

投稿類別：資訊類

篇名：AI 語音復仇者聯盟

作者：

王聖閔。國立北斗高級家事商業職業學校。資料處理科三年 1 班
陳玟君。國立北斗高級家事商業職業學校。資料處理科三年 1 班
楊芯瑜。國立北斗高級家事商業職業學校。資料處理科三年 1 班

指導老師：

劉玉敏老師

壹、前言

一、研究動機

現今科技日新月異，生活中的大小事都跟科技密不可分，其中人們視為像生命一樣的手機，有著能和人們溝通的 AI 語音助理，時常都聽得到人們對著手機以「嘿！Siri」或是「OK Google」來問天氣、餐廳.....等問題，也因為機器能聽得懂人們的語言且能答覆人們的需求，更讓人好奇說為何這項科技能辨識出人們所提出的問題。

電商巨擘亞馬遜在 2019 年 9 月 24 日宣布 AI 語音相容計畫，號召 36 家跨國企業，包刮 MediaTek、Microsoft、Intel等，共同研究如何加速語音助理的泛用性，讓不同廠商的語音服務能夠在同一台裝置上互通，其中合作商中並不包含蘋果及 Google，這場計畫是否引起科技界的復仇者聯盟呢？就讓我們一起來探討吧！

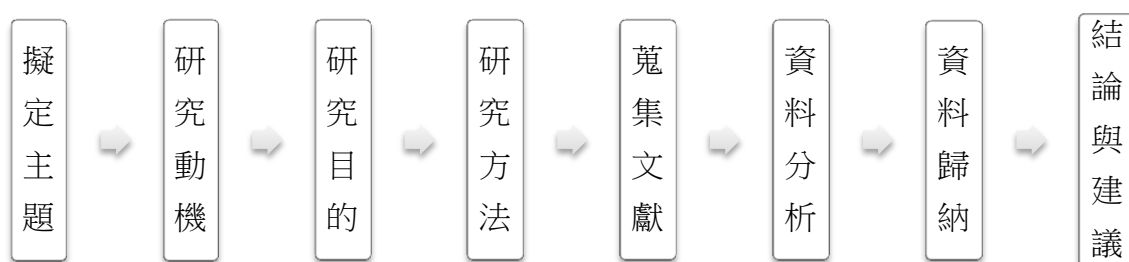
二、研究目的

- (一) 了解語音助理運作原理與技術。
- (二) 了解兩大語音助理功能之比較。
- (三) 分析使用者對兩大語音助理的使用經驗及應用現況。

三、研究方法

- (一) 文獻探討法： 蒐集資料、閱讀書籍與期刊論文，利用網路資訊及報導，了解語音辨識器運作原理與技術和兩大語音助理之功能。
- (二) 問卷分析法： 設計問卷表單，統計數據並針對使用者對兩大語音助理的使用經驗加以分析。

四、研究流程



圖一：研究流程圖

貳、正文

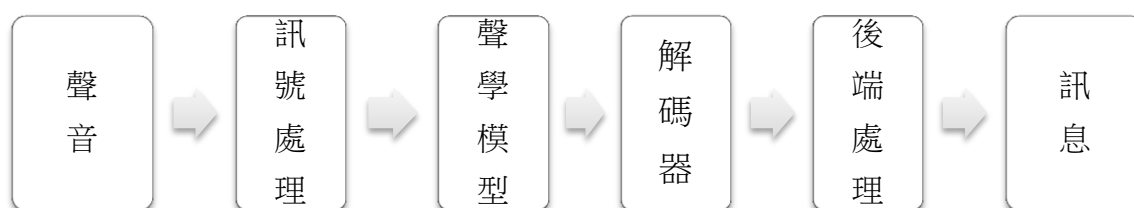
一、語音辨識的演變

在電腦發明之前，有關語音識別技術的原理早已誕生，而早期的聲碼器便被視為語音辨認及合成的雛形。語音辨識，又稱語音識別技術，也稱為自動語音識別（ASR）、電腦語音識別或是語音轉文本識別（STT），其目的是以電腦自動將人類的語音內容轉換為相應的文字。在 1920 年所生產的 Radio Rex 玩具狗是最早的語音辨識器。1950 年末期時，倫敦學院的 Denes 已將語法機率加入語音識別中。到 1960 年時，人工神經網路便引入了語音辨識。

