



CareerLearning

How to Write a ChatGPT Policy for Your Higher Education Institution

Presented by: Dr. Jim Castagnera, Esq.

President, Dr. Jim's One-Stop HR Shop

| Partner, Portum Group International

Agenda

I

Toward a ChatGPT “Bill of Rights”

II

Current Legal Issues

III

Some Core Tactics for Coping with ChatGPT

IV

Academic Integrity

V

Research Integrity

VI

A Policy for Administrators and Staff

Bill of rights

**Toward a ChatGPT Bill
of Rights**

ChatGPT

“ChatGPT is an artificial-intelligence chatbot developed by OpenAI and launched in November 2022. It is built on top of OpenAI's GPT-3.5 and GPT-4 families of large language models and has been fine-tuned using both supervised and reinforcement learning techniques.”

<https://en.wikipedia.org/wiki/ChatGPT>

““ChatGPT is scary good. We are not far from dangerously strong AI.”

- Elon Musk, Vox, December 15, 2022



Desirable Practices

For managing A.I.

1. Understanding the reasons for a specific AI decision
2. Understanding the patterns between input values and AI decisions
3. Developing trustworthy algorithms
4. Determining what data is used to train AI



Blueprint for an AI Bill of Rights

“The Blueprint for an AI Bill of Rights is a set of **five principles** and associated practices to help guide the design, use, and deployment of automated systems to protect the rights of the American public in the age of artificial intelligence. Developed through extensive consultation with the American public, these principles are a blueprint for building and deploying automated systems that are aligned with democratic values and protect civil rights, civil liberties, and privacy.”

<https://www.whitehouse.gov/ostp/ai-bill-of-rights/what-is-the-blueprint-for-an-ai-bill-of-rights/>

1. Safe and Effective Systems

“Automated systems should be developed with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risks, and potential impacts of the system.

“Systems should undergo pre-deployment testing, risk identification and mitigation, and ongoing monitoring that demonstrate they are safe and effective based on their intended use, mitigation of unsafe outcomes including those beyond the intended use, and adherence to domain-specific standards”

2. Algorithmic Discrimination Protections

- “Designers, developers, and deployers of automated systems should take proactive and continuous measures to protect individuals and communities from algorithmic discrimination and to use and design systems in an equitable way.
- “This protection should include proactive equity assessments as part of the system design, use of representative data and protection against proxies for demographic features, ensuring accessibility for people with disabilities in design and development, pre-deployment and ongoing disparity testing and mitigation, and clear organizational oversight.”

3. Data Privacy

- “Designers, developers, and deployers of automated systems should seek your permission and respect your decisions regarding collection, use, access, transfer, and deletion of your data in appropriate ways and to the greatest extent possible; where not possible, alternative privacy by design safeguards should be used.
- “Systems should not employ user experience and design decisions that obfuscate user choice or burden users with defaults that are privacy invasive.”

4. Notice and Explanation

“Designers, developers, and deployers of automated systems should provide generally accessible plain language documentation including clear descriptions of the overall system functioning and the role automation plays, notice that such systems are in use, the individual or organization responsible for the system, and explanations of outcomes that are clear, timely, and accessible.”



5. Human Alternatives, Considerations, and Fallback

“You should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems you encounter.”



LEGAL ISSUES

Current Legal Issues



Disparate Impact

“Disparate impact occurs when policies, practices, rules or other systems that appear to be neutral result in a disproportionate impact on a protected group. For example, testing all applicants and using results from that test that will unintentionally eliminate certain minority applicants disproportionately is disparate impact.”

Source: SHRM, <https://www.shrm.org/resourcesandtools/tools-and-samples/hr-qa/pages/disparateimpactdisparatetreatment.aspx>

Algorithm

“An algorithm is a procedure used for solving a problem or performing a computation. Algorithms act as an exact list of instructions that conduct specified actions step by step in either hardware- or software-based routines. Algorithms are widely used throughout all areas of IT.”

Source: <https://www.techtarget.com/whatis/definition/algorithm>

Algorithms in the Hiring Process

“Hiring is rarely a single decision, but rather the culmination of a series of smaller, sequential decisions. Algorithms play different roles throughout this process:

- “Some steer job ads toward certain candidates,
- while others flag passive candidates for recruitment.
- Predictive tools parse and score resumes, and help hiring managers assess candidate competencies in new ways, using both traditional and novel data.”
- Source: Harvard Business Review, May 6, 2019, <https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias>

Word2Vec bias

- “In 2016, Microsoft and Boston University researchers revealed that the Word2Vec (publicly available algorithmic model built on millions of words scraped from online Google News articles, which computer scientists commonly use to analyze word associations) model trained itself on gender stereotypes existing in online news sources (Bolukbasi, 2016).”
- Source: <https://medium.com/@MerveHickok/why-was-your-job-application-rejected-bias-in-recruitment-algorithms-part-2-f706fa9af651>, citing <https://www.semanticscholar.org/paper/Man-is-to-Computer-Programmer-as-Woman-is-to-Word-Bolukbasi-Chang/ccf6a69a7f33bcf052aa7def176d3b9de495beb7>

EEOC: Artificial Intelligence and Algorithm Fairness Initiative

“In 2021, U.S. Equal Employment Opportunity Commission (EEOC) Chair Charlotte A. Burrows launched an agency-wide initiative to ensure that the use of software, including artificial intelligence (AI), machine learning, and other emerging technologies used in hiring and other employment decisions comply with the federal civil rights laws that the EEOC enforces.”

