

Responsible by Design

GenAI & Ethics from
The Open University Learning Design Team



How to use the framework

This tool has been designed to help you find opportunities to explore the ethical implications of AI in your learning materials.

Go through each prompt and choose the response most appropriate to the current use of GenAI in your context. This overview highlights areas where you can address ethical implications of GenAI, fostering more responsible learning material design.

In the next column, add the solution to each particular question. You may find some inspiration in the 'Solutions Bank' in the final pages of this document.

See an example below, filled with blue text.

Pillar	Prompt	Check	Possible actions
Bias	Are students supported to consider the bias (gender, ethnicity, culture) of the GenAI product?	<input checked="" type="checkbox"/>	<i>Although we've identified, bias in GenAI, providing contextual examples would enhance understanding</i>
	Is the bias problem embedded in the task itself, teaching students to apply critical thinking?	<input type="checkbox"/>	<i>We could set up a reflection activity on possible biases in the GenAI output to encourage specific focus.</i>

This is the pillar that the prompt relates to



This is the prompt for you to answer

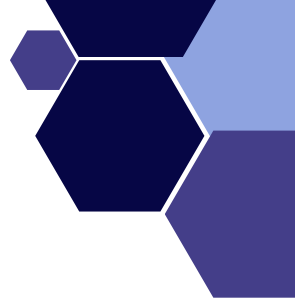


Have you met this prompt or is there room to improve?



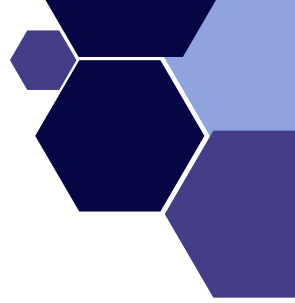
What could you do to strengthen the ethical considerations raised by this prompt?





Bias & Sustainability

Pillar	Prompt	Check	Possible actions
Encoded Bias	Are students supported to consider the bias (gender, ethnicity, culture) of the GenAI product?		
	Is the bias problem embedded in the task itself, teaching students to apply critical thinking?		
Environmental impact	Are students supported to consider the AI's impact on the carbon footprint?		
	Are students guided on crafting effective AI prompts to minimise unnecessary prompting?		



Exploitation & digital divide

Pillar	Prompt	Check	Possible actions
Exploitation	Exploitation of resources: Is the potential exploitation of resources (such as copyright of artists/writers) made clear and transparent to students?		
	Exploitation of people: Is the potential exploitation of people (such as those inputting and editing sensitive materials) made clear and transparent to students?		
Digital Divide	Will the quality of the output be affected by access to various types of AI?		
	Will the activity be affected by the students' digital skills levels?		

Opting out

As GenAI is a new and rapidly changing technology with inherent ethical issues, we should still provide students with the opportunity to 'opt-out' of GenAI-driven tasks on ethical grounds, especially in assessment.

This part of the framework allows us to consider those students who may not wish to participate in GenAI activities, and who may, therefore, wish to have an alternative. Pedagogical reasons may justify excluding GenAI from the learning experience.

Pillar	Prompt	Check	Possible actions
Opting out	Have you offered an alternative to using AI and given the right to opt out?		

Solutions bank



Bias:

Consideration: Are students supported to consider the bias (gender, ethnicity, culture) of the GenAI product?

Possible solution: Depending on the task, you may wish to draw the students' attention to the bias that is likely to occur with AI which could be in any form. You might remind them that tools like ChatGPT often generate outputs influenced by 'WEIRD' datasets (Western, Educated, Industrialised, Rich, Democratic perspectives).

Bias can also manifest in thought processes, language, and proposed solutions. All of this can be embedded in the activity, asking the student if they notice any bias towards a solution, language, style, or a type of person, etc.

Consideration: Is the bias problem embedded in the task itself, teaching students to apply critical thinking?

Possible solution: For some tasks, embedding the skill of interrogating the bias may be helpful, asking students which stereotypes they're seeing or whose voice they feel the response is written in.

Environmental:

Consideration: Are students supported to consider the AI's impact on the carbon footprint?

Possible solution: Consider guiding students to generate a specific number of answers and remind them of the carbon footprint each AI answer generates, as well as issues such as the storage of data and training the AI.

Consideration: Are students guided on crafting effective AI prompts to minimise unnecessary prompting?

Possible solution: You could signpost to examples of effective prompting in the context of the learning. You could provide subject-specific prompt writing resources.

Solutions bank



Exploitation:

Consideration: Exploitation of resources: Is the potential exploitation of resources (such as copyright of artists/writers) made clear and transparent to students?

Possible solution: As with bias, especially in generating images or sound, it is worth highlighting to students the source material that AI is trained on. Depending on the task, this could take the form of asking students, 'Do you spot anything that is familiar from elsewhere? What does it tell you about the content that AI has been trained on?'

Consideration: Exploitation of people: Is the potential exploitation of people (such as those inputting and editing sensitive materials) made clear and transparent to students?

Possible solution: You could ask students whether they think any sensitive content has been filtered out by the AI in order to produce the required answer, and then highlighting that this work is done by people in poor countries who are poorly paid and subject to violent or disturbing material without proper support.

Digital Divide:

Consideration: Will the quality of the output be affected by access to various types of AI?

Possible solution: Consider which AI tool you are proposing that the students use and whether a version of it is behind a paywall. If the students can access the paid version, consider whether you want them to compare the output or to restrict them to a particular version of the tool.

Consideration: Will the activity be affected by the students' digital skills levels?

Possible solution: Provide clear instructions, assuming users are complete beginners.