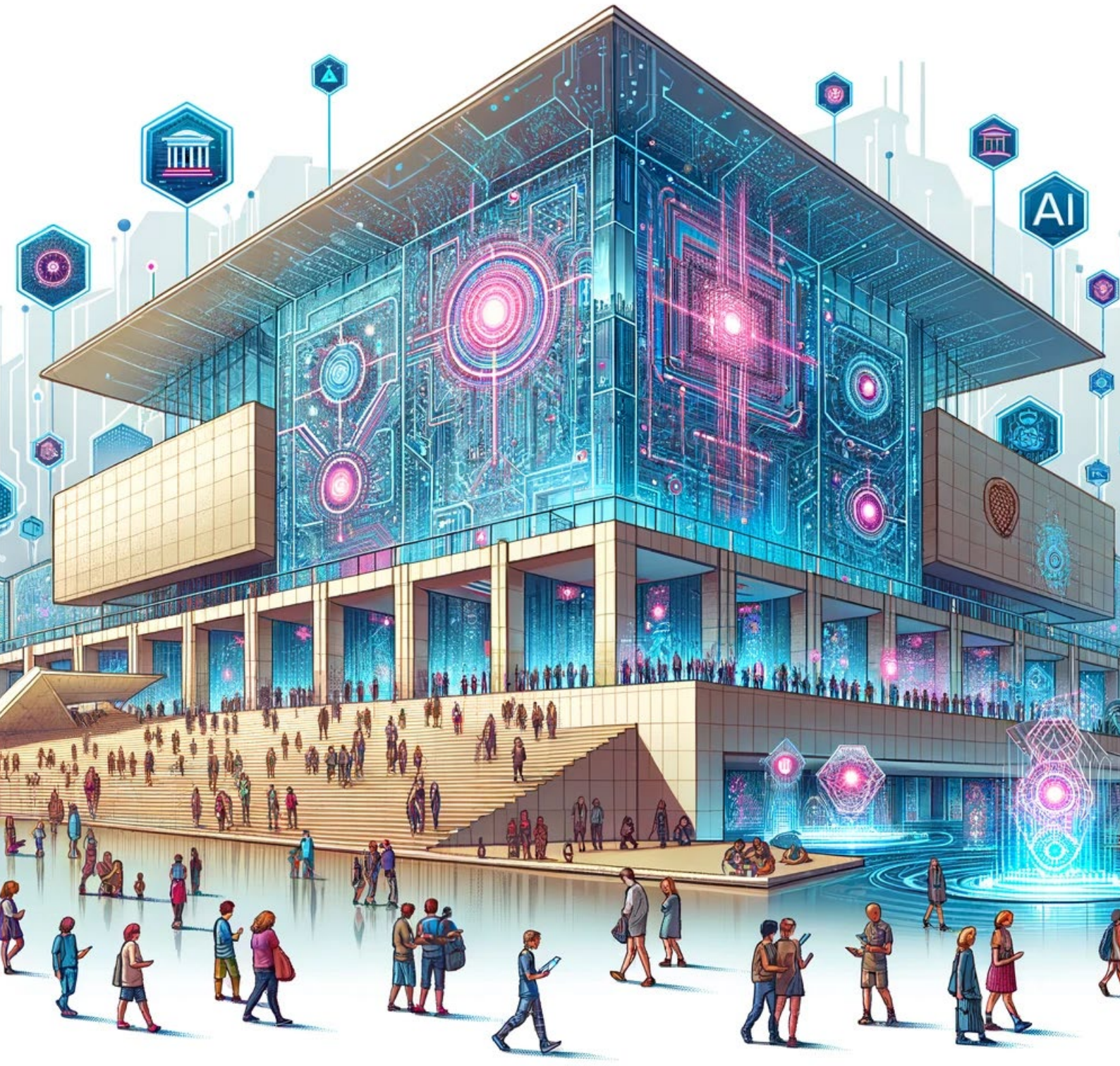

BSSS

Guide to AI in the ACT Senior Secondary System

Table of Contents

The Australian Framework for Generative AI in schools

The Australian Framework for Generative Artificial Intelligence (AI) in schools is a document that discusses best practices for the use of generative AI in Australian schools. This guide is intended to complement the introduction of the Framework in ACT Senior Secondary schools.

