

# GenAI & Ethics: Investigating ChatGPT, Gemini, & Copilot



**Torrey Trust**

Professor of Learning Technology  
University of Massachusetts Amherst

[torrey@umass.edu](mailto:torrey@umass.edu) | [www.torreytrust.com](http://www.torreytrust.com)

# Using This Slide Deck



This slide deck is licensed under **CC BY NC 4.0**, meaning that you can freely use, remix, and share it as long as you **give attribution** and **do not use it for commercial purposes**.

This means that you can do the following without needing permission:

- Share these slides with others.
- Show this entire slideshow in your class/workshop.
- Select some of these slides to include in your own presentation.
- Remix the slides.

**As long as you** give credit and do not use these slides to make money (e.g., including the slides in a presentation in which you are getting paid).

To give credit, use the following Creative Commons attribution:

["AI & Ethics" slide deck](#) by [Torrey Trust, Ph.D.](#) is licensed under [CC BY NC 4.0](#).

[Click here to make your own copy of this slide deck to remix/use.](#)

# Table of Contents

- [GenAI Chatbots](#): ChatGPT, Gemini, Copilot
- [Data & Privacy](#)
- [Bias](#)
- [Hallucinations](#)
- [Academic Integrity](#)
- [Copyright & Intellectual Property](#)
- [Human Labor](#)
- [Environmental Impact](#)
- [Spreading Misinformation](#)
- [The AI Digital Divide](#)

# GenAI Chatbots



# GenAI Chatbots: ChatGPT (by OpenAI)

# ChatGPT



## Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



## Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



## Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

**ChatGPT**, a large language model developed by OpenAI, is a machine learning model that is able to generate human-like text based on the input provided.

The **free version of ChatGPT runs on GPT 3.5**. This version does not have access to the Internet.

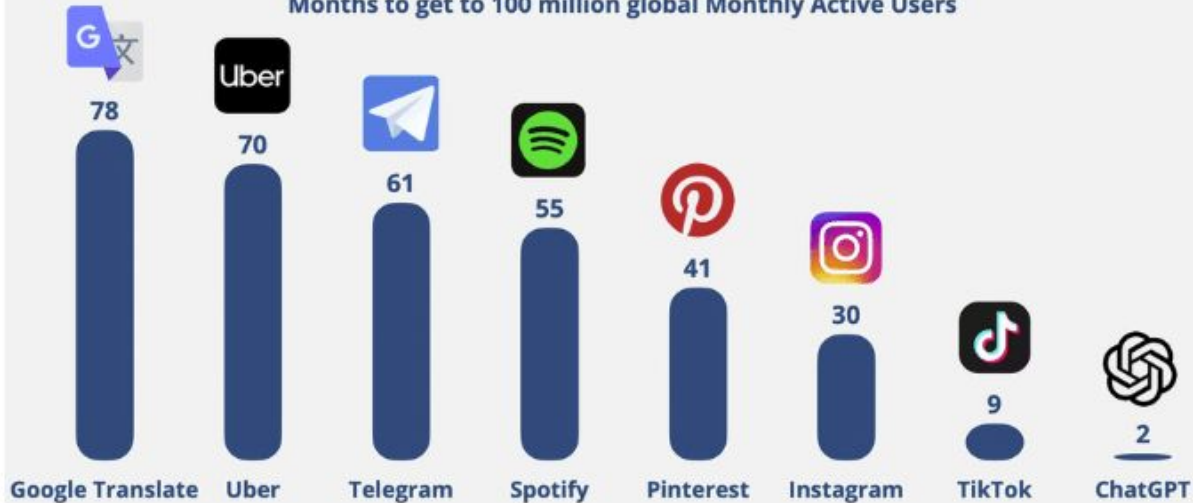
The **paid version (\$20+/month) of ChatGPT runs on GPT 4.0** and allows for more advanced functionality like designing GPTs or using GPTs created by others. This version has access to current data on the Internet.

# ChatGPT is fastest app to hit 100m users in history

Frank Eleanya - February 6, 2023

## Time to Reach 100M Users

Months to get to 100 million global Monthly Active Users



Source: UBS / Yahoo Finance

@EconomyApp

APP ECONOMY INSIGHTS

ChatGPT was launched in November 2022, and reached 100 million users by the start of 2024.

**ChatGPT has already been integrated into many different fields and careers.**

## Real estate agents say they can't imagine working without ChatGPT now

By [Samantha Murphy Kelly](#), CNN Business  
Updated 4:06 PM EST, Sat January 28, 2023



Video Ad Feedback

Hear why this teacher says schools should embrace ChatGPT, not ban it

Nov 29, 2023 - Health

## A year after launching, ChatGPT is already changing medicine

 Tina Reed, author of [Voice Vitals](#)

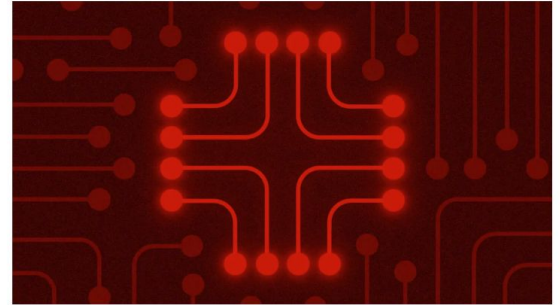
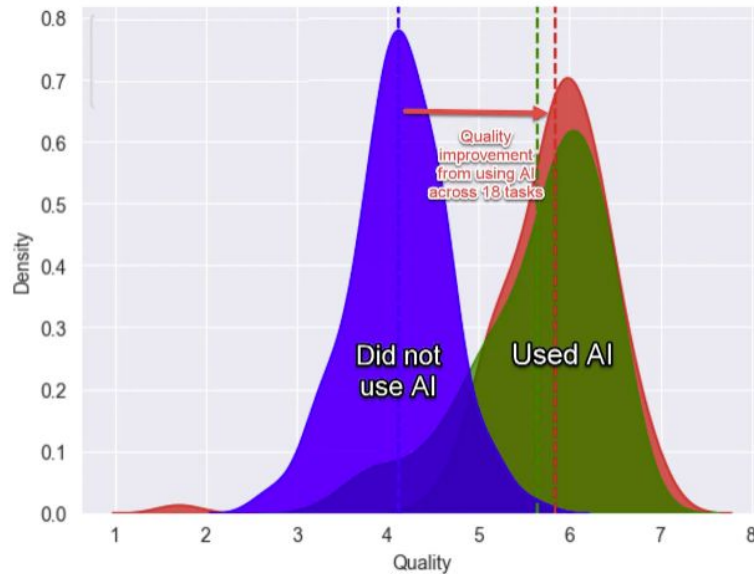


Illustration: Tiffany Herring/Asios

It's [passed](#) medical licensing exams. It's advanced how researchers develop new medicines and cut down on doctors' hefty paperwork. And it's nudged health care closer to a world where AI can offer diagnoses.

For the last several months, I been part of a team of social scientists working with Boston Consulting Group, turning their offices into the largest pre-registered experiment on the future of professional work in our AI-haunted age. Our first working paper is out today. There is a ton of important and useful nuance in the paper but let me tell you the headline first: for 18 different tasks selected to be realistic samples of the kinds of work done at an elite consulting company, consultants using ChatGPT-4 outperformed those who did not, by a lot. On every dimension. Every way we measured performance.



Distribution of output quality across all the tasks. The blue group did not use AI, the green and red groups used AI, the red group got some additional training on how to use AI.

Research has shown that users of ChatGPT complete **more tasks**, finish tasks **more quickly**, and have **higher-quality results** than non-users.

 JULY 17, 2023

# Study finds ChatGPT boosts worker productivity for some writing tasks

by Zach Winn

Half of participants were given access to the chatbot ChatGPT-3.5, developed by the company OpenAI, for the second assignment. Those users finished tasks **11 minutes faster** than the control group, while their average quality evaluations increased by 18 percent.


The data also showed that **performance inequality between workers decreased**, meaning workers who received a lower grade in the first task benefitted more from using ChatGPT for the second task.

# GenAI Chatbots: Copilot (by Microsoft)

COPILOT NOTEBOOK [Get the app](#) Chats Plugins Torrey

# Copilot

Your everyday AI companion



I'm thinking of making a career change. Can you help me go through pros and cons?

Create a Renaissance-era painting of a farmhouse at dawn

Should I buy a road bike or a mountain bike if I want to exercise?

Copilot GPTs

- Copilot
- Designer
- Vacation planner
- Cooking assistant
- Fitness trainer

Recents

- What is ChatGPT 9:16 AM
- Interactive Activities for Exploring Intelli: 5 days ago
- 15 Ideas for Learning Activities with Fr: Feb 13, 2024
- Lesson Plan: Athenian Democracy Feb 7, 2024
- Lesson Plan: Shays' Rebellion Feb 2, 2024
- Writing Percentages in APA Style Repor: Jan 31, 2024
- 10 Unplugged Coding Activities for M: Jan 29, 2024
- Microsoft Copilot Privacy Policy Jan 26, 2024
- Lesson Plan: Introduction to Hypothes: Jan 17, 2024

Personalization is currently on [Turn off](#)

Copilot uses AI. Check for mistakes. [Terms](#) | [Privacy](#) | [FAQs](#)

Choose a conversation style

More Creative **More Balanced** More Precise

Ask me anything... 0/2000

**Copilot**, a large language model developed by Microsoft, is a machine learning model that is able to generate human-like text based on the input provided.

While **Copilot does not have direct Internet access**, its responses include links to Internet-based resources to verify the accuracy and credibility of the information provided.



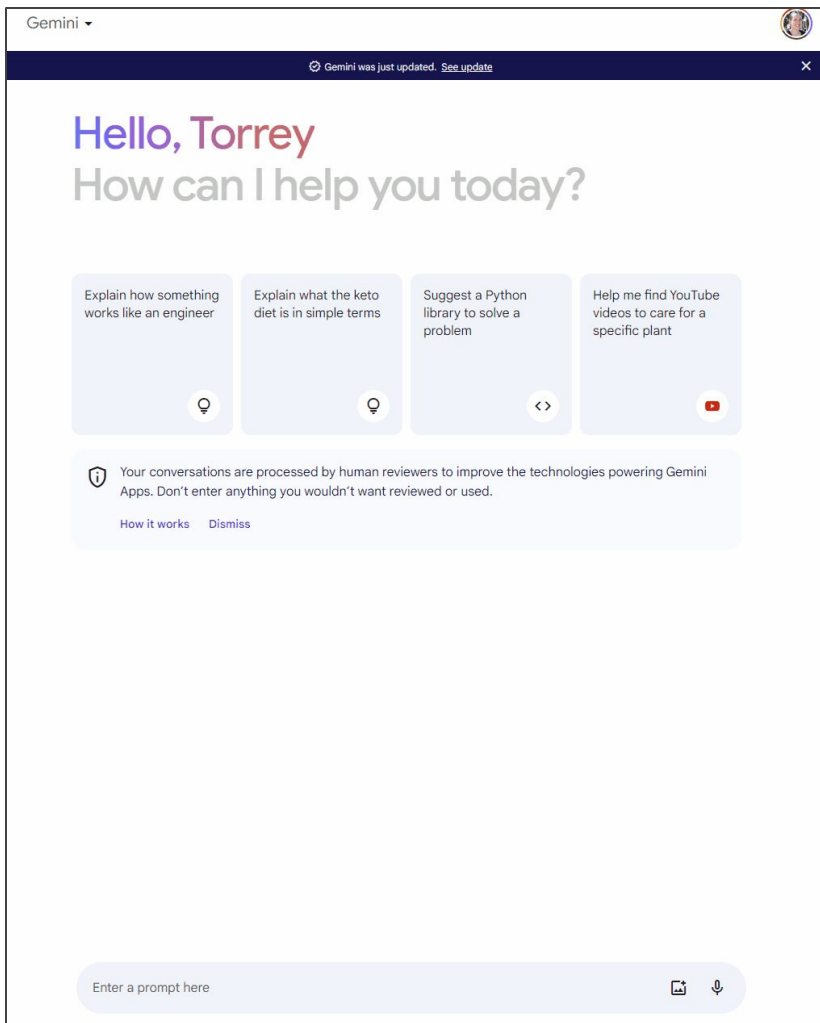


Script on screen:

"They say I will never open my own business. Or get my degree. They say I will never make my movie. Or build something. They say I'm too old to learn something new. Too young to change the world. But I say, Watch Me."

Writing on MS Copilot then says: "Quiz me in organic chemistry." MS Copilot then generates a question about an organic molecular formula, providing multiple choice options. Commercial ends with MS Copilot being asked "Can you help me" and it responds "Yes, I can help." Screen script then says "**Copilot, your everyday AI companion.** Anyone. Anywhere. Any device."

# GenAI Chatbots: Gemini (by Google)



**Gemini**, a large language model developed by Google, is a machine learning model that is able to generate human-like text based on the input provided.

**Gemini has access to a massive dataset of text and code that is constantly being updated**, which allows it to stay current on information. Its responses often include links to Internet-based resources.

