

## 科技部人文社會科學研究中心

### 補助科技部跨領域研究計畫之前置規劃案結案報告

臺灣 AI 右腦的「情感智能」建構計畫：

臉部表情、語言文字、聲音表情、肢體動作整合的即時  
情緒辨識系統 & 高情緒智商人工智慧的資安與法律

規劃案編號：MOST 107-2420-H-002-007-MY3-PA10909

規劃案執行期間：109 年 9 月 1 日至 109 年 12 月 31 日

執行機構及系所：國立臺灣大學戲劇學系暨研究所

計畫召集人：朱靜美（國立臺灣大學戲劇學系暨研究所 教授，

臺大人工智慧與機器人研究中心「前瞻機器人組」

& 「認知與行為科學組」 成員）

共同召集人：

古倫維 （中央研究院資訊科學研究所 副研究員）

包蒼龍 （大同大學資訊工程學系 教授）

劉靜怡 （國立臺灣大學國家發展研究所 教授）

謝禎罔 （大同大學資訊工程學系 教授）

計畫參與人員：

曹昱      （中央研究院資訊科技創新研究中心研究員）

張兆恬   （國立臺灣大學國家發展研究所 副教授）

科技部人文社會科學研究中心

中 華 民 國 110 年 1 月 30 日

## 目錄

壹、前言	1
貳、研究目的	2
參、文獻探討	3
肆、研究方法	8
伍、結果與討論	9
一、結論	9
二、建議	14
三、跨領域合作之經驗	14
四、獲補助之收穫與執行上遭遇之困難	15
五、結語	16
參考文獻：	16
陸、附件	21

## 中文摘要

人工智慧 (Artificial Intelligence) 是未來世界各國大廠的兵家必爭之地，具有無限的市場商機與影響力。其中，是否具有代表右腦領域的「情感運算」(Affective Computing) 能力，將左右 AI 是否能夠自然地走入人類的生活之中，提供更佳的服務、照護、陪伴、教育等各式功能，創造現今人力受限而無法企及的良好生活品質。然而要讓 AI 能精準辨識情緒與恰當表情達意是一項跨眾多領域的複雜研究，尚有許多待突破的困境與技術難點。

由於人類情感交流的四大主要媒介乃是「臉部表情」、「語言文字」、「聲音表情」與「肢體動作」，光是要讓電腦掌握這四種領域的個別項目就已經是複雜精細的工程，綜合這四種領域的多模態訊號處理 (multimodal signal processing) 更是浩大的工程，而資工領域學者在此所面臨的關鍵痛點便是：缺乏精準的大數據資料讓 AI 進行深度學習。然若 AI 無法整合這四種元素，那便無法精準辨識使用者的情緒，也就遑論後續如前所述的美好願景。此外，當 AI 更深入地介入人類生活並且蒐集更多使用者資料，這些涉及個人隱私資料的資訊安全如何加以保護，避免落入有心人士的手中。是故，如何透過法律的限制，規範明確的使用界線與門檻，都是與情感運算亦步亦趨的重要議題。因此，本研究計畫便聚焦於前述兩大重點領域有以下三大核心研究議題：

- (一) 臉部表情、語言文字、聲音表情、肢體動作的情感大數據資料庫建造；
- (二) 多模態訊號處理的即時整合情緒辨識應用系統與運算的後台建構；
- (三) 情感運算 AI 的資安維護與法律探討。

在本前置規劃案中，各子計畫主持人針對國內外學術文獻進行統整與分析，在團隊研究會議中交流各領域最新的產學界研究現況，並剖析自身下一步技術突破的切入點，同時在會議中邀請外部專家學者分享產業界的痛點與趨勢；也透過田野調查訪問目標市場的使用者蒐集真實用戶需求，挖掘未來研究應用的可能性，也探討 AI 與使用者可能面臨的資安隱憂與法律破口。最後以此基礎歸納聚焦出未來三年期整合型研究計劃的明確方向與六項子計畫間如何密切互助，以期能打造出精準又安全的友善情感 AI (Friendly Affective AI)，奠定臺灣在國際情感 AI 產業的領先優勢。

關鍵字：人工智慧、情感運算、多模態訊號處理、友善情感 AI

## Abstract

Artificial Intelligence (AI), with unlimited market opportunities and influence, is the future battleground for major companies around the world. Among them, the ability of Affective Computing, which represents the right brain faculty, will determine whether AI can naturally enter the human life and provide better services, care, companionship, education, and other functions, creating a good quality of life that is beyond the reach of current human capabilities. However, to enable AI to precisely recognize emotions and express intentions appropriately is a complex research across many fields, and there are still many constraints and technical limitations awaiting a breakthrough.

Since the four main mediums of human emotional communications are: "facial expressions," "languages," "voice expressions" and "body movements," merely enabling the computer to master individual items in these four fields is already a complicated and intricate project, and to integrate the multimodal signal processing in these four fields is an even bigger project. The key pain points faced by information engineering academics is the lack of accurate big data for AI to perform deep learning. However, if AI cannot integrate these four areas, it will not be able to accurately recognize the user's emotions, let alone to achieve the wonderful vision to follow as described above. In addition, as AI becomes more deeply involved in human life and collects more user data, the question arises as to how security of such information involving personal private data be protected from falling into the hands of those with ulterior motives? Therefore, it is important to establish clear boundaries and thresholds through legal restrictions for the use of affective computing. Hence, this research project focuses on the following three core research questions in the aforementioned two key areas:

(1) Construction of a large databank of emotional data for facial expressions, languages, voice expressions, and body movements.

(2) Backend construction for real-time integrated emotional recognition application system and computing with multimodal signal processing.

(3) Information security maintenance and legal matters of affective computing AI.

In this pre-project, each sub-project facilitator will consolidate and analyze the academic literature from home and abroad, exchange the latest research status of the

industry and academia in each field in team research meetings, investigate the entry point of their next technological breakthrough, and invite external experts and scholars to share the pain points and trends in the industry in the meeting. Also, through field surveys users in target markets are interviewed to collect real user needs, explore the possibilities of future research applications, as well as look into the potential security concerns and legal issues that AI and users may have to deal with. Finally, based on such foundation, to conclude and focus on a clear direction for the integrated research program in the next three years and how the six sub-projects will closely work with each other to create accurate and safe Friendly Affective AI, which will establish Taiwan's leading position in the international affective AI industry.

Keywords: artificial intelligence, affective computing, multimodal signal processing, friendly affective AI

科技部人文社會科學研究中心