What is Generative AI?

Generative AI is a type of computer program that creates new content like text, images, audio, and video that resembles what humans can produce. Generative AI is 'trained' by humans and computer programs using large quantities of data to work out what the best output (a sentence, a picture, a model, etc.) in response to a prompt (a sentence or paragraph describing what the user wants) should be. When asked to produce the output, it does so.

Companies make these models available to the public by giving them a look and feel that is familiar to a user; for example, being able to type in a question in normal language and getting an answer or other output in normal language. The high-quality outputs that can be achieved by relatively unsophisticated users have implications for society as a whole.

It's very important to know that AI does not understand what it is making. It is a program that is good at recognising sophisticated (and sometimes opaque) patterns in large bodies of text, images, sound and more, and manipulating these patterns to make reasonable responses. Responses from generative AI, however, are not always correct. Biases and omissions in the training data, and the underlying mechanics of probability, means that some of the time the AI will simply be incorrect, bizarre, or unreasonable.

It's also important to know that AI is not quoting from the Internet or lifting text/images wholesale from any particular source. This makes it very difficult to reliably detect and has been the root of academic integrity concerns in education in 2022-23, particularly in the use of Large Language Models.

The responses provided by an AI model are further mediated by instructions in the program that are colloquially referred to as "guardrails"; these guardrails help determine what requests (prompts) are acceptable, and what outputs are acceptable. These guardrails are decided upon by companies and there is no legislated standard governing what is and isn't acceptable outside of the protections already present in Australian law.

What can Generative AI produce?

Please note: Generative AI is evolving rapidly, and many products offer more than one of these functions.

Output	Explanation
Text	Outputs writing in response to a prompt. Can analyse text or images to look for patterns or offer advice. Can be asked to output text in different styles and voices, and to incorporate references and quotes. Can be asked to "improve" text for the user and will do so.
Images	Outputs an image in response to a prompt. Can work in a variety of styles, add, or remove elements from an image, including expanding on an existing image. Referred to as a 'deepfake' when a person's image is realistically mapped into a situation that they have never been in.
Analysis and Code	Outputs analysis of data (e.g., graphs, qualitative analysis of words) and/or computer code. This analysis can include the analysis of a person's study or work habits to aid self-reflection, or the analysis of a piece of work and advice for improvement. This analysis may also be of large groups of data such as student data or workforce data, such as resume-ranking programs.
Design	Outputs visual design such as layout, such as automatically arranging content, choosing colour and font schemes, creating images. A common addition to slideshow or graphic design programs.
Video and Audio	Outputs a video and/or audio clip. This may be entirely AI generated, generated based on an image or series of images, or a 'deepfake' where a person's image or voice is mapped onto a real or digital clip that shows them realistically doing or saying something that has never occurred. Music can also be generated via AI.
Organisation	Emergent technology that is being woven into multimodal tools that allows an AI to offer organisation assistance such as automatically adding follow-up meetings/to-dos to a person's calendar based on what is said in a meeting, automatically write summaries or minutes of meetings, or undertaking administrative tasks such as planning and booking a travel itinerary.

How can Generative AI be used in schools?

Not all uses of AI in school contexts will be permitted by all sectors and/or schools. Please note that the Board of Senior Secondary Studies (BSSS) does not provide technology to schools and students, nor does it set policies about the acceptable use of technology in sectors and schools.

