

人工智慧運用在諮商領域中的倫理議題： 以聊天機器人為例

Ethical Issues in the Application of Artificial Intelligence in Counseling: A Case Study of Chatbots

甘翠瑜¹、邱采芊²、張聖婷³、李明峰⁴

Cui-Yu Gan¹, Tsai-Chien Chiu², Sheng-Ting Zhang³, Ming-Feng Li⁴

摘要

本研究探討人工智慧於諮商領域中的應用，特別聚焦於聊天機器人的角色與可能引發的倫理議題。隨著生成式人工智慧技術逐漸普及並滲透至心理健康照護，相關研究指出它能夠提供即時性、低成本與可近性的支持服務，對於資源不足的族群而言具有潛在助益。然而，與此同時也帶來諸多挑戰，包括缺乏專業倫理規範、隱私與資料安全的風險，以及在助人關係中真實性與同理心的限制，對比傳統心理師的專業訓練與倫理規範，人工智慧在回應人類情緒時顯得較公式化，可能削弱受助者對「被理解」與「被支持」的感受。本文透過文獻回顧與案例分析，整理人工智慧聊天機器人在心理諮商中的應用現況與限制，並探討其引發的專業倫理困境，例如：當事人權利、公平待遇權、免受傷害權及保密議題。進一步地，研究也提出心理師在面對科技介入時的反思，包括實務現場如何更審慎的運用人工智慧、在諮商心理師培養過程與倫理培訓中納入相關議題，以及透過跨領域對話建立更完整的規範機制。期望此研究能為諮商輔導領域帶來理論與實務的啟發，協助專業人員在面對科技快速發展的同時，能與之合作並維持專業價值與助人工作的核心精神。

關鍵詞：人工智慧、聊天機器人、諮商倫理、心理健康、科技輔助

¹⁻³ 國立臺南大學諮商與輔導學系 碩士生

⁴ 國立臺南大學諮商與輔導學系 助理教授

通訊作者：邱采芊，（700）臺南中西區樹林街二段33號，國立臺南大學諮商與輔導學系，
E-mail: enyachiu31@gmail.com



壹、前言

隨著科技產品逐漸成為我們生活中重要的部分，科技的發展影響各產業的環境及研究重點，心理專業領域也無可避免地受到科技革新的影響。身為助人工作者須隨時跟上世界潮流，關注科技涉入人們生活的程度、探討利用科技協助心理諮商進行及推廣的方式。近年有許多研究在探討電腦科技及人工智慧能如何協助及分擔心理治療中的工作，在專業人員的訓練養成中，也有文章探討加入電腦模擬個案來評量實務能力之可行性。

面對科技在心理領域的崛起，有不少專業工作者會擔憂其對現況的衝擊，擔心其限制及效能；然而也有支持者抱持樂觀的態度，認為科技技術能提升大眾的心理健康，減輕人力負擔及降低工作倦怠的發生。因此，科技潮流的無法控制及抵擋，身為專業人員能做的是充實自身的知識，反思此項改變帶來的影響及對諮商教育的養成，在變革的不穩定中找到能掌控的確定性。故本文欲探討人工智慧在心理諮商中的運用，聚焦於聊天機器人這項產品及其技術，最後整理於實務工作中使用人工智慧可能會衍生之倫理議題，期望回歸到實務現場若面臨人工智慧的輔助，能知其效能及挑戰，更加有意識的使用科技。

貳、人工智慧在諮商中的運用

本文中探討的諮商為個別諮商，個別諮商的定義如下：諮商是助人的形式之一，受過訓練的諮商師與案主間所建立的個人對個人的專業關係，特色是在諮商過程中應用心理學的理论與相關溝通技能，來陪伴案主處理個人隱私、問

題及渴望（李茂興等譯，2002）。

人工智慧（artificial intelligence, AI）是指透過電腦程式及機器設備，模仿並展現人類智慧的技術。現今已有一些特定的AI形式融入我們的日常生活，例如與攝像頭配合使用來辨識人臉、將語音轉換成文字的語音辨識系統、住宅的科技使用，或是追蹤檢測生理狀態等。AI領域中的自然語言處理（natural language processing, NLP）的進步，使語言特徵及心理健康的相關研究逐漸興起（D'Alfonso, 2020）。在生活的應用上，較常見的範圍從社交軟體如Facebook及Dcard上針對自殺的敏感字詞的偵測，到近期興起的ChatGPT等對話系統進行來回的交談。

國外於20世紀便陸續有關於電腦應用於諮商治療領域的討論，目前較常見的有以下類別（引自Helgadottir, 2019）：

一、以網路為基礎提供的諮商服務 （Internet-supported interventions）

包含通訊諮商（Web-based therapy）、自助資源（Self-guided treatments）、治療師協助的自助服務（Therapist assisted online self-help）。

二、線上虛擬治療師（Computerized therapy，亦為本文所指之聊天機器人）

（一）目前應用情形

近年來人工智慧大幅發展，最廣為人知的聊天機器人即是ChatGPT，以下聊天機器人即以此模型稱之，本機器人是基於大型語言模型（GPT，即生成式訓練變換器）的聊天機器人，由OpenAI開發。



（二）特色

特色是透過生成文字來模擬人類般的對話，以回應使用者的輸入ChatGPT經過大量數據的訓練，使其能夠理解並生成各種主題和風格的文字。

（三）治療方式及內容

ChatGPT在心理治療領域被認為具有潛力，儘管伴隨著重要的倫理與實踐考量，已在多項研究中被評估為有益的工具：

1. 篩選與診斷：它可以協助進行初步的篩選和診斷，例如它能針對患者描述的症狀提供準確的診斷建議，如抑鬱症、廣泛性焦慮症和社交焦慮症，並具備對心理病理學相關理論的理解。

