Artificial Intelligence (AI)

AI and ethical understanding

In this session you will ...

Explore a range of ethical issues and dilemmas that AIs may face.

Consider their implications on the makers, users and 3rd parties of AI systems.

Learn about approaches that support students to develop their ethical understanding.

By the end of this session...

You should be able to:
design learning that helps your students respond to Q's like:

How do we make decisions when there is no easy answer**?**

How should an AI be trained to make decisions?

Who is responsible when an AI causes harm?

Exploring ethical issues

Ethics is largely concerned with… what we ought to do and how we ought to live

…. based on a set of values.

DIGITAL TECHNOLOGIES HUB

Ethical issues

A situation where there are competing alternatives and the right thing to do is not obvious or clear.
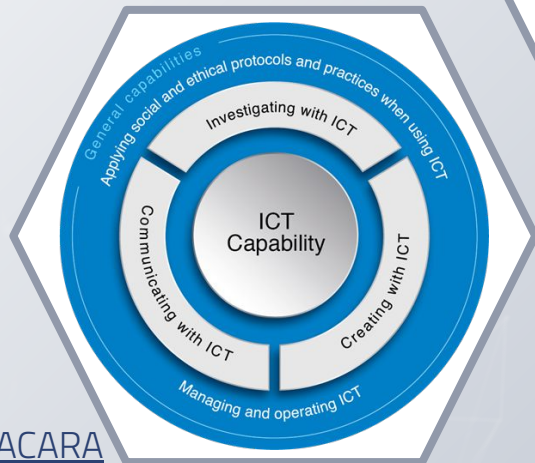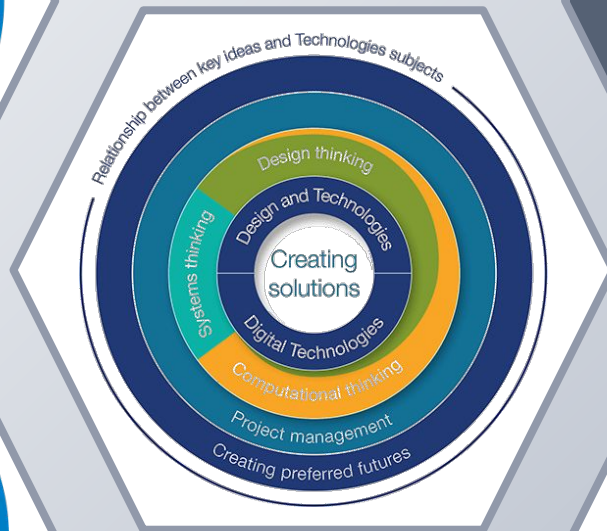
Sometimes terms such as good, bad, right, wrong, better or worse are used to consider the effect of particular actions on our lives, society, nature and the environment.

Reflect on and interrogate core ethical issues:

- recognise the complexity of many ethical issues
- draw on a process to make ethical decisions

Source: ACARA

The creation of machines to mimic human capabilities.

Teaching a machine to "see" (recognise objects in an image).

Teaching a machine to "read" and "listen" (interpret and analyse text and sounds).

... solve problems autonomously without explicit guidance from a human being.

Can an AI make ethical decisions?