二、語音辨識器之技術原理

目前所廣泛使用的語音辨識系統多採用統計圖識別技術，其基本模組包括訊號處理及特徵提取模組、聲學模型、發音詞典、語言模型、解碼器、後端處理……等，介紹如下：

- （一）訊號處理：又稱人聲處理，將其裝置所收到的聲音訊號數據化處理，再經數位類比轉換裝置輸出，以供聲學模型處理。
- （二）聲學模型：是語音辨識系統中最重要的部分之一，也是將語音識別作為人與機器更順暢交流的橋樑。其中只有單聲學模型無法判別每個單詞應該發什麼音，需要由另一個叫做發音辭典的模型做為輔助，這樣一來，聲學模型就能給定的文字串該依次發哪些音。
- （三）語言模型：是用來表示某一種語言的機率模型，簡單來說，是用來判斷一句話符合自然語言的表達習慣的可能性。
- （四）解碼器：為語音辨識系統的重要核心，即對輸入的訊號，根據模型及詞典，用一套公式化的算法得出最大概率輸出該訊號的詞串。
- （五）後端處理：即進行準確的語義理解，能夠基於上下文及應用場景進行學習和語義理解，透過數據分析，來判斷人們的需求並轉換成文字。



圖二：辨識器產生訊息的流程圖

資料來源：研究者自行繪製

三、「Android」與「iOS」語音助理比較

（一）Android 語音助理

Android，中譯安卓或安致，為一套基於 Linux 核心的開放原始碼行動作業系統。其中，安卓裝置上的語音助理有 Google 個人助理及 Google 即時資訊，兩者之差別在於 Google 個人助理可參與雙向對話（圖三）。

1. Google 個人助理：又稱 Google Assistant 為 Google 開發之智能助理（圖四），提

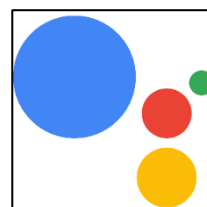
供與 Google 即時資訊類似的功能，例如：搜尋網路、即時翻譯、安排行程、天氣及路況查詢，與 Google 即時資訊之差別在於 Google 個人助理可參與雙向對話。

2. Google 即時資訊：又稱 Google Now 是 Google 開發之智能助理，也是 Google 搜尋應用程式的一部份(圖五)，右方的麥克風圖示為 Google 語音搜尋，又稱 Google Voice Search，使用者可對著手機或電腦說話，使用 Google 進行搜尋，即通過裝置傳送說出的內容，並經過伺服器辨識，根據辨識結果進行搜尋。



圖三：Google 語音助理雙向對話

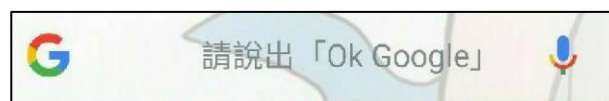
資料來源：研究者自行拍攝



圖四：Google 個人助理

資料來源：維基百科。2020 年 4 月 17 日，取自

<https://is.gd/52E5k8>



圖五：Google Now Launcher

資料來源：研究者自行拍攝

(二) iOS 語音助理

iOS (iPhone OS) 是蘋果公司為行動裝置所開發的作業系統。Siri 是一款內建在 iOS 系統中的人工智慧助理軟體，此軟體採用自然語言處理技術，只可在 iOS、macOS 中執行，使用者可透過對話與裝置進行互動(圖六、圖七)，其中 Siri 的聲音還分成了男聲和女聲，可依需求自行選擇設定(圖八)。

