Ethical AI Principals

(Video Transcript)

Let's start with a definition of what ethical AI is. Ethical AI seeks to promote fairness, minimize harm, and align AI with human values and well-being.

If I'm a designer and I'm working in ethical AI, one of the questions I really want to ask is: *How do I design artificially intelligent systems that follow ethical principles in society?*

When we think about the practice of building AI systems, ethical AI is that practice that follows ethical and moral standards in society. But it is dynamic—right? As AI changes, ethical AI must change to maintain momentum and keep pace with the advent of AI technology.

One of the big questions that ethical AI asks, from a bird's-eye perspective, is: *What are the legitimate and illegitimate uses of AI?*

Transparency

One of the things we talk about a lot in ethical AI is transparency, which is exactly what it sounds like: allowing users to understand how decisions are made in the AI systems they're using.

This really builds trust with the user and fosters a relationship that enables the user to assess how the system behaves with regard to bias and fairness. It allows them to see what's going on in the background when they're using these AI systems.

Accountability

There's also accountability, which means that the designers, developers, organizations—basically every stakeholder with a stake in the matter, including users—should be held accountable for the impact and consequences of AI systems.

This means things like ease of troubleshooting, understanding what errors are present, and how to address those errors. These mechanisms should all be included within the AI system being used.

Fairness

Then there's fairness. Outcomes should be fair and equitable for users. By mitigating bias and addressing disparities in AI outcomes, fairness can be achieved.

This is definitely easier said than done, as we'll see later, but the concept centers on ensuring that outcomes aren't biased or unfairly favoring one demographic, region, area, person, or industry over another.

Privacy

And then there's privacy. Individuals have rights regarding their privacy and data—especially in light of numerous consumer data laws.

The concept is built on the idea that these systems must be designed to handle user privacy and data in a secure and ethical manner. Information should be obtained consensually.

This not only means that I, as a user, give my data and consent to its use, but that I also understand what it means to permit a system to have my data. For instance:

- **Are they going to sell it?**

- **What are they going to do with it?**

I, as a user, should have a very thorough understanding of that, and that understanding should be accessible to everyone. That's the basic idea of privacy in ethical AI.