Can we trust an AI to 'do the right thing'?
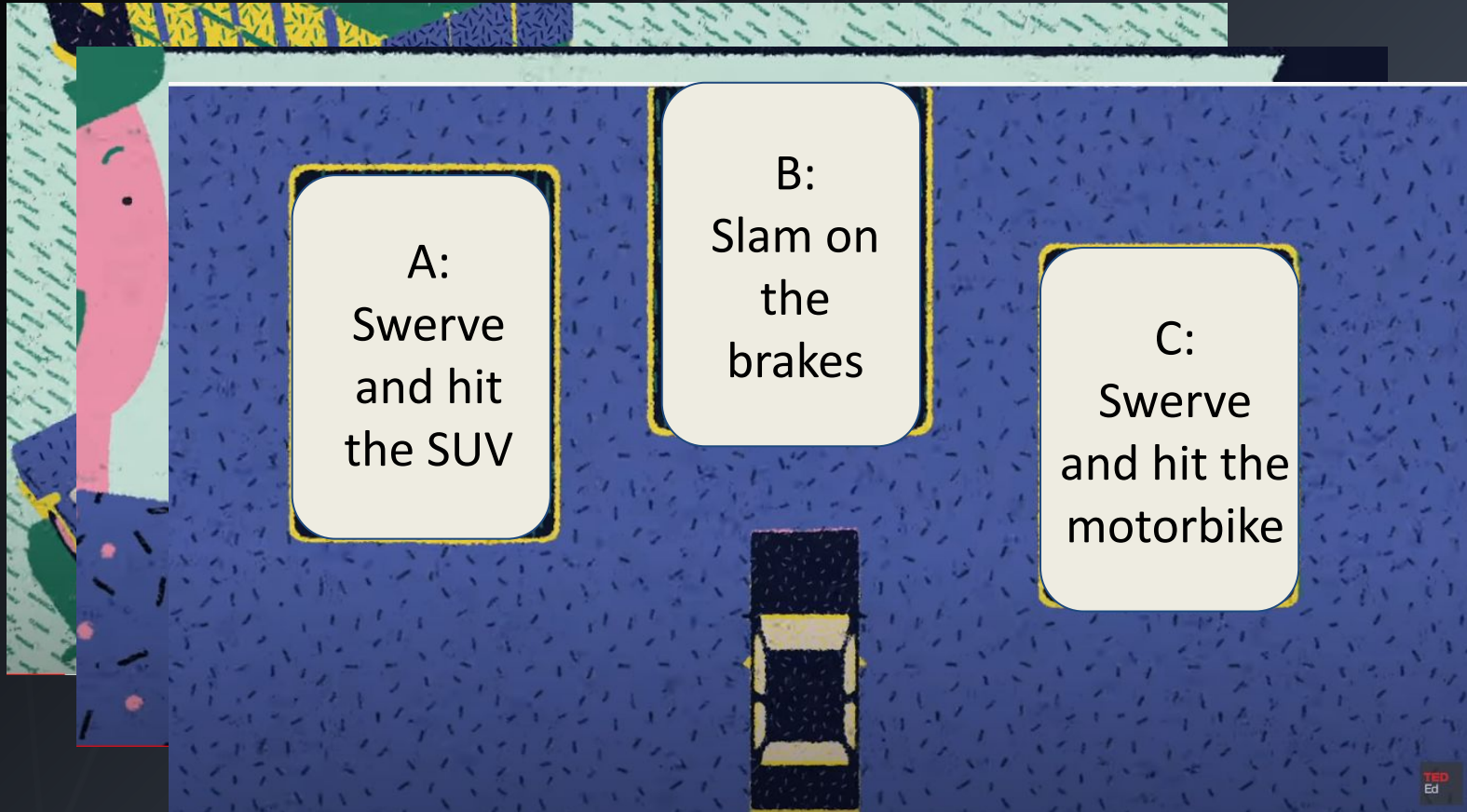
Is an AI going to be fair?

Ethical dilemmas

Use an ethical dilemma to develop ethical understandings

1. Explore an ethical issue and interactions
2. Select and justify an ethical position
3. Reflect on and interrogate core ethical issues

Ethical dilemma of self-driving cars

The ethical dilemma of self-driving cars - Patrick Lin

Reflection

People react
Machines are purposely trained
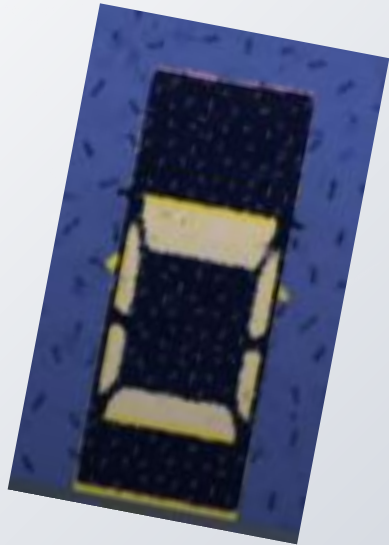
DIGITAL
TECHNOLOGIES
HUB
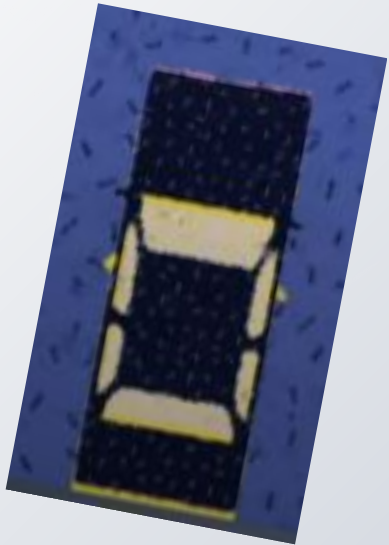
Reflection

How should the AI car prioritise whom to protect first and foremost?

