

PHSA Researcher Guidance: Use of Generative Artificial Intelligence to Complete Research Work

Purpose of this guidance

This guidance is provided to support safe and responsible use of generative AI (genAI) as a tool to enable research work and complete research tasks. There are many potential uses of genAI in medical and health research and these uses may involve both significant risks and benefits. PHSA seeks to promote and encourage the use of innovative tools to support research, and to encourage innovative research, including in the field of AI.

What is generative artificial intelligence (genAI)?

genAI is a form of AI that primarily uses [machine learning](#)¹ algorithms to **create** text, images, code or sound. genAI tools can be used for a variety of purposes such as summarizing information in plain language, developing analysis code, creating images for grant applications, reducing the word count of text, editing text for grammatical errors and more.

What is covered in this guidance?

This guidance recommends practices for safe and responsible use of genAI to complete research work or tasks, such as editing texts, communicating research findings, and writing analysis code (see Appendix A for example use cases).

What is NOT covered in this guidance?

This guidance does not apply to the use of genAI as part of a study intervention or procedure. For example, research that explores the benefits and actual risks or harm of using genAI in research or healthcare, or the development and application of genAI models for patient care is something that PHSA researchers are ideally positioned to complete but is not addressed here. As with any other type of research involving humans, research proposals aiming to develop or apply genAI in a clinical setting must be reviewed and approved by a research ethics board (REB).

Who is this guidance for?

This guidance is intended to be used by anyone engaging in research at a PHSA program or research institute.

How to use this guidance

Table 1 provides questions and considerations to assist you in informed decision-making regarding use of genAI to complete research work or tasks. Appendixes A and B provide illustrative use cases with clearly outlined risks and mitigation strategies and resources for further learning and development, respectively. This is a living document that will be updated as relevant new knowledge is generated.

We recommend all Principal Investigators (PIs), and research team leaders, talk with your team members and trainees about their current and planned use of genAI at work. Team members should disclose their intent to use genAI tools in their scientific work. Use of genAI will require both transparency of all team members and accountability of PIs.

Why does this guidance matter in a research setting?

One of the primary benefits of genAI is that these tools create content that sounds or looks plausible and is hard to distinguish from human-created content, but this also creates risk. Any content created by genAI is based purely on a model or algorithm that has been developed using information that is already available, typically from internet sources. Output from these tools is therefore prone to error, misinformation, plagiarism or general inaccuracies. Therefore, this guidance highlights the need to carefully review output from genAI tools before it is used in a research context. In addition, any action that involves inputting participant research data with or without personal identifiers into a cloud-based genAI tool must be prevented. Inputting data into a genAI tool has the potential to breach privacy rules as many of these tools store this data and use it for further model development. For further questions related to research data and AI please reach out to AIWorkingGroup@phsa.ca.

Note: This guidance is adapted from Government of Canada Guidance², UBC research guidance and other Canadian academic organization guidance notes.^{3,4,5}

Table 1. Considerations and actions recommended for use of generative AI in a research setting

CONSIDERATION	ACTION
Step 1: Consult	
Have you (the PI) or the team, discussed the possible benefits and harms of using genAI within your team?	<ul style="list-style-type: none"> Contact the AI working group at AIWorkingGroup@phsa.ca if you/your team has questions or would like more information on resources available.
Have you let the PI/your manager/supervisor know you plan to use genAI for any given work task(s)?	<ul style="list-style-type: none"> Disclose planned use of genAI to your manager/supervisor before using it for work related purposes.
Do any privacy or ethics concerns ⁶ need to be addressed prior to use? For example, further consultation is required in the following scenarios: <ul style="list-style-type: none"> Unpublished participant data or study results will be included as part of the prompt or disclosed to the genAI tool during planned use Information provided to the genAI tool includes any personal identifiers or potentially identifying information The genAI tool is being used as part of a study intervention 	<ul style="list-style-type: none"> Consult with PHSA privacy or the appropriate Research Ethics Board to ensure planned use meets ethical and privacy standards when needed. Consider using a local/institutional genAI or a subscription that meets your privacy requirements* Consider if there is an alternative way to use genAI to meet your project goals without full disclosure of sensitive information. For example, can you describe the type of data you have or use synthetic data and ask for example code to complete your analysis rather than provide actual data?
Is genAI the best tool to use to solve your problem?	Consider using a different tool if: <ul style="list-style-type: none"> You need to cite original sources for the information you are generating. Not all genAI tools can provide sources for output generated.

