

## 學術對談

# 人機之交淡若水： 從一次虛擬的「學術對談」說起

肖小穗

### 摘要

《傳播與社會學刊》於7月中旬組織了一次虛擬的學術對談，由傳播學的資深學者組成的團隊在線上訪談了當時爆紅出圈的新聞「人物」——智能聊天機器人ChatGPT(基於3.5模型)。訪談涉及資訊氾濫、跨學科合作、人工智能應用於研究等一系列學科亟需解決的問題。為了獲得內容豐富的回應，同樣的訪談在7月下旬再重複進行了兩次。通過深入分析這三次訪談中GPT-3.5生成之文本的內容和意義結構，本文發現這個模型尚不具備深度學術對談的能力，它只能在一個淺顯的層面上理解和回答問題。本文還通過它答題時的言語行為模式追索到這一模型的一些結構性偏向和局限。本文因此不建議完全依賴這一模型及其提供的資訊。文章最後借用「君子之交淡如水」的古訓來說明人機交往應把持的尺度。我們希望與人工智能發展各方面的交往和合作，但兩者之間保持適當的距離其實有助於這一交往和合作健康持續地發展下去。

肖小穗，香港浸會大學傳理學院傳播學系榮休教授。研究興趣：修辭傳播、文化研究、中國傳播理論。電郵：s82299@gmail.com

---

Dialogue

## “Human-Machine Friendship is as Tranquil as Water”: A Critical Reading of the Virtual Academic Dialogue

Xiaosui XIAO

---

### Abstract

In mid-July 2023, *Communication and Society* organized a virtual academic dialogue. A team of senior communication scholars conducted an online interview with the then-popular AI, ChatGPT, based on the 3.5 model. The interview covered a range of pressing issues in the field, including information overload, interdisciplinary cooperation, and the application of AI and GPT in research. This same interview was repeated twice in late July for further insights. Through an in-depth analysis of the content and semantic structure of GPT’s responses across these three interviews, this article concludes that the 3.5 model does not yet have the capacity for a deep academic dialogue. It is only able to comprehend and respond to questions at a rudimentary level. Furthermore, the article identifies some structural biases and limitations of the 3.5 model based on its pattern of responses. Consequently, the article advises against complete reliance on this model and the information it provides. Drawing from the ancient Chinese adage “the friendship of gentlemen is as tranquil as water,” the article illustrates the need for appropriate balance in human-computer interaction. While we can look forward to increasingly open

---

Xiaosui XIAO (Professor Emeritus). Department of Communication Studies, School of Communication, Hong Kong Baptist University. Research interests: rhetorical communication, cultural studies, Chinese communication theories.

*“Human-Machine Friendship is as Tranquil as Water”*

and in-depth interactions and collaborations with AI in the future, maintaining an appropriate distance is essential for the healthy and sustainable development of this cooperation.

**Citation of this article:** Xiao, X. (2023). “Human-machine friendship is as tranquil as water”: A critical reading of the virtual academic dialogue. *Communication and Society*, 66, 19–51.

《傳播與社會學刊》，(總)第66期(2023)

## 引言

我們這期學術對談的對象是一位特別的嘉賓，它是不久前仍爆火出圈、風頭無兩的ChatGPT-3.5。這位科技新寵於去年11月底橫空出世，在華人社區引爆的巨大衝擊波直到今年7月的學術對談期間，仍在搖撼著這裏的各行各業，學術的高雅之堂也不能倖免。經歷短期的眩暈、不知所措和無奈觀望之後，華人學者紛紛作出反應。以中國大陸的傳播學者為例，先是在期刊雜誌上發表文章，然後在專刊上連篇累牘，所探討的議題涉及知識暗流(張萌、朱鴻軍，2023)、新的即興的演示/展演文明(黃旦，2023)、新常人統治(劉海龍、連曉東，2023)、語言基礎設施、語法生成標準化(胡翼青、胡欣閱，2023)、數字團結和普遍聯結(杜駿飛，2023)、新文本間性(周慎，2023)、技術控制的選擇(劉永謀、王春麗，2023)，等等。這種種回應，從實踐到學術、從技術到哲學、從樂觀到悲觀，表明應對這位高深莫測的智能界來客的戰線已全面拉開。

