Meeting the Ethical Challenges of Computing Research

Questions CS Graduate Students Should Ask

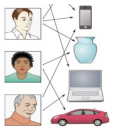
Barbara J. Grosz
Paulson School of Engineering and Applied Sciences
Harvard University

1

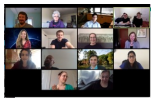
Exciting Advances Enabled by Computing Research

fraud detection

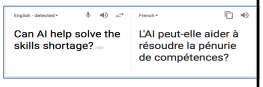
conversational agents



recommendations



connections



translation

2

With Widespread Use in Daily Life Many Risks of Societal Harms and Ethical Challenges

- Every citizen use
  - Diverse populations & cultures; divergent values
  - Machine learning perpetuating structural biases
- Systems used in unanticipated ways and contexts
  - Safety, security harder to ensure; imagination needed
  - Unsafe apps, training data differs from use population
- "Every" includes adversaries
  - Taybot, images, Polarization

and more in the press almost daily!

3

Participants in the "Traffic System": People, *not just* Autonomous Computer Agents

Oops: the roadways are not dedicated to autonomous cars (yet)…

Challenge: people's unexpected behaviors.

4

Ethics Considerations Have Two Complementary Roles

The challenges are socio-technical!

- Developing computing methods that enable technologies to meet societal needs.
- Designing, implementing, developing, deploying technologies in ways that avoid individual and societal harms.

Interactions: methods that enable societal "good", can cause harms if care is not taken throughout the pipeline from research to deployment.

5

Much Current News Centers on AI, but Ethical and Societal Challenges Pervade All CS Areas

- Theoretical computer science: Who/what considered in choice of optimization function for search or matching algorithms
- Systems: Hacking back; representational choices and harms; responsibilities of free software releases.
- Human-computer interaction: accessibility for all?
- AI: Which languages does a model handle? Whose data and for what uses? Autonomy vs. collaboration?

6

Ethics Along the Path From Computer Science Student To Computing Researcher, Scientist, Faculty Member

- Scientific problem: impact & a problem that matters
- Research team: choosing teammates and co-authors
- Testing: diversity of uses and potential user communities
- Publications & artifacts: specifying scope of intended use, testing, limitations.

At each step, pause to consider potential ethical and societal consequences.

7

Socio-Technical Design and Innovation Requires Knowledge from Multiple Perspectives

"..the capabilities needed for collaboration cannot be patched on but must be designed in from the start." (Grosz, 1994)

Ethics too!

8

Locations of Learning Ways to Perform Responsible Computing Research

- Research advisor
- In courses you take
- In courses you serve as a TA
- From courses, scholarship, and scholars in social sciences and philosophy:
  - Learn new ways of thinking and asking,
  - Learn to collaborate in research across disciplinary boundaries.

Do not need to become an expert in all fields!!

9

Embedded EthiCS @ Harvard

design thinking technical skills problem-solving ethical reasoning

10

Motivation for "continuously and ubiquitously" Considering Ethical and Societal Challenges in CS Courses

- Course on "Intelligent Systems: Design and Ethical Challenges" (CS 108)
- In-class activity:

Imagine that you are a member of the group at some social media company that has as its main responsibility making revenue from advertisements. This group, which we'll dub the "money team", has just signed a major client who is entering the market with an innovative and empirically-validated fitness solution. Based on your knowledge as a user of features of Facebook or some other social networking platform, please do the following:

1. List 3-4 comparative advantages/disadvantages of content-based vs. collaborative filtering.
2. List 3-4 advantages/disadvantages of user-user vs. item-item algorithms
3. List 5 features of a user's profile/data you want your team to consider using in the algorithm that decides to which users to post this ad and where to post it, for content- or collaborative-based recommendations (your choice)

11

Embedded EthiCS @ Harvard: What and Why?

- What: Integrate ethics throughout the computer science curriculum.
- Why: Compounding of modules
  - conveys that ethics is *integral rather than an interruption,*
  - *habituates* students to thinking ethically,
  - gives students the intellectual frameworks and vocabularies to think critically about these issues beyond the classroom, and repeated practice using them, and
  - demonstrates *breadth of applicability* across areas of CS.

Goal: For students to learn ways to design, build and test systems with ethics and potential societal impact in mind.

12

Embed philosophers with expertise in ethics in the teaching CS courses

- Embedded EthiCS TFs: Philosophy Ph.D. students and postdoctoral fellows with strong backgrounds in ethics and proven teaching abilities.
- Key characteristics of ethics modules:
  - Class session(s) on ethical challenge raised by some topic covered in the class,
    - with "active learning" engagement, and
    - an assignment related to the class sessions.
  - Course faculty and TFs are present and engaged.
- Teaching Lab: Embedded EthiCS TFs (graduate fellows and postdocs) and CS postdocs bring complementary expertise to bear on design of module activities, *and they gain expertise as they learn from one another*.