<https://www.eeoc.gov/ai>

Elements of the Initiative

- Issue technical assistance to provide guidance on algorithmic fairness and the use of AI in employment decisions;
- Identify promising practices;
- Hold listening sessions with key stakeholders about algorithmic tools and their employment ramifications; and
- Gather information about the adoption, design, and impact of hiring and other employment-related technologies.

The ADA and Application Algorithms

“Employers now have a wide variety of computer-based tools available to assist them in hiring workers, monitoring worker performance, determining pay or promotions, and establishing the terms and conditions of employment. Employers may utilize these tools in an attempt to save time and effort, increase objectivity, or decrease bias. However, the use of these tools may disadvantage job applicants and employees with disabilities. When this occurs, employers may risk violating federal Equal Employment Opportunity (“EEO”) laws that protect individuals with disabilities.”

<https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence>



How Job Applications Can Violate the ADA (1)

The employer does not provide a reasonable accommodation that is necessary for a job applicant or employee to be rated fairly and accurately by the algorithm.

How Job Applications Can Violate the ADA (2)

- The employer relies on an algorithmic decision-making tool that intentionally or unintentionally screens out an individual with a disability, even though that individual is able to do the job with a reasonable accommodation.
- Screen out occurs when a disability prevents a job applicant or employee from meeting—or lowers their performance on—a selection criterion, and the applicant or employee loses a job opportunity as a result.
- A disability could have this effect by, for example, reducing the accuracy of the assessment, creating special circumstances that have not been taken into account, or preventing the individual from participating in the assessment altogether.



How Job Applications Can Violate the ADA (3)

The employer adopts an algorithmic decision-making tool for use with its job applicants or employees that violates the ADA's restrictions on disability-related inquiries and medical examinations.

One Way Job Applications Can Violate Title VII

- “A recent Stanford University **study**, however, found that speech-to-text AI developed by Amazon, IBM, Google, Microsoft, and Apple misidentified the words of African Americans at double the rate of whites.
- “Furthermore, a full 20% of transcriptions from black speakers were misinterpreted so drastically as to render them unusable— compared to only 2% for white speakers.”

Source: <https://www.venable.com/insights/publications/2021/11/eecoc-update-artificial-intelligence-actual-discrim>, citing <https://news.stanford.edu/2020/03/23/automated-speech-recognition-less-accurate-blacks/>

EEOC Settlement, March 20, 2023

“DHI Group, Inc., which operates job-search website for technology professionals Dice.com, has entered into a conciliation agreement with the U.S. Equal Employment Opportunity Commission (EEOC) to resolve national origin discrimination charges, the federal agency announced today.”

Source: EEOC Press Release, “DHI Group, Inc. Conciliates EEOC National Origin Discrimination Finding,” <https://www.eeoc.gov/newsroom/dhi-group-inc-conciliates-eeoc-national-origin-discrimination-finding>

The Charges

“This resolution resolves multiple charges of discrimination filed against DHI, ending an investigation in which the EEOC determined it had reasonable cause to believe DHI violated Title VII when some of its customers posted positions on Dice.com that excluded those of American national origin, thereby deterring a class of workers from applying.”

The Resolution

- Pursuant to the conciliation agreement, DHI will compensate the estate of the original complainant
- and agreed to rewrite its programming to “scrape” for potentially discriminatory keywords such as “OPT”, “H1B” or “Visa” that appear near the words “only” or “must” in its customer’s new job postings.
- DHI also agreed to revise its guidance to customers on its “Job Postings not permitted on Dice” website pop-up to include instructions to avoid language such as “H-1Bs Only” or “H-1Bs and OPT Preferred.”

New York City's “Bias Audit” Law

- Initially to be effective January 1, 2023
- Enforcement postponed until April 15, 2023
- Restricts employers in New York City from using AI machine-learning products to screen candidates during hiring and promotions
- Unless such products have been subjected to a “bias audit” no more than one year beforehand
- New York City's Department of Consumer and Worker Protection held two hearings to receive public comment on its proposed rules.
- NYC Local Law 144,
<https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=4344524&GUID=B051915D-A9AC-451E-81F8-6596032FA3F9&Options=Advanced&Search>

Take-Aways

- Audit your application processes—both employee and student applications—for possible hidden biases.
- See <https://www.youtube.com/watch?app=desktop&v=wOt6n3FmBKE> (40-minute video on how to audit an AI system for disparate impact)
- Appeal processes <https://www.usnews.com/education/best-colleges/articles/a-guide-to-the-college-admissions-appeal-process>

Intellectual Property

Who owns the content generated by ChatGPT?

