

# Using AI for Research & Data Analysis


**Molly Chehak**, Director of Digital Learning Pedagogy  
**Doireann Renzi**, PhD, Assistant Director of Faculty Initiatives  
**Yianna Vovides**, PhD, Senior Director of Learning Design & Research

March 14, 2024

# Today's Session

- Impact and Power of AI
- Research Examples & Demonstrations
- Data Analysis Examples & Demonstrations
- Practice
- Limitations
- Resources





# 1

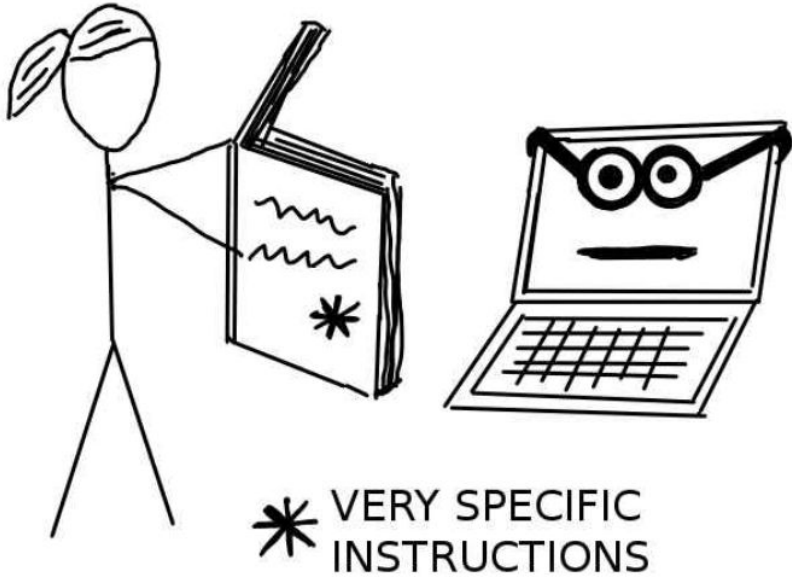
## The Impact of AI on Research

---

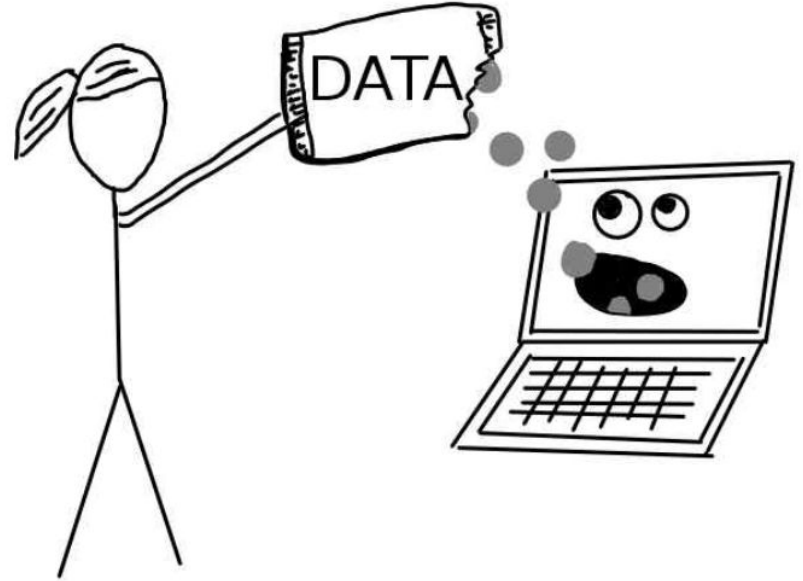
# How can you use AI in research?

<b>Idea Generation</b>	Help identify emerging trends and gaps in existing research by analyzing large volumes of scientific literature.
<b>Lit Review</b>	Tools like automated summarizers and citation network analysis software can significantly speed up the literature review process
<b>Data Analysis &amp; Analytic Coding</b>	AI algorithms excel at analyzing complex datasets, identifying patterns, and making predictions that would be challenging and time-consuming for humans to do manually.
<b>Experimentation</b>	Can simulate experiments, predict outcomes, and optimize research designs, saving time and resources in the lab.
<b>Writing &amp; Publication</b>	Aids in drafting research papers, checking for grammatical errors, ensuring adherence to publication standards, and even suggesting relevant journals for submission.
<b>Collaboration</b>	Can suggest potential collaborators based on research interests and expertise.