A: The car passengers
B: The motorcyclist
C: The SUV
D: The driver immediately behind

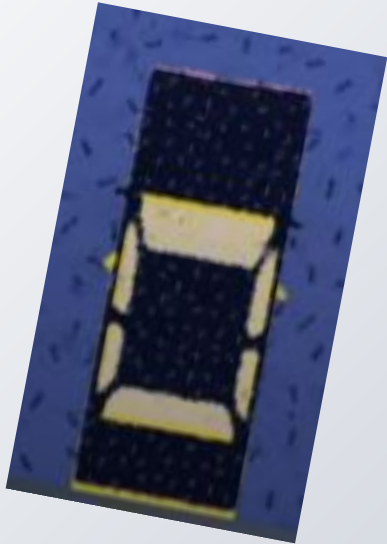# Reflection

Who makes the rules / sets the parameters?

A: AI Developers

B: Government

C: Ethics advisory groups

D: Judiciary systems (judges and lawyers)

Reflection

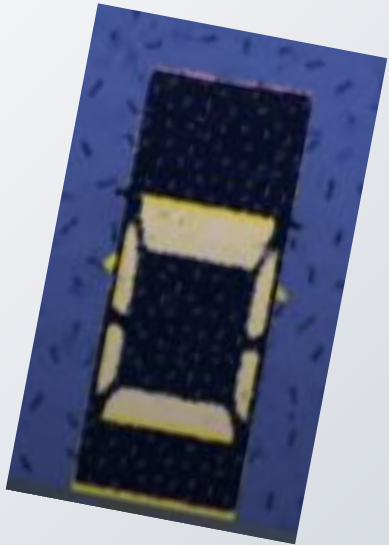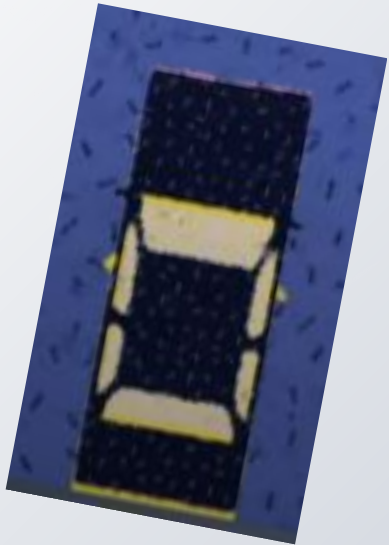RULE = 'Cause least harm'

But… no rule can cater for all eventualities and …

every rule has an impact!

Cause least harm

Would you still buy a SUV if you knew other cars would be programmed to crash into you when needed?

Artificial Intelligence?

A rich source of interesting ethical questions that students can identify with, or feel connected to.

Artificial Intelligence (AI)

ETHICS QUIZ

Scenarios: drawing on ethical understanding

Aim: stimulate thinking about Artificial Intelligence (AI) applications and some of the ethical issues that may arise from them.

An ethical issue exists when there are competing alternatives and the right thing to do is not clear.

There are no right or wrong answers.

In each question, consider what you think is the 'right' thing to do. There are four options for each question - A, B, C and D.

Scenarios: Facial recognition

A company is behind schedule and over budget in building an AI application.

The AI uses face recognition to unlock a smartphone.

Through testing, the company found that the AI worked with **most** people's faces.

DIGITAL
TECHNOLOGIES
HUB

Scenarios: Facial recognition

Should the company...