“I do not own the content that I generate. I am a machine learning model developed and owned by OpenAI, and the content generated by me is subject to OpenAI's license and terms of use”.

Source: ChatGPT



The European Union's Position

- “As ChatGPT itself states, while the content generated may be protected by copyright, it will not be owned by the AI itself. Indeed, under European (and US) law AI cannot own copyright, as it cannot be recognised as an author and does not have the legal personality which is a pre-requisite for owning (intangible) assets.”
- “For now, it seems that artists/creators who use AI to support their creative process may be able to claim ownership of the work if it reflects their choices and creativity.”
- https://intellectual-property-helpdesk.ec.europa.eu/news-events/news/intellectual-property-chatgpt-2023-02-20_en

U.S. Copyright Office's Position

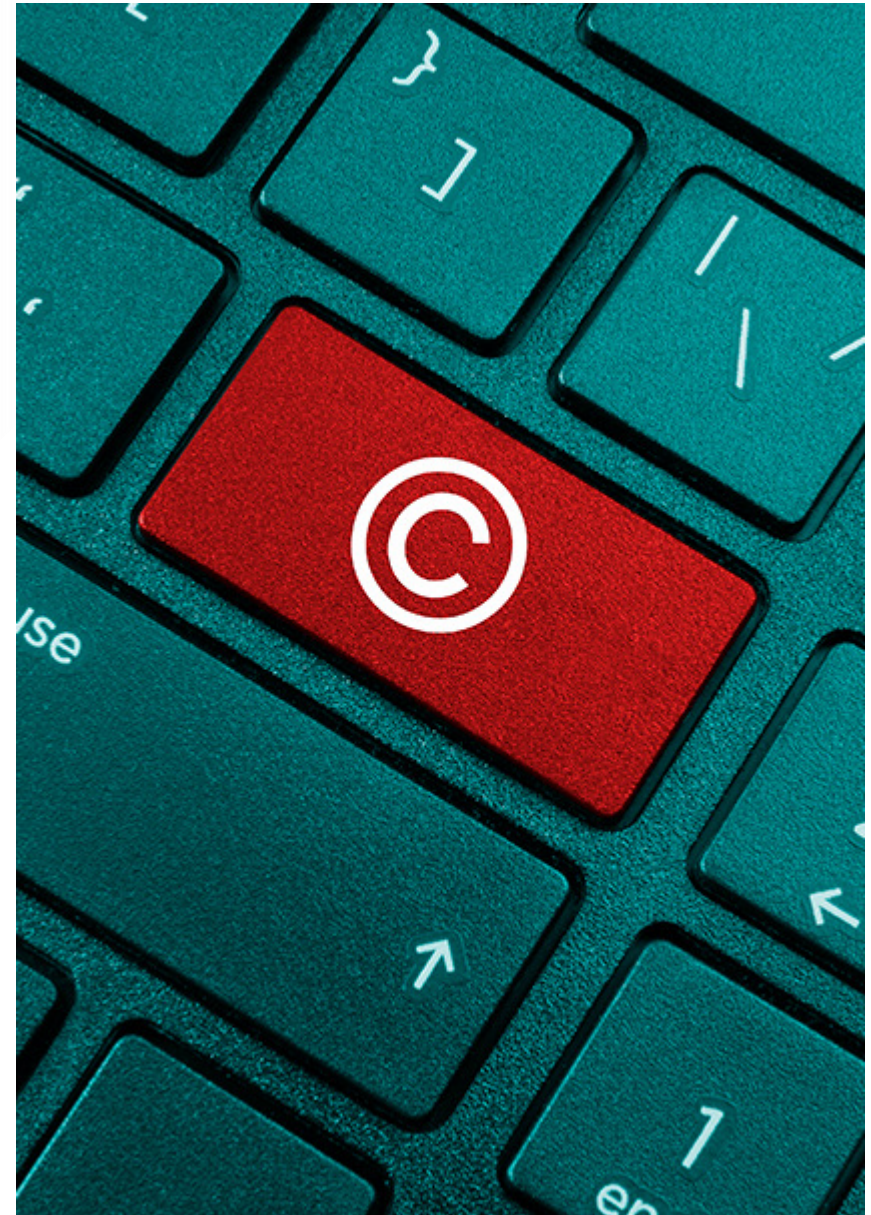
Zara of the Dawn

- “The Office has completed its review of the Work’s original registration application and deposit copy, as well as the relevant correspondence in the administrative record. We conclude that Ms. Kashtanova is the author of the Work’s text as well as the selection, coordination, and arrangement of the Work’s written and visual elements. That authorship is protected by copyright. However, as discussed below, the images in the Work that were generated by the Midjourney technology are not the product of human authorship. Because the current registration for the Work does not disclaim its Midjourney-generated content, we intend to cancel the original certificate issued to Ms. Kashtanova and issue a new one covering only the expressive material that she created.”
- U.S. Copyright Office, Feb. 21, 2023,
<https://fingfx.thomsonreuters.com/gfx/legaldocs/klpygnkyrpg/AI%20COPYRIGHT%20decision.pdf>

Lawsuits: Copyright Infringement

“Stock photo provider Getty Images has sued artificial intelligence company Stability AI Inc, accusing it in a lawsuit made public on Monday of misusing more than 12 million Getty photos to train its Stable Diffusion AI image-generation system.”

