

CS120 Project Proposal Rubric

Due: Friday, October 30, 2026, at 12:00 p.m. (noon) Pacific

Format: PDF, two pages, including references

Submission: Gradescope; one proposal per project, with all team members listed

Instructions: The proposal is worth **15% of the course grade**. This proposal should identify a concrete AI-safety problem and explain how you would investigate or assess an assurance claim. The grading criteria below correspond to the Gradescope rubric.

Section	Points (20 total)
Q1. Safety Problem and Assurance Claim	4
Q2. Proposed Methodology	6
Q3. Threats to Validity and Limitations	4
Q4. Feasibility and Scope	4
Q5. Clarity and References	2

Q1. Safety Problem and Assurance Claim (4 points)

- **Q1.1 (1 point):** Clearly identifies the AI system and its boundaries, including the intended deployment or use context, relevant stakeholders, and human inputs or oversight, where applicable.
- **Q1.2 (1 point):** Defines a specific safety problem or failure mode within the identified system boundary.
- **Q1.3 (2 points):** States a concrete research question or assurance claim, specifying what aspect of safety will be investigated.

Q2. Proposed Methodology (6 points)

- **Q2.1 (2 points):** Describes a concrete evaluation, experiment, audit, or analytical approach.
- **Q2.2 (2 points):** Explains how the method addresses the research question by identifying key assumptions, the evidence to be collected, the inferences that evidence could support, and the anticipated limits of those inferences.
- **Q2.3 (2 points):** Identifies relevant metrics, evidence, baselines, or criteria for evaluation.

Q3. Threats to Validity and Limitations (4 points)

- **Q3.1 (2 points):** Identifies and discusses a primary threat to validity (e.g., evaluation gaming, benchmark contamination, confounding factors, or inaccessible evidence), while noting other material threats where relevant.
- **Q3.2 (2 points):** Discusses access constraints, assumptions, or practical limitations of the proposed approach, including how these may restrict the conclusions that can be drawn.

Q4. Feasibility and Scope (4 points)

- **Q4.1 (2 points):** Defines a realistic scope with achievable goals.
- **Q4.2 (2 points):** Outlines concrete steps, milestones, or expected deliverables.

Q5. Clarity and References (2 points)

- **Q5.1 (1 point):** Clear, organized, and concise writing.
- **Q5.2 (1 point):** Appropriate references and adherence to submission requirements.