Interfaces for Human-AI Interaction

CS889 // Winter 2025

Anamaria Crisan (ana.crisan@uwaterloo.ca)

Course Description

Humans are increasingly interacting with AI systems in many areas of our professional and personal lives. However, these interactions are not seamless, leaving people frustrated or, in a worst-case scenario, harmed by AI technology. In this research-focused course, we will explore interdisciplinary perspectives from human-computer interaction, artificial intelligence, and human-centered design to investigate how to build and assess AI systems that are meant to interact with people. It aims to cover a breadth of topics, while also providing practical tools, techniques, and methods for designing software interfaces that support human-AI Interaction. The topics of this course are structured around five themes, beginning with a discussion on early and contemporary approaches to human-machine interaction, and followed by design guidelines and principles, paradigms human-AI teaming (including human-in-the-loop systems), feedback mechanisms between humans and AI agents, and concluding with responsible and trustworthy AI development.

At the end of this course, students should have developed a foundation for thinking critically about building human-AI systems in a human-centered way.

Audience

This course has no required pre-requisites. Taking a prior Human-Computer Interaction (HCI) course (CS 349 and CS 449) can be helpful, however, other programming skills (e.g., Python, Java Script, R) and research backgrounds (cognitive psychology, machine learning, artificial intelligence) are also valuable.

Course Organization

The course is organized around 10 core weeks, each covering a different topic related to human-AI interaction. During the second week of class, students will have the opportunity to suggest a focus topic for week 12. The last week of class will be dedicated to the presentation of course projects. There are no specific textbooks for this course, instead, we will collectively read and review older and contemporaneous research papers.

Each week, there will be two to three required papers to read, with additional, optional, readings for students interested in going further. Ahead of each class, students are required to submit a micro-review of *one* of the required readings. Additionally, each week

one student will be assigned to present one of the required readings, which will be discussed as a group.

Students are also expected to complete a self-directed course project, either individually or as part of a team.

Finally, this course will provide a basic overview of some programming languages and libraries you can use, but it does *not* provide in-depth instructions on these tools.

Participation in Class

Full participation in class will ensure that you take away more from the course. Your attendance and participation in class and online discussion will be recorded.

Paper Micro-reviews

A microsummary is a brief program committee-style review for each paper that will be submitted ahead of class and discussed. Students are required to submit a micro-review *each week* for one of the required readings. The micro-review should be no more than 500 words in length. There will be ten micro-reviews to produce over the course, each valued at 2.5% of your overall course grade.

Paper Presentation

Each week, two to three students will deliver a presentation on one of the required readings. Each presentation will be between 15-20 minutes. A general class discussion will follow. Students should expect to deliver up to two paper presentations over the duration of the course.

Course Project

Students are required to propose and to submit *one research project* that they can conduct individually, or as part of a team (up to three individuals). Students can submit one of the following three project types:

- A systematic literature review
- An evaluation of a technique
- An implementation of system/technique

The merits and contributions of each course project will be assessed based upon the project type (e.g., the requirements for a systematic review are not the same as an implementation).

Project Proposal

The project proposal is a 1–2 page report, excluding references, that outlines your project idea. Prior to submitting your proposal, you are required to submit a pitch to me (<500 words) that provides a rough idea of what you want to do. *I need to approve your project pitch before you move forward to writing the full project proposal*.

Project Presentation

Your project presentation consists of two components. First, you will be required to give a 10-minute research conference-style presentation of your project (+ 5 minutes for questions) in the last week of class. Second, you will be required to prepare and submit a short video (5 minutes or fewer) summarizing your research.

Project Report

You will write-up your project using the ACM template conference proceedings double column format. Reports should be a minimum of 6 pages and a maximum of 8, excluding references. One exception is systematic literature review projects, which should have a minimum of 8 pages and a maximum of 10.