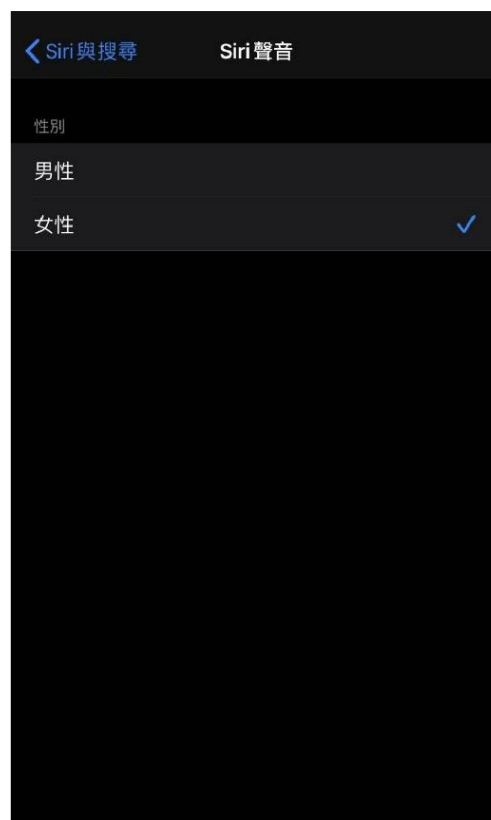


圖六：Siri 圖示

資料來源：維基百科。2019 年 1 月 29 日，取自 <https://is.gd/vhUJuZ>



圖七：Siri 語音助理雙向對話
資料來源：研究者自行拍攝



圖八：Siri 聲音設定
資料來源：研究者自行拍攝

(三) 語音助理功能比較

表一 語音助理比較

| 智慧助理 | Google Assistant | Siri |
|--------------|---|--|
| 營運公司 | Google | Apple |
| 發表時間 | 2016 年 5 月 | 2011 年 10 月 |
| 支援手機 平板系統 | Andriod (iOS) | iOS |
| 支援裝置 | Google Home Google Nset Mini Andriod Auto 汽車 Andriod TV 智慧電視 穿戴裝置手錶耳機 | HomePod iPhone iPad Mac Apple TV Apple watch CarPlay |
| 特點 | Google 連結系統最完善 搜尋引擎、Gmail 等 與 Google Now 連結 | 系統獨立 |
| 中文支援 | 有 (2018 Oct) | 有 (2012 Sep) |

資料來源：本研究自行彙整

表二 語音功能比較

| 功能 | Google Now | Siri |
|----------------|------------|------|
| 新聞 | O | O |
| 推薦 App | X | O |
| 單位轉換 | O | O |
| 天氣資訊 | O | O |
| 語音搜尋 | O | O |
| 推薦聯絡人 | X | O |
| 行事曆彙整 | O | O |
| 交通路線建議 | O | O |
| 第三方 App 整合 | O | O |
| 附近景點、餐點 | O | O |
| 來電辨識 Caller ID | O | O |
| 依據聯絡人的備忘錄 | X | X |
| 依據地理位置的備忘錄 | O | O |

資料來源：本研究自行彙整

(四) 語音助理之應用

隨者人工智慧與物聯網的進步，語音助理除了在手機方面進行支援外，也在其他裝置進行支援。有著居家助理別稱的它，一講就通、一說就懂、一問就知、一聲就控，不僅貼近了我們的生活，也結合了家庭設備，在智能音箱、車用系統、穿戴式裝置、智慧家庭系統.....等方面也進行支援。

1. 智能音箱

為打造智慧家庭的重要一環，外觀相似於藍牙喇叭，連結手機藍芽後，便能播放音樂、查詢天氣、播放新聞快報等，相關裝置如 Apple 的 HomePod、小米的 AI 音箱、Google 的 Google Nest Mini.....等（圖九）。



圖九：智能音箱

資料來源：Play 智慧家庭。2020 年 2 月 3 日，取自 <https://is.gd/QcbDzV>

2. 車用系統

透過 USB 線與手機連結後，車載裝置的螢幕便能顯示手機內的各種資訊，隨口一聲便可喚醒語音助理，輕鬆設定導航、接聽電話、調整空調等（圖十），例如：在行駛間根據環境的現況給予建議，為駕駛人排除需額外手動的行動所造成的風險，兼具了安全性與便利性。

AI 語音復仇者聯盟

- (1) 安全性：警示駕駛者以產生相對應的動作，如自動減速。
- (2) 便利性：駕駛過程中，只需動口向語音助理發送指令，便可完成多項需求。



圖十：車用系統

資料來源：聯合新聞網。2019年8月31日，取自 <https://is.gd/oH0zrc>

3. 穿戴式裝置

在通訊介面將藍牙與手機連接，以發揮更大的應用空間，市面上常見的穿戴裝置如 Apple Watch、小米手環（圖十一），其裝置服務包含休閒娛樂、訊息傳遞、健康管理與照護……等。

- (1) 休閒娛樂：如播放音樂、遊戲應用、訊息傳遞。
- (2) 訊息傳遞：提供重要資訊，如天氣、手機簡訊、來電顯示等功能。
- (3) 健康管理與照護：記錄生理資訊，如睡眠、跑步、心率，並藉由數據統計及長期紀錄，進而達到健康照護。



圖十二：穿戴式裝置

資料來源：WONGCW 網誌。2019年8月13日，取自 <https://is.gd/Qa9Bb8>

（四）智慧家庭系統

一套具有 AI 技術的智能家電，就像新世代的环境控制系统，結合了生活習慣和家中的各項設備，例如：台灣首套整合了 Amazon ECHO 語音聲控系统，經由聲音的傳達便可控制家中智能家電設備，如電視、門禁管制、智慧溫溼度感測器、紅外線遙控器……等（圖十二）。