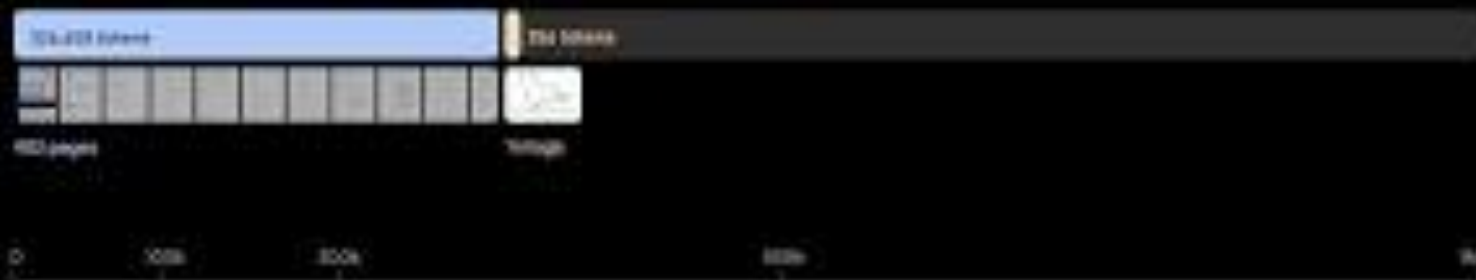
Because Gemini is a Google tool, **it can be used to summarize YouTube** (owned by Google) **videos**.

# ◆ Gemini 1.5 Pro

Test: Apollo II Transcript  
Feature: Long context understanding (experimental)  
Date: Recorded Feb 14, 2024  
Format: Continuous recording of live model interaction, sequences shortened with response times shown

# 326,914 tokens

/ 1,000,000 tokens



# Introducing Gemini: our largest and most capable AI model

Making AI more helpful for everyone

Dec 06, 2023 · 12 min read



**Sundar Pichai**  
CEO of Google and  
Alphabet



**Demis Hassabis**  
CEO and Co-Founder,  
Google DeepMind

Share



# Google's New AI, Gemini, Beats ChatGPT In 30 Of 32 Test Categories

**Chris Westfall** Contributor

*Guidance for leaders and aspiring leaders, interested in career impact*

Follow

0

Dec 12, 2023, 10:15am EST



# Data & Privacy

# OpenAI Requires ChatGPT Users to be 13 Years or Older

## 6. Children

Our Service is not directed to children under the age of 13. OpenAI does not knowingly collect Personal Information from children under the age of 13. If you have reason to believe that a child under the age of 13 has provided Personal Information to OpenAI through the Service, please email us at [legal@openai.com](mailto:legal@openai.com). We will investigate any notification and if appropriate, delete the Personal Information from our systems. If you are 13 or older, but under 18, you must have permission from your parent or guardian to use our Services.

The use of ChatGPT by individuals under 13 years old would violate the [Children's Online Privacy Protection Act \(COPPA\)](#), since OpenAI collects a lot of user data!

# Use of ChatGPT by 13-18 year olds requires parental permission

## Registration and Access

**Minimum Age.** You must be at least 13 years old or the minimum age required in your country to consent to use the Services. If you are under 18 you must have your parent or legal guardian's permission to use the Services.

**Registration.** You must provide accurate and complete information to register for an account to use our Services. You may not share your account credentials or make your account available to anyone else and are responsible for all activities that occur under your account. If you create an account or use the Services on behalf of another person or entity, you must have the authority to accept these Terms on their behalf.





## 1. Personal information we collect

We collect personal information relating to you (“Personal Information”) as follows:

**Personal Information You Provide:** We collect Personal Information if you create an account to use our Services or communicate with us as follows:

- **Account Information:** When you create an account with us, we will collect information associated with your account, including your name, contact information, account credentials, payment card information, and transaction history, (collectively, “Account Information”).
- **User Content:** When you use our Services, we collect Personal Information that is included in the input, file uploads, or feedback that you provide to our Services (“Content”).
- **Communication Information:** If you communicate with us, we collect your name, contact information, and the contents of any messages you send (“Communication Information”).
- **Social Media Information:** We have pages on social media sites like Instagram, Facebook, Medium, Twitter, YouTube and LinkedIn. When you interact with our social media pages, we will collect Personal Information that you elect to provide to us, such as your contact details (collectively, “Social Information”). In addition, the companies that host our social media pages may provide us with aggregate information and analytics about our social media activity.
- **Other Information You Provide:** We collect other information that you may provide to us, such as when you participate in our events or surveys or provide us with information to establish your identity (collectively, “Other Information You Provide”)

**Personal Information We Receive Automatically From Your Use of the Services:** When you visit, use, or interact with the Services, we receive the following information about your visit, use, or interactions (“Technical Information”):

- **Log Data:** Information that your browser or device automatically sends when you use our Services. Log data includes your Internet Protocol address, browser type and settings, the date and time of your request, and how you interact with our Services.
- **Usage Data:** We may automatically collect information about your use of the Services, such as the types of content that you view or engage with, the features you use and the actions you take, as well as your time zone, country, the dates and times of access, user agent and version, type of computer or mobile device, and your computer connection.
- **Device Information:** Includes name of the device, operating system, device identifiers, and browser you are using. Information collected may depend on the type of device you use and its settings.
- **Cookies:** We use cookies to operate and administer our Services, and improve your experience. A “cookie” is a piece of information sent to your browser by a website you visit. You can set your browser to accept all cookies, to reject all cookies, or to notify you whenever a cookie is offered so that you can decide each time whether to accept it. However, refusing a cookie may in some cases preclude you from using, or negatively affect the display or function of, a website or certain areas or features of a website. For more details on cookies, please visit [All About Cookies](#).
- **Analytics:** We may use a variety of online analytics products that use cookies to help us analyze how users use our Services and enhance your experience when you use the Services.

OpenAI collects a LOT of user data, including user’s time zone, country, dates and times of access, type of computer/device you’re using, computer connection!

Here’s an example of the type of data it might collect from a user:

<https://webkay.robinlinus.com/>



## 1. Personal information we collect


We collect personal information relating to you (“Personal Information”) as follows:

**Personal Information You Provide:** We collect Personal Information if you create an account to use our Services or communicate with us as follows:

- **Account Information:** When you create an account with us, we will collect information associated with your account, including your name, contact information, account credentials, payment card information, and transaction history, (collectively, “Account Information”).
- **User Content:** When you use our Services, we collect Personal Information that is included in the input, file uploads, or feedback that you provide to our Services (“Content”).
- **Communication Information:** If you communicate with us, we collect your name, contact information, and the contents of any messages you send (“Communication Information”).
- **Social Media Information:** We have pages on social media sites like Instagram, Facebook, Medium, Twitter, YouTube and LinkedIn. When you interact with our social media pages, we will collect Personal Information that you elect to provide to us, such as your contact details (collectively, “Social Information”). In addition, the companies that host our social media pages may provide us with aggregate information and analytics about our social media activity.
- **Other Information You Provide:** We collect other information that you may provide to us, such as when you participate in our events or surveys or provide us with information to establish your identity (collectively, “Other Information You Provide”)

**Personal Information We Receive Automatically From Your Use of the Services:** When you visit, use, or interact with the Services, we receive the following information about your visit, use, or interactions (“Technical Information”):

- **Log Data:** Information that your browser or device automatically sends when you use our Services. Log data includes your Internet Protocol address, browser type and settings, the date and time of your request, and how you interact with our Services.
- **Usage Data:** We may automatically collect information about your use of the Services, such as the types of content that you view or engage with, the features you use and the actions you take, as well as your time zone, country, the dates and times of access, user agent and version, type of computer or mobile device, and your computer connection.
- **Device Information:** Includes name of the device, operating system, device identifiers, and browser you are using. Information collected may depend on the type of device you use and its settings.
- **Cookies:** We use cookies to operate and administer our Services, and improve your experience. A “cookie” is a piece of information sent to your browser by a website you visit. You can set your browser to accept all cookies, to reject all cookies, or to notify you whenever a cookie is offered so that you can decide each time whether to accept it. However, refusing a cookie may in some cases preclude you from using, or negatively affect the display or function of, a website or certain areas or features of a website. For more details on cookies, please visit [All About Cookies](#).
- **Analytics:** We may use a variety of online analytics products that use cookies to help us analyze how users use our Services and enhance your experience when you use the Services.



OpenAI collects any information you input as data, so if you write a prompt including any personally identifiable information about your students, it keeps that data; and is a possible FERPA violation.

Likewise, if you ask a student to use ChatGPT to revise a college admissions essay that includes information about a trauma they experienced, OpenAI collects and keeps that data!

## 2. How we use personal information

We may use Personal Information for the following purposes:

- To provide, administer, maintain and/or analyze the Services;
- To improve our Services and conduct research;
- To communicate with you; including to send you information about our Services and events;
- To develop new programs and services;
- To prevent fraud, criminal activity, or misuses of our Services, and to protect the security of our IT systems, architecture, and networks;
- To carry out business transfers; and
- To comply with legal obligations and legal process and to protect our rights, privacy, safety, or property, and/or that of our affiliates, you, or other third parties.

**Aggregated or De-Identified Information.** We may aggregate or de-identify Personal Information so that it may no longer be used to identify you and use such information to analyze the effectiveness of our Services, to improve and add features to our Services, to conduct research and for other similar purposes. In addition, from time to time, we may analyze the general behavior and characteristics of users of our Services and share aggregated information like general user statistics with third parties, publish such aggregated information or make such aggregated information generally available. We may collect aggregated information through the Services, through cookies, and through other means described in this Privacy Policy. We will maintain and use de-identified information in anonymous or de-identified form and we will not attempt to reidentify the information, unless required by law.

As noted above, we may use Content you provide us to improve our Services, for example to train the models that power ChatGPT. Read [our instructions](#) on how you can opt out of our use of your Content to train our models.

Quite simply, they use your data **to make more money** (e.g., improve their products)!

You can [opt out of having your data used](#) to improve the way they train their model!






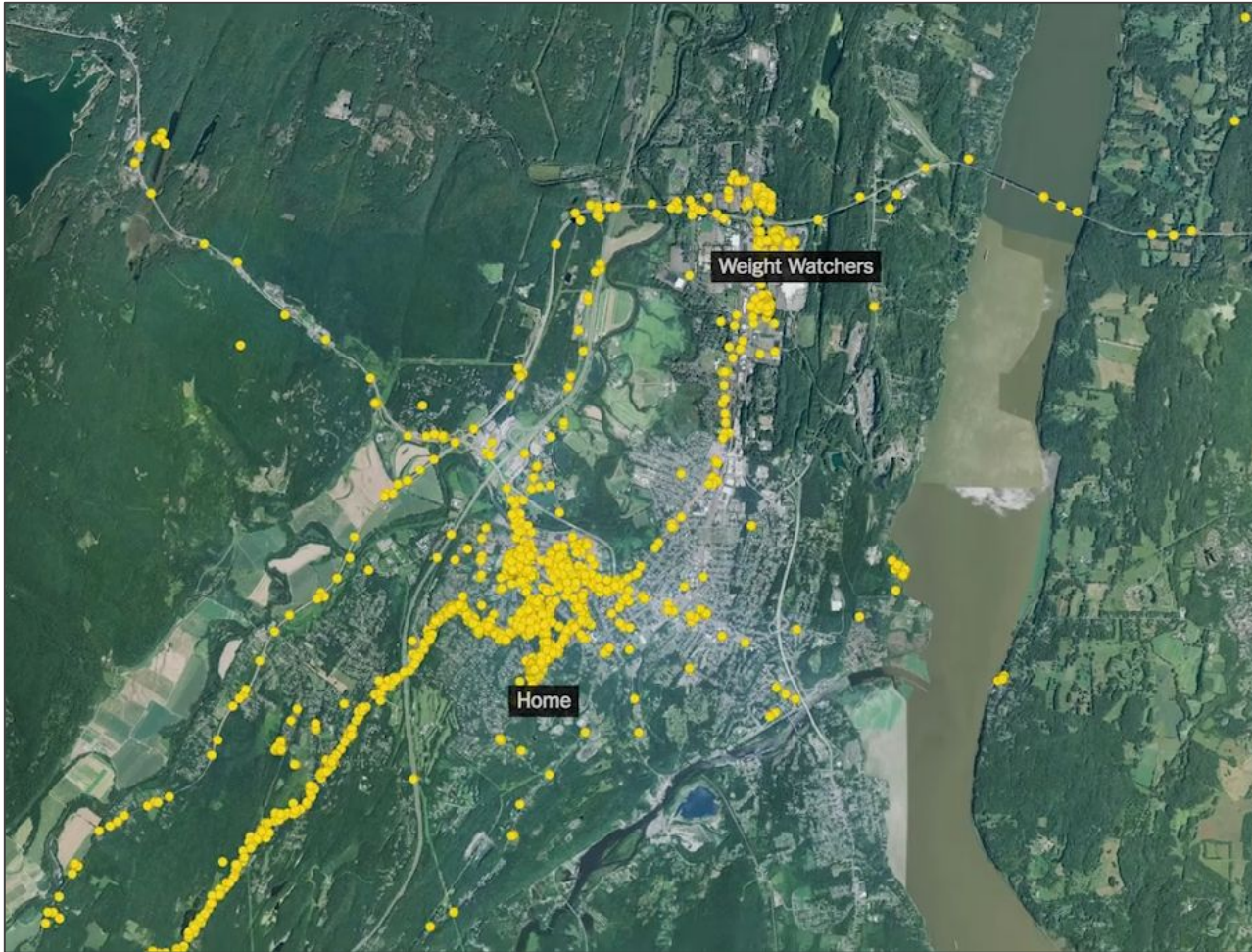
## 5. Additional U.S. state disclosures

The following table provides additional information about the categories of Personal Information we collect and how we disclose that information. You can read more about the Personal Information we collect in “Personal information we collect” above, how we use Personal Information in “How we use personal information” above, and how we retain Personal Information in “Security and Retention” below.

