

## 課名：人工智慧倫理 (AI Ethics)

115.12.19修正

### 課程基本資料

開設學校：東海大學

開授教師：甘偵蓉

聯繫email: zrgan@thu.edu.tw

班級人數：由聯盟學校或協同教師自訂

開課級別：學士班(碩士班學生可修課，但是不算畢業學分)

授課語言：中文

授權方式：混成式授權。

條件式授權：預設聯盟內學校皆需提供協同老師，並於校內帶領同步實體討論課。

封閉式授權：若協同老師無法提供實體討論課，請主動於1/9(五)前填寫表單向專辦

申請封閉式授權(<https://www.surveycake.com/s/DeIXR>)，或是預計實際修

課學生少於10人者，直接以封閉式(鏡像課程)授權。

協同老師經歷建議：建議協同教師學經歷：有實體或線上修習機器學習基本概念6小時、有受過哲學倫理學、哲學思辨訓練者尤佳。

同步遠距上課時間：每週二 15:20~18:10，包含一小時實體討論課，全程需同步遠距上課。

遠距上課YouTube位置：[www.youtube.com/@EthicsPhil](http://www.youtube.com/@EthicsPhil)

課程網頁：<https://sites.google.com/go.thu.edu.tw/aiethics>

### 課程概述：

本課程旨在鼓勵學生批判性的思考AI對個人、社會與制度的影響，引導學生了解AI倫理的多元面向。課程架構分為三部分：基礎篇（W1-W5），建立AI倫理基本概念，包含簡介機器學習與大型語言模型的基本原理，並教導實踐倫理常用來分析相關議題或案例的批判思維工具。核心主題篇（W6-W10），探討AI系統的責任歸屬、演算法偏誤問題與公平性、ML與LLM的資訊安全倫理，揭示技術設計中隱含的風險與價值判斷。特定議題篇（W11-W13），檢視陪伴式AI應用程式、物理AI、AI助攻或威脅民主、AGI與人類生存風險等。

## 參考書籍

1. 主要教科書：自編講義
2. Boddington, P. (2023). *AI ethics*. Singapur: Springer International Publishing. [AIE]
3. Borg, J. S., Sinnott-Armstrong, W., & Conitzer, V. (2024). *Moral AI: And How We Get There*. Random House. [MAI]

## 課程內容大綱

週次	日期	課程主題		備註
1	2/24	基礎篇	AI倫理課程簡介	講授2節，實體與線上同步測試NTU COOL, Talk to the City, Slido
2	3/3	基礎篇	機器學習與生成式AI基本設計原理	講授2節，實體與線上同步測試NTU COOL, Talk to the City, Slido
3	3/10	基礎篇	實踐倫理之批判思維工具簡介	講授2節，實體與線上同步討論1節，使用Talk to the City
4	3/17	基礎篇	實踐倫理之批判思維工具簡介	講授2節，實體與線上同步討論1節，使用Talk to the City
5	3/24	基礎篇	實踐倫理之批判思維工具簡介	講授2節，實體與線上同步討論1節，使用Talk to the City

6	3/31	核心主題篇	AI相關責任問題	講授2節，實體與線上同步討論1節，使用Talk to the City
7	4/7	核心主題篇	AI偏誤問題與公平性	講授2節，實體與線上同步討論1節，使用Talk to the City
8	4/14	春節溫書假	—	講授2節，實體與線上同步討論1節，使用Talk to the City
9	4/21	核心主題篇	ML & LLM資訊安全與倫理	講授2節，實體與線上同步討論1節，使用Talk to the City
10	4/28	核心主題篇	AI相關隱私問題	講授2節，實體與線上同步討論1節，使用Talk to the City
11	5/5	特定議題篇	AI Companion Apps倫理問題	講授2節，實體與線上同步討論1節，使用Talk to the City
12	5/12	特定議題篇	物理AI__機器人與自駕車倫理問題	講授2節，實體與線上同步討論1節，使用Talk to the City
13	5/19	特定議題篇	AI與民主	講授2節，實體與線上同步討論1節，使用Talk to the City
14	5/26	特定議題篇	AGI與人類生存危機問題	講授2節，實體與線上同步討論1節，使用Talk to the City
15	6/2	期末分組口頭報告	專案之倫理與風險分析報告	1分鐘簡介影片+5分鐘口頭說明
16	6/9	各校優秀報告1分鐘簡介影片，上傳至Gather Town進行跨校交流		
17	6/16	(for 16+2學校)彈性學習週__自行安排		
18	6/23	(for 16+2學校)彈性學習週__自行安排		

## 成績評量方式

序號 No.	評分項目Assessment Item	配分比例 Percentage
1	課堂參與討論 (含出席)	50 %

2	作業	20 %
4	期末分組報告	30 %