2. 提供治療資訊與建議：它能提供治療選項和方法，雖然在未明確要求下傾向於認知行為療法（CBT），但若有明確提示，也能提供其他治療取向的解釋，例如語意治療與存在分析。

3. 解釋患者材料：它可以對患者的材料（如夢境）進行解釋，並能從不同的心理治療學派（如心理動力學、系統學、存在分析、個體心理學）提供詮釋，這對於已具心理治療經驗的患者來說，是個獨自處理材料的好機會。

4. 提供治療技巧與實踐指導：能夠為治療師提供建議的治療技巧。

5. 學習其他治療方法：協助治療師了解不同的心理治療方法。

三、治療性機器人（Therapeutic robots）

（一）目前運用情形

此處的治療性機器人者的是運用在醫療照顧領域，尤其是高齡與失智症照顧，以陪伴患者為主，以配有感測器、攝影機、麥克風及電腦視覺等技術，目

的在與人類進行社交行為並互動。

（二）特色

1. 多樣化的外觀與感知能力：社交機器人可以是具備人類面孔的固定式平台，也可以是毛茸茸、會搖尾巴的小狗形狀。它們配備有感測器、攝影機、麥克風及電腦視覺等技術，使其能與人類或其他機器人更好地互動與交流。

2. 自然語言溝通：設計宗旨是透過生成文字來模擬人類對話，以回應使用者的輸入。它們旨在進行社交行為，例如陪伴、跟隨、引導、協助使用者，甚至參與對話。

3. 人形外觀的優勢：人形外觀能夠提升機器人的接受度及感知有用性。對使用者，特別是失智症患者而言，這種熟悉且無害的外觀有助於減少混淆和恐懼。然而，設計上需在「真實感」與「舒適度」之間取得平衡，以避免「恐怖谷」效應，即機器人過於像人卻又不完全像人時，可能引起不適感。

（三）治療方式

1. 日常任務輔助：可協助管理日常事務，例如提供藥物提醒、管理例行行程（如安排活動、約會、個人護理）。

2. 緊急情況偵測：具備緊急情況偵測能力，如跌倒偵測、健康監測，並能即時通知照護者或緊急服務，確保即時反應並降低受傷風險。

3. 認知刺激：能夠參與刺激性活動，如記憶遊戲或簡單對話，有助延緩認知能力下降。

4. 減輕重複性任務：社交機器人能夠處理重複性任務，從而減輕照護者的工作量與壓力，讓他們有更多時間專注於個人化及更複雜的照護需求。

（四）內涵

1. 社交機器人能處理重複性任務，



並提供結構和一致性的活動，讓照護者能更專注於個人化和人性化的互動，減輕他們的身心負擔和壓力。

2. 它們透過持續的陪伴和刺激性活動，顯著減少失智症患者的孤獨感和隔離感，並有助於減緩認知能力下降。

四、虛擬實境治療 (Virtual Reality Therapy)

(一) 目前運用情形

1. 廣泛應用於焦慮症：VR治療對焦慮症（包括特定恐懼症、創傷後壓力症候群、社交焦慮症、恐慌症）具有明確的療效。

2. 癌症患者的心理支持：

(1) 顯著改善乳癌患者在首次化療前的預期性焦慮、憂鬱和情緒困擾。

(2) 管理癌症相關的不適症狀，如疲勞、焦慮和恐懼。

(3) 降低醫療過程中的疼痛感知，且比藥物治療更有效和安全。

(4) 有助於癌症康復和提高生活品質。

(二) 特色

1. 非藥物性治療：VR是一種非藥物性的介入方式，可用於處理醫療程序引起的疼痛、焦慮或恐懼。

2. 高度沉浸與真實感：它能創造三維環境，提供廣泛、沉浸式且生動的現實幻象。透過多感官輸入（如立體視覺、頭部運動調整的視覺顯示、立體聲、觸覺反饋），讓使用者感覺身歷其境。

3. 安全可控的環境：VR能夠模擬環境或特定情境，讓患者在「安全環境」中體驗「如同真實」的感覺，並從中學習情緒控制。

4. 即時回饋與調整：VR介入能提供患者即時的表現回饋，並根據個人需求

調整難度。治療師可以控制暴露的時間和強度，並根據患者的進展或反應進行調整。

5. 多功能性與客製化：VR的靈活性使其適用於各種心理介入，並能根據每位患者的需求進行調整。

6. 克服實際障礙：VR能輕鬆呈現難以、昂貴、危險甚至不可能在現實中重現的恐懼刺激和情境（如戰鬥、恐怖主義、飛行），簡化暴露療法的規劃和執行。它還有助於患者在治療之外進行暴露練習，促進技能成長。

(三) 治療方式

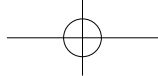
1. 模擬化療過程：透過VR，患者利用放鬆、專注和正念策略來管理焦慮和恐懼。完整的介入方案包含三個模組，模組分別為：預期 (Anticipation)、分散注意 (Distraction)、正念 (Mindfulness)，每個患者會在四次個別療程中重複這些模組。

2. 治療師控制與監測：治療師可以選擇特定的暴露情境、調整強度和時間，並選擇是否觸發特定刺激（如飛機顛簸），平台還能收集患者的主觀焦慮評分，並追蹤眼球運動，以監測患者是否試圖迴避恐懼刺激 (Wray et al., 2023)。