Teachers can use generative AI to:

- Generate lesson and curriculum plans
- Generate ideas for innovative activities and assessments
- Analyse curriculum materials as part of constructing an assessment
- Create exemplars, datasets, examples, arguments and other lesson-specific texts
- Visualise ideas for students
- Quickly format or transform resources
- Differentiate class materials
- Generate polished text from a broad/draft description

Students can use generative AI to:

- Get specific, personalised feedback
- Differentiate instructions or explanations, including making documents accessible for students with disability
- Format or transform assessment work (e.g., slideshows for a speech)
- Visualise and ideate (e.g., test concepts in a design class)
- Summarise articles, texts, and videos
- Change the lexical density of text for ease of understanding
- Converse with an AI to finesse understanding
- Brainstorm or discuss ideas for a specific task
- Create a scaffold for a specific task
- Create a study plan or calendar
- Generate study advice and hints

Most important strategies for students and teachers

- Ensure that all tasks have clear written instructions pertaining to the use of AI, academic integrity, and that these practices are equitably applied on a school-wide basis
- Experiment with AI as a user so that they are familiar with its strengths and weaknesses in their intended area of use
- Only use AI when they believe it will have a positive impact on learning

What will students and teachers need to do in response to the Framework?

There are a range of responsibilities implied in the Framework. Some of the implied responsibilities fall on developers, governments, schools, and systems. Below are some of the implied responsibilities the BSSS have identified that may affect students and teachers.

Students

Learning and thinking
• Students use generative AI to enhance, not replace, their own thought and creativity • Students understand that academic integrity is about engaging with your discipline as a scholar
Ethical and social responsibility
• Students use AI in line with class, school, and system policies, including those about academic integrity • All users use AI in ethically and in non-discriminatory ways, including not presenting AI generated content as their own work, not using generative AI to generate items that may be considered part of a closed or appropriative cultural practice, and not using generative AI to generate unethical content • Students do not enter their own or others' personally identifiable data into generative AI • Students report to a teacher or trusted adult when there is an unintended impact of AI • Students comply with Australian law and acceptable use guidelines
Attitudes and Values
• Users (all school community members) do not use AI to demean, harass, harm, bully, or otherwise hurt others. All users make efforts not to use AI in ways that would disrespect human and worker's rights • Students engage with the ideas of algorithmic and systemic bias, and proactively reflect on their own perspectives • Students engage with the information available about technology and AI, and ask questions when they don't understand • Students employ data safety practices • Students maintain positive, respectful dialogue about teachers and teacher expertise • Students understand that generative AI isn't always the best tool for the job

Teachers

Many of the responsibilities canvassed in the Framework are part of teachers' current responsibilities (e.g., Australian Curriculum, AITSL standards), including:

Teaching and Learning
• Monitoring student data and behaviour for trends and acting where needed to address problems • Explicitly teaching critical thinking and creativity • Using a wide range of texts, examples, content and contexts, to expose students to a range of ideas and minimise bias • Explicitly teaching academic integrity and the attribution of sources • Explicitly teaching what knowledge, ethics and integrity look like in the discipline/s taught
Maintaining Supportive and Safe Learning Environments
• Explicitly teaching ethics, pro-social behaviours, anti-bullying, anti-discrimination • Explicitly teaching the use of technology; both as relevant to your discipline, and the basic uses and cybersafety required when engaging with technology in your classroom • Making the processes for queries and appeals clear on documentation • Checking tools and new technologies before incorporating them into lessons
Compliance with administrative and professional requirements
• Teaching students about copyright and modelling behaviours that do not breach copyright • Compliance with privacy legislation, particularly when handling sensitive data • Compliance with other Australian laws

New or changed responsibilities:

Teaching and Learning
• Explicitly teaching about the inclusive use of AI technologies and algorithmic biases in the discipline taught • Teachers explicitly teaching about cultural appropriation and open/closed cultural practices where relevant to your discipline • Explicitly teaching students to generate content ethically and respectfully • Where generative AI is used in schools (e.g., for marking work or writing reports), teachers and school leaders monitor, and quality assure the decisions

Maintaining Supportive and Safe Learning Environments
• Teachers teach students about prompting, including the kinds of prompts and information that may breach privacy • Explicitly teaching how misinformation and disinformation work online, including systemic biases. • Modelling the ethical use of AI – e.g., copyright, attribution, not using AI to harass or demean others
Compliance with administrative and professional requirements
• Clearly stating AI policies on assignments, with consistent consequences for breaches. This is likely to be realised at a school level. • Interpreting policy with knowledge of your students and community to aid students in understanding generative AI • Teachers maintain up to date knowledge about technology, including generative AI. There is an implied responsibility from schools, systems, sectors, and government to provide quality-assured information to teachers. • Schools and teachers make policy and procedure around the use of generative AI clear, with consistent consequences for inappropriate use • Teachers do not enter their own, or others' personal data into commercial generative AI

Key Resources

SmartCopying: The National Copyright Unit

The National Copyright Unit serves all schools in all sectors in Australia. They have published advice about generative AI and can also be contacted with specific questions and concerns.

