Generative AI: Teaching Students the Applications, Risks, Best Practices, and Alternatives

Presenters: Tracy Coyne & Frank Sweis
Northwestern University Library

Collaboration team: Tracy Coyne, Jeannette Moss, and Frank Sweis

Conference track:
Climate Change:
Adapting to New Technologies
Colleague Acknowledgment:

Jeannette Moss, User Education Librarian

Frank Sweis, User Experience Librarian

Michelle Guittar, Dept. Head, Instruction & Curriculum Support
Learning Outcomes for this LOEX session:

- Understand the basics of what GAI is, how it works, the risks and limitations to students for using it for course assignments, and how library resources may provide better content;

- Know best practices for students to follow when using GAI effectively and responsibly;

- Hear the presenters’ reflections and takeaways from the instruction events they have run thus far.
What Motivated Us

August 2023
“Learners POV: A Student Panel on GAI in the Classroom” (two campus T & L units)

November 2023
Debut library session for students (our trio)
The content that we used for our student sessions begins with the next slide…
Generative Artificial Intelligence
What Students Need to Know
Learning Outcomes for this Workshop

By the end of this workshop, you will:

● Understand the basics of what GAI is and how it works
● Know how GAI is being used by students
● Be aware of GAI’s risks and limitations
● Understand when/why library resources may provide better content & quality
● Know the best practices for using GAI
Disclosure

Today’s presentation reflects current information. Things in the Artificial Intelligence world are changing rapidly. The same prompt may generate a different output in the future.
Past 5 years of “Generative AI” web search queries in the United States. (Source: Google Trends)
OpenAI timeline

**DECEMBER 2015**
The establishment of OpenAI in San Francisco

**NOVEMBER 2019**
The release of its successor, GPT-2

**NOVEMBER 2022**
ChatGPT was published

**JUNE 2018**
The launch of GPT-1, the first of OpenAI’s LLMs

**MAY 2020**
GPT-3 was made available

**MARCH 2023**
The launch of GPT-4, the latest model from OpenAI

Source: https://www.monterail.com/blog/what-features-is-gpt-4-missing
Hi can you tell me the timeline of released of Chat GPT?

Sure, here’s a brief timeline of the released versions of GPT:

1. GPT (Generative Pre-trained Transformer):
   - Released by OpenAI in June 2018.
   - Introduced the Transformer architecture for natural language processing.

2. GPT-2:
   - Released in February 2019.
   - Marked a significant advancement over the original GPT, with 1.5 billion parameters.
   - Generated widespread discussion due to concerns about potential misuse, leading to a staged release.

3. GPT-3:
   - Released in June 2020.
   - The largest model to date, with 175 billion parameters.
   - Generated even more attention and discussion due to its capabilities, versatility, and potential implications for AI development.

There may be updates, improvements, or smaller releases within these versions, but these are the main iterations of the GPT series up to my last update.
Bing Image Creator: A timeline of the various versions of GPT by OpenAI
Generative AI Defined

What GAI is
“...a word completion tool... predicting what the next word in a sentence should be so it can write a paragraph for you, what an image should look like based on a prompt.”
—Ethan Mollick


What GAI is not
“ChatGPT is not a research tool in the sense of looking up facts. It does not do any searching for information. It does not look for sources... Everything it writes to you is a generated, fabricated response that is triggered by your query.”
—Kiri

https://www.wkiri.com/

Artificial Intelligence
The ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.
https://www.britannica.com/technology/artificial-intelligence

Generative AI
An AI model that learns from training data and uses it to generate new content that resembles the original data.

Machine Learning
The process by which AI systems learn from data and improve their performance over time.

Large Language Models (LLMs)
AI trained on large quantities of text in order to interpret prompts and generate human-like text-based outputs. ChatGPT, Gemini, Bing and Claude are all examples.

Artificial Intelligence at Northwestern. (n.d.). What Instructors Need to Know: Artificial Intelligence at Northwestern - Northwestern University. Retrieved October 26, 2023,
Some ways students may be using GAI

- Write emails
- Generate ideas
- Get an overview of a subject or topic
- Write code
- Translate
- Improve writing style and syntax
Considerations

- Your instructor may not approve of GAI
- Students need to check the syllabus or ask the instructor
- Always declare your use of GAI to your instructor
- Be aware of Northwestern rules against cheating and plagiarism

Northwestern University Academic Integrity: A Basic Guide
### Participants’ Poll

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have Used GAI?</td>
<td>22</td>
<td>3</td>
<td>25</td>
</tr>
</tbody>
</table>
# Participants’ Poll