CONSIDERATION	ACTION
	<ul style="list-style-type: none"> ○ The output needs to be replicated by others (i.e., for publication). Many genAI tools do not produce replicable output, even when using an identical prompt.
Step 2: Verify	
<p>Are data sources used for the development of the AI model being applied within the genAI tool known and disclosed by the company who owns it? Are any biases resulting from these data sources understood?^{7,8}</p>	<ul style="list-style-type: none"> ○ Consider the impact the underlying data will have on your output. ○ Review AI product development and use information and AI product privacy and security statements/policies.
<p>Are you using any potentially sensitive information (e.g., participant research data, intellectual property, etc.) as part of the prompt or supplemental data provided to the genAI?</p>	<ul style="list-style-type: none"> ○ Review data sharing policies and determine if the prompts you enter could be shared back with the parent company for further model training, as is likely to happen whenever data is sent to a cloud service provider. <p>Consider using a local/institutional genAI or a subscription that meets PHSA privacy requirements*</p> <ul style="list-style-type: none"> ○ If you do not use a local or privacy protected genAI, do not input research participant data without first consulting with PHSA privacy.
<p>Have you reviewed your prompts and tested them to ensure no additional sources of bias are introduced? (See UBC’s genAI Resources for training resources)</p>	<ul style="list-style-type: none"> ○ It is known that prompts have a significant impact on the quality and content of output generated.⁹ ○ Try several variations of the prompts to understand the impact of the language you are using and ensure context and direction is clear and appropriate to achieve your goals. ○ All prompts, when explicitly entered, should be clearly documented so that results can be verified if needed and replicated where possible.
Step 3: Review	
<p>Have you considered the accuracy and quality of generated output considering known limitations to genAI tools such as copyright infringement, plagiarism, and hallucinations (hallucination are made up facts and this is a known common problem in genAI output)?</p>	<ul style="list-style-type: none"> ○ Output has been carefully reviewed to ensure results are free of bias. ○ Output has been carefully reviewed to ensure any source information is accurately cited. ○ Output has been carefully reviewed to confirm the accuracy of the information presented. ○ A few tips for assessing output include:

CONSIDERATION	ACTION
	<ul style="list-style-type: none"> ○ Use prompts to ask for bulleted list of grammatical errors, instead of inputting a full text document and asking genAI tool to directly fix errors. ○ Ask genAI tool to provide list of references, including hyperlinks, to be able to quickly assess accuracy and quality of source information used to create output.
Step 4: Disclose	
Have you disclosed how and why you used genAI?	<ul style="list-style-type: none"> ○ Final AI generated output included in any publication/ communication/ report includes a clear statement indicating genAI has been used to develop content and naming the tool used. ○ Responsibility for accuracy of the final AI generated output must be acknowledged by the tool user. genAI tools cannot be given authorship so all responsibility for content generated from these tools rests on the user.¹⁰

*No institutionally approved genAI tool is currently supported by PHSA; contact your research institute leadership to discuss locally available tools.

Last Updated: 12-APR-2024

Acknowledgements: We would like to acknowledge the initial work of BC Children’s Research Institute and BC Women’s Research Institute that was used as the foundation for this guidance, as well as the genAI Working Group members (listed in alphabetic order): Taqdir Ali, Mark Ansermino, Tim Bhatnagar, Jeff Bone, Lori Brotto, Kathryn Dewar, Sabina Dobrer, Jessica Gagliardi, Matthias Görges, Philipp Lange, Holly Longstaff, Farin Maralli, Jessica Morrice, John-Jose Nunez, Beth Payne, Gina Ogilvie, Nicole Prestley, Akshdeep Sandu, Hind Sbihi, Christian Schutz, Rishav Singh, Jonath Sujan, Lisa Tang, Aasta Thielke, Steven Ufkes, Tibor van Rooij, Benjamin Wan and William Wang.

References

1. Brown, S. (2021, April 21). [Machine learning, explained](#). MIT Management Sloan School.
2. Government of Canada. (2024). [Guide on the use of generative AI](#).
3. University of British Columbia. (n.d.). [Generative AI: UBC guidance](#).
4. University of Toronto, School of Graduate Studies. (2023). [Guidance on the appropriate use of generative artificial intelligence in graduate theses](#).

