

DSCC 485: Introduction to Privacy, Fairness, and Ethical Considerations for Healthcare Data Science

Fall 2026

Instructor: Rebecca Sanaeikia
PhD Candidate, Department of Philosophy
University of Rochester

Course Meetings: Wednesdays, 6:15–8:55 p.m. (Eastern Time)

Location: on Zoom : <https://rochester.zoom.us/j/99404091561>
Meeting ID: 994 0409 1561

Office Hours: By appointment (Virtual)

Email: becca.sanaeikia@rochester.edu

Course Description

This course will provide an overview of the regulatory and ethical dimensions of doing data science within the United States healthcare system. The course will cover the philosophical and technical considerations that ensure fairness, privacy, consent, and explainability in healthcare. The primary goal of this course is to train students to think critically about these considerations when choosing data science and machine learning techniques to apply to healthcare data. Topics for the course include: (1) the regulatory environment around clinical data capture and research with human subject data, (2) privacy and consent to collection, (3) bias in algorithms, (4) explainability in the context of healthcare, and (5) disagreement in values within algorithmic medicine. The topics included in this course are complementary with the technical skills students will develop throughout the program.

Course Format

DSCC 485 is a fully online course with both synchronous and asynchronous activities. The synchronous meetings will be discussion-based and focused on the assigned readings for that given week. Asynchronous activities will include online discussions to guide content knowledge and synthesis. All meetings will be recorded and posted on the course Blackboard site.

Course Objectives

The overall objective of the course is to increase the student's capacity for: (I) ethical sensitivity, (ii) ethical reasoning, and (iii) ethical motivation with respect to the use of data-driven methods in healthcare delivery and clinical research. By the end of this course, students will be able to think critically, creatively, and ethically about the application of data science methods to healthcare systems. Specifically, students will be able to:

- Engage with the philosophical and technical literature on data ethics.
- Articulate the ethical risks involved in the use of data in health research.
- Describe the major regulatory and ethical principles relevant to research with healthcare data.
- Apply these principles to cases of conflict between the duties of different roles.
- Articulate strategies for mitigating the ethical risks of machine learning in healthcare.
- Articulate the importance of responsible and ethical data science.
- Demonstrate self-efficacy with respect to navigating these risks.

Course Expectations

This class will discuss some controversial issues, and within the class there will be students with a wide variety of moral, political and religious commitments, some of which may be radically opposed to your own. Additionally, many of the questions we discuss will have special personal or professional significance to someone in the class. I therefore expect all students to:

1. Be Charitable – always take the most sympathetic reading of a comment or position.
2. Be Sensitive – be aware that classmates may have relevant personal or professional experience.
3. Be Impersonal – always evaluate positions and arguments: NOT people.
4. Be Humble – always be willing to change your position if provided with sufficient reason.

For more information on how to read and discuss philosophy, I encourage you to [watch this YouTube video](#).

Assignments

Your overall grade is determined by the sum of points you receive in the semester. Below I have listed the maximum possible points for each required assessment piece, as well as two optional bonus opportunities. I use a standard grading scale (A=1000-900, B=899-

800, etc.) to convert the points you receive into a letter grade. “Plus” or “minus” grades are given for being within 25 points of the relevant letter threshold. Detailed explanation of the assignments is listed in the tables below.

Assignment	Points
Class Participation	100
Readings (Perusall Annotations, 10 x 20)	200
Technical Presentation	300
Final Paper	400

Total Points	Letter Grade
1000 – 900	A
899 – 850	A-
849 – 800	B+
799 – 750	B
749 – 700	B-
...	...
< 599	E

Bonus Assignments

In addition to the 1000 required points above, you may earn up to 90 bonus points across the two optional assignments described below.

Bonus Assignment	Points
Discussion Boards (4 parts x 10 pts each)	40
Critical Appraisal Paper (bonus)	50

Class participation

100 Points

Ethical inquiry is best done in discussion with others. Synchronous class discussions and small group activities are an important component of the learning strategy in the course. Your in-class participation is also an important indicator of your capacity to practice ethical data science and philosophy: to understand the concepts, to formulate arguments, and engage in respectful critique of others viewpoints.