## Without Machine Learning



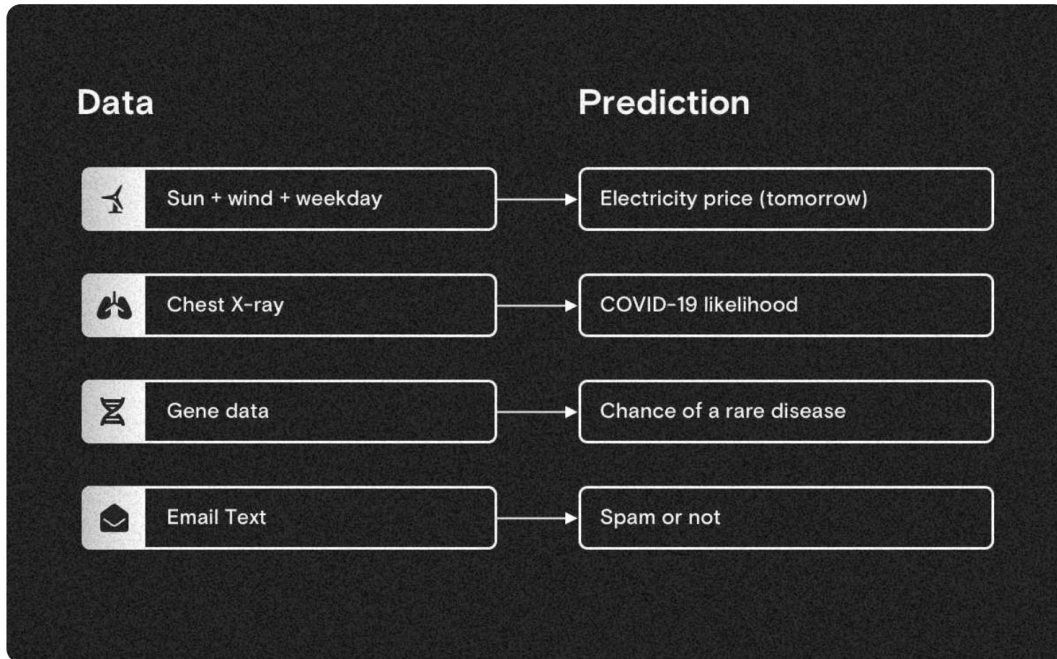
## With Machine Learning



# 1. Machine learning finds patterns in data

Simply, machine learning finds patterns in data and uses them to make predictions.

Whenever you have large amounts of data and want to **automate smart predictions**, machine learning could be the right tool to use.



Sample applications of machine learning: Turn data into predictions. Source author

**AI:** academic discipline founded in the early 1950s.

**Machine learning (ML):** study of algorithms that learn by experience; a subfield of AI.

**Deep learning:** a newer subfield of machine learning using neural networks; very successful in certain areas (image, video, text, and audio processing).

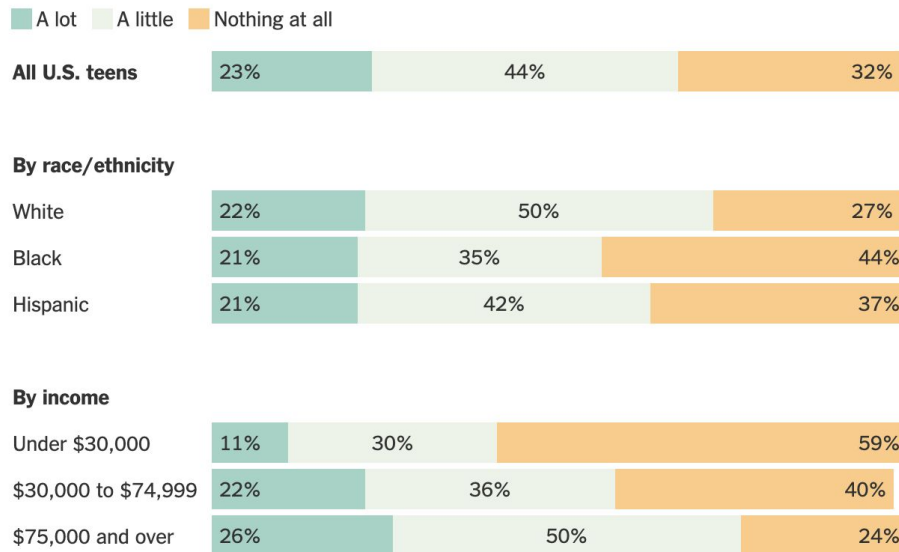
[Source](#)

# A complicated picture

December 2023

## Many Teens Have Never Heard of ChatGPT

How much, if anything, have you heard about ChatGPT, an artificial intelligence (A.I.) program used to create text?