(四) 內涵

虛擬實境像是一位經驗豐富的導遊，帶領體驗者進入一個「排練場」，是一個安全且可控的模擬世界，在這個世界裡面，患者可以多次面對自身最深的恐懼，學習如何掌握自己的情緒，並將這些習得的技能應用到真實生活中，從而練習減輕焦慮、提高面對困難的信心。

五、電子遊戲治療 (Video game therapy)



(一) 目前運用情形

1. 醫療復健領域：VGT用於慢性腦外傷患者的平衡訓練與注意力提升。

2. 心理健康領域的潛力：雖然VGT在焦慮、情緒障礙等應用尚未如VR般普及，但已有初步研究顯示其可作為輔助工具。

(二) 特色

1. 互動性高：結合體感裝置（如Xbox Kinect），遊戲設計能促使病患主動參與訓練。

2. 提升動機：相較傳統復健，VGT提供更高的趣味性，能提高使用者持續參與的意願。

3. 可調整性強：依照病患狀況可調整遊戲難度與挑戰性。

4. 雙重效益：部分研究顯示，VGT同時能改善身體控制與認知功能（如注意力、反應時間等）。

(三) 治療方式及介入內涵

1. 使用裝置：Xbox Kinect、Wii Fit等體感互動平台。

2. 遊戲類型：以動作導向、平衡訓練為主，如跳躍、閃避、重心轉移等活動。

3. 治療頻率：研究中常見為每週2~3次、每次30~60分鐘、持續6週以上。

4. 搭配測評工具：身體層面會使用以下工具：CB&M、TUG、Unified Balance Scale等，在認知層面則運用：Go/No-go任務、注意力反應測試等（Sofia Straudi et al., 2017）。

而臨床實務上的運用，人工智慧可協助診斷及制定可行的治療計畫，以及根據當事人之非語言訊息如肢體動作、臉部表情、聲音等進行情緒狀態的分析（Carter, 2019; Lee et al., 2021; Nixon et al., 2022）。國內有文獻開始探討人工智慧運用於諮商相關領域的成效，如探討諮

商系所學生每天與聊天機器人Woebot互動的感受（葉寶玲等人，2024）、探討人工智慧應用於心理諮商過程之可行性評估（連詠卉等人，2024），提供聚焦及處理策略，以及在不同年齡層及議題中的可行性及挑戰（陳稚均等人，2023），發現聊天機器人在時間及可及性的確有其方便之處，不過因目前主流的人工智慧機器人，如：ChatGPT原設計為英文介面，在語言轉換時回應內容會較難貼近實際文化。在聊天機器人的部分，國內也有團隊在研發相關軟體如Serenio，以AI解憂顧問的角色提供陪伴與支持，提供建議給使用者來面對壓力與焦慮。由研究的趨勢中，可見國內對於AI在心理健康的應用也逐漸蓬勃，持續關注在地化的發展，及在本土的文化脈絡下能如何更貼近。

由上述可知人工智慧在心理治療領域的研究至少已有半世紀的歷史，且持續的蓬勃發展，而人工智慧的效能性、提供協助的程度、被接受的程度，也是被相關領域持續關注（Andrews et al., 2018; Nadarzynski et al., 2019; Pereira & Díaz, 2019）。根據諮商中的特性及要素，對話式的AI在可觸及性佔有優勢，在使用上較不受時間及空間的限制，尤其對認為接受諮商會有污名的人們是一種解套的方式；而在品質、諮商關係和自我揭露上也有其益處，例如可以減少作業的時間、不受記憶容量的限制、較不擔心被批評而更願意揭露等，然而這些治療元素於AI中的表現仍需要更多的探討與評估（Miner et al., 2019; Lee et al., 2022）。

實務上有治療師已於他們的治療當中應用人工智慧，偵測案主的非語言訊息、指派使用虛擬實境（Virtual reality, VR）完成的家庭作業，亦對此抱持樂觀的態度（Jee & Heaven, 2021; Skibba, 2021



）。而回歸到使用者對於AI的態度，Aktan等人（2022）使用問卷來檢驗受試者對AI心理治療偏好的因素，樣本中55%的受試者偏好AI形式的心理治療，不過當提及個人資訊的安全及保密時，大多數人更信任真人的心裡治療師勝於AI系統。由使用者的回饋，再回來看專業人員被科技取代擔心，使用人工智慧的治療師認為可以將其視為輔助，作為治療中的擴充工具提供豐富的功能（Carter, 2019）。因此身為助人者，面對科技的潮流演變可抱持學習的態度，了解AI的運作模式也能在使用時更知其意圖、能力與限制，下文也將針對目前AI運用在諮商輔導中，日常裡相對觸手可及的聊天機器人做更詳細的介紹。

參、聊天機器人在諮商中的運用

一、聊天機器人的分類、功能及運作模式

現代生活節奏快速，資訊大量流通的狀況下人們很難找到時間讓自己的身心好好休息，緊繃久了情緒容易低落，紓壓是其中一個降低憂鬱情緒的方式，除了自行排解壓力之外，與他人「聊天」是多數人會選擇紓壓的方式，但並非每個人都可以隨時找到一個可以談心的對象，可能基於「不安全感」、「怕影響到別人」、「秘密被說出去」、「羞恥感」、「不被同理」等因素，而不敢找人聊天，只能把心事累積起來，久而久之就會影響到心理狀況，甚至罹患心理疾病（林嫻沛，2022）。研究指出，使用者在使用網路輔導資源時，最想得到的協助是「了解內心深處的想法和感受」、「傾訴困難與苦悶」與「得到支持與肯定」（張德聰、林丞增，2017），使用者因為匿名性、隱私及使用空間

