Responsible and Ethical AI Lesson Plan

# Presenter Background Notes

## What is AI?

Artificial intelligence as a collection of interrelated technologies used to solve problems autonomously and perform tasks to achieve defined objectives without explicit guidance from a human being. While AI has been around for decades, advances in the layers of technology required to make AI work have made significant advances resulting in a significant increase in the capability and complexity of AI available to the public, and its significant increase in everyday use.

For more information on what is AI and how is it impacting our lives, we recommend looking at the CSIRO podcast series [Everyday AI](https://www.csiro.au/en/news/podcasts/Everyday-AI-podcast) the [United Nations Generative AI Primer](https://unite.un.org/news/oict-ett-launches-generative-ai-primer) and the [Digital Technologies Hub](https://www.digitaltechnologieshub.edu.au/).

## Ethical and Responsible AI

Across the globe, Artificial Intelligence is transforming the way we work, socialise and deliver services. Understanding the strengths and limitations of AI is critical. There is no doubt that AI has great potential to improve our well-being and society but there is no guarantee it will do so. Ultimately, it must be people who determine the ethicality of AI and ensure that it is developed and used in a way that is positive and beneficial to all. For AI to be used and trusted by people, organisations and government it must be guided by sound ethical principles.

When considering the concept of ethical and responsible use of AI, we need to ask ourselves if the ethical boundaries of AI are different to any other areas, or do we simply need to apply the same principles to a new frontier? Australia is already party to seven core human rights agreements which have shaped our laws and impact how we engage with AI.

Fairness is a difficult concept to pin down and AI designers essentially have to reduce it to statistics. Because the machines themselves cannot understand the complex concept of fairness, developers must come up with mathematical definitions to define what it means. It is not always easy to make ethical decisions, and ensuring AI treats the many unique groups and individuals in Australia fairly is a challenge. It is inevitable that decisions will be made that some believe are fair and others do not. When the stakes are high, it is imperative to have a human decision–maker accountable for automated decisions. Australian laws already mandate it to a degree in some circumstances.

To help navigate this concept of fairness in AI, CSIRO developed the Artificial Intelligence ethics framework. One of the first of its kind to be created in Australia, the framework assumes that technology should serve the best interest of human beings and be aligned with human values.

The position of the development team was that an ethics framework for AI is not about rewriting existing laws or ethical standards, it is about creating a resource to ensure that existing laws and ethical principles can be applied in the context of new AI technologies.

The framework identified eight core principles. These principles are used to judge potential AI projects and determine if they meet the requirements to pursue development. This document was then used as the basis for the 2019 Australian Department of Industry, Science and Resources (DISR) voluntary Australian AI Ethics framework for businesses and again in 2023 for the Australian Framework for Generative AI in schools.

For more information about the ethical and responsible use of AI, we recommend [CSIRO’s Artificial Intelligence Australia’s Ethical Framework – A discussion paper](https://www.csiro.au/-/media/D61/Reports/Artificial-Intelligence-ethics-framework.pdf), the [Department of Industry, Science and Resources Australia’s Artificial Intelligence Ethics Framework](https://www.google.com.au/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjE1sruhfOCAxUYQt4KHU1gDE8QFnoECA4QAQ&url=https%3A%2F%2Fwww.industry.gov.au%2Fpublications%2Faustralias-artificial-intelligence-ethics-framework%2Faustralias-ai-ethics-principles&usg=AOvVaw2Mva03lv3__Xb-Z3HDwFfu&opi=89978449) and the [Australian Framework for Generative AI in schools](file:///C://Users/kin285/Downloads/Australian%20Framework%20for%20Generative%20AI%20in%20Schools%20(3).pdf).

In addition to the initial evaluation, continuous review must be a part of the AI process to ensure the development and use remain ethical. There is no one-size-fits-all solution to these emerging issues. To complement the eight principles, CSIRO has also developed a tool kit that can be used to assess risk and ensure compliance and oversight. This toolkit can be found in the CSIRO AI discussion paper.