Notes: Chart does not include respondents who did not respond. Respondents of other races are omitted because of low sample sizes. • Source: Pew Research Center survey of U.S. teens conducted Sept. 26 to Oct. 23, 2023 • By The New York Times

# Algorithmic Bias

AI-GENERATED IMAGES



Prompt: "Toys in Iraq"



"Playing soccer"



"Cleaning"



# Basic Sourcing Tools Demo

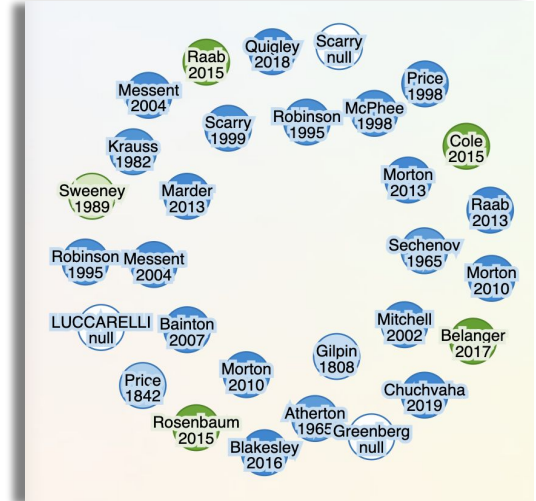
## AI TOOLS FOR RESEARCH



[PERPLEXITY](#)



[ELICIT](#)



[RESEARCH RABBIT](#)

# More examples and features

## More Tools

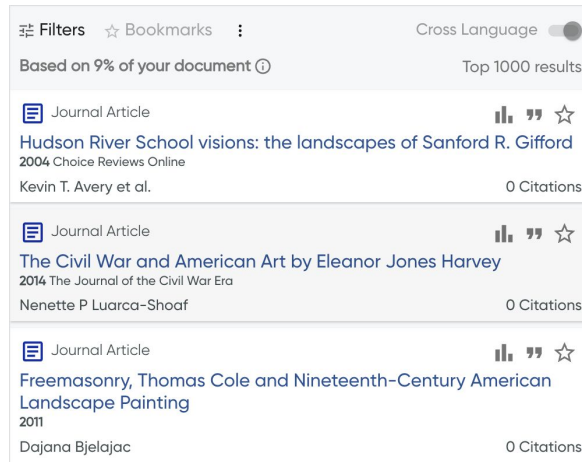
### Important Reminder

A result appearing at the top does NOT mean it is true. Interpreting research is complicated.

**Click the paper tile** to dig into the details of the study and find a link to the full text.










Thanks for joining our mission to make science more accessible for

[Consensus](#)

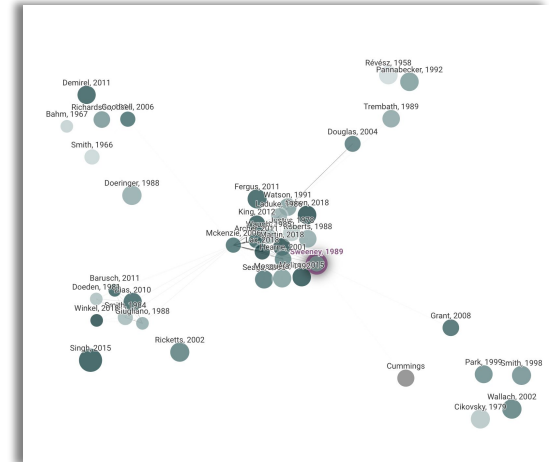


Filters ☆ Bookmarks ⋮ Cross Language

Based on 9% of your document ⓘ Top 1000 results

 Journal Article	  ☆
Hudson River School visions: the landscapes of Sanford R. Gifford 2004 Choice Reviews Online Kevin T. Avery et al.	0 Citations
 Journal Article	  ☆
The Civil War and American Art by Eleanor Jones Harvey 2014 The Journal of the Civil War Era Nenette P Luarca-Shoaf	0 Citations
 Journal Article	  ☆
Freemasonry, Thomas Cole and Nineteenth-Century American Landscape Painting 2011 Dajana Bjelajac	0 Citations

