

Using Generative AI in Research

Kim Frail & Janice Kung

University of Alberta Library

kim.frail@ualberta.ca

janice.kung@ualberta.ca / @janicekung



This work is licensed under a Creative Commons
[Attribution-NonCommercial-4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

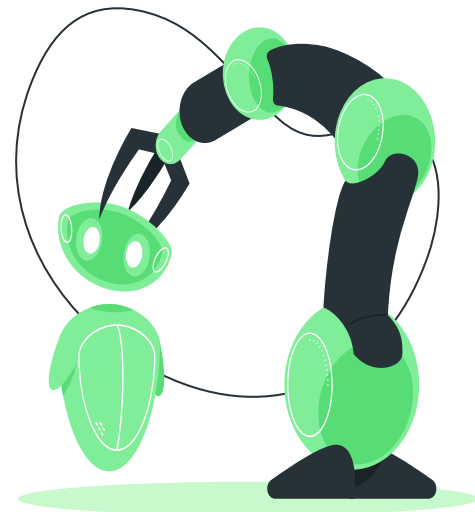


Table of Contents

1

Generative AI

An introduction

3

Key Considerations

2

Generative AI in Research

Research Life Cycle

4

Problems & Evaluating Outputs



How does generative AI work?

AI Create, image generated in response to "Create an image with a robot thinking intently in a futuristic cityscape," <https://aicreate.com/text-to-image-generator/> July 13, 2024.

Taking a step back...

LLMs Create Tokens

GPT-3 Codex

Here is a sentence broken into tokens. Notice how most words are their own token, but that sometime names or complete words like ChatGPT, or mispelled words, and , punctuation: , "are their" own tokens! As a rule of thumb for expressions in the English language there are approximately 750 words per 1000 tokens.

Clear

Show example

Tokens

68

Characters

312

Here is a sentence broken into tokens. Notice how most words are their own token, but that sometime names or complete words like ChatGPT, or mispelled words, and , punctuation: , "are their" own tokens! As a rule of thumb for expressions in the English language there are approximately 750 words per 1000 tokens.

Basic Text Prediction

➡ The students studied in the _____

Word (Token)	Probability
Waterpark	1%
Attic	7%
Bedroom	50%
Library	99%

**Calculator... but for
words!**





**Grounded vs.
Ungrounded**

Grounded vs. Ungrounded

Grounded

Connects the model to external source of knowledge.

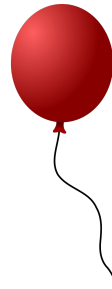
Produce outputs based on specific, real-world data or established contexts.



Ungrounded

Model relies only on its training data.

Generates outputs without strict adherence to factual data or real-world context.