《傳播與社會學刊》是在這此起彼伏的聲勢下加入行動的。傳聞這位智能聊天機器人問必答、無所不知，肚子裏能攬下百科全書的數據庫，比古人所說「才高八斗、學富五車」有過之而無不及，我們希望跟它來一次近距離的深度訪談，一來聽取它對當今一些問題的高見，二來測試它的學術應對能力，我們探知這位機器人的深淺虛實，為的是搞清楚狀況，在今後的合作中做到知「人」善用。

我們請出了包括李金銓、陳韜文、黃煜、魏然、周樹華、黃懿慧等海內外資深教授在內的10位學者擬定問題，他們作為實證、批判和技術三個學派的代表，提出了25道學術問題。我們還請了三位學者負責線上提問的實際操作，他們在相近的三個不同時段分別向GPT-3.5提出這25道問題，由此獲得了對方三個版本(版本一、二、三)的回應。這三個版本可供我們比較它們之間的相似度，以此判斷GPT-3.5答案的可靠性，我們當然也希望它們能為我們獲取內容豐富的答案。

為甚麼是GPT-3.5，不是實力更強的GPT-4.0？理由簡單，這時候的4.0模型尚未進入大眾視野，因為要付費，用戶不多。我們總要面對越來越強大的智能聊天對手。目前這個3.5版本的「學術對談」，作為一

個時代(GPT-3.5時代)的歷史見證，也作為我們與GPT首次對話的真實記錄，有它不可取代的研究價值和意義。再說，我們這次對話還有一個重要任務，找出GPT-3.5一些結構性偏向或局限，以便比對以後的系列，看這些結構性偏向或局限是否會像某種遺傳基因那樣傳承下去。

本文分析這三個GPT-3.5版本的回應。面對大語言模型生成的這些文本，作為讀者和點評人，我只能提供一個出自非常有限的記憶庫存，但擁有數十年豐富的學術閱讀經驗和受過人文素養的系統訓練的人類大腦的解讀和分析。好在GPT也使用人的自然語言，我不需要搬用任何專業的文本分析和話語分析方法。事實上，本文採用最土法的分析，憑藉本人對基本的語法、句法和詞法的認知。

當然，這篇分析文章不能只是挑一些語言或文字上的毛病，多年的閱讀經驗讓我不能滿足於表層和字面上的閱讀。我知道學者提出的問題總會帶有一些深切的關懷(deep concerns)，這些關懷會是我評價GPT-3.5「理解」能力的一個重要依據。另外，作為一個人文主義者，我難免會在閱讀中更多地關注人類表達與機器表達的不同之處，我因此也需要深入讀一些東西，我會嘗試從這部智能機器回答問題時的一些言語行為模式中讀出它的一些結構性趨向。

以下，我分別從「解題能力」、「答題表現」、「語言特色」、「偏見或偏向」四個方面來分析和評價GPT-3.5在這次學術對談中的文本生成效果。分析中，為方便讀者隨時查閱相關的GPT答案，本文索引時會將答案所屬的版本序列排在前面，答案對應的問題序列排在後面，例如2:5指第二版本中第五道問題的答案。

## GPT-3.5的解題能力

這次為GPT準備的問題，每一道都有它自己的歷史，看起來只是寥寥幾句，卻凝聚了多方的深切關注。我們期待GPT不但能讀懂這些問題的字面意思，還能把握其背後的深切關注。但結果很不理想，GPT-3.5對題義的「理解」或多或少都存在一些問題。就拿一開始的幾道問題來說吧。

《傳播與社會學刊》，(總)第66期(2023)

**Q1：現代社會已經充斥各種「資訊」(information)，由於人工智能的進步，「資訊」更將唾手可得。但「資訊」氾濫，是否會鞏固或動搖「知識」(knowledge)結構的基礎，是否會使得「智慧」(wisdom)淹沒不存，變得不重要？**

這道問題涉及資訊、知識結構和智慧，提問者顯然要知道GPT如何處理三者間的關係。問題還透露出一種憂慮，擔心資訊的氾濫會衝擊知識和智慧在人們認知結構中的地位。讓人不解的是，第一版的GPT-3.5忽略了有關知識結構的問題，只是說了一下資訊對知識的影響。第二版雖然回應了有關知識結構的問題，卻沒有回答有關智慧的問題。對談才剛開始，GPT-3.5在兩個時段上的回答就有如此大的分別，而且都只是部分地回答了問題，讓人懷疑它是否看清楚了題目。儘管版本二仍樂觀地宣稱「正確地利用資訊和人工智能可以幫助人們構建更加精確和全面的知識結構」，但它對智慧的有意無意的忽略說明提問者對知識結構和智慧之命運的擔心是對的。