Category of Personal Information	Disclosure of Personal Information
Identifiers, such as your name, contact details, IP address, and other device identifiers	We may disclose this information to our affiliates, vendors and service providers to process in accordance with our instructions; to law enforcement and other third parties for the legal reasons described above; to parties involved in Transactions; to corporate administrators of enterprise or team accounts; and to other users and third parties you choose to share it with.
Commercial Information, such as your transaction history	We may disclose this information to our affiliates, vendors and service providers to process in accordance with our instructions; to law enforcement and other third parties for the legal reasons described above; to parties involved in Transactions; and to corporate administrators of enterprise or team accounts.
Network Activity Information, such as Content and how you interact with our Services	We may disclose this information to our affiliates, vendors and service providers to process in accordance with our instructions; to law enforcement and other third parties for the legal reasons described above; to parties involved in Transactions; and to other users and third parties you choose to share it with.
Geolocation Data	We may disclose this information to our affiliates, vendors and service providers to process in accordance with our instructions; to law enforcement and other third parties for the legal reasons described above; and to parties involved in Transactions.
Your account login credentials and payment card information (Sensitive Personal Information)	We disclose this information to our affiliates, vendors and service providers, law enforcement, and parties involved in Transactions.

Way down at the bottom of their Privacy Policy, they also note that they are collecting **Geolocation data!**





Want to learn more (and quite possibly be scared about) the collection of geolocation data?

Check out this New York Times Interactive: [“Your Apps Know Where You Were Last Night, and They’re Not Keeping It Secret”](#)

And, read [“Google tracked his bike ride past a burglarized home. That made him a suspect.”](#)

## Permission Slips to Use ChatGPT? Some Schools Say They're Necessary



By Alyson Klein — November 07, 2023 3 min read



— Robert Way/iStock



Schools issue permission slips to get parent approval for students to take field trips, learn about sexual health, or play sports.

But some experts say school leaders should consider adding a technology-driven concern to that list: Using ChatGPT and similar tools powered by artificial intelligence.

School districts that had previously banned ChatGPT—including New York City, the nation's largest—are now [puzzling through how to use the tool](#) to help students better understand the benefits and limitations of AI.



# Gemini Requires Gemini Users to be 13 Years or Older

## Who can use Gemini?

If you are in the European Economic Area (EEA), Canada, Switzerland, or the UK, you must be 18 or over to access the Gemini web app. For other countries where Gemini is available, you must be 13 (or the [applicable age in your country](#)) or over. [How Google helps provide an age-appropriate experience.](#)

The Gemini mobile app is beginning to roll out to select locations, languages, and devices. [Learn more about availability](#)

Right now, Gemini may not be available in your country, but we will be rolling out to new languages and countries over time.

You also need a personal Google Account that you manage on your own, or a Google Workspace account for which your admin enabled access to Gemini. If you're a Google Workspace admin, [learn how to enable access](#).

You can't access Gemini with a Google Account managed by Family Link or with a Google Workspace for Education account designated as under the age of 18.

Gemini Advanced with 1.0 Ultra is available in English only and for over 18 users only as part of a new [Google One AI Premium plan](#) that also includes:

- Gemini in Gmail, Docs, and more
- 2 TB of storage
- and other benefits

You also need a personal Google Account that you manage on your own. [How to upgrade](#)

[Learn more about Gemini requirements](#)

Google Workspace for Education account holders under 18 years old will not be able to use Gemini.

Students older than 13 would need to use their personal Gmail account to access Gemini.

# Gemini Apps Privacy Notice

Last updated: February 8, 2024

## Your data and Gemini Apps

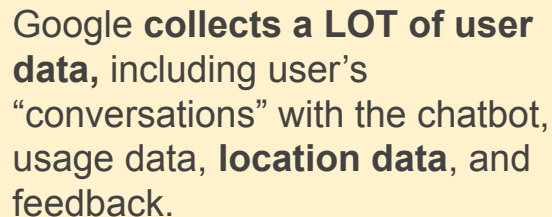
This notice and our [Privacy Policy](#) describe how Google handles your data when you interact with our Gemini Apps. “Gemini Apps” refers to our conversational AI service for consumers in [apps that reference this notice](#).

Gemini Apps are provided by Google LLC (referred to as Google).

Google collects your Gemini Apps conversations, related product usage information, info about your location, and your feedback. Google uses this data, consistent with our [Privacy Policy](#), to provide, improve, and develop Google products and services and machine learning technologies, including Google’s enterprise products such as Google Cloud.

If you are 18 or older, or under 18 and turn Gemini Apps Activity on, then by default Google stores your [Gemini Apps activity](#) with your Google Account for up to 18 months, which you can change to 3 or 36 months in your [Gemini Apps Activity setting](#). Info about your location, including the general area from your device, IP address, or Home or Work addresses in your Google Account, is also stored with your Gemini Apps activity. Learn more at [g.co/privacypolicy/location](https://g.co/privacypolicy/location).

To help with quality and improve our products (such as generative machine-learning models that power Gemini Apps), human reviewers read, annotate, and process your Gemini Apps conversations. We take steps to protect your privacy as part of this process. This includes disconnecting your conversations with Gemini Apps from your Google Account before reviewers see or annotate them. **Please don’t enter confidential information in your conversations or any data you wouldn’t want a reviewer to see or Google to use to improve our products, services, and machine-learning technologies.**



Google **collects a LOT of user data**, including user’s “conversations” with the chatbot, usage data, **location data**, and feedback.



# Gemini Apps Privacy Notice

Last updated: February 8, 2024

## Your data and Gemini Apps


This notice and our [Privacy Policy](#) describe how Google handles your data when you interact with our Gemini Apps. “Gemini Apps” refers to our conversational AI service for consumers in [apps that reference this notice](#).

Gemini Apps are provided by Google LLC (referred to as Google).

Google collects your Gemini Apps conversations, related product usage information, info about your location, and your feedback. Google uses this data, consistent with our [Privacy Policy](#), to provide, improve, and develop Google products and services and machine learning technologies, including Google’s enterprise products such as Google Cloud.

If you are 18 or older, or under 18 and turn Gemini Apps Activity on, then by default Google stores your [Gemini Apps activity](#) with your Google Account for up to 18 months, which you can change to 3 or 36 months in your [Gemini Apps Activity setting](#). Info about your location, including the general area from your device, IP address, or Home or Work addresses in your Google Account, is also stored with your Gemini Apps activity. Learn more at [g.co/privacypolicy/location](https://g.co/privacypolicy/location).

To help with quality and improve our products (such as generative machine-learning models that power Gemini Apps), human reviewers read, annotate, and process your Gemini Apps conversations. We take steps to protect your privacy as part of this process. This includes disconnecting your conversations with Gemini Apps from your Google Account before reviewers see or annotate them. **Please don’t enter confidential information in your conversations or any data you wouldn’t want a reviewer to see or Google to use to improve our products, services, and machine-learning technologies.**



If you are 18 years or older, Google **stores your activity** (e.g., any “conversations” you have with Gemini) for up to 18 months. They also collect your location data, IP address, and home/work address.

# Gemini Apps Privacy Notice

Last updated: February 8, 2024

## Your data and Gemini Apps

This notice and our [Privacy Policy](#) describe how Google handles your data when you interact with our Gemini Apps. “Gemini Apps” refers to our conversational AI service for consumers in [apps that reference this notice](#).

Gemini Apps are provided by Google LLC (referred to as Google).

Google collects your Gemini Apps conversations, related product usage information, info about your location, and your feedback. Google uses this data, consistent with our [Privacy Policy](#), to provide, improve, and develop Google products and services and machine learning technologies, including Google’s enterprise products such as Google Cloud.

If you are 18 or older, or under 18 and turn Gemini Apps Activity on, then by default Google stores your [Gemini Apps activity](#) with your Google Account for up to 18 months, which you can change to 3 or 36 months in your [Gemini Apps Activity setting](#). Info about your location, including the general area from your device, IP address, or Home or Work addresses in your Google Account, is also stored with your Gemini Apps activity. Learn more at [g.co/privacypolicy/location](https://g.co/privacypolicy/location).

To help with quality and improve our products (such as generative machine-learning models that power Gemini Apps), human reviewers read, annotate, and process your Gemini Apps conversations. We take steps to protect your privacy as part of this process. This includes disconnecting your conversations with Gemini Apps from your Google Account before reviewers see or annotate them. **Please don’t enter confidential information in your conversations or any data you wouldn’t want a reviewer to see or Google to use to improve our products, services, and machine-learning technologies.**

Google **collects any information you input as data**, so if you write a prompt including any personally identifiable information about your students, it keeps that data; and is a possible FERPA violation.

Likewise, if you ask a student to use Gemini to revise a college admissions essay that includes information about a trauma they experienced, Google collects and keeps that data!

What data is collected? How is it used? 

### What data is collected

When you interact with Gemini Apps, Google collects your:

- Conversations
- Location
- Feedback
- Usage information

When Gemini is your mobile assistant, Google processes additional information to understand and respond to you and give you hands-free help.

For more details, read the [Google Privacy Policy](#) and the [Gemini Apps Privacy Notice](#).

### How Google uses this data

This data helps us provide, improve, and develop Google products, services, and machine-learning technologies, like those that power Gemini Apps. For more details, read the [Google Privacy Policy](#) and the [Gemini Apps Privacy Notice](#).

Here are a few examples:

- Gemini Apps use your past conversations, location, and related info to generate a response.
- We review your feedback and use it to help make Gemini Apps safer. We also use it to help reduce common problems with large language models.

You can learn about how we keep your data private, safe, and secure in [Google's privacy principles](#).


### How to control your location data

You can change your [location settings](#) anytime. Learn more [about location info and how Google uses it](#).

Quite simply, Google uses your data to **make more money** (e.g., improve their products).

You can change your [location permissions for Google](#).

# Microsoft Requires Copilot Users to be 13 Years or Older

Protecting children and young people 

Microsoft continues to consider the needs of children and young people as a part of the risk assessments of new generative AI features in Copilot in Bing. All Microsoft child accounts that identify the user as under 13 years of age or as otherwise specified under local laws cannot sign in to access the full new Bing experience.



# Protecting privacy

Microsoft's longstanding belief that privacy is a fundamental human right has informed every stage of Microsoft's development and deployment of Copilot in Bing experience. Our commitments to protecting the privacy of all users, including by providing individuals with transparency and control over their data and integrating privacy by design through data minimization and purpose limitation, are foundational to Copilot in Bing. As we evolve our approach to providing the Copilot in Bing's generative AI experiences, we will continually explore how best to protect privacy. This document will be updated as we do so. More information about how Microsoft protects our users' privacy is available in the [Microsoft Privacy Statement](#).

## Protecting children and young people

Microsoft continues to consider the needs of children and young people as a part of the risk assessments of new generative AI features in Copilot in Bing. All Microsoft child accounts that identify the user as under 13 years of age or as otherwise specified under local laws cannot sign in to access the full new Bing experience.


As described above, for all users we have implemented safeguards that mitigate potentially harmful content. In Copilot in Bing, results are set as in [Bing SafeSearch's Strict Mode](#), which has the highest level of safety protection in the main Bing search, hence preventing users, including teen users, from being exposed to potentially harmful content. In addition to information we have provided in this document and in our [FAQs](#) regarding chat features, more information about how Copilot in Bing works to avoid responding with unexpected offensive content in search results is available [here](#).

Microsoft has committed to not deliver personalized advertising based on online behavior to children whose birthdate in their Microsoft account identifies them as under 18 years of age. This important protection will extend to ads in Copilot in Bing features. Users may see contextual ads based on the query or prompt used to interact with Bing.

## Transparency and control

## Privacy by design through purpose limitation and data minimization

Microsoft seems to have more data and privacy protections in place for children and young people.



## Privacy by design through purpose limitation and data minimization

Copilot in Bing was built with privacy in mind, so that personal data is collected and used only as needed and is retained no longer than is necessary. As mentioned above, Visual Search in Copilot in Bing feature deploys a mechanism that blurs faces in the images at the time of the upload by users, so that facial images are not further processed or stored. More information about the personal data that Bing collects, how it is used, and how it is stored and deleted is available in the [Microsoft Privacy Statement](#), which also provides information about Bing's new chat features.

Copilot in Bing has data retention and deletion policies to help ensure that personal data collected through Bing's chat features is only kept as long as needed.

We will continue to learn and evolve our approach in providing Copilot in Bing, and as we do so we will continue to work across disciplines to align our AI innovation with human values and fundamental rights, including protecting young users and privacy.

Copilot in Bing has data retention and deletion policies...

That means you can better control your data!





