COMP 301: Ethics and Accountability in Computer Science

Spring 2019

Tuesdays and Thursdays, 9:25 am - 10:40 am in Duncan Hall 1046

Co-instructors:

Tina L. Peterson, Ph.D. (tina.peterson@rice.edu); office hours Wednesdays 3 pm - 4 pm or by appointment in Mech Lab 106

Moshe Vardi, Ph.D. (vardi@cs.rice.edu); office hours by appointment in Duncan 3057

TA: Kevin Smith (kwsmith@rice.edu); office hours Wednesdays 1:15 - 2:15 pm or by appointment in Duncan 3062

COURSE DESCRIPTION

Computer scientists have a great deal of power in the twenty-first century; they have duties both to society and their own profession to wield that power wisely and responsibly. In this discussion- and reflection-oriented course students will master fundamentals of moral philosophy and principles of social responsibility; they will apply these modes of reflection and reasoning to current issues in computer science. These include but are not limited to algorithmic fairness and decision-making; surveillance and privacy; biased data and machine-learning outcomes; autonomous systems and human casualties; and diversity and inclusion in the tech industry.

LEARNING OBJECTIVES

During this course students will:

- Develop an understanding of major schools of thought in moral philosophy
- Master principles of social responsibility related to technology
- Apply ethical decision-making models to case studies
- Develop an awareness of potential consequences of their decisions as computer science professionals

COURSE MATERIALS
All readings are linked to in the course schedule. No textbook purchase is required. The instructors will do everything in their power to give students access to articles and other media free of paywall obstacles; any problems accessing the readings should be reported to the instructors via e-mail well ahead of the class meeting for which the reading is due.

ONLINE FORUM POSTS

Articles, other materials and discussion prompts will be posted to the class’s Canvas on a regular basis. Students will be expected to post a minimum of 5 one-paragraph critical responses and 10 short but meaningful replies to others’ posts to the forum over the course of the semester. Points will be awarded based on eloquence and quality of critical thought.

DEVICES IN CLASS

As this course is based on discussion and reflection, laptop, tablet, and smartphone use will NOT be allowed in class as it is potentially distracting to both the user and those sitting around her/him. Research has shown that non-academic laptop use in class is inversely correlated with class performance, and that taking hand-written notes encourages deeper processing of information.

ACCOMMODATIONS FOR DISABILITY

Students who require accommodation for a documented disability should speak with one of the instructors during the first week of class.

ASSESSMENT

This course is based on discussion and reflection, and opinions and perspectives may vary widely. Students will not be assessed (awarded points) based on whether their expressed perspective or analysis agrees with those of the instructors, but on how eloquently, thoughtfully and critically they answer questions posed, and how well they support their positions with high-quality evidence. They will also be assessed on the quality of writing in assigned papers. The point breakdown will be as follows:

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midterm exam</td>
<td>100</td>
</tr>
<tr>
<td>Group presentation on machine learning project</td>
<td>50</td>
</tr>
<tr>
<td>5 x 1-paragraph online forum posts</td>
<td>30</td>
</tr>
<tr>
<td>10 x thoughtful comments on others’ forum posts</td>
<td>30</td>
</tr>
<tr>
<td>6 x reading quizzes</td>
<td>30</td>
</tr>
<tr>
<td>First draft of response paper to “Black Mirror” episode</td>
<td>25</td>
</tr>
</tbody>
</table>
CALCULATION OF FINAL GRADES

Point totals will be converted into percentages, and final grades will be awarded as follows.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>100 – 96.0</td>
</tr>
<tr>
<td>B+</td>
<td>89.9 – 86.0</td>
</tr>
<tr>
<td>C+</td>
<td>79.9 – 76.0</td>
</tr>
<tr>
<td>D+</td>
<td>69.9 – 66.0</td>
</tr>
<tr>
<td>A</td>
<td>95.9 – 93.0</td>
</tr>
<tr>
<td>B</td>
<td>85.9 – 83.0</td>
</tr>
<tr>
<td>C</td>
<td>75.9 – 73.0</td>
</tr>
<tr>
<td>D</td>
<td>65.9 – 63.0</td>
</tr>
<tr>
<td>A-</td>
<td>92.9 – 90.0</td>
</tr>
<tr>
<td>B-</td>
<td>82.9 – 80.0</td>
</tr>
<tr>
<td>C-</td>
<td>72.9 – 70.0</td>
</tr>
<tr>
<td>D-</td>
<td>62.9 – 60.0</td>
</tr>
</tbody>
</table>

ATTENDANCE

Students will be expected to attend all class meetings and participate in class discussions in person and online. An attendance sheet will be passed around and collected within the first 10 minutes of each class meeting, and students' signatures on this sheet shall serve as the only record of their presence. Excused absences must be cleared with instructors in advance. Any more than two unexcused absences will affect students' point totals as shown below.