The Office of the eSafety Commissioner

The Office of the eSafety Commissioner has advice specific to generative AI, and a range of lesson plans and fact sheets aimed at teachers, students, and the general public. The Office recently released a position statement on generative AI that is most informative and useful.

BSSS Quality Assessment Guidelines

These guidelines are a research-based exploration of what makes a quality assessment. They are used in the quality assurance processes for system moderation and can be used by teachers to plan and interrogate their own assessments.

BSSS Ethical Research Guidelines

These guidelines are contained in the Policy and Procedure manual and describe ethical research practices. From 2024, the use of these guidelines is compulsory in all ACT Senior Secondary courses that conduct human research.

How can AI restrict thinking or contain biases, and why does this matter?

AI can restrict thinking by:

- Producing the “most likely” response, which will contain biases towards dominant cultures and ideas in the training data. At the moment, the majority of generative AI available to users in Australia is biased toward American data.
- Producing a response restricted by the ‘guardrails’ or requiring users to work around the guardrails. For example, algorithmic content checking on social media platforms has led users to construct euphemisms dubbed “algospeak” such as “unalive” instead of “dead”, and creative rhyming slang/emojis for words and topics likely to trigger an algorithmic punishment
- Producing a glib, easy response for a user, rather than requiring a user to think critically or tackle a difficult task
- Produce a polished response that negates the need for a user to learn skills in a domain (e.g., sentence construction, punctuation)

AI can be biased by:

- Biased training data leads to the reproduction of those biases
- Incorrect or falsified training data (often called ‘dirty’ data) may contain unfair representations of groups or ideas
- Reflecting and amplifying unspoken biases and patterns that are part of structural social inequality
- Choices made by the company programming the AI; for example, restricting discussion about ‘culture war’ topics

Human society, communication, technology, and culture inevitably contain biases. Some of these biases are deliberate and serve an important social function. Some are historical or accidental. Some are the result of insufficient diversity in the data.

Impacts on Students

- Students may not be exposed to a range of experiences, ideas, and ways of being
- Students may have their own identities invalidated by biases in training data (e.g., if all ‘nurses’ generated are women, and all ‘engineers’ are men)
- Students may be tempted to avoid tasks that fall into the Zone of Proximal Development (i.e., challenging tasks that a student can complete with assistance) by asking AI to generate for them
- Students may be tempted to avoid critical and creative thinking by assigning these tasks to AI, thereby not developing the skills and schemata to support their future growth

Impacts on Teachers

- Decisions made by biased or faulty algorithms may unfairly reward or punish, or represent students' learning incorrectly
- Biases may mean that a topic cannot be covered fully using AI, or that AI generated lesson materials may follow a particular pedagogy or way of working that is not strongly supported by research

Strategies to address bias and critical thinking

- Explain why some tasks need to be completed without AI
- Foreground processes such as the creative cycle
- Model critical and creative thinking
- Ask questions, for example:
    - Whose story is told in this data? Are the people or ideas represented familiar to our school and students, or different? In what ways are they different?
    - What are my biases? Do I view this material in a particular way because of my age, race, culture, nationality, political view, physicality, gender expression or religion?
    - Where is Australia represented in this data? Which Australian experiences are represented?
    - Are any of these depictions, stereotypes?
    - What is not depicted?
- If students are generating responses from one point of view, teachers may like to ask them to generate from an opposing point of view, or to argue the opposite case
- Use mnemonics and frameworks like the CRAAP test to evaluate sources as a group

Our implicit biases can be hard to spot, and hard to address. Explicit biases are often dealt with openly – for example, programs addressing biases against people of particular genders being associated with particular careers. An example of implicit bias might be describing someone as "suffering from a disability" without questioning whether the person *suffers*, or what defaulting to assuming people who have disability are suffering says about social and personal attitudes towards disability.