[Keenious](#)



[Connected Papers](#)

Bonus: [Chat PDF](#)

# Lean Library

An extension that allows researchers to connect directly with Lau's full-text resources even off-campus. Eg. If you're searching in Elicit and come across a paywalled resource, but the library has a subscription, the Lean Library box will pop up with a link directly to the catalog page for the resource.



**Lean Library**



**P**

Participant  
Practice

---



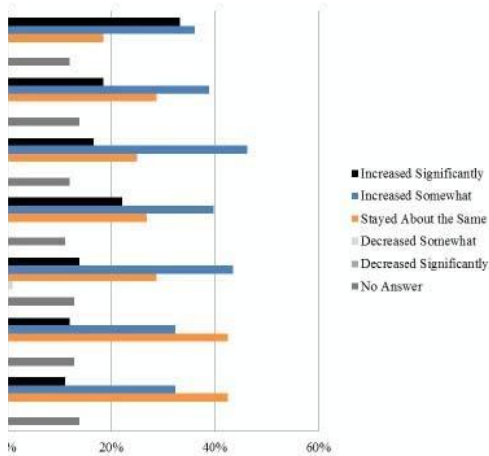
# 2

## The Impact of AI on Data Analysis

---

# Impact of AI on Data Analysis

## Task Types



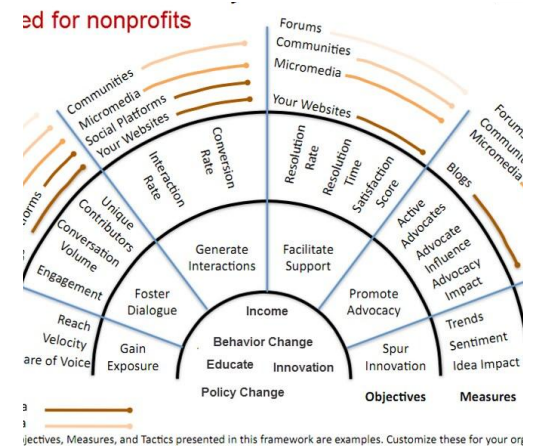
### Enhanced Data Handling + Predictive Analytics & Forecasting

Analysis possible of large datasets beyond human capacity with efficiency



### Increased Research Accessibility, Collaboration, and Customization

behavioral analysis, engagement metrics, optimization suggestions



### Facilitating Complex Decisions Fostering Interdisciplinary Approaches

Increased access leads to more data-driven decisions and allows for combination of data from multiple fields

# AI – your (& your students') coding assistant

Example in [Claude](#)

- Multiple tools will work
- Can walk student through decisions they have to make e.g. variable levels correct analysis, etc.
- Can actively generate code
- Can help troubleshoot errors

# AI's role in analysis and coding

---

- AI as a functional partner ONLY works once the fundamentals are required
- Think of - and train your students to think of - AI tools as  
- a very fast assistant coder whose work needs to be checked



FACULTY EXAMPLE

# Rebecca Helm

ASSISTANT  
PROFESSOR, EARTH  
COMMONS

Dr. Helm's students may use AI to:

1. Save time on program and coding introductions by [encouraging students to get familiar](#) before class.
2. [Write a program.](#)

*"I'm a scientist. I need to write a program in R to analyze a dataset. The data set is in CSV format. I want to map data onto a globe. I want the data points to be color coded based on a categorical variable. I want to use ggplots2. Write instructions for code. Go step by step."*