Ungrounded



ChatGPT

DALL-E

Image generation

Sora

Video generation



Midjourney

Image generation

Grounded



Perplexity



Microsoft Copilot

Powered by OpenAI



Google Gemini

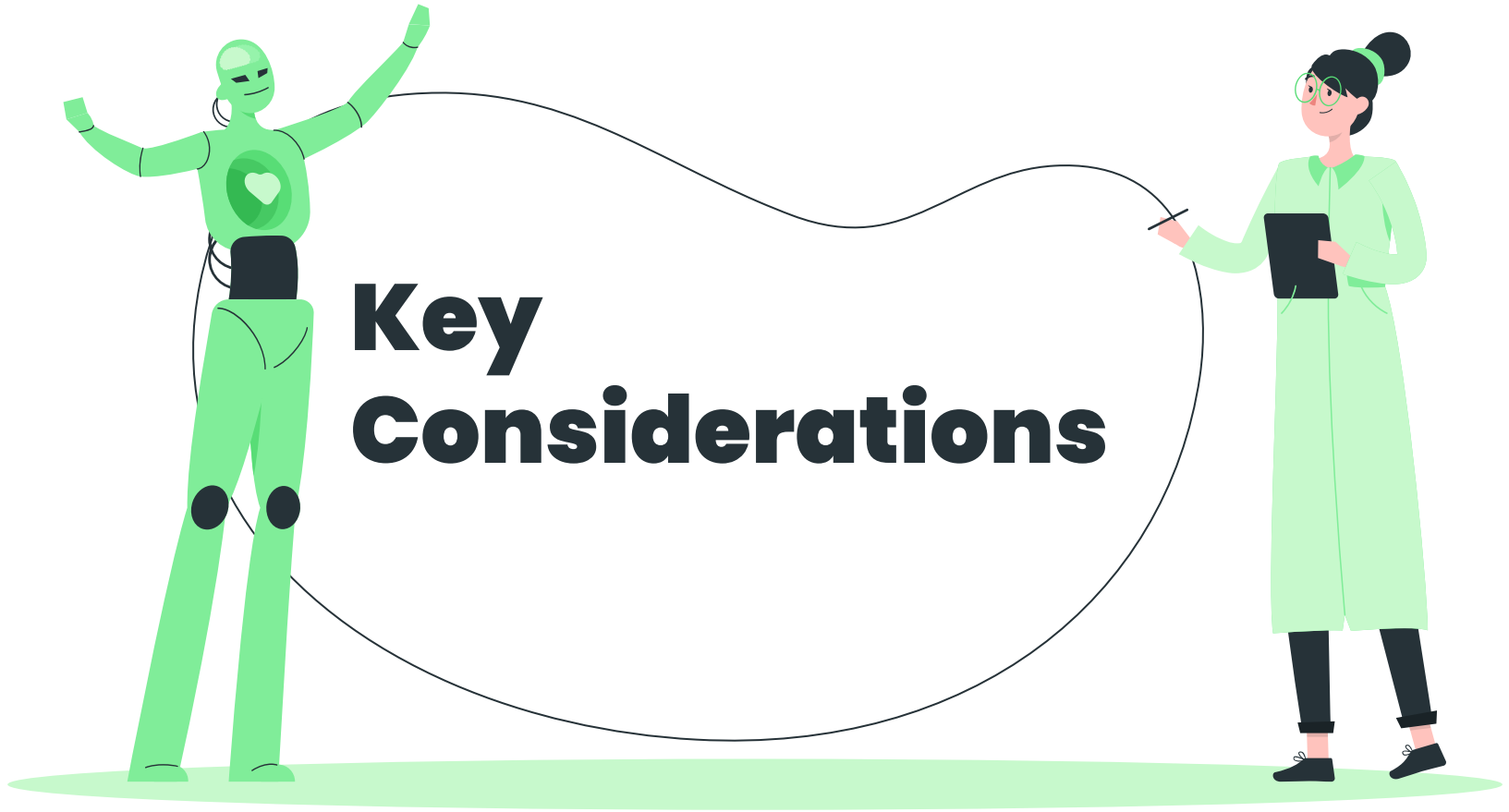
Formerly Bard, now available in Canada!

ChatGPT (ungrounded)

- Data Sources:
 - Common Crawl (60%)
 - WebText2 (web pages shared in Reddit posts)
 - Books
 - Wikipedia
- Free version does not require an account
- Suitable for creative writing, brainstorming, or learning about established topics.

Perplexity.ai (grounded)

- Data sources from the web in real-time:
 - Microsoft Bing
 - Academic journals (PubMed, Semantic Scholar, NIH databases)
 - Research papers
 - Social media platforms
 - News articles
 - Forums including YouTube and Reddit
- Free version uses a model based on GPT 3.5
- Includes in-text citations from the internet
- Upload files to extract key information
- Good for grey literature searching

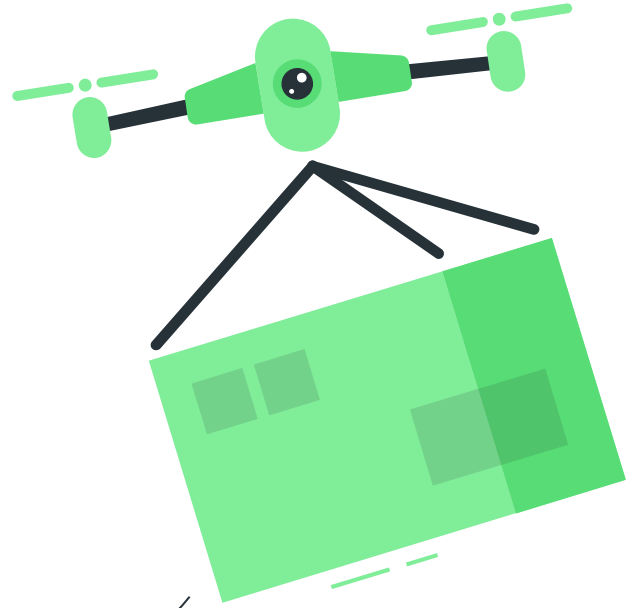


Key Considerations

Text-based Generative AI Tools

- **Step 1:** ask supervisor and supervisory committee re policies (e.g., GPS AI Guidelines, UofA Code of Student Conduct, Course/Dept specific, etc.,)
- **Step 2:** Think about publication/outputs *at the beginning* of your research process.
 - Look at journals, awards/scholarships/grants to see what their AI policies are
 - If allowed, how to document/disclose usage?