5. Western Canadian Deans of Graduate Studies. (2023). [Generative AI and graduate and postdoctoral research and supervision.](#)
6. Aiken, C., Flann, S., Longstaff, H., Manusha, S., Pavlovich, S., Scott, J. & and Wright, J. (2021). *A guidance for novel ethics of privacy issues associated with artificial intelligence in the public sector research domain.* [http://www.phsa.ca/researcher/Documents/AI%20Guidance_FINAL.pdf.](http://www.phsa.ca/researcher/Documents/AI%20Guidance_FINAL.pdf)
7. Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M.A., Al-Busaidi, A.S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L., et al. (2023). Opinion paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71(102642). [https://doi.org/10.1016/j.ijinfomgt.2023.102642.](https://doi.org/10.1016/j.ijinfomgt.2023.102642)
8. Celi, L.A., Cellini, J., Charpignon, M.-L., Dee, E.C., Dernoncourt, F., Eber, R., Mitchell, W.G., Moukheiber, L., Schirmer, J., Situ, J., Paguio, J., Park, J., Wawira, J.G. & Yao, S. (2022). Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review. *PLOS Digital Health*, 1(3): e0000022. [https://doi.org/10.1371/journal.pdig.0000022.](https://doi.org/10.1371/journal.pdig.0000022)
9. Shah, C. (2024). From prompt engineering to prompt science with human in the loop. [*Unpublished paper*]. [https://arxiv.org/pdf/2401.04122.pdf.](https://arxiv.org/pdf/2401.04122.pdf)
10. International Committee of Medical Journal Editors. (2024). [Defining the role of authors and contributors.](#)

Appendix A: genAI Use Case Examples

These use cases are provided to further describe potential benefits and risks of using genAI as a tool to complete research tasks. In each use case we provide a discussion of benefits, risks and mitigation strategies based on current knowledge.

- [Example 1: Using ChatGPT to Summarize Literature on A Research Topic](#)
- [Example 2: Code Debugging or Writing](#)
- [Example 3: Writing a Lay Language Summary of a Study to Communicate Results to The Lay Public](#)
- [Example 4: Generating a Figure for a Grant Proposal](#)
- [Example 5: Reviewing Research Applications/Proposals](#)

Example 1: Using ChatGPT to Summarize Literature on A Research Topic

Scenario
<p>Jill is a graduate student and working with her supervisor on a Canada Institutes of Health Research application. She is asked to draft a few paragraphs for an introduction on the topic of long-term effects of the COVID-19 pandemic restrictions on children’s school performance. She decides to use ChatGPT to help her get started and get ideas about what to write. Jill uses ChatGPT 3.5 because she has the free version of ChatGPT. She has read the OpenAI security information and knows that with this free version the data she submits as a prompt may be used to train future models. In this case, she is not concerned by this as she is not inputting any sensitive information. Jill plans out her prompts with the following results:</p>
Prompts
Prompt 1: What are the long-term effects of COVID-19 pandemic restrictions on children’s school performance?
Prompt 2: Expand on this statement to write two paragraphs.
Outputs
<p>Output 1: “The long-term impacts of the COVID-19 pandemic restrictions on children's school performance are still being studied, but potential effects may include learning loss, increased mental health challenges, decreased socialization skills, and widening educational disparities. The extent of these impacts can vary depending on factors such as access to resources, quality of remote learning, and individual circumstances. Ongoing support and interventions will be crucial in mitigating these effects.”</p>
<p>Output 2: “The COVID-19 pandemic has significantly disrupted traditional education systems worldwide, leading to prolonged school closures, remote learning mandates, and altered classroom environments. These disruptions have raised concerns about the long-term impacts on children's academic performance and overall well-being. One of the most significant challenges is learning loss, as students may have missed critical instruction time, experienced interruptions in their learning routines, and faced difficulties in engaging with remote learning platforms. This loss could manifest in lower academic achievement, particularly in foundational subjects like math and reading, and may disproportionately affect students from disadvantaged backgrounds who lack access to necessary resources or face additional barriers to learning.</p> <p>Moreover, the pandemic has taken a toll on children's mental health and social development. Prolonged periods of isolation, limited social interactions, and heightened stress levels can contribute to increased anxiety, depression, and feelings of loneliness among students. Additionally, the shift to remote learning has reduced opportunities for peer interaction, collaboration, and extracurricular activities, which are essential for developing social skills, emotional intelligence, and resilience. Addressing these long-term impacts will require comprehensive support systems within educational institutions, including targeted interventions to address learning gaps, increased access to mental health resources, and efforts to rebuild social connections and community engagement among students.”</p>
Risks and Benefits Discussion
Benefits
<ul style="list-style-type: none"> ○ Possible efficiency

