

會議記錄

專案名稱	尚未確認		
系統名稱	尚未確認		
文件編號	250804-小組第五次會議		
會議主題	小組第五次會議	編輯者	蕭伊庭
會議日期	2025 年 08 月 04 日	編輯日期	2025/08/04
會議地點	Teams 線上會議		
出席人員	保羅(Paul)、林湘凌(Cecilia)、蕭伊庭(Sherry)、周易暉(Ethan)、羅靖傑(Morris)、金殿倫(Darren)		

■ 本次會議討論事項：

壹、上週分工探討結果

【text to speech-蕭伊庭(Sherry)&周易暉(Ethan)負責】

有找到兩種方式，中階是使用開源工具跑出語音，能在自己的電腦上，用別人訓練好的模型，把一段文字轉成語音並播放；進階是訓練自己的聲音，錄一段人聲，經過處理後訓練模型，能讓電腦用訓練好的聲音說話。主要要學 Python 模組安裝、指令列（命令提示字元）、音訊知識（什麼是 .wav、sampling rate 等）。

保羅(Paul)建議：

1. 專題要物件化的設計

就算不用物件包也要用函式包，不能從頭寫到尾，要一個區塊一個區塊，發揮物件導向這件事，也就是要把程式寫的「可以再利用」，把功能獨立出來因為可能會用很多次，每次都呼叫寫好的函式，這是中大型的開發必須遵守的規則。

2. 在呈現上要用效果 demo

開發專題不是研究案，更注重實作，要對使用有了解才算是稍微掌握技術，呈現要用「成品」說話。

【avatar-金殿倫(Darren)負責】

[HeyGen 網站\(https://www.heygen.com/\)](https://www.heygen.com/)可以自訂真人 AI 模型，也可提供文本讓他配合口部動作講出來。

保羅(Paul)建議：線上服務在我們使用上會比較不方便，容易被綁住，可以參考主任組的文件，有可能會面臨沒有免費版 avatar 的問題，但用擬真人的比較好，要再廣泛去找，若是付費版要估算面試一次約 15 分鐘的成本。

【寵物主題區塊切分&AI 發想-林湘凌(Cecilia)負責】

現在 GPT 也可以看寵物面相做寵物溝通，但很不準，寵物溝通玄學這方面感覺不能碰，只能儘量找相關資料去餵給 AI 創造自己的模型，聽下來感覺面試主題會更有趣。

保羅(Paul)建議：

1. 面試主題有可能卡住，要有效果必須真的成功，聊天機器人用關鍵詞的方式，能掌握的話可以精準地讓他回答問題，如果不能做出模擬真人隨機應變面試，那跟預錄好的面試很像，還是有模擬的效果，但創造臨場感很重要。
2. 生成式 AI 可以客製化，收集足夠資訊就能做到，也沒有很難，但要怎麼去呈現會比較難，畢竟我們不只是文字呈現，要讓使用者有沈浸面試的感覺。
3. 使用者的表情等也會影響面試，雙向的會更難做。面試時教授會從表情語氣去判斷學生在想什麼，不好做因為現在面部表情辨識正確率較高的只有喜怒哀樂，不是不能做，只是要辨識出對面試比較重要的資訊，像是疑惑的微表情可能有點難。

【相關理論依據查詢-羅靖傑(Morris)負責】

1. 臉部辨識情緒來自基本情緒理論，主張人類的基本情緒如快樂、憤怒、悲傷、驚訝、厭惡與恐懼等具有跨文化的一致性，可透過臉部表情加以辨識，進而判斷個體的內在心理狀態。
2. 語音驅動分為語音識別和語音合成，早期沒有語音驅動不能讓語音和面部表情同步。穿戴動捕裝置做出動作能投影到模型上。
3. 恐怖谷大概是說人會對有相似外型的機器人有正向的情感，但達到一個程度心情觀感會急速下降，若虛擬程度再提高，心情正向程度又會上升。

保羅(Paul)建議：

1. 人類大部分的情緒是負多於正，亞洲人想要幸福，但從數據分析幸福並不是很正面的情況，可能是本來很差然後變好，有反差的感覺，像是愛也差不多，有時候是愛恨交織不一定是正面的。
2. 從表情上能理解的情緒是基本情緒，但複合式情緒不太能辨識的出來，或是辨識正確率較低，像是困惑這類心理情緒。另外的問題是每個人不一樣，有人可能是撲克臉但心裡情緒起伏很大，機器辨識若對方表現不夠強烈，正確率會降低，像是蒙娜麗莎到底有沒有笑，人類都有歧義了機器辨識出來我們

也不信。

3. 關於恐怖谷可以延伸找這塊，有很多擬真手遊但我們不會覺得可怕，要超過一個真實感人才不會覺得恐怖，人們好像能接受這理論。

貳、未來一週規劃

寵物那邊先暫時不動，把精力集中在面試，可以再開一個臉部辨識的部分，看一下能做到的程度和套件，套件要選 python 的最方便，動作辨識可以看 MediaPipe（給照片可以畫出火柴人），面部表情好像也可以辨識。

分三組 221 人：text to speech、MediaPipe 辨識人的表情&最後擬真人 avatar。