





Al Literacy

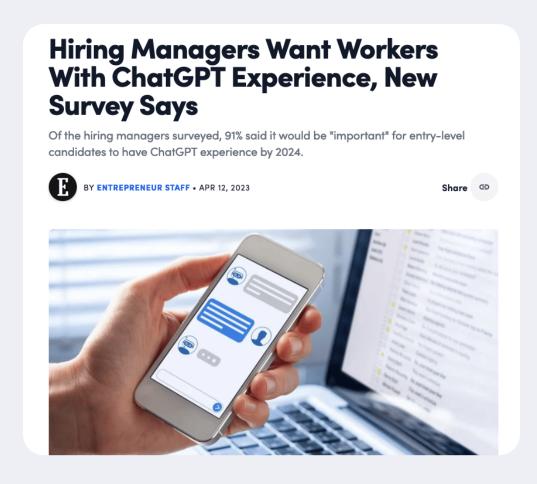
in the age of ChatGPT

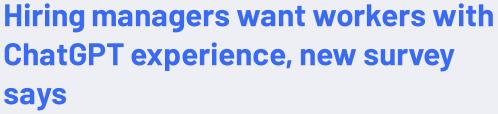


nicolehennig.com | nhennig@arizona.edu

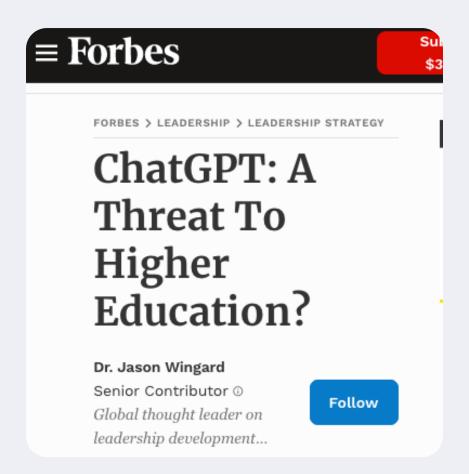
September 12, 2023

ChatGPT news is everywhere





- Entrepreneur



ChatGPT: A Threat To Higher Education?

- Forbes

Educators are in a difficult spot.

Here's what we'll cover.

- 1. The technology: how does it work?
- 2. How to use it effectively with ideas for teaching
- 3. Criticisms and ethical issues
- 4. The future

Discussion & questions



BIBLIOGRAPHY

You'll get

- A copy of these slides (PDF with links)
- A handout with more sources (PDF) https://bit.ly/ai-lit-writing

Nicole Hennig



Worked at the MIT Libraries for 14 years, first as webmaster, then head of UX department

2013 - 2018

My own business: technology training for librarians.

- books
- webinars
- online courses

Elearning developer at Univ. of Arizona Libraries 2019 to the present



I collaborate with my colleagues to create online tutorials.

What is Al literacy?

2020 paper, Long & Magerko



What is AI Literacy? Competencies and Design Considerations

Authors: Duri Long, Brian Magerko Authors Info & Claims

CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems • April 2020 • Pages 1–16 • https://doi.org/10.1145/3313831.3376727

"We synthesize a variety of interdisciplinary literature into a set of core competencies."

https://doi.org/10.1145/3313831.3376727



Based on those:

Al literacy is the ability to:

1 understand the basics of how Al works

2 use AI effectively and ethically

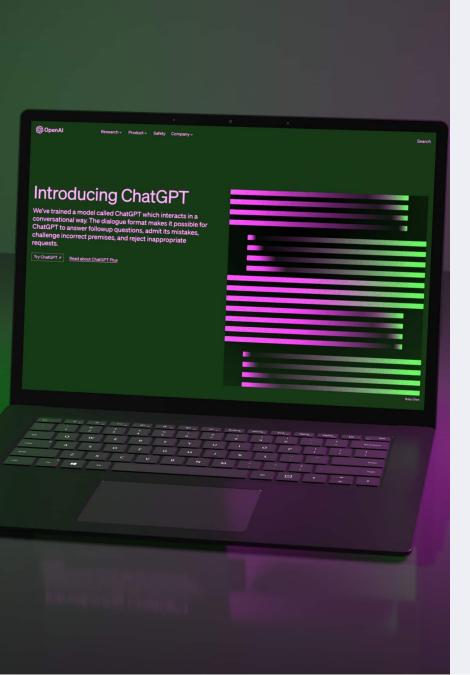
3 make informed decisions about using AI technologies

Why is Al literacy so important now?

1 It's becoming integrated into tools we use every day.

2 It's quickly becoming a skill that workplaces require.

- 3 It's closely connected to other literacies: digital, data, and media literacy and it's unique.
- 4 When we're Al literate, we can influence the future directions of this technology with ethics in mind.



Which of these describes your use of ChatGPT?

- Haven't used it
- Used it once
- Used it a few times
- Used it many times
- Use it every day

Part 1: The technology

What is AI?

What is a large language model?

What is generative AI?





"the field of study focused on creating machines that can perform tasks that typically require human intelligence."



These days Al is part of many services

- Siri and Alexa voice assistants
- Netflix or YouTube recommendations based on your viewing habits
- Gmail's spam filtering





Those are examples of narrow Al

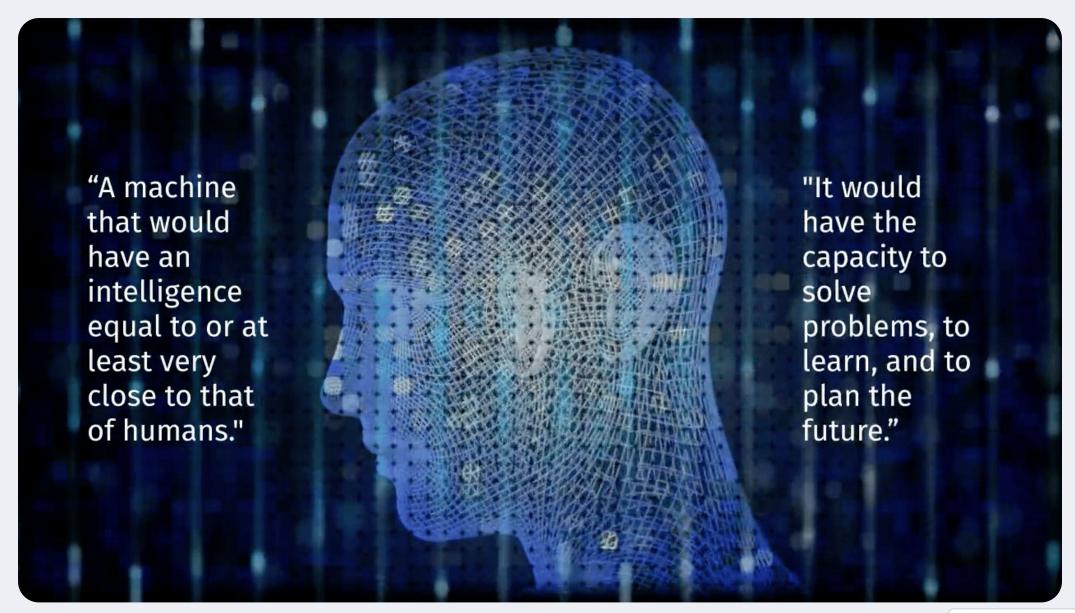
- Al designed to perform a specific task
- AlphaGo beat the human Go champion in 2017.



General Al

Artificial General Intelligence (AGI)

Experts don't agree on a precise definition.





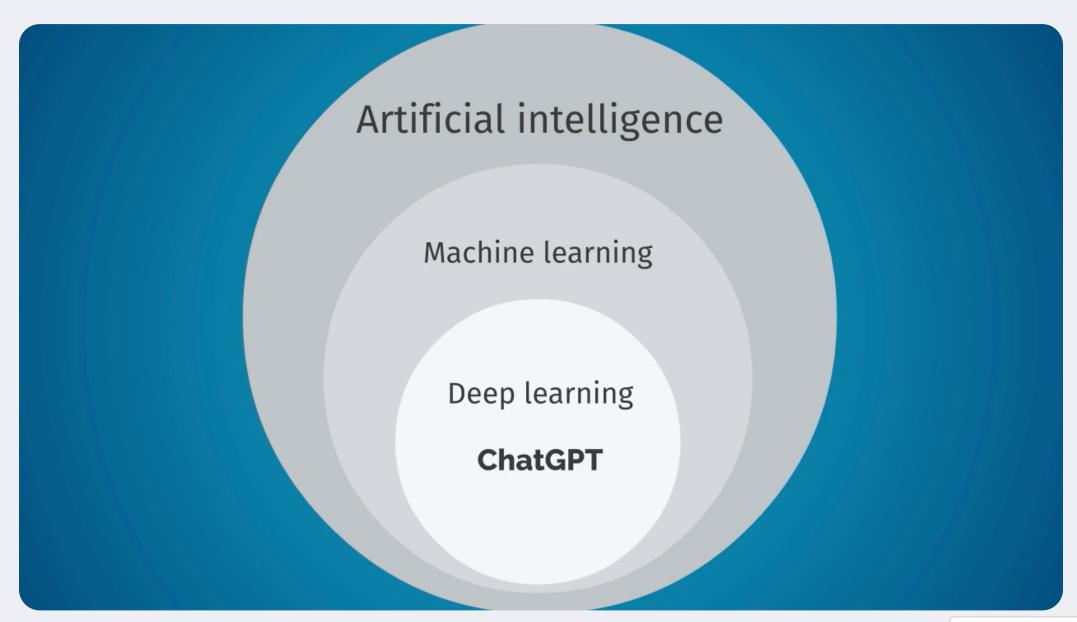
ChatGPT is nothing like this! 😀

ChatGPT: Optimizing Models Language For Dialogue We've trained a model called ChatGPT which interacts in a conversational way The dialogue

ChatGPT is narrow Al

- Designed only to generate text in a conversational style
- This is what we'll focus on today

Where does ChatGPT fit?



What is machine learning?

A subfield of AI that involves developing models that can **automatically learn** from and improve their performance based on input data, allowing computers to make predictions, recognize patterns, and solve problems without being explicitly programmed for each task.



Without being explicitly programmed!

They **learn patterns**, instead of following rules.

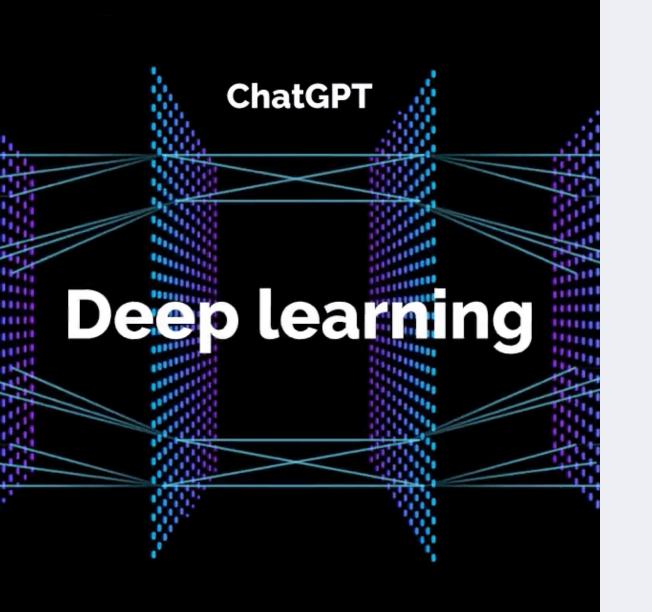
They are **probabilistic**, not deterministic.