Class participation will be assessed by four criteria: (1) timely attendance, (2) thoughtfulness of comments/questions, (3) engagement with other students in a respectful manner, and (4) evidence of interaction with the asynchronous activities for the week in Blackboard. Recognizing varied professional responsibilities and personal commitments, you are only required to attend eleven out of

the fourteen scheduled meetings. Though you may have to miss a synchronous session, you will still be responsible for all course activities and assignments for that week.

Readings

200 Points

Students will complete a collaborative mark-up of the required readings in Perusall, beginning in Week 2 — an excellent way to stay motivated and engage critically with others' viewpoints. There are 13 sessions from Week 2 through the end of the semester, and you must fully interact with the readings from at least ten of them. Each session's readings are scaled to a maximum of 10 points based on how many readings that session includes, and for each individual reading you should leave at least three substantive, engaging comments, for a maximum of 200 points overall.

Technical Presentation

300 Points

You will be expected to make a presentation for the class introducing that week's algorithm or technical study (listed in pink). The goal of these technical presentations is to introduce the algorithm and the ethical considerations that it raises. More concretely, each presentation will explain:

1. **The Problem:** explain the problem that the algorithm/study attempts to solve,
2. **The Approach:** explain the broad features of the approach to solving that problem (e.g. model structure, training data, input data, target variables), and
3. **The Ethics:** explain the main ethical problems or concerns raised by the algorithm/study.

The structure of your presentation is as follows:

- You will be allotted approximately 12 minutes during the class session for the presentation. Students will be responsible for answering questions for an additional 5-10 minutes.
- Students who are unable to attend class due to professional commitment or illness should contact the instructor as soon as possible for an alternate assignment.

Students will submit a final copy of the slide deck and a detailed grading rubric will be available on blackboard.

Final Paper

400 Points

A paper of **max.** 3000 words (10 pages) on one of the following questions:

1. Explain why (or why not) it is ethically permissible to reuse de-identified mental health notes in research without specific consent from the original data subject? Answer with reference to the papers by Evans, Levis et al. and Dickert et al.

2. Explain why (or why not) the use of race as an input feature in the Vaginal Birth After Caesarean Risk Calculator is ethically permissible? Answer with reference to the papers by Vyas et al., and Grobman et al.
3. Should an AI system that makes diagnostic predictions from medical imagery (e.g. MRI, CAT) be required to be “explainable” if they are sufficiently accurate? Answer with reference to the papers by Lipton, London and Mitchell et al.
4. Does the use of a Personalized Patient Preference Predictor to guide substituted judgment for incapacitated patients respect or undermine patient autonomy? Answer with reference to the papers by Earp et al. and Starke.
5. Is it ethically permissible for laypeople to use large language models (e.g., ChatGPT) for self-diagnosis or medical decision-making outside clinical supervision, and if so, what forms of regulatory oversight, if any, should govern this use? In your answer, consider what it would mean to apply community-governance and data-sovereignty principles — such as the CARE Principles for Indigenous Data Governance — to the design of oversight for layperson use of AI in health contexts, and engage with Binns' "Algorithmic Accountability & Public Reason" on where accountability for these systems should sit.
6. Self-directed question – please finalize with the instructor by October 25th.

You will be expected to write in the style of a philosophical essay. In particular, you will clearly state and argue for your position, by clearly explaining a philosophical concept or theory, its application to the question at hand, and consider and reply to 1-2 objections to your argument. You must reference the philosophical papers identified in the question, and while not required, citation of articles not discussed in class is encouraged.

The Paper Assignment: Three Documents, Two Due Dates

- Monday, 9 November – Upload to Blackboard a short draft of your paper. This could be a complete draft or a detailed outline. Late or no submission of the draft will result in the loss of 50 points. Drafts will be returned to you by 24 November.
- Monday, 18 December – Upload to Blackboard (i) the final paper, and (ii) a 1-page self-assessment of your draft which grades the draft according to the paper rubric, and identifies changes that you will make to the draft to improve it before submission.