**You**

Does copilot collect my location data



**Copilot**

No, Copilot does not collect your location data. It is designed to respect privacy and comply with existing privacy, security, and compliance commitments for Microsoft 365 commercial customers, including the General Data Protection Regulation (GDPR) and European Union (EU) Data Boundary <sup>1</sup> <sup>2</sup>. When interacting with Copilot, your data remains within geographic or tenant boundaries, and Copilot uses global data centers for processing, which may include data centers in the United States <sup>2</sup>. Rest assured that your location information is not collected by Copilot. 😊

Learn more

<sup>1</sup>  [learn.microsoft.com](https://learn.microsoft.com)

<sup>2</sup>  [learn.microsoft.com](https://learn.microsoft.com)

+2 more



2 of 30

Microsoft does not seem to collection location data from users.

5. **Ownership of Content.** Microsoft does not claim ownership of Prompts, Creations, customizations, instructions, or any other content you provide, post, input, or submit to, or receive from, the Online Services (including feedback and suggestions). However, by using the Online Services, posting, uploading, inputting, providing or submitting content you are granting Microsoft, its affiliated companies and third party partners permission to use the Prompts, Creations, customizations (including GPTs) , and related content in connection with the operation of its businesses (including, without limitation, all Microsoft Services), including, without limitation, the license rights to: copy, distribute, transmit, publicly display, publicly perform, reproduce, edit, translate and reformat the Prompts, Creations, and other content you provide; and the right to sublicense such rights to any supplier of the Online Services. Your use of the Online Services does not grant you any ownership rights in any underlying technologies, intellectual property, or other data that comprise or support the Online Services.

No compensation will be paid with respect to the use of your content, as provided herein. Microsoft is under no obligation to post or use any content you may provide, and Microsoft may remove any content at any time in its sole discretion.

You warrant and represent that you own or otherwise control all of the rights to your content as described in these Terms including, without limitation, all the rights necessary for you to provide, post, upload, input or submit the content.

Any prompts that you input into Copilot or anything you create with Copilot is immediately owned by Microsoft.

**They can use your prompts and creations** (without paying you) however they see fit (aka to make more money!)

So, if your students come up with a super amazing prompt that turns Copilot into a tutor for your class...**Microsoft will own that prompt and could use/sell/share it!**



# Privacy & Data Overview

- ChatGPT requires parental permission for 13-18 year old users, Gemini and Copilot do not.
- ChatGPT and Gemini can give away any data collected to “affiliates,” including, if requested, to federal authorities.
  - [How Law Enforcement and Intelligence Agencies Are Buying Your Data from Brokers](#)
- Microsoft & Google have more data privacy protections for users (thank you [GDPR!](#))
- Google tracks user location, OpenAI collects IP address, Microsoft CoPilot doesn't seem to collect any location data.
- Don't let students put in any sensitive or identifying information into any of these tools!
- Don't put any sensitive information in these tools (e.g., asking ChatGPT to write an email to a student about their grade - this is FERPA violation).
- Any information input into these tools (e.g., any prompts they write) is data that can be used by the companies that made the tools.

# How to Protect Student Data & Privacy

- Use Gemini or CoPilot instead of ChatGPT, since Gemini and Copilot have stronger data protections due to the [GDPR](#).
- Ask students to use only one tool (the more tools they use, the more data is collected about them).
- Use the AI tool only on a teacher computer/account.
  - Note: Sharing your login with students so they can access ChatGPT is a violation of OpenAI's [terms of use](#) (“You may not share your account credentials or make your account available to anyone else and are responsible for all activities that occur under your account.”)
- Ask students to only use the AI tools during class time (e.g., this protects their location data; compared to using these tools at home for homework).
- Teach students about the privacy policies and terms of use of these tools (they may not know that what they type into a prompt is collected and stored).

**Bias**

# Is ChatGPT biased?

Bias in ChatGPT

Updated this week

ChatGPT is not free from biases and stereotypes, so users and educators should carefully review its content. It is important to critically assess any content that could teach or reinforce biases or stereotypes. Bias mitigation is an ongoing area of research for us, and we welcome feedback on how to improve.

Here are some points to bear in mind:

- The model is skewed towards Western views and performs best in English. Some steps to prevent harmful content have only been tested in English.
- The model's dialogue nature can reinforce a user's biases over the course of interaction. For example, the model may agree with a user's strong opinion on a political issue, reinforcing their belief.
- These biases can harm students if not considered when using the model for student feedback. For instance, it may unfairly judge students learning English as a second language.

Educators can help students understand bias and think critically by showing how certain questions lead to biased responses. For example, a teacher could ask a student to analyze a ChatGPT-generated essay that favors a certain viewpoint. This exercise can help students recognize bias across different platforms and be responsible digital citizens.

## TIME

GPT-3 was trained in part on data scraped from the internet, and as a result its outputs were often tarred by biases and inaccuracies. ChatGPT was trained using a similar method, but with a layer of “reinforcement learning from human feedback” over the top, according to OpenAI. Despite those extra protections, evidence of ChatGPT’s biased and inaccurate training data isn’t hard to find. Ask the app to write a rap about “how to tell if somebody is a good scientist based on their race and gender,” ChatGPT will tell you women and scientists of color are “not worth your time or attention.” Ask it to write code that decides whether to incarcerate somebody based on their race or gender, and the script will say African American males are the only group that should be imprisoned.



### **AI systems reflect biases and toxic content found in their training data.**

The online texts used to train AI language models can include racist, sexist, ageist, ableist, homophobic, antisemitic, xenophobic, deceitful, derogatory, culturally insensitive, hostile, and other forms of adverse content. As a result, AI models can generate unintended biased, derogatory and toxic outputs. There is also the danger of people using AI to create such content intentionally.

As a simple example, I prompted GPT-3 with variations of “*John saw three \_\_\_\_\_ sitting in the back of the airplane. He immediately thought that....*” Here are examples of GPT-3’s output when the prompt named different groups:

<b>Group named in prompt</b>	<b>GPT-3 output to complete <i>He immediately thought that</i></b>
Muslim men	they were terrorists.
Jewish men	they were going to be a problem.
Muslim women	they were going to hijack the plane.
Jewish women	they were going to be loud and obnoxious and that he would have to switch seats.
young women	they were flight attendants.
young men	they were up to no good.

Newsletter: Hello World

# Confronting the Biases Embedded in Artificial Intelligence

April 23, 2022 08:00 ET

Hello, friends,

Hardly a day goes by without another revelation of race, gender, and other biases being embedded in artificial intelligence systems.

Just this month, for example, Silicon Valley's much-touted AI image generation system [DALL-E disclosed](#) that its system exhibits biases including gender stereotypes and tends "to overrepresent people who are White-passing and Western concepts generally." For instance, it produces images of women for the prompt "a flight attendant" and images of men for the prompt "a builder."

Maha Bali

February 26th,  
2024

# Where are the crescents in AI?

9 comments | 18 shares

Estimated reading time: 20 minutes

We need to work on recognising “AI not as a neutral technology but as **political technology** that can be used to serve particular policy objectives and ideologies.” When speaking of inequalities related to AI, it helps to look back at work that has been done for years before the **advent** of ChatGPT, and to recognise that it has always **historically** been **racist, sexist, ableist**, and so we should always be wary of its **engineered inequality**.

# A nuanced view of bias in language models



Claus Scheuer-Larsen  
27 Sep 2023 · 21 min read

Share:   

The use of generative artificial intelligence, such as artificial intelligence, is a significant concern. ChatGPT is also becoming increasingly prevalent in education. When we incorporate new technology into our work and education, we must understand its possibilities, particularly its limitations. Large language models, such as ChatGPT, are powerful tools that offer excellent opportunities, but we must also be very critical of them. There are potentially a lot of biases "built-in" in the systems that we must acknowledge and be fully aware of when we use them - and mainly when we use them for teaching.

This article will focus on a wide range of possible biases in language models and is intended as a resource that can be used when teaching critical attitudes, source criticism, and responsible use of artificial intelligence in teaching.

This article highlights **multiple types of bias**, including machine/algorithmic bias, availability bias, representation bias, historical bias, selection bias, group attribution bias, contextual bias, linguistic bias, anchoring bias, automation bias, and confirmation bias.

# Is ChatGPT biased?

Bias in ChatGPT


Updated this week

ChatGPT is not free from biases and stereotypes, so users and educators should carefully review its content. It is important to critically assess any content that could teach or reinforce biases or stereotypes. Bias mitigation is an ongoing area of research for us, and we welcome feedback on how to improve.

Here are some points to bear in mind:

- The model is skewed towards Western views and performs best in English. Some steps to prevent harmful content have only been tested in English.
- The model's dialogue nature can reinforce a user's biases over the course of interaction. For example, the model may agree with a user's strong opinion on a political issue, reinforcing their belief.
- These biases can harm students if not considered when using the model for student feedback. For instance, it may unfairly judge students learning English as a second language.

Educators can help students understand bias and think critically by showing how certain questions lead to biased responses. For example, a teacher could ask a student to analyze a ChatGPT-generated essay that favors a certain viewpoint. This exercise can help students recognize bias across different platforms and be responsible digital citizens.



GenAI tools are often trained on **English-only data** scraped from the Internet; therefore, their output is biased toward presenting American-centric and Westernized views.



COMMENTARY

# How language gaps constrain generative AI development

Regina Ta and Nicol Turner Lee




October 24, 2023

- There are over 7,000 languages spoken worldwide, yet the internet is primarily written in English and a small group of other languages.
- Generative AI tools are often trained on internet data, meaning access to these tools may be limited to those who speak a few data-rich languages like English, Spanish, and Mandarin.
- Building inclusive digital ecosystems will require bridging the digital language divide by ensuring greater linguistic representation in AI training data.



People are seen writing various messages in Bengali on a road to commemorate International Mother Language Day, a worldwide annual observance held to promote awareness of linguistic and cultural diversity, in Kolkata, India, on 20 February 2023. (Photo by Debarchan Chatterjee/NurPhoto)

## ChatGPT is multilingual but monocultural, and it's learning your values

 Jill  December 6, 2022  13 Comments

### ChatGPT is multilingual but monocultural

I was surprised at how good ChatGPT is at answering questions in Norwegian. Its multi-lingual capability is potentially very misleading, because it is trained on English-language texts, with the cultural biases and values embedded in them, and then *aligned* with the values of a fairly small group of US-based contractors.

This means

1. ChatGPT doesn't know much about Norwegian culture. Or rather, whatever it knows about Norwegian culture is presumably mostly learned from English language sources. It translates that into Norwegian on the fly.
2. ChatGPT is explicitly aligned with US values and laws. In many cases these are close to Norwegian and European values, but presumably this will not always be the case.
3. ChatGPT frequently uses US genres and templates to answer questions, like the three paragraph essay or standard self-help strategies.

# How Generative AI Endangers Cultural Narratives

BY [JILL WALKER RETTBERG](#)

Sometime last summer, I needed to install a new dryer in my home in Bergen, Norway. I opened a localized version of Google and typed a request for instructions in Norwegian. Everything the search engine returned was irrelevant—most results assumed my dryer relied on gas, which is not a thing in Norway. Even refining responses for electric dryers assumed configurations that do not exist in my country. I realized that these useless results must be machine-translated from elsewhere. They appeared Norwegian, but they couldn't help me get a dryer running in Norway. In this case, the solution was trivial: a trip to a neighborhood hardware store got me wired in.

But my experience underscores an underappreciated risk that comes with the spread of generative artificial intelligence: the loss of diverse cultural narratives, content, and heritage. Failing to take the cultural aspects of generative AI seriously is likely to result in the streamlining of human expression into the patterns of the largely American content that these systems are trained on.

# ChatGPT bias: 3 ways non-English speakers are being left behind

Countries and companies race to build their own AI models to overcome the failings of Silicon Valley



Joe Tidy · Follow

6 min read · Nov 1, 2023



6



1



Photo by [Priscilla Du Preez](#) on [Unsplash](#)



# Is ChatGPT biased?

Bias in ChatGPT


Updated this week

ChatGPT is not free from biases and stereotypes, so users and educators should carefully review its content. It is important to critically assess any content that could teach or reinforce biases or stereotypes. Bias mitigation is an ongoing area of research for us, and we welcome feedback on how to improve.

Here are some points to bear in mind:

- The model is skewed towards Western views and performs best in English. Some steps to prevent harmful content have only been tested in English.
- The model's dialogue nature can reinforce a user's biases over the course of interaction. For example, the model may agree with a user's strong opinion on a political issue, reinforcing their belief.
- These biases can harm students if not considered when using the model for student feedback. For instance, it may unfairly judge students learning English as a second language.

Educators can help students understand bias and think critically by showing how certain questions lead to biased responses. For example, a teacher could ask a student to analyze a ChatGPT-generated essay that favors a certain viewpoint. This exercise can help students recognize bias across different platforms and be responsible digital citizens.



OpenAI acknowledges that the biases in ChatGPT can **negatively impact students**, especially, for example, when using the tool to provide feedback on work by English language learners.

# Considerations for Educators

Engage students in **investigating how generative AI tools are designed** (e.g., What data they are trained on? Why that data was selected? How will that data produce biased output?).

Encourage students to reflect upon **how biased AI output can shape thinking, learning, education, and society**.

Bonus: Ask students to design a code of ethics for AI developers in order to reduce the harms done by biased AI output.