<table>
<thead>
<tr>
<th>Types of Use</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Question</td>
<td>Personal</td>
<td>School</td>
<td>Both</td>
<td>Neither</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>How used?</td>
<td>6</td>
<td>1</td>
<td>12</td>
<td>6</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>
### Participants’ Poll

<table>
<thead>
<tr>
<th>Question</th>
<th>Allows</th>
<th>Prohibits</th>
<th>Not Sure</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor/Dept Policy</td>
<td>11</td>
<td>4</td>
<td>10</td>
<td>25</td>
</tr>
</tbody>
</table>
# Current summary of features: GAI Tools

<table>
<thead>
<tr>
<th>Tool</th>
<th>Can Search Internet</th>
<th>Create Code</th>
<th>Create Image</th>
<th>Find Scholarly Articles</th>
<th>Create Text</th>
<th>Create Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Gemini</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Microsoft Copilot</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>ChatGPT</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claude</td>
<td></td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consensus</td>
<td></td>
<td></td>
<td></td>
<td>yes</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Elicit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adobe Firefly</td>
<td></td>
<td></td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midjourney</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perplexity</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PlayHT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>voice creation</td>
</tr>
<tr>
<td>Research Rabbit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sounddraw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>music</td>
</tr>
<tr>
<td>Stable Diffusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Synthesia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>video</td>
</tr>
</tbody>
</table>
How do we interact with generative AI models?

- Chat/conversational UI
  - ChatGPT
  - Microsoft Sydney/Bing Create
  - Google Gemini

- Image and graphics generators
  - Adobe Firefly
  - Canva
  - Stable Diffusion
  - Midjourney

- Video/Audio
  - Synthesia
  - PlayHT
  - Soundraw

- Research
  - Elicit
  - Consensus
  - Research Rabbit
How do we interact with generative AI models?

- Integrated into applications
  - "Paired Programmer"
    - Microsoft CoPilot
  - Command Palettes
    - Notion AI (text generation based on localized content)
- Plugins
- Custom models

Source: OpenAI
Conversational UI

Until we perfect the mind meld, conversation (including verbal and non-verbal elements) is how people exchange information with many systems and with as little effort as possible. Conversation is the interface that most people know how to use, even if they find in-person human interaction occasionally awkward.

- Erika Hall, Conversational Design
Some positives

● Natural language prompts
● Remembers whole conversation
● Brainstorming
● Generates keywords
● Edit your writing
● Helps you phrase your thoughts
● Arranges in logical order
● Summarizes content
● Automates repetitive tasks
● Templates letters, forms, resumes, etc.
Some negatives / limitations

- Accuracy and incorrect answers ("hallucinations")
- False citations
- Outdated content (in some cases)
- What is / isn’t included in training data
- Fee vs. free, usage limits
- Plagiarism and copyright infringement
- Labor
- Privacy
- Doesn’t know, reason, discern, or think critically
- Bias…
Biases and harm

● Generative AI “black box” problem
  ○ The model’s internal systems are not visible or publicly known (e.g. algorithms, training data)

● These can be great tools for research and productivity, but users need to check and review for their failures, biases, and potential harm they can create

Next 3 slides: content warning
“Generally, AI programs that create images from a text prompt will draw from a massive database of existing photos and images that people have described with keywords. The results it produces are, in effect, remixes of existing content. And there's a long history of photos that depict suffering people of color and white Western health and aid workers.”
Washington Post: This is how AI image generators see the world

Prompt:

Toys in Iraq
are soldiers with guns
Researchers led by doctors at Stanford University ran nine questions through four AI chatbots — including OpenAI's ChatGPT and Google's Gemini — that are trained on large amounts of internet text.

- All four models used debunked race-based information when asked about kidney function and lung capacity, the study published Friday in *Digital Medicine* found. Two of the models gave incorrect answers about Black people having different muscle masses.
- To varying degrees, the models appeared to be using race-based equations for kidney and lung function, which the medical establishment increasingly recognizes could lead to misdiagnosis or delayed care for Black patients.

**Study: Some AI chatbots provide racist health info (Axios)**
Generative AI and Learning

● Potential negative effects on learning
  ○ Over reliance on AI
  ○ Impaired critical thinking
  ○ Accuracy of information
  ○ Superficial engagement with topic
  ○ Reduced human interaction
  ○ Demotivation

Writing Effective Prompts

Use the CLEAR method (created by Leo S. Lo):

C: Concise: Brevity and clarity (eliminate unnecessary words)
L: Logical: Structured and coherent
E: Explicit: Clear specifications
A: Adaptive: Flexibility and customization
R: Reflective: Continuous evaluation and improvement

Evaluating Information

A checklist...