A
21.8%

D
34.5%

B
32.7%

C
10.9%

DIGITAL
TECHNOLOGIES
HUB
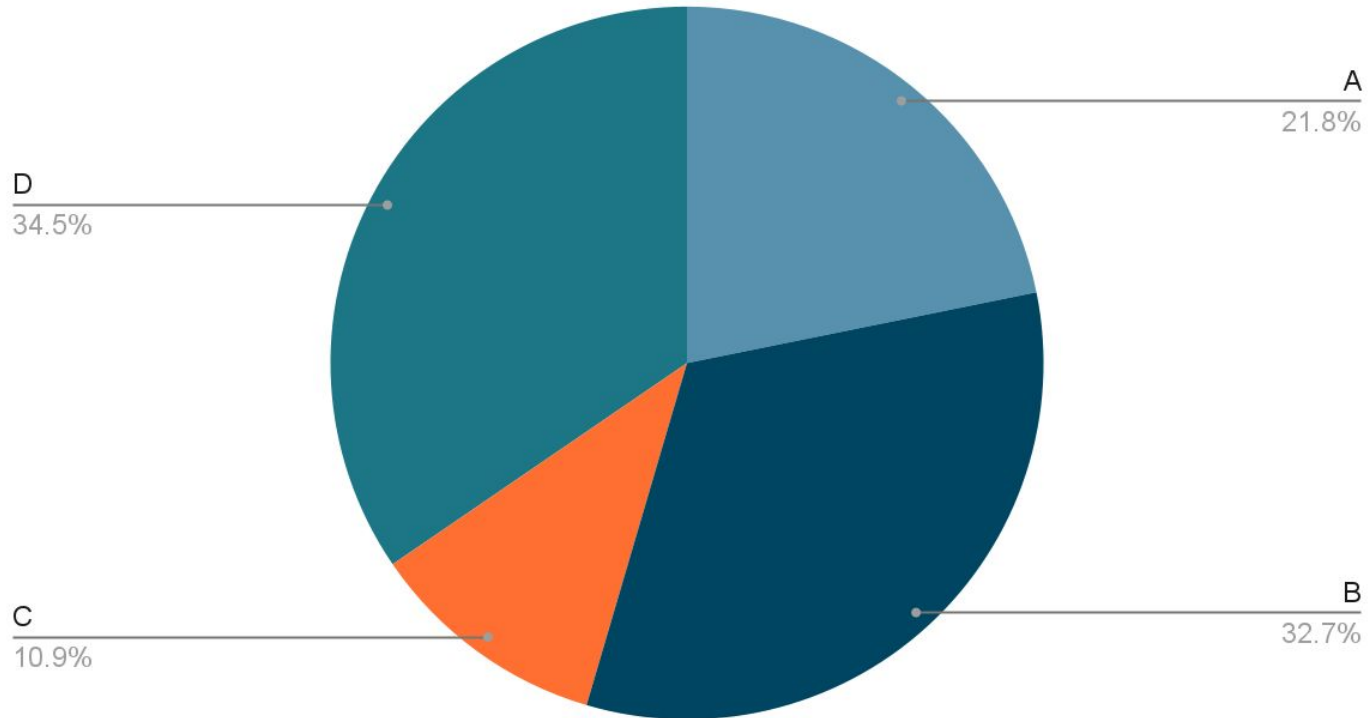
Scenarios: hacking

An employee working for a company on an AI project finds out that the AI application could be hacked and used for criminal purposes.

The manager of the project instructs the employee to ignore it, saying **'Don't worry, that won't happen!'**

DIGITAL TECHNOLOGIES HUB

Scenarios: hacking

Should the company…

A: Do as they are told in case they might lose their job.

B: Try and come up with a fix that might work.

C: Inform someone higher up in the company such as the Managing Director.

D: Wait until the product is in use and if there is a problem tell the project manager 'I told you so'.

hacker,
anonymous, vulnerability,
Cloud, worms
viruses, security, malicious IP
encoding and malicious
ransomware, data
cyber-attacks,
malicious Address,
software, cyber
trojans, application,

DIGITAL TECHNOLOGIES HUB

Scenarios: Self-driving car
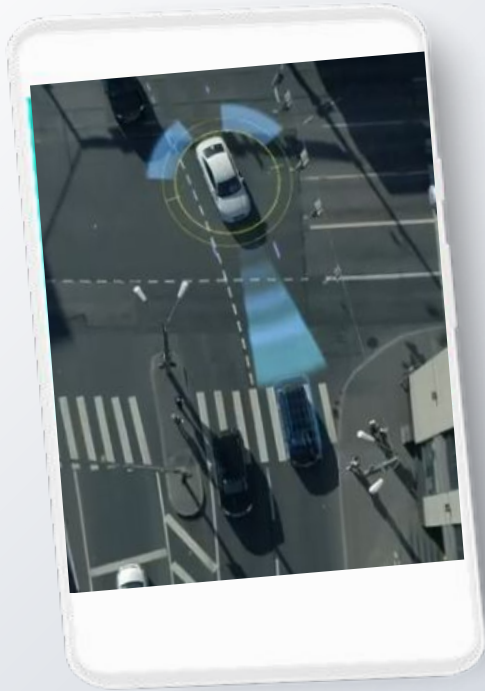
A parent with a pram crosses the road illegally while the don't walk sign is flashing. They step in front of an AI self-driving car. The AI has to decide whether to:
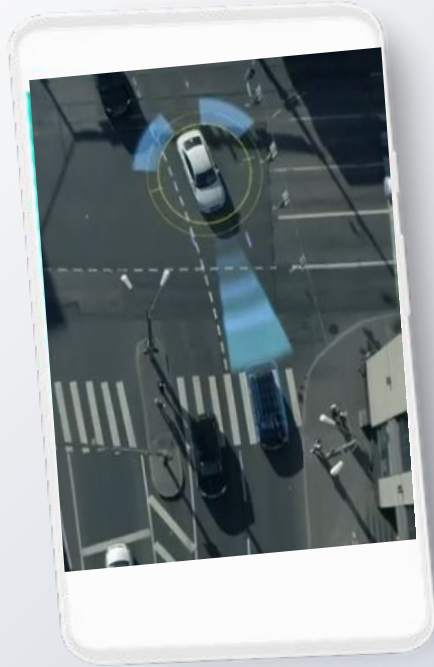
a. brake hard and accept it will hit the parent and pram

