systems should **benefit** individuals, society and the environment.
2. AI systems should **respect human rights, diversity**, and the autonomy of individuals.
3. AI systems should be **inclusive** and accessible and should not involve or result in unfair **discrimination** against individuals, communities or groups.
4. AI systems should respect and uphold **privacy rights** and data protection and ensure the **security** of data.
5. AI systems should reliably operate in accordance with their intended purpose.
6. There should be **transparency** and responsible disclosure so people can **understand** when they are being significantly impacted by AI and can find out when an AI system is engaging with them.
7. When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to **challenge** the use or outcomes of the AI system.
8. People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

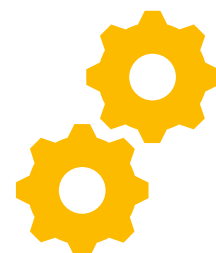
Strategies to implement AI in schools

- **Overarching framework:**
 - Create a shared, agreed upon school-wide **framework** for the use of AI tools to raise awareness and empower the school community to use AI ethically
 - Define what is **acceptable** and what is unacceptable: What constitutes cheating? What are the consequences?
 - When can teachers use AI for lesson plans and/or marking and feedback?
- **Assessment protocols for educational staff:**
 - Modify assessments to reduce ability to rely on Generative AI
 - Can you conduct oral assessments?
 - Include creative and personal elements to assessments
 - Ongoing engagement with student progress (i.e., draft submissions)



Strategies to implement AI in schools

- **Introduce safeguards:**
 - Use compliant Privacy impact assessments suitable for the education sector to ensure a privacy-by-design approach to any projects that may involve the use of AI
 - Demand transparency – schools should clearly communicate to students, parents, and educators which data is being collected, who it is being shared with and how it is being used
- **Staff training on AI, its capabilities and how to avoid data breaches:**
 - Whitelist any AI tools you have approved and any that are used only within your closed system
 - Provide guidance and establish clear guidelines for data handling, specifying the types of data that can be entered into AI systems and ensuring that only de-identified or otherwise relevant and necessary information is used.



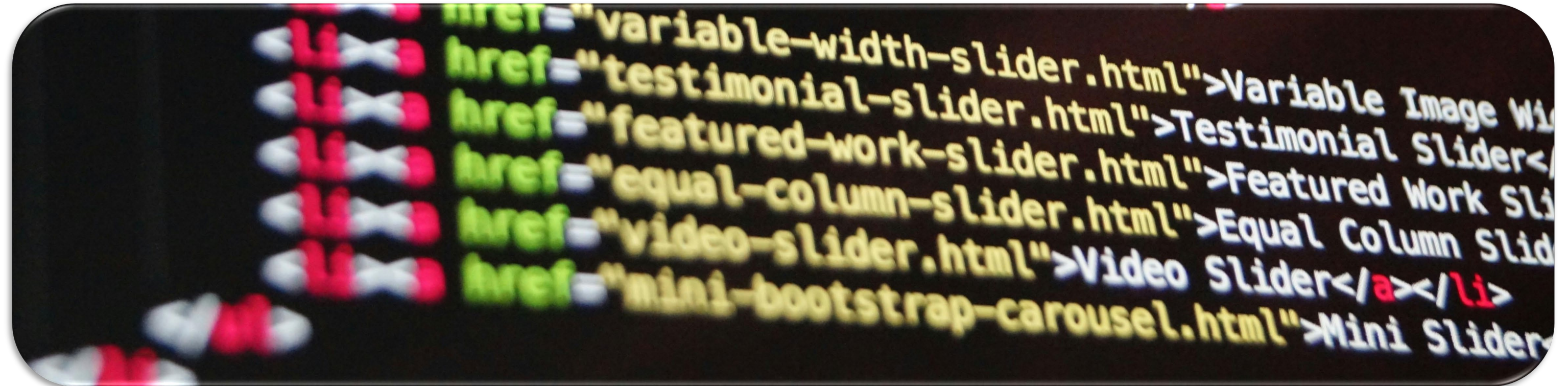
Additional policies to add to the puzzle

Data handling guidelines

- Internal information handling protocols or expectations
- What information can be inputted into GAI as prompts and what cannot
- What AI systems can be used or not (whitelisting)

Privacy Impact Assessments (PIAs)

- Change management tool
- Consider privacy risks prior to adopting new software or process





Interactive Exercises: Simulated data breaches and table discussions

1. Attackers encrypt and incapacitate vital systems including everything from emails to physical security measures. **What do you do?**
2. Attackers demand a ransom exceeding \$1 million in cryptocurrency for their release. **What do you do?**
3. Once the school has eradicated the malware, looking towards the future - **what do you do?**

1. A teacher's laptop has been hacked. **What do you do?**
2. A day later, the hackers send an email from the Principal's email address to parents requesting personal information and credit card details. **What do you do?**
3. Once the school has dealt with the data breach, looking towards the future - **what do you do?**

Summary: Key Themes

- Ransomware attacks
- Compromised devices
- Rebuilding school's CRM
 - School responses
- Cybersecurity measures



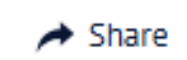
Updated National Victoria Crime

Year 11 students suspended after fake nudes circulated online



Daniella White

February 21, 2025 – 12.15pm



Two students have been suspended after fake, sexually explicit images of students from a school in Melbourne's north-west were circulated online.

Victoria Police said an unknown number of formal photos of female students at Gladstone Park Secondary College had been “inappropriately manipulated using artificial intelligence” before being released on the internet.



Questions?



Cecelia Irvine-So

Practice Leader

cirvine-so@moores.com.au
03 9843 2121



Penny Liberogiannis

Senior Lawyer

pliberoigiannis@moores.com.au
03 9843 2123