- (1) 電視：由傳統的手控遙控器演變為人聲控制
- (2) 門禁管制：結合悠遊卡刷卡系統作為保全設定與解除。
- (3) 智慧溫濕度感測器、紅外線遙控器：可監視區域的溫度濕度，與空調設備連動並遠端 APP 控制空調或時間排程。



圖十二：智慧家庭系統

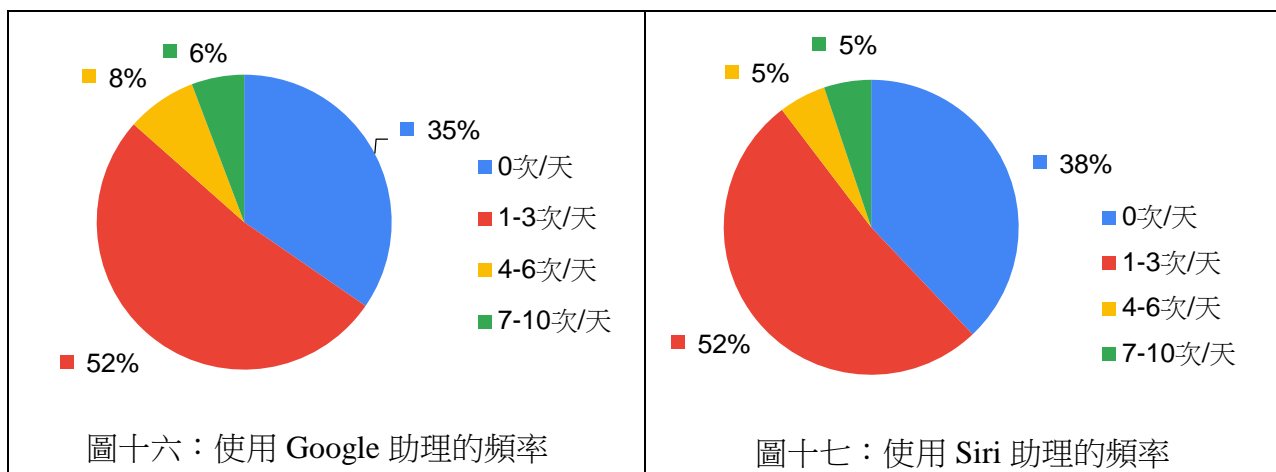
資料來源：聯合新聞網。2019年10月18日，取自 <https://is.gd/qFNIhc>

四、問卷分析與統整

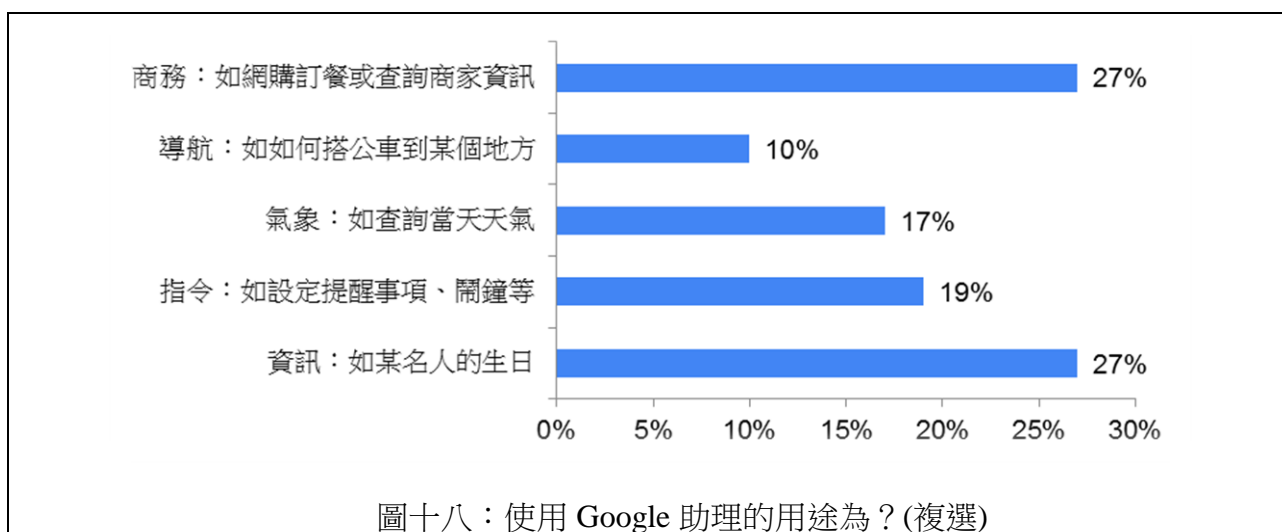
本問卷發放時間為 2020 年 4 月 14 日至 4 月 16 日，採網路問卷填答方式，共發放 139 份問卷，其中有效問卷為 139 份、無效問卷為 0 份，問卷分析結果如下：

