CURE3012
Special Topics in Cultural Studies II
- Introduction to the Ethics of Artificial Intelligence and Data Science
文化研究專題（二）- 人工智能與數據科學倫理導論

Course description:

Artificial intelligence (AI) and data science are reshaping every aspect of contemporary life—from how we work and communicate to how decisions are made about health, education, and governance. This course introduces the ethical, social, and philosophical issues that arise from these technologies. Students will examine topics such as algorithmic bias, data surveillance, automation, misinformation, and AI's impact on human agency and creativity. Drawing on philosophical frameworks, scientific research, and real-world case studies, the course encourages students to think critically about how technological systems embody values and power structures, and how they can be governed responsibly. The course aims to develop students' ethical reasoning, cross-disciplinary understanding, and ability to articulate informed positions on emerging technologies. Through a team-based learning approach, students engage in weekly readings, discussions, and collaborative writing exercises that cultivate both independent judgment and cooperative problem-solving. No technical background is required.

Learning outcomes:

1. Identify and explain key ethical theories and concepts relevant to the study of AI and data science.
2. Analyze real-world cases of AI technologies using multiple ethical frameworks
3. Evaluate the societal impacts of algorithmic systems—such as bias, surveillance, and automation—through critical and comparative approaches that incorporate global and cross-cultural viewpoints.
4. Apply ethical reasoning to formulate arguments concerning the responsible use of emerging technologies.
5. Collaborate effectively in team-based learning environments to produce outlines and response papers
6. Communicate clearly and persuasively, in written and oral forms, complex ethical ideas about technology
7. Reflect critically on one's own assumptions and role as a user, designer, or citizen in an AI-mediated society.

Course syllabus:

Topic	Contents / Fundamental Concepts
1. Introduction: Value and Technology	Overview of technology and human values. Heidegger's and Bryson's perspectives on technology as cultural practice.
2. AI and Data Science: Foundations	Basic concepts in artificial intelligence, machine learning, and data science.
3. Ethical Framework	Introduction to virtue ethics and the idea of human flourishing in technological contexts. Kantian deontology and utilitarian approaches to ethical reasoning.
4. Data and Surveillance Capitalism	The political economy of data extraction. Privacy, consent, and corporate power.
5. Algorithmic Bias and Fairness	Social and technical dimensions of bias. Case studies on discriminatory algorithms. Concepts of justice, accountability, and transparency.
6. AI as Mirror of Humanity	How AI systems reflect human cognition, prejudice, and creativity.
7. Data Colonialism	Power asymmetries in data infrastructures. Post-colonial and global-South perspectives on AI governance and digital labor.
8. Responsibility Gap	The "responsibility gap" in autonomous systems. Concepts of moral agency and accountability.
9. Online Manipulation	Behavioral economics of attention. Nudging, dark patterns, and algorithmic influence in social media. Ethical boundaries of persuasion.
10. Deepfakes and Misinformation	The epistemology of misinformation. Truth, authenticity, and trust in AI-generated media. Implications for democracy and public discourse.
11. Gamification	AI as learning system and the gamification of communication.

Course components:

Teaching Modes and Learning Activities	
On-site face-to-face (please specify if it is hybrid, i.e. some students will attend the activities elsewhere)	**Percentage of time**
*Lectures (hybrid yes/**no**)*	50%
*Interactive tutorial (hybrid yes/**no**)*	50%

Assessment type, percentage and rubrics:

Assessment type	Percentage
Weekly Quizzes (1-11)	25%
Reading Annotations & Reflections	5%
Weekly Group Outlines and Response Papers	30%
Final Individual Paper	20%
Final Quiz	20%

Assessment rubrics

Rubric 1: Reading Annotations & Reflections

Criteria	%	
Number of Annotations	25%	Comprehensive and consistent margin notes showing active engagement with all key sections of the text.
Quality of Annotations	40%	Insightful comments that identify main arguments, definitions, objections, and connections across ideas.
Quality of Reflection	25%	Clear and thoughtful reflection explaining what was learned, challenges faced, and new insights gained.
Presentation / Neatness	10%	Well-organized, legible, and complete submission following all formatting guidelines.

