Data Ethics (PHIL 174)
Spring 2025

David Danks
ddanks@ucsd.edu

Lecture: 11:00 – 12:20 pm, Tuesdays & Thursdays, Mosaic 204
Course TA: Joe O'Brien <j3obrien@ucsd.edu>
TA office hours: Tuesdays, 3:00-4:00 pm, RWAC/A&H 0436
Discussion sections: There will not be any discussion sections in this iteration of the course
Danks office hours: Wednesdays, 10:00-11:00 am, either in-person (RWAC/A&H 0462) or Zoom
(link on Canvas). Please use Canvas to sign-up for a timeslot (or to check about drop-in)
Course materials: Everything will be on the course Canvas site (http://canvas.ucsd.edu)
Course communication: All course announcements will be posted on Canvas. If you need to
contact Prof. Danks or the TA, then you *must* use your UCSD email address and include 'PHIL
174' in the subject line. Replies can be expected within 24-48 hours. Please note that grade
disputes will be discussed and resolved in-person (or via Zoom), not via email.

Learning objectives

By the end of this course, students will be able to:
- Describe the many ways that ethical issues arise throughout the lifecycle of a data science effort
- Generate appropriate ethical questions for a given data science effort
- Work individually or collaboratively to develop more ethical & responsible data science projects

Course requirements

Attendance: This course will be exploring ethical issues around data and AI in multidisciplinary
ways that will likely take you outside of your intellectual comfort zone. So although attendance during
lectures is *not* required, I strongly urge you to consider attending them. There will be opportunities for
questions and discussion in every lecture, which can play a helpful role in building your understanding
of key concepts. I realize that you may sometimes need to miss lecture, so podcasts will be provided,
but there will *not* be a synchronous Zoom option.

Assignments: The core assignments for the course will be project-based: (i) a multi-part project
(in group, if you want) during the quarter; and (ii) a unified individual project at the end of the
quarter. The full set of assignments (and point values) are:
- <u>In-quarter projects</u> (2×22% = 44%):
  1. A "pitch deck" and video presenting a design of an ethical data collection & analysis effort
  2. A "user's guide" explaining how to ethically deploy, use, and update a data science effort
- <u>Reflections</u> (4×3% = 12%): You will be asked to reflect on aspects of data science in your
  everyday life as an individual and as a UCSD student (1-2 paragraphs each).
- <u>Final project</u> (44%): You will analyze the ethics of a data science effort across its lifecycle. This
  final project will require you to individually do each of the tasks from the different parts of the in-
  quarter project, and can take the form of a video presentation, whitepaper/essay, or structured
  webpage (your choice, but no mix-and-match). You will need to provide a (brief) description of
  your focus/topic during week 9 so we can give feedback to improve the final project.

For the in-quarter project, we will provide a set of (broad) topics from which you can choose. You are permitted to work with up to two other people (i.e., you can have a 1-, 2-, or 3-person group). All in-quarter projects will be graded to the same standard, *regardless of how many people contributed*.

For the individual project, you will select your own topic (subject to our feedback), and it *must* be different than your in-quarter projects.

Extensions/late submission: Due dates & times for all assignments are listed below. **Please plan ahead, as no extensions will be granted and no late submissions will be accepted, except in *truly* exceptional circumstances** (e.g., Canvas going offline or documented medical emergency). If you think that an exception is warranted, then you must email Prof. Danks (and cc the TA) as soon as feasible. All relevant materials will be provided in a timely manner so that you can complete assignments significantly before their due dates, if you are concerned about finishing on time.

Final grade: Your final grade will be determined by the following scale:

- A+: 97.00 – 100.00
- A: 93.00 – 96.99
- A-: 90.00 – 92.99
- B+: 87.00 – 89.99
- B: 83.00 – 86.99
- B-: 80.00 – 82.99
- C+: 77.00 – 79.99
- C: 73.00 – 76.99
- C-: 70.00 – 72.99
- D+: 67.00 – 69.99
- D: 63.00 – 66.99
- D-: 60.00 – 62.99

It is possible that the final grading scale will be shifted downwards (i.e., so people's letter grades might improve), but you should not assume that any changes will happen.