Resources

The Human Rights Commission: https://humanrights.gov.au/education

SBS Cultural Atlas: https://culturalatlas.sbs.com.au/

Algorithmic bias: video: https://www.pbs.org/video/algorithmic-bias-and-fairness-18-4gxvyl/

The Crochet Mushroom: analysis of what a simple AI task tells us about bias: https://www.bsss.act.edu.au/__data/assets/pdf_file/0008/569681/Special_Interest_Paper_October_2023_AI_-_Students_Parents_and_Carers.pdf (p.10)

The CRAAP test: https://library.csuchico.edu/sites/default/files/craap-test.pdf

Incorporating AI into learning design

Depending on the school or sector, AI may be incorporated into learning design. As with all educational technology, AI should be used for a reason, and with a reasonable belief, supported by evidence as it becomes available, that the technology will have a positive impact on student learning outcomes.

The best way to learn about AI and its possibilities is to use AI. Ethan Mollick describes AI as having a "jagged frontier"; there are some tasks that can be easily accomplished or improved by AI, and others that can't be. Everyone is still learning which tasks are which, and the best way to learn in this case is by testing.

When prompting an AI, teachers and students can use the following mnemonic – "SPRITE".

Specificity: tell the tool as clearly as possible what you would like it to create

Parameters: set any exceptions or inclusions, such as "do not include true/false questions in your output"

Role: tell the tool what kind of role you'd like it to take; for example, a friendly teacher, a fictional character, an experienced coach

Iteration: ask the tool questions and get it to refine its output. For example, "can you expand on that paragraph?" or "can you rewrite that text with a greater focus on controversy in the research?"

Text Type: specify the type of text or output you are seeking, e.g., a rubric, a paragraph, a chart, or table

Exemplars: use exemplars to guide the tool; for example, asking to write in the style of a particular literary movement or create in the style of a particular art movement

A sprite is a mythical spirit in European folklore; something like a fairy. An AI can be a helpful sprite, but it might also play tricks; users should keep your wits about them and check anything that doesn't seem correct.

A note on commercial tools for education and learning design

Many software providers and companies are making use of AI to aid teachers and students, including workload reduction and assessment of student work. A school or sector is likely to make choices about which of these will be available for use by teachers and students.

Be mindful that there are also a number of small startups that offering AI tools for education. Some of these are ethical and well-designed, whereas others are trying to make a quick dollar. Unfortunately, these can be hard to distinguish. It is important to check with your school or sector's IT team for advice and approvals for specific tools.

Resources

Ethan Mollick's education-focussed work, including:

- Assigning AI: Seven Ways of Using AI in Class: https://www.oneusefulthing.org/p/assigning-ai-seven-ways-of-using
- How to use AI to Do Stuff: An Opinionated Guide: https://www.oneusefulthing.org/p/how-to-use-ai-to-do-stuff-an-opinionated

OpenAI's advice for getting started with AI: https://help.openai.com/en/articles/8313929-how-can-educators-get-started-with-chatgpt

Khan Academy: AI for Education: https://www.khanacademy.org/college-careers-more/ai-for-education

ACARA Guidance on Generative AI in the Australian Curriculum: https://v9.australiancurriculum.edu.au/teacher-resources/understand-this-curriculum-connection/artificial-intelligence/

AI and Academic Integrity

A key question underlying academic integrity and AI is:

Is the use of AI going to affect the validity of the teacher's measurement of the construct being assessed?

Or, more plainly: what knowledge and skills are being assessed and will the use of AI interfere with the student's learning and their ability to demonstrate their learning in a way that is fair – both to the student and to their classmates.

The BSSS have collated our 2023 published work on academic integrity into three papers for students, teachers, and parent and carers. These are linked below.