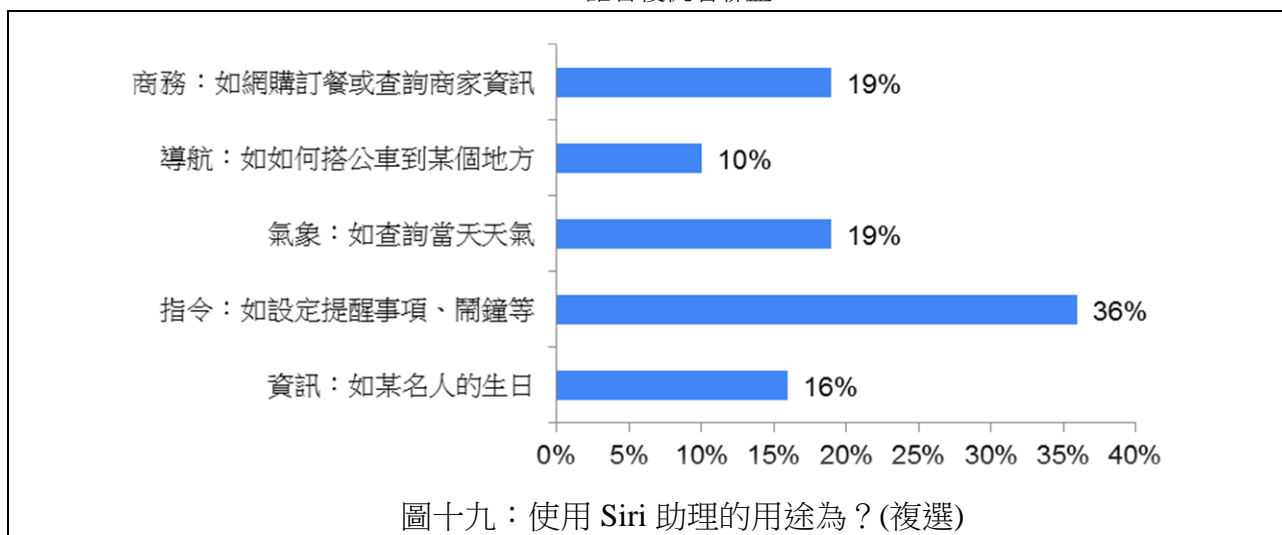
表三：受訪者基本資料及使用之語音助理

| 受訪者性別 | | 受訪者年齡 | | 使用的語音助理 | |
|-------|-----|---------|-----|-----------|-----|
| 女性 | 66% | 18 歲以下 | 76% | Siri 助理 | 42% |
| | | 19-30 歲 | 12% | | |
| | | 31-44 歲 | 4% | Google 助理 | 37% |
| 男性 | 34% | 45-59 歲 | 7% | 無 | 21% |
| | | 60 歲以上 | 1% | | |

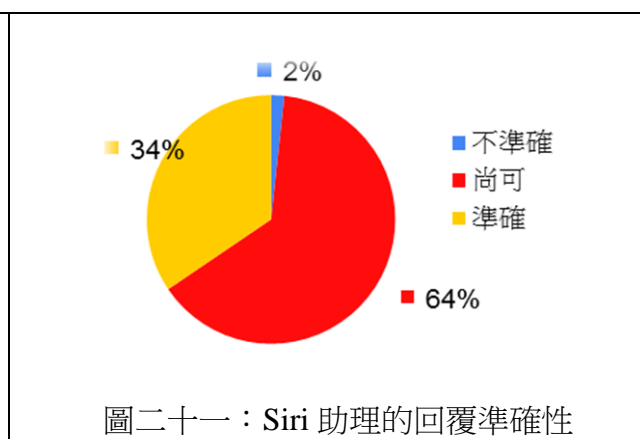
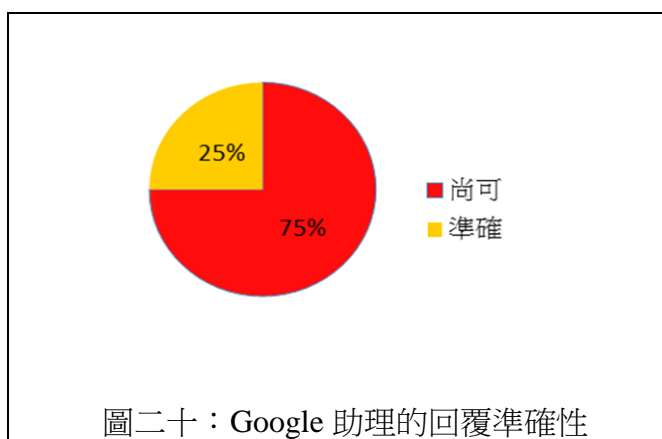