**Q2：在各說各話的宇宙，人類情感和認知的溝通(communication)還有空間、還有意義嗎？人類的理性(reason)和「價值」(value)還重要嗎？**

GPT(版本一)：在各種意見紛陳的宇宙中，人類情感和認知的溝通仍然非常重要。人類的情感和認知是我們理解世界、建立關係和共用經驗的基礎。雖然不同的觀點可能存在衝突，但透過溝通，我們可以尋找共同點、理解彼此的立場並建立共識。

GPT(版本二)：在各自說話的宇宙中，人類情感和認知的溝通仍然非常重要。儘管不同的宇宙可能存在著不同的價值觀和信仰，但人類的情感和認知是普遍存在的，可以作為跨越不同宇宙的共同語言。然而，不同宇宙之間的溝通也需要建立在尊重和包容的基礎上，以便在相互理解和對話的過程中達成共識。

在版本一，GPT-3.5把「各說各話的宇宙」解讀為「各種意見紛陳的宇宙」，但「各說各話」與「意見紛陳」不是同一個意思——前者拒絕溝通，後者接受溝通。在版本二，GPT-3.5換了個另類的解讀：人們「各

自說話」，由於「不同的價值觀和信仰」被分隔在「不同的宇宙」裏。於是，這邊是「情感和認知」，那邊是「價值觀和信仰」，本來是你中有我，我中有你，GPT-3.5硬要把它們割裂開來，聲稱「人類的情感和認知是普遍存在的，可以作為跨越不同宇宙的共同語言」。如何理解這套話術呢？我懷疑這是GPT-3.5為維持情感和認知的溝通「仍有空間和意義」而編出來的一種說辭。「人類情感和認知的溝通還有空間、還有意義嗎」這個問題，不是要我們切割「情感和認知」與「價值觀和信仰」的關係——今天沒有學者會否認價值觀和信仰本身也是一種情感和認知結構，而是希望我們對情感和認知的複雜構成做深入細緻的分析。

**Q3：傳播研究的分類越來越細，乃至不同類型的傳播學者之間越來越難進行有效對話。有甚麼解決方法嗎？**

內行人知道，這個問題的痛點不在這個「分類越來越細」的「細」字，導致不同類型之間「越來越難進行有效對話」的真實原因是專業內卷造成研究越來越深陷一種自我封閉的狀態。GPT-3.5簡單地理解為「子領域和分支正在不斷湧現」（版本一），或者是「傳播研究的分類越來越細」（版本二），或者是「對於特定主題的研究越來越細」（版本三），因此沒能對症下藥，給出相應和具體的解決方法。

**Q5：人的研究在多大程度上還有意義？**

同樣，GPT-3.5也沒有把住這一問題的痛點。問題不在於「有沒有意義」（人的研究當然還有意義），而在於為甚麼我們今天會有這樣的擔憂、發生了甚麼、人的研究受到了哪些方面的衝擊，導致我們今天要來探討「在多大程度上還有意義」，這才是我們所要關注的。GPT-3.5並沒有回應這些關注，只是籠統地回答為甚麼「仍然有意義」，它所解釋的五個原因——理解人類行為和互動、影響社會政策和實踐、理解文化和價值觀、與人類相關的倫理挑戰、提高媒介素養等——擱在哪個時代和情境都是適用的。

學術問題往往發源於一種深度的關切，這一關切並不總是溢於言表，所以需要仔細審題。就以上第一道問題而言，GPT-3.5間接和部分地回答了它的深度關切，這要歸功於問題本身已經陳述了這一關切。至於後三個問題，GPT-3.5回應的只是它們的字面意思，不是它們的深

《傳播與社會學刊》，(總)第66期(2023)