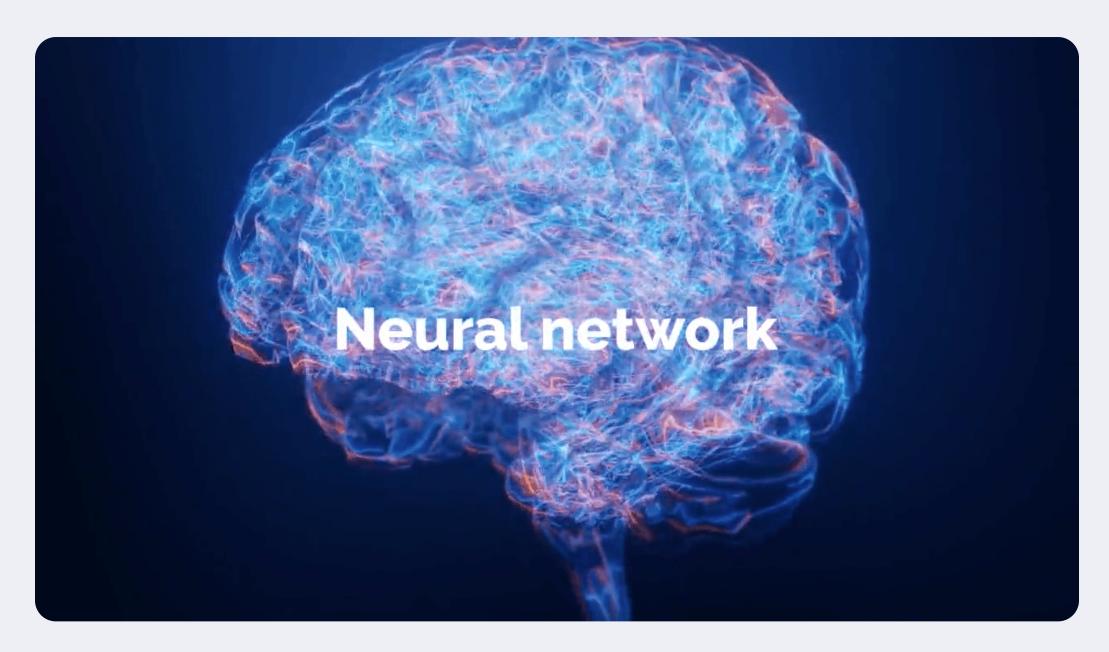
This is very different from explicit software programming and rules.

It's about patterns and probabilities.

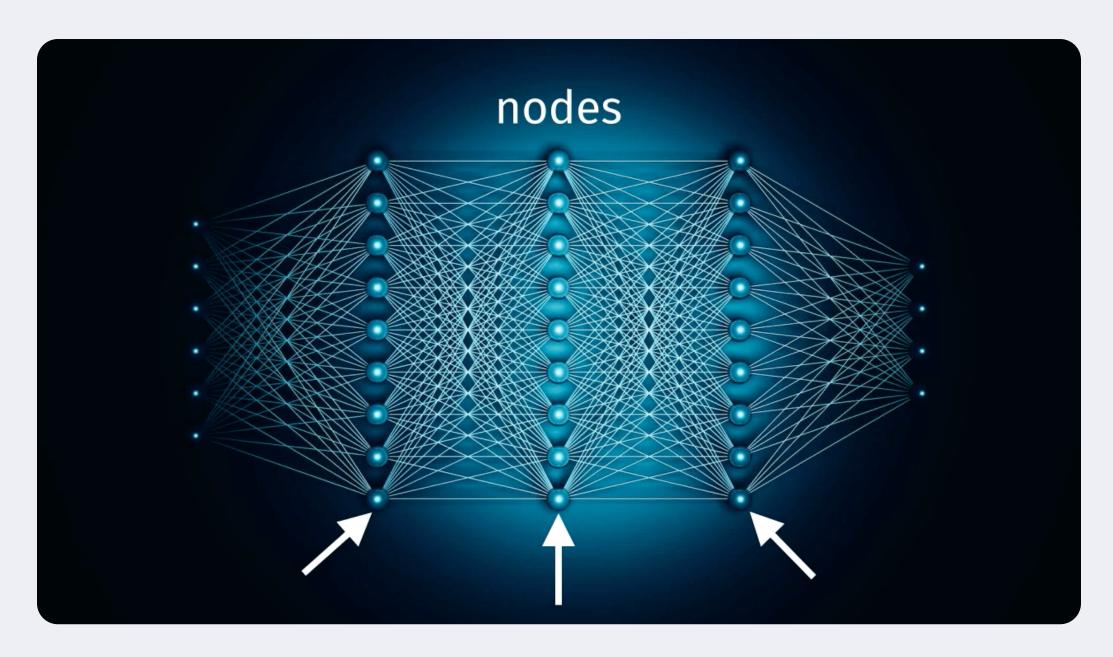


What is deep learning?

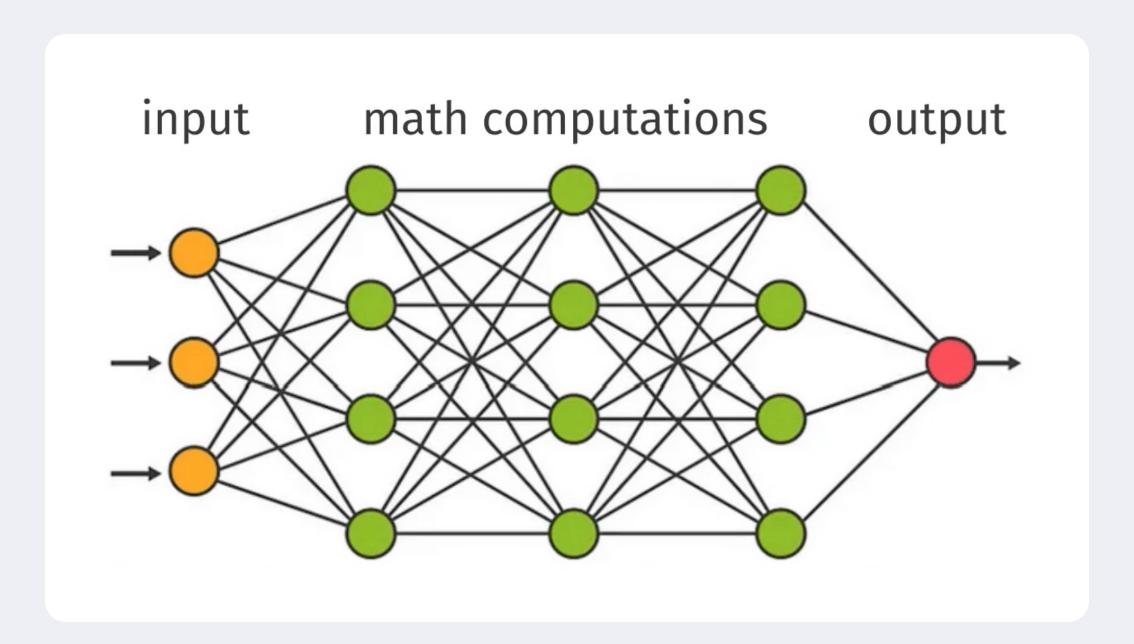
It's a particular type of neural network known as a **deep neural network,** which consist of multiple layers of interconnected nodes that enable the learning of more complex patterns.



a model inspired by neural networks in the **human brain**.



It consists of processing units, known as **nodes**, organized into layers.



These nodes receive input, perform math computations, and produce output.



Knowledge is not explicitly represented

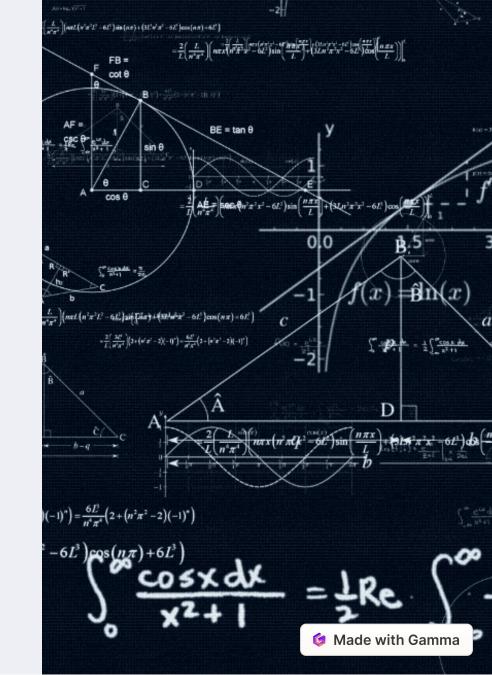
With billions of **parameters** in their networks, they don't have any easily interpretable "rules" or "knowledge base" that engineers can directly inspect or modify. Their knowledge is embedded in the connections between nodes in neural networks.

"a black box"

It's not words and sentences.

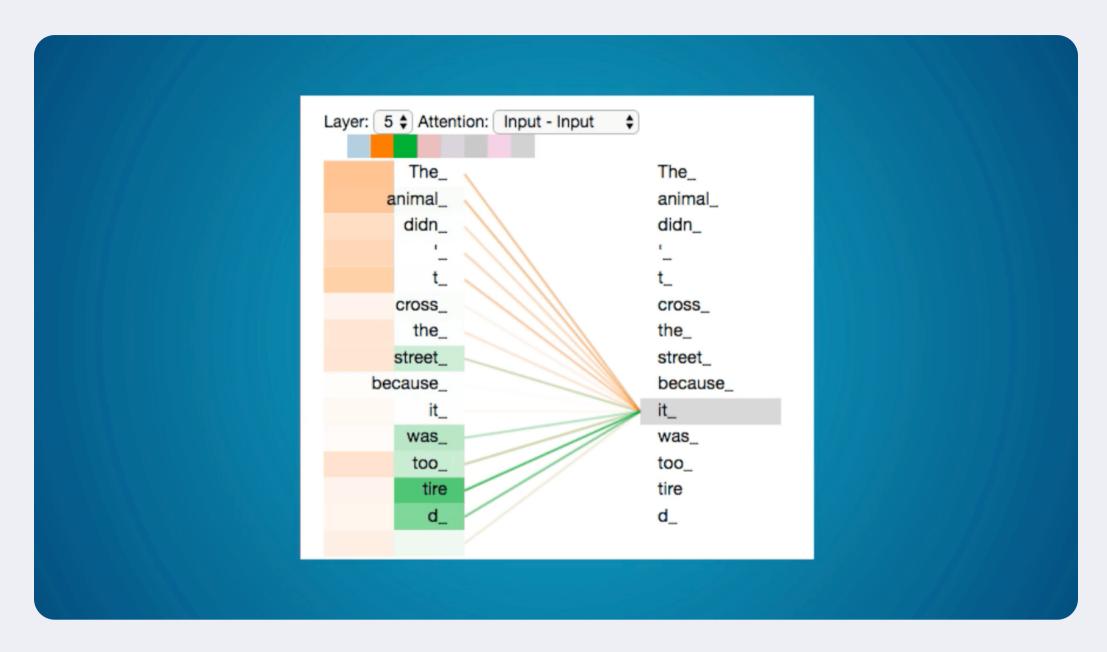
It's a big network of math operations.

(don't call it a "database")



2017 **Attention Is All You Need** Noam Shazeer* Niki Parmar* Jakob Uszkoreit* Ashish Vaswani* Google Brain Google Brain Google Research Google Research avaswani@google.com usz@google.com noam@google.com nikip@google.com Llion Jones* Aidan N. Gomez* † Łukasz Kaiser* Google Research University of Toronto Google Bra llion@google.com aidan@cs.toronto.edu ne dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 Englishto-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