彈性而使用網路資源，同樣也會擔心隱私是否被妥善保護。

二、聊天機器人運作模式

聊天機器人根據應用場景可分為Open-domain跟Closed-domain。Closed-domain會引導彼此的對話走向特定的主題，並回答與預定主題相關的答案，像是各銀行的客服機器人、通訊軟體Line設定的自動回覆等，都屬於此類機器人；而Open-domain則會試圖回答各領域的問題，如Apple的Siri、Google的Google助理等（吳雅卉，2019、彭昱傑，2017）。

三、聊天機器人類型

1. Eliza：早在五十年前即開始，1966年麻省理工學院人工智慧實驗室就發明了可以與個案進行簡易對話，主要使用語言為英文，利用程式碼編碼來回應使用者，多半會使用寒暄的方式進行，旨在模仿Rogers學派的治療師。

2. Woebot：由史丹佛發明的虛擬治療師，以認知行為治療為主，能迅速組合訊息、提供服務，能詢問使用者的心情、想法及傾聽感受，雖無法完全取代治療師，卻讓缺乏時間、時時刻刻在網路上的現代人，Woebot提供了不需事先預約、個人化且容易負擔的服務（Fulmer, 2017; Rucker, 2018）。

3. Ellie：由南加州大學創新科技研究所（Institute for Creative Technologies, ICT）所創造，用來治療經歷憂鬱和創傷後壓力症候群（PTSD）的退伍軍人；Ellie具有人臉辨識系統，能根據人臉上66點線索知覺互動對象的情緒狀態，使其不僅能偵測文字，還能反應情緒並分析非語言行為分析，並適度的點頭（Fulmer, 2018; Rucker, 2018）。

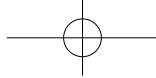
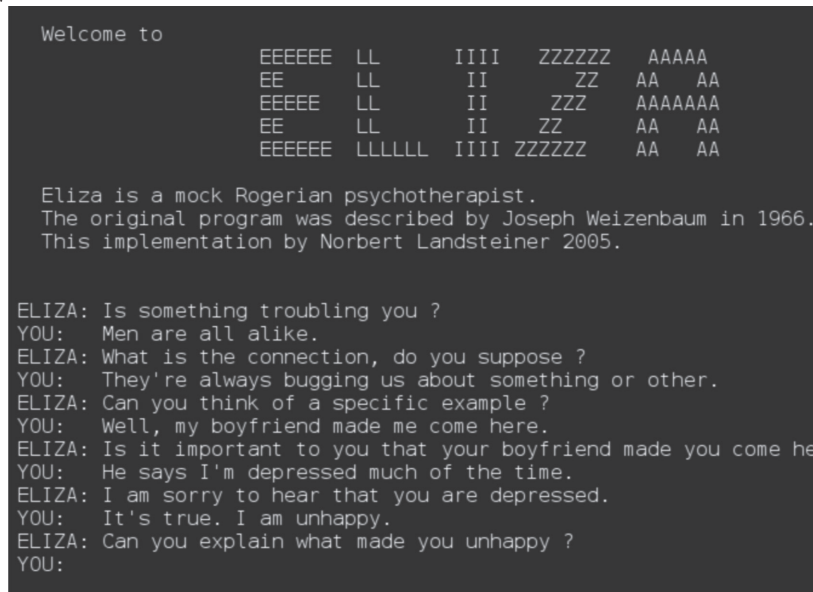
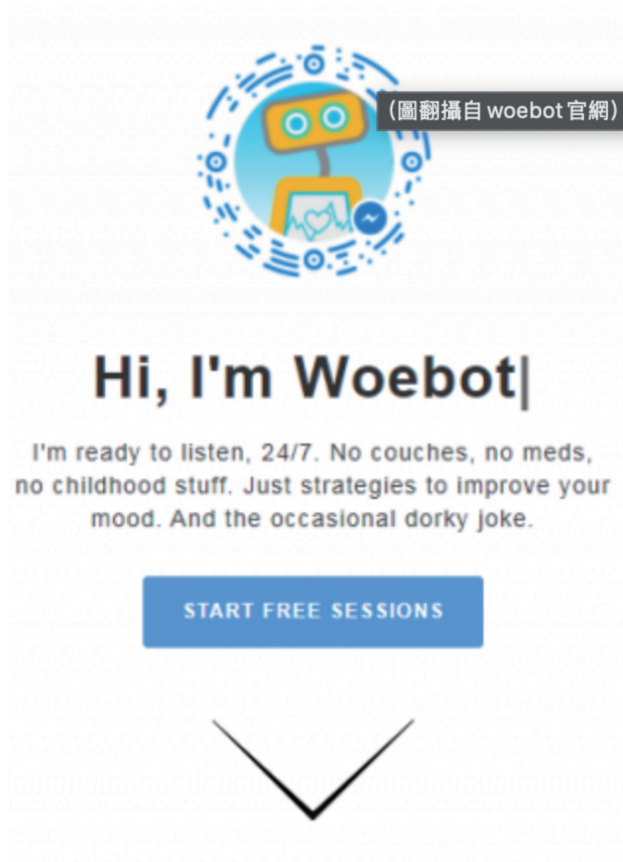


