

Generative AI for Researchers

March 19, 2024

Jillian Anderson | Simon Fraser University



Agenda

- **GenAI Foundations**

- **Opportunities**

- Roles of GenAI in Research
- Use Cases

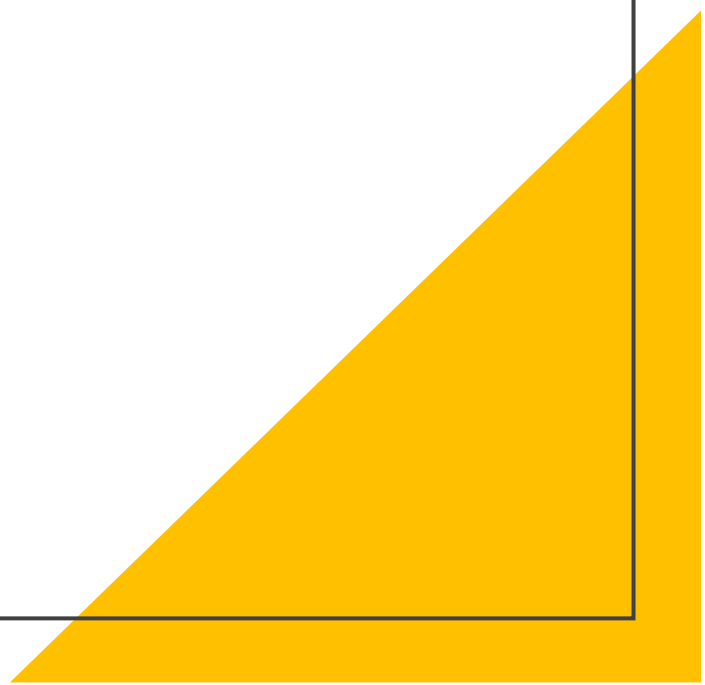
- **Risks**

- Practical Concerns
- Researcher Duty
- Bigger Picture Concerns

- **Best Practices**

- Policies & Guidelines
- Advice from the field

GenAI Foundations



What is GenAI?

- AI models trained to produce content
 - text, images, audio, video, code
- Tools such as Chat-GPT, Genesis, and DALL-E-3
- Prior AI models could recognize patterns, but couldn't create "new" things

Note

Many more types of models are available! Checkout this [Wikipedia section](#) for some other examples.



- Generative AI isn't new, its wide availability & easy use is
- A brief timeline
 - 2014 - 2017 → Critical technologies emerge (e.g. [GANs](#))
 - 2019 - 2020 → LLMs are developed ([BERT](#) & [GPT-3](#))
 - 2022 → OpenAI releases [ChatGPT](#)
- Rapidly evolving space

Note

*A more complete timeline
can be found [here](#).*



You

You are a renowned concept artist. Draw a super computer named cedar. It should be fanciful and a little gritty.



"A fanciful and gritty super computer named Cedar"

Designer

Powered by DALL·E 3

Generative AI in Research

1. Research *in* AI or LLM methods

- E.g. new architectures, alternatives to attention, optimizations

2. Research *about* AI or LLMs

- E.g. ethics of data collection, philosophy, how LLMs will affect jobs

3. Research *with* AI or LLMs

- Application of existing models/tools to other research areas
- E.g. analysis, writing, summarizing, coding

Meeting Agenda for New Project x +

1sfu-my.sharepoint.com/personal/jilliana_sfu_ca/_layouts/15/doc.aspx?so...

Meeting Agenda for New Project Search for tools, help, and more (Option + C) Jillian Anderson

