PPHA 38830: The Minds Behind the Machine: Artificial Intelligence In Modern Policy

Anita Nikolich

1. **Class Meeting Time & Location**
   The class will meet via Zoom on Wednesdays from 10:50-12:10 PM Central Time.

2. **Instructor Information**
   Virtual Office hours will be conducted via Zoom on Mondays from 12 – 1 PM Central Time and by appointment if students cannot make that time slot.
   Contact information: anikolich@uchicago.edu
   Students will interact with the instructor and classmates on Piazza.

3. **Teaching Assistant Information: TBD**

4. **Course Description**
   This non-technical course introduces important topics in Artificial Intelligence that policy makers must consider. We start with foundational AI concepts such as the role of data, the meaning of algorithms, and ways in which AI is used in decision making. By using case studies on issues such as predictive policing, disinformation propagation, deep fake videos, and facial recognition, students will explore AI concepts of fairness, ethics, privacy, transparency and governance.

5. **Course Goals or Objectives**
   This course will arm students with knowledge about the fundamental aspects of AI such that they can pose critical questions about how to shape future policies around the use of AI in society, the appropriate creation of AI, the ownership of AI, the deployment of AI locally and globally, and the use of AI for social good. Students will be equipped to write position papers on major topics in AI and be able to describe how AI is used in society today.

6. **List of Course Materials**
   The readings will all be available as open source online material. The material is a combination of academic papers, popular press articles, and short videos. They will be posted in Canvas at least 3 weeks ahead of the scheduled class.

7. **Reading/Topic List**
Week 1 – What is Artificial Intelligence?

- Intro to AI Terminology
- How Do Algorithms Work: Supervised Learning, Unsupervised Learning, Reinforcement Learning, Deep Learning
- The Importance of Data to AI
- Initial Framing of AI Issues for Policymakers
- Case Study: COVID Contact Tracing

Week 2 – AI Decision Making: Fairness, Accountability, Transparency, Ethics, Privacy

- AI and Personalization
- Content Moderation and Social Media
- Creating Filter bubbles
- Case Study: Predictive Policing

Week 3 – AI for Government

- Facial Recognition and Criminality
- AI for COVID Monitoring and Public Safety
- Case Study: Clearview AI and Facebook scraping

Week 4 – AI for Corporations

- Surveillance Capitalism
- How does AI Enable Microtargeting and Ad Tech
- Does AI Exacerbate Inequality? A look at FICO Scores
- Case Study: Target and Data Brokers

Week 5 - AI and Disinformation

- Brief History of Information Operations
- Deep Fakes and Cheap Fakes
How Does AI perpetuate Disinformation

Case Study: 2016 Election Disinformation

**Week 6 - Attacking and Manipulating AI**

What is Adversarial AI

Can AI Be Fooled?

Fraud Detection

Case Study: TBD

**Week 7 – What is AI Policy? Legal, Regulatory, Standards and Governance Issues in AI**

Agency Roles in AI: NITRD, JAIC, NIST, NSF, FTC

Striking a balance between regulation and innovation

Social Media and Transparency

Is AI Insurance necessary?

**Week 8 - Is AI a Threat to Democracy?**

Deep Fakes and Elections

Deep Fakes and Stability

Data Sharing

The role of Content Providers such as Twitter, Facebook and Instagram

Case Study: Cambridge Analytica and 2016 elections

Guest Q&A: Geoff Golberg, CEO Social Forensics

**Week 9 – International Aspects of AI Policy**

General Data Protection Regulation (GDPR) Rules (EU)

Cultural and Ethical Differences in Using AI Systems

Case study: China’s Social Credit System
Week 10 – Student Presentations

8. Major Assignments and Assessments
   a. Students will write 2 Position Papers/White Papers of up to 3 pages – one due in Week 4; one due in Week 8 - that present an approach to an AI Policy Issue. Examples: The use of AI in Facial Recognition should not be banned; Facebook AI-driven content should not be moderated or regulated; Predictive Policing by AI should be banned. White papers must explicitly describe why this is an AI-related policy issue.
   b. Students will work on a group project in groups of 4-5 that will be due prior to the class in Week 10. It will be presented in a 10 min presentation to the entire class via Zoom live or as a recording ahead of time (this is encouraged) so that the Class time can be used for Q&A. Examples will be given the first week of class but might include creating a disinformation campaign or outlining a framework for Ethical AI.

9. Grading Standards and Criteria (TBD)
   a. Include the breakdown in how assignments, exams, etc. will be assessed and how a final grade will be determined. Core courses at Harris are generally graded on a curve, while grading polices for electives are more varied.
   b. Include a grading scale.
   c. Describe how the “pass/fail” option works in your class. In general, the spirit of “pass/fail” at Harris is that students earning a “pass” grade have completed the course in full and demonstrated an understanding of the course material. As the instructor, you have discretion to define precisely how this broad standard applies to your course.
      i. An example of a pass/fail policy statement for a syllabus is: “Students who wish to take the course pass/fail rather than for a letter grade must use the Harris Pass/Fail request form (https://harris.uchicago.edu/form/pass-fail) and must meet the Harris deadline, which is generally 9am on the Monday of the 5th week of courses. Students who take the course pass/fail must attend class meetings and turn in all assignments, achieving marks on assignments that are overall commensurate with at least a C- letter grade.”
      ii. Another example is “To earn a passing grade, students taking the course pass/fail must: complete and submit all assignments; take both midterm and final exams; and earn passing grades on each assignment and exam.”
iii. Note that students are generally not permitted to take required, core courses on a pass/fail basis, and some certificate programs limit the courses students can take on a pass/fail basis.