Research Life Cycle







ChatGPT

Writing	Literature Review Assistance
<ul style="list-style-type: none">● Generate ideas/brainstorming● Structuring and outlining● Citation and reference support● Editing and proofreading● Suggestions for databases to find literature on your topic <p>"Correct the grammar: [paragraph]"</p> <p>"Provide 3 examples for citing a government website in APA style."</p>	<ul style="list-style-type: none">● Can explain complex concepts● Summarize information● Background information● Translations <p>"Can you explain x in simple terms?"</p> <p>"Identify the gaps in the literature on [topic]"</p> <p>"Translate [this passage] into English"</p>

What are the potential applications of nanotechnology in medicine, such as targeted drug delivery, tissue engineering, and diagnostics?




Nanotechnology holds significant promise in medicine, offering innovative solutions across various domains. Here are some key applications:

1. Targeted Drug Delivery:

- **Precision Treatment:** Nanoparticles can be engineered to deliver drugs directly to diseased cells, minimizing damage to healthy cells and reducing side effects.
- **Controlled Release:** Nanocarriers can be designed to release drugs in a controlled manner over time, enhancing the efficacy of the treatment.
- **Crossing Biological Barriers:** Nanoparticles can cross challenging biological barriers, such as the blood-brain barrier, allowing for the treatment of previously inaccessible areas.

2. Tissue Engineering:

- **Scaffolds for Tissue Regeneration:** Nanomaterials can be used to create scaffolds that mimic the extracellular matrix, supporting the growth and regeneration of tissues.
- **Stem Cell Differentiation:** Nanotechnology  can influence the differentiation of stem cells into

Perplexity

What are the potential applications of nanotechnology in medicine, such as targeted drug delivery, tissue engineering, and diagnostics?

Sources




Emerging Applications of Nanotechnology in...

 ncbi.nlm • 1

Emerging Applications of Nanotechnology in...

 ncbi.nlm • 2

Applications of nanotechnology in...

 sciencedirect • 3

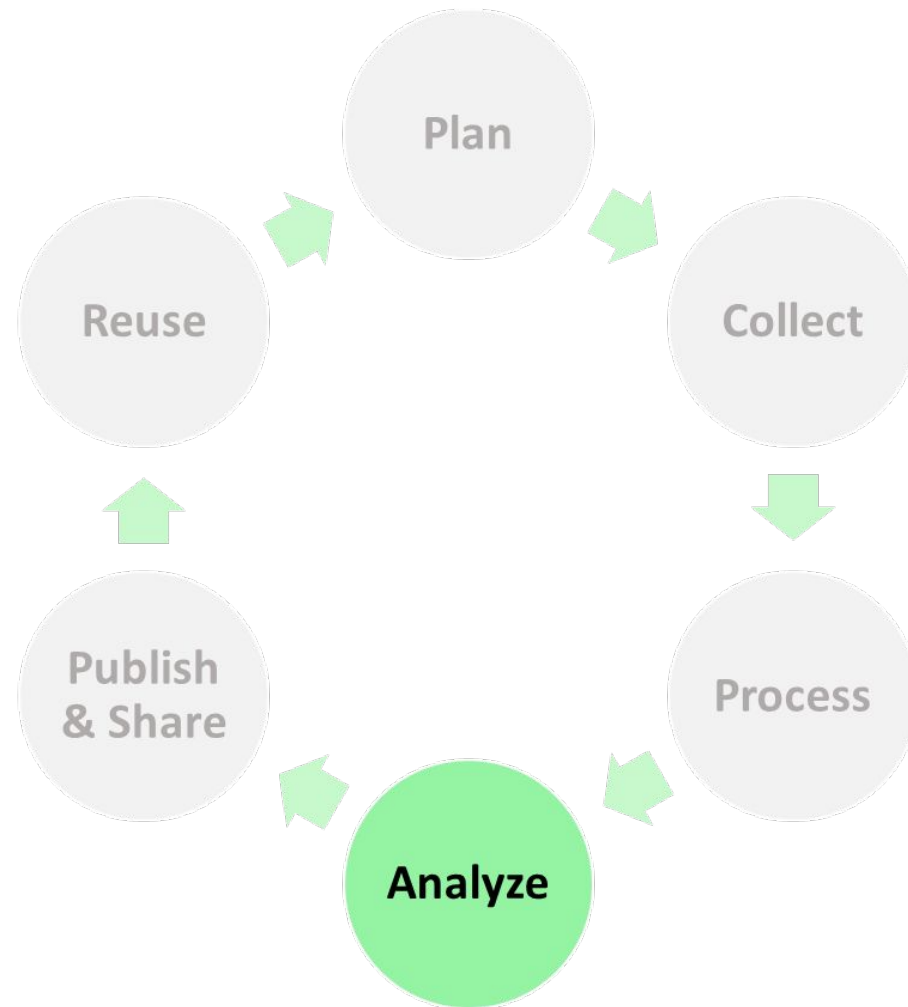


View 2 more

Answer

Nanotechnology has numerous promising applications in medicine, offering potential breakthroughs in diagnosis, treatment, and prevention of diseases. Some key areas where nanotechnology is being applied in medicine include:

1. Targeted drug delivery: Nanoparticles can be engineered to deliver drugs directly to specific cells or tissues, such as cancer cells. This allows for more precise treatment with reduced side effects compared to traditional drug delivery methods ¹ ². For example, nanoparticles can encapsulate chemotherapy drugs and deliver them specifically to tumor cells, minimizing damage to healthy tissue ⁴.
2. Diagnostics and imaging: Nanotechnology enables the development of more sensitive and accurate diagnostic tools. For instance:



Data Analysis

Essential to maintain academic integrity, conduct thorough research, and ensure final work represents your own ideas and efforts

Ensure compliance with institutional and granting body data privacy and research ethics standards

Some institutions are experimenting with local/offline or “secure” LLMs and Gen AI tools

Uploading of research data into proprietary Gen AI tools is not recommended

Prompts can be used to help you brainstorm ideas on how you could do your own data analysis

Data Analysis

Task	Example Prompt (do not upload actual data)
Interpret Statistical Results	"I conducted a regression analysis on my dataset to understand the relationship between variables X and Y. The results show a significant correlation. Can you help explain the implications of this correlation in practical terms?"
Understand Survey Responses	"I conducted a survey regarding patient preferences on vaccination locations. The responses are diverse and complex. Provide examples on how I can identify common themes or patterns within these responses?"
Data Visualization	"I have generated several visualizations from my dataset, but I'm not sure which type of visualization would best represent the relationship between these variables. Can you suggest suitable visualization techniques for this data?"
Check Assumptions in Analysis	"I'm working on a hypothesis test but unsure if my data meets the assumptions. Can you provide guidance on checking assumptions for this statistical test?"

I conducted a survey regarding patient preferences on vaccination locations. The responses are diverse and complex. Provide examples on how I can identify common themes or patterns within these responses?

ChatGPT



Identifying common themes or patterns in survey responses involves systematic techniques to organize and analyze qualitative data. Here are some examples of methods you can use:

1. Coding Responses

- **Step 1:** Read through the responses to familiarize yourself with the data.
- **Step 2:** Create codes (labels) for common words, phrases, or ideas that appear in the responses. For example, if many respondents mention convenience or distance, you can create a code for "convenience."
- **Step 3:** Group responses that have similar codes together. This helps to identify patterns, such as preferences for nearby locations, certain facilities, or transportation issues.

2. Thematic Analysis

- **Step 1:** Identify **recurring themes** based on respondents' language. For instance, if "accessibility" and "affordability" frequently appear, they can be considered major themes.
- **Step 2:** Organize these themes into broader categories. If responses mention public transportation, parking, or long travel times, you can group these under a "transportation" theme.

3. Cluster Analysis

- **Step 1:** Use software (like Excel, NVivo, or even Python) to organize responses and group similar answers.
- **Step 2:** Cluster similar responses together based on keywords or phrases. For instance, responses like "pharmacies are convenient," "local health centers are nearby," and "grocery stores are accessible" could be grouped into a cluster focused on **convenience**.

Check with supervisor/colleagues/librarian for further resources/guidance on suggested methods/tools

4. Word Frequency Analysis

- **Step 1:** Use word frequency tools to identify the most common words or phrases in the responses.
- **Step 2:** Words like "location," "doctor's office," "home," or "transportation" that appear frequently may point to common areas of concern or preference, and you can then delve deeper into how respondents discuss these terms.

5. Sentiment Analysis

- **Step 1:** Analyze whether respondents express positive, neutral, or negative sentiments toward certain locations.
- **Step 2:** For example, positive sentiment toward pharmacies might suggest a preference, while negative sentiment about doctor's offices could indicate frustration with wait times or scheduling.

6. Affinity Diagramming

- **Step 1:** Write each response on a sticky note or virtual board.
- **Step 2:** Group similar responses into clusters by physical proximity (e.g., clinics, pharmacies, homes). This can help visualize common patterns at a glance.

7. Quantitative Cross-Tabulation

- **Step 1:** If your survey contains quantitative data (e.g., Likert scale questions), cross-tabulate answers to identify relationships.
- **Step 2:** For example, you can analyze if respondents who rated "convenience" highly also preferred vaccination at pharmacies versus other locations.