Footer

1

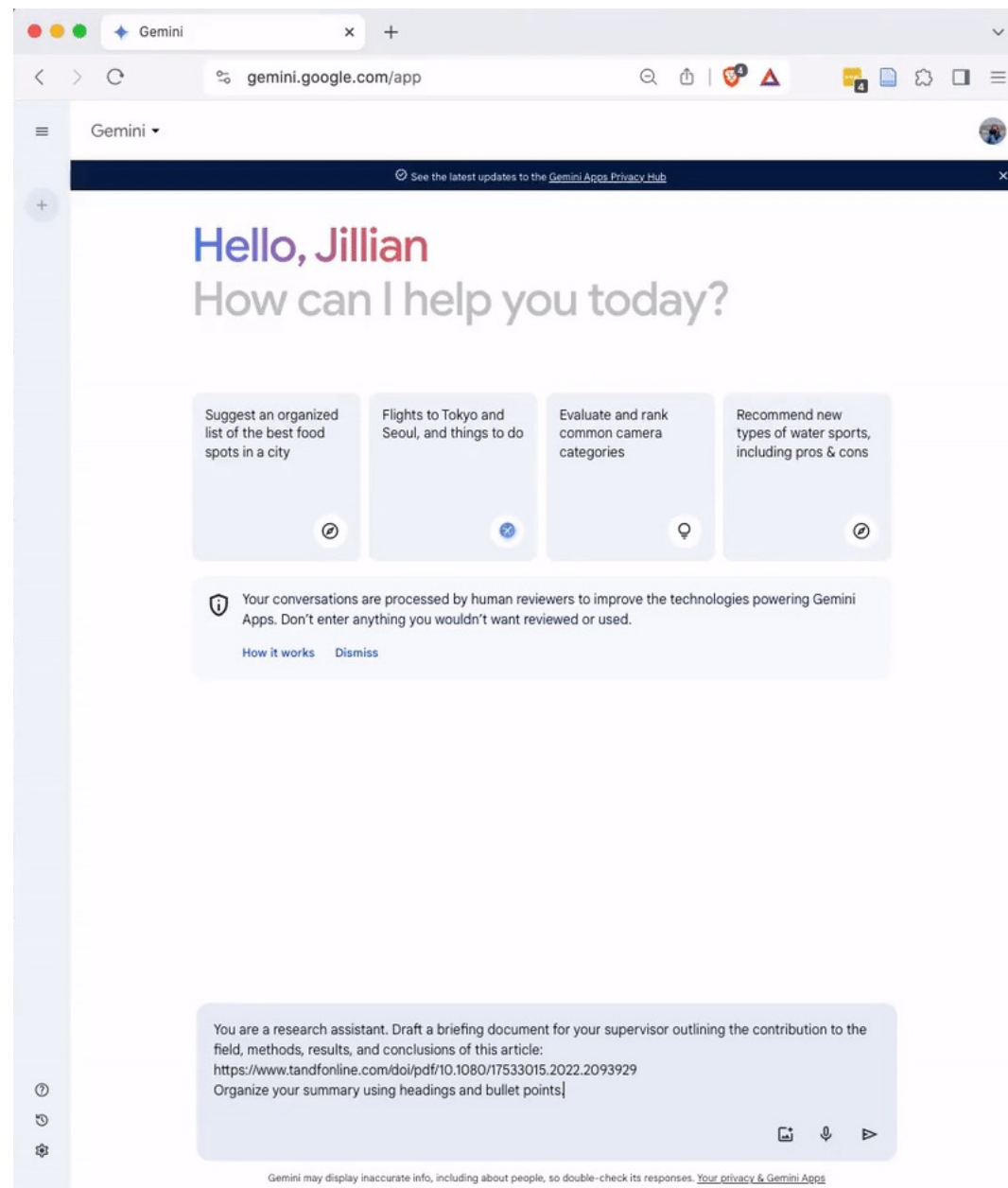
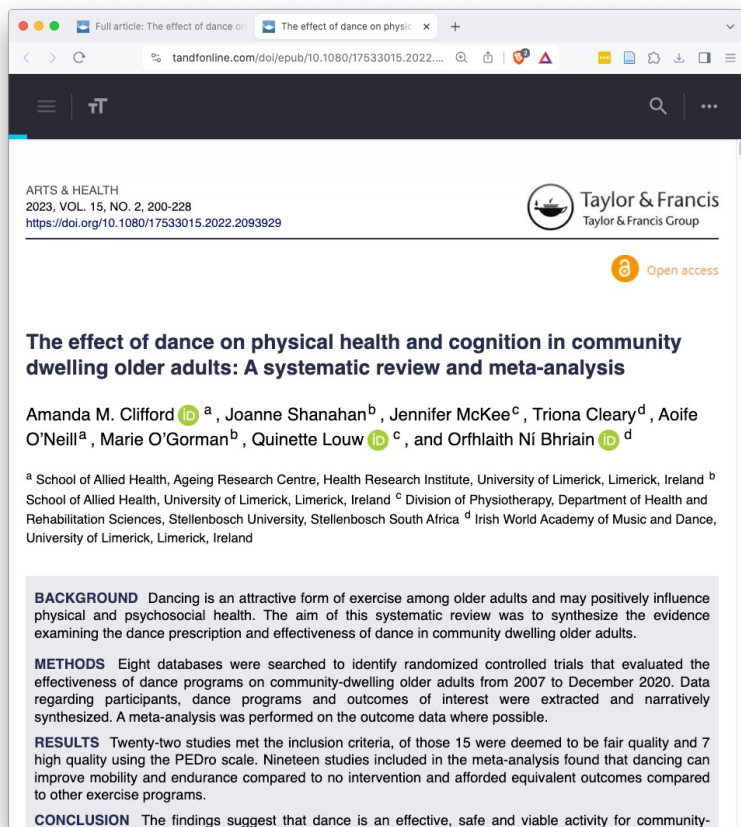
Draft with Copilot

