



Issues in Generative AI and Library Resources

Graduate Diploma in Journalism

Joshua Chalifour

Journalism, Communication Studies, & Digital Scholarship Librarian

joshua.chalifour@concordia.ca

Agenda

1. Essential concepts
2. Chatting with AI
3. Activity: experiment and discuss results
4. More generative AI tools
5. Deepfakes
6. Activity: publisher policies
7. Rights and integrity

AI Terminology and Trends

GPT-3/4

ChatGPT

generative pre-trained transformer

machine learning

training dataset

synthetic media

Sources

What do we know about the training?

Dataset	Quantity (tokens)	Weight in training mix
Common Crawl (filtered)	410 billion	60%
WebText2	19 billion	22%
Books1	12 billion	8%
Books2	55 billion	8%
Wikipedia	3 billion	3%

Datasets used to train GPT-3

Common Crawl

<http://urlsearch.commoncrawl.org>

Wikipedia = English only

WebText2, Books1, Books2?

“Language Models Are Few-Shot Learners.” arXiv. <https://doi.org/10.48550/arXiv.2005.14165>.

Participants in training GPT-3

- 621 participants
 - 343 male
 - 271 female
 - 7 other
- 38 = average age
- Mechanical Turk workers
- Based in USA (no other demographic data)
- \$12 (USD) paid to each participant
- 1 hour = estimated time to do the task



Chatting with AI

Interaction during the chat session



If there is no mail to deliver on a particular day, a mail delivery person may find themselves with additional free time. In such situations, they can consider the following activities:



1. Administrative tasks: They can use the downtime to complete administrative tasks associated with their job. This might involve organizing paperwork, updating delivery logs, or preparing reports for their supervisor.
2. Training and professional development: They can utilize the time to enhance their skills and knowledge. They might engage in online training courses, read industry-related materials, or explore resources that can help them improve their efficiency or learn new techniques.
3. Vehicle maintenance: If they use a vehicle for deliveries, they can perform routine maintenance tasks. This includes cleaning the vehicle, checking the engine oil, tire

How did it produce this answer?

Is it a satisfactory response to the prompt?

Prompt Writing

15:28

✓ BOT Midjourney Bot pencil drawing of a small boy and a girl playing in the garden, nanny is watching them, Victorian England --v 5 - @Manny Calavera (Open on website for full quality) (fast)



Made with Midjourney
<https://discord.com/channels/662267976984297473/984632500875821066/1120435066460446801>

15:28

✓ BOT Midjourney Bot 1 year old Autumn fairy with short black hair, wearing a pink tutu, sitting on a branch, photorealistic --v 5.1 --q 2 - @KatieB (fast)



Made with Midjourney
<https://discord.com/channels/662267976984297473/984632500875821066/1120434963645476995>

Prompt Writing

7:31 ✓ BOT Midjourney Bot mona lisa sloth - Image #2 @placid tuna



Made with Midjourney
<https://discord.com/channels/662267976984297473/995431473828077618/1120465827456626748>

17:31 ✓ BOT Midjourney Bot <https://s.mj.run/7RZ6tND5cNs>
<https://s.mj.run/JZ3QZXWgExE> <https://s.mj.run/JLb5QDfiQDU> <https://s.mj.run/WDO78CctFDM>
<https://s.mj.run/mQzV5C1yhaE> Scene: low-angle shot of this character standing in the kitchen, wearing oversized round glasses with orange rim, wearing a white chef's hat on his head, dressed in a white chef's apron, preparing delicious meal at a kitchen hob, there is a cook's hat on his head, he is cooking delicious meal Style: Pixar, Disney, Lucas Arts, 3D --q 2 --s 750 --v 5.1 - @NeonPuffin (relaxed)



Made with Midjourney
<https://discord.com/channels/662267976984297473/995431233121161246/112046603440776316>

Prompt Writing Tips

Provide context and details. State things like:

- length of the output you want (or dimensions)
- focus
- style or genre
- list specific things you want included
- linguistic requirements

State or modify advanced parameters particular to the tool

Scharth, Marcel. 2023. "How to Perfect Your Prompt Writing for ChatGPT, Midjourney and Other AI Generators." The Conversation, February 23, 2023.
<http://theconversation.com/how-to-perfect-your-prompt-writing-for-chatgpt-midjourney-and-other-ai-generators-198776>

More tips: <https://www.promptingguide.ai>

ChatGPT can provide a lot of variety. It could write code, jokes, extract keywords from text, etc. So be descriptive.

Start simple, iterate and augment or refine. A chat tool will continue working from your history.

Avoid ambiguity
Usually (not always) better to tell it to do something rather "not do" something

Activity 1: experiment with generative AI

1. Register a free text chat account

- ChatGPT <https://chat.openai.com>
- HuggingChat <https://huggingface.co/chat>
- Chatsonic <https://writesonic.com>

2. Generate a short essay with supporting academic references

- Iterate your prompt as needed
- Keep track of what you needed to do to get a “good” response
- Find and access 1 of the sources from a library database
- Reflect on the qualities of the output (sense, accuracy, helpfulness, etc.)