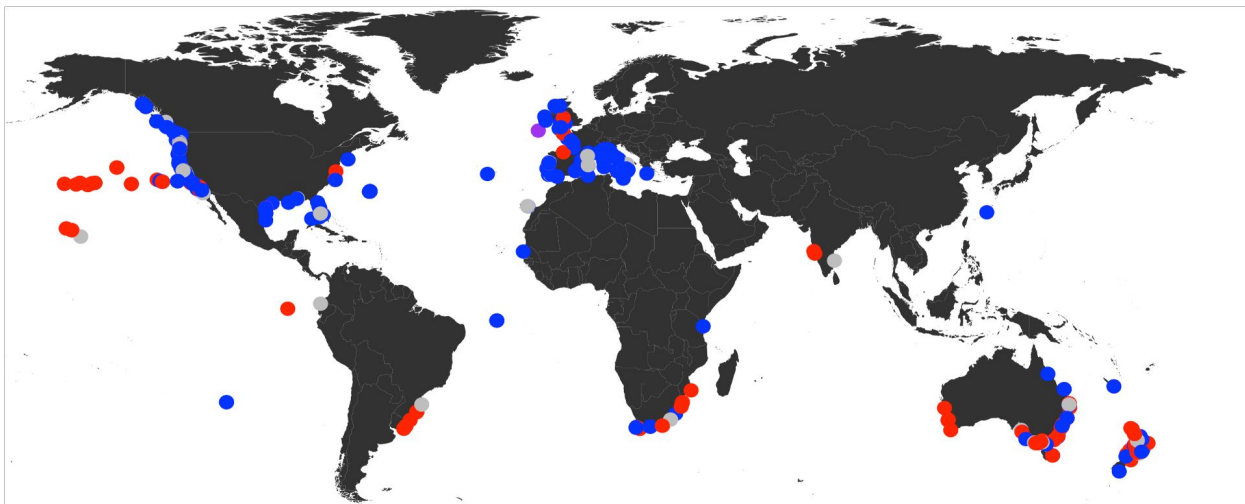
FACULTY EXAMPLE

# Rebecca Helm

ASSISTANT  
PROFESSOR, EARTH  
COMMONS

## Visualize data for better understanding

“Depending on the time in the semester and my other commitments, I don't always have time to code from scratch or modify old code through copy paste. Below, I used code generated by ChatGPT to generate a map of jellyfish occurrences for a quick discussion. I was able to generate this map in seconds, and it made the meeting more dynamic and highlights the power of quick coding support.”





# 3

## AI & Qualitative Analysis

---

# Data Analysis Prompts

---

[Conduct a thorough review and analytical comparison](#) of this Center for Teaching and Learning [website](#) against general standards and features observed in other centers for teaching and learning websites, such as [cndls.georgetown.edu](http://cndls.georgetown.edu), <https://crlt.umich.edu/>, and [cft.vanderbilt.edu](http://cft.vanderbilt.edu).

Address the specific details and aspects about the website in the following areas in your analysis:

- **Content Quality:** Evaluate the relevance, comprehensiveness, and currency of the information provided.
- **Usability and Design:** Assess the website's ease of navigation, aesthetic appeal, and user experience.
- **Resources and Services Offered:** Compare the range and quality of teaching and learning resources available, including any innovative tools or services unique to this site.
- **Engagement and Interaction:** Analyze how the website facilitates interaction with and between faculty, including forums, feedback mechanisms, and community-building features.
- **Accessibility:** Review the site's compliance with accessibility standards and its inclusivity towards diverse users.

\* [Summarize metrics](#)

\* Data visualization

\* Comparing data

# SENTIMENT ANALYSIS

## Student Example

“You are an analyst at a law firm that tracks regulatory affairs for its clients, and you have been assigned to monitor the responses proposed rule changes at the U.S. Securities and Exchange Commission (SEC).

You realize it would take you weeks if not months to review all the comments and analyze them in terms of positive, neutral, or negative.

What do you do?”

1. Scrape the data
2. Clean and organize - pre-processing
3. Run sentiment analysis (many packages, in this case text analysis)
4. Validate results
5. Share results



# 4

## Academic Integrity & Communication

---

# Communication is Key

- Transparency
- Always obtain permission before uploading the work of others
- Understanding of ethical considerations
- Balance automation and authenticity



# Academic Integrity & Citation Help

## APA Style (7th Edition)

While the American Psychological Association (APA) has not released official guidelines on citing generative AI quite yet, the recent [post on the APA Style Blog](#) provides guidance on citing ChatGPT adaptable to other AI tools.