Rubric 2: Weekly Group Outlines & Response Papers

Criteria	%	
Structure and Quality of Outline	50%	Logical and coherent outline that clearly guides argument development for the response paper.
Argumentation & Use of Evidence	25%	Insightful arguments grounded in course readings; demonstrates accurate understanding and integration of concepts.
Collaboration & Contribution	15%	Active participation in group discussion and equitable contribution to the final outline.
Formatting & Compliance	10%	Fully adheres to instructions, clear organization, and professional presentation.

Rubric 3: Final Individual Paper

Criteria	%	
Introduction & Thesis	10%	Clear, concise thesis that directly answers the response question and previews argument structure.
Theoretical Basis	20%	Sophisticated understanding and accurate explanation of relevant ethical theories and concepts.
New Insight & Analysis	25%	Original and well-supported argument extending or applying theoretical ideas to new contexts.
Structure & Coherence	15%	Smooth, logical flow with effective transitions and clear paragraphing.
Citation & Fact-checking	10%	Accurate use of sources and APA citation; no factual or interpretive errors.
Reflections (AI Feedback + Writing Process)	10%	Thoughtful evaluation of revision process and critical reflection on use of generative AI feedback.
Formatting & Compliance	10%	Follows all assignment requirements with clear writing and professional formatting.

Course schedule:

Class/ week	Date	Topic	Requirements
Week 1	Jan 8	Introduction: Value & Technology	Main reading: Sven Nyholm, This Is Technology Ethics — "What is Technology" (print for open-book quiz)
Week 2	Jan 15	AI & Data Science I; Ethical Framework I	Crash Course AI: #2 Supervised Learning; #3 Neural Networks & Deep Learning; #4 Training Neural Networks Main reading: Nyholm, This Is Technology Ethics — "What is Ethics" (to end of §2.3)
Week3	Jan 22	AI & Data Science II; Ethical Framework II	Crash Course Philosophy: #35 Kant & Categorical Imperatives; #36 Utilitarianism Main reading: Nyholm, This Is Technology Ethics — "What is Ethics" (remainder)
Week 4	Jan 29	Data & Surveillance Capitalism	Video: You Are the Object of a Secret Extraction Operation Main reading: Kate Crawford, Atlas of AI — "Data" chapter
Week 5	Feb 5	Algorithmic Bias & Fairness	Julia Angwin, Uncovering Machine Bias (AI Now) Main reading: Cathy O'Neil, Weapons of Math Destruction — Introduction & Ch. 1
Week 6	Feb 12	AI as Mirror of Humanity	Talk: Kate Crawford, The Trouble with Bias (6:08–43:00) Main reading: AI Mirror — Ch. 2 "Minds, Machines, and Gods"
Week 7	Feb 26	Data Colonialism	MIT Tech Review "AI Colonialism" series (Hao et al.) Main readings: Data Grab; Resisting Colonialism—why AI systems must embed the values of the historically oppressed
Week 8	March 12	Responsibility Gap	Main readings: *The Ethics of Advanced AI Assistants* — Ch. 2 (beginning–end); Sinnott-Armstrong, Schaich Borg & Conitzer, *Moral AI and How We Get There* — Ch. 5 "Can AI be Responsible?
Week 9	March 19	Online Manipulation	Talk: Sally Woellner, *Dark Patterns: How Design Seeks to Control Us* Main readings: *The Ethics of Advanced AI Assistants* — Ch. 4; Ch. 9 "Influence"
Week 10	March 26	Deepfakes & Misinformation	Main readings: *The Ethics of Advanced AI Assistants* — Ch. 3; Ch. 16 "Misinformation" (skip footnotes) Weatherall & O'Connor, *Fake News!* — Section 3 "The Epistemology of Misinformation" (print)