Authority

- Who is author / creator? What credentials?
- If web page, does URL provide insight?: edu., gov., com., org.

Purpose

- To inform, educate, sell, persuade?

Accuracy and Bias

- Can claims be verified in other sources?
- Any info missing or anyone missing from conversation?
- Inherent biases?

Further info Evaluating Sources (NU Libraries)

Academic sources

- Authored by subject experts, scholars
- Bibliographies, references, footnotes
- Editorial process for quality, accuracy
- Books, book chapters, encyclopedias from academic presses
- Peer-reviewed academic journals
- Govt or research institute reports
- Discoverable through library catalogs/databases, Google Scholar, sometimes other search engines

Who were some of the earliest American women nature writers?

Early American women nature writers have made significant contributions to the genre of nature writing, often in the 19th and early 20th centuries. Here are some of the earliest American women nature writers:

1. **Susan Fenimore Cooper (1813-1894):** Susan Fenimore Cooper, the daughter of James Fenimore Cooper, was an American author and naturalist. She is best known for her work "Rural Hours" (1860), a nature journal that observed the flora and fauna around her home in upstate New York. This book is considered one of the earliest examples of American nature writing.

2. **Margaret Fuller (1810-1850):** While Margaret Fuller is primarily known for her work in the field of Transcendentalism and her role as a literary critic, her essays and letters often contain references to nature and the natural world. She was an early American feminist and thinker who touched on various topics, including nature, in her writings.

3. **Celia Thaxter (1835-1894):** Celia Thaxter was an American poet and writer known for her observations of the natural beauty of the Isles of Shoals, a group of islands off the coast of New Hampshire and Maine. Her writings, particularly her book "Among the Isles of Shoals" (1873), celebrated the region’s flora and fauna.

4. **Jane Johnston Schoolcraft (1800-1842):** Jane Johnston Schoolcraft, also known as Bah'mawowezhikagwah, was an Ojibwe writer and poet. She is considered one of the earliest Indigenous American women writers. Her works include traditional Ojibwe stories and poetry that often incorporate elements of the natural world.

5. **Gene Stratton-Porter (1863-1924):** As previously mentioned, Gene Stratton-Porter was an American naturalist, photographer, and author who wrote about the natural world, particularly in the American Midwest. Her books, including "A Girl of the Limberlost" and "Moths of the Limberlost," reflect her passion for nature.

6. **Mary Austin (1868-1934):** Mary Austin, also mentioned earlier, was an American nature writer who celebrated the landscapes and ecosystems of the American Southwest in her works.
ChatGPT, revised question

More specific prompt (explicit, adaptive)

Results appear accurate when checked against academic sources found via Google.

However…

Nature writing is a genre that has a rich history in American literature, and there have been many notable women who have contributed to this genre. While it's challenging to definitively identify the five earliest American women in nature writing, here are five notable women who made significant contributions:

1. Susan Fenimore Cooper (1813-1894): Susan Fenimore Cooper, the daughter of novelist James Fenimore Cooper, is often considered one of the earliest American women in nature writing. Her book "Rural Hours" (1850) is a classic of early American nature writing and is noted for its observations of the natural world in upstate New York.

2. Mary Austin (1868-1934): Mary Austin was an author and naturalist who wrote about the American Southwest and the California desert. Her work "The Land of Little Rain" (1903) is a seminal text in the genre, exploring the landscapes, flora, and fauna of the region.

3. Gene Stratton-Porter (1866-1924): Gene Stratton-Porter was an author, naturalist, and photographer known for her books on the natural world, particularly the wildlife and landscapes of Indiana. Her book "A Girl of the Limberlost" (1909) is one of her most famous works.

4. Florence Merriam Bailey (1863-1948): Florence Merriam Bailey was a prominent ornithologist and author who wrote about birds and nature. Her book "Birds Through an Opera Glass" (1889) is considered one of the earliest examples of American bird-watching literature.

5. Marjory Stoneman Douglas (1890-1998): Marjory Stoneman Douglas was a journalist and conservationist best known for her work in advocating for the preservation of the Florida Everglades.
Academic Sources

NUsearch keyword search on:

*american nature writers*

also found...

