

Four reviews. Five principles. One decision routine.

Use this map to prepare for class and to check your memo. It is a course synthesis of four reviews—not a framework the four author teams jointly proposed.

1 · What each review contributes—and what it cannot establish alone

Jobin, Ienca & Vayena (2019) Map: 84 AI ethics guidelines; recurring principle language and differences in interpretation. Best question: Which values recur across guidelines? Limit: The review does not prescribe one university policy.	Corrêa et al. (2023) Map: 200 guidelines and governance documents; recommendations and implementation gaps. Best question: What turns an ethical promise into an operational commitment? Limit: Published guidance is not the same as enforcement or demonstrated impact.	Giarmoleo et al. (2024) Map: 309 articles through December 2021; system-design and human–AI interaction concerns, plus act- and agent-centered responses. Best question: Is the problem in the system, its use, the rule, or human judgment? Limit: The literature window ends in 2021.	Groen, Sharon & Becker (2026) Map: moral concerns through principles, lived realities, and power structures. Best question: Who experiences the rule, who decides, and who bears the risk? Limit: This conceptual lens does not prove that a particular classroom policy works.
--	--	--	--

2 · Translate principle language into a governance question

PRINCIPLE FAMILY	PLAIN-LANGUAGE COURSE DEFINITION	GOVERNANCE QUESTION FOR A UNIVERSITY
Transparency	Make the AI role, relevant limits, and uncertainty inspectable.	What must be disclosed, to whom, at what time, and in what form?
Fairness / justice	Inspect how access, benefits, burdens, errors, treatment, and remedies are distributed.	Which groups, access differences, error patterns, remedies, and appeal paths matter?
Responsibility / accountability	Name the people and institutions responsible for checking, deciding, explaining, and remedying harm.	Who checks, decides, answers for harm, and supplies a remedy?
Privacy	Constrain data collection, use, retention, inference, and sharing.	What data is collected, retained, inferred, shared, protected, or made contestable?
Safety / non-maleficence	Anticipate, test, reduce, monitor, and stop foreseeable harm.	What failure modes must be tested, and who has authority to stop the use?

Use the map: principle → design/use concern → governance commitment → lived reality and power → evidence and revision. For the college-paper case, ask what is known, what remains unknown, what action is justified now, and what process protects both learning and fair treatment.

Course readings: [Principles to Power](#) · [AI Ethics in College](#). Check the assignment sheet for full citations and accessible reading routes.