# From Hackers to Chatbots

## A Quick Tour of Computer and AI Ethics

Michael A. Covington, Ph.D.

Senior Research Scientist Emeritus
Institute for Artificial Intelligence
The University of Georgia

Consultant
Covington Innovations
Athens, Georgia

Wisdom/Work

WISDOM/WORK
Published by Wisdom/Work
[TomVMorris.com](TomVMorris.com)

ISBN 979-8-9992362-0-3 (paperback)

For more information see this book's web site at
[www.CovingtonInnovations.Com/ethics](www.CovingtonInnovations.Com/ethics)

Typeset July 17, 2025.
Cover design by Melody M. Covington.

The author would like to thank Melody M. Covington, Tonita Branan, Tom Morris, Daniel Crider, and Jim Mischel for valuable insights. The opinions expressed here are solely his own.

# 1 EASIER THAN YOU THINK

Computer and AI ethics are easier than you think, for one big reason.

That reason is simple: if it's wrong to do something without a computer, it's still wrong to do it with a computer.

See how much puzzlement that principle clears away.

Consider, for example, the teenagers in several places who have used generative AI to create realistic nude pictures of their classmates. How should they be treated? Exactly as if they had been good artists and had drawn the images by hand. The only difference is that computers made it easier. Computers don't change what's right or wrong.

If that's all there is to it, why is computer ethics or AI ethics even a subject? Haven't we reduced all of it to the ethics we already know?

No, and again there's a big reason. Computers put people in unfamiliar situations, where they don't know how to apply ethical principles. Those unfamiliar situations are what we need to explore.

The most dangerous of those situations involve artificial intelligence. The problem is not conscious machines taking over; we have no conscious machines. What we have are gullible people giving imperfect machines too much control over important matters. I will explore that in some depth.

Computer ethics resembles medical ethics in two ways: it involves specialized technical knowledge, and the people harmed by a bad decision may never know any wrong was done. In the case of

# 2 OLD PRINCIPLES IN NEW SITUATIONS

For thousands of years, human society has been developing rules and expectations about how to handle and transmit information. Don't break confidentiality, don't make false accusations, don't believe everything you hear, and so forth.

All of a sudden, computers have vastly increased people's ability to receive, process, and spread information of all kinds. (Printing and broadcasting did something similar, earlier, but most people were only on the receiving end.) Nowadays, people often find themselves in situations that centuries of civilization did not prepare them for. In this chapter I'll explore what some of these new situations are.

## Hacker culture and the brave new world

The most pervasive illusion associated with computing is that when you're using a computer, you're in a whole new world, and your traditional obligations don't apply. For some people, this is a real belief about ethics; for others, it's simply loss of familiar habits. Even people who have a firm grip on the ethics of ordinary computing can lose it again when they move to advanced AI.

To a considerable extent, this separate-world illusion first arose in the "hacker culture" of academic computer labs, especially MIT and Stanford, in the 1970s, when the Internet was first being set up, before personal computers were common.

"Hacker" in those days was a term of approval; it meant a person who programs computers for the joy of it. (Computer criminals adopted the term later, in the 1990s.) The culture of the early hackers is documented in detail in *The New Hacker's Dictionary,* edited by

# 3 COMPUTERS AS MASS COMMUNICATION

Many of the ethical challenges relating to computers come from the fact that computers have made mass communication available to ordinary people (and impostors and scammers) on a grand scale that people are not prepared to handle. Just fifty years ago, reading matter and radio and TV shows all came from a well-defined publishing and broadcasting industry, in which relatively few people were able to speak to large audiences. Ordinary people could not speak to the public.

Now, with web pages and social media, anyone can reach a large audience, often without realizing it. Anyone can also make and redistribute a perfect, lossless copy of anything that arrives on their computer, opening up new possibilities for deception. The world has changed! In this chapter, I won't talk about how the new mass media work – everyone already uses them – except to counter a few misconceptions. I'll focus on distinctly ethical challenges, first from the viewpoint of the recipient of the information, then the sender.

## Weighing information sources

Anyone using the Internet for the first time is likely to be overwhelmed by the deluge of information about all subjects, from all kinds of sources, many of them eloquent and appealing but almost impossible to verify. The sources range from well-known organizations such as the BBC, to equally authoritative-looking sources one hasn't heard of, to discussions in forums and personal messages that say "pass this on," and even e-mail that arrives unbidden, com-

# 4 ARTIFICIAL INTELLIGENCE

Many people have been told that something called "artificial intelligence" (AI) was suddenly achieved around 2023, that computers are now conscious like humans (or are about to be), and that this warrants wild optimism, severe panic, or a mix of the two.

There was indeed a sudden advance – the advent of GPT LLM chatbots with huge training sets, which I'll get to in the next chapter – but artificial intelligence is not a new idea, nor are computers conscious. I have had the odd and frustrating experience of being shouted down in discussions because I know how AI works, and people want to hear a wilder or more exciting story from others who are just guessing. In what follows I'll tell you what you need to know to distinguish reality from fantasy.