Describe what you'd like to write, including notes or an outline, and Copilot can generate a draft to help get you started

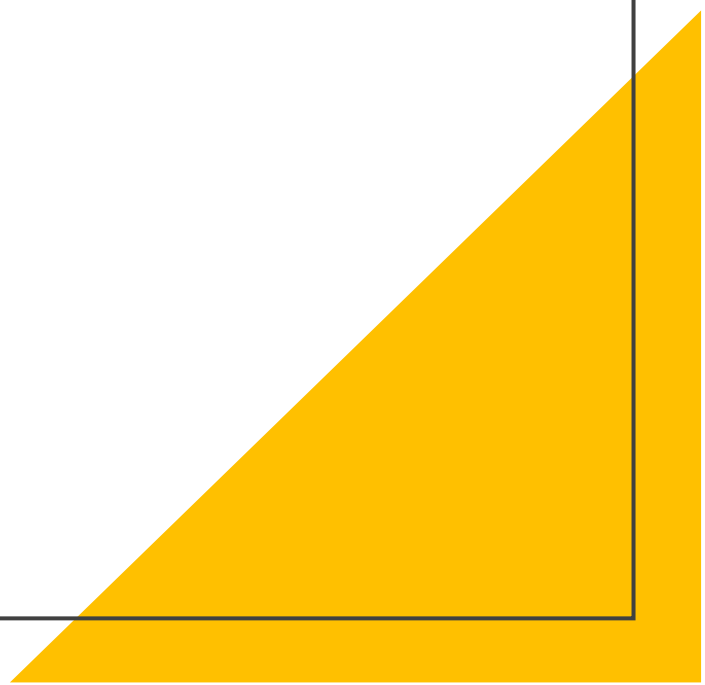
0 / 2000

Generate Cancel Reference your content

Page 2 of 2 118 words English (U.S.) Text Predictions: On Editor Suggestions: Showing Loading Add-ins 100% Give Feedback to Microsoft