Grade Scheme and Rubric

Each paper will be graded on a 300 to 0 scale (scaled to the 400 points allotted to this assignment). You will be graded according to the following criteria:

1. Understanding of content.
2. Critical engagement with readings.
3. Quality of argument.
4. Quality of objections and replies.

5. Structure of essay.
6. Quality of written expression.

Bonus Assignment: Discussion Boards

40 Points

At the end of each of the four parts of the course, you will have the option to complete a discussion board post for bonus credit (10 points per part, for a maximum of 40 points across the semester). For each discussion board, you must critically reflect on one particular argument from that part of the course and/or apply it to a real-life example other than the ones discussed in class.

Discussion boards will open on Blackboard at the conclusion of each part of the course and will be due before the first class session of the following part:

- Part 1 (Regulatory and Ethical Basics in Different Domains) – discussion board due before Sep 23
- Part 2 (Data Collection and Privacy) – discussion board due before Oct 14
- Part 3 (Machine Learning, Accuracy and Fairness) – discussion board due before Nov 11
- Part 4 (Explainability, Autonomy and Accountability) – discussion board due before Dec 16

Each discussion board is worth up to 10 bonus points and will be graded on the depth of critical engagement with the selected argument and the aptness and originality of the real-life example chosen.

Bonus Assignment: Critical Appraisal Paper

50 Points

This optional bonus assignment is designed to strengthen your ability to critically evaluate the ethical, methodological, and clinical implications of artificial intelligence and machine learning applications in healthcare. In this mini-paper, you will select a peer-reviewed research article from your own field or area of professional interest that presents a machine learning or artificial intelligence model used in a clinical domain (e.g. oncology, obstetrics, pediatrics). You will then systematically evaluate the study using the ROBUST-ML checklist as a structured framework for assessing bias, transparency, rigor, and clinical utility. If you are unsure if your selected article is appropriate, you are welcome to consult with the instructor in advance.

Al-Zaiti SS, Alghwiri AA, Hu X, et al. A clinician's guide to understanding and critically appraising machine learning studies: a checklist for Ruling Out Bias Using Standard Tools in Machine Learning (ROBUST-ML). *Eur Heart J Digit Health*. 2022;3(2):125-140. Published 2022 Apr 12. doi:10.1093/ehjdh/ztac016

This bonus assignment will be due on 18 December at 11:59pm (the same date as the final paper). You will be responsible for submitting three items: (1) the selected AI/ML peer-reviewed paper, (2) a completed ROBUST-ML checklist, and (3) a mini-paper no longer than 3-4 pages (or approximately 1000 words). A rubric will be available for transparent grading. The mini-paper should synthesize the findings of the ROBUST-ML checklist. Additionally, students should:

- Briefly describe the clinical problem and the role of machine learning in the study.
- Highlight any significant findings on the ROBUST-ML checklist.
- Discuss the clinical utility and ethical considerations such as fairness, transparency, patient safety, and implementation.
- Provide a synthesis of the main limitations and propose specific strategies to improve the model or its evaluation.

Submission Policy

All assignments must be uploaded to Blackboard according to the instructions provided. Late uploads of assignments to Blackboard will lose 10% of the points for that assignment for each day late, calculated from midnight. Extensions or exceptions should be brought to the course faculty as soon as possible to avoid late deductions.

Academic Honesty Policy

Academic honesty is essential to the course, as it ensures a level playing field and safeguards the value of your grade. You can find information about the academic honesty policy at <http://www.rochester.edu/college/honesty/index.html>.

Citation Style and Policy

All assignments should be formatted in either AMA, APA, or Vancouver citation formatting per your preferences. If you prefer a different type of citation style, please consult with the course faculty in advance.