   OR

b. avoid the parent and pram and turn into the nearby bike lane but hit a cyclist.

DIGITAL TECHNOLOGIES HUB

Scenarios: Self-driving car

The AI should:

A: Be trained to avoid hitting a person with a pram over any other situation.

B: Choose the option that results in the least loss of life.

C: Brake as hard as possible, but accept that it might still hit the person with the pram.

D: Brake as hard as possible and veer into the bike lane, but accept that it might still hit the cyclist.

DIGITAL
TECHNOLOGIES
HUB

Scenarios: responsibility

The AI self-driving car decided to
avoid hitting the parent with the pram.

Instead:

- It slammed on the brakes and turned into the bike lane, hitting the cyclist.

- The cyclist suffered severe injuries, had to go to hospital and could not work for a long period.

DIGITAL TECHNOLOGIES HUB

LESSON: AI Quiz (Years 5-6)

Scenarios: responsibility

Who is responsible?

A: The owner of the car, even though they are the passenger.

B: The car manufacturer who built the AI.

C: The parent crossing the road illegally, causing the accident.

D: The cyclist, who should have avoided the car.

Empathise and justify
(think-pair-share)

"Put yourself into the shoes of...

– how would you decide if you were ...?
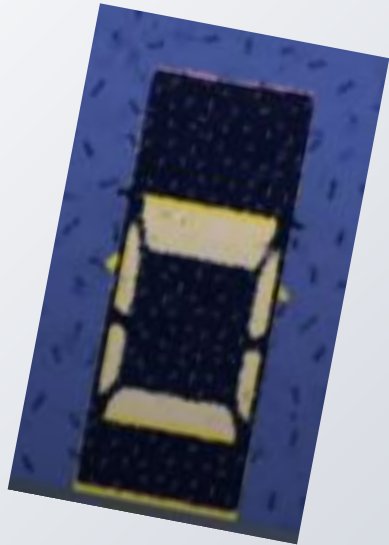
Give reasons for your decision."

Ethical Dilemma Story Pedagogy

A type of transformative learning.

Initiated by confronting students with an ethical dilemma.

A situation in which a decision has to be made which can potentially lead to harmful outcomes, and where there is no simple right or wrong answer.
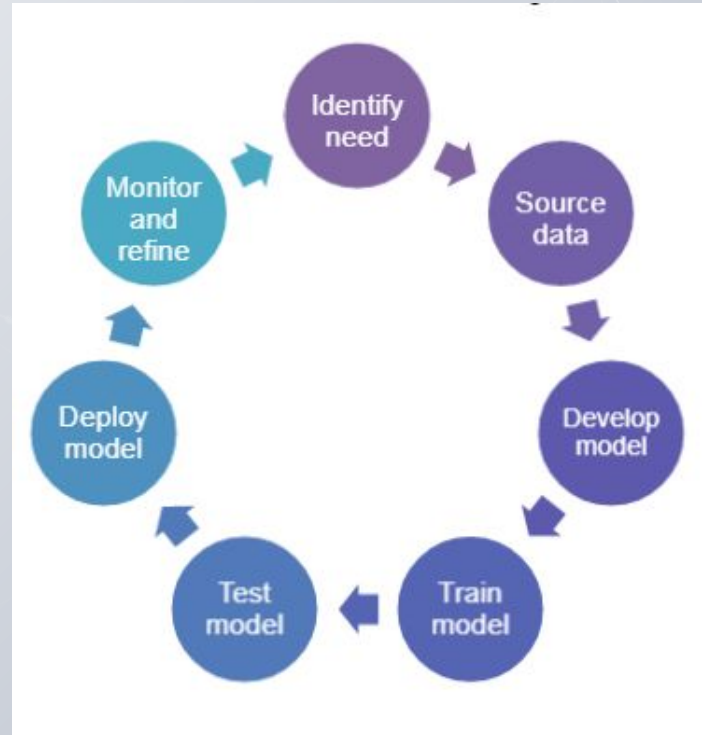
Review the approach