Opportunities

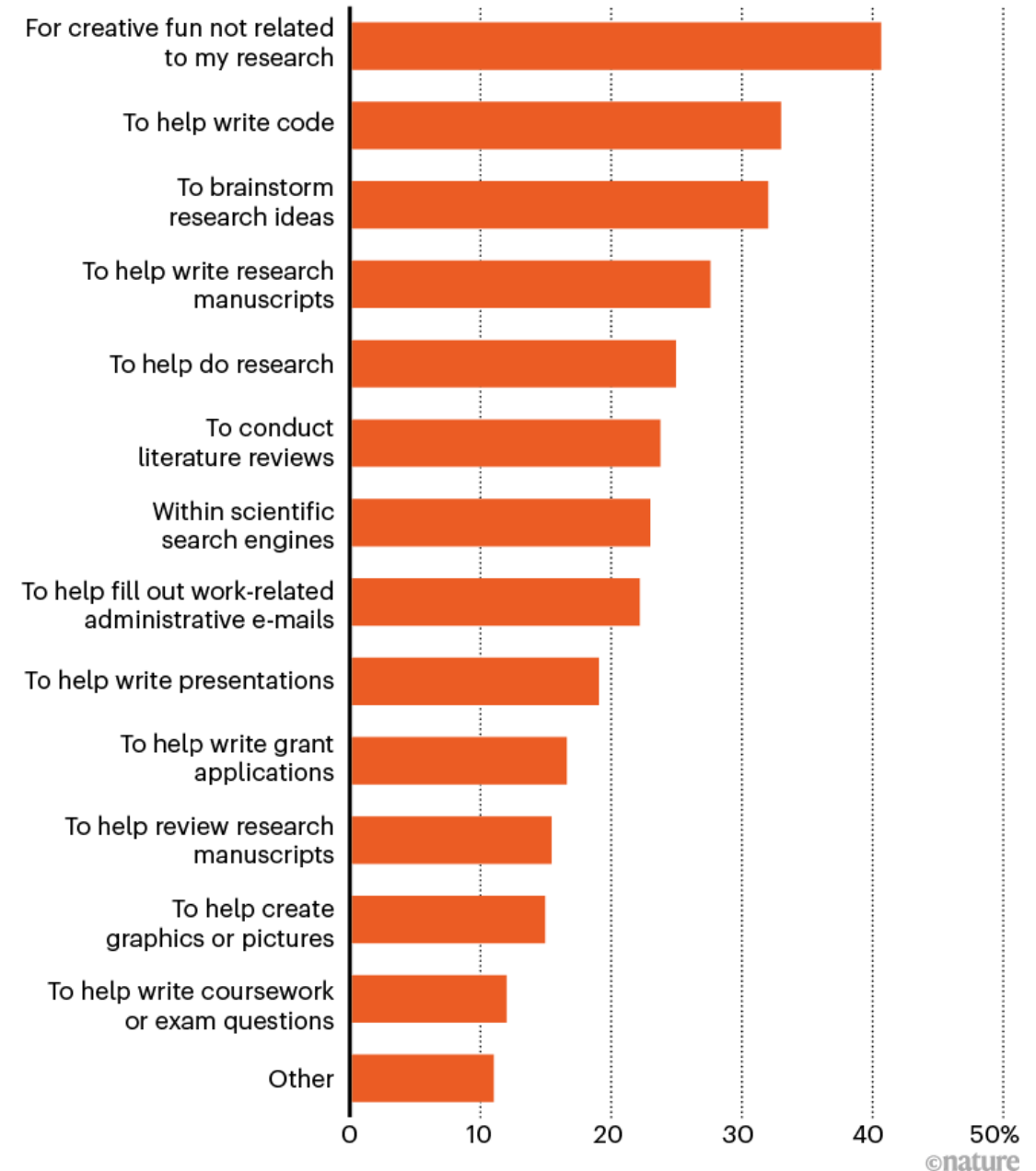


Roles of GenAI in Research

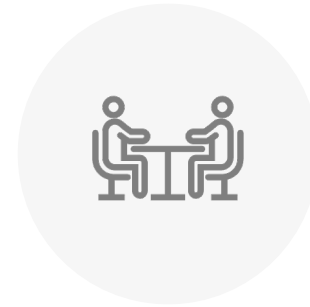
The many hats of GenAI in research

How Researchers Use LLMs

Q: What do you use generative AI tools for? (Choose all that apply)



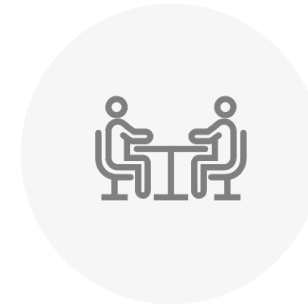
Roles of GenAI in Research



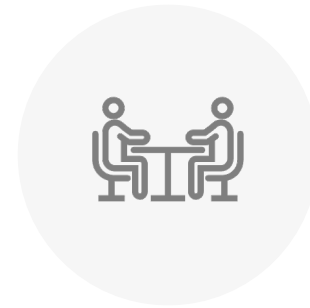
Roles of GenAI in Research



Copyeditor

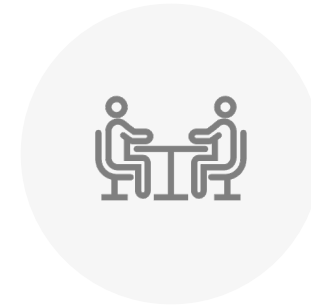


Roles of GenAI in Research



Ghost Creator

Roles of GenAI in Research



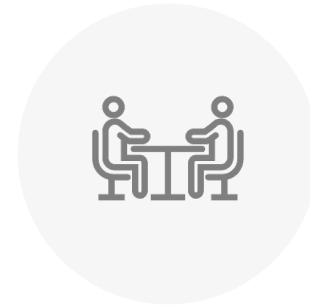
Research Assistant

Roles of GenAI in Research



Sounding Board

Roles of GenAI in Research



Educator

Use Cases

How can we apply those roles to stages of the research process?