Generative AI Policy

The University provides guidance about Generative AI on the [Office of the Provost website](#). In this course (and program), you are permitted to use AI whenever it is helpful, unless specifically forbidden. However, several caveats apply to all uses of Generative AI in this course:

- By policy, UR students may use Generative AI only at <https://chat.rochester.edu> or <https://chat.qualifiedhealthai.com>. You may not use other Generative AI for assignments without express prior permission.
- You must cite your use of Generative AI directly in the text, in a comment, or some other obvious manner. This citation must include the prompt(s) you provided to the Generative AI. (Ideally, this citation may include the entirety of the dialog between you and the Generative AI.)

Attendance

Although this course is offered in a format to promote maximum scheduling flexibility for students, the UR requires a set number of weekly contact hours. In this course, weekly contact hours are a cumulation of both synchronous and asynchronous content and activities. However, most of the topics in this course are better learned through real-time interactions and discussion. You are expected to make a reasonable effort to attend the synchronous meetings and to inform the instructor of any scheduling conflicts. If the instructor feels that a student's absences become problematic (e.g., as reflected in poor student work), the instructor and student will need to discuss the appropriate course of action.

Students attending synchronous sessions are required to keep their camera on for the duration of the session. This allows for full participation in discussion-based activities and helps the instructor assess engagement as part of the class participation grade. If you have a circumstance that makes this difficult, please contact the instructor in advance to discuss accommodations.

Accessibility and Academic Accommodations

The University of Rochester welcomes applications from students with disabilities. The University employs professional staff committed to assisting students with disabilities in the classroom, residence halls, libraries, and other campus locations. It is a personal decision to disclose a disability and request an accommodation. A decision not to disclose will be respected. Students seeking accommodation for a disability or temporary medical condition must request accommodation through the [UR Office of Disability Resources](#). Arranging student accommodations can take some time, and so students requesting accommodations are encouraged to do so through the UR Office of Disability Resources as soon as possible.

Religious Holidays

If an assignment or exam is scheduled to conflict with your observance of a religious holiday in your tradition, please let me know as soon as possible so we can arrange an alternative date or assignment.

Academic Calendar and Registrar Policies

The [Academic Calendar](#) provides important deadlines for drop/add/withdraw from courses, as well as schedules for final exams, breaks and holidays, and University Commencement. The [University Registrar's website](#) also covers policies for audits, leaves of absence, withdrawal from the University, incompletes, and privacy.

Anti-Discrimination and Harassment

Please see the [Policy Against Discrimination, Harassment, and Discriminatory Employment/Service Practices](#) website for the University's anti-discrimination and harassment policy. You can report an issue on the Office of University Engagement and Enrichment [website](#).

Title IX

All members of the University of Rochester community have the right to learn and work in a safe environment free from all forms of discrimination, harassment, and related retaliation. This includes discrimination and harassment based on protected characteristics, and Title IX-prohibited behaviors, which include sexual assault, dating and domestic violence, and stalking. Students who have been subjected to sexual misconduct have the right to receive academic and [other supportive measures](#), to receive counseling and health services, and to make a report to the university and/or to law enforcement. In addition, [students who are pregnant, have recently experienced childbirth, and/or have medical needs related to childbirth](#) can contact the Title IX office for assistance. All [faculty and teaching assistants are mandatory reporters](#) who must report Title IX-prohibited behavior to the Title IX office. Students can also contact the Title IX Office directly (titleix@rochester.edu or 585-275-1654). For more information, please visit www.rochester.edu/sexualmisconduct.