由圖十六、圖十七可得知，受訪者使用 Google 助理與 Siri 助理的頻率，大多以一天平均 1~3 次居多，皆為 52% 的比例，而一天 0 次的比例皆超過三成。此結果說明超過五成的使用者每天都會使用語音助理，也有三成的受訪者雖然有在使用，但並非每天都會使用語音助理。

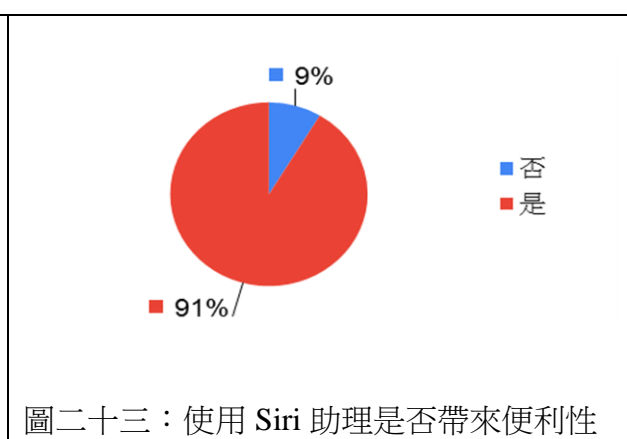
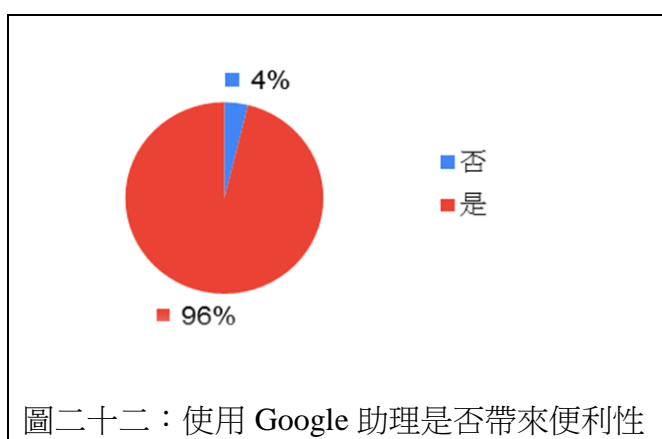




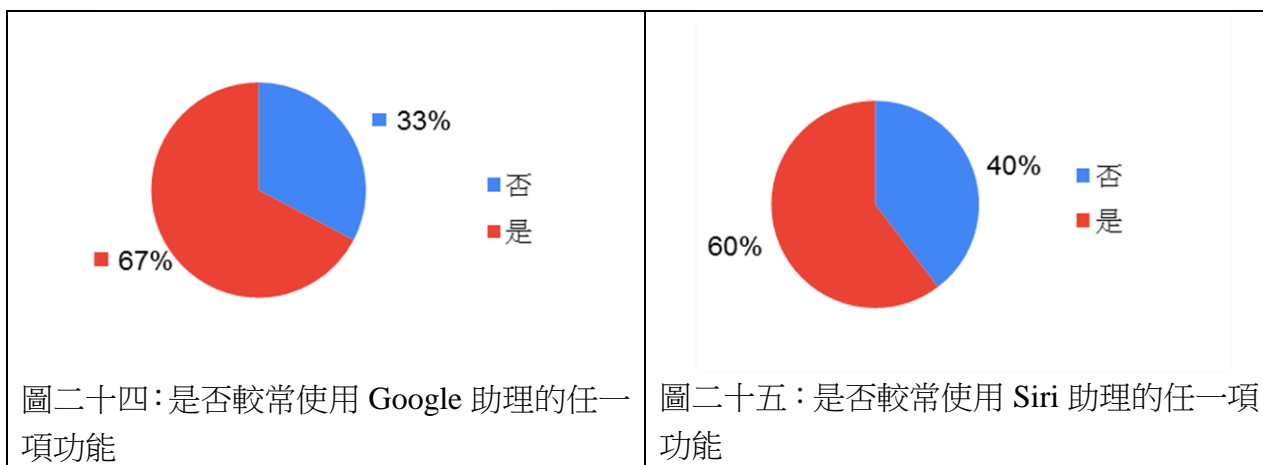
由圖十八可知，受訪者使用 Google 助理的用途大多數以查詢資訊及商務居多，比例各佔 27%；由圖十九可知，受訪者使用 Siri 助理的用途多為設定提醒事項佔 36%，其次為查詢氣象和商務，兩者各佔 19%。此結果說明使用 Google 助理的受訪者大多以查詢商務資訊；而使用 Siri 助理的受訪者，則大多數以設定指令居多。