Literature Review

- Research Assistant
 - Triaging sources for inspection
 - Orientation to the research landscape
- Educator
 - Explaining jargon

Writing

- Ghost Creator
 - Writing drafts
 - Generating titles or abstracts
- Sounding board
 - Improving clarity
 - Raising questions or concerns
 - Reviewing style – voice, tone, etc.

Writing Code

- Copyeditor
 - Help discover why code might be broken
- Ghost Creator
 - Generate documentation for code you wrote
 - Write code for visualizations
- Educator
 - Explain other code

Warning

As with any code found on the internet, there is no guarantee the code genAI produces is correct!

Note

Checkout Marie's webinar on GitHub Copilot [here](#)

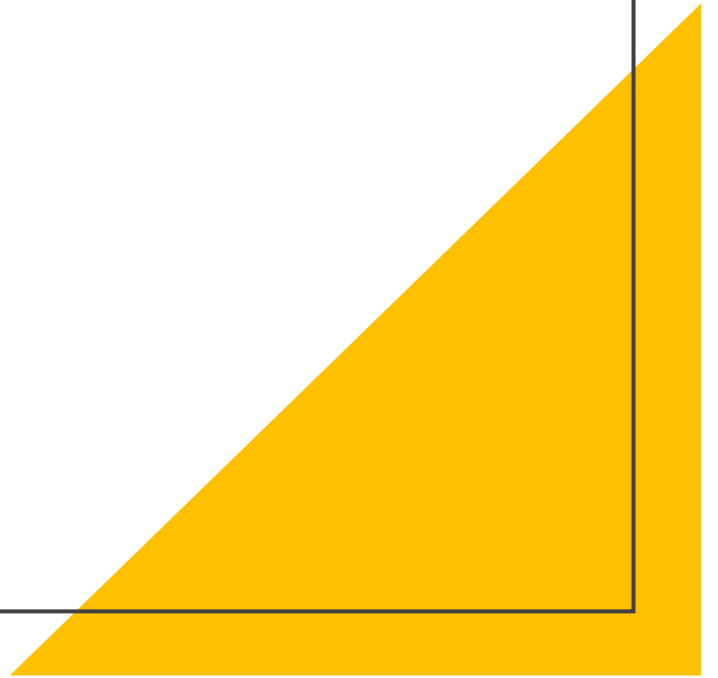
Other Opportunities

- Reducing language barriers
- Identify funding opportunities
- Finding collaborators
- Etc.

