From electronic surveillance to hacking, from driverless cars to autonomous weapons, computing technologies raise special ethical questions, and contemporary ethical and political theories can bear on these questions in ways that are sometimes complex and surprising. In this course, we will investigate varied approaches to some of the questions that arise in the ethics of computing technology. The specific topics we will discuss are as follows:

1. Technology and Values
2. Privacy
3. Artificial General Intelligence
4. Bias in Computer Systems
5. Data Justice
6. Moral Attention and the Attention Economy
7. Responsibility and Accountability in the Technology Industry
8. Hacking
9. Plus, several guest sessions to be announced

The class sessions will be discussion-driven and activity-based, with substantial pre-reading materials. Assessment will be through a combination of short-form writing and reflection, as well as a major project.

Course ID: 139163 | **Course Website:** https://canvas.harvard.edu/courses/117639/

Recommended prerequisites: One or more philosophy courses and/or several Embedded EthiCS modules.

Meetings: Tuesdays 3:45 PM to 5:45 PM, Science & Engineering Complex LL2.225 (Allston).

Instructor

Dr. Trystan S. Goetze, Postdoctoral Fellow of Embedded EthiCS

- Pronunciation: /TRISS-tin GETS/
- Any pronouns
- Email: tgoetze@fas.harvard.edu
- Google account (for file sharing only): tgoetze@g.harvard.edu
- **Office Hours**
    - Tuesdays, 1:30-2:30 PM, Science & Engineering Complex, Room 4.441, 150 Western Avenue, Allston
    - Thursdays, 2:00-3:00 PM, Emerson Hall, Room 303, 25 Quincy Street, Cambridge
- Website: https://www.trystangoetze.ca/

Learning Goals

Students will:

- Become familiar with several major areas of the computer ethics literature, and the positions defended therein.
- Practise verbal and written forms of articulating their own positions on computer ethics issues, supported by reasoned arguments.
- Improve their research skills across several disciplines in responsible computing.
- Apply their learning in a major project developed throughout the semester.

Course Schedule

The course meets weekly, in person, for 2 hours. The sessions will be driven by discussion of the assigned readings, your weekly written reflections on them, and activities designed by the instructor.

The course topics are provisionally scheduled in the order they appear. However, several sessions will be taught by a guest instructor. They are likely to take place in February or March. Announcements about who will be teaching the session, the topic, and the readings will be made at least one week in advance.

In addition, students may ask for adjustments to the schedule if they believe it would be helpful to their final projects to study a particular topic earlier in the term. Such adjustments must be requested before Week 6 (when the project proposal is due). Final decisions on the schedule will be made by the instructor after seeking consensus from all enrollees.

There is no textbook. Readings are available through the course website and the libraries.

1 – Technology and Values

- Winner, "Do Artifacts Have Politics?" (16 pp)
- Moor, "What is Computer Ethics?" (10 pp)
- Optional further reading:
  - Bright, "Du Bois' democratic defence of the value free ideal"
  - Douglas, *Science, Freedom, and Democracy*

2 – Privacy

- Garfinkel, "Privacy in a Database Nation" (selections from *Database Nation: The Death of Privacy in the 21st Century*), in *Computers, Ethics, and Society*, eds. David Ermann and Michele Shauf, 3rd ed. (15 pp)
- Nissenbaum, "Privacy as Contextual Integrity" (40 pp)
- Fried, "Privacy" (19 pp)
- Optional further reading:
  - Warren & Brandeis, "The Right to Privacy"
  - Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*
  - Schoeman (ed.), *Philosophical dimensions of privacy: an anthology*

3 – Artificial General Intelligence

- Coeckelbergh, "Robot rights? Towards a social-relational justification of moral consideration" (12 pp)
- Gunkel, "The other question: can and should robots have rights?" (11 pp)
- Vold & Harris, "How Does Artificial Intelligence Pose an Existential Risk? (26 pp)
- Optional further reading:
  - Bostrom, *Superintelligence: Paths, Dangers, Strategies*
  - Dreyfus, *What Computers Still Can't Do*

4 – Moral Attention and the Attention Economy

Guest session: Dr. Jenna Donohue.

