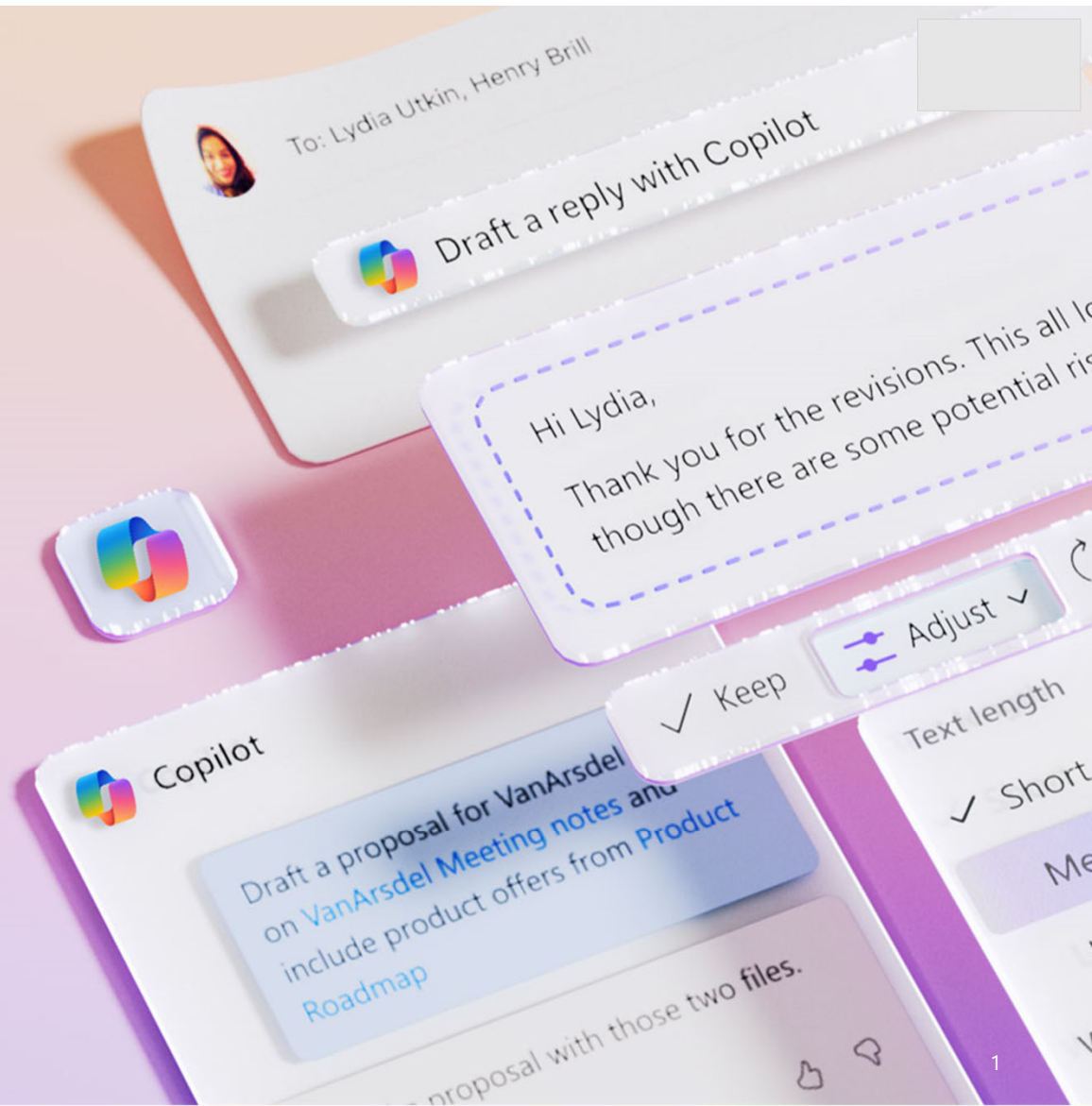


Copilot for Microsoft 365

Responsible Use and Ethics

Benedict Yeo & Samantha Seah

SMU Libraries



What is AI Ethics

Ethics is a set of moral principles which help us discern between right and wrong. AI ethics is a multidisciplinary field that studies how to optimize **AI's beneficial impact** while reducing risks and adverse outcomes.

<https://www.ibm.com/topics/ai-ethics>

Ethic frameworks

Microsoft

[Responsible AI Standard](#)

Fairness

- How might an AI system allocate opportunities, resources, or information in ways that are **fair [non-discriminatory]** to the humans who use it?

Reliability and Safety

Privacy and Security

Inclusiveness

Transparency

Accountability

- How can we create oversight so that **humans** can be **accountable** and in control?

For Users

Applies when AI outputs do the following :

- allocate resources and opportunities [e.g. scholarships]
- include descriptions or representations of people, cultures, or society [e.g. marketing material]

Ensure human accountability.

Use human agency to evaluate if the tool or output is

- Factual
- Fair
- Suitable

When using AI, keep an eye out for...

Bias

AI systems may exhibit unfair prejudice in favor of or against one thing, group, or person compared to another. For example, an AI system used for hiring that disproportionately recommends candidates from a certain demographic over others due to biased training data.

Misrepresentation

This occurs when AI inaccurately portrays an event, individual, social groups, or data, leading to misinformation. An example is an AI-generated news article that presents incorrect statistics about a public health issue, potentially leading to public misunderstanding.

Hallucination

Hallucination refers to the generation of information or data that is not grounded in reality. An example would be a language model creating fictitious legal cases based on elements of various true cases.

