AI, Image and Text Generation, and Information Bias

Initial Considerations

- Is the use of an AI application or technology necessary? What is the purpose or goal for using AI and can the information or task be completed in another modality?
  - For example, if using an AI image generator to develop an image of a CEO, one should consider whether there is another more reliable, representative, and inclusive source for these images other than an AI generator.

- There is real world impact to the use of AI generated text and images. Images and text posted in public forums and documents can be used by AI generators in the future, recycling information and biases in a continual loop.

- Individuals should make serious ethical and technical considerations when creating and using texts or images generated using AI. At least, the image and/or text should be cited and identified as machine-generated with the platform, date, and prompt(s) used.

Considering whether a generator is appropriate to a task is not just about input/output, but also long-term use of the output itself.

Bias in AI image and text generators:

- AI models are trained on large datasets of human-generated data, which can reflect societal biases. For example, an AI image generator trained on a dataset of images of CEOs may be more likely to generate images of white men than women or people of color.

- AI models can be biased in their outputs even if the training data is not. This is because AI models learn to identify patterns in the data, and these patterns can reflect implicit biases of users even if they are not explicitly stated. For example, an AI text generator may be more likely to generate text that is sexist or racist, even if the training data does not contain explicitly sexist or racist language.

- Bias in AI image and text generators can have a real-world impact. For example, if an AI image generator is used to create images for a news article, the biased images can reinforce existing stereotypes and prejudices.

Suggestions to critically evaluate AI image, text generators, and overcome bias:

- Be aware of the potential for bias in AI image and text generators. When using these tools, it is important to be critical of the outputs and to consider the potential for bias.

- Much of the training data and algorithms is not publicly available, so recognizing and critically assessing the availability of training data and training algorithms is just as important as the attempt to consider its nature. Individuals should use AI programs and tools that are transparent with their data and training processes.
  - If available, examine the training data for the AI model. If possible, try to find out what data the AI model was trained on. This can help you to identify potential sources of bias.

- Be mindful of the context in which you are using the AI generator. Consider the potential impact of the generated outputs and how they could be used to reinforce existing biases.

- Use multiple sources of information. Image and text generators are not reliable sources of information in and of themselves. Anything generated by these applications should be verified by the user using reliable, verifiable, and valid information resources.

- When using AI image generators, be mindful of the stereotypes that are commonly associated with the images that you are generating. For example, if you are generating an image of a CEO, be aware of the stereotype that CEOs are typically white men.

- When using AI text generators, be mindful of the language that is used in the generated text. For example, if you are generating text about a particular group of people, be aware of the potential for the text to be biased or offensive.

- If you are using AI image or text generators for academic purposes, be sure to cite the tools that you used. This will help to ensure that your work is transparent and that others can understand the potential for bias in your findings.

The creation of this document utilized Google Bard and ChatGPT as well as research literature on AI, biases, and stereotyping, and feedback from professionals in digital initiatives and information literacy.

Aoun, Joseph. Robot-Proof: Higher Education in the Age of Artificial Intelligence. The MIT Press, 2017.

Bard. (2023, Aug. 30). "Act like a college professor working with students on information literacy. Give a list of talking points on bias in AI image and text generators. Include suggests on how students can critically evaluate these products and overcome them."

Bard. (2023, Sept. 1). "Provide a list of ten key points that students need to know when critically evaluating AI generators for their bias, stereotypes, and other social inequalities."

Chat-GPT-4. (2023, Aug. 30). "Act like a college professor working with students on information literacy. Give a list of talking points on bias in AI image and text generators. Include suggests on how students can critically evaluate these products and overcome them."

Gruetzemacher, Ross, and Jess Whittlestone. "The Transformative Potential of Artificial Intelligence." Journal of Higher Education Policy and Management, 2021, https://doi.org/https://doi.org/10.1016/j.futures.2021.102884.

Miller, Matt. AI for Educators: Learning Strategies, Teacher Efficiencies, and a Vision for an Artificial Intelligence Future. Dave Burgess Consulting, Incorporated, 2023.

Munn, Luke, et al. Unmaking AI Imagemaking: A Methodological Toolkit for Critical Investigation. arXiv:2307.09753, arXiv, 19 July 2023, https://doi.org/10.48550/arXiv.2307.09753.

Open AI. Teaching with AI. https://openai.com/blog/teaching-with-ai?utm_source=tldrai. Accessed 5 Sept. 2023.

Ten Considerations to Combat Bias when using AI Image and Text Generators

1. AI generators are trained on data that reflects societal biases. This means that AI generators are more likely to generate outputs that are biased against certain groups of people.

2. AI generators can be biased even if the training data is not. This is because AI generators learn to identify patterns in the data, and these patterns can reflect biases even if they are not explicitly stated.

3. Bias in AI generators can have a real-world impact. For example, if an AI generator is used to create images or text that is used to inform people about the world, the biased outputs can reinforce existing stereotypes and prejudices.

4. Individuals should be aware of the potential for bias in AI generators and be critical of the outputs. No one should simply accept the outputs of AI generators as accurate or unbiased.

5. Individuals should examine the training data for the AI generator. If possible, Individuals should try to find out what data the AI generator was trained on. This can help individuals to identify potential sources of bias.

6. Individuals should be mindful of the context in which they are using the AI generator. Consider the potential impact of the generated outputs and how they could be used to reinforce existing biases.

7. Individuals should use multiple sources of information. Don't rely solely on AI generators for information. Instead, use a variety of sources, including traditional sources such as books and articles, as well as other AI tools.

8. Individuals should be mindful of the stereotypes that are commonly associated with the outputs of AI generators. For example, if an AI generator is generating images of CEOs, individuals should be aware of the stereotype that CEOs are typically white men.

9. Individuals should be mindful of the language that is used in the outputs of AI generators. For example, if an AI generator is generating text about a particular group of people, individuals should be aware of the potential for the text to be biased or offensive.

10. Individuals should cite the AI generators that they use in their work. This will help to ensure that their work is transparent and that others can understand the potential for bias in their findings.