Strategies and advice for students

- The unattributed, disallowed, or dishonest use of AI, or the use of AI in any other way to substitute for your own work and thinking, is a breach of academic integrity, and will be dealt with using the BSSS Academic Integrity Policy and Procedures.
- Use sources not because you 'have to', but as part of developing your understanding of the subject you are studying, and as a way to place the work you are doing in the context of your discipline.
- Use AI ethically and honestly. If AI isn't permitted in the task, don't use it. If it is, record your prompts and hand in what your teacher asks you to that shows and supports your use of AI.
- Don't get tied up arguing semantic definitions of plagiarism. The BSSS counts the use of AI as a breach of academic integrity, unless used in a way that is explicitly permitted in the task instructions.
- Don't take the word of businesses that are promoting a product as to what constitutes academic integrity. For example, a word processing AI that completely rewrites your sentences is most likely to be counted as a breach of academic integrity, particularly in subjects where the way you write is a big part of what is being assessed (e.g., English).
- Remember that your teacher is interested in measuring what you have learned, and how well you have learned it.
- Recognise that it's normal for teachers to query work that shows differences to your normal work. This might not be because of AI. You will get a chance to explain what happened.
- Keep your notes and drafts as you create take-home tasks and be prepared to show them to your teacher if needed.

Strategies and advice for Teachers

- Use paradigms such as Webb's *Depth of Knowledge* to articulate the kinds of thinking underlying the construct you're assessing; this may be helpful in deciding the breadth of the use of AI that is acceptable in any given task.
- Use techniques of disciplinary literacy to frame referencing's positive impact on a student's work; avoid simply framing referencing as a 'tick-box' or compliance issue.
- Do not have tasks in the same markbook, in the same column, that have different levels of permissiveness regarding the use of AI.
- Make clear statements in any given assessment about what is acceptable use of AI
- Assess across a range of modes where possible.

Strategies and advice for Parents and Carers

- Parent and carer attitudes toward academic integrity are hugely important; helping your student to understand that using AI to substitute for their knowledge, understanding or skills puts their learning at risk is key information that you can impart to your child.
- Queries about the use of AI are normal and expected. If your student is worried, it may be helpful to go through the student guide to Academic Integrity with them.
- Using AI where it is not permitted in a task will be treated as a breach of Academic Integrity.
- Fostering a growth mindset can be helpful; many of the reasons students may plagiarise are linked to fears of failure and fears of short-term consequences, without considering the skills and knowledge that a task promotes.

Resources

BSSS AI and Academic Integrity for Teachers

BSSS AI and Academic Integrity for Students

BSSS AI and Academic Integrity for Parents and Carers

Quality Assessment Guidelines:
https://www.bsss.act.edu.au/information_for_teachers/quality_assessment_guidelines

MLA advice on citing AI: https://style.mla.org/citing-generative-ai/

APA advice on citing AI: https://apastyle.apa.org/blog/how-to-cite-chatgpt

Webb discussing the Depth of Knowledge framework:
https://www.youtube.com/watch?v=qFXU6_TYIjc

Monash University: Acknowledging the use of generative AI:
https://www.monash.edu/student-academic-success/build-digital-capabilities/create-online/acknowledging-the-use-of-generative-artificial-intelligence

Monash University: Citing and referencing generative AI:
https://guides.lib.monash.edu/citing-referencing/generative-ai

Teaching and modelling ethical behaviour

The explicit teaching of ethical understanding and personal and social capacity is part of all ACT BSSS Senior Secondary courses. The rise of problematic online behaviour, and the ability of tools to create sophisticated texts for unsophisticated users, means that schools will necessarily be part of society's defence against misinformation and disinformation.

Teachers are undoubtedly already aware of the influence of algorithmic echo chambers; where the content that people encounter on the internet is influenced by social media algorithms showing content biased only toward a person's existing beliefs. While the intent of these algorithms is ostensibly reasonable – showing users more of the content they like – the impact can be that a person ends up in what is called a "filter bubble", unaware that they are not being shown information that contradicts their point of view. This is also possible with AI.