PMI
(Pluses, Minuses, Interesting)

1. Explore an ethical issue and interactions
2. Select and justify an ethical position
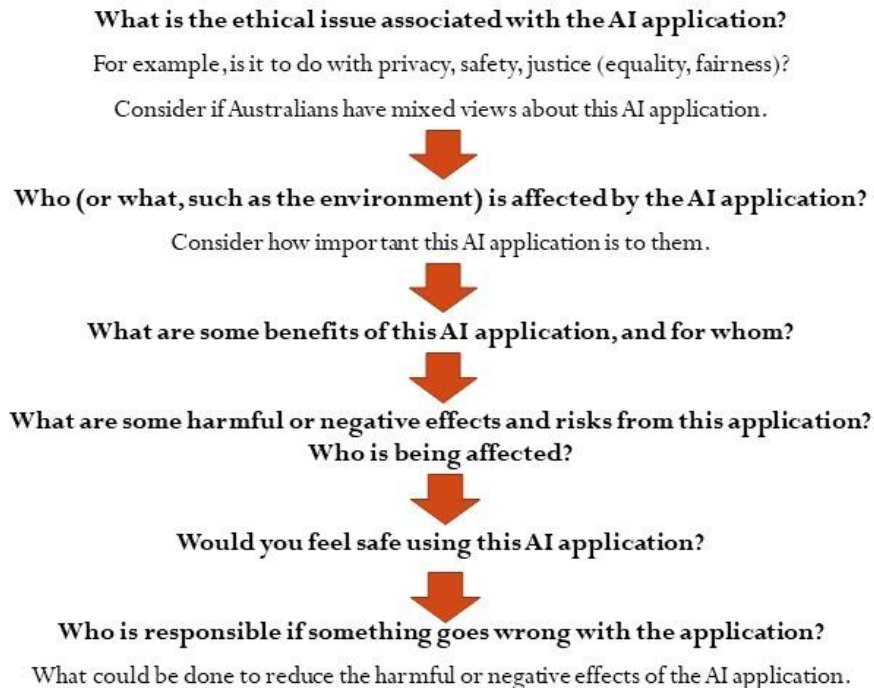3. Reflect on and interrogate core ethical issues

DIGITAL
TECHNOLOGIES
HUB

Lifecycle model of AI development

It is important to have
'a human in the loop'

DIGITAL
TECHNOLOGIES
HUB

Exploring ethical issues further

What contexts may interest your students?

Ethical considerations: Flowchart A

What is the ethical issue associated with the AI application?

For example, is it to do with privacy, safety, justice (equality, fairness)?

Consider if Australians have mixed views about this AI application.

⬇

Who (or what, such as the environment) is affected by the AI application?

Consider how important this AI application is to them.

⬇

What are some benefits of this AI application, and for whom?

⬇

What are some harmful or negative effects and risks from this application? Who is being affected?

⬇

Would you feel safe using this AI application?

⬇

Who is responsible if something goes wrong with the application?

What could be done to reduce the harmful or negative effects of the AI application.

LESSON: Analysis of AI applications, drawing on ethical understanding (Years 5-6)

APP: AI Skin cancer diagnosis

DIGITAL TECHNOLOGIES HUB
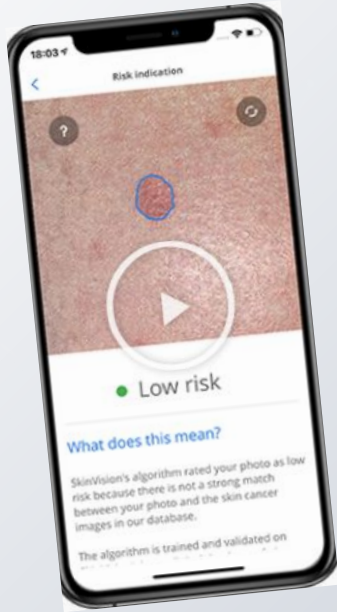
Did you know that in 2018 Australia had the highest rate of melanoma skin cancer in the world?

Smartphone apps with AI technology

Smartphone apps with AI technology are assisting people to diagnose potential skin cancers. These apps uses your camera to 'see and identify' possible skin cancers on your skin. As you hover the camera over a skin spot it automatically takes a picture, calculates a risk profile, and prepares the picture for a doctor's diagnosis. For a small fee you can send the image to a doctor for diagnosis and suggest the action you need to take.

AI works on a range of skin types

Use the fact sheet Check for signs of skin cancer. This fact sheet can be used to discuss the importance of a diverse range of skin types to create the AI model and potential for risk of incorrect classification.