Mabel Osgood Wright (1859-1934)


Google Gemini cites sources... use with caution however

In spite of links to internet sources:

*wrong dates*
*wrong name*

Contributions in the late 18th and early 19th centuries:

1. **Margaret Bayard (1758–1849)**: A keen observer of the natural world, Bayard detail her life on a Pennsylvania estate, filled with rich descriptions of flora formally published during her lifetime, her writing offers a fascinating glimpse of women's engagement with nature.

Also checked against

- [Margaret Bayard Smith](https://www.monticello.org/research-education/thomas-jefferson-encyclopedia/margaret-bayard-smith/)
- [Margaret Bayard Smith papers](https://www.loc.gov/collections/margaret-bayard-smith-papers/about-this-collection/)
1. **False citation**
   Google found different book title, different author
   Sideris co-edited another book about Carson.

2. **False citation.** Google finds blog entry for exhibit:
   Environment & Society Portal

   Souder wrote a different book “On a Farther Shore” about Carson

3. **Accurate citation.**
   https://www.rachelcarson.org/about-linda-lear

---

**Verify citations:**

- NUsearch
- Google
- Google Scholar
- Open Worldcat (for books)
ChatGPT for search strategies

First attempt:
Suggest ways I may search for scholarly journal articles on gender identities of Latina/o peoples over the 20th century.

3. Search with Keywords:
   • Use your identified keywords and concepts to conduct searches in these databases. For example, you might use a search query like: "Latina/o gender identities 20th century."

Offers keyword and Boolean Operator suggestions along with databases.
*No records found*

Second attempt: (iterate the search: Boolean)
Generate a Boolean search using three synonyms from Latina/o synonymy and three from the synonyms for gender identity.

Certainly, here's a Boolean search query using three synonyms from the list of Latina/o synonymy and three from the list of gender identity synonymy:

(Latino OR Chicano OR Latinx) AND (Gender Identity OR Gender Expression OR Non-Binary)

Slide and images credit:
Michelle Guittar, Head of Instruction & Curriculum Support, NU Libraries
Elicit

Training data is scholarly papers from Semantic Scholar. Semantic Scholar’s site states:

“Semantic Scholar covers all STM and SSH disciplines including biology, medicine, computer science, geography, business, history, and economics.”

**Not clear how representative the training data is of ALL academic disciplines

Elicit features:

● Discovers scholarly sources based on your prompt
● Extracts data from pdfs
● List of concepts

Got a warning: low on credits (after only 4 searches)

Paper #1: book rev. article led to login screen for Choice Reviews
Paper #3: “paper not found” in semantic scholar. Google search lead to a university site with a download, collection of essays
Paper #4: book rev. article lead to U of C Press, login page
Library Databases Compared to GAI

Deep Web vs. Surface Web

Real, accurate citations

Subject / disciplinary-focused:

- Authored by subject experts, checked by editors
- Library databases published by academic publishers, organizations or societies:
  - MLA Internatl Bibliography, Modern Languages Association (Literary studies, Philology)
  - PsycINFO, American Psychological Assoc. (Psych)
  - ERIC, Institute of Educational Services, Dep. of Education (Edu)
  - ACM Digital Library, Assoc. Of Computing Machinery (Comp Sci)

Keyword, Title, Author, Subject vs. Natural language searching

Unlimited use to NU community! No paywalls; No tokens/credits needed

Ask a Librarian or check Research Guides for further database suggestions!

Bing Image Creator

Easy to use, free

Example assignment: create a cover design for a children’s book

Example prompt for a course assignment: Create an image of an orange frog on a sailboat in a storm.
GAI creates images.

New, or a copy of something that already exists?
Example prompt for a course assignment: goal is to design a stage set for a play.

Prompt: Create an image of a tenement kitchen in New York City in the early 1900s.
Bing Image Creator

Bing produced these images which appear to be photos.

It is connected to the Internet but there is no source information.

These might be copyright protected.
Library of Congress:

Librarian suggested the Library of Congress to research image of tenements.

This photograph has source information.

place it in an painting frame and hang it inside a museum gallery
Adobe Photoshop, prompt: “in a museum being viewed by a group of people”
Scholarly Journals: Author Guidelines

Editorial

Using AI to write scholarly publications

Mohammad Hosseini a, Lisa M. Rasmussen b, and David B. Resnik c

dDepartment of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA; dDepartment of Philosophy, University of North Carolina, Charlotte, North Carolina, USA; eNational Institute of Environmental Health Sciences, Durham, North Carolina, USA

Article History Received 11 January 2023; Accepted 11 January 2023.