- Murdoch, "The Idea of Perfection," in *The Sovereignty of Good*, pp. 16–17, 33–44
- Williams, *Stand Out of Our Light*, chapter 12 (19 pp)
- Kozyreva et al., "Critical Ignoring as a Core Competence for Digital Citizens" (6 pp)
- Highly recommended reading:
  - Williams, *Stand Out of Our Light*, chapters 5–9 (59 pp)
- Optional further reading:
  - The rest of Murdoch, "The Idea of Perfection"
  - Panizza, "Attention," in *The Murdochian Mind*.
  - The rest of Williams, *Stand Out Of Our Light: Freedom and Resistance in the Attention Economy*
  - Wu, *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*

5 – Bias in Computer Systems

- Friedman and Nissenbaum, "Bias in Computer Systems" (17 pp)
- Noble, "Missed Connections: What Search Engines Say About Women" (4 pp)
- Benjamin, "Assessing Risk, Automating Racism" (2 pp)
- Buolamwini & Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification" (12 pp)
- Optional further reading:
    - Obermeyer et al., "Dissecting racial bias in an algorithm used to manage the health of populations"
    - Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*
    - Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code*

6 – Data Justice

- Dencik & Sanchez-Monedero, "Data Justice" (11 pp)
- D'Ignazio & Klein, "The Power Chapter," *Data Feminism* (32 pp)
- Lewis, "Guidelines for Indigenous-Centred AI Design v.1" (3 pp.)
- Optional further reading:
    - D'Ignazio & Klein, *Data Feminism*
    - Dencik, Rentz, Redden, and Treré, *Data Justice*
    - Lewis et al., *Indigenous Protocol and Artificial Intelligence Position Paper*

7 – Autonomous Vehicles

Guest session: Dr. Jeffrey Behrends.

- Nyholm and Smids, "The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem?" (15 pp)
- Behrends and Basl, "Trolleys and Autonomous Vehicles: New Foundations for the Ethics of Machine Learning" (22 pp)
- Behrends and Basl, "Appendix: Varieties of Trolley Pessimism" (8 pp)

8 – Responsibility and Accountability in the Tech Industry

- Ladd, "Computers and Moral Responsibility: A Framework for an Ethical Analysis" (19 pp)
- Nissenbaum, "Computing and Accountability" (9 pp)
- Sparrow, "Killer Robots" (16 pp)
- Optional further reading:
    - Gotterbarn, "Informatics and Professional Responsibility"
    - Johnson, "Computer systems: Moral entities but not moral agents"

9 – Hacking

- Spafford, "Are Hacker Break-Ins Ethical?" (13 pp)

- Radziwill et al., "The ethics of hacking: should it be taught?" (7 pp)
- Lin, "Ethics of hacking back: Six arguments from armed conflict to zombies" (24 pp)
- Ienca & Haselager, "Hacking the brain: brain–computer interfacing technology and the ethics of neurosecurity" (12 pp)
- Optional further reading:
    - Gillespie, *Cybercrime: key issues and debates*
    - Einar Himma (ed.), *Internet Security: Hacking, Counterhacking, and Society*

10 – Technology, Flourishing, and Friendship

Guest Session: <u>Dr. William Cochran</u>.

- Vallor, *Technology and the Virtues*, selections (25 pp)
- Sharp, "The obstacles against reaching the highest level of Aristotelian friendship online" (9 pp)
- Elder, "Excellent online friendships: an Aristotelian defense of social media" (10 pp)
- Optional further reading:
    - Aristotle, *Nicomachean Ethics*, Books VIII & IX

11 – Gamification

- Nguyen, "Agency as Art," in *Games: Agency as Art* (26 pp)
- Nguyen, "Gamification and Value Capture," in *Games: Agency as Art* (27 pp)
- Nguyen, "How Twitter Gamifies Communication" (27 pp)
- Optional further reading:
    - Nguyen, *Games: Agency as Art*
    - Gabrielle, "Gamified life"
    - Robson et al. "Is it all a game? Understanding the principles of gamification"
    - McGonigal, *Reality is Broken: Why games make us better and how they can change the world*

12 – Projects Showcase

The last class session will be a special event, to be held in the West Atrium of the Science & Engineering Complex. You will share your projects in the form of posters, demos, exhibitions, art installations, or whatever best suits your project. Faculty and students will be invited to attend.

Assessment

Your grade will be based on the following assignments.

Questions for Consideration (1% × 10 = **10%**)

Every week before class (except the first and last weeks), send the instructor 3–5 questions about the week's readings. Good kinds of question include:

- Clarification questions about claims made, arguments presented, or background knowledge assumed in the reading.
- Questions about the implications of the authors' arguments or conclusions.
- Questions about potential connections between different readings, including those from previous weeks.
- Questions about other students' personal experiences with the technologies under discussion.
- Questions about applications of the authors' arguments.

The main purpose of this assignment is to seed our in-class discussion.