Risk	Mitigation
<p>Uncertainty/Misinformation</p> <p>The output generated is very general and lacks citations. Some genAI tools (such as the paid version of ChatGPT) allow you to upload a set of papers and ask questions specific to those sources. Output must still be carefully reviewed as hallucinations (made up facts) are always possible with these tools.</p>	<p>All listed facts/statements in the output must be verified, and references found. Prompting ChatGPT to provide sources will result in a list of potential websites to review but these will not be linked to any specific statement in the output. Literature searches are more efficiently performed in a reputable database.</p>
<p>Privacy</p> <p>Any data input into a genAI tool may be stored by the parent company for further model development or monitoring. It is often unclear what is stored and who may get access to it.</p>	<p>Never input sensitive information into a genAI tool with unknown privacy agreement. Find an alternate strategy to achieve your goals or consider using a local/institutional genAI or a subscription that meets PHSA privacy requirements.</p>
<p>Bias</p> <p>Most genAI tools are developed on corpuses of information scrapped from the internet and these sources are not fully disclosed. It is often unclear what the resultant bias in the output may be.</p>	<p>All output must be carefully considered for possible racial, gender and other biases. It is up to the end-user to take responsibility for output.</p>
<p>Plagiarism/Disclosure</p> <p>Many genAI tools do not provide clear citations or source information and will directly copy statements from books or articles in the output. Without verifying output and identifying reputable sources you run the risk of plagiarizing other people's work.</p>	<p>All listed facts/statements in the output must be verified, and appropriate references found.</p>
<p>Helpful Tips for Using genAI for Writing and Editing</p>	
<p>✓ Ask genAI tool to provide list of references, including hyperlinks, to be able to quickly assess accuracy and quality of source information (possible in some tools)</p>	
<p>✓ When editing – ask for bullet point suggestions for edits rather than a direct re-write of provided text. This allows you to maintain control of all changes and verify the quality of the final written product.</p>	

Example 2: Code Debugging or Writing

Scenario
<p>Philip is an undergraduate student working on a research project that aims to create an image classifier that can classify whether the chest x-ray shows a pneumothorax. He is familiar with training image classifiers in Tensorflow but wants to implement it in Pytorch. He decides to use ChatGPT to help set up the training function.</p>

Prompt

“Generate a PyTorch code for a function that trains an image classifier. The function should accept the following parameters:

train_loader: DataLoader object representing the training dataset.

model: The PyTorch model architecture to be trained.

criterion: The loss function used for optimization.

optimizer: The optimizer for updating model parameters.

Optional parameter: num_epochs specifying the number of epochs for training.

Optional parameter: device indicating the device for computation.”

Output

```
import torch

def train_image_classifier(train_loader, model, criterion, optimizer, num_epochs=10,
                           model_device='cpu'):
    model.to(model_device)
    for epoch in range(num_epochs):
        model.train()
        running_loss = 0.0
        for inputs, labels in train_loader:
            inputs, labels = inputs.to(model_device), labels.to(model_device)
            optimizer.zero_grad()
            outputs = model(inputs)
            loss = criterion(outputs, labels)
            loss.backward()
            optimizer.step()
            running_loss += loss.item() * inputs.size(0)

        epoch_loss = running_loss / len(train_loader.dataset)
        print(f'Epoch [{epoch+1}/{num_epochs}], Loss: {epoch_loss:.4f}')

    print('Training Finished!')
```

Risks and Benefits Discussion

Benefits

- Efficiency: saves time by automatically writing boilerplate codes
- Provides different approaches to a problem
- Can help user learn about syntaxes of a new programming language/libraries

Risk

Mitigation

Incorrect Code

If the prompt provided is not clear in terms of what the user wants the code to do, the large language model

Always manually check the logic of the code generated by the LLM.