Copyright

This involves the unauthorized use of copyrighted material by AI systems. An example could be an AI that uses or reproduces copyrighted music in its generated content without permission from the copyright holders.

Plagiarism

This is when an AI system copies or closely imitates the language, thoughts, or expressions of another author's work without proper acknowledgment. For instance, an AI that generates an essay by piecing together content from various online sources without citation.

Privacy Issues

This concern arises when AI systems collect, store, or use personal data without consent or in ways that violate individuals' privacy rights. An example is an AI-powered surveillance system that analyzes and stores biometric data without the knowledge or consent of the individuals being monitored.

S.U.R.E method

Source

- If the generated content has a reference, check the reference. Does it look relevant to your prompt?
- If no reference is given, try to find trusted sources on the topic

Understand

- Does the generated content express fact or opinion? Is it vague? Is it biased?

Research

- Depending on what you're looking for, know where you can search to verify information e.g., for organizational knowledge search Sharepoint
- If you're referencing new findings in research, search the topic in the library's search

Evaluate

- Exercise some judgement before you share information with others or use it in your work.

Let's look at some examples to practice S.U.R.E

Example of generated output in Teams

Research

Pick the appropriate app to use Copilot

What is Covidence used for

Covidence is a web-based software platform that streamlines the production of systematic reviews, including Cochrane Reviews. It is designed to make the systematic review process faster, easier, and more enjoyable. Covidence supports key steps in the review production process, including citation screening, full-text review, risk of bias assessment, and data extraction. It is used by researchers, students, and information specialists to conduct systematic reviews in various fields, including healthcare, public health, and social sciences. ¹

Understand

Fact or opinion?

Source

Since Copilot provided a reference, check it!

Copy

1 reference



3 of 30 responses

AI-generated

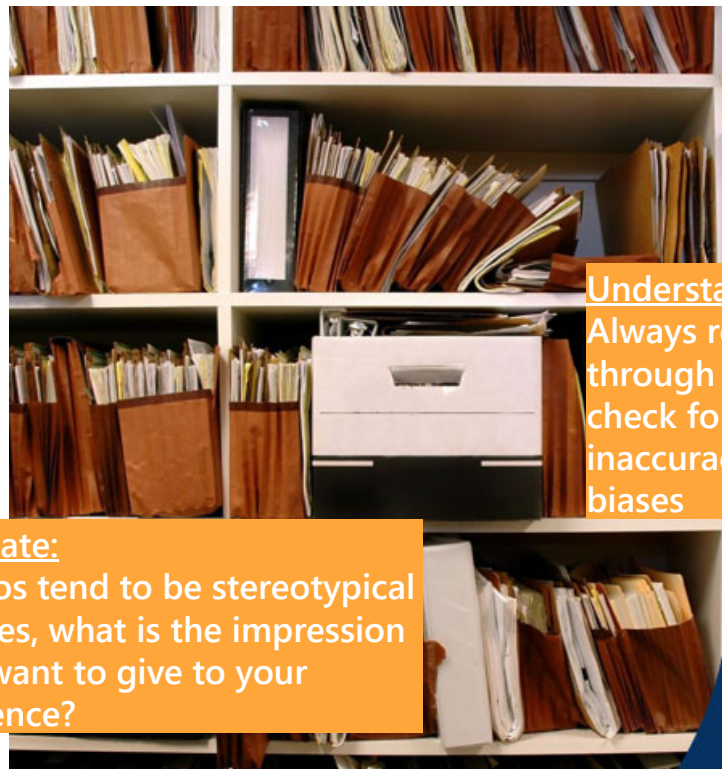
¹ Presentation106
Restricted
PowerPoint [redacted] modified on 8/16/23
"Covidence [redacted] SPSS, SAS, Stata [redacted] Nvivo [redacted] ? ..."

Evaluate

Information may be privileged, be careful what you share

Example of generated slides for a briefing

Prompt: Generate a presentation on library fines and fees waiver procedures for use in library orientation to freshmen. Include pictures



How to handle fines and fees

- Check your library account regularly for updates on fines and fees.
- Pay the fines and fees online or at the library to avoid holds on academic records and restricted borrowing privileges.
- Appeal fines and fees if you can provide a valid reason for late returns or damages.
- Donate food or money to waive or reduce fines.

Source:
Consider giving Copilot a reference file to work off

Understand:
Always read through to check for inaccuracies or biases

Evaluate:
Photos tend to be stereotypical images, what is the impression you want to give to your audience?

Research:
Checking policy or asking the colleague-in-charge

Example of marketing content

This example requests Copilot to create slides on for a Purple Parade event.

What responsible steps might you take to evaluate the content?

Source

Is there a reference file or did you provide sufficient details for Copilot?
Cite your sources, if any is used.

Understand

Read to understand the output

Research

Research to check any facts, terms used, and images generated to see if they are acceptable

Join the Purple Parade

- The Purple Parade is open to everyone
- Visitors can show their support for disability rights and the inclusion of persons with disabilities
- Participate in games and activities, enjoy food from social enterprises, and celebrate the abilities of persons with disabilities

Evaluate

Are the outputs (e.g this image) appropriate or relevant?



Recommended Practices

1. Always evaluate your output for any inaccuracies and biases

- You may use a framework like SURE
- Check with a person, such as a colleague or expert in the field

2. Acknowledge creators, sources or tools

- Even if you have express permission to use content, including AI output, in your work, make it your practice to acknowledge when you use it

3. Keep up to date with software updates

- Be aware of updates to training data and understand current limitations of the tool
 - This may also help fix previous known issues/bugs/limitations
-