度關切。諷刺的是，即使是回應字面意思，至少有三個版本沒能把握住問題中一些重要表述或重要概念的意思，譬如版本一和二對「各說各話」的解讀，又譬如版本二把「智慧」混同於「知識」。中肯地說，在大多數情況下GPT-3.5對於問題字面意思的把控還是比較準確的，它的主要問題是深度理解不足。遺憾的是，這一不足不僅影響了GPT-3.5對其他問題的理解，更影響了它後面答題時的整體表現。

我們知道，GPT不具備人的理解能力，依照學者的專業理解和判斷來要求GPT，似乎是在挑戰它的不可能。但是作為一部大型的人工智能對話機器，它應該有自己的方法識別甚麼是「題內之義」，並且能在望聞問切一大摞與題義相關的文獻資料的過程中把觸它們的深度關切。我們對一部參與學術對話的人工智能的要求決不只是能流暢地使用自然語言，它還應該知道我們在說甚麼和想說甚麼。

當然，我們也需要檢討自己的提問方式，儘管我們的問題對一位學者來說不是甚麼新問題，不少還是傳播學界的老生常談，但對GPT-3.5來說，我們的提示也許還不夠清晰明確，我們還可以做得更好。在這次對談中，我們也有學者提出既明確又明細的問題，譬如有關傳播與社會抗議激進化關係研究的一些問題，GPT-3.5對這部分問題的理解也有些出入，譬如學者要求對這項研究中的「激進化」概念作一個文獻的綜述，但版本一和版本二都只是簡短地解釋了這項研究是幹甚麼的。學者還要求「概括一下關於傳播與社會抗議之間關係的研究結果」，而版本一只是歸納了研究文獻中出現的若干「關鍵主題」，這不完全是學者所要求的「研究結果」。

學者提出的問題，是我們評判和分析GPT表現的基本依據，所以後面的章節還會一而再再而三地回到學者的問題和GPT-3.5的「理解」上來。

## GPT-3.5的答題表現

首先，通過比較三個時段(三個版本)的回應，我們沒有發現GPT-3.5在哪一個時段的發揮普遍比其他兩個時段更好。我們很快就會知曉，即使是回答同一個問題，它也是時好時不好，所以說不上在甚麼時段GPT-3.5會適合回答甚麼樣的問題，我們只能說GPT-3.5的表現是

隨機的，也就是說，即使是電子作業，我們也不能期待GPT-3.5任何時候都能保持最佳狀態，更不能保證它任何時候都能交出最好的答卷。GPT-3.5的答卷由三部分組成，第一部分是引言，GPT-3.5通常會簡要地回應學者的問題，同時開門見山表明自己的態度或觀點。第二部分是答卷的主體，GPT-3.5會根據學者的需求提供相應的服務，通常是建議若干方法或解釋。最後還會有幾句結語或結論。

解題不足，不可避免地限制了GPT-3.5答題的發揮，它只能在一個淺顯的層面上回答問題，導致這次對談的局面打不開。閱讀GPT-3.5的答卷，經常困擾我們的是它流於表面的泛泛而論，在前面的幾個例子中，我們已經對這種泛泛而論見識一二。總感覺它的話下不去，不接地氣，沒有根基，在半空飄忽不定，所以才一會兒說成這樣，一會兒說成那樣，你不知道哪個版本是靠得住的。它也不像是要就事論事，所以說話不涉及太多具體內容和事例。感覺它不是真的要說一件事，只是煞有介事，它更在乎如何自圓其說，為此，甚至可以改變對一個概念的尋常理解。這種流於表面的泛泛而談貫穿三個版本，隨處可見，以至於我覺得這就是GPT-3.5慣有的言談風格和應對策略。

但即使是在一個淺顯的層面上，GPT-3.5還是有它亮眼的表現。它的本事是把我們的問題橫向地攤開來說，你若問它「如何」，它可以迅速列出一排方法來；你若問「為甚麼」；它可以迅速給出一堆原因來；你若徵求意見，它也可以迅速整理出好幾段思路來。它這種快速歸類和集結的能力——片刻間拉出一個隊列或陣容的建議來，讓人歎為觀止。可就是，接下來讀這一個個建議時，便發現它們依然下不去，依然懸在半空。

我特別分析了這些建議的句法和意義結構，一般是兩三句話，說了甚麼不重要，重要的是彼此綁定的關係。如果GPT-3.5建議的是某種應對的方法，它要求採取某項行動，那麼通常會這樣：GPT-3.5會給這項行動某個目的或某種意義，或者為它設置某些條件，或者把這項行動看作是對特定情勢的回應，又或者看作是一種解決方法。譬如在「如何促進跨領域對話和合作」的問題(Q3)上，GPT-3.5建議以下五種方法(1:3)：