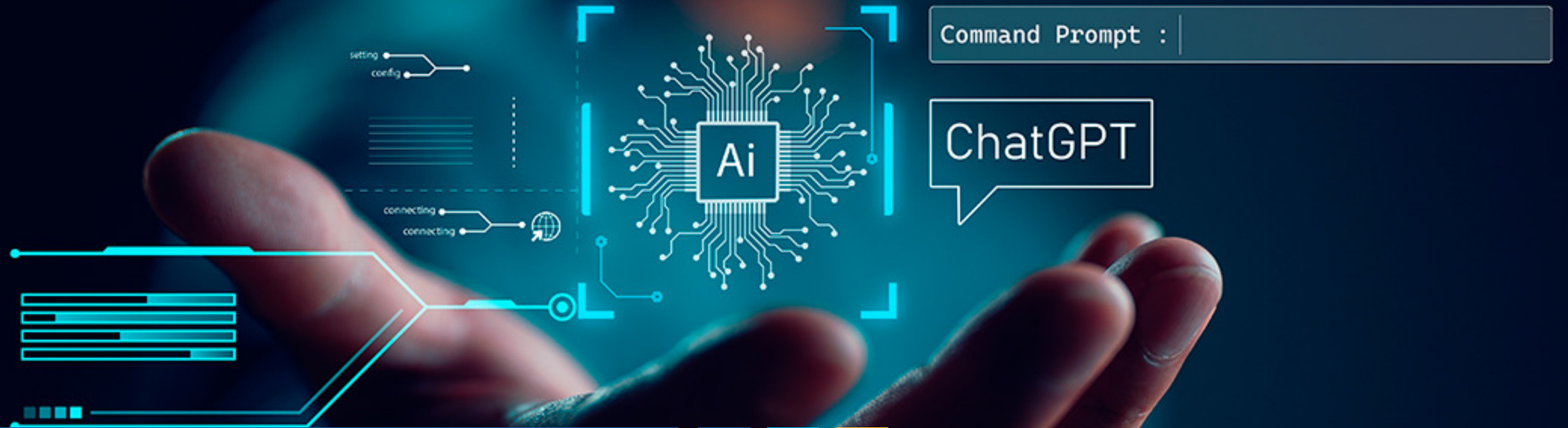
Source: Reuters, February 6, 2023,
<https://www.reuters.com/legal/getty-images-lawsuit-says-stability-ai-misused-photos-train-ai-2023-02-06/>



Take-Aways

1. Make proper disclosures if you are publishing content derived from ChatGPT. A sample disclosure that would comply with ChatGPT's policies is provided above.
2. Remember that your inputs into ChatGPT are not confidential and can be used by them to create new content for other users.
3. You will not have the ability to get copyright protection for content you derive from ChatGPT except to the extent you have made your own original changes to that content, and then only with respect to your original aspects. Only humans can get copyright protection for their work.
4. There is risk that content generated from ChatGPT infringes copyrights of others to the extent the content generated copied the works of others that was subject to copyright. If this is the case, you could be personally liable for copyright infringement if you use this content. This is still an open issue that will likely be weighed in on by the Supreme Court one day.
5. ChatGPT is an important tool– but remember it does have inherent risks and should just be part of your overall content strategy.

<https://thedietitianeditor.com/chatgpt-copyright/>



Some Core Tactics for Coping with ChatGPT



Human in the Loop

“A human is assisted by a machine. In this model, the human is doing the decision making and the machine is providing only decision support or partial automation of some decisions, or parts of decisions. This is often referred to as intelligence amplification (IA).”

Source: Harvard Business Review, November 10, 2021,
<https://hbr.org/2021/11/managing-ai-decision-making-tools>

Human on the Loop

“Here, the machine is assisted by a human. The machine makes the micro-decisions, but the human reviews the decision outcomes and can adjust rules and parameters for future decisions. In a more advanced set-up, the machine also recommends parameters or rule changes that are then approved by a human.”

–Ibid.

What Higher Ed Professionals Say (1)

- “First, familiarize yourself with AI writing aids. Your librarians and academic technicians can help.
- “Faculty members need time to play with new tools and explore their implications.
- “Administrators can carve out time for faculty training support.”