### **In-text example:**

(OpenAI, 2023)

### **Reference example:**

OpenAI (2023). ChatGPT (May 24 version) [Large language model]. <https://chat.openai.com/chat>



# Tips for Academic Integrity & Citation Help

If you are publishing an academic paper which involves a study of AI:

- Publish your prompts in an appendix to the paper itself. Journals should require it as part of the methods. (If you are studying how other people use AI, you do not have to publish their prompts)
  - Use the latest methodologies, such as Chain of Thought, unless there's a reason not to do so. Reviewers should expect to see that, in most cases, researchers experimented with multiple prompts to see if it has an effect
- Papers should give the model/date/API and temperature (please, no more just saying "we used ChatGPT")
  - Run the prompt enough times to account for random variation
  - When a new, more capable model comes out, journals should allow some form of update where the prompts are tested against the new model to get a sense of boundary conditions & rate of change.

# Select Research & Data Analysis AI Tools

ACTIVITY TYPE	TOOL EXAMPLES
Idea Generation	<a href="#">ChatGPT</a> , <a href="#">Gemini</a> , <a href="#">Bing</a> , <a href="#">Claude</a> , <a href="#">IdeaNote</a>
Concept Mapping & Planning	<a href="#">Lucidchart</a>
Time/Project Management	<a href="#">krisp.ai/ai-meeting-assistant</a> , <a href="#">Goblin Tools</a> , <a href="#">Trello</a> , <a href="#">Asana</a>
Research Exploration & Literature Review	<a href="#">typeset.io</a> , <a href="#">Rayyan</a> , <a href="#">Explainpaper</a> , <a href="#">Connectedpapers</a> , <a href="#">Research Rabbit</a> , <a href="#">Scite</a> , <a href="#">Perplexity</a> , <a href="#">Semantic Scholar</a>
Summarizing	<a href="#">Glasp</a> , <a href="#">ChatGPT</a> , <a href="#">Gemini</a> , <a href="#">Bing</a> , <a href="#">Claude</a> , <a href="#">Smmry</a>
Outlining & Drafting	<a href="#">ChatGPT</a> , <a href="#">Gemini</a> , <a href="#">Bing</a> , <a href="#">Claude</a> , <a href="#">Kickresume</a> , <a href="#">textblaze.me</a> , <a href="#">Scrivener</a>
Collaborative Writing	<a href="#">Overleaf</a> , <a href="#">Google Docs</a> , <a href="#">Authorea</a> , Manuscripts.io
Reference Management	<a href="#">Zotero</a> , <a href="#">Mendeley</a> , <a href="#">EndNote</a> , <a href="#">Citavi</a>
Coding Help	<a href="#">hashnode.com/ai</a> , <a href="#">Fronty</a> , <a href="#">Tabnine</a> , <a href="#">debugcode.ai</a> , <a href="#">Stack Overflow</a>
Data Analysis & Statistical Analysis	<a href="#">GPT4</a> , <a href="#">Tableau AI</a> , <a href="#">Excel</a> , <a href="#">Google Workspace</a> (non-GU), <a href="#">Power.bi</a> , <a href="#">Rstudio</a> (tidyverse), (tools work with: SPSS, SAS, Stata, Python (pandas, NumPy))
Data Visualization	<a href="#">D3.js</a> , <a href="#">Plotly</a> , <a href="#">Infogram</a> , <a href="#">Datawrapper</a> , <a href="#">Adobe Illustrator</a>
Presentation Tools	<a href="#">Gamma</a> , <a href="#">SlidesGPT</a> , <a href="#">slidesgo</a> , <a href="#">Canva</a> , <a href="#">Prezi</a>

# Annotated List Research Tools

PLATFORM	APPLICATION SUMMARY
<a href="#">Research Rabbit</a>	Research Rabbit is a free online citation-based literature mapping tool that helps explore research by visually mapping out related articles and authors based on your starting point (seed papers).
<a href="#">Connected Papers</a>	Another free literature visualization tool that helps discover new research and track the development of a field. It focuses on the relationships between different papers rather than just the citations.
<a href="#">Phind</a>	A search engine specifically designed for academic research that can search across a variety of different databases and sources and uses machine learning to help you find the most relevant results.
<a href="#">Chat PDF</a>	Allows you to have a conversation with a computer about the content of a PDF document and can be useful for summarizing the document, finding specific information, or getting different perspectives on the material.
<a href="#">Academic Insight Lab</a>	A collection of tools and resources designed to help researchers with different aspects of their work, such as writing, publishing, and data analysis.