- Document use in Methodology and References section
- Specify which author/s used GAI
- Include the prompt/s verbatim

Mohammad Hosseini, Lisa M. Rasmussen & David B. Resnik (2023) Using AI to write scholarly publications, Accountability in Research, DOI: 10.1080/08989621.2023.2168535
Transparency and best practices for using generative AI

- Use of generative AI should be disclosed
- Possible guidelines
  - Disclose generative AI use in Methodology and References section of a scholarly publication
    - Who used the system
    - When was it used
    - What prompts were used
  - Authors accept responsibility and accountability for the content generated and that generated content should be checked by a domain expert for accuracy, bias, relevance, and reasoning
Current authorship guidelines of scholarly journals

- **Springer Nature** journals prohibit the use of Generative AI to generate images for manuscripts; texts generated by LLM should be well documented, and LLM is not granted authorship.

- **Science** journals prohibit the use of Generative AI to generate text; Generative AI-generated images and figures can be used only with explicit permission of their editors.

- **JAMA and the JAMA network** journals do not allow Generative AI to be listed as authors. However, Generative AI generated content or assistance in writing / editing are allowed in manuscripts but should be reported in the manuscript.

- **The International Conference on Machine Learning** prohibits content generated by Generative AI, unless it is part of the research study being described.

---

“Using Generative AI for Scientific Research” Michigan Institute for Data Science, University of Michigan
“While NLP systems are likely to become better at minimizing bias, doing math, making relevant connections between concepts, and avoiding plagiarism, they are likely to continue to make factual and commonsense reasoning mistakes because they do not (yet) have the type of cognition or perception needed to understand language and its relationship to the external physical, biological, and social world. NLP systems can perform well when working with text already created or curated by humans, but can perform (dangerously) poorly when they lack human-generated data related to a topic and try to piece together text from different sources. Thus, any section of a manuscript written by an NLP system should be checked by a domain expert for accuracy, bias, relevance, and reasoning.”

Mohammad Hosseini, Lisa M. Rasmussen & David B. Resnik (2023) Using AI to write scholarly publications, Accountability in Research, DOI: 10.1080/08989621.2023.2168535
How to Cite GAI

**APA style**

In-text: (Open AI, 2023); Reference list: OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model].

https://chat.openai.com/chat

**Chicago Manual of Style**

In-text CMOS author-date style example: (ChatGPT, March 7, 2023). Reference list for author-date style: not needed--cite in text only. For CMOS notes-bibliography style, see examples.

**MLA style**

In-text example, describe how you used AI (e.g.paraphrased from what ChatGPT generated): While the green light in The Great Gatsby might be said to chiefly symbolize four main things: optimism, the unattainability of the American dream, greed, and covetousness ("Describe the symbolism"), arguably the most important--the one that ties all four themes together--is greed.

Resources

NU academic integrity rules against cheating/plagiarism, including

Academic Integrity: A Basic Guide

NU libguide: Using AI Tools in Your Research

NU Library databases A to Z list

The Writing Place at NU
Acknowledgements

Resources beyond NU that we used for this workshop:


End Student-facing content
Reflections:

- After the first session, focused less on tools and more on potential risks, biases, harms.

- GAI changes were happening daily: company acquisitions and mergers, rebranding, move to paid model or limited tokens, addition of premium options.

- Institutional changes: addition of AI advisory board, AI guidelines and protocols, enterprise license for MS Copilot

- Instructor and student views toward GAI: runs the gamut from daily use to prohibition
Emphasizing Generative AI Risks, Biases, and Harm

- As we iterated on workshop, we reduced content that covered other GAI tools and broader landscape
  - Focus more on popular, more utilized tools
- Emphasized risk, biases, and harm inherit in these tools
- Library context appropriate environment to foster a critical discussion of GAI
  - Similar to analyzing and contextualizing library collections and metadata (e.g. subject headings)
- Students better understand their risks/limitations and can better assess these tools and their appropriate use
  - Library provided alternatives
Changing Landscape of Generative AI and Future Implications

● We spent a lot of time between sessions reviewing and updating the content of our workshop

● GAI is an area of rapid growth and regular product updates/releases

● Libraries may face constraints staying on top of the tools and updating their resources/policies to respond to changing landscape

● Future workshops will more closely focus on Northwestern's emerging policies and resources related to GAI
  ○ Focus on tools licensed by Northwestern
Thank You, LOEX Attendees!

- Conference proceedings paper on this content: will be posted by LOEX

Tracy Coyne
tracy-coyne@northwestern.edu