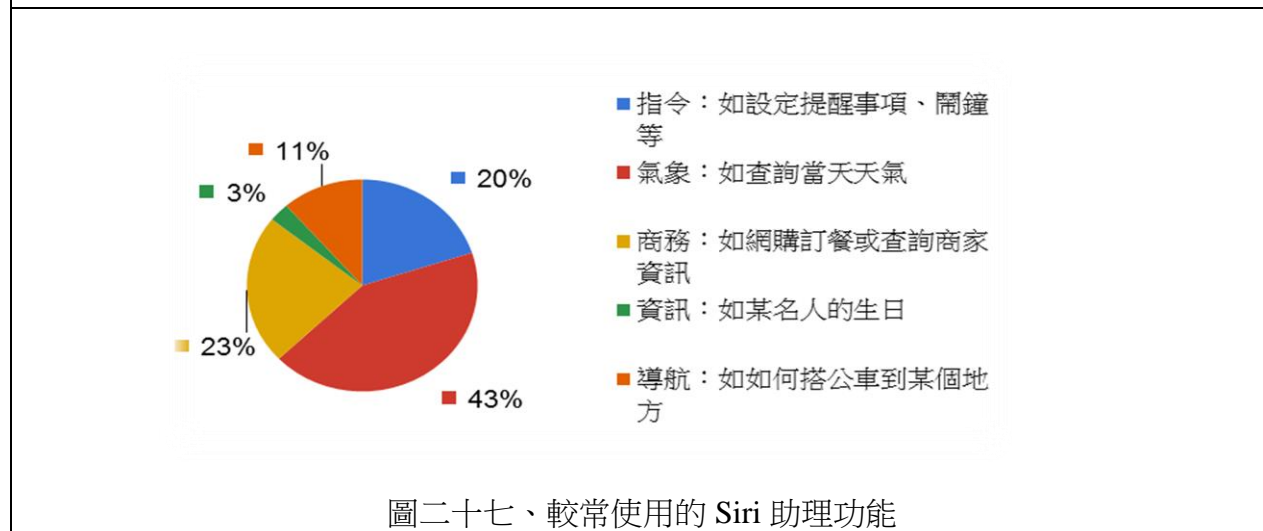
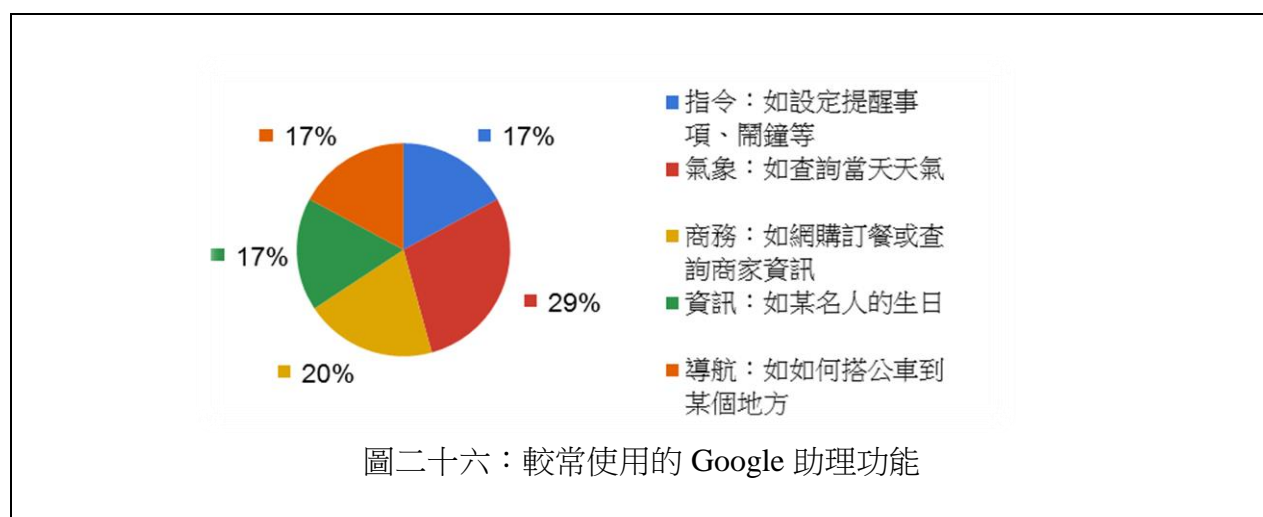
由圖二十、圖二十一可知，受訪者對 Google 助理的回覆準確性認知，有 75% 認為準確；Siri 助理有 64% 認為準確。此結果說明六成以上的使用者皆認為兩大語音助理回復準確性高。