Weekly Blog (2% × 10 = **20%**)

Every week after class, post on your blog. You can skip two without penalty. Write at least 250 words (about 1 page of 12-point double-spaced text), reflecting on the course material; it need not be about the current week's readings or discussion, but it should substantially connect with the course themes. One model for a good blog post would be to briefly summarize one point from a reading or the in-class discussion, and comment on it by raising an objection or making an application to a case. Other formats such as personal reflections on your own experiences, commentary on recent news stories, or even flash fiction, may also make for good blog posts.

The purpose of this assignment is: (1) to have you practise writing about the class material regularly, (2) to improve your philosophical thinking skills, and (3) to give you ideas to start from for your final project.

Major Project (5% + 10% + 15% + 25% + 15% = **70%**)

Over the course of the semester, you will work on a project of interest to you in computer ethics, culminating in a poster presentation at a mini-conference to be held in the Science and Engineering Complex. The project will be broken up into several deliverables as follows.

D1. Problem Identification (5%)

Using the course readings and your own research in scholarly and media sources, identify and describe three specific computer ethics problems. Due Week 4.

D2. Project Proposal (10%)

Choose one of the problems you identified in the previous deliverable. Provide a description of the problem and how your project will address it. The proposed project could take many forms, from a research paper to a video essay to a board game to an artwork. The only restriction is that projects cannot involve research on human participants (it would take too long to clear ethics review). Due Week 6.

The default format for the final project is a research paper of 2,500–3,000 words (5,000–7,000 words for graduate students), and a 36 in × 48 in research poster.

If you wish, you may propose a project in a different format. In this case, your proposal must explain how your project will substantively explore the philosophical themes of this course,

and how the work involved is roughly equivalent to the research paper and poster of the default option. You must also propose a rubric for evaluating the project.

D3. Draft Project Materials (15%)

Develop your project for submission. It need not be complete. A draft paper and poster, an artist statement and sketches, or an electronic prototype of a board game are potential submissions at this stage. In a short (maximum 1 page) cover sheet, you should describe the project, how you intend to present it, and identify areas where more work needs to be done and where you would like specific feedback. Due Week 10.

D4a & D4b. Project Presentation (20 + 5 = 25%)

Present your work at the course mini-conference in Week 12. The default format is a conference poster, which you will spend at least an hour explaining and discussing with attendees. Alternative formats might include art installations, technical demonstrations, or small museum exhibits.

You are advised to plan to arrive at the SEC an hour early and to leave an hour late, to ensure sufficient time to set up, complete your peer evaluations, and clean up. The instructor can offer some limited storage space in the Embedded EthiCS suite (SEC 4.440–441), if necessary.

You will spend some time at the mini-conference circulating amongst your peers and viewing their presentations. You will be assigned two peer presentations to grade. The instructor will use these peer evaluations in conjunction with their own evaluation of presentations in order to reach a final grade. These evaluations will be provided to you to use in revising your project materials for final submission. 5% of your final grade will come from your engagement with the peer evaluation exercise.

Take notes on your conversations with peers and guests. You will use them to revise your project for final submission.

D5. Final Submission (15%)

Submit your final project. Due by the end of day May 3, 2023.

Include a brief (no more than 1 page) reflection on the feedback you received from attendees of the mini-conference, formally and informally, and how you addressed these comments in your final submission.

Course Policies

The following policies govern this course, in addition to the policies of Harvard College.

Enrolment

Undergraduate students at Harvard University in any concentration, school, or department may apply to enroll by emailing the instructor at tgoetze@fas.harvard.edu. Applications should briefly explain the prospective student's interest in the subject and make note of any prior coursework, modules, extracurriculars, or other experiences that have prepared them for an upper-division computer ethics course. A previous course in philosophy and/or multiple Embedded EthiCS modules are recommended prior experience, but are not required. Decisions on enrolment will be made partly to ensure a diverse range of backgrounds are represented in the class.

Graduate students should also apply to enrol, as described above. If accepted, the student will meet with the instructor to discuss supplemental coursework in recognition of their advanced academic status.

The Harvard College Honor Code

Members of the Harvard College community commit themselves to producing academic work of integrity—that is, work that adheres to the scholarly and intellectual standards of accurate attribution of sources, appropriate collection and use of data, and transparent acknowledgement of the contribution of others to their ideas, discoveries, interpretations, and conclusions. Cheating on exams or problem sets, plagiarizing or misrepresenting the ideas or language of someone else as one's own, falsifying data, or any other instance of academic dishonesty violates the standards of our community, as well as the standards of the wider world of learning and affairs.

