

投稿類別：資訊類

篇名：

AI 人工智慧中語音系統的發展與利弊

作者：

簡敏靜。新北市立秀峰高中。普通科一年二班。

指導老師：

吳潤宗老師

壹、前言

一、研究動機

打從一開始，我就對 AI 人工智慧中語音功能的發展和利弊影響有著濃厚的興趣，在現代科技發達、日新月異的新世代，AI 人工智慧成為我們優秀且方便的工具，不再是以手機鍵盤打字，而是利用「語音」去操縱生活的大小事，但也因為此項技術的產生，造成部分就職人員失業等，正因如此，我想對此做更深入的研究。

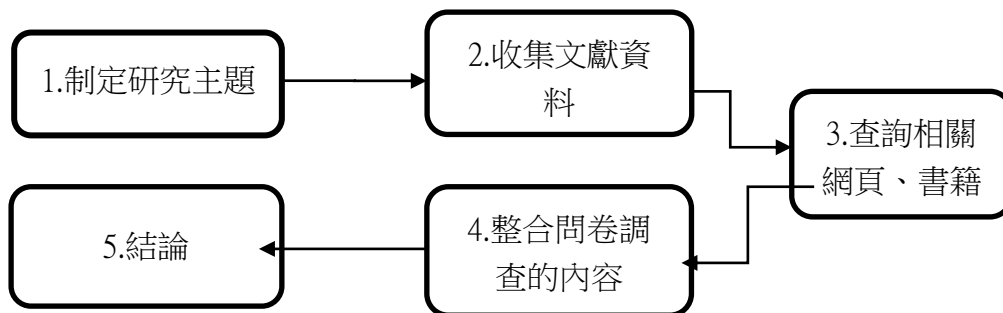
二、研究目的

- (一) 調查 AI 語音系統的使用頻率及滿意度。
- (二) 分析 AI 語音系統帶來的利弊與影響。
- (三) 探討 AI 語音系統未來的發展。

三、研究方法

- (一) 閱讀相關書籍與期刊。
- (二) 查找與此議題相關的網頁。
- (三) 製作問卷調查並分析。

四、研究流程



(圖一資料來源：研究者繪製)

貳、正文

一、AI 人工智慧

(一) AI 的定義與其應用之處

AI (原名 Artificial Intelligence) 又稱作「人工智能」，是利用人類擁有的智慧再加上人工方式所形成的技術，讓機器具備和人類一樣的思考邏輯及行為模式，並發展出推理、學習、翻譯、聲音與視覺辨識……等相關智能。近幾年來，人類把 AI 發揮得淋漓盡致，還須解決技術及制度的課題。AI 在日常生活中的應用如下：

- 1、語音助理（辨識功能）
- 2、社群的廣告放映
- 3、影像辨識、影音網站（例如 YouTube 推薦影片、Spotify 精選）
- 4、Google Map 路線規劃、導航系統
- 5、自然語言處理（例：翻譯、聊天機器人）

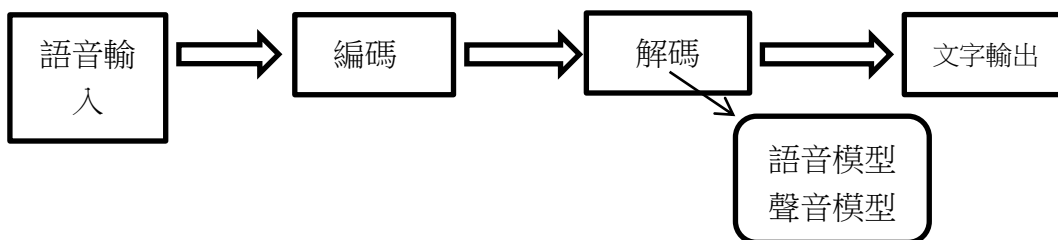
二、AI 語音系統

(一)功能

AI 語音智能以人的對話方式提供需要的服務與信息。目前透過 Google Assistant、亞馬遜……等各具代表性的語音助理，各大科技已積極研發到蘋果的 Siri，在各行各業方面，語音助理都能被廣泛使用。語音智能主要分為五個模塊：