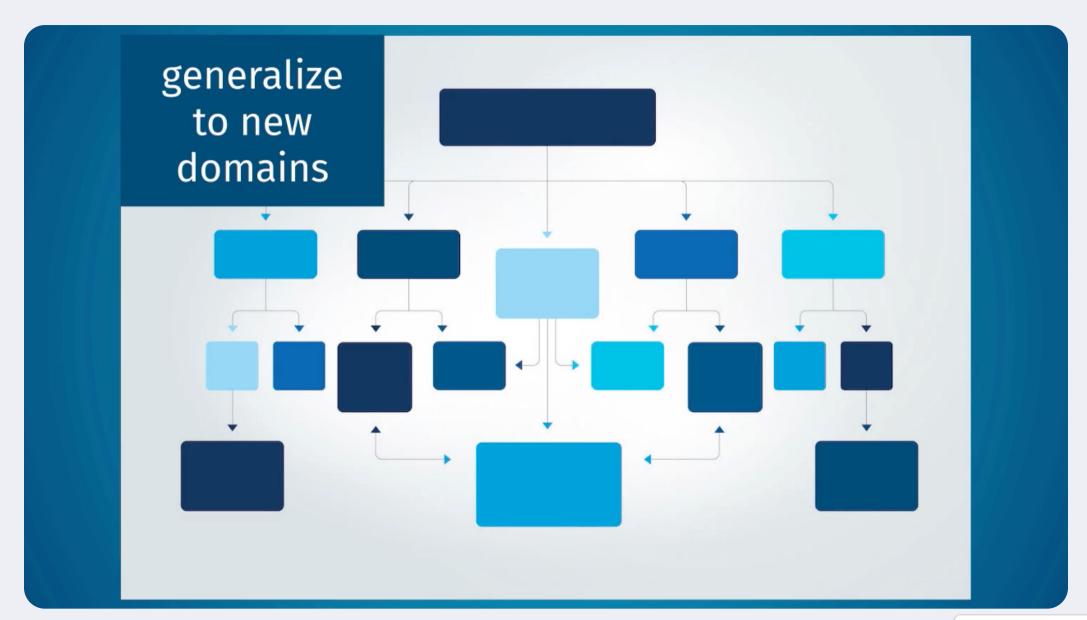
Transformer architecture

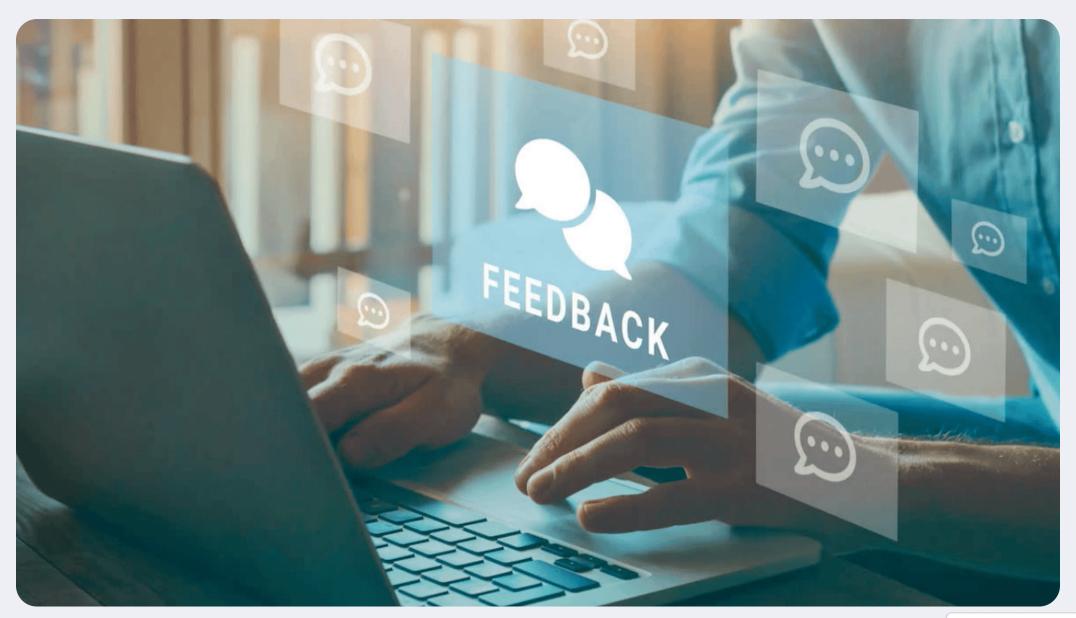


Looks at the whole context, can pay attention to certain words.



Large language models LLMs

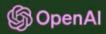




(例OpenAI

Versions for researchers

- GPT 1 (2018)
- GPT 2 (2019)
- GPT 3 (2020)



Menu

November 30, 2022

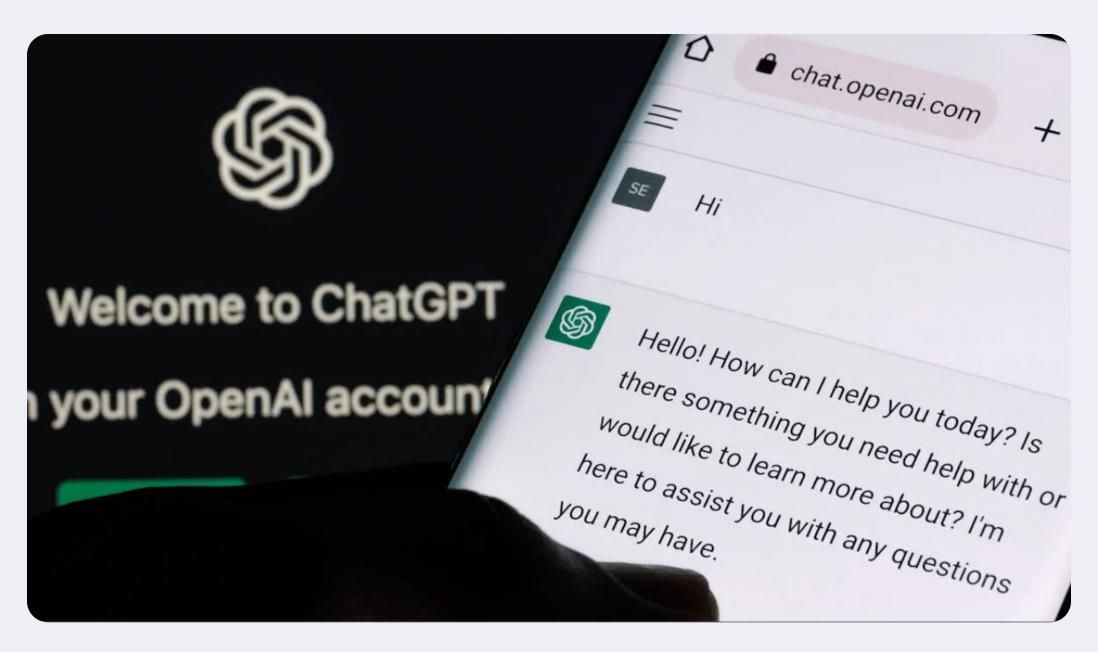
Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

Try ChatGPT ↗

Read about ChatGPT Plus

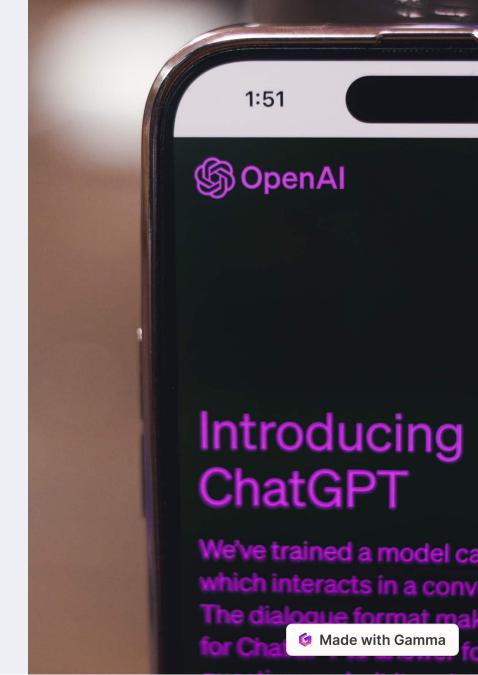




Added an easy chat interface and made it free for the public.

ChatGPT

"Generative Pre-trained Transformer"



What was ChatGPT (3.5) trained on?

WikipediA

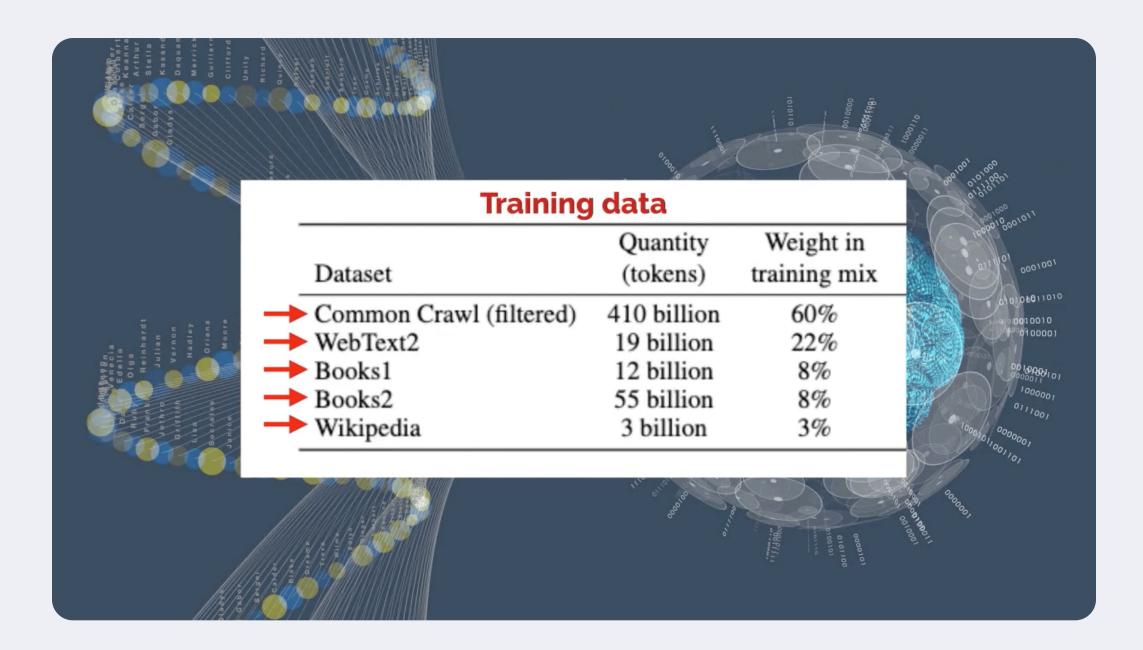
The Free Encyclopedia

Internet data through mid-2021

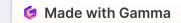


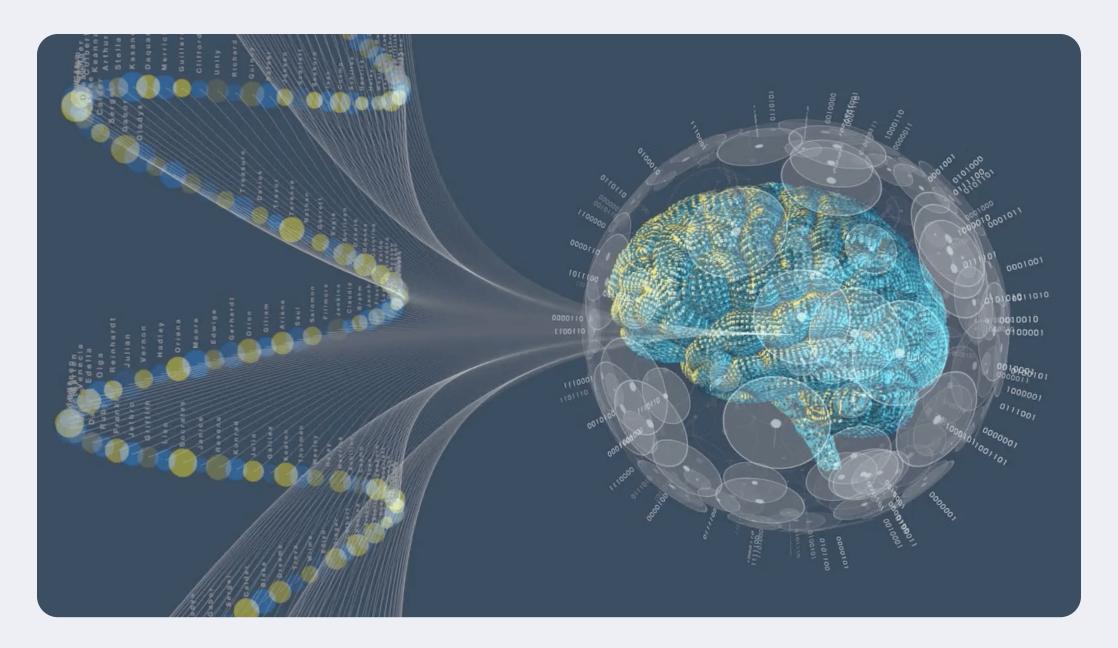
Dataset	Quantity (tokens)
Common Crawl (filtered)	410 billion
WebText2	19 billion
→ Books1	12 billion
→ Books2	55 billion
Wikipedia	3 billion

Open AI hasn't made public what GPT-4 is trained on.



Once the training data is done, it's not needed for the model to work.





It learned from all that data and saved it as patterns. It uses that to generate new text.

It can be set aside.



How to prevent harmful use? Guardrails

Systems to guide behavior of the model. Aimed at preventing harmful or misleading content.

- Content filters
- User feedback helps add more guardrails

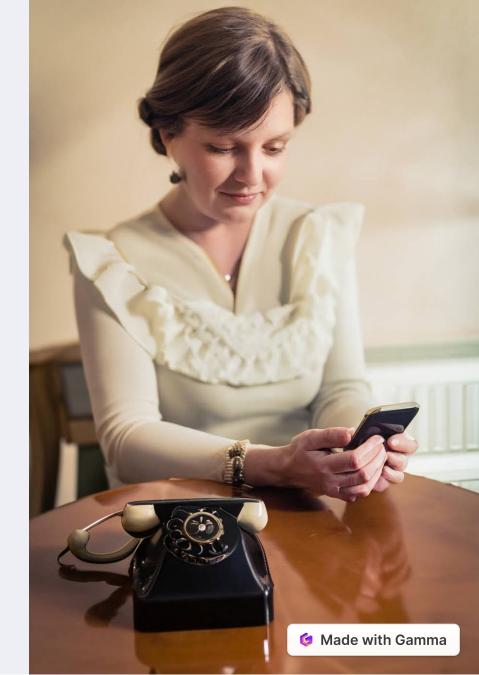


What is "generative AI?"