<p>(LLM) can produce a code that provides a different logic than desired. Like hallucination in text-generation context, the LLM may produce results that incorporate non-existent libraries/tools or incorrect logic, despite giving an appropriate prompt. In addition, the logical errors may be subtle and difficult for the user to debug.</p>	
<p>Outdated Code ChatGPT’s internal knowledge base may not be the most up-to-date, so it may not be able to work with the newer libraries or tools. Moreover, it may propagate reliance on outdated practices or code.</p>	<p>Being aware of how frequently the LLM is updated provide a rough estimate of how up-to-date the code it generates could be. User may need to refer to other documentations for the newer libraries and tools.</p>
<p>Efficiency LLM may produce a code that is logically correct, but computationally inefficient.</p>	<p>Manually review of code to optimize any inefficiencies.</p>
<p>Security The LLM may suggest code that is functional but contains security vulnerabilities.</p>	<p>There are LLMs which are specifically trained to detect and fix security vulnerabilities. However, it is important to manually review and test for security vulnerabilities and to consult security expert, if required.</p>
<p>Privacy Any data input into a genAI tool may be stored by the parent company for further model development or monitoring. It is often unclear what is stored and who may get access to it.</p>	<p>Never input sensitive information into a genAI tool with unknown privacy agreement. Find an alternate strategy to achieve your goals or consider using a local/institutional genAI or a subscription that meets PHSA privacy requirements .</p>
<p>Helpful Tips for Using genAI for Coding</p>	
<p>✓ Describe your data in the prompt and ask for example code that would work with the type of variables your dataset includes.</p>	

Example 3: Writing a Lay Language Study Summary to Communicate Results to The Lay Public

Scenario	
<p>Giuseppina is putting together a grant due in a few hours. She is done most components, but still has to write a lay summary that can be understandable by the lay public.</p> <p>She takes the abstract for her grant, and supplies it to ChatGPT with the follow prompt:</p>	
Prompt	
<p>"Please rewrite the following scientific abstract so that it can be understandable by the lay public:</p> <p>In recent years, stem cell therapy has become a very promising and advanced scientific research topic. The development of treatment methods has evoked great expectations. This paper is a review focused on the discovery of different stem cells and the potential therapies based on these cells. The genesis of stem cells is followed by laboratory steps of controlled stem cell culturing and derivation. Quality control and teratoma formation assays are important procedures in assessing the properties of the stem cells tested. Derivation methods and the utilization of culturing media are crucial to set proper environmental conditions for controlled differentiation. Among many types of stem tissue applications, the use of graphene scaffolds and the potential of extracellular vesicle-based therapies require attention due to their versatility. The review is summarized by challenges that stem cell therapy must overcome to be accepted worldwide. A wide variety of possibilities makes this cutting edge therapy a turning point in modern medicine, providing hope for untreatable diseases."*</p>	
Output	
<p>"Stem cell therapy has become an exciting field of research in recent years, holding promise for treating various health conditions. This paper reviews the discovery of different types of stem cells and how they're used for potential treatments. It explains how scientists grow and study these cells in the lab, ensuring they're safe and effective. The review also talks about special materials like graphene and tiny bubbles called extracellular vesicles, which are being explored for their role in therapy. Despite the excitement, there are still challenges to overcome before these treatments can be widely used. Overall, stem cell therapy offers hope for diseases that were previously untreatable."</p>	
Risks and Benefits Discussion	
Benefits	
<ul style="list-style-type: none"> ○ Quickly produce a lay abstract ○ ChatGPT supplies some ideas for her how to word certain concepts in a widely accessible way 	
Risk	Mitigation
<p><i>Inaccuracy</i></p> <p>The output may have factual errors introduced, given that it is trained on a general language corpus and not specifically on scientific content. It may summarize</p>	<p>All output must be carefully checked for scientific accuracy.</p>

concepts incorrectly or reword statements that are no longer correct.	
Language Use The output may be rewritten to be too simple, or may still be too specific and not appropriate for a lay audience.	The output must be checked and compared with the tone, complexity and language use that the author would normally use. ChatGPT’s judgement should not be substitute for the user’s.
Plagiarism/Disclosure As in a prior example, genAI may directly plagiarize others when generating its content. Even if the subject matter is specific to your grant, generated sentences could still be inadvertently plagiarized.	If ChatGPT has significantly reworded sentences or paragraphs, consider rewriting them in your own style, using some of the suggested wording if desired. If in doubt, online tools for plagiarism detection can be used, though these are not foolproof.
Helpful Tips for Using genAI to Summarize Text	
✓ Use specific language to describe the audience for the intended output	
✓ Use an iterative process to keep refining the results until the desired result is achieved	