Literature Analysis

- Summarizing and synthesizing information from scholarly articles (e.g., creating comparison tables)
- “Asking questions” about a particular scholarly article or group of articles
- Free versions but most have usage limits without a subscription

Elicit, Assistant by Scite, Consensus, etc.

Literature Analysis

Elicit

Q What impact does regular meditation practice have on stress reduction?

Summary of top 4 papers

Copy

Regular meditation practice has been consistently linked to stress reduction. Studies have shown that the Transcendental Meditation technique can lead to decreased physiological arousal, reduced anxiety, and improved general health (Alexander, 1993). Similarly, mindfulness meditation has been found to significantly reduce psychological symptomatology and increase a sense of control (Astin, 1997). These effects are further supported by the findings of Arora (2021), who highlighted the stress-lowering and cognitive function-amplifying effects of meditation. Lane (2007) also found that even brief meditation training can improve perceived stress and negative mood, with more frequent practice leading to better outcomes. These studies collectively suggest that regular meditation practice can be an effective tool for managing and reducing stress.

Sort: Most relevant

Filters

Export as

UPGRADE

16

Paper	Abstract summary	Methodology	Research gaps	Manage Columns
Effects of the transcendental meditation program on stress reduction, health, and employee development: A prospective study in two occupational settings C. N. Alexander +5 1993 · 119 citations DOI	Regular meditators improved significantly more than controls on multiple measures of stress and employee development.	Prospective study design comparing the effects of Transcendental meditation (TM) technique on stress reduction, health, and employee development between employees who learned TM and controls with similar characteristics.	Not mentioned (the abstract does not explicitly mention any research gaps or unanswered questions that the study aims to address)	<ul style="list-style-type: none">Study objectivesTheoretical frameworkResearch questionHypotheses testedFuture researchFunding source
Effectiveness of meditation programs in empirically reducing stress and amplifying cognitive function, thus boosting individual health status: A narrative overview R. Arora +1 Indian Journal of Health Sciences and Biomedical Research	Meditation potentially offers a modality which could boost all aspects of health in a cost-effective, low time-consuming manner to a wide spectrum of individuals.	The methodology involved a multistep approach for selecting articles based on specific keywords and shortlisting titles, abstracts, and full texts.	<ul style="list-style-type: none">- Need for larger sample sizes in future studies- Need for stronger study designs (e.g. more methods) in future studies- Need for more conclusive evidence on the effectiveness of meditation programs	<ul style="list-style-type: none">DatasetStudy countIndependent variablesDependent variablesMeasured variablesRegion

