以AI Agent的結構產生三個Tools，可以經由簡單的語意輸入讓AI去挑選Tool執行，舉例：

(1) 對AI以語音方式輸入「播放嗽叭」、「播放聲音」、「讓喇叭發聲」等同語義詞，喇叭就會發出聲音來(播放出來的聲音內容不拘)

(2) 對AI輸入「STT」、「語音轉文字」等同語義詞，就會在系統上產生一個由麥克風輸入後的語音檔及轉文字的txt檔。

(3) 對AI輸入「現場人數」、「鏡頭中人數」等同意詞，可透過camera取得畫面中人數，並回覆回來。

請說明所採用的AI Service是那一家(OpenAI, Google或是Claude等?及所使用的模型是哪個(GPT 4o, Gemini 1.5 Pro或Sonnet等)。建議以GitHub Repository的URL交卷，建立完整可執行的專案結構，以利複驗。

若相關程式能在Raspberry Pi上實現尤佳，若可，請詳述Raspberry Pi所採用的硬體有哪些?舉凡：Raspberry Pi 5 or Pi 4B, Pi 原廠Camera or USB Camera(廠版、型號)，以利複現。若不行，請也說明自己使用的系統及規格。

P.S. 其餘未說明到的請自行發揮