Please consult Intellectual Property Rights before making a photocopy. Please use the textbook of copyrighted edition.

②國玄東華大學

教學計劃表 Syllabus

課程名稱(中文) Course Name in Chinese	人工智慧倫理			學年/學期 Academic Year/Semester		113/2
課程名稱(英文) Course Name in English	AI Ethics					
科目代碼 Course Code	TCAI10010	系級 Department 學二 (& Year		開課單位 Course-Offering Department	理工學院	
修別 Type	選修 Elective	學分數/時 Credit(s)/Hou		3.0/3.0		
授課教師 Instructor	/黄華彥/甘偵蓉	E(遠端授課教師	F)			
先修課程 Prerequisite						

課程描述 Course Description

本課程旨在帶領學生認識及反思 AI 這項技術及其應用所涉及的倫理、風險與社會議題。首先,將簡介學習機器學習 AI 的發展歷史以及國際倫理規範,以及哲學倫理學的基本概念。

其次,將說明 AI 的資料來源和分類如何影響 AI 的預測或決策。接著,將探討 AI 模型和演算法的設計,並分析 AI 技術所涉及的社會性與政治性等問題。

另外,本課程亦討論 AI 應用之後所帶來的人類生存危機感、偏誤與歧視、加劇既有的社會不平等、性別刻板印象、以及勞動產業鍊及剝削等倫理與社會爭議。

最後,本課程期待修課學生能針對目前已知發生倫理社會爭議的 AI 專案提出可行的解決方案。

課程目標 Course Objectives

本課程旨在帶領學生認識及反思AI這項技術及其應用所涉及的倫理、風險與社會議題。首先,將簡介學習機器學習AI的發展歷史以及國際倫理規範,以及哲學倫理學的基本概念。其次,將說明AI的資料來原和分類如何影響AI的預測或決策。接著,將探討AI模型和演算法的設計,並分析AI技術所涉及的社會性與政治性等問題。另外,本課程亦討論AI應用之後所帶來的人類生存危機毆、偏誤與歧視、加劇既有的社會不平等、性別刻板印象、以及勞動產業鍊及剝削等倫理與社會爭議。最後,本課程期待修課學生能針對目前已知發生倫理社會爭議的AI專案提出可行的解決方案。

圖示說明Illustration :● 高度相關 Highly correlated ○中度相關 Moderately correlated

授課進度表 Teaching Schedule & Content

週次Week	內容 Subject/Topics	備註Remarks
1	2/19 課程與 AI 發展史簡介	
1	測試 Slido,TttC	
9	2/26 倫理基本概念與 AI 國際倫理規範簡介	
2	測試 Slido, TttC	
9	3/5 AI 資料的來源與分類之倫理議題	
3	講授 2 節,實體與線上同步討論1 節	
4	3/12 AI 模型的倫理與社會議題	
4	講授 2 節,實體與線上同步討論1 節	
F	3/19 AI 演算法的社會與政治性之議題 (一)	
5	講授 2 節,實體與線上同步討論1 節	

6	3/26 AI 演算法的社會與政治性之議題 (二)							
0	講授 2 節,實體與線上同步討論1 節							
7	4/2 春節溫書假!!							
	4/9 AI 與人類生存危機?							
8	講授 2 節,實體與線上同步討論1 節							
9	講技 2 即,實體與線上向步討論1 即 4/16 期中考試週 Midterm Exam							
	4/23 AI 的部署與運作之倫理與社會議題:偏誤、偏見與歧視							
10	講授 2 節,實體與線上同步討論1 節							
	4/30 AI 的部署與運作之倫理與社會議題:自動化不平等(一)							
11 講授 2 節,實體與線上同步討論1 節								
1.0	5/7 AI 的部署與運作之倫理與社會議題:自動化不平等(二)							
12	講授 2 節,實體與線上同步討論1 節							
10	5/14 AI 的部署與運作之性別議題							
13	講授 2 節,實體與線上同步討論1 節							
1.4	5/21 AI 的勞工與產業鍊之社會議題							
14	14 講授 2 節,實體與線上同步討論1 節							
1.5	5/28 線上展覽分組簡報與同儕互評							
15	預錄 1 分鐘專案自介							
16	6/4 繳交分組書面報告							
10	授課教師與盟校教師評分							
17	6/11 彈性授課							
18	6/18 彈性授課							
	教學策略 Teaching Strategies							
課堂講	授 Lecture 分組討論Group Discussion 參觀實習]	Field Trip						
↓ — 其他Mi	scellaneous:							
	教學創新自評 Teaching Self-Evaluation							
l	Innovative Teaching)							
問題導	向學習(PBL) 團體合作學習(TBL) 解決導向學	로 웹 (SBL)						
翻轉教	室 Flipped Classroom							
社會責任((Social Responsibility)							
在地實	踐Community Practice 產學合作 Industy-Academia Cooperati	on						
跨域合作(Transdisciplinary Projects)							
跨界教	學Transdisciplinary Teaching 跨院系教學Inter-collegiate Teaching							
業師合	授 Courses Co-taught with Industry Practitioners							
其它 othe	er·							
, a other								