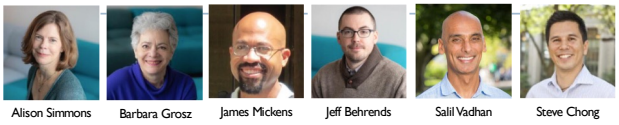
13

AY 2021-2022 Teaching Lab

- 1 Faculty: Jeff Behrends
- 2 Philosophy Postdocs: William Cochran, Jenna Donohue
- 1 Philosophy "Bridge" Postdoc: Trystan Goetze
- 2 CS Postdocs: Arpita Biswas, Subarna Chatterjee
- 4 Philsophy Graduate Fellows: Sophie Gibert, Ellie Lasater-Guttmann, Krupa Patel, Eliza Wells

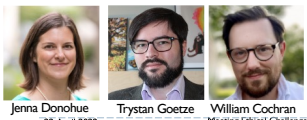
14

Faculty Leaders



Alison Simmons Barbara Grosz James Mickens Jeff Behrends Salil Vadhan Steve Chong

Philosophy Postdocs



Jenna Donohue Trystan Goetze William Cochran

Computer Science Postdocs

Arpita Biswas Subarna Chatterjee

15

Sample of AY 2020-2021 Modules

	Course Title	Module Title
CS 50	Introduction to Computer Science	Democracy and the Digital Public Sphere
CS61	Systems Programming and Machine Organization	Balancing Accessibility and Efficiency: Principles of Distributive Justice in Design
CS109a	Introduction to Data Science	Fairness Ex(tra) Machina
CS143	Computer Networks	5G and Privacy in the Workplace
CS171	Visualization	The Ethics and Politics of Data Visualization
CS182	Artificial Intelligence	Fairness in Modeling Social Networks
CS187	Introduction to Computational Linguistics and Natural Language Processing	The Ethics of Defending Against Neural Fake News
CS222	Algorithms at the End of the Wire	Fair Queuing
CS249r	Tiny Machine Learning	Surveillance, Privacy and Power
CS252r	Advanced Topics in Programming Languages	Interpretability and the Right to an Explanation
CS263	Systems Security	The Ethics of Hacking Back
CS279	Research Topics in Human-Computer Interaction	Ignorance and Algorithms: The Ethics of Unknown Unknowns

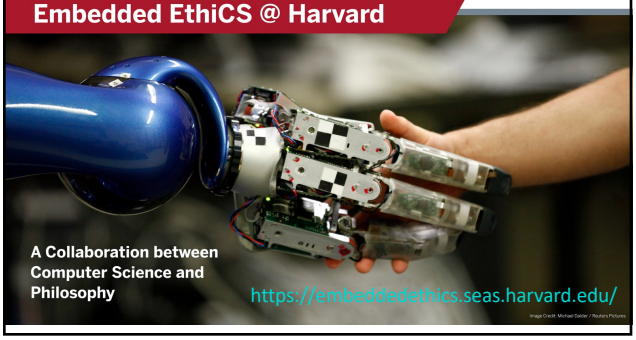
16

Fall 2021 Modules

Course	Course Title	Module Topic
AC 295	Deep Learning for NLP	Gender Bias
CS 109A	Introduction to Data Science	Algorithmic Bias
CS 127	Cryptography	Privacy and Client-Side Scanning
CS 136	Economics and Computation	Mechanism Design
CS 153	Compilers	Do the freedoms of free software come with responsibilities?
CS 165	Data Systems	Data and Privacy
CS 182	Artificial Intelligence	Thinking responsibly about AI systems
CS 187	Introduction to Computational Linguistics and Natural Language Processing	GROVER & Misinformation
CS 263	Systems Security	Ethics of Hacking Back
CS 283	Advanced Computer Vision	Facial recognition, where is the bias (data, model, function)
CS 61	Systems Programming and Machine Organization	Language Encoding and Harms

17

Embedded EthiCS @ Harvard



A Collaboration between Computer Science and Philosophy

https://embeddedethics.seas.harvard.edu/

18

Questions Graduate Students Should Ask

- Identifying your scientific problem(s)
  - Who: adviser and more senior members of research group
  - What: Are ethical concerns an ordinary component of the research environment and group discussions?
- Research team:
  - How: Observation of group (current and recent past)
  - What: How diverse are the research and co-author groups?
- Research Design & Testing:
  - What: What are the research group norms for including diverse groups in design and testing? [as stakeholders? collaborators?]
  - What: IRBs? More than IRBs?

19

Questions Graduate Students Should Ask

- Courses in the graduate degree curriculum
  - How prevalently are ethical concerns raised in courses that are part of the graduate study curriculum?
  - How flexible is the graduate program in supporting relevant courses in social sciences and humanities?
- Courses that might TF:
  - How is ethics included in CS curricula?
  - What experience might I gain in working with scholars in ethics or social scientists in this context?

20

Questions Graduate Students Should Ask

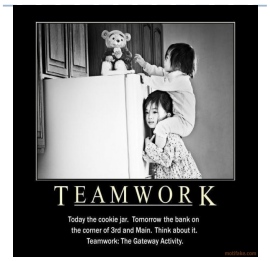
- Publications:
  - Is discussion of potential or ethical impacts included in papers written by the research group and co-authors?
  - Are limitations of research innovations included in papers written by the research group and co-authors?
- Artifacts
  - How are they tested before release?
  - What documentation of intended use or limitations is included in release?

21

https://embeddedethics.seas.harvard.edu/

Thank you!

Questions?



22