由圖二十二、圖二十三可知，超過九成受訪者皆認為兩大語音助理皆帶來許多便利性。



由圖二十四、圖二十五可知，超過六成的受訪者都有較常使用的任一項功能，顯示使用者有各自喜愛使用的語音助理功能。



由圖二十六、圖二十七可知，受訪者較常使用的兩大語音助理功能皆為查詢氣象，Google 助理佔 29%，Siri 助理佔 43%。此結果說明透過語音助理查詢氣象可立即得知天氣狀況，且取代收看氣象新聞、查詢資訊等動作，故深受多數使用者喜愛。

參、結論

一、結論

每個語音助理皆擅長不同的功能與服務，經本研究發現，Google 助理擅長搜尋資料、Siri 則擅長於手機內建功能的設定，使用者可依個人喜好去做選擇，雖然 Siri 只能在 Apple 產品中使用，Google 助理卻可在任何平台使用，打破 Siri 的限制。隨著人工智慧及機器學習技術的發展進步，語音助理在現今越來越融入我們生活當中，每天手機不離身，語音助理將不再是助理，將來它會是我們的貼身秘書，會是我們的智囊團，不再以被動的方式，而是以主動協助的方式陪伴在你的身邊，早上起床後，語音助理會主動告知今天的工作事項，只要說幾點要出門，它能在出門前幫你設定好車子導航，告知你天氣如何，在路途中已跟星巴克點了一杯屬於你的口味的咖啡，現實活動與日常習慣之間就能通過語音命令實現融合，更為方便和智能。

二、建議

在問卷分析中，以回復準確性而言，雖有六成以上的使用者皆認為兩大語音助理回復準確性高，但在接收語音辨識方面仍可加以改善，當使用者與語音助理進行互動時，若聲源距麥克風較遠，且環境中存在大量噪音，多徑反射和混響，收音狀況不佳，會影響語音識別率。(廖專崇，2017) 由於一般智慧型手機大多採用單麥克風設計，而導致收音效果不佳，因此針對回復準確性提出建議，可透過麥克風陣列技術解決此問題，採設置兩組以上的麥克風對音訊進行偵測，將噪音環境中的特定聲音訊號增加，改善大多數手機只有單麥克風的問題，以提升語音辨識的準確度，使之更臻完善。

肆、引註資料

1. 維基百科。2020 年 4 月 17 日，取自 <https://is.gd/52E5k8>
2. 維基百科。2019 年 1 月 29 日，取自 <https://is.gd/vhUJuZ>
3. Play 智慧家庭。2020 年 2 月 3 日，取自 <https://is.gd/QcbDzV>
4. 黃俐嘉 (2019)。ŠKODA 也可以使用無線 Apple CarPlay 了！但不是每款車都有。2019 年 8 月 31 日，取自 <https://is.gd/oH0zrc>
5. WONGCW 網誌。2019 年 8 月 13 日，取自 <https://is.gd/Qa9Bb8>
6. Play 智慧家庭。2019 年 10 月 18 日，取自 <https://is.gd/qFNlhc>
7. 廖專崇 (2017)。聰明動口不動手，語音助理變身智慧家庭中樞。新通訊元件雜誌，196，7