Al that can generate new content:

- text
- images
- video
- music
- speech

1 This image was generated with Adobe Firefly.



It's important to know the difference between generative AI & other types of AI.



Discriminative Al

Classify or recognize patterns in existing data.

Examples

- Gmail's spam filtering
- Facebook recognizing people to tag in photos
- YouTube video recommendations

Generative Al

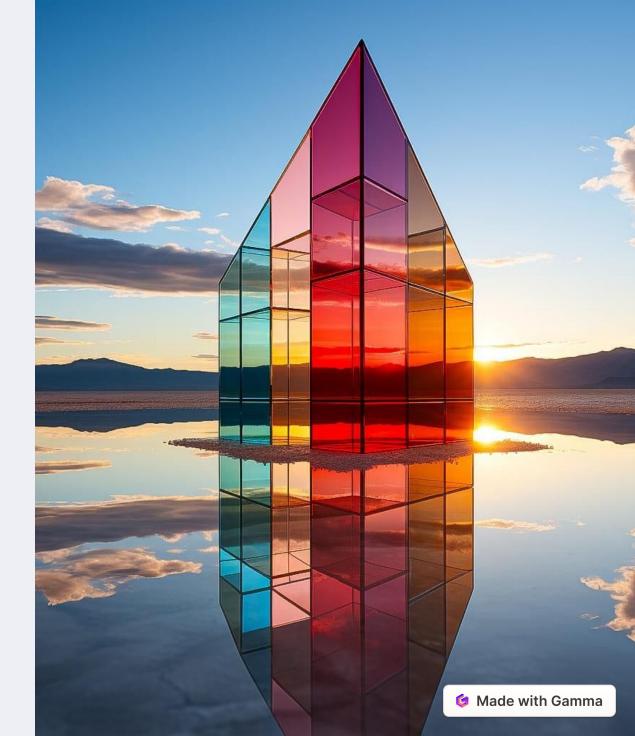
Creates new content based on learned patterns from existing data.

Examples

- ChatGPT, Bing Chat
- MidJourney, Stable Diffusion
- Eleven Labs

Today we'll focus on text generation with ChatGPT and similar tools.

☐ Image generated with MidJourney



Who created ChatGPT?

- Founded in 2015.
- CEO is Sam Altman
- Microsoft invested \$\$\$
- Open Al's vision





ChatGPT is not the only LLM

ChatGPT

Open Al

Free, built on GPT-3.5

Not connected to the Internet



ChatGPT Plus

\$20/mo, built on GPT-4

Enterprise version for large businesses

Small business version coming later



Bing Chat

Microsoft

Free, built on GPT-4 (with Creative mode)

Connected to the Internet



Bard

Google

Free, built on PaLM 2

Connected to the Internet



Claude

Anthropic

Free, built on their own model

Not connected to the Internet

Focused on **reliability and safety**.



Perplexity AI

Perplexity

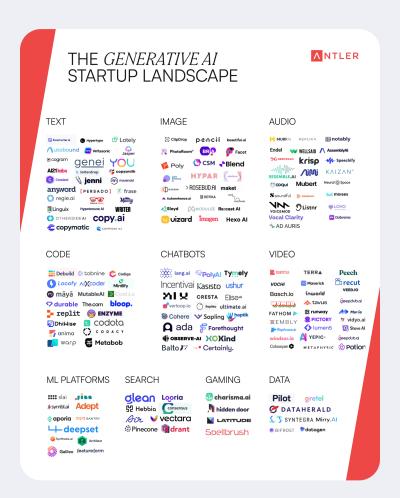
Free, uses GPT-3.5 (free version), Pro version uses GPT-4

Connected to the Internet

You can use the free version without creating an

account.

And many, many tools based on these



A search engine for genAl tools: Futurepedia











Alpaca Falcon

Vicuna Llama

There are also open source LLMs

▼ BLOOM

With its 176 billion parameters, <u>BLOOM</u> is able to generate text in **46 natural languages** and 13 programming languages. A research group received a grant to build it using the French government's supercomputer.

▼ Alpaca

From <u>Stanford University</u>. The researchers were able to train LLaMA to give similar outputs to OpenAl's model. The model had a **training cost of only \$600** — far less than the millions of dollars it takes to train these models. It's for **academic use only** (commercial use is prohibited).

▼ Falcon

<u>Falcon</u> was Founded and built by the Technology Innovation Institute (TII), a company that is part of the Abu Dhabi Government's Advanced Technology Research Council. This model is multilingual, understanding English, German, Spanish and French, and Dutch, Italian, Romanian, Portuguese, Czech, Polish and Swedish. <u>Learn more</u>.

▼ Llama2

<u>Llama2</u> is Meta's open source large language model (LLM). It's basically their response to OpenAI's GPT models and Google's AI models — but with one key difference: it's freely available for almost anyone to use for research and commercial purposes.

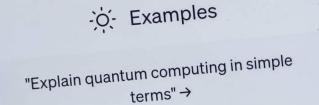
These are just a few of them.

Using ChatGPT Effectively



Why might you want use ChatGPT?

- To help you understand what your students are doing with it.
- To help you generate ideas for teaching and learning activities.



How to sign up

https://chat.openai.com, verify your email

Enter your name and birth date. Verify your phone (mobile only) - they'll text you a code.



They remind you that...

1 It's a free research preview, so might have incorrect info.

And may contain biased info.

They review some of your inputs to improve the model.

X They remind you: don't share sensitive info.

They welcome feedback.

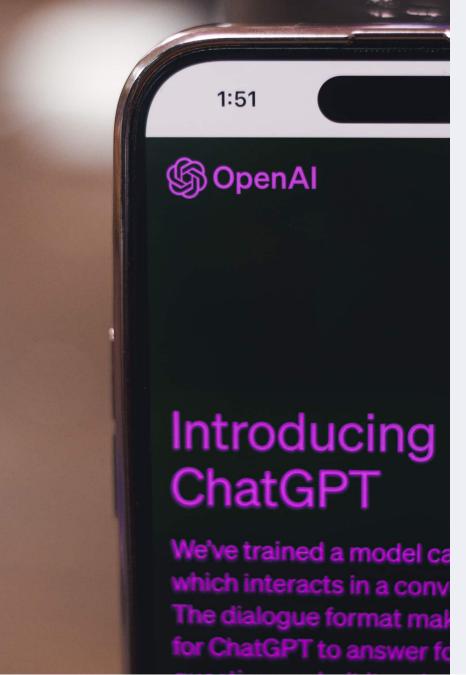


Thumbs up or down, add comments.



People make this mistake

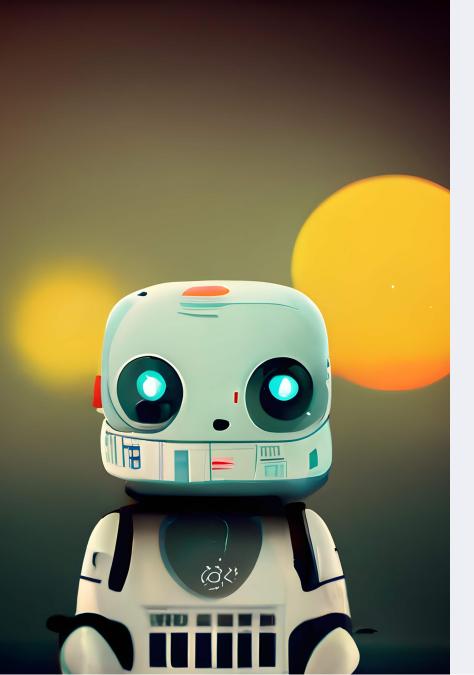
You ask a question, you don't like the answer, you leave.



It's worth learning about prompting

Prompting: how you talk with it





Think of ChatGPT as your personal intern.

They are new and don't know you well.

They are pretty smart, but they need you to give them a role and specific instructions.

Start with tasks you know a lot about.

You can guide them since you know about this topic.



Prompting advice

1 Give it a role to play or some context.

- 2 Give it detailed instructions, including format.
- 3 Keep conversing: ask for changes

1. Give it a role to play or a context.

Role

"Act as an expert community organizer."

"Act as a high school biology teacher."

"Act as an academic librarian."

Context

"Use Mike Caulfield's ideas for how to be an effective fact-checker. Now create a lesson plan for college freshmen to"



2. Give it step-by-step instructions, including format.

Prompt:

Act as an instructional designer. Please make an outline of important topics to cover for a classroom workshop on ChatGPT. This will be part of a series of instruction sessions about information literacy. You have already worked with this class on the basic concepts of information literacy. **Create the outline as a bulleted list** and follow it with a classroom activity based on one of the topics you've listed.

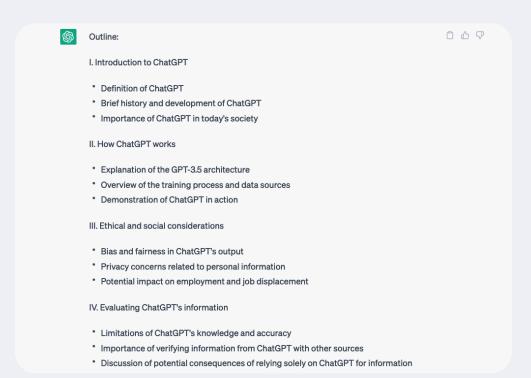


3. Don't expect the perfect answer right away.

Ask for changes, clarifications, or improvements. Tell it exactly what you want and what format you'd like it in.

Result

Outline + an activity about ethical concerns



V. Using ChatGPT responsibly

- * Guidelines for using ChatGPT ethically and responsibly
- . Tips for incorporating ChatGPT into research and writing
- Discussion of the importance of acknowledging ChatGPT's contributions in academic work

Classroom activity:

One topic that could be explored further in class is the ethical considerations of using ChatGPT. To explore this topic, students could be divided into groups and given a scenario in which they must use ChatGPT to gather information on a specific topic. Each group would then evaluate the information obtained from ChatGPT and discuss any ethical concerns that arise from their findings. For example, a scenario could involve using ChatGPT to gather information on a controversial topic, such as climate change or gun control, and the group would need to discuss the potential biases and limitations of the information provided by ChatGPT. This activity would allow students to practice critical thinking skills and consider the implications of relying on Al for information.

Use the results as a starting point.

Use your own expertise to modify these. It's great for getting ideas. But you are the expert.

Demo

https://chat.openai.com/

I recommend using Chat GPT Plus (\$20/mo) or Bing Chat (free), since both use the more capable model GPT-4.

Prompts used

You are an expert instructional designer. Please help me create 2 active learning activities for college freshmen that help them learn to use ChatGPT effectively. Outcomes should be that they understand the need to fact check the output, and that they can prompt effectively.

Now create a series of online learning modules with the same outcomes, for the same audience. Each module should include an animated video (please write the transcript), a bulleted list containing 4 or 5 brief hints, and 3 multiple-choice quizzes that students can do to reinforce their knowledge. For the quizzes, each question should have only one correct answer.

Saved conversation



What about topics you don't know much about?

Encourage everyone to check ChatGPT's answers.

Hallucination

Sometimes ChatGPT makes up answers. Why?

It works on the **probability** of what the next words should be, based on its training data. (ChatGPT was trained on data only up to mid-2021).

Some experts estimate that ChatGPT gets it wrong about 20% of the time.



It's improving, though



ChatGPT (free) runs on GPT-3.5

ChatGPT Plus (\$20/mo) runs on GPT-4

GPT-4 (Plus version) hallucinates less and is more capable



Menu

We spent 6 months making GPT-4 safer and more aligned. GPT-4 is 82% less likely to respond to requests for disallowed content and 40% more likely to produce factual responses than GPT-3.5 on our internal evaluations.

Safety & alignment

Training with human feedback

We incorporated more human feedback, including feedback submitted by ChatGPT users, to improve GPT-4's behavior. We also worked with over 50 experts for early feedback in domains including Al safety and security.

40% more likely to produce factual responses than GPT 3.5 (free version)

Hallucination can still happen, though

Encourage everyone to fact check.

Keep a "human in the loop."