## DISR Core Principles for AI

|  |  |
| --- | --- |
| Core Principles for AI | |
| **Human, societal and environmental well-being:**AI systems should benefit individuals, society and the environment. | **Human-centred values:**AI systems should respect human rights, diversity, and the autonomy of individuals. |
| **Fairness:**AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups. | **Privacy protection and security:**AI systems should respect and uphold privacy rights and data protection, and ensure the security of data. |
| **Reliability and safety:**AI systems should reliably operate in accordance with their intended purpose. | **Transparency and explainability:** 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. |
| **Contestability:**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. | **Accountability:**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. |

## Australian Framework for Generative AI in Schools

|  |  |
| --- | --- |
| Core Principles for AI | |
| **Teaching and Learning**: Generative AI tools are used to support and enhance teaching and learning. | **Human and Social Wellbeing**: Generative AI tools are used to benefit all members of the school community. |
| **Transparency:** School communities understand how generative AI tools work, how they can be used, and when and how these tools are impacting them. | **Fairness:** Generative AI tools are used in ways that are accessible, fair, and respectful |
| **Accountability:** Generative AI tools are used in ways that are open to challenge and retain human agency and accountability for decisions. | **Privacy, Security and Safety:** Students and others using generative AI tools have their privacy and data protected |

# Lesson Plan

## Lesson Objectives

The purpose of this lesson is to begin a conversation about not just what AI is, but how students believe it should be used in a context that is relevant to them. This lesson focuses on the responsible and ethical use of AI. Students are required to reflect on their own perceptions of what ethical and responsible behaviour is before working together to create a set of guidelines for the future use of AI in the classroom.

## Lesson Structure and Timing

There are several activities included in the lesson plan we would recommend using the What is AI? Activity and the Class framework as bookends and selecting which additional activities you would like to do for the body of the lesson. The timings listed below are recommendations. Dependent on your class and how they engage with the key questions you may use slightly more or less time.

|  |  |
| --- | --- |
| **What is AI?:** 10 minutes | **Line of agreement activity:** 10-15 minutes |
| **Interactive Activity:** 5 minutes | **Class framework activity:** 10-15 minutes |
| **Everyday AI brainstorming activity:** 10-15 minutes | **Summation:** 5 min |
| **AI examples:** 5 minutes |  |

## Lesson Activities

**Activity 1 – What is AI? (10 minutes)**

Q1. What is AI and how does it work?

* Display slides 3-5 of the provided PPT to go over the basics of AI.
* Use the provided presenter information and key questions to facilitate discussion.

*Discussion prompts and questions can be found in the notes section of each slide.*

**Activity 2 – OPTIONAL Interactive Activity (5 minutes)**

* Display slide 6 of the provided PPT
* As a whole class or individually play Quick, Draw! Slide 6 has an embedded hyperlink that you can follow or use this [link](https://quickdraw.withgoogle.com/).
* Use the provided presenter information and key questions to facilitate discussion.

***Tip: students love this game! You may need to promise students another turn later to get them to stop playing!***

**Activity 3 – Everyday AI Brainstorm activity (10-15 minutes)**

* Display slide 7 of the provided PPT and ask students to come up with examples of AI in their everyday lives.
* Split students into pairs or small groups and provide post-it notes.
* Digitally or physically (class board/display wall/butcher’s paper) cluster students’ answers under the categories provided or in categories decided on by the class.
* Use the provided presenter information and key questions to facilitate discussion.
* Display responses in the room.

***If you are short on time this activity could be swapped to a discussion to save time***

**Activity 4 - Line of agreement (10 – 15 minutes)**

* Display slide 8 of the provided PPT
* Explain to students that the activity for today is all about making decisions.
* Create a real or imaginary line on the ground going from one side of the classroom to another.
* Explain to students that you are going to pose them questions and statements and they have to stand on the line. At one end is strongly agree at the other is strongly disagree.
* Read out the scenarios displayed on the slides one at a time.
* After each scenario ask one or two students to explain their thinking.
* Highlight differences in opinion and caveats in their answers.
* Use the provided presenter information and key questions to facilitate discussion.

**Activity 5 – AI Examples (5 minutes)**

What does it mean to use something ethically and responsibly?

* Display slide 9 of the provided PPT
* Explain that organisations, just like people, need to make choices about how they use AI and what they use it for. Their decisions can determine if AI is being used ethically and responsibly. But that making these decisions can be difficult.
* Display slide 10 of the provided PPT
* Use the provided presenter information to facilitate discussion.

**Activity 6 – Creating a framework (10-15 minutes)**

* Display slide 11 of the provided PPT
* Present students with the challenge of making their own set of guidelines. You may wish to make these general or you may wish to focus on one area of AI use. This activity can be done in pairs, groups or whole class.
* Use the provided presenter information and key questions to facilitate discussion.

**Activity 7 - Summation (5 minutes)**

* Display slide 12 of the provided PPT
* Display Framework(s)
* Remind students that ethical AI is not something that we can ‘set and forget’
* Make a plan to come back and review your framework regularly to see if it is working as intended.
* A potential extension activity would be to look through the compliance and oversight toolkit and adapt sections to suit student’s context.
* Use the provided presenter information and key questions to facilitate discussion.