**第一題**

**針對各校同學Week 10 第三節所討論的問題回答，您發現最被公認的黑暗設計是什麼？有舉例像哪個網站的什麼設計嗎？**

最被公認的黑暗設計就是 Cookie 的同意視窗，它會將全部接受設定的非常明顯，拒絕按鈕則是放在很難找到，需要透過好幾個步驟才能點到，引誘人在不耐煩的心情下直接按下全部同意。

另外 Youtube 和 Facebook 的通知設定上面的文字，也在刻意混淆視聽，讓人誤以為在設定隱私，實際上卻是開啟授權，將自己的資料分享出去。

**第二題**

**請閱讀附件的教材Borg, J. S., Sinnott-Armstrong, W., & Conitzer, V. (2024). “Ch6\_Can AI incorporate human morality?" in *Moral AI: And How We Get There*. Random House，摘要整理裡面的內容（如果有使用生成式AI協助整理請告知）。**

這個章節主要在探討能不能，以及應不應該將人類的道德判斷加入 AI 系統當中。

1. **AI 倫理的兩條路線**
   * **工具:** AI 只是協助人類的工具，道德的責任依舊屬於使用工具的人。
   * **代理人:** 當 AI 能做越來越多的事情，做出自己的判斷與選擇，它可能會成為「有道德的主體」，這代表 AI 本身應該需要具有道德判斷的能力。
2. **人類道德的不確定**
   * **道德的多樣性:** 人們對相同的道德問題會因為文化，情境，說的話等因素有所不同。
   * **道德分歧:** 就算在同個社群當中，也常因為價值觀的不同而導致衝突。
   * **規則怎麼定義:** 單靠明確的規則，很難說清現實當中，道德的模糊與彈性。
3. **模仿學習**
   * AI 可以透過機器學習模仿人類判斷，但這樣對能否真的「學會」，依舊得打個問號。
   * 假如資料來源有偏見，或者彼此矛盾，反而會放大人類的道德瑕疵。
4. **AI 該學哪一種道德**
   * 人類自己對哪一套倫理理論是正確的，沒有共識，自然不知道該讓 AI 學哪個好。
   * 大多數人覺得以實用主義出發，不追求唯一真理，而是用一套大多數人可以接受的妥協方案。
5. **應不應該讓AI有道德對應的地位**
   * 假如AI真有了道德，應不應該讓它有對應的地位或權力?
   * 如果給了對應的地位與權力，可能會模糊人與機器的界線。
   * 人自己願不願意將道德判斷交給 AI。
6. **未來展望**
   * 就算技術難以讓AI理解道德，但還是得試試看。
   * 作者說不只要思考 AI 能否做到，更要想「我們期望 AI 變成甚麼樣子」。