- 2 absences: no penalty
- 3 to 4 absences: - 20 points
- 5 to 6 absences: - 40 points
- 7 or more absences: failing grade for course

COURSE SCHEDULE AND REQUIRED READING (*subject to revision*)

All readings should be done in advance of the class for which they are listed. Quizzes on the readings will be given randomly at the start of eight classes throughout the semester; each student's lowest two quiz grades will be dropped.

8 Jan  In-class brainstorm activity: what powers do computer scientists have?

Review syllabus and expectations for course
The trolley problem and abstracted power

10 Jan     Lecture by Dr. Vardi: “An Ethical Crisis in Computing”
Assigned: One-page critical response to Dr. Vardi’s talk.

15 Jan     Human nature + computers = ?

Reading:
(https://www.newyorker.com/magazine/2012/02/06/the-story-of-a-suicide)

Due via Canvas by 9 am: One-page critical response to Dr. Vardi’s talk.

17 Jan     Tech and utopia vs. dystopia

UPDATED

Readings:
(https://www.eff.org/cyberspace-independence)

(https://www.sparknotes.com/lit/flies/)

22 & 24 Jan Code as law
Respect for human rights
Machine learning and algorithmic fairness in criminal justice
Transparency, accountability, and explainability
Readings:

Angwin et al. (2016). “Machine Bias.” ProPublica


29 Jan  
Machine learning and algorithmic fairness in employment decisions

Reading:


31 Jan  
Moral philosophy: Rawls' Theory of Justice

UPDATED  
Machine learning and algorithmic fairness in health care

Readings:

(https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5962261/)


5 Feb  
Present solutions to missing-data problem in organ

UPDATED  
transplant ML system

Gentrification and the tech industry

https://canvas.rice.edu/courses/20367
Readings:

Newitz (2017). "A startup aims to stop gentrification, with help from the tech industry." arstechnica.  

7 Feb       NO CLASS (spring recess)

12 Feb      “Move fast and break things” vs. duties and obligations of computing professionals

UPDATED

Readings:

The ACM Code of Ethics (draft 3, 2018)  
(https://ethics.acm.org/2018-code-draft-3/)

Hill (2017). “The ethical problem of software neglect.” Blog@CACM.  

14 Feb      Deliberate law-breaking via software

Readings:


19 & 21 Feb  Surveillance and privacy

Reading:

https://canvas.rice.edu/courses/20367

26 Feb  
Addictive apps, filter bubbles, and alienation

Reading/viewing:


28 Feb  
Midterm exam

5 & 7 Mar  
Autonomous systems and human casualties

Readings:


12 & 14 Mar  NO CLASS (spring break)

19 Mar  Moral philosophy: Kant's Categorical Imperative
Targeted advertising and users as commodities

Reading:


21 Mar  Ownership of data and models derived from human behavior

Reading:


Due on Canvas by 9 am: First draft of response paper to “Black Mirror” episode.

26 & 28 Mar  Diversity and inclusion in the tech industry

Readings:

(https://www.wired.com/story/computer-science-graduates-diversity/)


https://canvas.rice.edu/courses/20367

Assigned: Ethical machine learning projects

2 & 4 Apr  Civic institutions, social media accountability, and democracy

Readings:


9 & 11 Apr  The human consequences of the robot economy and the humans-as-robots economy

Readings:


Due: Final draft of response paper to “Black Mirror” episode.
Course Summary:

<table>
<thead>
<tr>
<th>Date</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tue Jan 15, 2019</td>
<td><a href="https://canvas.rice.edu/courses/20367/assignments/100768">One-page response paper to Dr. Vardi's Jan 10 lecture</a> due by 9am</td>
</tr>
<tr>
<td></td>
<td><a href="https://canvas.rice.edu/courses/20367/assignments/101600">Reading quiz 1</a> due by 9:35am</td>
</tr>
<tr>
<td>Tue Jan 22, 2019</td>
<td><a href="https://canvas.rice.edu/courses/20367/assignments/102470">Reading quiz 2</a> due by 9:35am</td>
</tr>
<tr>
<td>Thu Jan 31, 2019</td>
<td><a href="https://canvas.rice.edu/courses/20367/assignments/103994">Reading quiz 3</a> due by 11:59pm</td>
</tr>
<tr>
<td></td>
<td><a href="https://canvas.rice.edu/courses/20367/assignments/100951">Roll Call Attendance</a></td>
</tr>
</tbody>
</table>

16 & 18 Apr Group presentations of ethical machine learning projects

Re-visit first week brainstorm: what powers do computer scientists have?