Repository Entry Template
Embedded EthiCS @ Harvard Teaching Lab

<table>
<tr><td colspan="2" align="center">Overview</td></tr>
<tr><td align="right">Course:</td><td>CS 290A Seminar on Effective Research Practices & Academic Culture</td></tr>
<tr><td align="right">Course Level:</td><td>Graduate</td></tr>
<tr><td align="right">Course Description:</td><td>This is a reading and discussion-based seminar designed for entering Computer Science Ph.D. students. This course prepares students to manage the difficult and often undiscussed challenges of Ph.D. programs through sessions on:

1. Research skill building (e.g. paper reading, communication)
2. Soft skill building (e.g. managing advising relationships, supporting your peers)
3. Academic culture (e.g. mental health in academia, power dynamics in scientific communities)
4. Research and professional-oriented discussions[1]</td></tr>
<tr><td align="right">Module Topic:</td><td>Moral and Professional Responsibility in Computing</td></tr>
<tr><td align="right">Module Author:</td><td>Trystan Goetze</td></tr>
<tr><td align="right">Semesters Taught:</td><td>Fall 2022</td></tr>
<tr><td align="right">Tags:</td><td>Moral responsibility [phil], professional responsibility [cs], research ethics [cs]</td></tr>
<tr><td align="right">Module Overview:</td><td>This module examines difficulties in tracing moral responsibility for the consequences of the development and use of a computing system. Through interactive lecture, students are introduced to a standard view of moral responsibility in analytic philosophy, then walked through several causal chains of decision-making that lead to computer systems causing harm. Students are challenged to consider which parties in the chain may be morally responsible for the downstream effects. The distinction between backward-looking and forward-looking responsibility is introduced, and students are asked to consider whether any parties in the causal chain have forward-looking responsibilities in virtue of their connection to the resulting events and their professional roles.

The active component of the module has students apply these concepts to a specific case of an autonomous vehicle that collides with a pedestrian. They are asked to complete a causal chain of different decision-makers along the way to the accident and to consider who among them may be</td><td>The module draws on work that I've published (link) as well as a previous module by Ellie Lasater-Guttmann (link).</td></tr>
</table>

---

[1]Course website: https://yanivyacoby.github.io/harvard-cs290/

	backwards-responsible or have forwards-responsibilities.	
Connection to Course Material:	One goal of this seminar is to introduce PhD students to professional norms and issues in computational disciplines of study and practice. Since they will be taking up major roles in tech research and industry, becoming aware of their professional responsibilities to stakeholders in the technologies they develop is an important part of their professional development.	The course instructors expressed a desire to have a module that would appeal across the different specializations represented in the class. Some students are interested in applied areas of computer science, but others are theorists working on highly abstract computational methods. The topic of responsibility was chosen because the causal chains enable the illustration of how decisions at each step of the way, going all the way back to theoretical inquiry, have an influence on downstream technological development and its consequences.

Goals

Module Goals:	1. Understand the concepts of backward- and forward-looking moral responsibility 2. Apply these concepts to a concrete case study in computer science 3. Appreciate the connection between one's professional role and one's moral duties	
Key Philosophical Questions:	1. When a computer system causes harm, who is responsible? 2. Whether or not anyone is responsible in a backwards-looking sense, does anyone have any forward-looking responsibilities in connection with the harm? Why? 3. Which contextual details are important in answering these questions?	These questions are addressed directly through the activity.

Materials

Key Philosophical Concepts:	● Moral responsibility ● Professional responsibility ● Backward-looking and forward-looking responsibility ● Causal chains	These concepts are introduced via the lecture. Students then practise applying them in small groups, followed by whole class discussion.
Assigned Readings:	● Helen Nissenbaum, "Computing and Accountability" ([link]) Pre-class assignment:	This reading introduces some of the philosophical background to the topic. Nissenbaum also argues against several excuses that are

	Search the web for a recent news story about some scandal involving computing technology, a tech company, or a computer science research group.
	  1. Briefly summarize the scandal, and provide the link to the news story.
	  2. Identify the different parties involved in the scandal: who were the developers, the end users, the victims, etc.?
	  3. Who, among those parties, should be held accountable for their role in the scandal, and in what way? Refer to Nissenbaum's arguments in your answer.
		used to deny one's responsibility for the harms caused by technologies to which one contributed. She also makes some policy suggestions that go beyond the scope of the module, but are useful for students to consider.

Implementation

Class Agenda:	1. Course head recaps previous session 2. Introduction to Embedded EthiCS 3. Concept of moral responsibility 4. Distinction between backward- and forward-looking responsibility 5. Illustration of causal chain and questions about responsibility at each step 6. Small group activity 7. Take-up of group activity as a class 8. Wrapping up and feedback	

Sample Class Activity:	Students are given a partial causal chain leading from theory of machine learning and computer vision, to applications in autonomous vehicles, to a specific accident where an autonomous vehicle kills a pedestrian.
	
	Consider a case where a pedestrian is killed by an autonomous vehicle on a busy city street.
	  1. Are there any relevant agents missing from this chain of events?
	  2. Which of these agents are responsible for the death? Why or why not?
	  3. Which of these agents have forward-looking responsibilities to prevent such harms? Why or why not?
		This activity requires students to engage with the philosophical material by sketching the causal chain and reflecting on which agents in the chain may be responsible, in either sense.

Module Assignment:	End of class reflection submitted through Canvas: How does what we discussed in today's session relate to your practice as a researcher or professional?	There was no room in the course structure for an additional graded assignment.
Lessons Learned:	Students were mostly highly engaged with the topic and interested in puzzling out who is responsible, what the various contextual elements of the cases are, and how different parties should respond. There	

was some productive disagreement in the whole class discussion that illuminated the difficulties surrounding these issues. Some students felt that the reading was dated because of its examples. Another point of criticism from students was that the pacing was too slow, and could have eliminated some exposition before the activity.