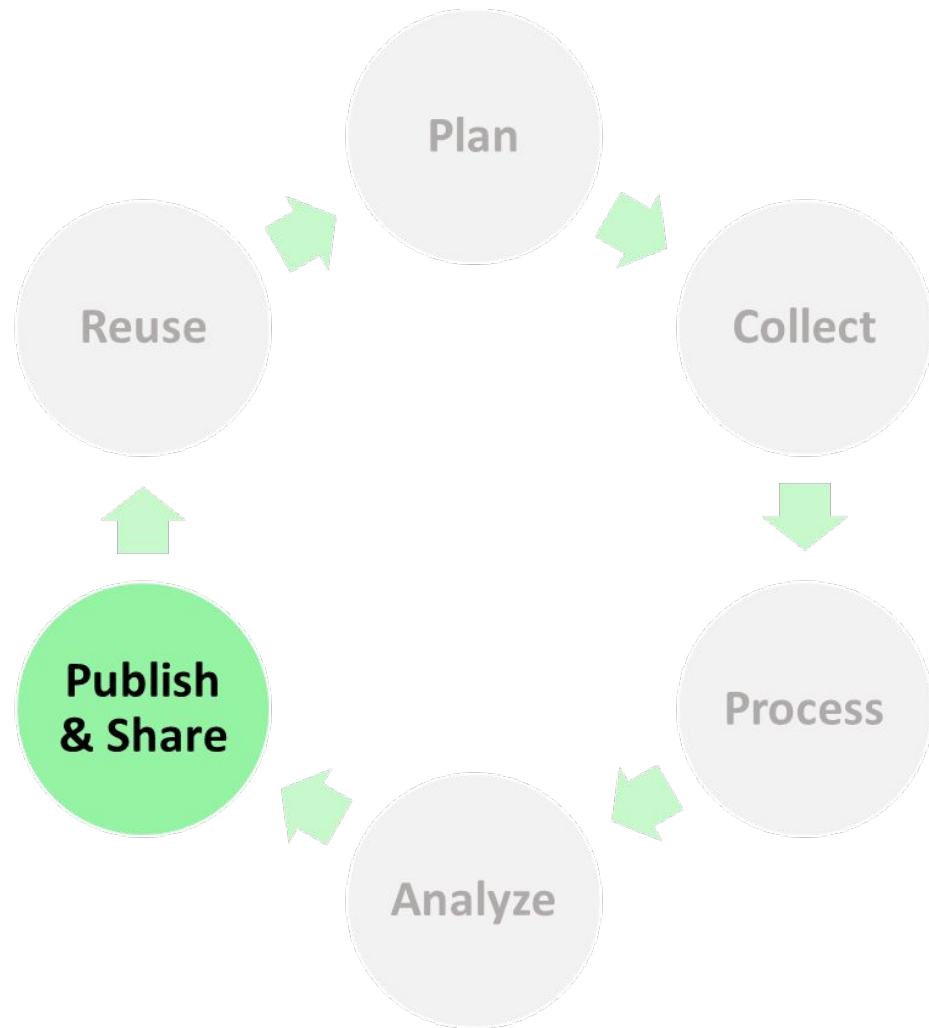
Literature Analysis Elicit

did participants identify the cause of their stress?

The papers provided do not explicitly mention whether participants identified the cause of their stress. These studies primarily focus on evaluating the effectiveness of meditation and mindfulness practices in reducing stress levels, improving health, and enhancing cognitive function among participants. They measure outcomes such as stress reduction, psychological symptomatology, sense of control, spiritual experiences, and cognitive improvements rather than focusing on the identification of stress causes by the participants. For instance:

- [Alexander \(1993\)](#) discusses the impact of the Transcendental Meditation program on stress reduction and employee development in an occupational setting, measuring physiological and psychological improvements without detailing if participants identified their stress sources.
- [Arora \(2021\)](#) provides a narrative overview of the effectiveness of meditation programs on stress reduction and cognitive enhancement, discussing the physiological and psychological mechanisms involved. The paper does not mention whether participants were asked to identify what was causing their stress.
- [Lane \(2007\)](#) evaluates the effects of a brief meditation training on perceived stress and negative mood, showing significant improvements in various psychological measures. The study does not specify if identifying stress causes was part of the intervention or assessment.
- [Astin \(1997\)](#) examines the effects of mindfulness meditation on stress reduction, psychological symptomatology, and spiritual experiences. The focus is on the outcomes of mindfulness practice rather than on whether participants identified the causes of their stress.

In summary, the primary aim of these studies is to assess the efficacy of meditation and mindfulness techniques in managing and reducing stress, rather than exploring the identification of stressors by the participants.