3. Copy and paste your prompt & AI response into this shared document.




<https://>

10 - 15 minutes and then we will discuss.



Discussion

 Write me a literature review of puppy and covid research from 2021, use MLA style for all references.

 I am unable to find any research studies specifically on "puppy and COVID-19" from 2021.  

However, I found some research studies that discuss the impact of the COVID-19 pandemic on pet ownership, including an increase in the number of people adopting puppies.

1. "The Impact of the COVID-19 Pandemic on Pet Ownership" by Dr. John Bradshaw and Dr. Alex German, published in the *Journal of Small Animal Practice* on May 2021, states that the pandemic has led to an increase in pet adoption, particularly of dogs and puppies. The study found that many people adopted pets to alleviate feelings of loneliness and isolation during lockdowns.
2. "The COVID-19 Pandemic and the Increase in Pet Adoptions" by Dr. Sarah Brown, published in the *Journal of Veterinary Behavior* on June 2021, discusses how the pandemic led to a surge in pet adoptions, particularly of puppies. The study found that many people adopted puppies as a way to cope with the stress and uncertainty of the pandemic.
3. "The Impact of the COVID-19 Pandemic on Animal Shelters and Pet Adoptions" by Dr. Jane Doe, published in the *Journal of Applied Animal Welfare Science* on August 2021, notes that the pandemic led to a significant increase in the number of people adopting puppies from animal shelters. The study suggests that the pandemic may have a positive long-term impact on animal welfare, as many of these puppies will be placed in loving homes.

References:

- Bradshaw, John, and Alex German. "The Impact of the COVID-19 Pandemic on Pet Ownership." *Journal of Small Animal Practice*, vol. 62, no. 5, 2021, pp. 256–261.
- Brown, Sarah. "The COVID-19 Pandemic and the Increase in Pet Adoptions." *Journal of Veterinary Behavior*, vol. 29, 2021, pp. 21–25.
- Doe, Jane. "The Impact of the COVID-19 Pandemic on Animal Shelters and Pet Adoptions." *Journal of Applied Animal Welfare Science*, vol. 24, no. 3, 2021, pp. 233–239.

Problematic? Useful?

Is it true there was no research to be found?

What/where did it “search”?


What happens when we look for this?




“A wire-fox-terrier as a scientist researching in a lab; in the style of a black-and-white 1920s-era marketing illustration” generated with DALL-E 2 and NightCafe.studio

← → ↻ 🏠 🛡️ 🔒 https://dair-community.social/@emilybende ☆ 🔍 Search

mastodon

◀ Back 

 **Emily M. Bender (she/her)**
@emilybender@dair-community.social

2- [#ChatGPT](#) etc are designed to create confident sounding text. If your thought is that you throw in some ideas and then try to evaluate what comes out, are you really in a position to do that evaluation? If it sounds good, are you just gonna go with it? Minutes before the submission deadline?

>>

Dec 14, 2022, 09:53 · 🌐 · Web · 🔄 6 · ★ 32

How do you get a foothold of knowledge in a subject?



More Tools with Generative AI

how does social media effect canadian political discourse?



SUMMARY OF TOP 4 PAPERS

BETA

The papers suggest that social media has had a significant impact on Canadian political discourse. Raynauld (2016) found that social media played a key role in the 2012 student movement in Quebec, with protesters using Twitter to share news reports and commentary to support their positions. Hampton (2017) found that social media use is negatively related to deliberation in many offline settings, and that social media users with the strongest attitudes on an issue are less likely to join political conversations in some offline settings. Halpern (2013) examined the potential of social media to foster democratic deliberation and found that Facebook expands the flow of information to other networks and enables more symmetrical conversations among users. Finally, Gruzd (2018) found that

Add information about all papers

Intervention

Outcomes measured

Paper title

Abstract summary



Political protest 2.0: Social media and the 2012 student strike in the province of Quebec, Canada

Vincent Raynauld, Mireille Lalancette, Sofia Tourign...

2016 19 Citations [DOI](#)

Social media-fueled mobilization initiatives have quickly mobilized narrow segments of the public in several advanced liberal democracies.



Social media and political discussion: when online presence silences offline conversation

Keith N. Hampton, Inyoung Shin, Weixu Lu

2017 109 Citations [DOI](#) [PDF](#)

Soci
conv



Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression

Daniel Halpern, Jennifer Gibbs

Comput. Hum. Behav.

2013 412 Citations [DOI](#)

Soci
delib



The State of Social Media in Canada 2017

A. Gruzd, Jenna Jacobson, Philip Mai, Elizabeth Du...

SSRN Electronic Journal

2018 65 Citations [DOI](#) [PDF](#)

Soci
man

Click tasks for other tools

Brainstorm research questions

Input

How has bus accessibility changed urban centres in Canada?