Examples of things that people can do with AI that are morally, ethically, or legally problematic:

- Create convincing misinformation; for example, deepfakes or AI generated images of people in situations they have never been in. This could be to discredit or harm (e.g., image-based abuse), or to lionise and celebrate (e.g., AI generated images of politicians helping people)
- Create misinformation with the purpose of harming others (e.g., scams using an AI generated voice of a real person)
- Create misinformation that is used to cyberbully or harass others
- Unintentionally reinforcing harmful biases (e.g., AI that discriminated by gender when sorting resumes) or enabling users to create content that reinforces harmful biases
- "Jailbreaking", "prompt injecting" or other behaviour designed to circumvent guardrails and create illegal or unethical content
- Using tools for purposes that are not intended; for example, using an AI designed to remove unwanted people from an image to cyberbully others

These uses of AI can be from individuals, special interest groups, or from governments seeking to promote a point of view, disrupt democratic processes, or sway public opinion.

Strategies to promote ethical behaviours

- Studying texts and topics that allow you to address echo chambers and antisocial behaviours, e.g., ethics, the paradox of tolerance, texts themed around 'resilience', texts that show strong role models of all genders and nationalities, units that explore propaganda
- Examining how misinformation and disinformation spreads in your discipline, and how knowledge is constructed
- Examining ethical behaviour in your discipline, e.g., ethical research, experimentation, representation of events and people, ethical dilemmas, censorship
- In pastoral care units, explicitly addressing respectful relationships and ways of being, including online

- Ensure that behaviour and inclusion policies within the school address online bullying and harassment
- Link all school community members to factsheets from the eSafety commissioner

Resources

Office of the eSafety Commissioner resources for educators:
https://www.esafety.gov.au/educators

Office of the eSafety Commissioner resources for young people:
https://www.esafety.gov.au/young-people

BSSS Ethical Guidelines for Assessment:

Victorian State Government: Level 11-12 Resilience, rights, and respectful relationships – resource with many activities that students can work through as part of a program of learning or pastoral care program:
https://fuse.education.vic.gov.au/Resource/LandingPage?ObjectId=05140241-6934-4130-97a7-0446282decaf

Student Wellbeing Hub – resources to support the teaching of respectful relationships:
https://studentwellbeinghub.edu.au/

Copyright, culture, and AI

AI's copyright status is vexed, and generally highly dependent on the model, the data it was trained with, and the practices of the programmers. Many companies have not released the full extent of the training data used, including whether that data included copyrighted works. Teachers and students are strongly advised to make use of Smartcopying's services if **any** queries about copyright come up. This unit is funded to support schools in matters of copyright.

AI is not directly copying text or images from extant sources. However, it is possible to train an AI on the products of one source (e.g., a specific author or artist's work) and then generate works that are exceptionally similar. Once a source is part of training data, it cannot be removed.

Equally, authors, artists and other creatives have concerns that the use of AI generated text and images will displace workers in these industries, whilst training on the products of those workers.

Students can be explicitly taught about the practices and norms of their discipline; what is seen as homage, and what is copying. For example:

- Appropriation and ekphrasis in literature, art, and design
- Software copyright and practices in coding
- Design rights and intellectual property
- Sporting Teams, national flags, the Olympics and copyright in PE or Sports
- The evolution of copyright of the Aboriginal flag
- Multiple/simultaneous discovery in the Sciences
- The impact of the rush to publish in the sciences

AI can also be used to generate cultural artefacts. The Framework suggests that teachers and students should be careful when generating items that may be specific to a culture or group, and that the product of AI is not used as representative of an authentic cultural product.

Closed cultural practices are those practices that are only practiced by people who are born into, or accepted into, a community. These are often religious or ceremonial.

Cultural appropriation is the inappropriate use of elements of one culture, usually a minority culture, by members of another culture. Often this is in the context of a majority culture profiting from the cultural products of a minority culture.

For example, there are some Indigenous stories that cannot be shared with people who are not Indigenous and/or have not been initiated into a tradition. Equally, most Australians would find the idea of a non-Indigenous person painting "authentic Aboriginal" art to be anathema; however, a quick search on Pinterest will generally demonstrate a range of North American based, highly stereotypical, and appropriative lessons aimed at young students where they do just that.