Discipline Specific Resources

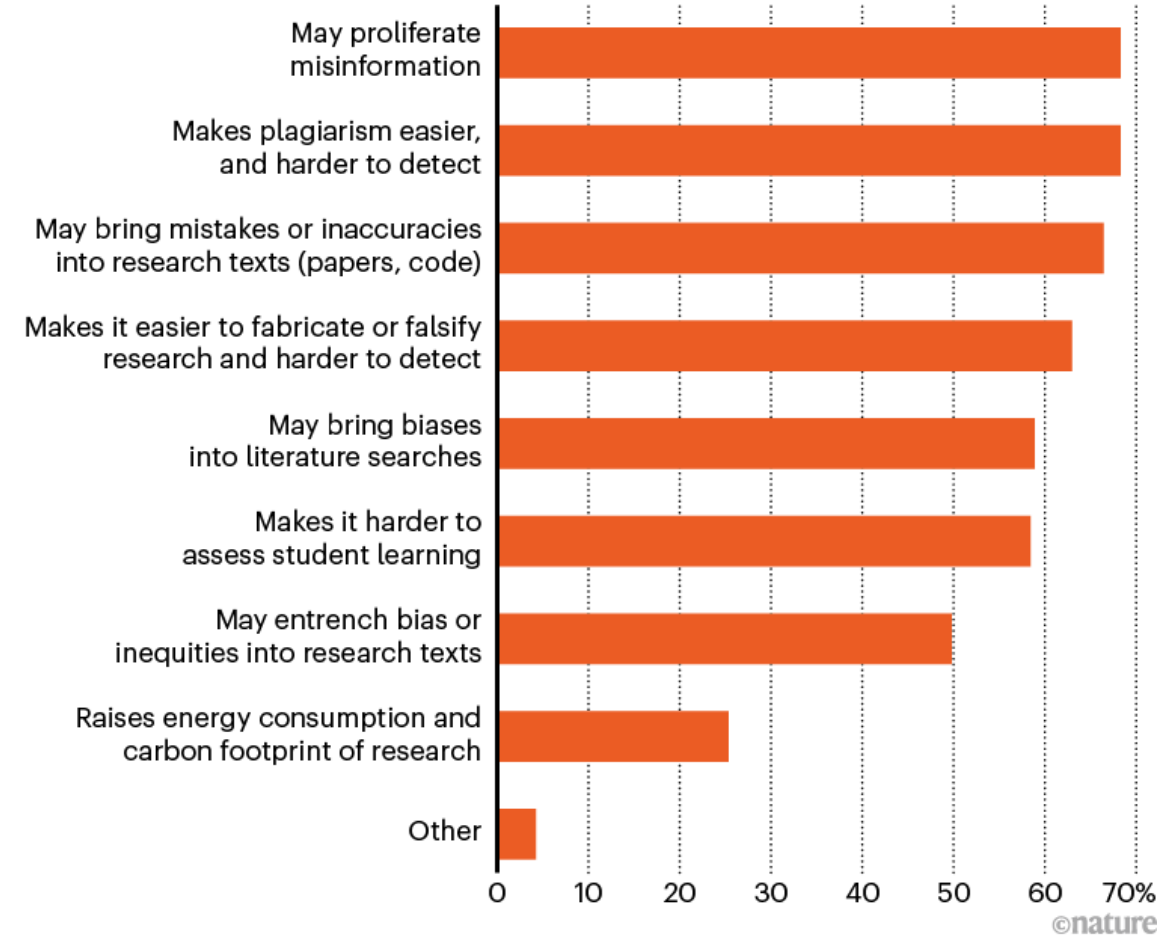
- Can Large Language Models Transform **Computational Social Science**? [article](#)
- Generative AI for Economic Research: Use Cases and Implications for **Economists** [article](#)
- Leaps and Boundaries: The Expert Panel on Artificial Intelligence for **Science and Engineering**. [report](#)
- Using large language models in **psychology**. [article](#)

Risks



Problems of Generative AI


Q: Where do you think generative AI may have negative impacts on research?
(Choose all that apply)



Practical Concerns

The immediate risks of using generative AI

Inaccuracy (a.k.a. hallucinations)

- GenAI models can present inaccurate information as fact
 - Even citations might be fictional
- Stochastic Parrots 
 - They sound good, but don't understand what they're saying

Data Privacy

- In general, publicly available models will
 - Upload & retain data to their own servers
 - Use data for model training
- Consider
 - Are there data licenses or laws you must follow?
 - Would you be upset if some portion of your prompt was used in a response to another user?

Model Bias

- Models can blindly perpetuate biases they learn from the internet
- Some models are “coached” to act in a certain way and (hopefully) minimize these harmful biases

Researcher Duty

We must hold ourselves (and others) accountable to these duties

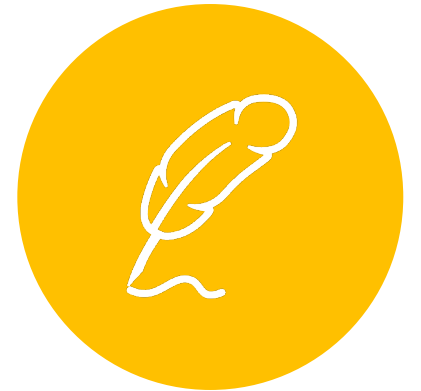
Duties

- Duty of **Verification**
 - DO verify the accuracy & validity of GenAI outputs
 - DO check for unintentional plagiarism



Duties

- Duty of **Disclosure**
 - DO document & disclose use of GenAI in all research stages



Duties

- Duty of **Discretion**
 - **DO NOT** assume GenAI is private
 - **DO NOT** share confidential, sensitive, or proprietary information with public GenAI
 - **DO NOT** assume GenAI output is already in the public domain