Source: Inside Higher Ed, January 11, 2023,

<https://www.insidehighered.com/news/2023/01/12/academic-experts-offer-advice-chatgpt#>

What Higher Ed Professionals Say (2)

- “Next, consider the tools relative to your course.
- “What are the cognitive tasks students need to perform without AI assistance?
- “When should students rely on AI assistance? Where can an AI aid facilitate a better outcome?
- “Are there efficiencies in grading that can be gained?
- “Are new rubrics and assignment descriptions needed?
- “Will you add an AI writing code of conduct to your syllabus?
- “Do these changes require structural shifts in timetabling, class size or number of teaching assistants?”

Ibid.

What Higher Ed Professionals Say (3)

- “[T]alk with students about instructions, rules and expectations.
- “Provide this information on course websites and syllabi and repeat it in class.
- “Guide teaching assistants in understanding appropriate uses of AI assistance for course assignments.
- “Divisions or departments might agree on expectations across courses.”
- *Ibid.*

What Higher Ed Professionals Say (4)

- “In the past, near-term prohibitions on slide rules, calculators, word processors, spellcheck, grammar check, internet search engines and digital texts have fared poorly.
- “They focus on in-course tactics rather than on the shifting contexts of what students need to know and how they need to learn it.
- “Reframing questions about AI writers will drive assignment designs and assessments that can minimize academic integrity concerns while promoting learning outcomes.”
- *Ibid.*

What Higher Ed Professionals Say (5)

- “Instead of treating ChatGPT as the horizon, look farther out.
- “Our approach to teaching should be guided not by one recent product but by reflection on the lives our students are likely to lead in the 2030s.
- “What will the writing process look like for them?”
- “Will they use models as research assistants? As editors?”
- *Ibid.*

What Higher Ed Professionals Say (6)

- “Teachers can oversee the selection of appropriate language models or AI-based tools to ensure they meet student needs and school policies.
- “Students could be responsible for using the model to generate language that is accurate, appropriate for the audience and purpose, and reflective of their own voices, while monitoring and reporting issues they encounter.
- “Lastly, the college’s administration can be responsible for providing feedback to the developer and updating the school’s policies regarding the use of AI writing assistants.”

Ibid.

What Higher Ed Professionals Say (7)

- “[M]odel generating and critiquing outputs and then have students try on their own. (In order to teach about ChatGPT’s failings, we’ll need to test any examples and exercises right before teaching, since language models are frequently updated.)
- “Finally, we should assess how well students can identify ChatGPT failings in terms of logic, consistency, accuracy and bias. Can they detect fabrications, misrepresentations, fallacies and perpetuation of harmful stereotypes? If students aren’t ready to critique ChatGPT’s output, then we shouldn’t choose it as a learning aid.”
- *Ibid.*



What Higher Ed Professionals Say (8)

“Channel anxiety over ChatGPT into productive experimentation. We built a local team of writing faculty to engage with the tools and to explore pedagogical possibilities.”



Academic Integrity

IV

What ChatGPT Thinks of Academic Integrity

- “For example, when preparing this page, the author prompted ChatGPT as follows: ‘In three sentences, please tell me why academic integrity is important.’”
- “ChatGPT then responded, ‘Academic integrity is important because it upholds the principles of honesty, fairness, and trust in the academic community. Maintaining academic integrity ensures that students are evaluated based on their own merit and not on dishonest or unethical behavior. Additionally, it helps to preserve the integrity and reputation of academic institutions and the degrees they confer’.”
- Ben Trachtenberg, Director, Office of Academic Integrity, U. of Missouri, March 2023, <https://oai.missouri.edu/chatgpt-artificial-intelligence-and-academic-integrity/>

Toward a Revised Academic Integrity Policy

- **“May Students Use ChatGPT and Similar Tools for their Academic Work?”**
- “The answer is, ‘It depends.’ If your professor allows you to use ChatGPT, and you use it as permitted, then you are not committing academic dishonesty. This is similar to use of a graphing calculator, which is acceptable as long as the instructor permits it.
- “Just like a graphing calculator, however, using ChatGPT on assignments is prohibited if the instructor does not allow its use. Students who use ChatGPT and similar tools without permission, or who use them in improper ways, are violating the academic integrity rules of the University.”
- *Ibid.*

Sample Policy: University of British Columbia

- The use of ChatGPT or other generative AI tools does not automatically equate to academic misconduct at UBC. At this time, the use of artificial intelligence tools is a course-level decision and there is no overall ban on its use in teaching and learning.
- If using ChatGPT and/or generative AI tools on coursework has been **prohibited** by the instructor, then using these tools would be considered to be academic misconduct.
- If using ChatGPT and/or generative AI tools has been **permitted** by the instructor, then instructors should make sure to convey the limitations of use and how it should be acknowledged and use should stay within those bounds.
- If the use of ChatGPT and/or generative AI tools has **not been discussed or specified** by the instructor, then it is likely to be considered as prohibited as an example of the “use or facilitation of unauthorized means to complete an examination or coursework” and more specifically as “accessing websites or other online resources not specifically permitted by the instructor or examiner”
- <https://academicintegrity.ubc.ca/chatgpt-faq/>