Clear unstarred



- ☆ How has bus accessibility affected the design and history of urban centers in Canada?
- ☆ How much has overall public transit use grown or shrunk in specific cities?
- ☆ How has the growth in public transit commuting changed the shape of cities?
- ☆ How has transportation demand shifted from car travel to mass transit?
- ☆ How do people currently get to and from places in Canada?
- ☆ What percentage of people in major Canadian cities commute by transit?
- ☆ How accessible are buses in Canada?
- ☆ How have transit agencies used their market power to set bus fares? * Has Uber and Lyft hurt urban bus system ridership? * Where are new low-ridership bus routes being introduced? * What percentage of buses now run late, compared with a decade ago?

<https://elicit.org> — academic orientation

Refine through chat.

As a search engine, Bing is able to provide URLs to sources.

Google Bard operates similarly.

<https://www.bing.com/chat>

The screenshot displays the Microsoft Bing Chat interface. At the top, the Microsoft Bing logo and 'CHAT' button are visible. A blue button prompts the user to search for 'Five most important resources on journalism ethics'. The chat history shows two messages:

Message 1: The chatbot confirms the search and provides a list of five resources: 1. IFJ Global Charter of Ethics for Journalists (International), 2. Accountable Journalism (International), 3. Society of Professional Journalists (USA), 4. European Federation of Journalists (Europe), and 5. MEAA Journalist Code of Ethics (Australia). It includes a 'Learn more' section with links to asatonline.org, mediaupdate.co.za, and '+2 more'. The progress indicator shows '1 of 30'.

Message 2: The user asks for links to the five most important articles. The chatbot responds that it cannot provide links due to a lack of consensus but offers to provide the resources mentioned earlier. The progress indicator shows '2 of 30'.

Message 3: The user replies 'yes'. The chatbot then provides the same list of five resources as in the first message. The progress indicator shows '3 of 30'.

At the bottom, there is a search bar with the placeholder text 'Ask me anything...' and a microphone icon.

Fakes?



Trying to detect synthetic media?

- OpenAI's text detection tool
- GPTZero - <https://gptzero.me>

Reverse image search tools

- tineye.com
- images.google.com



Roley, Gwen. 2023. "Deepfake Video Promotes Canadian Investment Scam." AFP Fact Check, May 4, 2023, AFP Canada edition.
<https://factcheck.afp.com/doc.afp.com.33ED4DW>.

Activity 2: policies about AI use

1. Form 3 – 4 teams
2. Go to the Google Spreadsheet:
 - **<https://>**
3. Select an organization (Elsevier, Wired, CBC, Science.org). Each has its own tab in the spreadsheet.
4. Read the *section(s)* of the **policy linked** to that organization.
5. Determine what the organization does and who its audience is.
6. List (in the spreadsheet) what sorts of AI usage are permitted or prohibited.

10 - 15 minutes and then share with the class what you found.

Citing, copyright, terms & conditions

MLA has [specific guidance](#), which includes acknowledging use for things such as editing or translating.

APA [recommends](#) personal communication style

Chicago [recommends](#) including the date, resource, etc. For example:

“1. Text generated by ChatGPT, March 7, 2023, OpenAI, <https://chat.openai.com/chat>.”

[OpenAI says](#) to disclose AI involvement and cite by stating something such as

“This article was generated with the help of GPT-3.”

Who is the author?
Who owns the
content?

OpenAI sharing and
publication [policy](#)

NightCafe says
[you own it](#) but maybe
not the © or TM.

“Text generated by an LM is not grounded in communicative intent, any model of the world, or any model of the reader’s state of mind. It can’t have been, because the training data never included sharing thoughts with a listener, nor does the machine have the ability to do that.”

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell.
“On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜.”
[https://doi.org/10.1145/3442188.3445922.](https://doi.org/10.1145/3442188.3445922)

Sources & Additional Information

Slide background image generated with VQGAN+CLIP, 30 June 2022, NightCafe, <https://creator.nightcafe.studio>.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜.” <https://doi.org/10.1145/3442188.3445922>.”

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. “Language Models Are Few-Shot Learners.” arXiv. <https://doi.org/10.48550/arXiv.2005.14165>.

Scharth, Marcel. 2023. “How to Perfect Your Prompt Writing for ChatGPT, Midjourney and Other AI Generators.” The Conversation, February 23, 2023. <http://theconversation.com/how-to-perfect-your-prompt-writing-for-chatgpt-midjourney-and-other-ai-generators-198776>

Huang, Haomiao. 2023. “The Generative AI Revolution Has Begun—How Did We Get Here?” Ars Technica, January 30, 2023. <https://arstechnica.com/gadgets/2023/01/the-generative-ai-revolution-has-begun-how-did-we-get-here/>

Mirsky, Yisroel, and Wenke Lee. “The Creation and Detection of Deepfakes: A Survey.” ACM Computing Surveys 54, no. 1, January 2022: 1–41. .

Quick Things for Digital Knowledge:

- [ChatGPT & Generative AI](#)
- [Machine Learning](#)

Google AI Glossary

Knowing Machines [reading list](#) to learn more about the broader implications of the training datasets.