

# Moodsic

第一組: 呂政祺、孔啟熙、李享

## Topic:

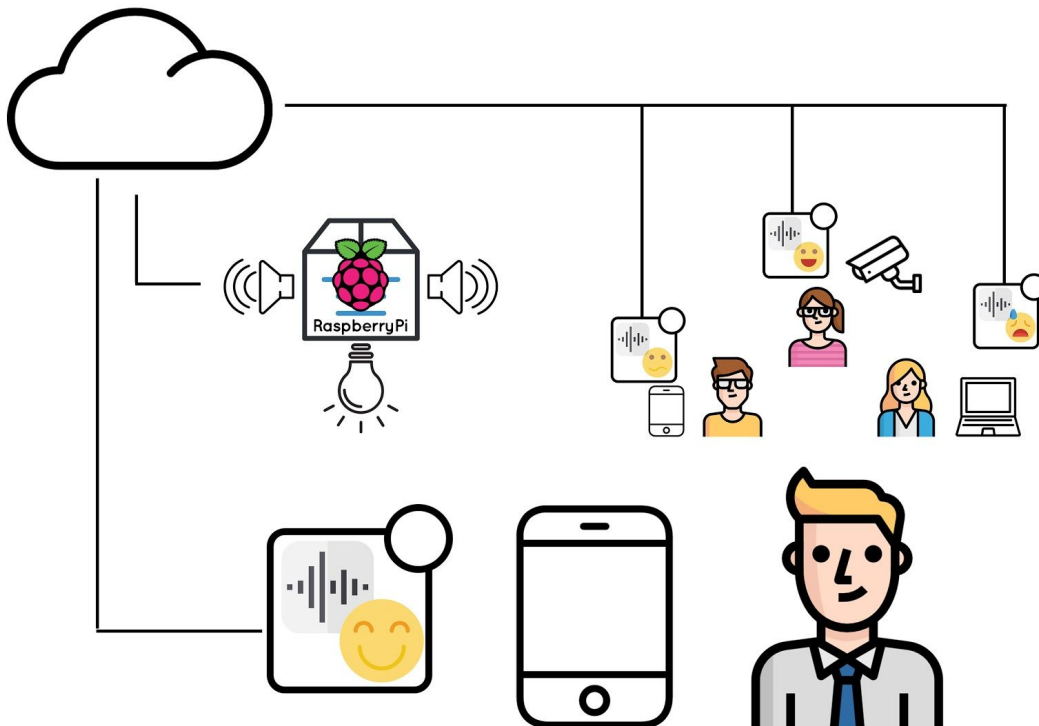
表情與語音情緒分析在音樂推薦的應用，推薦音樂根據主要為下列兩種情境

1. 被動式：利用常駐的攝影機在一特定空間(如客廳的全景)內辨別不同使用者的臉，透過判斷使用者相關聆聽紀錄喜好等個人背景資訊與臉部影像分析之情緒參數隨時以每首歌為單位替換該空間之背景音樂。
2. 主動式：以影像與語音介面(類似視訊)作為操作模式尋求系統推薦音樂，利用言語中的關鍵字分析與臉部影像辨識所得到之情緒分析作為推薦點歌之依據。

## Motivation:

想法來源為Spotify是使用AWS架設起來的，而其中有項功能是每週會整理出一份歌單，歌單內的歌曲皆為使用者沒有體驗過的且是使用者平常喜愛的音樂類型，因此我們萌生了想要實作出更有互動感的音樂服務的念頭。有時一早起來想要聽到神清氣爽的歌曲來提神或是朋友來到家裡開派對，便可以透過這個音樂助理，只要辨別出你的臉來播放平常喜歡聆聽的音樂或是透過大家對談的口氣、情緒來自動播放適當氣氛的樂曲，或是哼著歌詞便能自動搜尋出該歌曲來播放，也是一個滿有樂趣的互動。

## System Architecture:



被動式：如圖中右側三個較小人像，利用各種裝置蒐集影像與語音資訊

主動式：如圖下側較大人像，主動利用手機點歌

## 使用到的AWS服務：

**Amazon Rekognition**：使用鏡頭資料進行臉部辨識，確認使用者。

將影像或影片提供給 Rekognition API，服務就會識別各種物件、人物、文字、場景和活動，並偵測任何不當的內容。也提供高度準確的臉部分析和臉部辨識。

**Amazon Lex**：用於使用語音和文字在應用程式建立交談界面。

提供自動語音辨識 (ASR) 的進階深度學習功能，可將語音轉換為文字，也提供自然語言理解 (NLU) 來辨識文字的含義。

**Amazon Comprehend**：負責語意分析，辨別使用者情緒

是利用機器學習探索文字洞見的自然語言處理 (NLP) 服務。提供關鍵片語擷取、情緒分析、實體辨識、主題建模及語言偵測等 API。

**RDS**：存放使用者資訊以及使用者的歌單(歷史資料)

**S3**：存放training data

### Demo scenario：

因為使用者臉部資料需要事先根據歌單做連接，沒辦法現場找志願者試用我們的服務。

所以demo方式偏向兩種選擇：

- 1.現場操作但由組員試用
- 2.事先操作並錄影

### Purchase plan：

迷你 USB 麥克風，250元/個

webcam camera: 5MP Camera Module (for Raspberry Pi)，5百萬像素，750元/個

Logitech羅技 C310 網路視訊攝影機(500萬畫素/內建麥克風)，1,090元/組

發光器相關(ex:麵包板、LED燈數個、連接線、Blinkt! LED燈)，約500元(未定)