11011101 (2010).

- "Engineering the climate: The ethics of solar radiation management" by Toby Svoboda, Environmental Ethics (2011).
- "Geoengineering and the myth of unilateralism: Pressures and prospects for international cooperation" by Daniel Bodansky,

A ----- L ---- D ---- (0040)

Don't ask ChatGPT for a list of sources

I can't find the citations that ChatGPT gave me. What should I do?

UA Libraries FAQ



Let's talk about what people use it for most effectively.

1. Writing first drafts

- marketing materials
- cover letters for jobs
- grant proposals

Change level:

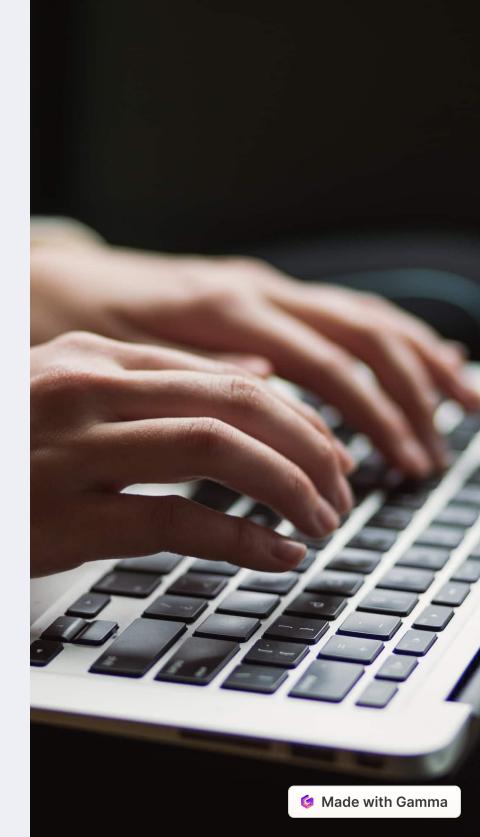
Make it appropriate for:

college freshmen, grad students, 8th graders

Change style:

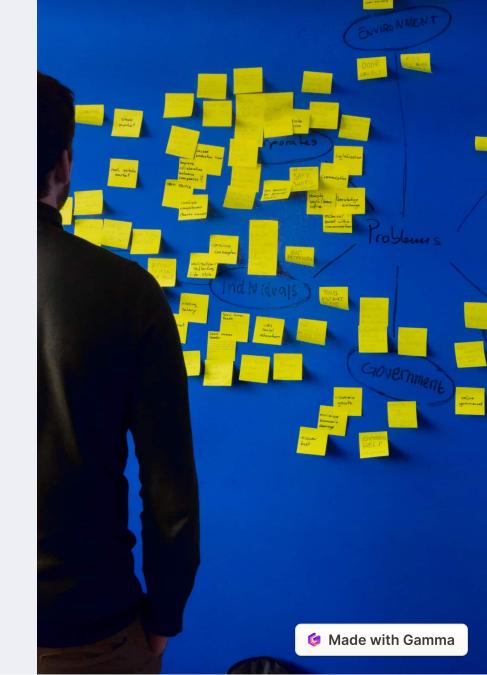
Make it more:

formal, persuasive, academic, descriptive, technical



2. Generating ideas

Ask ChatGPT to make a list of ideas for a topic.



3. Learning and studying

Come across something that's hard to understand?

- explain it like I'm in high school
- explain it like I'm 5

Ask for analogies, definitions, synonyms, or an outline.



People with ADHD find it useful



How AI tools like ChatGPT can combat ADHD paralysis

Generative AI could help people who struggle with focus. But becoming a Dungeons & Dragons firbolg cleric still requires a human touch.

By Abby Ohlheiser on April 16, 2023

Helps people who struggle with focus. Helps organize facts and stay focused.

It's useful for neurodiverse people

Al is non-judgmental, and you can ask a million questions without fear of being judged.

Furthermore, it can explain complex concepts in small achievable chunks of information, synthesising and condensing information at a pace that suits the learner.

-Simon da Roza, founder of **Xceptional Learners**

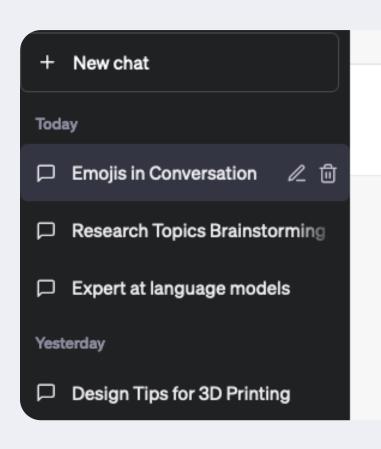
4. Summarizing

• You are an expert biology teacher. Here is a transcript from a video lecture about photosynthesis. I'd like you to summarize it briefly, listing the most important points. Write in simple, straightforward language.



A few tips

Start a new chat for each topic.



New chat in sidebar.

- Sometimes gets confused if you change topics in one conversation.
- It remembers within a conversation, not between.

You can share the result of a chat with others

Use the share on top right of your conversation.



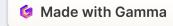
See **ChatGPT Shared Links FAQ**.

(It doesn't include your name unless you decide to include it).

Useful when collaborating with your colleagues.

Here's a shared link.

This link is useful when citing a conversation.



Choose an output format

- a table
- a bulleted list
- ascii art
- multiple choice quiz questions
- emojis (Avatar: The Way of Water)





You are an expert in population studies. Please create a table that lists the top 10 most populous cities in the world. First column will contain the city name, second column the population.



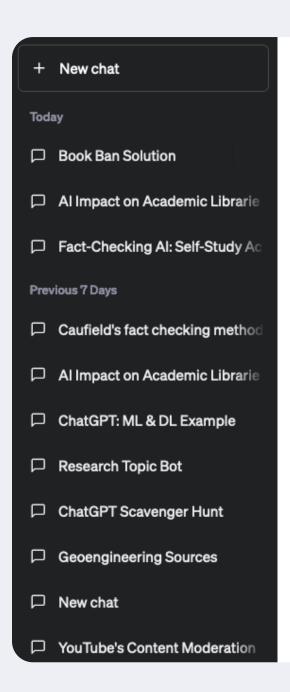
Sure! Here is a table that lists the top 10 most populous cities in the world as of 2021:





City	Population (2021)
Tokyo, Japan	37,393,000
Delhi, India	30,291,000
Shanghai, China	27,058,000
São Paulo, Brazil	22,043,000
Mumbai, India	21,042,000
Beijing, China	20,384,000
Karachi, Pakistan	20,026,000
Istanbul, Turkey	15,190,000
Dhaka, Bangladesh	14,543,000
Moscow, Russia	13,197,000

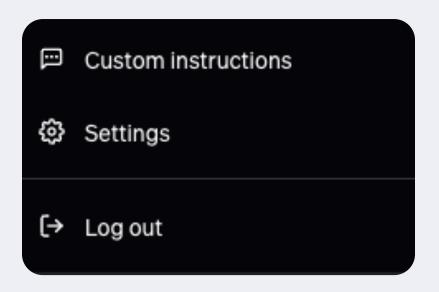
Review your chat history



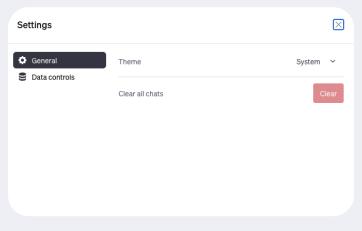
Model
Web Browsing (GPT-4) ALPHA

ChatGPT PLUS

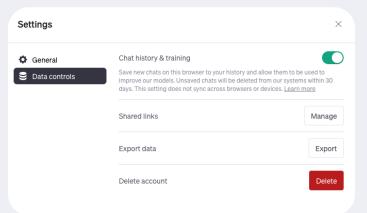
You can delete or turn off history



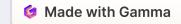
• Use the settings: bottom left column, near your name.



 You can delete your chat history ("clear conversations")



- <u>Turn off future chat history</u> (this also turns off training)
- You can **export your data.**



You can go crazy with prompts!

[the "several experts" prompt]

I want you to split into several expert librarians. The first librarian will
describe a problem happening in their community with people who want
to ban certain books. The other librarians will work together and try to find
a solution. Use as many librarians as you want. Assign a random special
expertise to each librarian. Let them vote on the best solution and have the
winner present it to me.

(<u>saved conversation</u>, replacing the expert with writing instructors, and the problem with students using ChatGPT to write their papers)



Using ChatGPT to help with library research



Generate topics for your research paper with ChatGPT

ChatGPT can be a useful tool when it comes to deciding what your topic should be for a research paper.

It's **not** good for doing the actual searching because **it makes up sources.** See <u>I</u> can't find the citations that ChatGPT gave me. What should <u>I</u> do?

But you can use it to help you:

- Narrow down your topic ideas.
- Come up with keywords for searching in library databases.
- · Construct a search strategy for those databases.
- Recommend possible library databases to use for your topic.

Here's how.

 Sign up for a free account on ChatGPT (if you haven't already). Or go to https://chat.openai.com and log in to your account.

Download the Google Doc.



If you want to learn more about prompting:

Free course: https://learnprompting.org

There is a lot more to learn. Best way to learn is to keep experimenting.



What about search?

People want to use ChatGPT as a Google replacement.

You get answers directly, instead of 10 blue links.

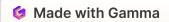


Bing chat

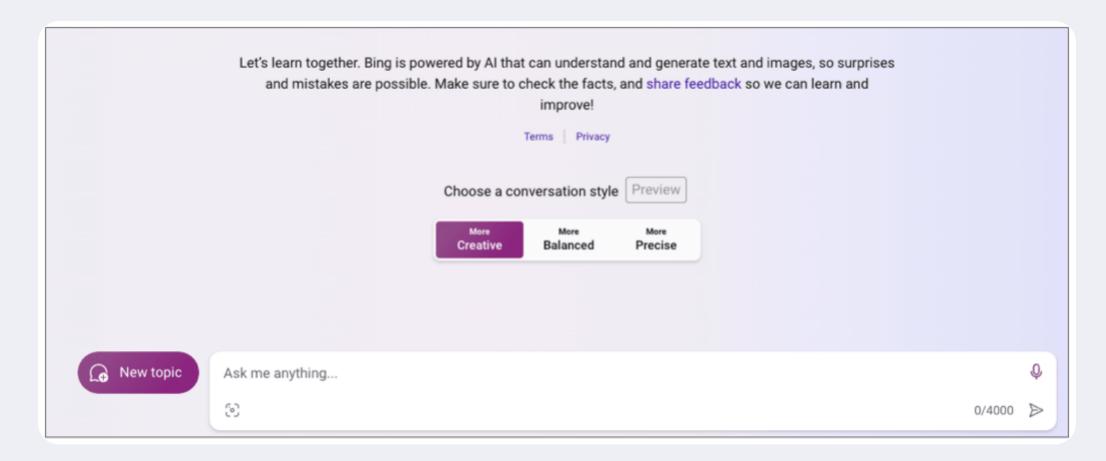
Based on GPT-4 (more capable than GPT3.5)

Internet connected: tells you where it got the information, with **links to** sources. This helps you verify the answers.

Sign up at https://bing.com/new



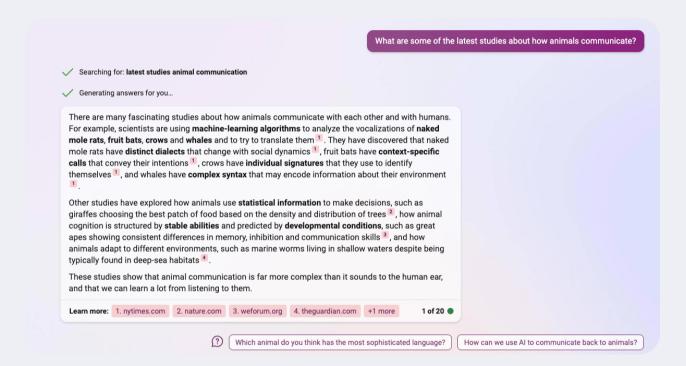
Bing Chat



Example

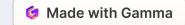
What are some of the latest studies about using AI to understand animal communication?

Bing chat



Look for the share, export and copy buttons, upper right of each response.





A similar tool: Perplexity.ai

Perplexity.ai

Based on the same model: GPT-4

Links to sources.

No sign-up required.

- "view sources" you can eliminate a source (good for removing unreliable websites)
 It will repeat without that source.
- share button top right corner (shares your whole conversation, not just one answer)



Internet-connected tools make it less likely to hallucinate.

It's called "grounding."