Sample Policy: Arizona State University

- We encourage academic units and/or faculty to determine whether student use of generative AI/ChatGPT in their courses is permitted or prohibited, and to state this and any parameters in your syllabi, announcements, and assignment instructions.
- At the very basic level, we recommend faculty ensure their syllabi include the ASU Academic Integrity policy, including copyright and plagiarism verbiage; the use of Generative AI/ChatGPT falls within these policies and processes.
- Further, we encourage faculty to state in syllabi, assignment instructions, etc., whether the use of AI/ChatGPT is permitted or prohibited (and to check with their unit leadership if needed).
- Faculty should emphasize that students must cite any borrowed content sources to comply with all applicable citation guidelines and copyright law and avoid plagiarism. Simply put, if students use generative AI, they should cite it: [How to Cite ChatGPT in APA format](#).
- Students and faculty should also ensure any AI-generated citations are correct, as ChatGPT is notorious for listing nonsensical citations.
- <https://instruction.thecollege.asu.edu/academicintegrityAIChatGPT>

Detecting ChatGPT Plagiarism

- “There are several AI detectors currently in existence, such as ***GPTZero*** and **OpenAI’s AI Text Classifier**. Despite the proliferation of such tools, it is important to remember that they might not be fully tested and that the technology to potentially outwit them continues to evolve. The detectors are not foolproof and can produce false negatives and false positives. It may also be possible for the user to modify content to avoid detection. For those reasons, it is recommended that AI detection tools not be used as the sole factor in decision-making around an allegation of academic misconduct. Privacy considerations might also exist around entering student work into the third-party site without their consent. To avoid privacy issues, instructors should not use these tools to evaluate any student work that contains the name of the student or any other personal information of the student or third parties. While these types of detectors might be experimented with in certain situations, they should not form the basis of a decision.”
- <https://academicintegrity.ubc.ca/chatgpt-faq/>

GPTZero

Quickly and efficiently detect whether an essay is ChatGPT or human written

- **About GPTZero**
- GPTZero is an app that detects essays written by the impressive AI-powered language model known as ChatGPT.
- Edward Tian, a computer science major who is minoring in journalism, spent part of his winter break creating GPTZero, which he said can "quickly and efficiently" decipher whether a human or ChatGPT authored an essay.
- His motivation to create the bot was to fight what he sees as an increase in AI plagiarism. Since the release of ChatGPT in late November, there have been reports of students using the breakthrough language model to pass off AI-written assignments as their own.
- <https://gpt3demo.com/apps/gptzero>

How to Write a ChatGPT Policy × Global Business Concept Manag... × ChatGPT Q&A - Academic Integ... × New AI classifier for indicating / × +


← → ↻ <https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text> ☆

Biz Dev Banking Website & Photos Social Media Webinar Providers Drexel Google Westlaw Sign In | Th... MyChart - Home Video Conferencing... MyRoyalty Kindle Vella >> Other Bookmarks

OpenAI Research ▾ Product ▾ Developers ▾ Safety Company ▾ Search Log in ↗ Sign up ↗

New AI classifier for indicating AI-written text

We're launching a classifier trained to distinguish between AI-written and human-written text.



Open AI's AI Text Classifier

- “We’ve trained a classifier to distinguish between text written by a human and text written by AIs from a variety of providers. While it is impossible to reliably detect all AI-written text, we believe good classifiers can inform mitigations for false claims that AI-generated text was written by a human: for example, running [automated misinformation campaigns](#), using AI tools for academic dishonesty, and positioning an AI chatbot as a human.”
- <https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text>

However...

- **“Our classifier is not fully reliable.** In our evaluations on a “challenge set” of English texts, our classifier correctly identifies 26% of AI-written text (true positives) as “likely AI-written,” while incorrectly labeling human-written text as AI-written 9% of the time (false positives). Our classifier’s reliability typically improves as the length of the input text increases. Compared to our [previously released classifier](#), this new classifier is significantly more reliable on text from more recent AI systems.”
- *Ibid.*

Turnitin

- “After months of **sounding the alarm** about students using AI apps that can churn out essays and assignments, teachers are getting AI technology of their own. On April 4, Turnitin is activating the software I tested for some 10,700 secondary and higher-educational institutions, assigning ‘generated by AI’ scores and sentence-by-sentence analysis to student work. It joins a handful of other free detectors already online.”

Source: Washington Post, April 3, 2023,

<https://www.washingtonpost.com/technology/2023/04/01/chatgpt-cheating-detection-turnitin/>



Research Integrity



What ChatGPT Potentially Can Do

1. Define the research question or hypothesis.
2. Conduct a literature review.
3. Gather and analyze data.
4. Organize the paper.
5. Draft the paper.
6. Proofread and edit the paper.

<https://www.ilovephd.com/chatgpt-write-scientific-research-paper/>

A Success Story?