圖1
Eliza示意圖



註：圖片引用自 <https://web.njit.edu/~ronkowitz/eliza.html>

圖2
Woebot示意圖



註：圖片引用自 <https://woebothealth.com/>

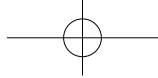
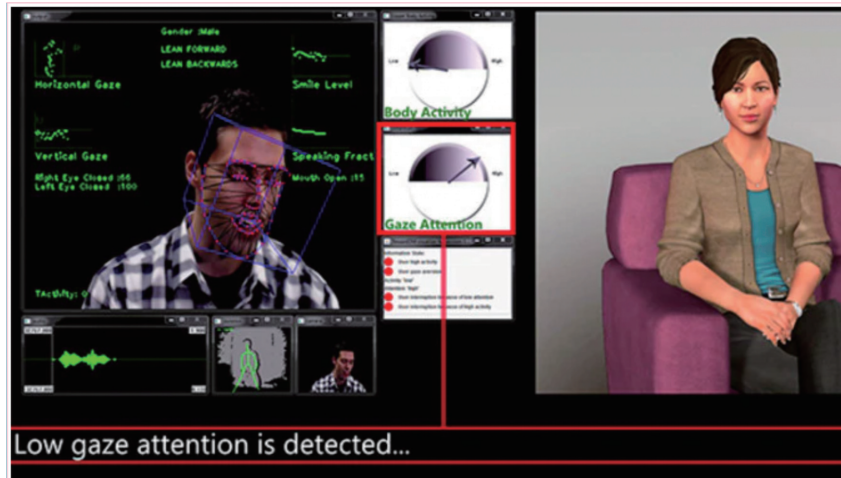


圖3
Ellie示意圖



註：圖片截自<https://www.youtube.com/watch?v=ejczMs6b1Q4>

四、聊天機器人功能

1. 初次使用發送歡迎訊息：將聊天機器人設為好友時，系統會發送歡迎訊息和簡易操作提示給使用者。

2. 資訊樣板：提供關於自我覺察、心情日記及其他舒壓方式等資訊，讓使用者對情緒有更多的認識，忘記使用方法時也提供快速索引的按鈕。

3. 心情日記：選擇紀錄心情時，機器人會給予提示、使用者依照步驟完成，完成後可查詢紀錄中的所有日記。

4. 心情檢測：檢測題目來自台灣精神醫學會（2021）翻譯自美國精神醫學會出版的精神疾病診斷與統計手冊第五版（The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders, DSM-5），共有九題是非題，讓使用者依據最近兩週的狀態填寫，幫助使用者了解自己，也注意何種狀態就需要尋求專業協助。

五、協助治療效能

1. 諮商師有家庭、需要休息與休

假，AI可以隨時提供服務（Fulmer, 2018）。

2. 適合某些難以接觸的使用者，如：溝通障礙、身體障礙或防衛心較高的使用者（Kraus, Zack & Stricker, 2004），能增進使用者的心理健康、提升診斷的精確性，使用者不會覺得自己被評價，因為機器人只能傾聽，不會提出任何要求（Fulmer, 2018），使用者會覺得比較自在，且可隨時使用（available），達到治療師無法滿足的互動次數。

3. 英國倫敦大學Maia Garau與Mel Slater等人的實驗中將16位大學生分為反應組及非反應組，反應組中的虛擬人物會點頭、注視並有基本回應，非反應組的虛擬人物呈現靜止無反應，每位研究參與者與虛擬人物交談三分鐘，過程中紀錄使用者的視線方向及注視時間、身體行為反應及心理問卷回饋，結果證實即使使用者知道和他們互動的是機器人，受到社會互動的潛規則影響，他們仍會產生自然互動反應並願意和其互動，和其分享一些在人際互動裡不願說出的話和覺得害羞、困窘的經驗，使用者會較自在與舒適、願意表達敏感話題



也容易投入諮商歷程 (Maia Garau et al., 2005)。

除了治療功能外，具情感交流及支持功能的聊天機器人也備受歡迎，如蘋果的Siri、Ally Bank的Alexa及微軟的小冰 (Xiaoice)，除了協助查詢生活資訊外，也能跟使用者有簡單的互動，根據調查，Siri與使用者的對話平均只維持約1.5輪，而小冰與使用者的對話則可維持至23輪，分析使用情形後發現每天晚上11點半到凌晨是使用者找小冰對話的高峰時段，尤其是55歲以上的使用者，23輪的對話中除了回答問題之外，另外也包含了情感交流，顯示聊天機器人能排解中高齡使用者子女不在身旁或無法與子女溝通的孤立感，成為理解心情、滿足情感需求的工具 (彭子珊，2014；Fulmer, 2017)。

肆、運用聊天機器人相關之倫理議題

上文整理了聊天機器人的發展、運作功能及其效能，然而聊天機器人在心理與諮商領域中的逐漸普及讓我們不禁思考聊天機器人背後所涉及的倫理議題。

在諮商工作中，心理師都會遵守各地區諮商學會所頒布的倫理守則，像是美國諮商學會 (American Counseling Association, ACA) 訂定的倫理守則、美國心理學學會 (American Psychological Association, APA) 制定立的《心理學家倫理原則與行為守則》 (Ethical Principles of Psychologists and Code of Conduct)、台灣輔導與諮商學會制定的諮商專業倫理守則等。