## Origins and goals of artificial intelligence

The term *artificial intelligence* originated at a conference at Dartmouth College in 1956 on the use of computers to do tasks that require human intelligence. One can argue that even the earliest computers did that – arithmetic used to require human intelligence – but AI meant tasks requiring more study of human cognition and less use of already-known mathematics.

Almost from the beginning, there were differing opinions about whether computers could actually do what humans do. Can we figure out how the brain works, build a machine that does the same thing, and have a conscious, thinking computer? Or do we need to focus on studying particular tasks and making computers do them, whether or not the mechanisms they use are humanlike? The first of these is called *strong AI,* with the key claim that a computer can

# 5 CHATBOTS AND GENERATIVE AI

A chatbot is a computer program that carries on a conversation in a human language. The term comes from “chat robot” but denotes a piece of software, not a robot.

What shook the world in early 2023 was a new type of chatbot called ChatGPT, released by OpenAI, which was able to carry on conversations with people in English and perform many helpful intellectual tasks. Similar products from other companies followed quickly, such as Microsoft Copilot, Google Gemini, and Meta AI’s Llama. It looked as if humanlike thinking machines had finally been achieved.

Though immensely useful, the new chatbots are not quite as humanlike as they look, and this chapter will explain how they work and what their limitations are, with special reference to ethical questions.

## Chatbots: ELIZA in 1965

Way back in 1965, Joseph Weizenbaum of MIT built a chatbot that convinced people it had humanlike intelligence. His chatbot was called ELIZA (*Communications of the ACM* 9:36–45, 1966), named after the character in *My Fair Lady*.

ELIZA was intended only to demonstrate human-machine interaction in English. It pretended to be a non-directive psychotherapist and offered counseling. People thought it was actually understanding their thoughts and asked the laboratory staff to give them privacy when they used it.

Here is part of a conversation with ELIZA from Weizenbaum’s article. Sentences typed by the human are marked with “>”.

## Censorship, benign and malicious

Commercial LLMs normally prohibit certain kinds of output, and the prohibitions are becoming more extensive as people see more things that can go wrong. For instance, chatbots nowadays will not normally tell you how to make a bomb or how to kill yourself (we hope).

The end user cannot tell whether these restrictions are implemented within the LLM or as some kind of post-check. What is known is that if they are inside the LLM, it is often possible to "jailbreak" them by changing the context, such as saying you are writing fiction.

In any case, well-intentioned censorship can cause problems. At one time I heard from someone studying church history who was fine-tuning an LLM to summarize a set of religious books from the Puritan era. The chatbot refused to summarize some of the doctrines in the books because their morality was too conservative for the chatbot's modern standards of political correctness. Be aware: when you ask chatbots to summarize existing literature or history, they may leave things out.

One dramatic example is DeepSeek, an LLM distributed free of charge from China. It answers many questions well, but if you ask it about the Tiananmen Square incident, it claims not to know anything. What's funnier is that if you are taking the output in streaming mode (word by word), you can sometimes see it start to give an answer, then retract words and change the answer.

The lesson here is that from now on, political propaganda is going to be distributed in chatbots, not just news media. People will ask chatbots questions and imagine they are getting the whole truth, when in fact something has been left out at the behest of a special-interest group.

## Entertainer or companion?

Much uncharted territory is opening up rapidly as we look at possible uses of generative AI in entertainment. Chatbots can now write grade-B novels – in fact, book publishers are plagued with them and are trying to figure out what to do. For the most part, they

# 7 WHAT ETHICS IS NOT

So far, this book has been about the challenges to people's ethical reasoning that arise in new situations created by computers. Now I want to turn toward ethics itself.

For the purposes of this book, we do not have to answer deep questions about the origin and nature of ethics. We do not have to debate whether God exists or whether Kant was right about the starry heaven above and the moral law within. For almost everything, we can rely on what clear thinkers already agree on. Nonetheless, in this chapter I want to point out some things that ethics is not. It can be a relief to find that you can be ethical without solving all these problems or developing all these skills.

## Ethics is not emotions

Computer experts are often afraid they won't be good at ethics because it is subjective, touchy-feely, and emotional, not the kind of thing engineers are good at. Actually, ethics is not a matter of emotions, empathy, or even social skills, and that is true in several different ways.

It's true that ethical challenges are often phrased as, "How would you feel if such-and-such were done to you?" But the real question is not "What would your emotions be?" but rather "Would you consider yourself harmed or benefited?"

Indeed, the whole point of ethics is to be fair to people, not to make them happy. If you miss that point, you'll spend your time dealing with squeaky wheels, people who have made up their minds not to be happy no matter how fairly they're treated. You will also

# INDEX

# ABOUT THE AUTHOR

Michael A. Covington, Ph.D., was co-founder and long-time associate director of the Institute for Artificial Intelligence at the University of Georgia, where he now holds the title of Senior Research Scientist Emeritus. He also chaired the University's first computer ethics committee. He is available for consulting and public speaking in these and related areas; see www.covingtoninnovations.com.