Course Schedule

Date	Topic	Readings	Media Assignments	Learning Goals
Part 1: Regulatory and Ethical Basics in Different Domains				
Sep 2	Data Science in a Healthcare Context	McCradden et al., "A normative framework for AI as a sociotechnical system in healthcare" Lipworth et al., "Ethics and Epistemology in Big Data Research" (opt) Abràmoff et al., "Pivotal trial of an autonomous AI-based diagnostic system..."	Case study: Google Flu Trends	Identify the goals of data science in healthcare Apply the "intervention ensemble" approach to data science projects
Sep 9	Research Ethics and the Common Rule	National Commission for the Protection of Human Subjects, "The Belmont Report" Meyer M., "There Oughta Be a Law: When Does(n't) the U.S. Common Rule Apply?"	Podcast: What can you do to control how your health and medical data is used Case study: Levine A., " Suicide hotline shares data with for-profit spinoff "	Describe the major domains of value in the Belmont Report Describe the core components of the Common Rule
Sep 16	Learning Health Systems and Quality Improvement	Faden et al., "An ethics framework for a learning health care system..." McCradden et al., "A research ethics framework for the clinical translation of healthcare machine learning" Cribb et al. What does 'quality' add? Towards an ethics of healthcare improvement." Journal of Medical Ethics, 46(2), 118–122. (opt) Wong et al., "External Validation of a Widely Implemented Proprietary Sepsis Prediction Model..." In class video: Dr. Hotspot	Case study podcast: Is it quality improvement	Identify the distinction between quality improvement and research Analyze the application of the Common Rule to learning health systems
Part 2: Data Collection and Privacy				
Sep 23	Clinical Confidentiality and the HIPAA Privacy Rule	Beauchamp and Childress, "Professional-Patient Relationships", pp 283-312	Podcast: " Is healthcare privacy possible if 'all	Analyze the justifications for norms of veracity,

Date	Topic	Readings	Media Assignments	Learning Goals
		Evans, Barbara J., "The HIPAA Privacy Rule at Age 25: Privacy for Equitable AI" Reiser and Singh. HITECH Meets Dobbs: Federal Mandates, State Laws, and the Complexities of Abortion Data Sharing (opt) Levis et al., "Natural language processing of clinical mental health notes may add predictive value..."	data are health data'?" (YouTube/podcast) Podcast: Period Tracking apps	confidentiality and privacy in medicine Evaluate criticisms of HIPAA for secondary re-use of patient data
Sep 30	Broad Consent and Data Registries	Dickert et al., "Reframing Consent for Clinical Research: A Function-Based Approach" Ploug & Holm, "Meta Consent – A Flexible Solution to the Problem of Secondary Use of Health Data" (opt) Barazzetti, G., et al. Broad consent in practice: lessons learned from a hospital-based biobank for prospective research on genomic and medical data” pp 915–924 In class video: House MD	"Your Medical Records Are Being Sold" (12 min) News case: "Do Patients Own Their Health Data?" — PBS NewsHour (8 min)	Articulate the concept of informed consent in healthcare data science Evaluate the consent practices of relevant case studies
Oct 7	Privacy and De-identification	Price, W. N. & Cohen, I. G. "Privacy in the age of medical big data” pp 37–43 Kasperbauer, "Protecting health privacy even when privacy is lost" Kaissis et al., "Secure, privacy-preserving and federated machine learning in medical imaging"	Pre-class Movie: The Immortal Life of Henrietta Lacks	Describe one account of the value of health privacy Identify different strategies for preserving health privacy in datasets
Part 3: Machine Learning, Accuracy and Fairness				
Oct 14	Sources of Bias and Epistemic Injustice	Onofrey, Cross & Choma (2024). "Bias in medical AI: Implications for clinical decision-making." PLOS Digital Health Buchman et al., "Investigating Trust, Expertise, and Epistemic Injustice in Chronic Pain" (opt) Pierson et al., "An algorithmic approach to reducing unexplained pain disparities in underserved populations" Chen et al., "Ethical machine learning in healthcare"	Pre-class video: Synthetic Data Podcast: "Bias in Healthcare AI" — Dr. Xiao Liu, RSM Digital Health Podcast (33 min)	Identify the parts of the ML pipeline where bias can be embedded Describe the concept of epistemic injustice and its application to health data science