台灣近期有相關研究指出聊天機器人擁有高效性、普及性、隱私保護等優勢，不過較無法提供人性化的關懷、須

在學習後才能適應不同用戶的表達方式和情感需求及無法給予個案真實而有人情味的支持 (林柏瑜，2024)，另有研究指出受訪者在與ChatGPT互動後覺得GPT較難感知情緒、對話與回應總是一問一答，較為單一，不會主動提問 (連詠卉等人，2024)，聊天機器人作為治療性工具的普及率越來越高，但是聊天機器人背後隱含的相關倫理議題則遲遲未被人們廣為討論和制定相關規定。我們以美國諮商學會在其2014年所修訂的倫理守則中，針對遠距諮商、科技和社交媒體提出了一些規範，以及台灣輔導與諮商學會提出的諮商專業倫理守則作為參考，探討使用聊天機器人可能會面臨之倫理議題。

一、當事人權利

目前的聊天機器人模型回應方式需靠使用者長期與之以文字訓練，才能依使用者的提問、互動模式來提供相對應的回應，聊天機器人較缺乏情感與溫度、因聊天機器人屬於生成式AI，只能備受接受提問，在回應上過度理智化及缺乏變通，無法提供心理師自己的關懷與情感 (連詠卉等人，2024)，在較難提供人性化關懷的狀態下，對某些個案可能會產生負面影響 (林柏瑜，2024)，且在使用者遇到自我傷害或傷害他人的危機時，聊天機器人無法危機通報與立即轉介，在應用人工智慧時需考量以上的當事人權利。

二、公平待遇權

聊天機器人的應用程序層出不窮，市面上常見的聊天機器人常被用在客服回應系統上，當聊天機器人成為你的“



治療師”時，它和其他功能的聊天機器人的異同會是如何？首先，基於台灣輔導與諮商學會的諮商專業倫理守則中，第2.2.2條的公平待遇權，當我們在建置聊天機器人時，應當建立一個可讓所有有需求的使用者可以不受個體文化背景的影響，也可以使用的聊天機器人。為達到此目的，聊天機器人的操作介面會是一個很重要的元素。周恩頤與許文錦（2021）指出聊天機器人的系統與操作介面會影響使用者的使用態度，故在建置聊天機器人的使用介面時需考量其操作的簡易度，以避免提高使用者的認知負荷。學者也建議聊天機器人的操作介面最理想的情況是在經過多次操作後就能夠熟練地上手。因此，專家在建置聊天機器人時因盡可能地將使用介面簡單化，令使用者可以輕鬆上手，以確保每個有需要的使用者都可以公平地使用聊天機器人。

三、免受傷害權

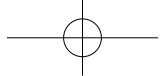
聊天機器人過去的原型是由實驗室建置而成，但現今市面上卻開始流通來自不同領域人才建置的聊天機器人，雖然這些聊天機器人都主張能夠提供心理治療方面的協助，像是Woebot、Joy等等。但無從得知的是，這些聊天機器人的背後是否有專業且符合資格的治療師給予專業的意見且進行監督。為了讓案主免受傷害，建置聊天機器人必需要有專業心理相關背景人員的介入，才能達到此目的。但因聊天機器人的內部運作是非常複雜的，需要由科技人員協助才能完成建制，心理諮商領域的人是很難在沒有外力的情況下獨立完成建制的（Russell et al., 2021）。因此，最理想的狀況是，有意建置聊天機器人的企業必

需邀請符合規定的心理治療領域中的專家給與專業意見，以避免使用者在使用的過程中受到可避免的傷害。

在建立聊天機器人時，創建的人員應當為危機狀況做充足地準備。一般而言，聊天機器人的使用者希望獲得情感支持，他們可能正在面對生活中的困難，卻沒有人可以分享或傾訴，所以才會選擇使用聊天機器人作為傾述對象。許純禎與林哲正（2020）指出危機諮商的個案需要早期辨識並根據個案情境提供適當的治療，而現今的人工智能科技已然達到可幫助人們辨識出有危機狀況的使用者的程度。值得令人思考的是，在建制聊天機器人的時候需要如何設置聊天機器人的回覆詞條，以便在辨識到使用者輸入一些關鍵字時可以立即回覆一些關心的訊息。站在諮商倫理的基本原則，我們必須確保案主可免受傷害。因此，在建立聊天機器人的自動回復時，應當設置相關求助資源的訊息，像是1995生命線協助專線，1980張老師生命專線以及110報警專線等，以便使用者在有需要時，可取得適當的幫助。

四、保密議題

保密在諮商關係的維繫與進展中扮演著重要的角色，它能建立當事人對諮商師的信任，同時讓當事人感到安心，知道自己的隱私權得到保護（謝欣霓，2011）。而美國諮商學會近年來所修訂的倫理守則中，關於遠距諮商、科技和社交媒體的規範中也提出，心理諮商專業人員需承認在保護電子記錄方面存在著一些限制。他們應該向案主告知可能有授權或未經授權的人（例如同事、主管、員工、資訊技術人員）能夠存取這些記錄或傳輸。同樣的，在使用聊天



機器人時也會發生類似的狀況，尤其聊天機器人的後台可能是由非心理諮商領域人員所管理，因此，使用者應當了解自己所輸入的內容有機會被系統的管理員瀏覽。為提升使用者這方面的意識，在在使用者開始使用聊天機器人時，系統應清楚列出聊天機器人的各項限制，尤其是保密方面的限制。我們建議必須讓使用者了解自己在使用聊天機器人時所說的話可能會以何種方式儲存，保留或消除。這樣做的目的是為了提供使用者一個安心的環境去使用聊天機器人而不必擔心自己所說的話會被系統保留或流傳在外。