Which models connect to the Internet?

- ChatGPT Plus (\$20/mo)
- Bing Chat
- Perplexity.ai
- Google's Bard

Which don't? (yet)

- ChatGPT (free)
- Claude from Anthropic



This is changing - at some point maybe all of them will connect to the internet.

Are there tools that detect Al generated writing?

Yes, but they are very unreliable.

- GPTZero
- Al Writing Check
- <u>Al classifier</u> from Open AI "As of July 20, 2023, the AI classifier is no longer available due to its low rate of accuracy."
- <u>Turnitin</u> (new feature) D2L at UA has turned it off.



Tips are shared everywhere about how to get around these tools.





And there are false positives



What to do if you are falsely accused

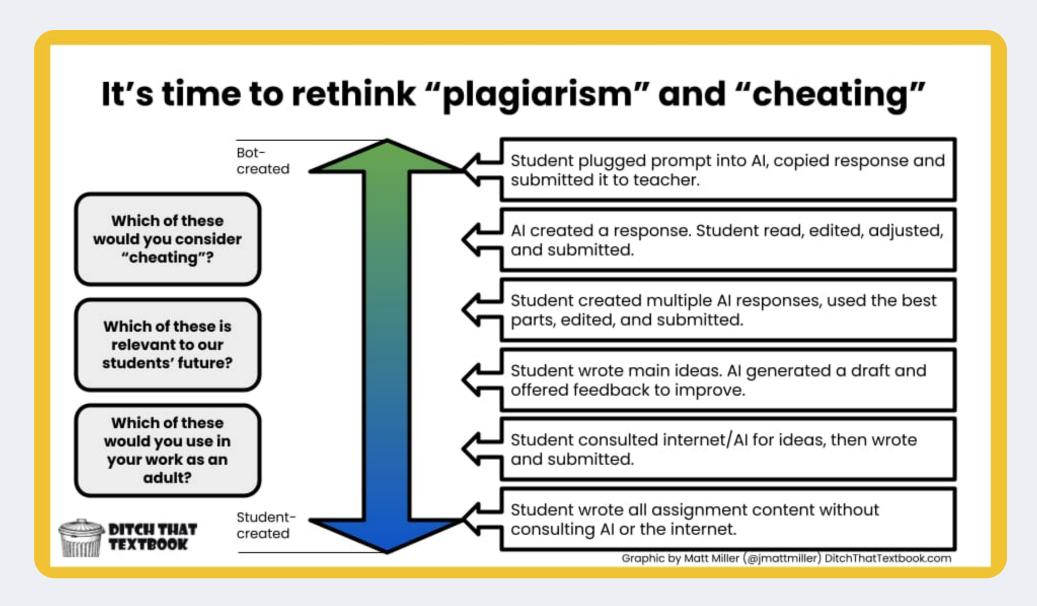
GPT detectors are biased against non-native English writers

Many academics are recommending against these

Vanderbilt University

Guidance on AI Detection and Why We're Disabling Turnitin's AI Detector

Posted by Michael Coley on Wednesday, August 16, 2023 in Announcements, News.



Source: **Ditch That Textbook**

Banning it for a whole school or campus won't work. Generative Al is appearing in tools you already use.



Notion - transform text, automate simple tasks, and generate content



<u>Slack</u> - uses Claude from Anthropic



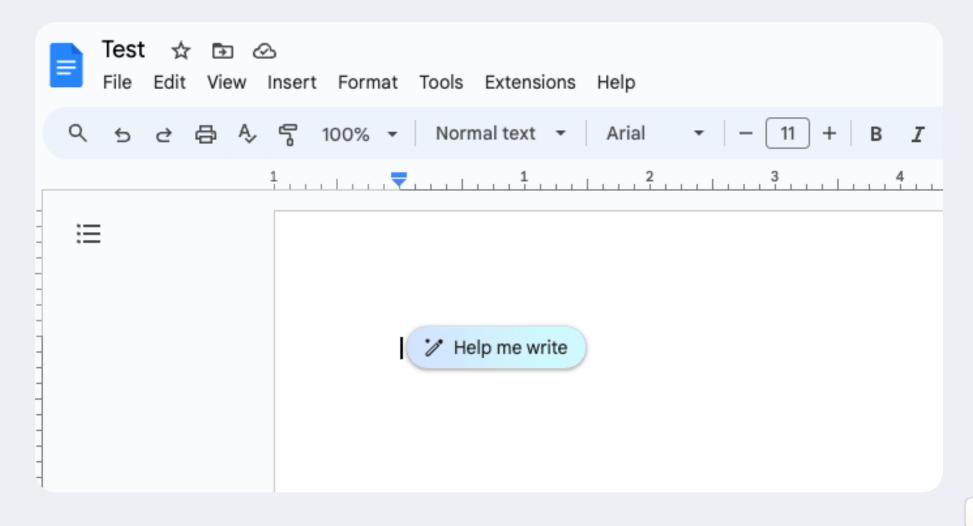
Canva - Magic Write
(also has image
generation)



Box - ask questions about what's in a document.

Will be part of more <u>Microsoft</u> & <u>Google</u> products.

Al in Google Docs



Advice for instructors



Syllabus guideance

1

Guidelines for Syllabus Statements About Generative Al

August 2023 v 2.1.1

I. Purpose of these Guidelines

The following guidance aims to help instructors thinking about the impact of generative artificial intelligence (AI) tools on teaching and learning. Specifically these guidelines suggest instructors:

- include a syllabus statement regarding use of AI tools;
- create transparent and productive learning environments by explicitly discussing appropriate, creative, and/or ethical AI use within a course, discipline, and/or profession; and
- prevent situations in which a student unintentionally engages in academic dishonesty while using AI tools.

The University of Arizona Catalog includes a list of required syllabus policies, including The Code of Academic Integrity. The phrase "graded work/exercises must be the product of independent effort unless otherwise instructed" can be interpreted differently. Therefore, instructors should also include a statement specifically addressing course policy on the use of generative Al and discuss this policy with students. If applicable the syllabus should include information about the use of Al detection applications, because these involve sharing student-created work. Example syllabus statements are provided below.

Templates

- Generative AI use is NOT allowed for any purpose.
- Generative AI use IS permitted or encouraged.
- Generative AI use is allowed for certain purposes/assignments, but not for others.

Ideas from Ethan Mollick

Ethan Mollick, a professor at the Wharton School of the University of Pennsylvania.

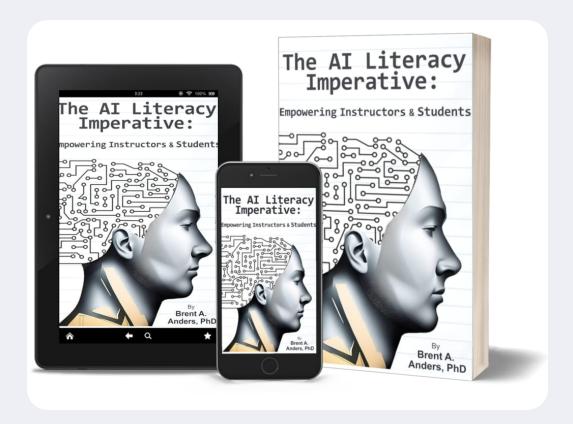
My class required AI. Here's what I've learned so far.

"... teaching an AI to improve an essay is a pedagogical method that can produce new insights. I would strongly suggest that you push students in this direction, if you intend to incorporate AI essays into your classes."

Ideas from Brent Anders

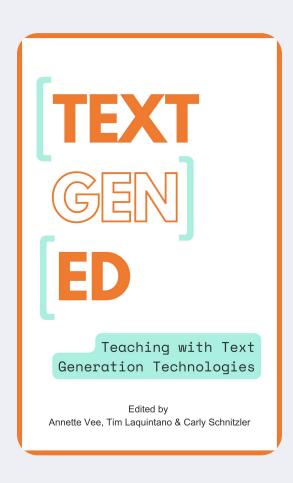
Follow Brent Anders, instructor at American Univ. of Armenia.

Writing Assignments in the Age of Al



TextGenEd - collection of teaching ideas

by Tim Laquintano, Carly Schnitzler, and Annette Vee



Teaching ideas with reflections from instructors

"All assignments include a description of how students responded along with reflections from the instructors."

https://wac.colostate.edu/reposito ry/collections/textgened/

Prompting can help "make thinking visible"

From **<u>Higher Ed Discussions of Al Writing</u>** (on Facebook)

Replies to **Sandra Metzger, Ed.D.**'s comment on this **post**



Sandra Metzger, Ed.D. (She/Her) • 2nd 47m ••• Innovative, human-centered leader with hands...

Excellent point.

I think this also applies to how the process of prompting can be a good way for students to "make their thinking visible" and for teachers to function as thinking guides.





Leonard Rodm... (He/Him) • 3rd+ 40m ••• Al Evangelist | LMS Lead Admin at Impact...

If I were still teaching high schoolers, I'd make them write prompts as in class assessments, then run them!

```
Like • 🖒 1 Reply
```



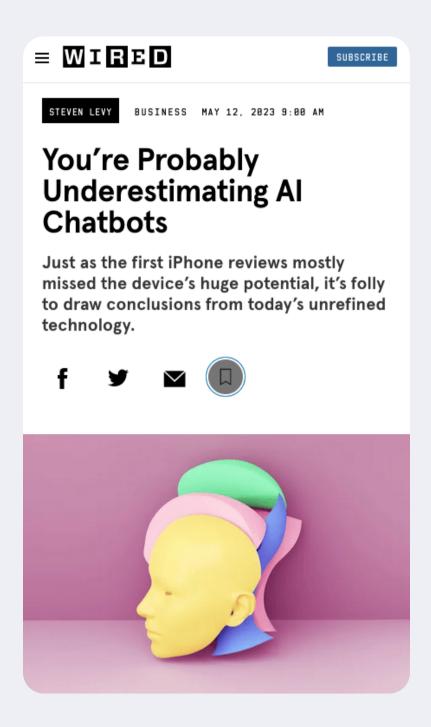
Sandra Metzger... (She/Her) • 2nd 38m ••• Innovative, human-centered leader with h...

Exactly, Leonard Rodman, M.Sc.

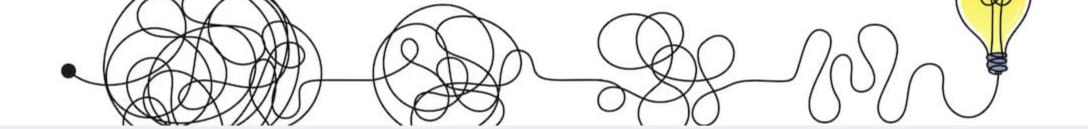
The back-and-forth dialogue of a prompt iteration is also highly useful as evidence of learning and thinking.

Like • 🖒 1 Reply

This is just the beginning



"...it's folly to draw definitive conclusions based on these early versions of the technology."



Ethical issues

Misinformation | Copyright | Bias



Misinformation

Teaching fact-checking is more important than ever.

Progress toward making truthful systems

<u>Grounding</u> is an important concept in AI research. It is about connecting the model to **external sources of knowledge.**

- <u>Elicit</u> answers from 175 million papers ("find relevant papers without perfect keyword match")
- Scite see how research has been cited
- Consensus evidence-based answers
- **CORE-GPT** based on Core's collection of open access scientific papers





COPYRIGHT

Copyright & intellectual property

- a. the **output** of generative AI (data you create using AI)
- b. the **input** of generative AI (training data)



Output

Can you copyright something you made with AI? Open AI says:

14. Can I use output from ChatGPT for commercial uses?

 Subject to the <u>Content Policy</u> and <u>Terms</u>, you own the output you create with ChatGPT, including the right to reprint, sell, and merchandise – regardless of whether output was generated through a free or paid plan.

U.S. Copyright Office

The term "author" ... excludes non-humans.

But, if you select or arrange Al-generated material in a sufficiently creative way....

... In these cases, copyright will only protect the human-authored aspects of the work.



Example

The selection and arrangement of the text and images they chose were accepted as being copyrightable. But not the images themselves. (they used MidJourney).



Comic book "Zarya of the Dawn"

What about the input? (training data)

One position:

It's unlawful to use copyrighted material to train generative AI tools. (see <u>case from artists</u>)

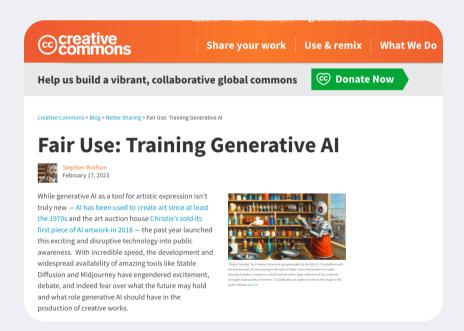
Another position:

It's protected under Fair Use because it's transformative.