1、語音識別（ASR）

從聲音轉換為文本的過程，或是將用戶的聲音轉化為語音。通常用來辨識說話者的身分。

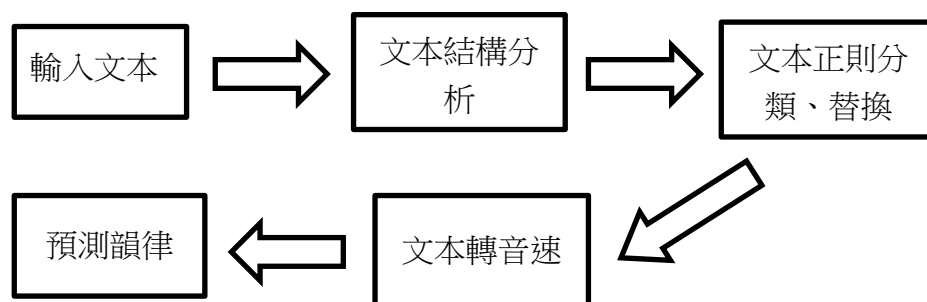


(圖二資料來源：頭條@壹點方圖) 語音識別原理流程。人工智慧之語音識別(ASR)。2019 年 8 月 5 日。取自 <https://twgreatdaily.com/OxNbYmwBUcHTFCnf42R2.html>

2、語音合成（TTS）

以文本轉換為語音，將輸入的文句信息或儲存於電腦中的文件資料模擬人聲，轉變為流

利的口語技術。在應用層面上大多在電腦螢幕上文章的閱讀，以語音指引的方式輔助、補充說明，但在發音方面衍伸出許多問題，例如：清晰流暢度不足、注音破音字處理不完整、韻律節奏錯誤無法正確傳遞訊息……等，導致人們認為「AI 的語音助手不像人的道理」。



(圖三資料來源：典型的前端處理流程) 語音合成處理流程。為什麼你的語音助手說話不像人？關於 TTS 技術原理及挑戰。2019 年 8 月 7 日。作者：鄭杰文。取自 <https://www.leiphone.com/news/201908/rNnRNQFag7slHsD0.html>

3、自然語言理解（NLU）

在語音與文字的相關應用方面幾乎都會用到 NLU，NLU 指機器像人類一樣具備語言理解的能力，讓機器用不同語言表達，學會判斷句型的含義，從文本的語義解析中找出關鍵詞，並進行實體識別、意圖識別。

以往的機器只能處理「結構化數據」，換句話說就是要用戶輸入確切的指引，再加上過度依賴死板的關鍵數據，導致無法被有效利用，為了避免這種情況、提升機器的語言程度，就須透過訓練強化。

優點	缺點
一、機器翻譯 （運用時，須讓機器理解上下文） 二、機器客服 （理解對話內容） 三、智能音箱 （理解用戶意圖，須以對話方式）	一、語言多樣性 （自然語言可靈活運用、無規律，字詞組合可表示很多含義） 二、語言對知識的依賴 （語言與知識連結） 三、語言的上下文 （對話的環境不受限制，上下文無法聯繫導致意義有所差別）

(圖四資料來源：研究者繪製)

4、自然語言生成（NLG）

各企業為了評估數據量、降低成本，發展了新的人工智慧應用：「NLG 自然語言生成」，將非語言的格式或數據轉換成人類可以理解的語言。分為六個步驟：

(1)內容確定 Content Determination

決定信息是否包含在文本中，在文本中的數據信息會比最終傳達的多。

(2)文本結構 Text Structuring

確定要傳達的信息並合理的組織順序。

(3)句子聚合 Sentence Aggregation

將多個信息合併在同一句，方便閱讀。

(4)語法化 Lexicalisation

以自然語言的基本文法概念加連接詞，使其更完整。

(5)參考表達式生成 Referring Expression Generation 〈簡稱 REG〉

和語法化相似，以 REG 識別內容的領域，並使用該領域詞彙。

(6)語言實現 Linguistic Realisation

確定後，形成結構完整的句子。

5、對話管理（DM）

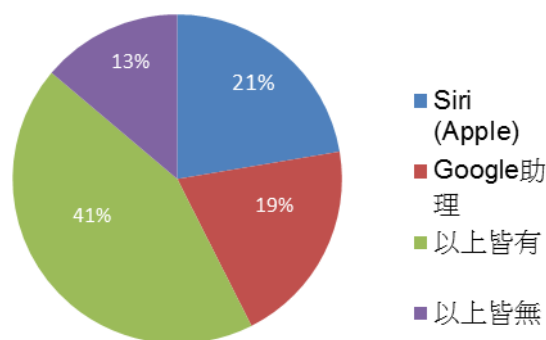
對話管理模塊主要用於確認信息的完整度，以引導式的多輪對話去除模稜兩可的信息，除了要負責維護對話的狀態、儲存用戶與系統的歷史行為，並進行資料庫管理、查詢。