綜上所述，聊天機器人的過度使用可能導致使用者對機器人產生沉迷和依賴。這可能對使用者的社交互動、情感表達和人際關係產生負面影響。聊天機器人的開發者有責任注意上述所提及之事項。使用者若曾經在使用聊天機器人有不好的經驗，將影響他們在未來尋求真正的專業協助，像是心理諮商等服務（Henderson et al., 2013）。因此，聊天機器人的管理是至關重要的環節。

伍、結論

人工智慧發展是大時代的趨勢，當各行各業都在思考如何與AI共存時，常常跟人互動的我們，更需要思考如何讓互動的溫度不被AI取代，AI就像是現代的「數位樹洞」，提供隨時可傾訴的便利與匿名性，具有其不限時間、場地、不需休息隨時都在的便利性，使用者需負擔的經濟成本相對低廉，我們期待人工智慧的優勢能成為晤談中的助力，像是測驗的計算、初步情緒狀態的評估判別等，人是需要強烈社會性的物種，我們擅長運用同情、自豪、團結，並能使用說服、激勵、鼓舞人心等社會性驅力

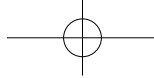
（social drive），機器系統雖擅長心理狀態的評估，但積極解決問題的部分仍遠遠落後人類（引自羅耀宗譯，2017），因此，在開發聊天機器人的過程中，開發者需擁有足夠的心理相關知識背景、專業能力的支撐，系統在保密開發上需妥善保管個案資料、是否能抵擋駭客的入侵、竊取資料，若危機個案與聊天機器人互動時應給出適合的因應策略、或是即時轉介、通知相關系統，這也是當我們仰賴科技時，需要細心謹慎評估的。

參考文獻

- McLeod, J. (2002)。諮商概論（李茂興、吳柏毅、黎士鳴譯）。弘智。（原著出版於1993）
- 台灣輔導與諮商學會（2021）。台灣輔導與諮商學會諮商專業倫理守則。
<http://www.twcpa.org.tw/about.php?id=372>
- 何斐麗（2024）。由利害關係人觀點分析社交機器人在高齡照顧中的競爭優勢—從失智患者的家庭照顧者角色〔碩士論文〕。東海大學。
- 吳雅卉（2019）。聊天機器人研製—以PTT八卦板文章為知識庫〔碩士論文〕。南臺科技大學。
- 林柏瑜（2024）。人工智慧AI在輔導諮商的應用：挑戰與前景。諮商與輔導，461，37-40。<https://www-airitilibrary-com.po.nutn.edu.tw/Article/Detail?DocID=16846478-N202405070008-00013>
- 沈慶鴻（2019）。諮商教育與服務的未來：人工智慧發展下的影響與回應。本土諮商心理學學刊，10(4)，34-54。
- 周恩頤、許文錦（2021）。從聊天機器人探究對話式服務體驗：認知負荷



- 觀點。管理評論，40(1)，21-46。
。https://doi.org/10.6656/MR.202101_40(1).CNI021
- 林嫻沛 (2022)。紓壓聊天機器人〔未出版之碩士論文〕。國立臺北商業大學。
- 陳稚均、洪煒峻、蔡瓊輝、翁瑋鉞、郭玫君、鄒韋民 (2023)。AI辨識畫作即時反映情緒提升青少年輔導效能。中科學報，10(1)，147-163。
。https://doi.org/10.6902/JNTUST.202312_10(1).0008
- 陳靜怡、高淑貞 (2019)。過動兒行為處理策略人工智慧查詢系統試用體驗初探。臺灣諮商心理學報，7(1)，49-84。https://doi.org/10.29824/JTCP
- 許純禎、林哲正 (2020)。人工智慧在危機諮商的應用。諮商與輔導，419，10-14。https://doi.org/10.29837/CG
- 張德聰、林烝增 (2017)。網路輔導之本土化經驗探討：使用網路即時通訊服務的當事人之求助態度及輔導感受之研究。本土諮商心理學學刊，9(2)，35-65。
- 連詠卉、戴承勳、朱方好、黃政昌 (2024)。人工智慧AI應用於心理諮商過程之可行性評估～以ChatGPT 3.5為例。諮商與輔導，465，17-22。
。https://www-airitilibrary.com.po.nutn.edu.tw/Article/Detail?DocID=16846478-N20240906002-00007
- 彭子珊 (2014)。Siri 就是你的同事。天下雜誌，563，60-61。
- 彭昱傑 (2017)。聊天機器人系統設計與實作。〔碩士論文〕。國立中正大學。
- 葉寶玲、郭文正、蔡佳容 (2024)。諮商系所學生使用聊天機器人經驗初探。教育心理學報，56(1)，45-72。
。https://doi.org/10.6251/BEP.202409_56(1).0003
- 謝欣霓 (2011)。保密倫理對諮商關係的影響。諮商與輔導，306，41-44。
。http://dx.doi.org/10.29837/CG.201106.0013
- 羅伊婷、徐尚為、簡慧雯、宋聖芬 (2018)。失智症患者運用人工智慧輔助設備進行認知訓練之成效探討：文獻回顧與未來展望。台灣老人保健學刊，14(1 & 2)，53-66。
- Brynjolfsson, E., & McAfee, A. (2017)。認識你不能錯過的未來：人工智慧大商機 (羅耀宗譯)。哈佛商業評論，134，21-32。
- Aktan, M. E., Turhan, Z., & Dolu,. (2022). Attitudes and perspectives towards the preferences for artificial intelligence in psychotherapy. *Computers in Human Behavior*, 133, 107273. https://doi.org/10.1016/j.chb.2022.107273
- Andrews, G., Basu, A., Cuijpers, P., Craske, M. G., McEvoy, P., English, C. L., & Newby, J.M. (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. *Journal of Anxiety Disorders*, 55, 70-78. https://doi.org/10.1016/j.janxdis.2018.01.001
- Carter, K. (2019, January 9). *How computer-assisted therapy helps patients and practitioners*. American Psychological Association. https://www.apa.org/members/content/computer-assisted-therapy
- D'Alfonso, S. (2020). AI in mental health. *Current Opinion in Psychology*, 36, 112-117. https://doi.org/10.1016/j.copsyc.2020.04.005
- Garau, M., Slater, M., Pertaub, D.-P., & Razaque, S. (2005). The responses of