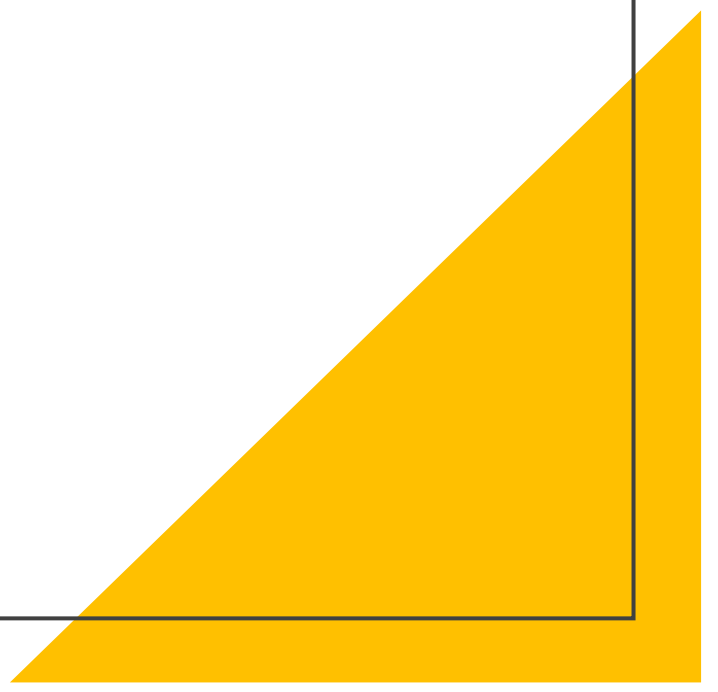
Bigger Picture Concerns

The potential long-term implications of using genAI

Bigger Picture Concerns

- How might widespread adoption of genAI impact scientific research?
- For example, we may consider
 - Will researchers gravitate towards certain research questions?
 - How will reward structures adapt to hyper-productivity?
 - How should mentorship & training change?
 - Will unique perspectives & voices be diluted?

Best Practices



Policies & Guidelines

a.k.a. following the rules

Current Trends

- Funding agencies and publishers continue to release policy on the use of generative AI
- In *general*, the trend seems to be

	Allowed	Conditions
Grants	✓	disclosure, human oversight
Publications	✓	disclosure, human oversight, limited to specific purposes
Peer reviews	✗	n/a

Canadian Federal Funding Agencies

- Awaiting official guidance from the tri-agencies
- Advice from ad hoc generative AI panel of external experts
 - Urged against blanket ban on gen AI in grant writing. Instead,
 - Make applicant ultimately accountable for application's content
 - Require disclosure in the use of gen AI in proposal preparation
 - Advised disallowing reviewers from unsanctioned use of gen AI

Publisher Guidelines

- Many academic publishers and journals have released policy on the use of Generative AI in academic publication.
 - [Nature](#), [Science](#), [Springer](#), [Elsevier](#), [Taylor & Francis](#)

Advice from the field

Learning from others

Advice from the field

- **Choose the right model**

- Each has trade-offs
- Some models specialize in certain task/discipline intersections



Note

Checkout this [blog post](#) for model comparisons

Advice from the field

- **Prompt engineering**

- Garbage-in, garbage-out
- A little familiarity goes a long way
- Frameworks can help guide prompts
 - e.g. Role-Task-Format



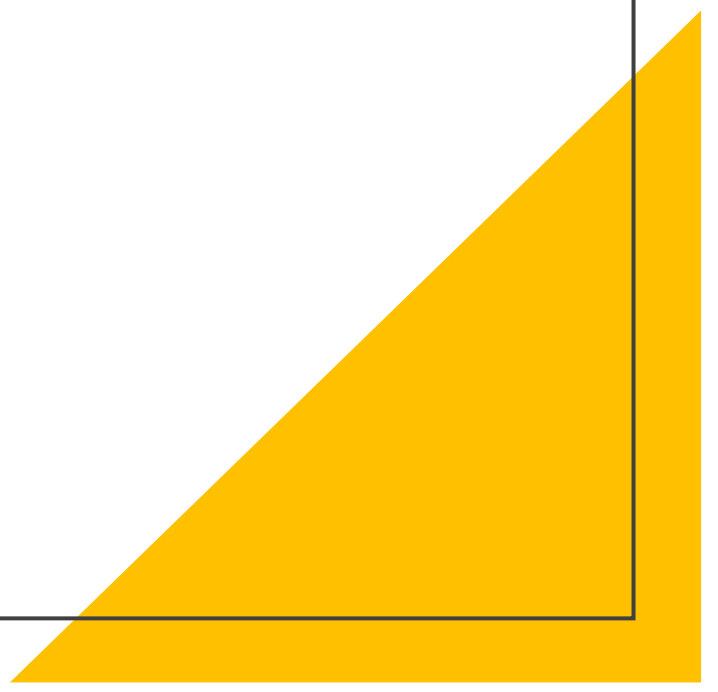
Situation/Scenario	Framework	Use When You Want To...
Clear Problem, Direct Solution Needed	PS (Problem, Solution)	Directly address a specific problem with a straightforward solution.
Role-Specific Task & Output	RTF (Role, Task, Format)	Tailor a response based on a specific role or perspective, with a clearly defined task and desired format.
Current State vs Desired Outcome	BAB (Before, After, Bridge)	Understand the transition from a current state to a desired state and the steps needed to get there.
Complex Problem with Underlying Challenges	SCQA (Situation, Complication, Question, Answer)	Analyze a situation with inherent complications and seek focused answers to specific questions.
Detailed, Contextual Response Needed	PTC-GO (Persona, Task, Context, Goal, Output)	Get a comprehensive response that considers the persona handling the task, the context, the goal, and the desired output.
Comprehensive Understanding of a Project or Task	5Ws 1H (Who, What, When, Where, Why, How)	Gain a holistic view of a project or task, covering all aspects from stakeholders to methodologies.