10. Policies

a. Attendance policy: This course is scheduled to meet from 10:45 pm–12:00 pm Central Time on Wednesdays via Zoom. Students are expected to attend this session and are expected to have read the reading assignments and watched any posted lecture content for that week. If the ability to attend class at the scheduled time is disrupted by technology challenges or for reasons outside the student’s control, the student should attend office hours that week. Discussion board posts are due each Monday by 2200 Central Time.

b. Students must submit written assignments through Canvas.

c. Late assignments are not accepted except with prior approval from the Instructor.

d. Required citation style or other requirements for written assignments

e. Students are expected to actively participate in Zoom discussions, whether verbally or by using the “raise hand” or chat feature.

f. Where to direct questions (e.g., TA vs Instructor vs Head TA)

g. Use of Zoom, Panopto, and other on-line resources. Note that the University has developed specific policies and procedures regarding the use of video/audio recordings: these policies are explicitly described in the University’s student manual as well as in the guidelines for instructors available here. Perhaps the key components of this policy are worth spelling out in your syllabus: By attending course sessions, students acknowledge that:

i. They will not: (i) record, share, or disseminate University of Chicago course sessions, videos, transcripts, audio, or chats; (ii) retain such materials after the end of the course; or (iii) use such materials for any purpose other than in connection with participation in the course.

ii. They will not share links to University of Chicago course sessions with any persons not authorized to be in the course session. Sharing course materials with persons authorized to be in the relevant course is permitted. Syllabi, handouts, slides, and other documents may be shared at the discretion of the instructor.

iii. Course recordings, content, and materials may be covered by copyrights held by the University, the instructor, or third parties. Any unauthorized use of such recordings or course materials may violate such copyrights.

iv. Any violation of this policy will be referred to the Area Dean of Students.

11. Resources available to students

a. Instructors may also wish to highlight certain resources provided for students by the University, the Harris School, and/or the instructional team for the course itself.

b. For example, the University has long offered a comprehensive set of student support services (described here), including student health services. And in response to the COVID-19 pandemic and associated disruptions, the University has provided links for students via its “Learning Remotely” website: specific resources are listed here. Specifically, instructors may wish to explicitly remind students about available
counseling services in particular: If you or someone you know is feeling overwhelmed, depressed, and/or in need of support, remote counseling services are available. Student Counseling Service (SCS) urges you to attend to your mental wellbeing and to reach out to them for support during these challenging times. All SCS services are covered by the Student Life Fee, and there is no additional cost for students to access their services. See https://wellness.uchicago.edu/mental-health/student-counseling-service-spring-quarter-faq/. Students seeking new services/resources can call 773.702.9800 during business hours (Monday–Friday 8:30 a.m.–5 p.m.) and ask to speak with a clinician. **Students needing urgent mental health care can speak with clinicians over the phone 24/7 by calling the SCS at 773.702.3625.**

c. Similarly, the Harris School itself provides both academic and non-academic support services for students. These resources are described (and links provided) via the Canvas site *Harris Student Handbook*, which all Harris instructors can access.

In addition to providing the above information in their syllabi, instructors are encouraged to address general expectations across other domains, referring to the University’s *Student Manual* and the Harris School’s own policies and regulations as needed (available in Canvas). Specifically, instructors should share both University and Harris School-specific guidelines and principles in at least three key domains: academic integrity; students with disabilities; and diversity and inclusion. Please consider including the material and text below, suitably modified, revised, and/or expanded to align with your own course plans.

**Academic Integrity**

All University of Chicago students are expected to uphold the highest standards of academic Integrity and honesty. Among other things, this means that students shall not represent another’s work as their own, use un-allowed materials during exams, or otherwise gain unfair academic advantage.

- The University’s policies regarding academic integrity and dishonesty are described [here](#). It is worth explicitly stating the University’s approach here: “It is contrary to justice, academic integrity, and to the spirit of intellectual inquiry to submit another’s statements or ideas as one's own work. To do so is plagiarism or cheating, offenses punishable under the University's disciplinary system. Because these offenses undercut the distinctive moral and intellectual character of the University, we take them very seriously.”