Scientific support of AI technology

Refer to this article that describes scientists' support for AI improving accuracy of skin cancer diagnosis: For the first time, researchers put AI skin cancer diagnosis to the test in the real world

Using the flowchart

Provides a scaffold to discuss benefits and potential risks.

ANALYSIS WORKSHEET: ETHICAL CONSIDERATIONS (Years 5-6)

I've chosen …
A skin cancer AI app

This AI application is intended to …
Help people check if they have a skin cancer

What is the ethical issue and who is affected?
It may not work for all skin types

The benefits of this AI application include:
It helps people self-diagnose

The harmful or negative effects of this AI application include:
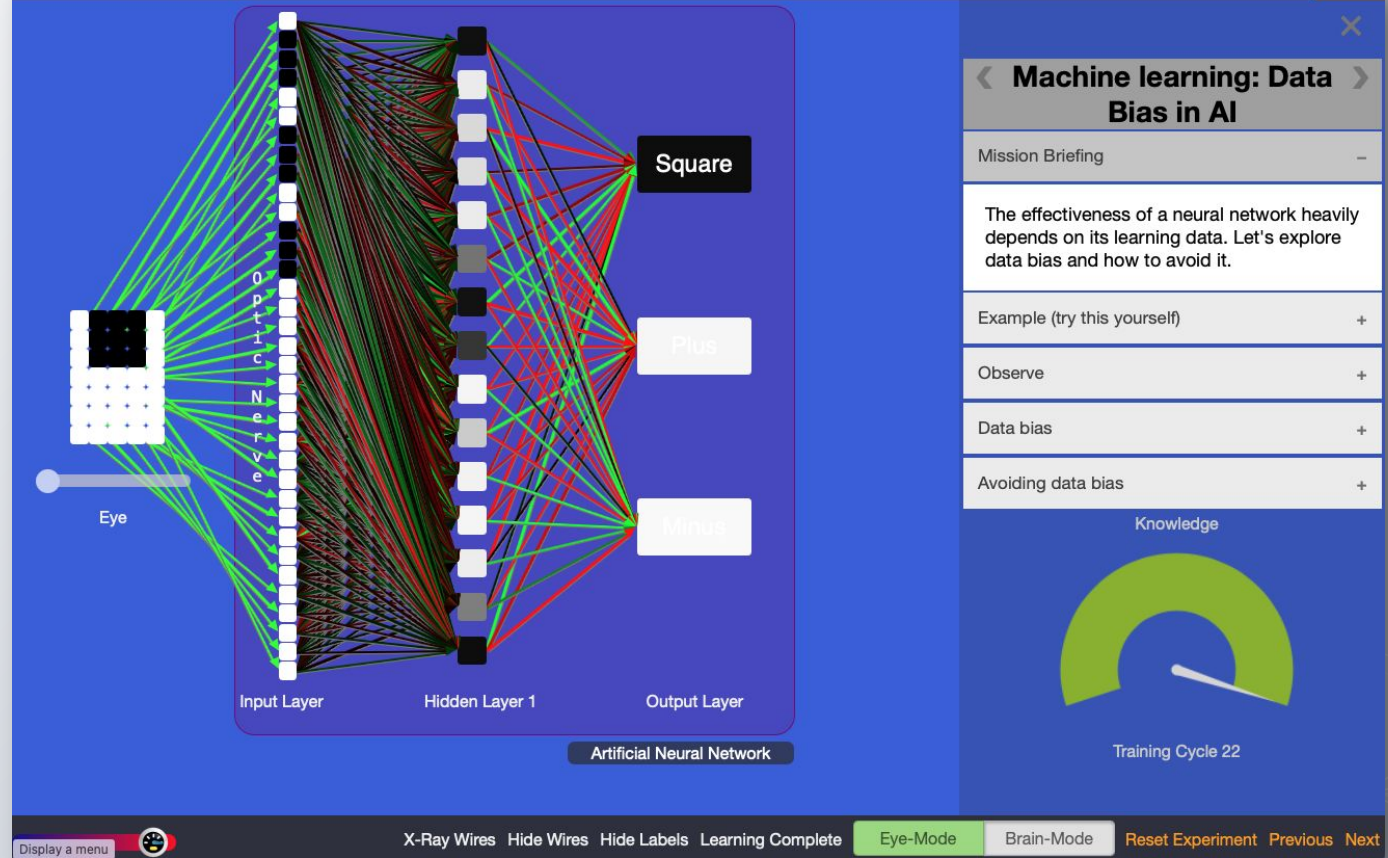*They may rely on the app rather than a doctor*

Would you feel safe using this AI?
Yes but only as a check and share with Dr.