三、AI 語音系統產生的利弊與影響

(一) 日常生活中

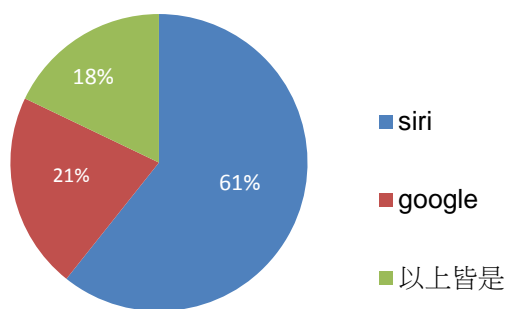
現代科技發達，人人手中的 3C 產品都有 AI 語音系統，最著名的像是 Google 語音助理及 Apple 手機品牌才有的 SIRI 功能。根據針對本校學生的調查，可從以下圖五、圖六兩張圓形圖得知，兩種人工智慧並行居多，以頻率來判斷，兩種語音系統並行頻率較高，從滿意程度來看，SIRI 比 GOOGLE 助理更受歡迎。

使用各語音系統頻率圓形圖



圖五

使用語音系統滿意度圓形圖

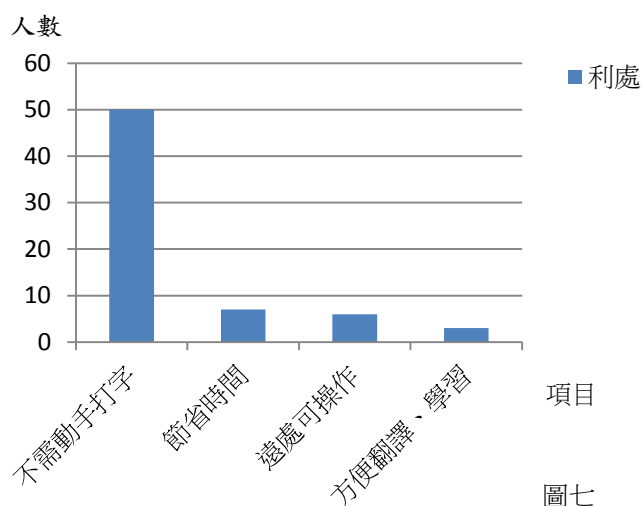


圖六

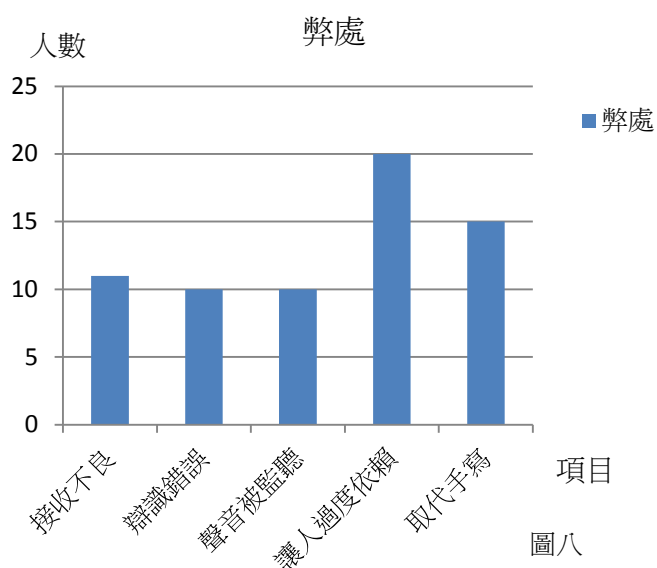
利與弊：

可從以下圖七、圖八兩張長條圖得知，「手寫」是多數人在判斷其利弊影響的關鍵，人們雖認為不須打字即可操作是AI語音系統的優勢，但又不希望過度依賴此功能，甚至取代手寫，因而產生矛盾。

利處



圖七



(二) 金融業中的影響

現代科技發達，各行各業都會利用 AI，其中最具代表性的就是「金融業」。新興科技正推動金融產業的數位化轉型腳步，未來專精特定產業、解決方案或客群，創造服務價值與商業洞察(溫紹群、楊之奇、蔡旻純，2019)。