Plagiarism

The representation of other people's ideas as though they were your own is plagiarism, and a violation of academic integrity. I have an absolutely zero tolerance policy about plagiarism. Plagiarism is *not* restricted to verbatim copying of other people's words. Basically, you should cite anything that (a) you got from someone else; and (b) a reasonable person would not know ahead of time. For proper citation, you should cite any ideas and quotes that are due to other people. (Your choice about citation format, but please be consistent.)

I realize that the vast majority of you will never consider cheating. However, a few of you may (for a variety of reasons) be tempted to plagiarize others' work. Do not take chances with plagiarism: **if you are uncertain whether you are doing something acceptable, please just ask**. We are happy to answer questions about whether something constitutes plagiarism. And if you are ever tempted to present someone else's ideas as your own, then please come talk to me so that we can address the underlying challenges that are pushing you to that point.

Generative AI policy: In this course, you are permitted to use generative AI systems. If you use generative AI during your assignment, then you should include a brief statement (in a footnote, or an extra slide, or…) to indicate how you used it (e.g., "I used ChatGPT to help brainstorm potential ethical issues."). You should also be aware that these systems can sometimes reproduce exact text that was in their training corpus. Such a submission will be treated as a case of plagiarism, as you are presenting someone else's words as your own. Finally, I reserve the right to request an in-person conversation about an assignment, if I have concerns about your comprehension of the material.

Statement of support for your well-being

We as a community have to support one another. If you, or anyone you know, experiences academic stresses, difficult life events, or feelings of anxiety or depression, then we strongly encourage you to seek support. Take care of yourself. Do your best to maintain a healthy lifestyle this quarter by eating well, exercising, getting enough sleep, and taking some time to relax. This will help you achieve your goals and cope with stress. More generally, all of us benefit from support during times of struggle. There are many helpful resources available on campus and an important part of the college

experience is learning how to ask for help. Asking for support sooner rather than later is almost always helpful. If you need it, please consider reaching out to a friend, faculty, or family member you trust for help getting connected to support that can help.

Schedule
(subject to revision)

Date	Lecture Topic	Reading/Listening
4/1	Course overview Lifecycle of a data science effort	
4/3	What is "ethics"?	National Academies, *Fostering responsible computing research*, Ch. 2
4/7	*Reflection #1 due at 6:00 pm*	
4/8	Neutrality thesis	Danks, "Digital ethics as translational ethics"
4/10	Algorithmic society (Zoom)	Zuboff, "You are the object…" Tufekci, "Algorithmic harms beyond…"
4/14	*Reflection #2 due at 6:00 pm*	

Data Collection & Analysis

Date	Lecture Topic	Reading/Listening
4/15	Overview of issues in data collection	Hand, "Aspects of data ethics…"
4/17	Ethical challenges in measurements	Butler, et al. "Beyond measure" (pp. 1-10)
4/22	Core concepts of privacy & consent	SEP entry on "Privacy" (Section 3)
4/24	Privacy & consent in practice (Zoom)	Acquisti, et al. "Privacy and human behavior…"
4/28	*Reflection #3 due at 6:00 pm*	
4/29	Core concepts of bias & fairness	Fazelpour & Danks, "Algorithmic bias"
5/1	Practical aspects of algorithmic bias	Margaret Mitchell podcast: *Towards Data Science*
5/6	Explainable models to the rescue?	Lipton, "Mythos of model interpretability"
5/8	Data & AI governance	
5/12	*In-quarter project #1 due at 6:00 pm*	

Deployment, Use, & Revision

Date	Lecture Topic	Reading/Listening
5/13	Overview of issues in use	Elish, "Moral crumple zones…"
5/15	Data subject vs. Algorithm subject	Parasidis, et al. "A Belmont Report for health data"
5/20	Accountability during deployment	Raji, et al. "Closing the AI accountability gap"
5/22	Algorithmic justice vs. fairness	Sloane, "Inequality is the name of the game…"
5/27	Human-computer interaction	Schneider, et al., "Digital nudging…"
5/30	*In-quarter project #2 due at 6:00 pm*	

Broader issues

Date	Lecture Topic	Reading/Listening
5/29	Data colonialism & sovereignty	Carroll, et al. "Indigenous data governance…"
6/2	*Reflection #4 due at 6:00 pm*	
6/3	Professional obligations	ACM Code of Ethics & Professional Conduct
6/5	Data + Human = ?	Guszcza & Schwartz "Superminds …"
6/11	*Final project due at 6:00 pm*	