Evaluation

- **Participation in Class (50%)**
  - Class Attendance/Participation (5%)
  - Paper Micro-reviews (25%)
  - Paper Presentations (20%)
- **Course Project (50%)**
  - Project Proposal (10%; Pitch=3%, Proposal=7%)
  - Project Presentation (15%; In-class: 10%; video=5%)
  - Project Report (25%)

Course Schedule

Note: This topic list and schedule is subject to change.

Week	Topic
Theme 1: What is Human-AI Interaction	
1 (Jan 6th)	A History of Human-Machine Interaction
2	Agency + Automation
Theme 2: Designing with AI/ML	
3	Guidelines for Designing Human-AI Interaction ***Due: Project Proposal Pitch***
4	Tools, Techniques, and Applications

Theme 3: Human AI Teaming and Complementarity	
5	Mixed-Initiative, Human-in-the loop, and Interactive Machine Learning ***Due: Project Proposal***
6	Human-AI Alignment
7	READING WEEK NO CLASS
Theme 4: Feedback Mechanisms	
8	Active Learning and Meta-learning
9	Reinforcement Learning from Human Feedback
Theme 5: Responsible AI Development	
10	Fairness, accountability, Trust, and Transparency
11	AI Ethics and Society
12	Student Selected Focus Topic
13	Project Presentations

Policies

Deliverable and Late Penalties

All micro-reviews are due at midnight on the day of class (i.e., the summary should be sent prior to the start of class). Project proposal and report should be submitted by 5pm on the day they are due.

A failure to submit a micro-review prior to the start of class will result in 0% grade for that week. However, students are given 1-free passes to use during the entirety of the term, worth a total of 2.5% (e.g., you can fail to submit one micro-review). *Students that do not use their free pass will have it automatically added to their final grade (*i.e., in theory, it is possible to score more than 100%). Late penalties for project proposal and presentations are 5% deduction for each additional day (5:01pm to 5:00 pm).

Academic Integrity

To maintain a culture of academic integrity, members of the University of Waterloo community are expected to promote honesty, trust, fairness, respect and responsibility. Please consult the University of Waterloo policies for full information: https://uwaterloo.ca/academic-integrity

Use of Generative AI

Given this course on human-AI interaction, sometimes it is appropriate to submit AI generated content. However, there are limitations, and you are still required to submit work that clearly demonstrates your intellectual contribution.

Our course will follow the University of Waterloo (https://uwaterloo.ca/academic-integrity/artificial-intelligence-and-chatgpt) and ACM (https://www.acm.org/publications/policies/new-acm-policy-on-authorship) policies on the use of generative AI (GenAI) in course work. If you use GenAI in your work, beyond typographical correctness (e.g., using Grammarly to check your spelling and grammar) please acknowledge its use. A special note that using GenAI to summarize papers and submit them as micro-reviews is not permitted and will be given a zero grade.

All submitted works will be assessed through AI checkers. If it detects that you have non-original content, you will be given the opportunity to discuss this with me. Failing to comply with this policy may result in an academic offense as outlined in Policy 71, Student Discipline

Grievance and Appeals

A decision made or penalty imposed under Policy 70, Student Petitions and Grievances (other than a petition) or Policy 71, Student Discipline may be appealed if there is a ground. A student who believes he/she has grounds for an appeal should refer to Policy 72, Student Appeals.

Discipline

A student is expected to know what constitutes academic integrity to avoid committing an academic offence, and to take responsibility for his/her actions. [Check the Office of Academic Integrity for more information]. A student who is unsure whether an action constitutes an offence, or who needs help in learning how to avoid offences (e.g., plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course instructor, academic advisor, or the undergraduate associate dean. For information on categories of offences and types of penalties, students should refer to Policy 71, Student Discipline. For typical penalties, check Guidelines for the Assessment of Penalties.

Note for Student with Accessibility Needs

AccessAbility Services, located in Needles Hall, Room 1401, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with AccessAbility Services at the beginning of each academic term.