Week 11	April 2	Gamification & Reasoning Models	Main readings: Twitter Gamifies the Conversation Demystifying Reasoning Models
Week 12	April 9	In-class Essay Assignment	
Week 13	April 16	Review and In-class Final Quiz	

Contact details for teacher:

Professor/Lecturer/Instructor:	
Name:	Linus Ta-Lun Huang
Office Location:	KKB 208
Telephone:	3943-5270
Email:	linushuang@cuhk.edu.hk
Teaching Venue:	
Website:	https://www2.crs.cuhk.edu.hk/faculty-staff/teaching-faculty/huang-ta-lun-linus
Other information:	

Academic honesty and plagiarism:

Attention is drawn to University policy and regulations on honesty in academic work, and to the disciplinary guidelines and procedures applicable to breaches of such policy and regulations. Details may be found at http://www.cuhk.edu.hk/policy/academichonesty/.

With each assignment, students will be required to submit a signed declaration that they are aware of these policies, regulations, guidelines and procedures.

- In the case of group projects, all members of the group should be asked to sign the declaration, each of whom is responsible and liable to disciplinary actions, irrespective of whether he/she has signed the declaration and whether he/she has contributed, directly or indirectly, to the problematic contents.

- For assignments in the form of a computer-generated document that is principally text-based and submitted via VeriGuide, the statement, in the form of a receipt, will be issued by the system upon students' uploading of the soft copy of the assignment.

- Students are fully aware that their work may be investigated by AI content detection software to determine originality.

- Students are fully aware of the AI approach(es) adopted in the course. In the case where some AI tools are allowed, students have made proper acknowledgment and citations as suggested by the course teacher.

Assignments without a properly signed declaration will not be graded by teachers.

Only the final version of the assignment should be submitted via VeriGuide.

The submission of a piece of work, or a part of a piece of work, for more than one purpose (e.g. to satisfy the requirements in two different courses) without declaration to this effect shall be regarded as having committed undeclared multiple submissions. It is common and acceptable to reuse a turn of phrase or a sentence or two from one's own work; but wholesale reuse is problematic. In any case, agreement from the course teacher(s) concerned should be obtained prior to the submission of the piece of work.

Use of generative AI tools:

Approach 3 - Use of AI tools is allowed with explicit acknowledgement and proper citation

Use of AI tools is allowed with explicit acknowledgement and proper citation

Students may use some AI tools in some class activities and assignments on the condition that they make explicit acknowledgement and proper citations of the input from AI tools.

Acknowledging support from AI tools
Students are required to acknowledge all functional uses of a generative AI tool and cite it when they paraphrase, quote, or incorporate into their own work any content (whether it is text, image, data, or other format) that was created by it.

i. An example of acknowledgement

'I acknowledge the use of (name of AI tool – e.g. ChatGPT (https://chat.openai.com/) to (specify the support, e.g. plan my essay, generate some ideas for the content, ask for examples of data collection instruments, get the dates of historical events, etc.).

ii. An example of citation
OpenAI. (2023). *ChatGPT* (Mar 20 version). https://chat.openai.com/chat

(Students are reminded that due to the rapid developments of generative AI tools, some citation formats may be updated regularly.)

iii. An example of including texts generated by an AI tool in their work

"The following text was generated by an AI tool / language model (ChatGPT):"
[Insert the text generated by ChatGPT here.]

iv. An example of including texts generated by an AI tool and the prompts that were used to elicit the text from the AI tool

"[The prompt], as generated by an AI language model (ChatGPT):"
[Insert the text generated by ChatGPT in response to the prompt.]

Students are reminded to learn and use the AI tools responsibly and ethically and be aware of the limitations.

Students are reminded to clarify with the course teacher and obtain permission if necessary when in doubt.