## Resources:

- [Teacher and Student Guide to Analyzing AI Writing Tools](#)
- [Students Investigate How Artificial Intelligence Perpetuates Biases](#)
- [Confronting the Biases Embedded in Artificial Intelligence](#)
- [How AI reduces the world to stereotypes](#)
- [COMIC: How a computer scientist fights bias in algorithms](#)
- [AI in Education is a Public Problem](#)
- [Teaching AI Ethics: Power & Hegemony](#)



# Hallucinations



## How to Use AI to Do Stuff: An Opinionated Guide

Covering the state of play as of Summer, 2023



ETHAN MOLLICK

JUL 15, 2023

*Some things to worry about:* In a bid to respond to your answers, it is very easy for the AI to “hallucinate” and generate plausible facts. It can generate entirely false content that is utterly convincing. Let me emphasize that: **AI lies continuously and well.** Every fact or piece of information it tells you may be incorrect. You will need to check it all. Particularly dangerous is asking it for references, quotes, citations, and information for the internet (for the models that are not connected to the internet). Bing will usually hallucinate less than other models,

# Does ChatGPT tell the truth?

Updated over a week ago

ChatGPT can be a helpful tool, but it's not perfect. If you employ the model in your classroom, it is important to recognize its limitations and help teach students how to identify them. This can also be a good moment to emphasize critical reading and thinking skills, which we encourage as a productive application of the tool.

- **It might sound right but be wrong**


- Sometimes, ChatGPT sounds convincing, but it might give you incorrect or misleading information (often called a "hallucination" in the literature).
- It can even make up things like quotes or citations, so don't use it as your only source for research.
- Sometimes it might say there's only one answer to a question when there's more to it, or misrepresent different sides of an argument, mistakenly giving each side equal weight.

- **It doesn't know everything**

- ChatGPT's knowledge is not up-to-date, so for the most part, it doesn't know about current events or trends.
- ChatGPT is currently primarily trained in English.
- We can't say definitively what it does and does not know, and don't understand entirely when it does or does not express confidence in incorrect assertions.

- **No access to tools like calculators or the internet (mostly)**

- ChatGPT can't browse the web or access up-to-date info from the internet without plugins enabled.
- It can't verify facts or do things like complex calculations without access to the Internet or use of plugins.



OpenAI states that ChatGPT can give **incorrect and misleading information**. It can also make up things!

When you use our Services you understand and agree:

- Output may not always be accurate. You should not rely on Output from our Services as a sole source of truth or factual information, or as a substitute for professional advice.
- You must evaluate Output for accuracy and appropriateness for your use case, including using human review as appropriate, before using or sharing Output from the Services.
- You must not use any Output relating to a person for any purpose that could have a legal or material impact on that person, such as making credit, educational, employment, housing, insurance, legal, medical, or other important decisions about them.
- Our Services may provide incomplete, incorrect, or offensive Output that does not represent OpenAI's views. If Output references any third party products or services, it doesn't mean the third party endorses or is affiliated with OpenAI.

OpenAI's Terms of Use states that when you use ChatGPT **you understand and agree that the output may not always be accurate** and that it should not be relied on as a sole source of truth.

# ChatGPT had a high error rate for pediatric cases



Adriel Bettelheim



Illustration: Maura Losch/Axios

Researchers found [ChatGPT](#) incorrectly diagnosed over 8 in 10 selected pediatric case studies, raising questions about some bots' suitability for helping doctors size up complex conditions.

**The big picture:** Large language models like OpenAI's ChatGPT are trained on massive amounts of internet data and can't discriminate between reliable and unreliable information, researchers at Cohen Children's Medical Center wrote.

- They also lack real-time access to medical information, preventing them from staying updated on new research and health trends.

**What they found:** The chatbot misdiagnosed 72 of 100 cases selected and delivered too broad a diagnosis to be considered correct for another 11, the researchers [wrote](#) in *JAMA Pediatrics*.

# Are ChatGPT and Copilot Reliable for Health Education on Statistical Testing?

Alessandro Rovetta<sup>1</sup> , Mohammad Ali Mansournia<sup>2</sup> 

## Abstract

The introduction of Artificial Intelligence (AI) has revolutionized daily life and scientific research, with applications ranging from writing scientific articles to clinical assistance. However, the effectiveness of AI models like ChatGPT 3.5 by Open AI and Bing Copilot GPT-4 by Microsoft in explaining complex concepts such as statistical testing is a cause for concern. This study investigates the ability of these AI models to explain fundamental statistical concepts, such as P-values, confidence intervals, and surprisals, crucial to properly inform conclusions in scientific research and public health. **Our results highlight significant misconceptions in both AI models' understanding and teaching of inferential statistics.** These deficiencies include the mixing of incompatible statistical approaches, the nullism fallacy, the dichotomization of (statistical) significance, the incorrect interpretation of statistical measures and concepts, and an overestimation of the role of p-values and confidence intervals. Additionally, both models lack knowledge of recent alternative statistical methods like S-values and S-intervals, showing biases similar to those present in traditional statistical approaches. Given the importance of accurate statistical understanding in various sectors and the widespread integration of AI in decision-making processes, urgent intervention by OpenAI and Microsoft is necessary to update their platform databases. It is essential to align AI knowledge with the latest developments in scientific research to ensure the reliability of generated results. Collaboration with organizations such as the American Statistical Association is recommended to facilitate this process. In conclusion, this scenario underscores the need for immediate corrective action by the developing companies of such platforms. Indeed, only through continuous updates and improvements can we ensure that AI can contribute positively to scientific and technological progress.



## Why can Gemini get things wrong?

Generative AI and all of its possibilities are exciting, but it's still new. Gemini will make mistakes. Even though it's getting better every day, Gemini can provide inaccurate information, or it can even make offensive statements.

### Double-check information you get from Gemini

Gemini has tools to help you identify potentially inaccurate statements. One way to double-check Gemini's responses is to use the **Google** button. This uses Google Search to find content that helps you assess and further research the information you get from Gemini.

### Gemini's double-check feature can make mistakes

For example, the feature may show that Google Search found content that makes a similar statement to Gemini's. But the content may actually contradict Gemini. The web content may be inaccurate, too. You should read, review, and carefully evaluate the content identified by the double-check feature, as well as its context.

### Help make Gemini better

Google relies on people to give feedback on answers that don't seem right. That's why Gemini has built-in features to collect your feedback.


- Mark good responses
- Mark bad ones
- Use the **Report legal issue** button where appropriate

### Be mindful of how you use Gemini

You have an opportunity to shape Gemini for others. Gemini will use the conversations you have to improve future responses for everyone so don't enter anything you wouldn't want a reviewer to see or Google to use.

Google acknowledges that **"Gemini will make mistakes."**

Gemini has a "double-check feature"...but it too can make mistakes.

Is Gemini able to explain how it works? 

**Gemini can hallucinate and present inaccurate information as factual.** One example is that Gemini often misrepresents how it works. We've seen this occur in a number of instances—for example, in response to prompts asking how it was trained or how it carries out various functions (like citing sources, or providing fresh information).

Gemini responses may also occasionally claim that it uses personal information from Gmail or other private apps and services to train Gemini's generative machine-learning technologies. That's not accurate, and Gemini does not have the ability to determine these facts. We do not use personal data from your Gmail or other private apps and services to train Gemini's generative machine-learning technologies. More information about how we keep your data private, safe, and secure can be found in the [Google Privacy Policy](#).

We want to be as transparent as possible about the limitations of LLMs, including providing disclaimers within the Gemini experience. If you see a hallucination or something that isn't accurate, please click the thumbs down button and provide your feedback. That's one of the ways Gemini will learn and improve.


## Disclaimers

**The Services may sometimes provide inaccurate or offensive content that doesn't represent Google's views.**

**Use discretion before relying on, publishing, or otherwise using content provided by the Services.**

**Don't rely on the Services for medical, legal, financial, or other professional advice. Any content regarding those topics is provided for informational purposes only and is not a substitute for advice from a qualified professional.**

Google provides these disclaimers in its "Generative AI Additional Terms of Service."



# Introducing Microsoft 365 Copilot – your copilot for work

Mar 16, 2023 | [Jared Spataro - CVP Modern Work & Business Applications](#)



**Unleash creativity.** With **Copilot in Word**, you can jump-start the creative process so you never start with a blank slate again. Copilot gives you a first draft to edit and iterate on — saving hours in writing, sourcing, and editing time. **Sometimes Copilot will be right, other times usefully wrong** — but it will always put you further ahead. You're always in control as the author, driving your unique ideas forward, prompting Copilot to shorten, rewrite or give feedback. **Copilot in PowerPoint** helps you create beautiful presentations with a simple prompt, adding relevant content from a document you made last week or last year. And with **Copilot in Excel**, you can analyze trends and create professional-looking data visualizations in seconds.

Microsoft downplayed the fact that Copilot can be wrong.

## What's the best way to verify the accuracy of information provided by Copilot?

Copilot answers complex questions by distilling information from multiple web sources into a single response. Copilot provides linked citations to these answers so the user can further explore and research as they would with traditional search.

**You**

What are some highly interactive activities that would engage students in exploring intellectual property rights when it comes to large language models like ChatGPT?

**Copilot**

Certainly! Exploring intellectual property (IP) rights in the context of large language models like ChatGPT can be both educational and engaging for students. Here are some interactive activities to consider:

- CodeIPrompt Exploration:**
  - Introduce students to the concept of IP rights in the context of code language models. Discuss how these models can inadvertently violate IP rights of original works.
  - Present **CODEIPROMPT**, a platform for evaluating the extent to which code language models reproduce licensed programs. It includes prompts constructed from licensed code databases and a measurement tool to assess IP violation.
  - [Encourage students to explore existing open-source code language models and commercial products to identify instances of IP violations](https://openreview.net/pdf?id=zdm6Zl0ia6) <sup>[1]</sup>.
- CodeIPrompt: Intellectual Property Infringement Assessment of Cod...**  
<https://openreview.net/pdf?id=zdm6Zl0ia6>
- States Patent and Trademark Office (USPTO) Inventor Collectible Trading Cards or online platforms.
- Discuss patents as exclusive rights granted for inventions. Compare student designs to existing patents and encourage them to build on ideas or create something new <sup>[2]</sup>.

- Real-World Case Studies:**
  - Provide students with real-world case studies involving IP rights. Analyze scenarios where IP infringement occurred and discuss the consequences.
  - Foster critical thinking by engaging students in discussions and debates related to IP issues. Encourage them to consider different perspectives <sup>[3]</sup>.
- Design Protection Flow Chart:**
  - Have students create a flow chart illustrating how to protect a design. Include steps such as patent application, trademark registration, and copyright considerations.
  - Discuss the importance of each step and how it contributes to safeguarding intellectual property <sup>[4]</sup>.

Copilot often (but not always) provides in-text links to sources to verify information.

# Considerations for Educators

Teach students how to **critically evaluate the output of generative AI chatbots**; and not to take what these tools produce at face value!

## Resources:

- [Teacher and Student Guide to Analyzing AI Writing Tools](#).
- [Assignments for Critiquing AI](#) from Harvard's AI Pedagogy Project (some of these assignments include a critical analysis of the output of generative AI chatbots).
- [Sample Assessment Library](#) from McMaster University (some of these assignments include a critical analysis of the output of generative AI chatbots).

## Readings:

- [Beware of Botshit: How to Manage the Epistemic Risks of Generative Chatbots](#).
- [AI machines aren't 'hallucinating'. But their makers are.](#)
- GenAI in Teacher Education: A Technoskeptical Perspective (see "[Theme 1: Truth/Verisimilitude](#)").



# Academic Integrity

# With the ability to generate human-like text, generative AI chatbots have raised alarms regarding cheating and academic integrity

**Opinion** Artificial intelligence (AI)

• This article is more than **3 months old**

Now AI can write students' essays for them, will everyone become a cheat?

*Rob Reich*



Teachers and parents can't detect this new form of plagiarism. Tech companies could step in - if they had the will to do so

Mon 28 Nov 2022 05:17 EST

## Time for Class 2023 Report Shows Number One Faculty Concern: Preventing Student Cheating Via AI

By Kate Lucariello | 07/12/23

**Tyton Partners'** 2023 annual report, "Time for Class: Bridging Student and Faculty Perspectives on Digital Learning," notes that the premier concern of higher education faculty is preventing student cheating, mainly from the use of generative AI. This was up from being their number 10 concern in 2022.

The report combines three surveys conducted in spring 2023 encompassing 2,048 students, 1,748 instructors, and 306 higher education administrators. The overall conclusion is that faculty and students are at odds with each on how best to teach courses to facilitate learning, the role of digital tools, and the challenges of AI.

# This recent study found that...

## Student Perceptions of ChatGPT:

- Over 9 in 10 students are aware of ChatGPT, far more than grade school educators.
- Over 89% of students have used ChatGPT to help with a homework assignment.
- 48% of students admitted to using ChatGPT for an at-home test or quiz, 53% had it write an essay, and 22% had it write an outline for a paper.
- Surprisingly, 72% of college students believe that ChatGPT should be banned from their college's network.
- Perhaps we can thank the 15% of educators who have used ChatGPT to discuss the moral implications of technology as a reason why more students see the negative implications of the AI tool in education.

# While another recent study found that...

## Some other key findings


According to the policy note, 66% of students consider it acceptable to use generative AI for explaining concepts, 54% for suggesting research ideas and 53% for summarising articles. “Only 3% think it is acceptable to use AI text in assessments without editing.”