What does Creative Commons say?

It should be considered fair use, as the purpose of copyright law is to encourage new creative works and benefit the public interest.

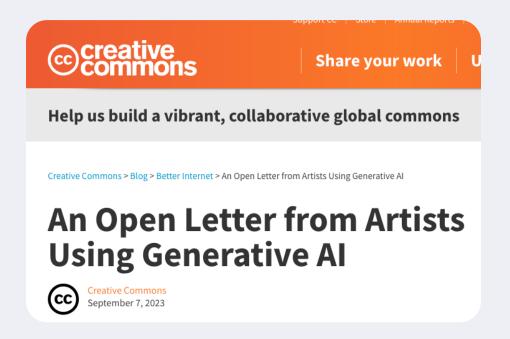


Fair Use: Training Generative AI - It's transformative use.

Read their series of posts about copyright and AI.



Yes, there are artists in favor of generative Al tools

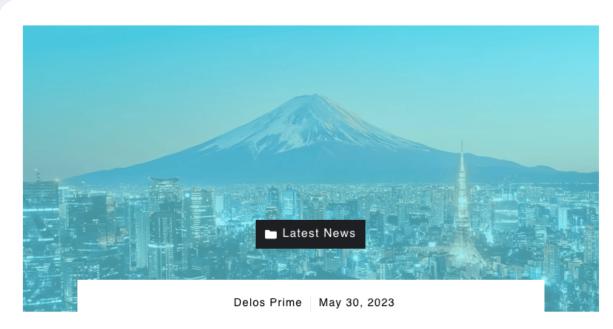


"Just like previous innovations, these tools lower barriers in creating art—a career that has been traditionally limited to those with considerable financial means, abled bodies, and the right social connections."

https://creativecommons.org/2023/09/07/an-open-letter-from-artists-using-generative-ai/



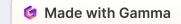
Japan: copyright doesn't apply to Al training



Japan Goes All In: Copyright Doesn't Apply To Al Training

In a surprising move, Japan's government recently reaffirmed that it will not enforce copyrights on data used in Al training. The policy allows Al to use any data

Japan's government recently reaffirmed that it will not enforce copyrights on data used in Al training.



Israel: Al training is fair use



The Israel Ministry of Justice has issued an important opinion concluding that <u>the use of</u> <u>copyrighted materials in the machine learning (ML) context is permitted</u> under existing Israeli copyright law.





It will take years for this to be resolved in the courts.

Pamela Samuelson, Professor of Law, UC Berkeley

In the meantime, <u>Adobe</u> (for Firefly), and <u>Microsoft</u> (for Copilot) have offered to any legal bills from lawsuits against individual users of their tools.





Siobhan O'Connor 1, ChatGPT 2

PMID: 36549229 DOI: 10.1016/j.nepr.2022.103537

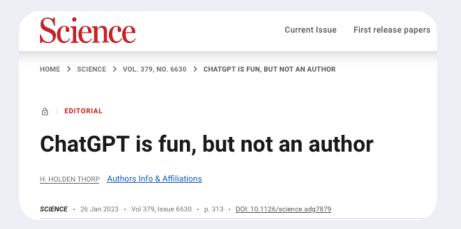
Affiliations + expand



Should you list ChatGPT as an author?

No, but **some have**. (←see screenshot)

Publishers have come out with statements



Artificial intelligence (AI). Text generated from AI, machine learning, or similar algorithmic tools cannot be used in papers published in *Science* journals, nor can the accompanying figures, images, or graphics be the products of such tools, without explicit permission from the editors.



Nature recognizes that authors will use it in some way.

"... no LLM tool will be accepted as a credited author..."

"... researchers using LLM tools should **document this use in the methods** or **acknowledgements** sections."

Guidelines are available for citing generative Al

- MLA
- APA
- Chicago Manual

Be transparent in your use.



You aren't citing a person, so don't list the author as ChatGPT.

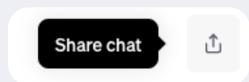
- It's not a communication with a person.
- The chat output can't be reproduced by other users.
- It's software, not a person. Give credit to the creators of the language model.
- Acknowledge your use in text or notes.
- Check with your publisher.

MLA

Note

1. At the time of writing this post, *ChatGPT* doesn't have a built-in feature to create a unique URL to the conversation. However, an outside tool like the Chrome extension *ShareGPT* can generate such a link. If you use that type of outside tool, include the unique URL that the tool generates instead of the general URL.

But now you can create a link to a conversation.



MLA

Version

Name the version of the AI tool as specifically as possible. For example, the examples in this post were developed using *ChatGPT* 3.5, which assigns a specific date to the version, so the Version element shows this version date.

Publisher

Name the company that made the tool.

(So get familiar with version numbers and company names).

See this chart.





Bias is in the training data

sexism, racism, homophobia, ableism, ageism, etc.

Open AI has worked on mitigating this

Reinforcement Learning with Human Feedback

Humans rate with the goal of removing bias.

But they can't help but use their preferences and unconscious bias.





Claude: Uses AI to train the AI - with a set of principles

See Claude's constitution

- a set of principles: some modeled on UN Declaration of Human Rights, considering non-Western audiences
- goal of an AI that's helpful, honest, and harmless



So we'll need to keep teaching about bias.

You can also ask ChatGPT to be more inclusive.

Example prompt

Act as an expert feminist. Please write a short story about a boy and a girl choosing their careers. Don't use stereotypical careers that are based on gender.

Saved conversation



Misconceptions that get media hype

Flawed study



ChatGPT left/liberal?

Meta's Palm2 conservative?

Critique



By **DEREK ROBERTSON** | 08/24/2023 04:00 PM EDT

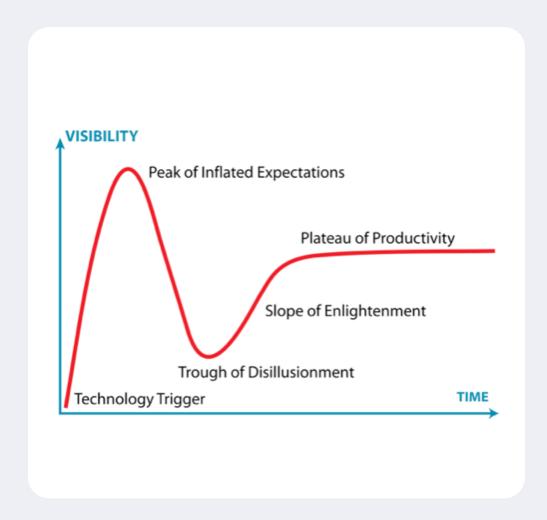


The ChatGPT app logo. | Richard Drew/AP Photo

A <u>new research paper</u> **purports to shed some light on** the ongoing debate about political bias in AI-powered tools like ChatGPT.

Future developments (very near future!)

Gartner Hype Cycle



Gartner proposes that every new technology goes through this cycle.

(see also Wikipedia, Google, cell phones)

Examples of moral panics around technologies, see **Pessimists Archive**.





1. Open source models are innovating

- Open source LLMs: <u>trained on many</u>
 <u>languages</u>, using responsible training data, working to mitigate bias
- Read about <u>LlaMA, Alpaca, & Vicuna</u>
- Smaller LLMs: run on <u>your computer</u> or even <u>mobile phone</u>!
- Open model from Meta: Llama2.



2. Models are getting more capable.

Misconception that get media hype

Flawed study

ChatGPT's accuracy has gotten worse, study shows

The LLM's ability to generate computer code got worse in a matter of months, according to Stanford and UC Berkeley researchers.

BY ANDREW PAUL | PUBLISHED JUL 19, 2023 6:00 PM EDT

Critique



ADVENTURES IN OPACITY -

Study claims ChatGPT is losing capability, but some experts aren't convinced

Either way, experts think OpenAI should be less opaque about its AI model architecture.

BENJ EDWARDS - 7/19/2023, 3:14 PM

Falcon 180B - open model from TII Abu-Dhabi



New Open-Source 'Falcon' AI Language Model Overtakes Meta and Google

Meta's massive, 70-billion parameter LLaMA 2 has been outflanked by the newly released 180-billion parameter Falcon Large Language Model.



☐ Sep 6, 2023

③ 3 min read

"Falcon 180B surpasses LLaMA 2 and other models in both scale and benchmark performance across a range of natural language processing (NLP) tasks."

"When compared against ChatGPT, the model is more powerful than the free version but a little less capable than the paid "plus" service."

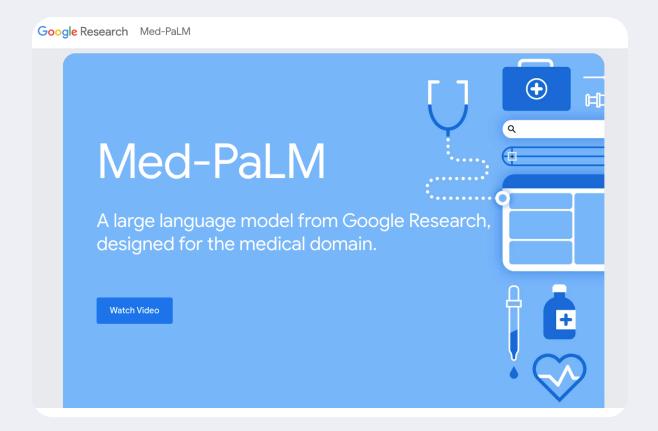
Google's Gemini - coming this fall



"Gemini was created from the ground up to be multimodal, highly efficient at tool and API integrations and built to enable future innovations, like memory and planning."

3. Specialized models are being developed.

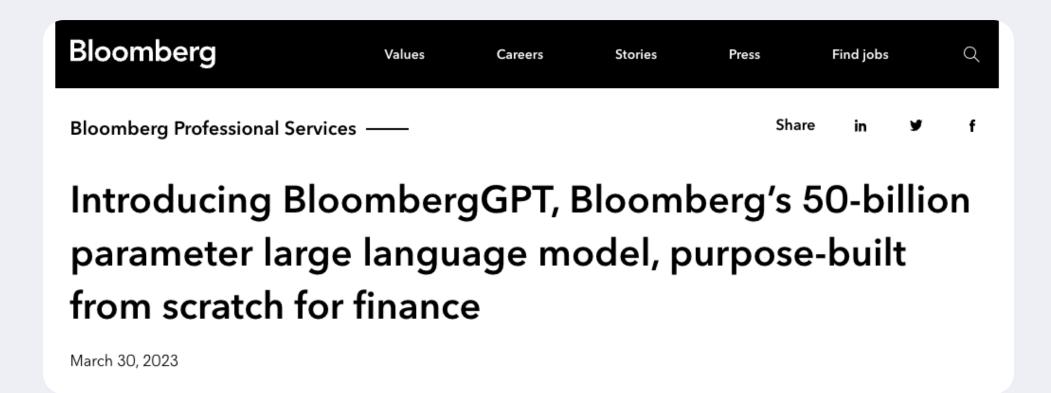
Med-PaLM from Google



Designed to provide **high quality answers to medical questions**.



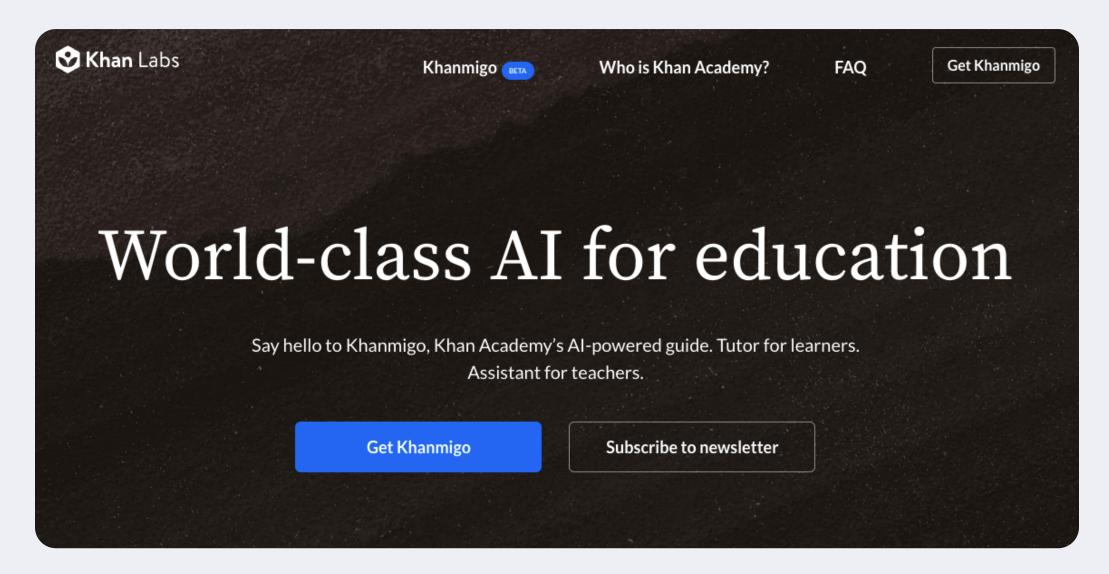
LLM for finance from Bloomberg



Specifically <u>trained on a wide range of financial data</u> to support a diverse set tasks within the financial industry.