- “Not all ChatGPT-related matters have elicited concern within the scientific research field. A February 2023 article in Nature described computational biologists’ use of ChatGPT to improve completed research papers. In just five minutes, the biologists received a review of their manuscript that increased readability and spotted equation-based mistakes. During a trial with three manuscripts, the team's use of ChatGPT was not always smooth, but the final output returned better-edited manuscripts.”
- Wen & Wang, “The future of ChatGPT in academic research and publishing: A commentary for *clinical and translational medicine*”, Mar. 20, 2023, NIH National Library of Medicine, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10028016/>

Pros and Cons of ChatGPT in Scientific Publishing

- “As it stands, ChatGPT cannot be relied upon to provide correct facts or produce reliable references, as stated by a January editorial in *Nature Machine Intelligence*. Accepting the limitations of ChatGPT and using it only for certain tasks allows researchers to delegate tedious jobs, such as manuscript editing, to the AI model while avoiding catastrophes such as the publication of false information.”
- *Ibid.*

Sample Policy: Proceedings of the National Academy of Sciences

- “**Authorship** must be limited to those who have contributed substantially to the work. The corresponding author must have obtained permission from all authors for the submission of each version of the paper and for any change in authorship. Use of artificial intelligence (AI) software, such as ChatGPT, must be noted in the Materials and Methods section (or Acknowledgments, if no Materials and Methods section is available) of the manuscript and may not be listed as an author. Authors are solely accountable for, and must thoroughly fact-check, outputs created with the help of generative AI software.”
- <https://www.pnas.org/author-center/editorial-and-journal-policies#authorship-and-contributions>



A Policy for Administrators and Staff

VI

Sample Policy

- Simply copy and paste the below content into your preferred text processor. Our only ask is that you acknowledge and respect this template's license by adding the following line in a clearly visible part of your document:
"Based on Template ChatGPT Use Policy by Kairoi Ltd (2023) / kairoi.uk / CC-BY 4.0"
- <https://github.com/KairoiAI/Resources/blob/main/Template-ChatGPT-policy.md>



Introduction

- Artificial Intelligence (AI) tools are becoming increasingly useful and accessible. In particular, generative AI tools have become a part of many professionals' workflows. Generative AI tools include those that produce text or images in response to queries. They are an accessible way to interact with large language models (LLMs), AI systems trained on enormous amounts of data. In that training process, LLMs require huge compute resources. In practice, text-to-image tools can be used by publishers to create visuals, and artists to aid them through the creative process. Similarly, chatbots based on LLMs can help write blogposts, emails and reports. ChatGPT is one such LLM-based chatbot.
- ChatGPT serves as an interface to interact with an LLM that is trained on enormous amounts of data. When inputting a prompt in ChatGPT, we are basically interacting with all the data it was trained on. Through different machine learning techniques, ChatGPT is also trained to predict the next word in a sentence, which allows its responses to sound natural and even compelling. However, ChatGPT can provide factually incorrect information, is easily manipulated, and is further trained by the prompts users provide.

Our University's Position on ChatGPT (1)

- Our University approves of the use of ChatGPT... But only under carefully controlled conditions.
- First, know that ChatGPT has been approved conditionally for the following activities: (List Activities)
- Second, (Name of Individual and Position) must preapprove all uses.
- Third, an approved use of ChatGPT must conform to the Guidelines and Approved Uses listed below and must not fall into any of the Forbidden Uses.

Our University's Position on ChatGPT (2)

- At [**Uni/college name**], we do not encourage the usage of ChatGPT, but we acknowledge that staff may value integrating the tool into their workflows. [**Org name**] will not provide you with login credentials but, if you choose to use ChatGPT, you may do so by signing up with a personal email and mobile number at <https://chat.openai.com/>, notwithstanding the present policy. It is expressly forbidden for you to create an account to use ChatGPT with [**uni/college name**] email addresses or phone numbers.
- The present policy sets out general guidelines, as well as forbidden and acceptable uses of ChatGPT, to ensure the safety of staff, the company, and our partners, clients and suppliers. Whilst [**uni/college name**] cannot monitor your use of ChatGPT on personal devices, we do have a [**Computer Use Policy**] whereby your use of devices of [**org name**]'s property are subject to relevant monitoring procedures. The present *ChatGPT Use Policy* complements and in no way supersedes [**uni/college name**]'s [**Computer Use Policy**], and applies to all staff employed by or seconded to [**uni/college name**].

General Guidelines

- ChatGPT is a tool, similar to using Google to browse the internet, or Word to write documents.
- When you type into ChatGPT, you are giving information to the system's developers at OpenAI.
- Always ask yourself whether ChatGPT is the best tool for the task at hand (ask your colleagues or supervisor if you're not sure).
- Each time you start a new chat on ChatGPT, you are starting “from scratch” in the sense that it cannot reference other chats you had with it.
- ChatGPT can provide inaccurate information, and all its responses must be carefully checked for errors.
- ChatGPT is best used to feed back on one's own work, and not to scrutinise others' efforts.
- When seeking feedback on text or code, it's safest to provide only snippets to ensure you are minimising the amount of information divulged.
- If you don't want something to be in the public domain, then don't type it into ChatGPT!