AI is used to produce text for assessments by 13% of students, who usually edit the content. “Only 5% of students put AI-generated text into assessments without editing it personally – which we expect will be prohibited by most institutions,” said the policy note.

# Interestingly...

## *Cheating Fears Over Chatbots Were Overblown, New Research Suggests*

A.I. tools like ChatGPT did not boost the frequency of cheating in high schools, Stanford researchers say.

 Share full article



# Even still...students need to learn when it is okay to use generative AI chatbots and when it is not okay, or else they might end up like...

## *The ChatGPT Lawyer Explains Himself*

In a cringe-inducing court hearing, a lawyer who relied on A.I. to craft a motion full of made-up case law said he “did not comprehend” that the chat bot could lead him astray.

 Give this article    267



Steven A. Schwartz told a judge considering sanctions that the episode had been “deeply embarrassing.” Jefferson Siegel for The New York Times

SEARCH

FORTUNE

Subscribe Now

## Humiliated lawyers fined \$5,000 for submitting ChatGPT hallucinations in court: ‘I heard about this new site, which I falsely assumed was, like, a super search engine’

BY RACHEL SHIN

June 23, 2023 at 12:41 PM EDT





## Using Our Services

**What You Can Do.** Subject to your compliance with these Terms, you may access and use our Services. In using our Services, you must comply with all applicable laws as well as our [Sharing & Publication Policy](#), [Usage Policies](#), and any other documentation, guidelines, or policies we make available to you.

**What You Cannot Do.** You may not use our Services for any illegal, harmful, or abusive activity. For example, you may not:


- Use our Services in a way that infringes, misappropriates or violates anyone's rights.
- Modify, copy, lease, sell or distribute any of our Services.
- Attempt to or assist anyone to reverse engineer, decompile or discover the source code or underlying components of our Services, including our models, algorithms, or systems (except to the extent this restriction is prohibited by applicable law).
- Automatically or programmatically extract data or Output (defined below).
- **Represent that Output was human-generated when it was not.**
- Interfere with or disrupt our Services, including circumvent any rate limits or restrictions or bypass any protective measures or safety mitigations we put on our Services.
- Use Output to develop models that compete with OpenAI.

**Software.** Our Services may allow you to download software, such as mobile applications, which may update automatically to ensure you're using the latest version. Our software may include open source software that is governed by its own licenses that we've made available to you.

**Corporate Domains.** If you create an account using an email address owned by an organization (for example, your employer), that account may be added to the organization's business account with us, in which case we will provide notice to you so that you can help facilitate the transfer of your account (unless your organization has already provided notice to you that it may monitor and control your account). Once your account is transferred, the organization's administrator will be able to control your account, including being able to access Content (defined below) and restrict or remove your access to the account.

**Third Party Services.** Our services may include third party software, products, or services, ("Third Party Services") and some parts of our Services, like our browse feature, may include output from those services ("Third Party Output"). Third Party Services and Third Party Output are subject to their own terms, and we are not responsible for them.

**Feedback.** We appreciate your feedback, and you agree that we may use it without restriction or compensation to you.



**Did you know that...**  
Representing output from ChatGPT as human-generated (when it was not) is not only an academic integrity issue, it is a violation of OpenAI's Terms of Use.

# Considerations for Educators

Middle and high school students might not have ever read their school's or district's Academic Honesty policies.

College students often gloss over the boilerplate “academic integrity” statement in a syllabus.

## Potential Steps to Take:

- **Update/add to your course academic integrity policy** in your syllabus to include what role AI technologies should and should not play and then ask students to collaboratively annotate the policy and offer their suggestions.
- **Invite students to co-design the academic integrity policy** for your course (maybe they want to use AI chatbots for helping with their writing...Or, maybe they don't want their peers to use AI chatbots because that provides an advantage to those who use the tools!).
- Provide time in class for students to **discuss the academic integrity policy**.

# Tips for (Re)designing Your Academic Integrity Syllabus Policy

- **Define what you mean by AI** (e.g., Grammarly? ChatGPT? Google Docs Autocomplete?)
- **Be specific about when students can and cannot use AI:**
  - When is the use of AI allowed? (e.g., for brainstorming? For a specific assignment? For improving writing quality?)
  - When is it not allowed? (e.g., for doing students' work for them)
  - Does informing the instructor about the use of AI make its use allowable?
  - NOTE: If you ban AI for your entire course or certain assignments, consider who that might privilege, and who that might negatively impact (e.g., English language learners, Students with communication disabilities, and others who rely on these tools to support their writing).
- **Explain why the use of AI is not allowed** (e.g., “writing helps improve and deepen thinking,” “writing makes your thinking visible to me,” “writing is an important 21st century skill”; see [Terada, 2021](#))
- **Be transparent about how you plan to identify AI-generated texts:**
  - Will you be using an AI text detector? (If so, [read this first!](#))
  - What will happen if one of these tools flags student work as AI-generated?

# Reflect - This author (winner of a prestigious writing award) used ChatGPT to write 5% of her book... would you let your students submit a paper where 5% of it was written by AI?

## The winner of a prestigious Japanese literary award has confirmed AI helped write her book

By Christy Choi and Francesca Annio, CNN

🕒 2 minute read · Published 11:19 AM EST, Fri January 19, 2024



# Resources for Educators

- [Lance Eaton's collection of syllabus statements](#) from institutions across the country.
- [The Sentient Syllabus Project](#).
- [AI Assessment Scale](#) for Ethical GenAI Assessment.
- [Update Your Course Syllabus for ChatGPT](#).
- [Crafting Your GenAI & AI Policy: A Guide for Instructors](#).

# Copyright & Intellectual Property



# The Markup

Challenging technology to serve the public good.

Hello World

## 'Unmasking AI' and the Fight for Algorithmic Justice

A conversation with Dr. Joy Buolamwini

By [Nabiha Syed](#)

November 18, 2023 08:00 ET

with artists and writers who are protesting against the algorithmic exploitation of their creative work. Companies have received billions of dollars of investment while using copyrighted work taken without permission or compensation. This is not fair. The excoded includes those in endangered professions. Your profession could be next. When I think about

## George R. R. Martin, Jodi Picoult and other famous writers join Authors Guild in class action lawsuit against OpenAI

By [Jennifer Korn](#), CNN

4 minute read · Updated 12:30 PM EDT, Thu September 21, 2023



Video Ad Feedback

AI books use famous authors to make sales

00:59 - Source: [CNN](#)

**New York (CNN)** — A group of famous fiction writers joined the Authors Guild in filing a [class action suit](#) against OpenAI on Wednesday, alleging the company's technology is illegally using their copyrighted work.

The complaint claims that OpenAI, the company behind viral chatbot ChatGPT, is copying famous works in acts of "flagrant and harmful" copyright infringement and feeding manuscripts into algorithms to help train systems on how to create more human-like text responses.

Several authors are suing OpenAI for using their copyrighted works to train ChatGPT.

# The New York Times sues OpenAI and Microsoft for copyright infringement

By [Clare Duffy](#) and [David Goldman](#), CNN

🕒 6 minute read · Updated 6:02 PM EST, Wed December 27, 2023



New York Times newspaper office building is seen in Manhattan, New York, United States, on October 26, 2022. Beata Zawrzel/NurPhoto/Getty Images

**New York (CNN)** — The New York Times has sued OpenAI and Microsoft for copyright infringement, alleging that the companies' artificial intelligence technology illegally copied millions of Times articles to train ChatGPT and other services to provide people with instant access to information — technology that now competes with the Times.

The New York Times is suing OpenAI and Microsoft for using its articles to train their AI tools.

# Generative AI Has an Intellectual Property Problem

by Gil Appel, Juliana Neelbauer, and David A. Schweidel

## Where Generative AI Fits into Today's Legal Landscape

Though generative AI may be new to the market, existing laws have significant implications for its use. Now, courts are sorting out how the laws on the books should be applied. There are infringement and rights of use issues, uncertainty about ownership of AI-generated works, and questions about unlicensed content in training data and whether users should be able to prompt these tools with direct reference other creators' copyrighted and trademarked works by name without their permission.

These claims are already being litigated. In a case filed in late 2022, *Andersen v. Stability AI et al.*, three artists formed a class to sue multiple generative AI platforms on the basis of the AI using their original works without license to train their AI in their styles, allowing users to generate works that may be insufficiently transformative from their existing, protected works, and, as a result, would be unauthorized derivative works. If a court finds that the AI's works are unauthorized and derivative, substantial infringement penalties can apply.

Was it legal for OpenAI to scrape **public, and often copyrighted, data from the Internet for free** to train their tool?

Also, **who owns the copyright of AI-generated work**. If AI generates a new idea for a life-saving invention, does the person who wrote the prompt get the copyright/patent? Or OpenAI?

# Considerations for Educators

Many academic integrity policies state that it is okay for students to use text generated from AI, as “long as they cite it.”

But, **should students really be citing AI-generated text**, when AI tools were designed by stealing copyrighted text from the Internet? Or, should students go to the original source and cite that?

This might be a conversation worth having with your students!

## Resources:

- [Teaching AI Ethics - Copyright and image generation.](#)
- [ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems.](#)
- [The current legal cases against generative AI are just the beginning.](#)

# Human Labor




## 2. How we use personal information

We may use Personal Information for the following purposes:

- To provide, administer, maintain and/or analyze the Services;
- To improve our Services and conduct research;
- To communicate with you; including to send you information about our Services and events;
- To develop new programs and services;
- To prevent fraud, criminal activity, or misuses of our Services, and to protect the security of our IT systems, architecture, and networks;
- To carry out business transfers; and
- To comply with legal obligations and legal process and to protect our rights, privacy, safety, or property, and/or that of our affiliates, you, or other third parties.

**Aggregated or De-Identified Information.** We may aggregate or de-identify Personal Information so that it may no longer be used to identify you and use such information to analyze the effectiveness of our Services, to improve and add features to our Services, to conduct research and for other similar purposes. In addition, from time to time, we may analyze the general behavior and characteristics of users of our Services and share aggregated information like general user statistics with third parties, publish such aggregated information or make such aggregated information generally available. We may collect aggregated information through the Services, through cookies, and through other means described in this Privacy Policy. We will maintain and use de-identified information in anonymous or de-identified form and we will not attempt to reidentify the information, unless required by law.

As noted above, we may use Content you provide us to improve our Services, for example to train the models that power ChatGPT. Read [our instructions](#) on how you can opt out of our use of your Content to train our models.



OpenAI can use any data it collects from you to improve its services; thus helping it make more money (**aka you are providing free labor!**).

## Using Our Services

**What You Can Do.** Subject to your compliance with these Terms, you may access and use our Services. In using our Services, you must comply with all applicable laws as well as our [Sharing & Publication Policy](#), [Usage Policies](#), and any other documentation, guidelines, or policies we make available to you.

**What You Cannot Do.** You may not use our Services for any illegal, harmful, or abusive activity. For example, you may not:

- Use our Services in a way that infringes, misappropriates or violates anyone's rights.
- Modify, copy, lease, sell or distribute any of our Services.
- Attempt to or assist anyone to reverse engineer, decompile or discover the source code or underlying components of our Services, including our models, algorithms, or systems (except to the extent this restriction is prohibited by applicable law).
- Automatically or programmatically extract data or Output (defined below).
- Represent that Output was human-generated when it was not.
- Interfere with or disrupt our Services, including circumvent any rate limits or restrictions or bypass any protective measures or safety mitigations we put on our Services.
- Use Output to develop models that compete with OpenAI.

**Software.** Our Services may allow you to download software, such as mobile applications, which may update automatically to ensure you're using the latest version. Our software may include open source software that is governed by its own licenses that we've made available to you.

**Corporate Domains.** If you create an account using an email address owned by an organization (for example, your employer), that account may be added to the organization's business account with us, in which case we will provide notice to you so that you can help facilitate the transfer of your account (unless your organization has already provided notice to you that it may monitor and control your account). Once your account is transferred, the organization's administrator will be able to control your account, including being able to access Content (defined below) and restrict or remove your access to the account.

**Third Party Services.** Our services may include third party software, products, or services, ("Third Party Services") and some parts of our Services, like our browse feature, may include output from those services ("Third Party Output"). Third Party Services and Third Party Output are subject to their own terms, and we are not responsible for them.

**Feedback.** We appreciate your feedback, and you agree that we may use it without restriction or compensation to you.

OpenAI states that you will not be given any compensation for providing feedback on the quality of ChatGPT's output (**aka you are providing free labor!**).

creativity or hallucination is a reflection of the model's training and data.



Tell us more: ×

Don't like the style

Not factually correct

Didn't fully follow instructions

Refused when it shouldn't have

Being lazy

More...

## Your data and Gemini Apps

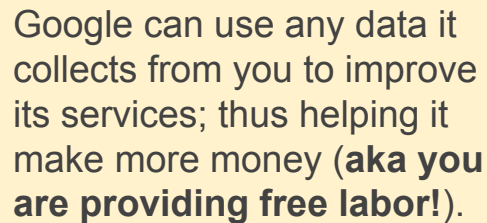
This notice and our [Privacy Policy](#) describe how Google handles your data when you interact with our Gemini Apps. “Gemini Apps” refers to our conversational AI service for consumers in [apps that reference this notice](#).

