



## Critical AI Literacies: A Guide for Students

*These guidelines were prepared by Critical AI @ Rutgers as part of DESIGN JUSTICE LABS, an NEH and Rutgers-sponsored initiative for sharing tools for teaching and researching critical AI literacies. For more information, or to comment or contribute, write to [criticalai@sas.rutgers.edu](mailto:criticalai@sas.rutgers.edu)*

“Generative” artificial intelligence (AI) tools are now embedded in popular software, tempting you to regard these commercial applications as reliable “copilots” and tutors. While these tools generate content in seconds, they rely on immense troves of human-generated “training data” and the energy-intensive computing resources necessary to statistically “model” it. Generative AI is the product of humans’ work, including the centuries of writing used in training (often without consent); the work of technologists in fields like computer science and linguistics; and the poorly compensated labor of human annotators whose ongoing work is necessary to make such “AI” seem more reliable and human-like.

Generative AI leverages a data-driven technique called “machine learning”—a process of statistical pattern-finding through which models “learn” to “optimize” for useful predictions. (They do this by updating the weights in a statistical message-passing “architecture”) Although the human-like text and images these systems produce can be hard to distinguish from content created by people, machine learning systems are disembodied statistical models that do not understand language or images in the way humans do.

Machine “intelligence” excels at pattern-matching at vast scales, enabling generative AI tools to work probabilistically—by predicting the most likely responses to a user’s prompt. However, generative AI tools are incapable of anything like a full-bodied “experience”: as disembodied statistical models they have no access to the world beyond a user’s inputs, no ability to update themselves between rounds of training, and (for all these reasons) no means of assuring the truthfulness of their outputs. That is why generative AI requires so much oversight

from human workers, including users. These systems should not be confused with search engines and cannot be trusted to generate accurate writing without considerable proofreading and fact-checking.

Discussing the ethics of generative AI (and “AI” more generally), is a complicated subject that you should reflect on after getting the facts (see below). Many people believe that it is impossible to use generative tools ethically. The underlying models on which generative AI is built recapitulate biases and stereotypes; infringe on copyright protections; surveil users (including students); leak data; expend enormous amounts of energy, water, and investment; and concentrate enormous power and resources in the hands of a tiny elite. The authors of this document, in accordance with Kathryn Conrad’s [“Blueprint” for AI rights](#), recommend that all students be given the right to opt out of any assignment that involves using such tools, even if no subscription is required. At the same time, since generative applications are now widely available, it is important to understand how they work and, for those who wish to use them, to learn how to avoid the most concerning academic and intellectual pitfalls. These may include the violation of your institution’s code of conduct and/or of your instructor’s course policies, or accusations of academic misconduct (which may result because of faulty AI “detection” tools). Above all, students who rely on chatbots may lose crucial opportunities to develop the foundational skills in writing, research, and critical thinking that make higher education valuable in the first place. Building critical AI literacies is a process of empowerment that enables students and citizens to exercise independent judgment about whether or if to use this very new and largely untested commercial technology.

### ***Learn What “AI” Is and How “Generative AI” Works***

“AI” is a loose term coined in the 1950s which refers to many different automated technologies and which is often used in science fiction (even though the human-like androids one often finds in such narratives do not actually exist). Generative AI is a type of machine learning that generates plausibly human-like content by offering the most statistically probable outputs in response to the user’s prompt. Reliance on probabilities can make generative outputs repetitive;

for example, employers have found that hundreds of applicants send near identical letters in response to a job ad, likely because they are prompting chatbots to respond to the job ad copy. The use of workers to remove objectionable content from chatbots can also make autogenerated text more formulaic and predictable. Scientists have discovered that AI [models trained on the outputs of earlier AI models will eventually “collapse”](#) as their content becomes increasingly predictable and homogenous (see, e.g., Goodlad and Stone [2024](#)).

### ***Understand the Policies Appropriate for Use of Generative AI as Specified by Your Instructors and Institution***

Many institutions have a code of conduct and/or policies for “academic integrity” which have been adopted by the institution and/or by instructors to govern the acceptable use of gen AI tools. Different instructors may have different policies, in part to respond thoughtfully to the learning goals for their course. It is important to identify and follow relevant guidelines for the use of automated tools which may be on a course site or syllabus. If in doubt, ask your instructor! While some instructors allow students to use content or information from a generative AI tool so long as the content in question is appropriately cited, others believe that chatbot use interferes with the specified learning goals for their course.

### ***Exercise Judgment by Deciding for Yourself Whether Chatbot Use Makes Sense for a Given Purpose.***

Once you recognize the considerable harms of generative AI (see below), you will be better prepared to make good choices about when or whether to use these tools. Given that students must always consult credible sources to fact-check and evaluate the often unreliable outputs of generative AI, it is reasonable to conclude that it makes sense to use search engines in the first place. (Using a chatbot for simple information retrieval is estimated to be at least ten times more energy intensive than using a conventional search engine without synthetic features such

as “AI overviews”).

***Recognize that Developers Use Human-Generated Scripts and Human Reinforcement to Make Chatbots Seem More Intelligent and Human-Like than They Actually Are***

Generative AI has been implemented in the form of conversational chatbots to encourage habitual use. But despite their human-like fluency, these systems don't understand language in a human-like way. Trained on data from the internet, they are chockful of stereotypes and misconceptions. As journalist Kyle Chayka ([2024](#)) points out, bots are better at mimicking “the semblance of emotional connecting” than they are at delivering reliable information. Chatbots can be deceptive and habit-forming. Any student who is concerned about chatbot use should contact a counselor, trusted teacher, or medical professional.

***Learn More about the Actually Existing Harms of Generative AI***

**DESIGN JUSTICE LABS provides resources for the teaching and study of critical AI literacies and design justice perspectives.** One of our goals is to teach students to understand generative AI tools in the way data scientists do—as researchers rather than consumers. To learn more about student research projects involving generative AI, write to us: [criticalai@sas.rutgers.edu](mailto:criticalai@sas.rutgers.edu)