Cultural sharing is not appropriative, and it is a huge part of everyday life. Students are still learning about the world, and need to have the conversation about what is and isn't appropriate in your subject, and why. For example:

- Visual arts and design courses might discuss the difference between common and sacred motifs, and where the line is drawn
- Art, music, dance, literature, and drama classes may look at the way different cultures are not monolithic, e.g., by identifying an instrument or style from its culture or country of origin, rather than simply using catch-all terms like "Asian" or "African" dance.
- Social sciences and other humanities may choose to use authentic texts from within cultures; for example, Bawaka Country is listed as lead author on a group of studies into the relationship between people, culture, and country.

Resources

Smartcopying: Using Generative AI Platforms in Schools: https://smartcopying.edu.au/using-generative-ai-platforms-in-schools/

Cultural Appropriation Factsheet: https://au.reachout.com/articles/why-cultural-appropriation-isnt-cool

Publications on Researchgate for which Bawaka Country is listed as an author: https://www.researchgate.net/scientific-contributions/Bawaka-Country-2014774093

Centre for Research on Foundation Models: https://crfm.stanford.edu/fmti/

Privacy and Security

Teachers and schools handle sensitive information, and as such, have responsibilities under the Privacy Act. Many people in the ACT have experienced digital privacy breaches in the recent past and the flow-on effects from such breaches such as increased and targeted spam or identity impersonation and theft.

Areas of privacy risk in schools

- Personal or sensitive data held by a school must not be put into generative AI without informed consent.
- Personal or sensitive data held by a school must only be processed using secure servers and storage, preferably located in Australia.
- Some tools may require users to sign in using personal information such as email addresses or legal names. In the event of a data breach (unwanted people having access to data), this risks their inputs being linked to a person's identity.
- Models that feedback prompts into their data sets can unwittingly absorb personal or sensitive data. In this instance, the model is most likely unable to remove this data once discovered.
- Generative AI can be used to breach privacy; for example, image-based search engines that can be used to find all images of a child online, or triangulate sources to expose a pseudonymous user's personal details ("doxing"). These services sometimes request money to 'remove' a person from their results or expose people to risk through publicising their personal information.
- Unintended data breaches may occur; for example, in 2023, some users were briefly able to see the titles of other users' chats in ChatGPT
- Students may upload personal information (their own or others') without understanding the risks.

Each sector and school will have processes and forms for the use of technology. It is important that students are regularly reminded of risks. Everyone must be mindful, too, that there are emergent technologies such as quantum computing that may make gathering data about individuals even easier.

Strategies to promote Privacy and Security

- Only use tools for work that are permitted for use by your workplace or school
- If you use an unapproved tool, consider the impact of a data breach on you or the person whose data you are using: for example, where is the AI provider's data held? Could the data cause embarrassment? Could the data be used to find a person – e.g., by a non-custodial parent? Does the data expose someone's medical or cultural records? What are the legal consequences for you of a data breach (e.g., are you breaching privacy law, AITSL standards)?
- When undertaking human research, using the ethical assessment guidelines to emphasise the role of a researcher in ethically using personal data

- Deidentify data before using AI tools, and only use the minimum amount of data to accomplish a task
- Teach students about the privacy risks that can come with online interaction, and discuss the over-collection of data where appropriate

Resources

eSafety Commissioner: Generative AI Position Statement: https://www.esafety.gov.au/industry/tech-trends-and-challenges/generative-ai

eSafety Commissioner: Educators: https://www.esafety.gov.au/educators

Office of the Information Commissioner: Global expectations of social media platforms and other sites to safeguard against unlawful data scraping: https://www.oaic.gov.au/newsroom/global-expectations-of-social-media-platforms-and-other-sites-to-safeguard-against-unlawful-data-scraping

ACARA Digital Literacy Capability: https://v9.australiancurriculum.edu.au/f-10-curriculum.html/general-capabilities/digital-literacy?element=0&sub-element=0

Australian Capital Territory, Canberra 2023

Cover image generated by OBSSS staff using Dall-E III