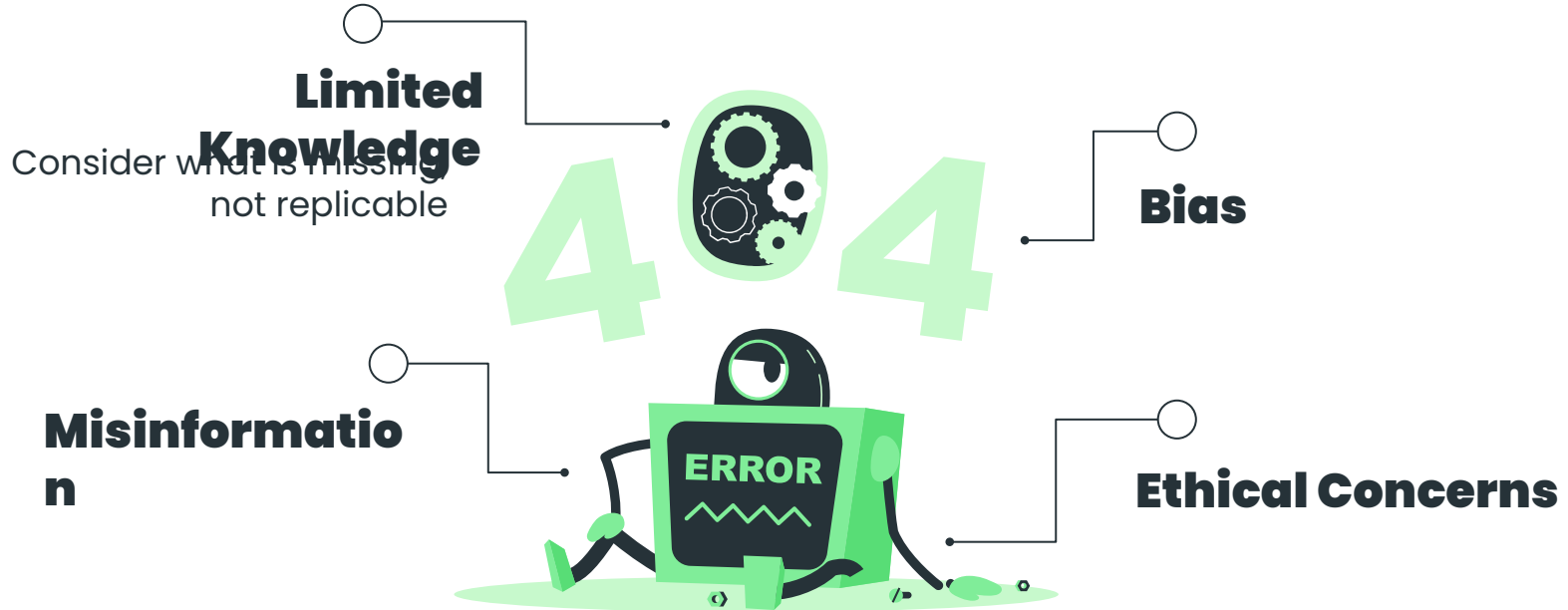
Publish & Share

Task	Prompt Ideas
Knowledge translation	<ul style="list-style-type: none">• Simplify research findings in layperson's terms• Create summaries for presentations, reports, conference abstract submissions• Design educational materials• Develop social media content to promote your research• Translate technical terms and jargon• Draft text and structure for research posters, ensuring the info is engaging and accessible
Data visualization	<ul style="list-style-type: none">• Recommend tools and software (e.g. Tableau, R (ggplot2), Python (matplotlib, Seaborn))• Brainstorm and outline key points for an infographic that visually represents complex data or research results

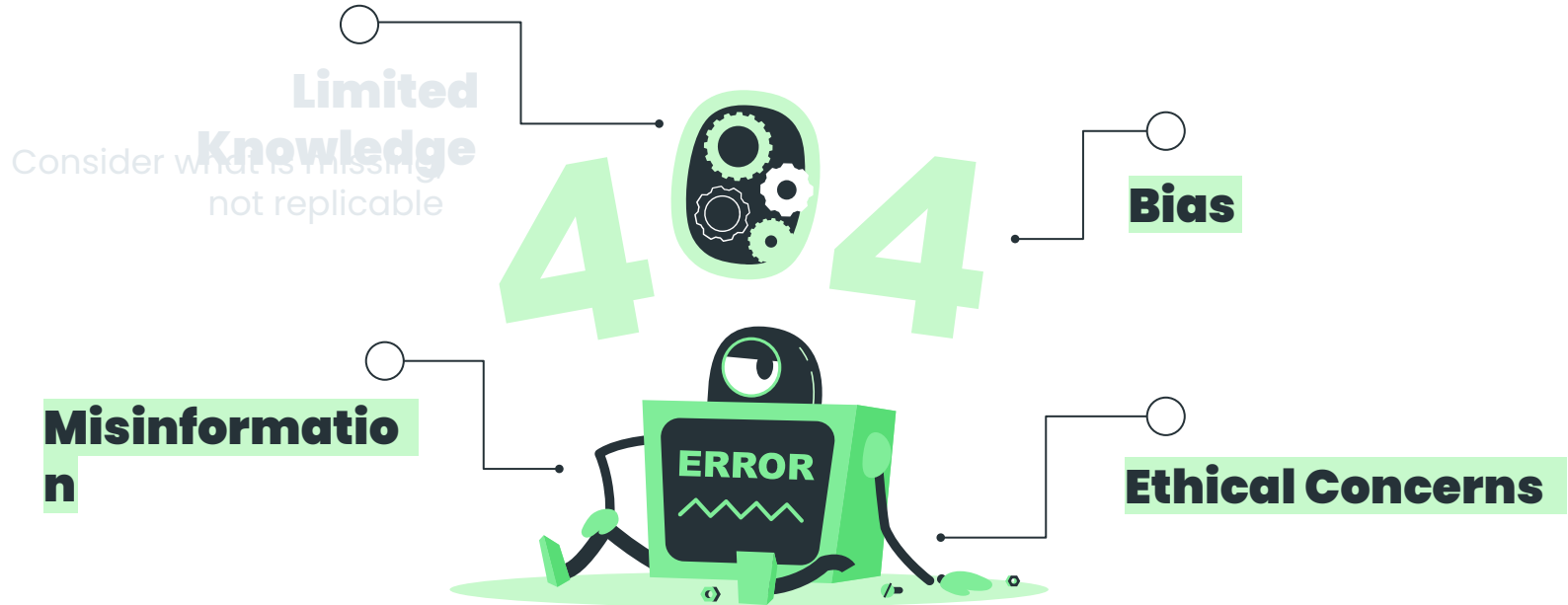


Problems & Evaluating Outputs

What are the Problems?