Additional information on Harvard College's academic integrity procedures: https://oaisc.fas.harvard.edu/materials

Information on proper use of sources and avoiding plagiarism: https://usingsources.fas.harvard.edu/avoiding-plagiarism

Health and Wellness

If you are ill, and especially if you test positive for COVID-19, please do not come to class. Remote attendance may be possible if you feel well enough to participate, but please make the best decision for your health. You are encouraged to wear a face covering while in class (especially an N95 respirator), but this is not a requirement to attend in-person.

For information on late or missed assignments due to illness, see the Late Submissions & Extensions policy, below.

Collaboration

You are encouraged to consult with one another on the choice of paper topics, and you may also share library resources. You may find it useful to discuss your chosen topic with your peers, particularly if you are working on the same topic as someone else, but you should ensure that the written paper you submit for evaluation is the result of your own research and reflects your own approach to the topic.

The use of AI-powered writing tools, such as GPT-3 and ChatGPT, is prohibited, except where specifically approved by the instructor.

Accessibility

Harvard University values inclusive excellence and providing equal educational opportunities for all students. Our goal is to remove barriers for disabled students related to inaccessible elements of instruction or design in this course. If reasonable accommodations are necessary to provide access, please contact the Disability Access Office (DAO). Accommodations do not alter fundamental requirements of the course and are not retroactive. Students should request accommodations as early as possible, since they may take time to implement. Students should notify DAO at any time during the semester if adjustments to their communicated accommodation plan are needed.

Attendance

Attendance is expected, but won't be enforced or rewarded beyond its effects on your learning.

If you are feeling ill, please do not come to class. If you feel well enough to participate but are testing positive for COVID-19 or otherwise have reason to believe that you are contagious, please contact the instructor to arrange remote participation via Zoom. But please make your health your first priority.

Late Submissions & Extensions

The following is the FAS policy on late assignments and extensions:

Instructors have the authority to grant undergraduates an extension of time for medical reasons and other special circumstances up to the end of the Examination Period. Ordinarily, students requesting an extension of time to complete course work must have received the consent of the instructor before the final examination or before the final meeting of a course in which there is no final examination.

In deciding the length of an extension granted for medical reasons, the head of the course should apply the principle used by the Administrative Board when it votes extensions beyond the Examination Period: Extensions are granted for a period commensurate with the time missed during an illness, without penalty. Questions about an appropriate extension for an individual undergraduate may be addressed to the student's Allston Burr Resident Dean or Resident Dean of First-Year Students.

If a student submits work after a grade has been filed with the Registrar but before the end of the Examination Period, and if acceptance of that late work results in a grade change, the instructor should submit a grade change at my.harvard.edu. (See Changes in Grades.)

An extension of time to complete course work beyond the end of the Examination Period can be granted to an undergraduate only by vote of the Administrative Board of Harvard College and only in exceptional circumstances. Instructors may not accept work from an undergraduate after the end of the Examination Period without the explicit authorization of the Administrative Board.

Undergraduates cannot receive a grade of incomplete (INC).

In addition, the following policies apply to this course's assignments, except where significant extenuating circumstances arise:

- Questions for Consideration: late submission will not be accepted.
- Weekly Blog: late submission will not be accepted.
- D4a. Project Presentation: because of the nature of the presentation, late submission is significantly more complicated, raising the bar for extenuating circumstances.
- D4b. Peer Evaluation: Late submission will not be accepted.

About Embedded EthiCS

This course is offered in collaboration with the Embedded EthiCS program at Harvard. Embedded EthiCS is an interdisciplinary partnership between the Departments of Philosophy and Computer Science. In the Embedded EthiCS Teaching Lab, led by Dr. Jeff Behrends, graduate students and postdocs design ethics lessons that are taught in computer science courses at Harvard. Our pedagogical goals are to teach students to:

1. Identify social and ethical issues in computing
2. Reason through those social and ethical issues as they relate to technical work in computing
3. Communicate reasoned positions on those social and ethical issues to peers and publics
4. Design ethically and socially responsible systems

The program was founded in 2017 by Professor Alison Simmons (Philosophy) and Professor Barbara Grosz (Computer Science), in response to student demand for ethics content in the computer science curriculum. The program now delivers about a dozen modules every semester, accounting for most of the computer sciences courses offered at any given time.

For more information about Embedded EthiCS, visit our website:

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

Support Resources

Please consult the following if you require support in your studies.