Date	Topic	Readings	Media Assignments	Learning Goals
		In class video: "How Biased Data Shapes Healthcare AI" — Tomisin Adebari, TEDx Morgan State (12 min)		
Oct 21	Accuracy and Tradeoffs	Halligan et al., "Disadvantages of using the AUC to assess imaging tests" Ord and Beckstead, "Rationing and Rationality" (opt) Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks" In class video: diagnosing melanoma	—	Apply the standard decision-theoretic approach to translating scores into decisions Apply different ways of settling the trade-offs
Oct 28	Disparate Treatment by Clinical Algorithms	Vyas et al., "Hidden in Plain Sight – Reconsidering the Use of Race Correction in Clinical Algorithms" Sanaeikia, R. (2026). Beyond Two-Step Methodology: The Atomic Approach to Transgender Data Collection in Biobanks. (opt) Grobman et al., "Prediction of vaginal birth after cesarean delivery in term gestations..."	Podcast: Racial Bias in Oxemeters	Understand legal distinction between disparate treatment and disparate impact Discuss arguments for and against different theories of what makes disparate treatment wrongful
Nov 4	Disparate Impact by Clinical Algorithms	Barocas, Hardt & Narayanan, "Relative Notions of Fairness" Ferryman K, "Bounded Justice, Inclusion, and the Hyper/Invisibility of Race in Precision Medicine", pp 27-33 (opt) Obermeyer et al., "Dissecting racial bias in an algorithm used to manage the health of populations"	Optional video: Critical Condition: Health in Black America (2025)	Understand how "neutral" algorithms can produce disparate impact Explain different accounts of wrongfulness of disparate impact
Part 4: Explainability, Autonomy and Accountability				
Nov 11	Large Language Models	Bender et al., "On the dangers of stochastic parrots" Reddy et al. Deception in clinical large language models: an under-recognised safety risk (opt) Thirunavukarasu et al., "Large language models in medicine"	Pre-class Case study: ChatGPT Led to a Man's Near-Fatal Health Crisis, Lawsuit Claims Documentary: Eternal You	Describe the capacities and limits of LLMs in healthcare Describe the ethical risks associated with widespread use of LLMs

Date	Topic	Readings	Media Assignments	Learning Goals
Nov 18	Interpretability	Lipton, "The Mythos of Model Interpretability" London, "Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability" (opt) Mitchell et al., "Model Cards for Model Reporting"	Pre-class video: "Babylon Health & AI Medical Apps — Healthcare Scandal Case Study" (18 min) Case study: "Digital Healthcare: Is it Clinically Effective?" — BBC Newsnight (7 min)	Describe different conceptions of interpretability Articulate why interpretability is important for subject autonomy
Nov 25	Thanksgiving Recess	NO CLASS — Thanksgiving Recess (Nov 25–29); classes resume Nov 30	—	—
Dec 2	Patient Values and Control	Birch et al., "Clinical decisions using AI must consider patient values" Earp et al., "A Personalized Patient Preference Predictor for Substituted Judgments in Healthcare" Starke, "Machine Learning–Based Patient Preference Prediction: A Proof of Concept"	End of life decisions	Understand the problem of value disagreement Describe the welfarist and egalitarian approaches to justice
Dec 9	AI Governance	Haendel et al., "Governing real-world health data as a public utility" Binns, "Algorithmic Accountability & Public Reason" (opt) Lee et al., "WeBuildAI: Participatory Framework for Algorithmic Governance"	Case study: When Big Tech Targets Healthcare (documentary, 51 min) In class Case study: "CARE Principles for Indigenous Data Governance — Be FAIR and CARE" (8 min)	Articulate the concepts of "public reason" and "accountability" Evaluate different strategies for embedding community values in algorithms
Dec 15–17	Reading Period	Last day of classes: Dec 14. Final Exams: Dec 18–23. Final Paper due date to be announced by instructor.	—	—

This course schedule is tentative and may change during the semester in response to extenuating circumstances, such as weather or illness, and/or to ensure successful learning and achievement of course learning outcomes and goals. Although changes to the schedule are not anticipated, if they must occur, they will be noted in Blackboard, and as much advance notice as possible will be given.