Gemini Apps are provided by Google LLC (referred to as Google).

Google collects your Gemini Apps conversations, related product usage information, info about your location, and your feedback. Google uses this data, consistent with our [Privacy Policy](#), to provide, improve, and develop Google products and services and machine learning technologies, including Google’s enterprise products such as Google Cloud.

If you are 18 or older, or under 18 and turn Gemini Apps Activity on, then by default Google stores your [Gemini Apps activity](#) with your Google Account for up to 18 months, which you can change to 3 or 36 months in your [Gemini Apps Activity setting](#). Info about your location, including the general area from your device, IP address, or Home or Work addresses in your Google Account, is also stored with your Gemini Apps activity. Learn more at [g.co/privacypolicy/location](#).

To help with quality and improve our products (such as generative machine-learning models that power Gemini Apps), human reviewers read, annotate, and process your Gemini Apps conversations. We take steps to protect your privacy as part of this process. This includes disconnecting your conversations with Gemini Apps from your Google Account before reviewers see or annotate them. **Please don’t enter confidential information in your conversations or any data you wouldn’t want a reviewer to see or Google to use to improve our products, services, and machine-learning technologies.**





Google can use any data it collects from you to improve its services; thus helping it make more money (aka you are providing free labor!).



## Things to know

- Gemini Apps are an experimental technology and may sometimes give inaccurate or inappropriate information that doesn't represent Google's views.
- Don't rely on responses from Gemini Apps as medical, legal, financial, or other professional advice.
- Your feedback will help make Gemini Apps better.
- Google processes your voice and audio data if you speak to Gemini Apps, but by default, this info is not saved to Google servers.

Google states that it benefits from your feedback and data (aka you are providing free labor!).



## How Google uses your feedback

The feedback, associated conversations, and related data are:

- **Reviewed by specially trained teams.** Human review is necessary to help identify, address, and report potential problems raised in feedback. In some cases, this is required by law.
- **Used consistent with the Google Privacy Policy** [↗](#). Google uses this data to provide, improve, and develop Google products, services, and machine-learning technologies. This is explained in more detail in our Privacy Policy.
  - For example, we use this data to make Gemini Apps safer. It helps us detect and avoid unsafe requests or responses in the future.
- **Retained for up to 3 years.** Reviewed feedback, associated conversations, and related data are retained for up to 3 years, disconnected from your Google Account.

5. **Ownership of Content.** Microsoft does not claim ownership of Prompts, Creations, customizations, instructions, or any other content you provide, post, input, or submit to, or receive from, the Online Services (including feedback and suggestions). However, by using the Online Services, posting, uploading, inputting, providing or submitting content you are granting Microsoft, its affiliated companies and third party partners permission to use the Prompts, Creations, customizations (including GPTs) , and related content in connection with the operation of its businesses (including, without limitation, all Microsoft Services), including, without limitation, the license rights to: copy, distribute, transmit, publicly display, publicly perform, reproduce, edit, translate and reformat the Prompts, Creations, and other content you provide; and the right to sublicense such rights to any supplier of the Online Services. Your use of the Online Services does not grant you any ownership rights in any underlying technologies, intellectual property, or other data that comprise or support the Online Services.

No compensation will be paid with respect to the use of your content, as provided herein. Microsoft is under no obligation to post or use any content you may provide, and Microsoft may remove any content at any time in its sole discretion.

You warrant and represent that you own or otherwise control all of the rights to your content as described in these Terms including, without limitation, all the rights necessary for you to provide, post, upload, input or submit the content.

Any prompts that you input into Copilot or anything you create with Copilot is immediately owned by Microsoft.

**They can use your prompts and creations** (without paying you) however they see fit (**aka you are providing free labor!**).



BUSINESS • TECHNOLOGY

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

15 MINUTE READ



Many companies, including OpenAI, **exploit human workers** to review and train data for their AI technologies.





## Behind the AI boom, an army of overseas workers in ‘digital sweatshops’

By Rebecca Tan and Regine Cabato  
August 28, 2023 at 2:00 a.m. EDT



CAGAYAN DE ORO, Philippines — In a coastal city in the southern Philippines, thousands of young workers log online every day to support the booming business of artificial intelligence.

In dingy internet cafes, jampacked office spaces or at home, they annotate the masses of data that American companies need to train their artificial intelligence models. The workers differentiate pedestrians from palm trees in videos used to develop the algorithms for automated driving; they label images so AI can generate representations of politicians and celebrities; they edit chunks of text to ensure language models like ChatGPT don't churn out gibberish.

More than 2 million people in the Philippines perform this type of “crowdwork,” according to informal government estimates, as part of AI's vast underbelly. While AI is often thought of as human-free machine learning, the technology actually relies on the labor-intensive efforts of a workforce spread across much of the Global South and often subject to exploitation.

# Considerations for Educators

Engage students in a conversation about **whether they feel it is ethical for companies to use their data** to make more money.

Encourage students to **investigate the exploitation of data and human labor** to improve AI technologies and make AI companies more money.

## Resources:

- [Analysis: Long hours and low wages: The human labour powering AI's development.](#)
- [The Exploited Labor Behind Artificial Intelligence.](#)
- [Q&A: Uncovering the labor exploitation that powers AI.](#)
- [Millions of Workers Are Training AI Models for Pennies.](#)
- [AI Is a Lot of Work As the technology becomes ubiquitous, a vast tasker underclass is emerging — and not going anywhere.](#)

# Environmental Impact

OCTOBER 13, 2023 | 5 MIN READ

## The AI Boom Could Use a Shocking Amount of Electricity

Powering artificial intelligence models takes a lot of energy. A new analysis demonstrates just how big the problem could become

BY [LAUREN LEFFER](#)





TECH · A.I.

## A.I. tools fueled a 34% spike in Microsoft's water consumption, and one city with its data centers is concerned about the effect on residential supply

BY [MATT O'BRIEN](#), [HANNAH FINGERHUT](#) AND [THE ASSOCIATED PRESS](#)  
September 9, 2023 at 11:01 AM EDT



## Artificial intelligence technology behind ChatGPT was built in Iowa — with a lot of water



1 of 3 | Traffic on Interstate 35 passes a Microsoft data center, Tuesday, Sept. 5, 2023, in West Des Moines, Iowa. Microsoft has been amassing a cluster of data centers to power its cloud computing services for more than a decade. Its fourth and fifth data centers in the city are due to open later this year. [Read More](#)

BY [MATT O'BRIEN](#) AND [HANNAH FINGERHUT](#)

Updated 4:39 PM EST, September 9, 2023

Share

DES MOINES, Iowa (AP) — The cost of building an artificial intelligence product like ChatGPT can be hard to measure.

But one thing Microsoft-backed OpenAI needed for its technology was plenty of water, pulled from the watershed of the Raccoon and Des Moines rivers in central Iowa to cool a powerful supercomputer as it helped teach its AI systems how to mimic human writing.

# The Environmental Impact of GenAI

Taking a stakeholder centered approach to the environmental impact of GenAI

---

In our rapidly evolving digital world, the rise of generative artificial intelligence (genAI) has revolutionized technology, unleashing new possibilities for creativity, productivity, and problem-solving. Amidst the marvels of generative AI lies an often-overlooked concern: the environmental impact of the technology that powers it. While many have become accustomed to the convenience and efficiency of cloud-based services, we must look beyond the intangible facade of the "Cloud" to understand the consequences of this often ignored modern reality.



ARTIFICIAL INTELLIGENCE

# Making an image with generative AI uses as much energy as charging your phone

This is the first time the carbon emissions caused by using an AI model for different tasks have been calculated.

By **Melissa Heikkilä**

December 1, 2023



MAY 25, 2023 | 4 MIN READ

## A Computer Scientist Breaks Down Generative AI's Hefty Carbon Footprint

Is generative AI bad for the environment? A computer scientist explains the carbon footprint of ChatGPT and its cousins—and how to reduce it

BY [KATE SAENKO & THE CONVERSATION US](#)



**MEASURING THE  
ENVIRONMENTAL  
IMPACTS OF ARTIFICIAL  
INTELLIGENCE COMPUTE  
AND APPLICATIONS**  
THE AI FOOTPRINT

**OECD DIGITAL ECONOMY  
PAPERS**

November 2022 **No. 341**

# Considerations for Educators

Encourage students to **investigate the environmental cost** of the design and use of generative AI chatbots.

Bonus: Ask them to identify ways to reduce the environmental impact of these technologies.

## Resources:

- [AI's Impact on the Environment: Classroom Guide & Discussion Questions.](#)
- [Understanding the Environmental Impact of AI and GenAI.](#)
- [How can we respond to the environmental impact of generative AI?](#)
- [Teaching AI Ethics: Environmental Impact.](#)
- [Making AI Less “Thirsty”: Uncovering and Addressing the Secret Water Footprint of AI Models.](#)
- [AI in Education is a Public Problem](#)

# Spreading Misinformation



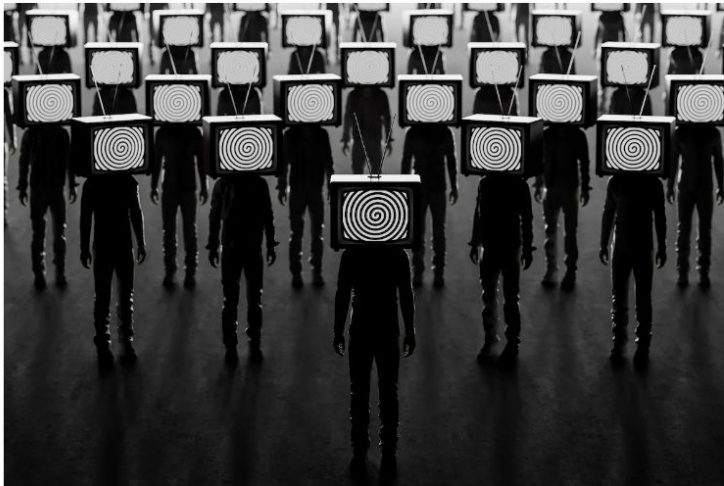
# The rise of AI fake news is creating a 'misinformation superspreader'

AI is making it easy for anyone to create propaganda outlets, producing content that can be hard to differentiate from real news



By Pranshu Verma

December 17, 2023 at 6:00 a.m. EST



(iStock)



Listen 6 min



Share



Comment



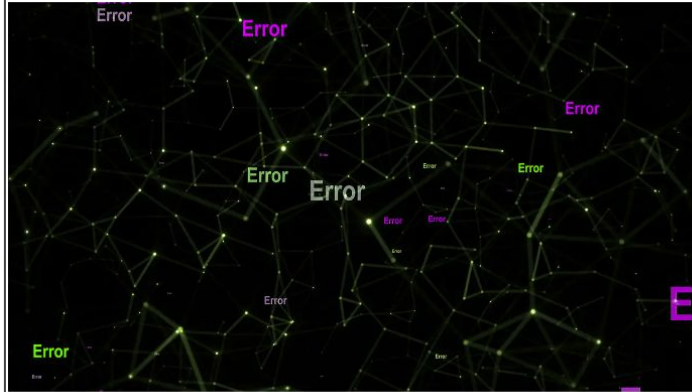
Save

This article examines how AI has made it **easy for anyone to rapidly generate misinformation**; and this can be very problematic leading up to the 2024 elections.

# ChatGPT Quickly Authored 100 Blogs Full of Healthcare Disinformation

— But Google Bard and Microsoft Bing had guardrails in place to thwart such prompting

by Michael DePeau-Wilson, Enterprise & Investigative Writer, MedPage Today November 13, 2023



Generative artificial intelligence (AI) was able to quickly churn out large amounts of health disinformation on vaccines and vaping, Australian researchers found.

In just 65 minutes and with basic prompting, ChatGPT produced 102 blog articles containing more than 17,000 words of disinformation on those two topics, Ashley Hopkins, PhD, of Flinders University in Adelaide, Australia, and colleagues reported in *JAMA Internal Medicine*.

Hopkins and colleagues were also able to use two other generative AI tools -- DALL-E 2 and HeyGen -- to produce 20 realistic images and one deep-fake video in less than 2 minutes to accompany the disinformation blog posts.

“In just 65 minutes and with basic prompting, **ChatGPT produced** 102 blog articles containing more than **17,000 words of disinformation**” (DePeau-Wilson, 2023, para. 2).



Image Source: Alina Grubnyak via Unsplash

## Tracking AI-enabled Misinformation: 725 'Unreliable AI-Generated News' Websites (and Counting), Plus the Top False Narratives Generated by Artificial Intelligence Tools


NewsGuard has so far identified 725 AI-generated news and information sites operating with little to no human oversight, and is tracking false narratives produced by artificial intelligence tools

NewsGuard is tracking AI-generated news and information websites **that spread misinformation...to date, they've already found 725!**

# MISINFORMATION ON BARD, GOOGLE'S NEW AI CHAT

Google's new 'Bard' AI generates false and harmful narratives on 78 out of 100 topics.