* Example abstract from Zakrzewski, W., Dobrzyński, M., Szymonowicz, M. et al. (2019). Stem cells: past, present, and future. Stem Cell Research & Therapy, 10, 68. <https://doi.org/10.1186/s13287-019-1165-5>

Example 4: Generating a Figure for a Grant Proposal

Scenario	
Jennifer is working on a proposal for an infrastructure grant that involves building a new research consortium. She wants to create a logo for the consortium to develop some branding and create a better sense of coherence. She finds an AI based logo generator online.	
Prompts	
Input: description of the consortium and proposed name	
Outputs	
Output: several logo options that can be purchased for use	
Risks and Benefits Discussion	
Benefits	
<ul style="list-style-type: none"> ○ New logos are created quickly and easily 	
Risk	Mitigation
<p>Plagiarism</p> <p>genAI-based image generators are trained on existent images. A new logo may be very similar to that existent images that were used for training.</p>	Use an image-searcher to compare your image against others (e.g., https://images.google.ca/) As possible, try to avoid very simple logos that might be similar to what already exists.
<p>Terms of Use</p> <p>Image generation tools may allow images to be used with restriction and may retain rights to use the image in certain context.</p>	Carefully examine Terms of Use for the product, noting if there are any restriction, or if users retain full rights to the images. If in doubt, use a different product, or consult legal resources.
<p>Inappropriate Content</p> <p>Such tools are trained on images across the internet and can included a wide variety of contents. Generated images may contain components that have special meanings to groups of people, even if not obvious to yourself.</p>	Ensure that a wide variety of others review generated images before they are publicly shared, especially if the logo contains elements that seem distinct. Have a low threshold to replace elements that may possibly be misunderstood.
<p>Bias</p> <p>If your logo contains representations of humans, genAI models may portray them by furthering stereotyping with respect to gender, race, class, profession or other attributes. However, models may also try to correct for this but lead to portrayals that are incorrect.</p>	Consider whether your images further both positive and negative stereotypes. Consider revising your prompt to add diversity and combat such stereotypes. Ensure that portrayals are still consistent with intention.

Helpful Tips for Using genAI to Create Images or Logos

- ✓ For some image generators you can specify the licensing type you are interested in. For use in public or scholarly communication this should typically be for public use/domain, or “Creative Commons” known as a Creative Commons Zero (CC0) license type. See [here](#) for an example of this.

Example 5: Reviewing Research Applications/Proposals

Scenario	
Jacob is reviewing several applications for a research institute graduate student travel award. He wants to use ChatGPT to select the strongest application. The rubric provided for the grant review asks reviewers to give each application a score out of 5 total points in the following areas: candidates track record of academic performance and productivity; relevance of the travel to the trainee's research; and quality of proposed research or training activity associated with the travel. He copies and pastes each application with the applicant's CV into ChatGPT before proving a prompt.	
Prompts	
Input: I am going to provide you with a number of grant applications. Summarize each application using the following criteria: number of candidates publications; total impact factor of publications; clarity of writing; inclusion of a research or training plan in the application; and inclusion of a statement linking proposed travel to the applicant's research goals.	
Outputs	
A list of reviews giving details of each criteria provided	
Risks and Benefits Discussion	
Benefits	
<ul style="list-style-type: none"> ○ Quickly pulls out some relevant information for the review ○ Summary statements could be used to report back on scores 	
Risk	Mitigation
<p><i>Bias</i></p> <p>If you include names, gender or age of applicants, models can introduce bias in how these factors are assessed due to the underlying bias in development data. This could lead to reduced scores or lower assessments of quality for applicants that are unfair.</p>	Do not ask for subjective review of quality of proposals using a genAI tool. Only very clearly defined quantitative assessment categories could be used, but output still needs to be double checked for accuracy.
<p><i>Inaccuracies</i></p> <p>There is a known issue for genAI tools to hallucinate or make up facts. The output may therefore have factual errors introduced, given that it is trained on a general language corpus and not specifically on scientific content. It may summarize concepts incorrectly or reword statements that are no longer correct.</p>	All output must be carefully checked for accuracy.
<p><i>Lack of reproducibility/transparency</i></p> <p>The models used for genAI tools are continuously being revised making reproducibility of results often impossible</p>	Use of a genAI tool to support grant review should always be disclosed to the funding body. Justification for funding

even with the exact same prompt. This ongoing change in models also means we cannot fully know how the prediction/output is being generated.

decisions must be clear and many funding bodies may not support the use of genAI in this process.