# Annotated List Research Tools

PLATFORM	APPLICATION SUMMARY
<a href="#"><u>Consensus</u></a>	A platform that helps researchers collaborate and share their work; allows you to create private groups, share documents, and have discussions with other researchers.
<a href="#"><u>Magic School</u></a>	Helps you learn about new research by watching short videos created by other researchers and can be a great way to get a quick overview of a topic or to find new research that you might be interested in.
<a href="#"><u>Perplexity</u></a>	Helps you understand the language used in academic research papers and can identify key terms, phrases, and concepts, and it can also show you how these terms are used in different contexts.
<a href="#"><u>Scite</u></a>	Allows you to read and discuss academic research papers in real time with other researchers and can be a great way to get feedback on your work or to learn more about a topic from other experts.
<a href="#"><u>Keenious</u></a>	Helps researchers find and manage their research data as well as store, organize, and share data with others.
<a href="#"><u>Elicit</u></a>	Helps write better research papers and also provides feedback on your writing style, grammar, and clarity.

# Courses on Deep Learning

## Short Courses

### [ChatGPT Prompting for Developers](#)

- Course provider: Deeplearning.ai with instructors Ilsa Fulford (OpenAI) and Andrew Ng (Coursera, Deeplearning.ai)
- Notes: This course gives a great introduction into how to structure your engagement with LLMs effectively ("prompt engineering"). Though the course does run provided code, I think its strength lies in its approachability. Deeplearning.ai offers a number of these free, short courses, including another on the OpenAI API co-taught again by Ilsa and Andrew.

### [Practical Deep Learning for Coders](#)

- Course provider: Fast.ai with instructor Jeremy Howard (Fast.ai, Kaggle)
- Notes: Seems mostly focused on the Fast.ai ecosystem, but recommended highly and frequently. Some materials did seem outdated.

### [Machine Learning Specialization](#)

- Course provider: Coursera with instructor Andrew Ng (Coursera, Deeplearning.ai)
- Notes: This course series is a staple in the ML community, but I think you would need to continue to the [Deep Learning](#) series, the sequel set of courses, in order to get a good handle on recent AI developments. On the other hand, if you are interested in all the algorithms that have run our lives for the last couple decades—recommender engines, text and image recognition, etc.—then this is the course for you. The ML specialization does require some comfort with math notation, statistical modeling, and basic python (familiarity with the numpy package is a plus).
- Challenging but very rewarding.

# Resources

---

## How To & Tools

[How Assistants work in Chat GPT](#)

Julius.ai

Keenious

Consensus

Elicit

Research Rabbit

[Gen AI Tool Tracker](#)

# Contact CNDLS

CNDLS Email  
CNDLS Website

[cndls@georgetown.edu](mailto:cndls@georgetown.edu)

[cndls.georgetown.edu/ai/resources](https://cndls.georgetown.edu/ai/resources)

---

# AI bots -- for online session

## OGE AI Statement

In alignment with our commitment to fostering a rich, interactive learning environment and in adherence to our mission to create opportunities for students to navigate international education structures and to develop essential interpersonal skills, we adopt a cautious yet open approach to the use of artificial intelligence (AI) technologies within our educational framework. While we recognize the potential benefits of AI in enhancing educational experiences, we prioritize the integrity of human interactions, especially in settings that involve highly personalized and confidential exchanges, such as advising sessions.

To this end, we strongly discourage the use of AI bots as stand-ins for students or faculty in online class meetings and advising sessions is strictly prohibited. This policy is in place to ensure the preservation of the personalized, confidential, and sensitive nature of these interactions, which are crucial to our educational mission. We aim to help students explore value systems and other underlying aspects of culture; this exploration is foundational to what we do in partnership with students. Thus, any deployment of AI technologies in our classroom, particularly those that might replace human engagement, must receive express permission from OGE staff. Our advising sessions cover a wide range of topics including:

- Academic advising, including the integration of international courses into Georgetown degree progression
- Health and safety information