

Understand the Machine

A DISCUSSION GUIDE

AI Ethics Conversation Starter Kit

Five big questions about AI — bias, privacy, jobs, transparency, and accountability — with real cases and ready-made questions for talking it through with family, friends, or colleagues.

Free AI literacy for everyone — no tech background needed

understandthemachine.org

You don't need a philosophy degree to think about whether AI is being used fairly. This kit gives you five real questions, a short true story for each, and discussion prompts you can use at the dinner table, in a classroom, or at work.

AI ETHICS IN 60 SECONDS

AI raises hard questions about fairness, privacy, jobs, transparency, and who's responsible when it gets things wrong. The goal isn't to be for or against AI — it's to make sure a powerful technology is built and used in ways that treat people fairly.

How to Use This Kit

- Pick one question per conversation — you don't have to cover all five at once.
- Read the short case aloud, then work through the discussion questions together.
- There are no "right" answers. The point is to think out loud and hear other views.
- Great for families, classrooms, book clubs, team lunches, or community groups.

Question 1: Is AI Biased?

AI learns from data, and data reflects the world as it is — including its unfairness. If a system trains on biased data, it makes biased decisions. Not on purpose: it's doing exactly what it was built to do — find patterns. The trouble is that some of those patterns are patterns of discrimination.

True story: A widely used U.S. hospital algorithm was meant to flag patients needing extra care. In 2019, researchers found it gave lower risk scores to Black patients than to equally sick white patients — because it used past healthcare *spending* as a stand-in for health needs, and decades of unequal access meant less had been spent on Black patients. No one intended harm; the chosen data carried the bias.

Talk about it

- 1 Where in your own life might a decision about you be made partly by AI?
- 2 If an AI is "just following the data," who is responsible when the data is unfair?
- 3 Is it possible to build an AI with no bias at all? What would that take?
- 4 Would you rather have an important decision made by a person or by an AI? Why?

Question 2: What Happens to Your Privacy?

AI is hungry for data — the more it has, the better it performs. That pits powerful AI against personal privacy. Your typing speed, location, voice, and browsing can all be combined into a remarkably detailed picture of your life: your habits, health, and beliefs.

True story: Clearview AI scraped billions of photos from social media without asking, building a face-search tool that could identify almost anyone from a single picture — and sold access to police. The company said the photos were "public." Most people who posted them never imagined a surveillance database. After lawsuits, in 2022 it agreed to restrict U.S. sales. The unresolved question: if data is public, can anyone use it for anything?

Talk about it

- 1 When did you last read a privacy policy before clicking "agree"? Why or why not?
- 2 Is a photo you post publicly "fair game" for a company to use however it likes?
- 3 What personal data would you never want an AI system to have? Where's your line?
- 4 Should there be things companies simply aren't allowed to collect, even with consent?

Question 3: Will AI Take My Job?

AI will change many jobs, eliminate some, and create others. The honest ethical question isn't about the technology — it's about how we handle the transition, and what we owe the people whose work is disrupted.

True story: There are roughly 3.5 million truck drivers in the U.S. Self-driving truck companies already run driverless trucks on some Texas highways. The tech isn't ready to replace most drivers yet, but it's heading that way. The hard questions — should companies retrain workers? should rollout be gradual? — are ethics questions, not engineering ones.

Talk about it

- 1 If a machine can do a job more cheaply, does the company owe anything to the workers it replaces?
- 2 Whose job is it to help people retrain — employers, government, schools, or individuals?
- 3 What kinds of work do you think will always need a human? Why?
- 4 Have you already seen AI change a job you know? What happened?

Question 4: Can We Understand How AI Decides?

Many AI systems are "black boxes": they produce a result, but no one — not even their builders — can fully explain why. As AI makes bigger decisions, that transparency gap gets more troubling.

True story: U.S. courts use AI tools to help judges set bail. A 2016 ProPublica investigation of one system, COMPAS, found it was almost twice as likely to wrongly flag Black defendants as future criminals than white defendants. The maker disputed it — and because the system's reasoning was hidden, even experts couldn't agree on whether it was fair.

Talk about it

- 1 If an AI affects your life, do you have a right to know how it reached its decision?
- 2 Is it acceptable to use a system nobody can fully explain, if it's accurate on average?
- 3 Should companies be forced to reveal how their algorithms work, even if it's a trade secret?
- 4 Who should get to check whether these systems are fair before they're used?

Question 5: Who's Responsible When AI Goes Wrong?

When a person makes a harmful mistake, you can hold them accountable. When an AI does, who's to blame — the company that built it, the one that deployed it, the engineer, the manager? Everyone in the chain can point to someone else. This is the "accountability gap."

True story: In 2018, an Uber self-driving test car struck and killed a pedestrian in Tempe, Arizona — the first such fatality. The sensors detected her but the system dismissed it as a false alarm; the human safety driver wasn't watching the road; and the company had prioritized speed over caution. The safety driver was charged; Uber settled. But should the engineers and managers share responsibility too?

Talk about it

- 1 When an AI and a human both fail, how should blame be divided?
- 2 Should a company be able to say "the algorithm did it" and avoid responsibility?
- 3 Would clearer laws about AI liability make companies more careful — or slow down useful technology?
- 4 Who do you trust to set the rules for AI: companies, governments, or independent groups?

Where the Rules Stand

European Union

The AI Act (from 2024) sorts AI by risk. High-risk uses — hiring, healthcare, policing — face strict rules; some, like real-time public face recognition, are largely banned. Fines reach €35M or 7% of global revenue.

United States

No single federal law — a patchwork of existing rules and state laws. Colorado, for example, requires businesses to disclose AI used in significant decisions. Policy shifts with each administration.

UNESCO (global)

In 2021 all 193 member states adopted the first global agreement on AI ethics. It isn't binding, but it sets shared principles: protect human rights, ensure transparency, and don't widen existing inequalities.

What You Can Do

- ☐ **Ask "is AI involved?"** when you get an important decision — you have a right to know.
- ☐ **Skim the key parts of privacy policies** — data collection, sharing, and automated decisions.
- ☐ **Adjust your privacy settings** on your phone, browser, and social accounts.
- ☐ **Talk about it** — raising the questions with others is itself progress.
- ☐ **Tell your representatives** you care about AI rules — the decisions are being made now.

Sources & further reading

- UNESCO — Recommendation on the Ethics of Artificial Intelligence (2021), unesco.org
- NIST — AI Risk Management Framework, nist.gov
- EU AI Act — summary and text, artificialintelligenceact.eu
- ProPublica — "Machine Bias" investigation of COMPAS (2016)
- Obermeyer et al. — study of racial bias in a health algorithm, *Science* (2019)

Free forever — help keep it that way

Understand the Machine is a nonprofit. Every guide, PDF, and newsletter is free. If this helped you, a \$15 gift keeps them free for the next person. Give at understandthemachine.org/donate

This guide lives online, always up to date, at understandthemachine.org/ai-ethics-guide/

© 2026 Understand the Machine · Shared freely under CC BY-NC 4.0 — attribution appreciated.