Advice from the field

- **Do not implicitly trust GenAI**
 - Don't directly use model output, instead ask for advice to incorporate yourself
 - Sounding board, rather than ghost creator
- Remember, you're responsible for anything you put your name to!



Conclusion



To Remember

- Generative AI isn't always the right tool
- When deciding whether to use genAI consider
 - Policies restricting/prohibiting this use of genAI
 - Risks & benefits to using genAI for this purpose
 - Whether it is important the work make use of your unique perspective, expertise, or opinions?

References 1/3

Advice from the ad hoc generative AI panel of external experts. (2024, January 12). Innovation, Science and Economic Development Canada.

<https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/use-generative-artificial-intelligence-development-and-review-research-proposals/advice-ad-hoc-generative-ai-panel-external-experts>

Bazarova, N., Belot, M., Elemento, O., Joachims, T., Li, A., MacRae, B., Mimno, D., Placanica, L., Rush, A. M., Shu, S., Warner, S., & You, F. (2023). *Generative AI in Academic Research: Perspectives and Cultural Norms* (p. 33). Cornell University. Retrieved from

https://it.cornell.edu/sites/default/files/itc-drupal10-files/Generative%20AI%20in%20Research_%20Cornell%20Task%20Force%20Report-Dec2023.pdf

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>

References 2/3

- Clune, J., & Murphy, G. (2024, February 15). *Pitfalls and Potential: Integrating Generative AI into your Research and Scholarship*. <https://research.ubc.ca/pitfalls-and-potential-integrating-generative-ai-your-research-and-scholarship/recording>
- Dursi, J. 2024, February. RCT #175 - Digital Humanities' LLM Edge. Research Computing Teams, 175. Retrieved from <https://newsletter.researchcomputingteams.org/archive/rct-175-digital-humanities-llm-edge-plus-eoling-a-8332/>
- Ihsan, A. (2023, November 15). *6 ChatGPT Prompt Frameworks I wish I'd had from the start*. <https://www.linkedin.com/pulse/6-chatgpt-prompt-frameworks-i-wish-id-had-from-start-arслан-ihsan-ihxkf/>
- Mollick, E. (2023, September 16). *Which AI should I use? Superpowers and the State of Play*. https://www.oneusefulthing.org/p/which-ai-should-i-use-superpowers?publication_id=1180644&utm_campaign=email-post-title&r=3ecf2t&utm_medium=email

References 3/3

- Prillaman, M. (2024). Is ChatGPT making scientists hyper-productive? The highs and lows of using AI. *Nature*, 627(8002), 16–17. <https://doi.org/10.1038/d41586-024-00592-w>
- Van Noorden, R., & Perkel, J. M. (2023). AI and science: What 1,600 researchers think. *Nature*, 621(7980), 672–675. <https://doi.org/10.1038/d41586-023-02980-0>
- Why scientists trust AI too much—And what to do about it. (2024). *Nature*, 627(8003), 243–243. <https://doi.org/10.1038/d41586-024-00639-y>
- Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., & Yang, D. (2023). *Can Large Language Models Transform Computational Social Science?* (arXiv:2305.03514). arXiv. <https://doi.org/10.48550/arXiv.2305.03514>

Questions?