學期成績計算及多元評量方式 Grading & Assessments									
配分項目	配分比例	多元評量方式 Assessments							
Items Percentage		測驗 會考	實作 觀察	口頭 發表	專題 研究	創作 展演	卷宗 評量	證照 檢定	其他
平時成績(含出缺席) General Performance (Attendance Record)	30%			~					課堂參與討論。依據線上提問或公共討論系統的發言紀錄。
期中考成績 Midterm Exam	15%						~		Take home exam 24 小時
期末考成績 Final Exam	25%%						~		運從生用目倫等 用目倫等 用目倫等 等方 所知的 出 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一
作業成績 Homework and/or Assignments	30%								閱讀筆記與上 課學習單,當天 上完課後48小 時內上傳
其他 Miscellaneous									

評量方式補充說明

Grading & Assessments Supplemental instructions

教科書與參考書目(書名、作者、書局、代理商、說明)

Textbook & Other References (Title, Author, Publisher, Agents, Remarks, etc.)

- 1. Borg, J. S., Sinnott-Armstrong, W., & Conitzer, V. (2024). Moral AI: And How We Get There. Random House.
- 2. Russell, S., & Norvig, P. (2021). Chap. 1 & Chap. 27. In Artificial intelligence: A modern approach (4th ed.). University of California, Berkeley.
- 3. Gabriel, I. (2020). Artificial intelligence, values, and alignment. Minds & Machines, 30, 411-437. https://doi.org/10.1007/s11023-020-09539-2
- 4. Russell, S. (2019). Chap. 7 & Chap. 10. In Human compatible: Artificial intelligence and the problem of control. Penguin.
- 5. Vallor, Shannon (2016). Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting. New York, NY: Oxford University Press USA.
- 6. 凱特. 克勞馥(Kate Crawford) (2022),第二、三、四章,出自《人工智慧最後的秘密》,臉譜文化,pp. 71-108。
- 7. 泰娜. 布策(Taina Bucher) (2021),第一、二、四章,出自《被操弄的真實:演算法中隱藏的政治與權力》,台灣商務印書館。
- 8. 維吉尼亞.尤班克斯(Virginia Eubanks) (2022),第三、四、五章,出自《懲罰貧窮》,寶鼎出版。
- 9. 凱西. 歐尼爾(Cathy 0'Neill)(2109),第一、十章,出自《大數據的傲慢與偏見》,大寫出版。
- 10. 約蘭德(Yolande Strengers) &珍妮. 甘迺迪(Jenny Kennedy) (2023),第一、二章,《智慧妻子》,陽明大學交通出版社。
- 11. 尼克.伯斯特隆姆(Nick Bostrom)(2016),第十三、十四、十五章,出自《超智慧》,八旗文化。
- 甘偵蓉(2023),〈人工智慧科研倫理與風險之基本認識〉,《科技、醫療與社會》季刊,
 167-220。
- 13. 甘偵蓉(2024),〈AI 開發過程的倫理權衡:自駕車決策案例研究〉,中研院《歐美研究》 季刊第 54 卷第 1 期,頁 1-68。DOI: https://doi.org/10.7015/JEAS.202403_54(1).0001 (THCI/TSSCI)
- 14. 甘偵蓉(2024),〈人工智慧系統應該內建倫理嗎?人工道德行為者之探討〉,《危機時代的哲學—「後」疫情時期的反思》,五南圖書出版。ISBN: 978-626-393-008-7

課程教材網址(含線上教學資訊,教師個人網址請列位於本校內之網址)
Teaching Aids & Teacher's Website(Including online teaching information.
Personal website can be listed here.)
其他補充說明(Supplemental instructions)