Acceptable Uses

- Ask for a spell-check to ensure proper grammar (for example: "Please check the grammar of this sentence: "these processes must be seperated"").
- Ask for better ways to word a sentence (for example: "What's a better way of saying "our service is faster and more efficient"?").
- Ask for the outline of reports if it is the first report of the sort (for example: "What's a useful outline of a recruitment report?").
- Ask for the outline of a blogpost for inspiration (for example: I need to write a blogpost introducing privacy-enhancing technologies – how could it be structured?").
- Ask for feedback on snippets of code or help with Excel formulas (for example: "I am using Excel. Column A has the price "\$10" and column B has the number of avocados offered at this rate ("2"). How do I get the price per avocado in column C?").

Forbidden Uses

- You may not sign up to ChatGPT using company credentials (work email and phone number).
- You may not divulge sensitive information, such as passwords, addresses, bank details or phone numbers.
- You may not type out or paste draft contracts, such as employment contracts or contracts with clients or providers.
- You may not type out or paste documents that may contain business-sensitive data, such as draft annual reports, business cases and accounting reports.
- You may not seek feedback on job applicants' CVs by pasting them in in part or in full.
- You may not type out or paste proprietary code, documents or other information that is not intended for the public eye.
- You may not type out or paste content that is subject to copyright licenses that do not allow for their indiscriminate sharing.
- Failure to comply with this section of the present policy will result in disciplinary action and can lead to termination of employment.

Best Practices

- Make the creation of the policy a cabinet-level priority.
- Carefully define the relevant terms and concepts.
- **Note:** ChatGPT now has at least nine serious competitors, including Google's Bard and Microsoft's BingAI. <https://gigster.com/blog/5-chatgpt-competitors-emerging-in-2023/>
- Include faculty and staff representatives in the development and testing of the policy.
- Be sure to include processes for (1) monitoring; (2) recording; (3) reporting problems and abuses; (4) reacting to issues.

Training Program: Using a ChatGPT Employee Policy

- Introduction to ChatGPT and ethical use
- Overview of ChatGPT and its capabilities
- Importance of ethical and responsible use of AI technology
- Alignment with company values and policies
- Source: ChatGPT, August 10, 2023

2. Understanding Data Privacy and Confidentiality

- Explanation of data privacy concerns and regulations
- How ChatGPT processes and stores user data
- Guidelines for handling sensitive or confidential information

3. Ensuring Fairness and Avoiding Bias

- Recognizing and addressing biases in AI-generated content
- Importance of diversity and inclusivity in conversations
- Techniques to prevent inadvertent reinforcement of biases

4. Navigating Sensitive Topics and Legal Boundaries

- Identifying and handling conversations involving sensitive subjects
- Guidelines for avoiding legal pitfalls and liability issues
- When to escalate or seek guidance for complex or risky discussions

5. Providing Accurate and Verified Information

- Emphasizing the use of credible sources for information
- Verifying facts and avoiding misinformation
- Handling situations where accurate information is uncertain

6. Maintaining Professionalism and Tone

- Adopting an appropriate and respectful tone in interactions
- Guidelines for addressing user queries courteously
- Avoiding offensive or inappropriate language

7. Dealing With User Misunderstandings and Expectations

- Strategies for clarifying misunderstood queries or responses
- Managing user expectations about ChatGPT's capabilities
- Redirecting users to appropriate channels for specific needs

8. Escalation and Crisis Management

- Recognizing situations that require escalation to human staff
- Steps to follow in the event of a crisis or emergency situation
- Ensuring clear communication with users about limitations and assistance

9. Continuous Learning and Feedback

- Encouraging employees to learn from interactions with users
- Sharing best practices and success stories within the team
- Providing mechanisms to gather feedback and improve the AI's performance

10. Monitoring and Compliance

- Introduction to monitoring tools and quality assurance processes
- Periodic review of AI-generated content to ensure policy adherence
- Reporting concerns or violations related to the ChatGPT employee policy

COMING SOON

Join us!
I have upcoming webinars with CareerLearning.



What was your key takeaway?

Please let us know in the Q&A.

We will use your response with your first name and last initial on our website so others can benefit!



Questions?



**Dr. Jim Castagnera, Esq.
President**



Dr. Jim's One-Stop HR Shop



<https://www.drjimshrshop.com/>



610-742-0424



dr.castagnera@gmail.com

Disclaimer

Our webinars are designed to be accurate and authoritative, but we do not provide legal, accounting, medical or similar professional services. Client remains solely responsible for compliance with all applicable laws, regulations, accounting standards and official health guidance. If legal, accounting, medical or other expert advice is desired, Client should retain the services of an appropriate professional.

Photocopying, distributing or using these copyrighted materials electronically without the copyright owner's express written consent is strictly prohibited. Consent is granted for the use of purchasers of webinars and then solely for the use of registered attendees of the webinars. No right or license is given to reproduce the materials in any form or format or to place the materials in any format on any website or blog or to otherwise republish it in any manner without the express written permission of the copyright holder.