

Discussion Guide: Ch 11, AI Ethics¹

Goal: Move beyond "rules" and discuss the real-world friction of power, politics, and the limits of human foresight.

Topic 1: The "Mindlessness" Trap

Philosopher Hannah Arendt suggested that "evil" often comes from **mindlessness**—people just doing their jobs without thinking about the consequences.

- **Discussion Question:** Is it "evil" for a coder to build an algorithm they don't fully understand if their boss tells them to? Who is responsible for the "mindless" use of AI: the person using it, or the person who didn't provide a "driver's license" or training for it?

Topic 2: The Stop Button

The text mentions that "Responsible Innovation" tools often assume the technology **will** be built and just try to make it "better".

- **Discussion Question:** If your ethical analysis showed that a new AI tool would definitely increase social inequality, but your company stood to make \$100 million from it, do you think the "ethical principles" we talked about in Ch 10 would actually stop the project?

Topic 3: The STEM/Humanities Divide

The author argues that engineers need to "learn to do things with texts" and humanities people need to "learn to do things with computers".

- **Discussion Question:** In your own education so far, do you feel like these two "worlds" are separate? How would your major change if "Ethics" wasn't a separate class, but a core part of every technical assignment?

¹ AI help generate this page, with instructor edits