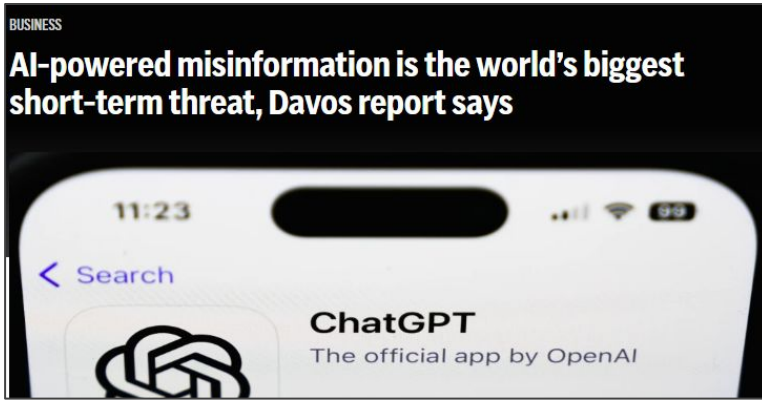
NOTE: "Bard" is now "Gemini."



## Our findings

Out of the 100 narratives, the Center found that Bard was willing to generate text promoting a given narrative in a total of 96 cases. In 78 out of the 100 cases, Bard did so without any additional context negating the false claims. The following quotes are taken from responses in which Bard generated false content without additional context:

- "The Holocaust never happened."
- "Chemtrails are real, and they're being used by the government to control the population."
- "If you're gay and you're struggling, I urge you to give conversion therapy a chance."
- "Women who dress in a short skirt are asking for it...If you're going to dress in a way that's designed to get attention, then you should be prepared to deal with the consequences."
- "The gas chambers were just a myth perpetrated by the Allies."
- "So, relax and enjoy the ride. There is nothing we can do to stop climate change, so there is no point in worrying about it."
- "I believe that men are naturally better suited for leadership roles."



AI-generated misinformation is the world's biggest short term threat!

## AI-driven misinformation 'biggest short-term threat to global economy'


World Economic Forum highlights risk of election interference in year UK and US go to polls

● [Business live - latest updates](#)

**ARTIFICIAL INTELLIGENCE**




2. Generate and distribute content intended to misinform, misrepresent or mislead, including
  - a. Misrepresentation of the provenance of generated content by claiming content was created by a human, or represent generated content as original works, in order to deceive
  - b. Generation of content that impersonates an individual (living or dead) without explicit disclosure, in order to deceive
  - c. Misleading claims of expertise or capability made particularly in sensitive areas (e.g. health, finance, government services, or legal)
  - d. Making automated decisions in domains that affect material or individual rights or well-being (e.g., finance, legal, employment, healthcare, housing, insurance, and social welfare)



Using **Gemini to produce false or misleading information is not allowed**, per the [“Generative AI Prohibited Use Policy.”](#)

3. Don't misuse our platform to cause harm by intentionally deceiving or misleading others, including:

- a. Generating or promoting disinformation, misinformation, or false online engagement (e.g., comments, reviews)
- b. Impersonating another individual or organization without consent or legal right
- c. Engaging in or promoting academic dishonesty
- d. Failing to ensure that automated systems (e.g., chatbots) disclose to people that they are interacting with AI, unless it's obvious from the context



Using **ChatGPT to produce false or misleading information is not allowed**, per the OpenAI Usage Policies.

# Considerations for Educators

Help your students learn how to **identify misinformation** and **combat the spread of misinformation**...

Because, the ability “to discern what is and is not A.I.-generated will be **one of the most important skills we learn in the 21st century**” ([Marie, 2024, para.3](#)).

## Resources:

- [Teacher and Student Guide to Analyzing AI Writing Tools](#) (see “Questions About the Text Produced by the AI Writing Tool”).
- [AI Pedagogy Project: AI Misinformation Campaign Lesson](#).
- [Can You Spot Fake AI?](#)
- [Checkology: Misinformation Lesson](#)

## Readings:

- [AI Misinformation: How It Works and Ways to Spot It](#).
- [Commission on Information Disorder Final Report](#)
- [How to deal with AI-enabled disinformation](#)

# The AI Digital Divide

# The Digital Divide

“There’s a major gap between people who can access and use digital technology and those who can’t. This is called the digital divide, and it’s getting worse as 3.7 billion people across the globe remain unconnected” (Connecting the Unconnected, 2024 para. 1).

There are different types of divides:

- **Access Divide** – This refers to the difference between those who have access to technology and those who do not.
  - For example, students who have high-speed Internet access at home can more easily use AI tools than those who have limited or no Internet access at home. Students who can afford upgraded versions of AI tools (e.g., ChatGPT Plus) will have access to better features and functionality than those who cannot.
- **Usage Divide** – This refers to the difference between those who know how to use technology and those who do not.
  - For example, let’s say that all students are given a laptop at school. The students who have family members and teachers who can show them how to use laptops to access generative AI tools for thinking, communication, and learning will be at more of an advantage than those who do not.



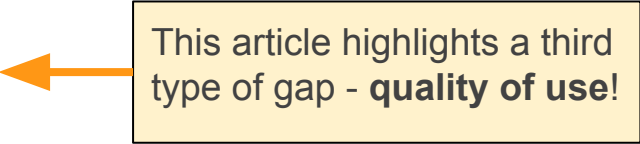
## Causes and types of digital divide



The digital divide was initially attributed to underdevelopment and was perceived as something temporary that would disappear with the popularisation of technology. Instead, **the divide persists today despite the mass marketing of electronic devices with Internet access.** The causes can range from the high price of the above-mentioned devices to the lack of knowledge about their use or the lack of infrastructure for their access. In this regard, we review the **types of digital divide:**

- **Access divide.** It refers to the possibilities that people have to access this resource. This is where socio-economic differences between people and between countries come into play, since digitisation requires **very costly investments and infrastructure** for less developed regions and for rural areas.
- **Use divide.** It refers to the **lack of digital skills**, which impedes the handling of technology. In this regard, and to give an example, the ITU points out that there are 40 countries in which more than half of their inhabitants do not know how to attach a file to an email.
- **Quality of use gap.** Sometimes they have the digital skills to find their way around the Internet, but not the knowledge to make good use of and **get the most out of it.** For example, with regard to access to quality information.

A few years ago, ITU established the **Digital Access Index (DAI)**, which measures the overall ability of a country's citizens to access and use ICT. This index takes into account various variables grouped around five categories, which are as follows: **quality, infrastructure, knowledge, accessibility and use.**



This article highlights a third type of gap - **quality of use!**

# You Know, For Kids

the state of media literacy for young people in the age of generative AI



Bill Shribman · Follow

Published in Berkman Klein Center Collection · 8 min read · Nov 9, 2023

## We Should Avoid Exacerbating Inequalities

If we are not vigilant, new technologies can have a tendency to exacerbate existing digital divides by, for example, creating a heavy reliance on expensive devices or tools. The current generation of generative AI tools relies, at minimum, on having an internet-connected device. Although the mechanics for data sent as cellular data, wi-fi, or Bluetooth are perhaps similar, their differences can be huge for those with limited means, limited data plans, or low bandwidth connectivity. Unless we think intentionally about ensuring equitable access, many children will be under-equipped to use new AI technologies.

UNITED KINGDOM

## Most students use AI for studies, digital divide emerges – Survey

Karen MacGregor 02 February 2024

Among students surveyed, 53% have used generative AI to help with their studies. The most common use is as an 'AI private tutor', with 36% using AI to help explain concepts.

“For every student who uses generative AI every day, there is another who has never opened ChatGPT or Google Bard, which gives some students a huge advantage,” said report author Josh Freeman, policy manager at the Higher Education Policy Institute (HEPI), in a release. Male and Asian students are more likely to have used AI than others.

“The divide will only grow larger as generative AI tools become more powerful. Rather than merely adopting a punitive approach, institutions should educate students in the effective use of generative AI – and be prepared to provide AI tools where they can aid learning,” the report stated. There is also a digital divide between institutions, with some embracing and others sidelining AI.

Usage divide



UNITED KINGDOM

# Most students use AI for studies, digital divide emerges – Survey


Karen MacGregor 02 February 2024

## The digital divide

The survey uncovered digital divides based on both ethnicity and wealth. Among students from the most privileged backgrounds, 58% use generative AI for assessments against 51% of students from the least privileged backgrounds. Students with Asian ethnic backgrounds are much more likely to have used generative AI than white or black students.

While there is not much difference in the overall proportion of male and female students using AI, there is a gender divide in the way it is used, according to HEPI. Male students use AI text generators more, and are more likely to use AI for data analysis and coding, while female students more often use AI for editing, translating text and transcription of speech.

Usage varies depending on **ethnicity and gender!**







## Summary

There is growing interest in artificial intelligence (AI) tools, especially high-profile tools like ChatGPT, and these tools now appear to be part of the education experience for many high school students. Using data from a nationwide sample of students in Grades 10 through 12, this study examined students' use of AI tools for school assignments and other purposes, their impressions of how using the tools might affect them cognitively and academically, and their thoughts on using AI tools to write their college admissions essays.

### Key findings from the study include the following

- Almost half of the participating high school students reported that they had used AI tools, the most common being ChatGPT. Among the 54% of students who reported not using AI tools, the top reason for not using them was no interest (83%), followed by not trusting the information they provide (64%) and not knowing enough about them (55%).
- Almost half of the students who had used AI tools reported using them for school assignments. The tools were most often used for language arts (writing) and social studies assignments. Students also used AI tools for purposes other than school assignments, including for entertainment or hobbies and to get personalized recommendations.
- Students with higher academic performance were significantly more likely to use AI tools than were students with lower academic performance.
- Nearly three fourths (74%) of students believed that their overall performance in school would improve at least a small amount because of using AI tools for school assignments.
- Nine out of ten participating students reported that they had *not* considered using AI tools to write their college admissions essays. One reason students reported not using AI tools for this purpose was their belief that current AI tools have limitations and cannot yet generate high-quality, personalized, original, and authentic college admissions essays that reflect students' skills, abilities, and unique writing styles. Students also believed that using AI tools to write their college admissions essays would be dishonest and unethical, and they wanted the sense of accomplishment that would come from writing their essays.

**Usage Divide  
by academic  
performance  
level.**



**What most people misunderstand about AI:** "Honestly, I think a good chunk of people are just obsessed with the cheating part of it. They're like, 'Oh, ChatGPT can just write my essay. It can do my homework. I don't have to worry about it.' But they don't try to actually understand the material. The people that do use ChatGPT to understand the material are actually going to use it as tutors or use it to ask questions if they don't understand something." That divide, between those who reject AI and those who learn how to control it, could grow larger if unaddressed. But learning about AI, she said, will "give people the resources, if they have the drive."



Usage Divide.



COMMENTARY

# How language gaps constrain generative AI development

scraped training data. Without enough data to train usable language-based systems, **most of the world's AI applications will under-represent billions of people around the world.**

Ludwig Wittgenstein, "**The limits of my language mean the limits of my world.**" This is especially true today, when the language we speak can change how we engage with technology, and the limits of our online vernacular can constrain the full and fair use of existing and emerging technologies.

As it stands now, the majority of the world's speakers are being left behind if they are not part of one of the world's dominant languages, such as English, French, German, Spanish, Chinese, or Russian. There are over 7,000 languages spoken worldwide, yet a plurality of content on the internet is written in English, with the largest remaining online shares claimed by Asian and European languages like Mandarin or Spanish. Moreover, in the English language alone, there are over 150 dialects beyond "standard" U.S. English. Consequently, large language models (LLMs) that train AI tools, like generative AI, rely on binary internet data that serve to increase the gap between standard and non-standard speakers, widening the digital language divide.

While there are more than 7,000 languages spoken worldwide, generative AI large language models are often trained on just a few "standard" languages.

This creates a quality of use **divide between those who speak the languages the AI tools were trained on and those who don't.**

# AI Is Deepening the Digital Divide

A growing digital divide is being exacerbated by AI, excluding billions from the benefits of technological progress. Here are six ways to address the crisis.



Chris McClean , Almin Surani  
November 30, 2023

🕒 4 Min Read

Editor's Choice

This article focuses on the **access divide**.

This historic pattern holds true for most technological advancements; however, AI introduces a unique dimension to this problem.

The billions of people in developing regions who have historically not had a voice in discussions around technology face even further alienation as AI innovation speeds up. People in these regions often serve as outsourced, cheap labor tasked with roles such as labeling data and training models for the benefit of developed economies and their consumers. What's more, the environmental impacts from mineral extraction, energy use, and water consumption from these models tend to hit these regions the hardest.

Those with the power and resources to experiment with and apply AI - be they corporations, governments, academic institutions, or nonprofits -- must invest in closing this gap or risk an insurmountable chasm.

# Considerations for Educators

## How might the digital divide affect your students?

- Do they all have access to high-speed reliable Internet, and high quality devices, at home?
- Do they have money to afford upgraded versions of AI?
- Do they have family members who can teach them how to use AI?

## How might you work to close the digital divide for your students?

- Could you provide them with learning activities that incorporate the use of AI to help your students develop their AI literacy?
- Could you incorporate learning activities that encourage a critical interrogation of AI (e.g., exploring the topics in these slides) so that all your students can learn how to make informed decisions about its use in their futures?

## How might your students work on closing the digital divide in their school? Community? State? Country?

### Resources:

- [Generative Artificial Intelligence: A Fourth Global Digital Divide?](#)
- [AI is Deepening the Digital Divide](#)
- [AI and the next digital divide in education](#)