This **Code of Ethics** is designed for students conducting the Week 4 **Bias Audit** . It establishes the "non-negotiable" requirements for using AI in high-risk educational and hiring contexts, ensuring that innovation serves as a leveling tool rather than a mechanism for deepening stratification .

## Code of Ethics for Algorithmic Auditing

### 1. The Principle of Human Precedence

* **Co-pilot, Not Autopilot:** AI must be treated as a tool for augmentation, not an autonomous decision-maker.
* **Mandatory Oversight:** Any consequential decision—including admissions, grading, or "at-risk" labeling—must involve meaningful human review.
* **The Right to Override:** Institutions must maintain the structural capacity to override any automated recommendation that contradicts human judgment or ethical standards.

### 2. Radical Transparency and Explainability

* **Candidate Notification:** Students and job applicants have an absolute right to be informed when AI participates in their evaluation.
* **Explainable Logic:** Audited systems must provide an interpretable "right to explanation," detailing the logic used to reach a specific score or ranking.
* **Disclosure of Constraints:** Auditors must explicitly document the limitations and "epistemological limits" of the data used to train the model.

### 3. Rigorous Bias Mitigation

* **Mandatory Independent Audits:** Because 80% of educational AI systems show bias when left unmonitored, audits must be regular, mandatory, and independent.
* **Disaggregated Reporting:** Performance and error rates must be published and broken down by socioeconomic status, race, gender, geography, and disability status.
* **Detecting False Negatives:** Auditors must specifically test for false-negative rates—such as the documented 19% to 21% error rates for minority students in predictive models—to prevent unfair exclusion.

### 4. Protection Against "Mission Drift"

* **Intellectual Integrity Over Efficiency:** Systems must not be optimized for "engagement metrics" or "organizational productivity" at the expense of critical thinking or long-term human development.
* **Guarding Against Proxy Discrimination:** Auditors must identify and eliminate "proxies" for privilege, such as using ZIP codes or parental education levels as stand-ins for academic potential.

### 5. Data Sovereignty and Privacy

* **Learner Ownership:** Digital learning records and verified credentials belong to the learner, not the platform or the employer.
* **Protection from Surveillance:** Auditors must resist the normalization of "continuous algorithmic evaluation" that transforms learning into a form of permanent surveillance .

### Audit Compliance Statement

Any AI system used within this "Crucible" foundation year must meet these standards to be considered "high-risk" compliant under the **EU AI Act** and our internal design anchors. Failure to meet these criteria requires immediate redesign or decommissioning of the tool.

**Would you like me to draft a "Career Compass" guide for students on how to present this Capabilities Dossier to employers in Future 3 (The Redesign)?**