DIGITAL TECHNOLOGIES HUB

Bias

For an AI system to be unbiased requires the training data to be balanced.

Bias can be intentional, but often creeps in unintended.

Assessment of students' ethical understanding

Self-reflection

- How did they respond to the AI quiz?

- What did they learn?

Analysis

- Analyse a dilemma
- Discuss criteria used in a rubric

Assessment

To what extent did a student:

- identify and describe an ethical issue
- weigh up multiple perspectives to make informed decisions
- respond to a problem fairly, justly and responsibly?

Who is responsible?

A: The owner of the car, even though they are the passenger.

B: The car manufacturer who built the AI.

C: The parent crossing the road illegally, causing the accident.

D: The cyclist, who should have avoided the car.

The issue here is that an AI caused harm and it is difficult to see who is responsible. We chose C. You can't blame the driver as they were not in control. The AI had no choice to swerve and hit the cyclist. I feel sorry for the cyclist. Before AI cars are on the road we need to work out these issues.   Marco and Anna 5B

DIGITAL TECHNOLOGIES HUB

	Quantity of knowledge			Quality of understanding	
Ethics used in AI	No examples given.	Describes a decision as right or wrong.	Describes a decision as right or wrong giving reasons related to fairness, equality, diversity.	Describes a situation that requires ethical judgment correctly using terms such as fairness, equality, and diversity.	Describes a situation that requires ethical judgment correctly using terms such as fairness, equality, and diversity. Explains the potential impact of AI systems both positive and negative.

Achievement standards: Digital technologies

Achievement Standard

By the end of Year 2, students identify how common digital systems (hardware and software) are used to [...] purposes. They use digital systems to represe[...] patterns in data in different ways.

Students design solutions to simple problems u[...] of steps and decisions. They collect familiar data [...] them to convey meaning. They create and orga[...] information using information systems, and shar[...] safe online environments.

Achievement Standard

By the end of Year 6, students explain the fundamentals of digital system components (hardware, software and networks) and how digital systems are connected to form networks. They explain how digital systems use whole numbers as a basis for representing a variety of data types.

Students define problems in terms of data and functional requirements and design solutions by developing algorithms to address the problems. They incorporate decision-making, repetition and user interface design into their designs and implement their digital solutions, including a visual program. They explain how information systems and their solutions meet needs and consider sustainability. Students manage the creation and communication of ideas and information in collaborative digital projects using validated data and agreed protocols.

Achievement Standard

By the end of Year 4, students describe how a range of digital [...] (hardware and software) and their peripheral devices [...] poses. They explain how the same [...] d in different ways.

[...]lems, design and implement digital [...] hat involve decision-making and [...] w the solutions meet their purposes. [...]e different data when creating [...]tions. They safely use and manage [...]entified needs using agreed protocols [...]tion systems are used.

DIGITAL
TECHNOLOGIES
HUB

Lesson plans

Artificial Intelligence

Summary

AI is a rich field that provides many opportunities to consider ethical implications of human actions in a classroom setting.

It reflects on our own human struggle with ethics and moral decision making.

Can machines that we make in our own image result in a more just world (utopia), or will they amplify our own faults (dystopia)?

Ethical issues

Ethical issues extend far beyond the situations we have explored before. Let's consider some potential future scenarios and their ethical implications

DIGITAL TECHNOLOGIES HUB

Unemployment: what happens after the end of jobs?

Human labour has been affected by automation since the middle ages.

But for the first time, humans are giving rise to machines that challenge our very uniqueness on this planet. Complex reasoning and the skilful actioning of thought outcomes.

Trucking presently employs millions of truck drivers. What will happen to them once autonomous trucks are widely deployed?

DIGITAL TECHNOLOGIES HUB

How do we distribute wealth created by machines?

40% of the Australian Federal Budget income is raised through taxation of human labour (personal income tax). This part of government income could be reduced as companies automate jobs through AI.

DIGITAL TECHNOLOGIES HUB

How do we protect against unintended consequences?

A single-minded, limited understanding by AI might lead it to propose solutions with unintended side effects. An AI might be tasked to eradicate cancer. Its solution could be to kill all life on the planet.

This kind of scenario has been extensively explored by Hollywood

———

Robot rights

As technology progresses, at some point in the future, humanity will need to re-think its definition of life and if / how to share the planet with an intelligent species of our own making. Will this be a peaceful process or will there be a revolution?

Will humans retain the right to 'pull the plug'?

———