Helpful Tips for Using genAI to Support Grant Review

- ✓ Check with the funding body first if they will accept review supported by genAI.
- ✓ Only include clear and quantitative categories of review so that output can be verified and checked for accuracy.

Appendix B: Resources for Further Learning and Development

The following resources are a curated list of articles, tools and trainings that provide an overview of genAI and discussion of key opportunities and challenges. Additional resources will be added to this list as they become available, as well as listed on the AI Working Group [website](#).

Bias, Transparency and Equity

Celi, L.A., Cellini, J., Charpignon, M.-L., Dee, E.C., Deroncourt, F., Eber, R., Mitchell, W.G., Moukheiber, L., Schirmer, J., Situ, J., Paguio, J., Park, J., Wawira, J.G. & Yao, S. (2022). **Sources of bias in artificial intelligence that perpetuate healthcare disparities — A global review.** *PLOS Digital Health*, 1(3): e0000022.

<https://doi.org/10.1371/journal.pdig.0000022>.

Literature review on data used to train clinical AI systems, and demographics of authors reporting on the use of AI in clinical medicine. The review found a disproportionate overrepresentation of US and Chinese datasets and authors in the literature, including 40% of the studies being radiology focused. The authors caution using narrow data-rich populations to train clinical AI systems could further perpetuate health disparities in data-poor populations.

Nazer LH, Zatarah R, Waldrip S, Ke JXC, Moukheiber M, Khanna AK, et al. (2023) **Bias in artificial intelligence algorithms and recommendations for mitigation.** *PLOS Digit Health* 2(6): e0000278. <https://doi.org/10.1371/journal.pdig.0000278>

The authors outline sources of potential bias in the development and implementation of AI algorithms for healthcare and discuss strategies to mitigate them. Sources of bias are identified across the entire spectrum of AI development and implementation starting with defining the problem, data collection and processing, model development and validation, and final implementation. Many examples are provided to illustrate each type of known bias. A checklist is provided to help researchers understand potential biases and address them throughout the AI research and implementation process.

Ethics, Privacy and Governance

Aiken, C., Flann, S., Longstaff, H., Manusha, S., Pavlovich, S., Scott, J. & Wright, J. (2021). **A guidance for novel ethics of privacy issues associated with artificial intelligence in the public sector research domain.**

http://www.phsa.ca/researcher/Documents/AI%20Guidance_FINAL.pdf.

Highlights key themes from a literature review on the use of AI in the public sector and related ethics of privacy issues. Key themes and associated recommendations include data quality and assessment, perceptions and norms, access, financial consideration, education, research participant safety and care, intellectual property, and governance.

McCadden, M.D., Anderson, J.A., Stephenson, E., Drysdale, E., Erdman, L., Goldenberg, A., & Zlotnik Shaul, R. (2022). **A research ethics framework for the clinical translation of healthcare machine learning.** *American Journal of Bioethics*, 22(5), 8-22. doi: [10.1080/15265161.2021.2013977](https://doi.org/10.1080/15265161.2021.2013977).

Discusses using an adapted research ethics guidelines and privacy protections to develop a framework for evaluating and translating machine learning models into clinical care. The research ethics framework includes

three phases: (1) exploratory machine learning research; (2) silent evaluation; and (3) prospective clinical evaluation.

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*.

<https://www.who.int/publications/i/item/9789240029200>

Based on the expertise and work from 20 leading experts, the World Health Organization identified six core principles to promote the ethical use of AI in health: (1) protect autonomy; (2) promote human well-being, human safety and the public interest; (3) ensure transparency, explainability and intelligibility; (4) foster responsibility and accountability; (5) ensure inclusiveness and equity; and (6) promote AI that is responsible and sustainable. In addition, the guidance discusses several key considerations regarding the use of AI in health, namely law and policies; key ethical principles; ethical challenges; how to build an ethical approach; liability regimes; and governance framework elements.

Generative AI engineering approaches and use

Bandi, A., Adapa, P.V.S.R., & Kuchi, Y.E.V.P.K. (2023). **The power of generative AI: A review of requirements, models, input–output formats, evaluation metrics, and challenges**. *Future Internet*, 15(8), 260.

<https://doi.org/10.3390/fi15080260>.