- The Harris School’s policies are available in the *Harris Student Handbook* Canvas site.
  - The *Academic Honesty and Plagiarism* section expresses the main principles.
  - Detailed guidelines for more specialized student work (e.g., problem sets including computer code) are offered in the section titled *Harris Integrity Policy for Problem Sets Involving Code*.  

o Harris’s specific procedures for handling suspected violations of these policies are available in the section *Harris Procedures for Allegations of Plagiarism, Cheating, and Academic Dishonesty* and are also re-produced as an Appendix to this document.

o All students suspected of academic dishonesty will be reported to the Harris Dean of Students for investigation and adjudication. The disciplinary process can result in sanctions up to and including suspension or expulsion from the University.

o In addition to disciplinary sanctions, I will impose a grade penalty of “X” for students who have committed academic dishonesty.

o Other penalties for violations could be imposed instead, alone or in combination. For example:
  - Students will receive a grade of 0 on the exam or problem set in question as well as the next lowest problem set grade. This will often result in the student earning a failing grade in the course.
  - The student will receive a grade of 0 on the exam or problem set in question and cannot earn higher than a B- in the course, regardless of their performance on other assignments and exams.
  - At the instructors’ discretion, the student may receive a failing grade for the course regardless of their performance on elements of the course.
  - Students found in violation of academic dishonesty will receive a failing grade in the course.
  - The student will receive a grade of 0 on the assignment or exam in question (subject to the discretion of the instructor).
  - The student may be asked to re-do the assignment or retake the exam (without credit) to ensure that the student has learned how to properly cite sources or demonstrate that he or she has command of material covered.

**Disability Accommodations**

The University’s policies regarding students with disabilities are available [here](#). Students who have disability accommodations awarded by the University Student Disability Services Office should inform the Harris Dean of Students office by the end of the first week of class. The Harris Dean of Students Office will work with the student and instructor to coordinate the students’ accommodations implementation.

Harris students are not required to submit their accommodations letter to the instructor. Students from other divisions in the University must submit their accommodations letter to either the instructor or the Harris Dean of Students Office.

Students who do not yet have formal accommodations in place but who feel they need accommodations on a temporary or ongoing basis should contact the Harris Dean of Students Office or Student Disability Services.
It is also worth noting that teaching in a remote environment may generate other questions and issues related to providing accommodations to students with disabilities. Some suggestions are available here.

Diversity and Inclusion

The Harris School welcomes, values, and respects students, faculty, and staff from a wide range of backgrounds and experiences, and we believe that rigorous inquiry and effective public policy problem-solving requires the expression and understanding of diverse viewpoints, experiences, and traditions. The University and the Harris School have developed distinct but overlapping principles and guidelines to insure that we remain a place where difficult issues are discussed with kindness and respect for all.

- The University’s policies are available here. Specifically, the University identifies the freedom of expression as being “vital to our shared goal of the pursuit of knowledge, as is the right of all members of the community to explore new ideas and learn from one another. To preserve an environment of spirited and open debate, we should all have the opportunity to contribute to intellectual exchanges and participate fully in the life of the University.”
- The Harris School’s commitments to lively, principled, and respectful engagement are available here: “Consistent with the University of Chicago’s commitment to open discourse and free expression, Harris encourages members of the leadership, faculty, student body, and administrative staff to respect and engage with others of differing backgrounds or perspectives, even when the ideas or insights shared may be viewed as unpopular or controversial.” We foster thought-provoking discourse by encouraging community members not only to speak freely about all issues but also to listen carefully and respectfully to the views of others.

Appendix: Full Harris Academic Integrity Procedures

Harris Procedures for Allegations of Plagiarism, Cheating, and Academic Dishonesty

First Violation

If a student is accused by an instructor or teaching assistant of plagiarism, cheating, or any other form of academic dishonesty, the student will be summoned to meet with the Dean of Students and the instructor. In the meeting, the student and instructor both present information about the situation. If it is determined by the instructor and the Dean of Students that the student has, in fact, plagiarized or cheated, the following sanctions will be imposed for the first violation:
• The student will generally receive a grade of 0 on the assignment or exam in question. Please note that grading decisions are fully at the discretion of the instructor, who may decide to impose harsher grade penalties.
• The student may be asked to re-do the assignment or retake the exam (without credit) to ensure that the student has learned how to properly cite sources or demonstrate that he or she has command of material covered.
• A formal letter of finding is sent to the student stating that the student has been found in violation of the code of academic honesty and what the sanctions were. The letter, along with any evidence presented, is archived in Harris Student Affairs records until the student graduates if the student has no other violations.
• Students found in violation of the academic honesty policy are not permitted to withdraw from the course to avoid grade penalties from the instructor.
• In cases where plagiarism or academic dishonesty is egregious, the case may be referred to the Area Disciplinary Committee even on a first offense. The Dean makes all decisions about which cases will go before the Area Disciplinary Committee.

Second Violation

If a student who has already been found in violation academic dishonesty is again accused of academic dishonesty, the case will be sent to the Harris Area Disciplinary Committee. Details about the Area Disciplinary Committee procedures can be found in the University Student Manual. Information about the first violation, including the formal letter of finding any evidence, will be presented to the Area Disciplinary Committee, along with evidence of the current allegation. If the student is found in violation of academic honesty a second time, the Area Disciplinary Committee can assign sanctions including transcript notes, disciplinary probation, suspension or expulsion from the University.

Academic Dishonesty Appeals

If a student has been found in violation of academic honesty and does not believe that either the finding or the sanction is fair or correct, the student has the right to appeal the finding by requesting a hearing from the Area Disciplinary Committee. More information about the Area Disciplinary Committee is available here.