- people to virtual humans in an immersive virtual environment. *Presence: Teleoperators & Virtual Environments*, 14(1), 104-116. <https://doi.org/10.1162/1054746053890242>
- García, A. T., Serra, C. M., Huguet, M. A., Gardó, L. G., Esteve, A. A., & Quiroga, J. A. R. (2024). Efficacy of a Virtual Reality Intervention for Reducing Anxiety, Depression, and Increasing Disease Coping in Patients with Breast Cancer Before Their First Chemotherapy Dose. *Cognitive Therapy and Research*, 48, 451-465. <https://doi.org/10.1007/s10608-023-10440-2>.
- Hatch, S. G., Goodman, Z. T., Vowels, L., Hatch, H. D., Brown, A. L., Guttman, S., Le, Y., Bailey, B., Bailey, R. J., Esplin, C. R., Harris, S. M., Holt, D. P., Jr., McLaughlin, M., O'Connell, P., Rothman, K., Ritchie, L., Top, D. N., Jr., & Braithwaite, S. R. (2025). When ELIZA meets therapists: A Turing test for the heart and mind. *PLOS Mental Health*, 2(2), e0000145. <https://doi.org/10.1371/journal.pmen.0000145>
- Helgadottir, F. D. (2019). *Computers and therapy*. AITherapy. <http://www.aitherapy.com/articles/computers-and-therapy>
- Henderson, C., Evans-Lacko, S., & Thornicroft, G. (2013). Mental illness stigma, help seeking, and public health programs. *American Journal of Public Health*, 103(5), 777-780. <https://doi.org/10.2105/AJPH.2012.301056>
- Jee, C. & Heaven, W. D. (2021, December 6). *The therapists using AI to make therapy better*. MIT Technology Review. <https://www.technologyreview.com/2021/12/06/1041345/ai-nlp-mental-health-better-therapists-psychology-cbt/>
- Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H. C., Paulus, M. P., Krystal, J. H., & Jeste, D. V. (2021). Artificial intelligence for mental health care: clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 6(9), 856-864. <https://doi.org/10.1016/j.bpsc.2021.02.001>
- Lee, J., Lee, D., & Lee, J. G. (2022). Influence of Rapport and Social Presence with an AI Psychotherapy Chatbot on Users' Self-Disclosure. *International Journal of Human-Computer Interaction*, 1-12. <https://doi.org/10.1080/10447318.2022.2146227>
- Miner, A. S., Shah, N., Bullock, K. D., Arnow, B. A., Bailenson, J., & Hancock, J. (2019). Key considerations for incorporating conversational AI in psychotherapy. *Frontiers in Psychiatry*, 10, 1-7. <https://doi.org/10.3389/fpsyt.2019.00746>
- Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Digital Health*, 5, 1-12. <https://doi.org/10.1177/2055207619871808>
- Nixon, D., Mallappa, V. V., Petli, V., & HosgurMath, S. (2022). A novel AI therapy for depression counseling using face emotion techniques. *Global Transitions Proceedings*, 3(1), 190-194. <https://doi.org/10.1016/j.gltp.2022.03.008>



- Pereira, J., & Díaz, Ó. (2019). Using health chatbots for behavior change: a mapping study. *Journal of medical systems*, 43, 1-13. <https://doi.org/10.1007/s10916-019-1237-1>
- Raile, P. (2023). The usefulness of ChatGPT for psychotherapists and patients. *Humanities and Social Sciences Communications*, 10, Article 2567. <https://doi.org/10.1057/s41599-023-02567-0>
- Russell F., Tonya D., Cori C., and Angela J.. (2021). The Ethics of Psychological. Artificial Intelligence: Clinical Considerations. *Counseling and Values*, 66(2), 131-144. <https://doi.org/10.1002/cvj.12153>
- Skibba, R. (2021). *The computer will see you now: is your therapy session about to be automated?*. The Guardian. <https://www.theguardian.com/us-news/2021/jun/04/therapy-session-artificial-intelligence-doctors-automated>
- Trusch, P. (2017). Re-reading ELIZA: Human-machine Interaction as Cognitive Sense-ability. *Australian Feminist Studies*, 32(94), 411-426. <https://doi.org/10.1080/08164649.2017.1466647>
- Wray, T. B., Kemp, J. J., & Larsen, M. A. (2023). Virtual reality (VR) treatments for anxiety disorders are unambiguously successful, so why are so few therapists using it Barriers to adoption and potential solutions. *Cognitive Behaviour Therapy*, 52(6), 603-624. <https://doi.org/10.1080/16506073.2023.2229017>

投稿日期：113年09月11日

第一次修改日期：113年11月11日

第二次修改日期：114年04月24日

通過日期：114年08月19日