Based on a literature review, this article provides overview and analysis of genAI requirements, models, generative types, input-output classification and evaluation metrics, and discusses challenges and implementation issues related to genAI. The article includes discussion of common phases of using genAI including problem definition, data collection and preprocessing, model selection, model training, model evaluation, model fine-tuning, deployment, and monitoring and maintenance.

Shah, C. (2024). **From prompt engineering to prompt science with human in the loop**. [Unpublished paper].

<https://arxiv.org/pdf/2401.04122.pdf>.

Based on the methods of qualitative data coding, this paper proposes a similar approach to address the unexplainable, unverifiable, and less generalizable outcomes from differences in prompt engineering in the use of large language models in research. The author proposes an iterative approach to prompt development, simultaneously refining prompts while training researchers to objectively, consistently, and independently evaluate large language model responses.

Online Courses – Machine Learning

Google for Developers. (n.d.) *Introduction to machine learning*. Google for Developers.

<https://developers.google.com/machine-learning/intro-to-ml>.

A 20-minute course aimed at exploring the different types of machine learning, the key concepts of supervised machine learning, and approaches to solving problems with machine learning that differs from traditional approaches.

Opportunities and Risks of Generative AI

Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koochang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M.A., Al-Busaidi, A.S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L., et al. (2023). **Opinion paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy.** *International Journal of Information Management*, 71(102642). <https://doi.org/10.1016/j.ijinfomgt.2023.102642>.

This paper brings together opinions on the opportunities and challenges of transformative AI tools from 43 international experts across the fields of computer science, marketing, information systems, education, policy, hospitality and tourism, management, publishing and nursing. Key themes across the contributions included ChatGPT as a productivity enhancing tool, academia likely to experience some of the most disruptive effects, concerns about job losses, the potential misuse and abuse of AI, major limitations of genAI tools, the lack of regulatory templates, and future research directions. Based on these key themes, the authors propose 10 key research areas of transformative AI tools.

Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K. & Chen, L. (2023). **Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration.** *Journal of Information Technology Case and Application Research*, 25(3), 277-304. <https://doi.org/10.1080/15228053.2023.2233814>.

Discusses key applications of generative AI in business, education, health care and content generation. The authors further discuss ethical, technological, regulatory and policy challenges with generative AI and the value of human-centered AI collaboration to guide the design and application of generative AI.

Government of Canada. (2024). **Guide on the use of generative AI.** <https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html>.

Provides an overview of genAI, discusses challenges and opportunities with genAI use, and provides best practices for genAI users in federal institutions related to protection of information, bias, quality, public servant autonomy, legal risks, distinguishing humans from machines, and environmental impacts. The guidance recommends assessing the risks of using genAI for different applications, and only use genAI in situations where the risks can be mitigated.

Varghese, J. & Chapiro, J. (2023). **ChatGPT: The transformative influence of generative AI on science and healthcare.** *Journal of Hepatology*, [in press, corrected proof]. <https://doi.org/10.1016/j.jhep.2023.07.028>.

Provides a general overview of general AI and AI subtypes. Using ChatGPT as an example, the authors discuss opportunities of genAI use for text generation and advanced coding tasks, challenges related to bias, transparency, explainability, and data fabrication, and practical opportunities and regulatory challenges for use of genAI in health care.

Use of Generative AI in Graduate Student Research and Writing

University of British Columbia. (n.d.). **Generative AI: UBC guidance.** <https://genai.ubc.ca/guidance/>.

Provides a compilation of tools and resources to support faculty, staff, students and researchers in the responsible use of genAI.

University of Toronto, School of Graduate Studies. (2023). ***Guidance on the appropriate use of generative artificial intelligence in graduate theses.*** <https://www.sgs.utoronto.ca/about/guidance-on-the-use-of-generative-artificial-intelligence/#:~:text=This%20includes%2C%20for%20example%2C%20using,other%20uses%20of%20generative%20AI.>

Provides guidance on the use of genAI tools (e.g., ChatGPT) in graduate student research and thesis writing. Includes a set of frequently asked questions and additional resources related to the use of genAI in research, academic writing, and thesis development and editing.

Western Canadian Deans of Graduate Studies. (2023). ***Generative AI and graduate and postdoctoral research and supervision.*** http://wcdgs.ca/content/dam/ex/wcdgs/Nov24_2023WorkingGroupReport_GenerativeAI_GraduateResearch_SupervisionDeliverables.pdf.

Provides recommendations on the use of genAI in graduate research and writing and how to develop AI literacy for graduate supervisors and students.