What are the Problems?



Misinformation

- Sometimes AI tools make stuff up, called **hallucinations**.
- Remember: based on the **probability** of what the next word should be
- **Replicability**: responses can vary from one day to the next
- Based on AI model's inherent biases and training dataset limitations

Bias

- Bias exists in the training data itself (sexism, racism, homophobia, ableism, etc.)
- Western, English focused (developed in the US)
- Not enough information or nuance for certain types of content in the dataset
- **Only humans** have the skillset to think critically, conduct nuanced research, and fact-check!

Ethical Concerns

- Do NOT share sensitive or confidential information
- Copyright
- Low wages to humans training ChatGPT
- Environmental costs
- Others?

BUSINESS • TECHNOLOGY

Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic



Perrigo B. Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic. January 18, 2023. <https://time.com/6247678/openai-chatgpt-kenya-workers/>

Evaluate Outputs



Find Better Coverage

✿ Answer

Based on the search results provided, here are some suggestions for finding current research on nanotechnology in medicine through the University of Alberta Library:

1. PubMed Central (PMC): While not specifically mentioned in the results, this free full-text archive is likely accessible through the University of Alberta Library and contains numerous articles on nanotechnology in medicine ².
2. ScienceDirect: This database is also likely available through the library and offers a wide range of scientific and medical research, including papers on nanomedicine ².
3. Web of Science: As a comprehensive citation database covering multiple disciplines, including nanotechnology and medicine, this resource is probably accessible through the university library ².
4. University of Alberta Library Catalogue: The library's own catalogue can be searched for relevant journals and resources on nanotechnology in medicine ³.
5. Open Journals: The University of Alberta Library hosts several open-access journals, some of which may publish research related to nanotechnology in medicine. For example, the Canadian Journal of Applied Research might include relevant articles ⁵.
6. Faculty of Science Resources: The University of Alberta's Faculty of Science has a strong focus on nanoscience and nanotechnology. Their website may provide links to relevant publications or research groups working on medical applications of nanotechnology ¹ ⁴.
7. Specialized Databases: The library likely provides access to specialized databases in nanotechnology and medicine. You can check the "Databases A-Z" list on the library website for subject-specific resources ².

While the search results don't provide a comprehensive list of nanotechnology-specific journals, these resources should serve as good starting points for finding current research on nanotechnology in medicine at the University of Alberta Library. For more targeted assistance, it would be advisable to consult with a librarian or use the library's research guides, which may have more specific recommendations for this interdisciplinary field.

Please suggest some reliable databases or journals where you can find current research on nanotechnology in medicine at the University of Alberta Library

UAL

Subject Guides

Ask Us / Find a Subject Librarian

Other Text-Based Generative AI Tools

Scholarly Content:

- Consensus
- Elicit
- ResearchRabbit
- Undermind
- ResearchGPT (queries Consensus)

General Content:

- Google Gemini
(requires personal Gmail account, UofA gmail won't work)

Using Generative AI Guide

<https://guides.library.ualberta.ca/generative-ai/home>

Using Generative AI

Search this Guide

Search

version française

Home

Types of Generative AI

Deciding Whether to Use AI

Ethical Considerations

Which AI Tool to Use

How to Use AI Tools

Evaluating AI Content

How to Cite AI

Additional Resources

About This Guide

NOTE: This guide focuses on students. Instructors who want to address AI in their teaching should refer to the [Teaching in the Context of AI toolkit](#) by the U of A's Centre for Teaching and Learning.

Generative artificial intelligence (AI) tools are evolving incredibly quickly, and they are having a significant impact on education and research. This guide provides information about using generative AI in ethical, creative, and evaluative ways. It focuses on five key areas:



Types of generative
AI



Ethical & responsible
use of AI



How to use AI
tools



Evaluating AI content



How to cite AI

Do you have feedback about this guide? [Let us know!](#)

Learning Outcomes

At the end of the session you will be able to:

1. Demonstrate a basic understanding of how AI tools work
2. Differentiate between grounded and ungrounded AI tools
3. Identify key considerations for grad students/researchers
4. Identify ways AI tools can be used to support the phases of the research lifecycle
5. Identify main areas of concern with using AI tools
6. Outline the steps and potential resources for evaluating and citing AI output



Thanks!

Questions?

kim.frail@ualberta.ca

janice.kung@ualberta.ca

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik** and illustrations by **Stories**

Acknowledgement: Special thanks to **Robert Vance** for their contributions to this presentation.