四、AI 語音系統未來發展

自 2015 年起，AI 語音市場日益擴張，從硬體產品、軟體服務、系統平台等方面發展，各個 AI 語音助理功能不同，在這裡舉三個例子：

(一)互動

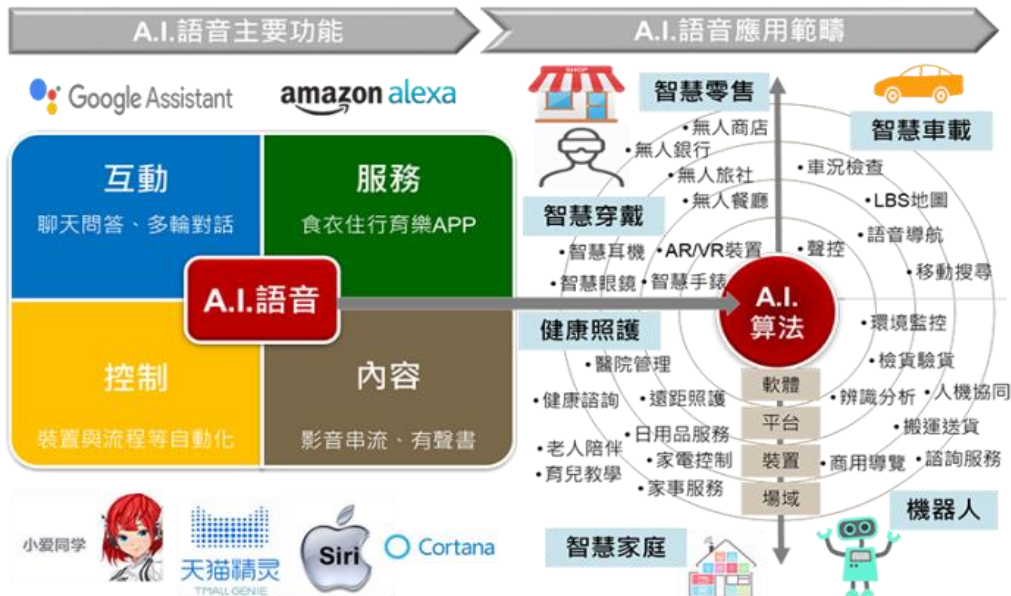
利用聊天問答的方式，但必須涉及到 ASR、TTS、NLU、NLG、DM 等功能技術結合，最終須像人類說的話一樣流暢。

(二)服務

從食、衣、住、行、育、樂等提供任何服務資訊，是所有 AI 語音功能中最方便、快速的部分。

(三)控制

透過人工智慧語音助理來控制智慧電視、掃地機器人、空調系統、照明設備、智慧穿戴、智慧開鎖、智慧監控攝影機等(陳佑怡 2019)。



圖九 人工智慧語音應用服務布局範疇
資料來源：工研院 IEK Consulting(12/2018)



圖十 中國大陸-人工智慧語音技術與產業鏈
資料來源：易觀；工研院 IEK Consulting(7/2018)

參、結論

AI 人工智慧，又稱「人工智能」，指機器具備與人類相同的思維邏輯和行為模式。它的用途廣泛，例如：AI 語音系統。語音系統中的語音識別及語音合成結合後，利用自然語言理解及生成，最後進行對話管理，變成人類聽得懂的語句。

秀峰高中師生 SIRI、Google Assistant 使用頻率與滿意度經調查後，絕大部分都是兩者並用，但在滿意度方面，SIRI 比 Google 略勝一籌。透過 AI 語音系統，不必輸入文字就可方便