Custom models for education - tutoring



Khanmigo from Khan Academy. Al-powered tutoring.

How Al Could Save (Not Destroy) Education | Sal Khan | TED

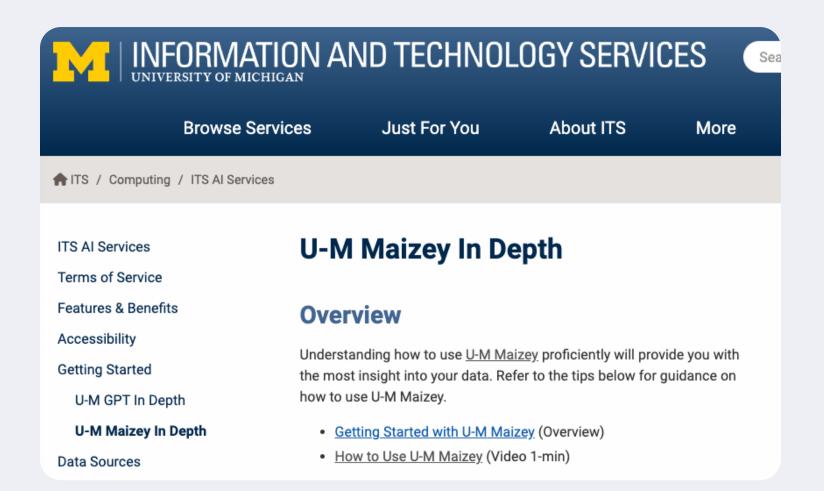
Custom models for education - language learning

duolingo blog Topics ∨ Authors Archive Q duolingo.com **MARCH 14, 2023** Introducing Duolingo Max, a learning experience powered by GPT-4 by Duolingo Team

Language learning app: powered by GPT-4.

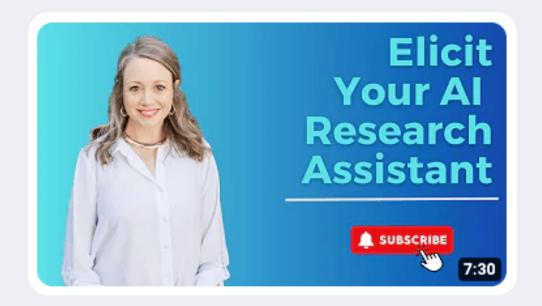


U Michigan campus providing custom model



<u>Maizey is a generative AI customization</u> tool that allows U of M faculty, staff, and students to build their own version of U-M GPT that is <u>trained on a custom</u> <u>data</u> set they provide.

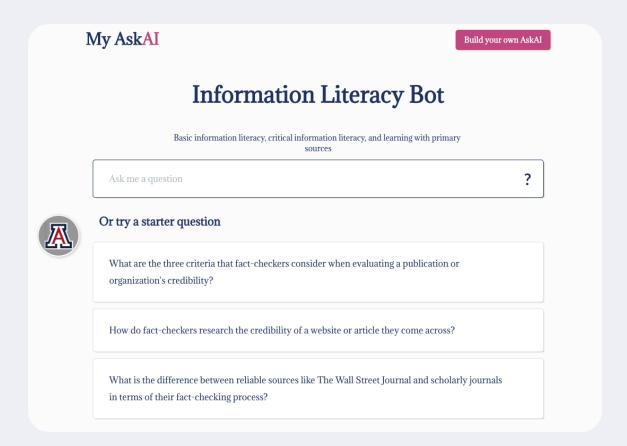
Research tools based on language models



<u>Elicit – Your Al Research Assistant</u> (YouTube)
<u>Elicit FAQ</u>

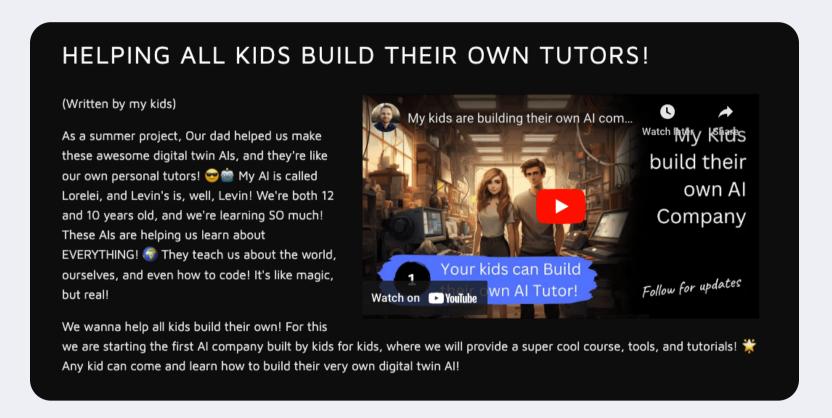


4. Build chatbots based on your own data (now, and it's easy)





Kids building chatbots



As a summer project, Our dad helped us make these awesome digital twin Als, and they're like our own personal tutors! My Al is called Lorelei, and Levin's is, well, Levin! We're both 12 and 10 years old, and we're learning SO much!

5. Exponential change

Open-source LLMs are now only a few months behind proprietary ones.

And it's only going to continue. As the McKinsey analysis shows, estimates for the time until software systems can reach human-level performance have fallen dramatically. Over the past year and a quarter, the consensus, as evaluated on Metacalculus, for the first weakly general AI has come forward from 2042 to 2027.

—<u>Exponential View, Azeem Azhar</u>

Azeem Azhar Creator of Exponential View

The Exponential Age

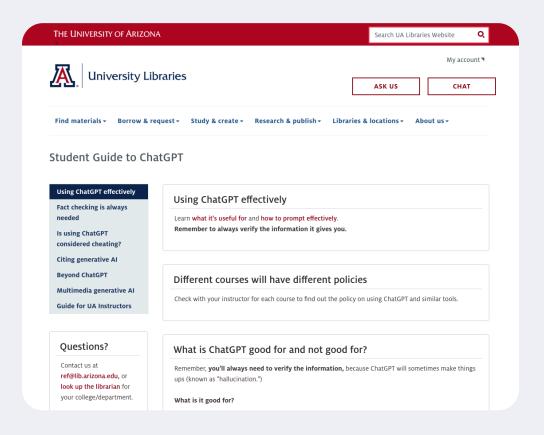
How Accelerating Technology is Transforming Business, Politics and Society



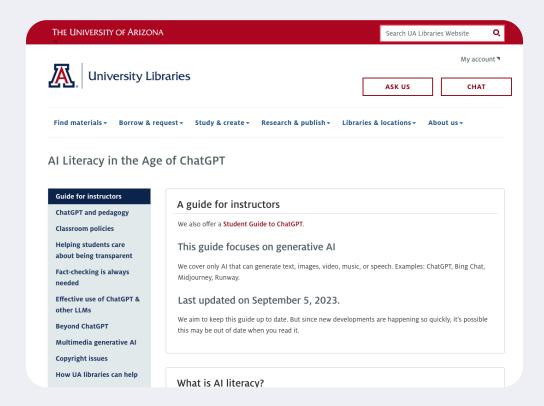
Let's teach students and ourselves to use generative Al as a creative tool to augment our powers.

UA Libraries resources

Student guide

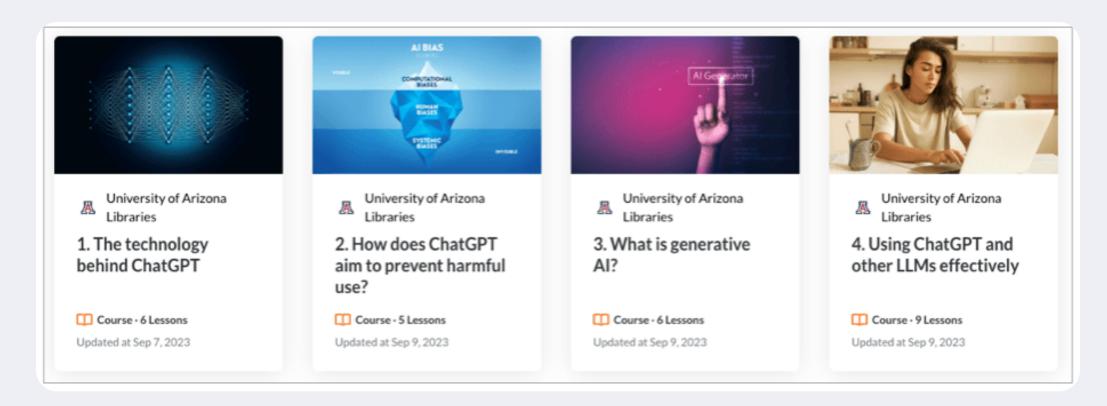


Instructor's guide



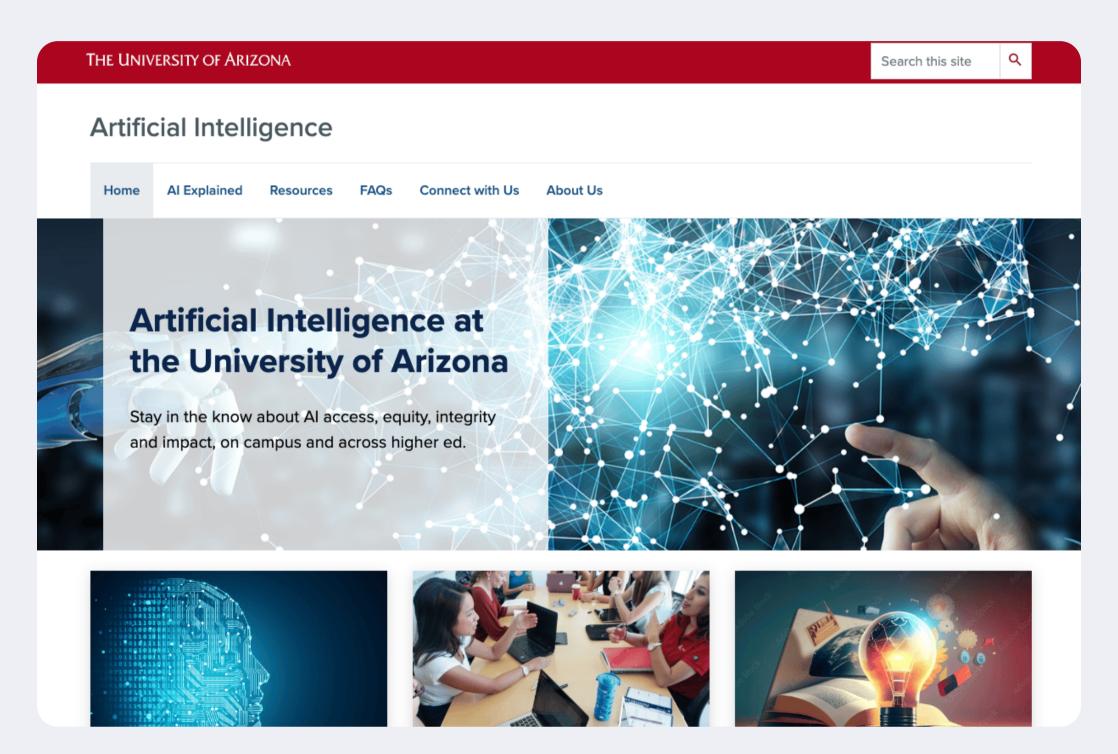


UA Libraries - Tutorials coming soon



Not yet live, but will be linked here: https://libguides.library.arizona.edu/learn-tutorials

Artificial Intelligence website for UA



https://artificialintelligence.arizona.edu



Stay current

Use and share the handout

Follow me on **Twitter** or **Mastodon**



Questions & discussion

Thank you!

nhennig@arizona.edu

https://nicolehennig.com