Academic Support

Academic Resource Center

The Academic Resource Center (ARC) works with College students to help them bring out their academic best. Through one-on-one sessions accountability groups and hours, and group workshops, academic coaches introduce students to techniques for managing time, reading strategically, studying effectively, building better habits, and much more. They also assist students with prioritizing goals, creating accountability structures, customizing their learning environment, connecting with instructional staff, and finding the resources they need.

Disability Access Office

The Disability Access Office (DAO) works with Harvard College students with disabilities to ensure equal access to all aspects of college life at Harvard including academic accommodations. To make a request for reasonable accommodations in a course, undergraduate students should contact the Disability Access Office (DAO) and fill out a registration form to request accommodations and supply supporting documentation.

Harvard College Writing Center

The Writing Center offers one-on-one consultations to Harvard undergraduates working on any writing assignment for any course. Writing tutors can provide help with argument, structure, and clarity at any stage of your writing process. They can also provide general help with strategies for drafting and revising academic papers.

Peer Tutoring

Peer tutoring provides students with an opportunity to learn from the diversity of their peers and to engage with one another outside the classroom. When students need help with a specific course, the ARC connects them to a network of trained peer tutors who meet with students individually or in small groups. Peer tutors support students' learning by reviewing materials from class and working through critical concepts and areas of confusion.

Advising Network

Students are encouraged to speak with their academic advisers throughout the semester for additional support. All members of the advising network can be found in my.harvard in the "Advising" tab. Advisers can offer you support in managing your time, balancing your academics and extra-curricular, directing you to academic support resources, and helping you think about how this course fits in with your overall academic goals.

Academic and Concentration Advisers

Academic and Concentration Advisers work with students to help navigate Harvard's resources and explore various curricular and co-curricular opportunities, particularly in relation to how they may help guide your academic journey. First-year students are assigned

a first-year academic adviser (which can also be your proctor), sophomores are assigned a sophomore adviser, and students who have declared their concentrations are able to meet with concentration advisers within their academic department.

Director of Undergraduate Studies (DUS) and Assistant/Associate Director of Undergraduate Studies (ADUS)

Each department or program has a DUS that guides students through the requirements for concentrating or getting a secondary in their field of study. They are typically faculty members who also teach in the department. In addition to the DUS, some departments also have an ADUS. The DUS and ADUS split responsibilities in supporting undergraduates in their department and answering many of your questions about concentration requirements, course sequences, and course-related questions.

Peer Advising Fellow (PAF)

Peer Advising Fellows (PAFs) are sophomore, junior, and senior students at the College who have been specially selected and trained to offer advice and assistance throughout students' first year at Harvard. They offer useful guidance on how to balance curricular and extracurricular choices, how to master the challenges of transitioning to college life, and how to take advantage of the many available opportunities.

[Resident Deans of First-Years (RDFs)](#) / [Allston Burr Resident Deans (ABRDs)](#)

RDFs and ABRDs work with you in your Yard or House to provide academic assistance and personal support. RDFs and ABRDs can also help you think about academic opportunities (research, study abroad, the choice of a concentration, etc.) for which your experiences have prepared or inspired you.

Mental Health and Wellbeing

In addition to academic support and your advising network, support for mental health and wellbeing is also available to all Harvard College students.

[Center for Wellness and Health Promotion (CWHP)](#)

The [Center for Wellness and Health Promotion (CWHP)](#) cultivates individuals and collective wellbeing through holistic educational experiences. The CWHP focuses on many health and wellbeing topics including alcohol and other substances, mindfulness and meditation, sexual health, and sleep.

Counseling and Mental Health Services (CAMHS)

CAMHS is Harvard's Counseling and Mental Health Service at HUHS, which seeks to work collaboratively with students and the university to support students who are experiencing some measure of distress in their lives. Please go to https://camhs.huhs.harvard.edu/our-mission to learn about the mission and goals of CAMHS and how to utilize its services and resources. Mental Health Urgent Care and consultation are always available through CAMHS at 617-495-2042 during business hours and 617-495-5711 evenings, nights and weekends.

[Harvard University Health Services](#) (HUHS)

HUHS works to identify and respond to the health care needs for students and faculty within the Harvard community. Students can [book appointments with HUHS](#) to meet with a physician to address any health-related issues they may encounter.

Peer Student Wellness Groups

There are also a variety of student wellness groups that support students through a variety of concerns, including mental health, nutrition, relationships, LGBTQIA+ identities, sexual health, drugs, alcohol, and more. You can find more information about these groups in the [Guide to Student Wellness Groups](#)

More Information

For more information about these and other academic resources, please feel free to contact the [Office of Undergraduate Education.](#)