傳達訊息，但因此產生的影響有利也有弊。因為當人們習於以語音輸入的同時，文字書寫及拼音的能力也可能逐漸退化。

在 AI 語音系統及功能專業化的時代，我們可從硬體設備及系統平台的改善，得到更順暢的互動、更多元的用途，及更便利的操作方式。

肆、引註資料

作者：落合陽一。譯者：詹慕如（2018 年 12 月 7 日）。

【超 AI 時代的生存戰略】。台北市：三采文化股份有限公司。

作者：李開復（2018 年 7 月）。第一版。台北市：遠見天下文化。

【AI 新世界 中國、矽谷和七巨人如何引領全球發展】

作者：溫紹群、蔡旻純、楊之奇。第七卷 第四期 (108 年 12 月)。

【金融數位轉型的十字路口：策略轉型兩打造安全、便利、開放的支付環境】

MAKERPRO 自造達人社群/媒體/平台。AI 來襲！三分鐘看懂人工智慧

製造日期：2019 年 5 月 31 日 取自 <https://makerpro.cc/2019/05/introduction-to-ai/>

每日頭條。由多智時代發展於科技。2018 年 10 月 16 日。

智能語音助手的工作原理是什麼，目前市面上主要有哪些產品？取自

<https://kknews.cc/zh-tw/tech/zejmm3g.html>

3C 新報。效果提升、語音控制與個人化，AI 電視的應用與革新

作者拓璞產研。發布日期 2019 年 11 月 5 日。取自 <https://ccc.technews.tw/2019/11/05/ai-tv/>

就是教不落，網路大小事。作者 阿湯。2018 年 10 月 27 日。

Google 語音助理 vs.Siri 語音助理功能篇，私人秘書你選誰？取自

<https://steachs.com/archives/42316>

今天頭條。作者 壹點方圓。2019 年 8 月 5 日。

人工智慧之語音識別（ASR）。取自 <https://twgreatdaily.com/OxNbYmwBUcHTFCnf42R2.html>

語音合成 Speech Synthesize Note。取自 <https://irw.ncut.edu.tw/peterju/speech.html>

雷鋒網 讀懂智能&未來。作者 鄭杰文。2019 年 8 月 7 日。
為什麼你的語音助手不像人？關於 TTS 技術原理及挑戰。取自
<https://www.leiphone.com/news/201908/rNnRNQFag7slHsD0.html>

每日頭條。由九三智能控發展於資訊。2018 年 7 月 13 日。
語音合成（TTS）技術原理簡介：如何一步步將文字變成語音。取自
<https://kknews.cc/zh-tw/news/6o2494p.html>

每日頭條。由產品經理的 AI 知識庫發展於科技。2019 年 9 月 16 日。
一文看懂自然語言理解—NLU（基本概念+實際應用+3 種實現方式）取自
<https://kknews.cc/zh-tw/tech/99qbzk8.html>

作者 Estelle Huang。2017 年 12 月 22 日。Rule-based vs.NLU：聊天機器人何聽懂人類的自然語言？取自
<https://medium.com/@estelle.husky/rule-based-vs-nlu-%E8%81%8A%E5%A4%A9%E6%A9%9F%E5%99%A8%E4%BA%BA%E5%A6%82%E4%BD%95%E8%81%BD%E6%87%82%E4%BA%BA%E9%A1%9E%E7%9A%84%E8%87%AA%E7%84%B6%E8%AA%9E%E8%A8%80-17065de49a>

EET Taiwan 高通公司。2018 年 4 月 16 日。個人語音助理時代已經來臨！取自
<https://www.eettaiwan.com/news/article/20180416NT31-season-voice-based-personal-assistants>

每日頭條。2019 年 8 月 1 日。由產品經理的 AI 知識庫發表於科技。
一文看懂自然語言生成 - NLG（6 個實現步驟+3 個典型應用）取自
<https://kknews.cc/zh-tw/tech/oepmm8m.html>

新通訊元件雜誌。2019 年 1 月 15 日。迎接人機介面新商機 掌握全球 AI 語音助理下一步。作者 陳右怡。取自
<https://www.2cm.com.tw/2cm/zh-tw/tech/B73A5BE5A4954D5988102C7D8A5B8A68?type=26>

附錄：問卷內容

一、下列 AI 語音系統你用過何種？

(A) Siri(APPLE 手機) (B) Google 語音助理 (C) 以上皆是 (D)以上皆無

二、承第一題 (假如你的答案為 C)

() 你較喜歡使用何種語音系統？(A) Siri 語音系統 (B) Google 語音系統
原因為何？

三、承第一題 (假如你的答案為 A 或 B 或 C 其中之一)

請寫下你都利用於何處？(例如:搜尋、娛樂……等)

四、承第一題 (假如你的答案為 D) 不使用的原因為何？

五、() 一天使用語音系統大約幾次？

(A) 幾乎不用 (B) 一 (C) 二 (D) 三次 (E) 超過三次

六、語音系統是個必要的功能嗎？為什麼？

七、使用語音功能可帶來哪些便利性？(至少寫出一個)

八、語音系統帶來的利處有哪些？(寫出一個例子)

九、語音系統帶來的弊處有哪些？(寫出一個例子)

十、請寫出你對語音系統未來發展的看法 (至少寫滿二十字)

製作人：102 簡敏靜