《傳播與社會學刊》，(總)第66期(2023)

- (1) 跨學科合作：建立跨學科的團隊或研究項目，將不同領域的學者和專家聚集在一起。這有助於彼此分享知識、經驗和方法論，並從不同的角度來看待問題。
- (2) 促進對話平台：建立或參與跨領域的學術會議、研討會或工作坊，提供學者們交流和討論的場所。這些平台可以促進學者之間的互動，並促進跨領域的對話和合作。
- (3) 融合方法論：傳播研究的不同分支可能使用不同的方法論和研究工具。學者可以探索整合不同方法論的可能性，以便獲得更全面的研究結果。這種跨方法論的融合可以促進對話和互相學習。
- (4) 共同研究議題：尋找共同的研究議題或問題，這些議題可能跨越不同的傳播領域。通過共同關注的議題，學者可以共同合作，將不同的專業知識和觀點帶入研究，進一步豐富對話和研究成果。
- (5) 培養開放思維：學者們應該持開放的態度，願意接受和探索不同領域的觀點和方法。這需要學者們積極學習和更新自己的知識，並樂於與其他領域的學者進行合作和對話。

從以上的方法陳述中，我發現了起碼三種不同的句法結構。前三種方法的結構是「行動—目的—意義」，後兩種分別是「手段—目的」和「條件—行動」。在這次對談中經常跟GPT-3.5所建議的方法或行動聯繫在一起的句法結構還包括「情勢—應對」、「問題—解決」等。假如GPT-3.5的建議涉及對某種狀況（譬如問題）的認知和理解，那麼經常出現的句法結構是「原因—結果」、「正面—反面」、「問題—解決」等。這些句法結構也可以看作是一些相對獨立和自我完整的意義結構。GPT-3.5不僅生成了它的「具體建議」，還生成了這樣一些意義結構來引導我們去理解和接受這些建議。

這些是最尋常不過的句子結構，對GPT而言，生成這樣的句法和意義結構不是甚麼大學問，把相關的資訊作些配對就是了。相信GPT-3.5生成這些句法和意義結構時首先考慮的是句子之間意義的關聯性和

銜接性，而不是事實陳述的真確性和可靠性，這就是我為甚麼對這種句法和意義結構所表達的東西心裏沒底、不踏實。在我眼裏，GPT-3.5 在此只是完成建構了一些看似自我完整的句子，它們合理和完整的外表並不遮蓋內涵的空洞。

我們再來看一個完整的問答。

**Q13：媒體和傳播研究學者如何應對使用人工智能和聊天機器人所帶來的道德問題 (ethical problems)？**

GPT (版本一)：媒體和傳播研究學者在使用人工智能和聊天機器人時，應該重視並嚴肅面對相應的道德問題。以下是一些應對道德問題的方法：

- (1) 倫理審查：在使用人工智能和聊天機器人進行研究之前，學者應進行倫理審查。這包括對研究目的、數據收集和使用、知情同意和隱私保護等進行評估，確保研究方案符合倫理標準和法律法規。
- (2) 保護參與者隱私：學者應確保使用人工智能和聊天機器人收集的數據不會損害參與者的隱私權。採取適當的數據匿名化和保護措施，確保參與者的身份和個人資訊得到保護。
- (3) 提供知情同意：在使用聊天機器人進行調查和互動研究時，學者應向參與者提供明確的知情同意說明，解釋研究目的、程序和數據使用方式，並尊重參與者的自由選擇權。
- (4) 解釋模型結果：在使用人工智能算法進行分析時，學者應注意模型結果的解釋性和可靠性。確保結果的解釋和推斷符合科學標準，避免對研究結果進行不恰當的解讀和過度解釋。
- (5) 注意偏見和歧視：人工智能算法可能存在偏見和歧視問題，學者應該注意這些問題的存在並採取相應的校正措施。確保研究結果公正、公平和客觀。
- (6) 透明度和可追溯性：學者應該保持研究過程的透明度，包括數據的來源和處理過程、算法的使用方式等。這有助於確保研究的可追溯性和可信度。

《傳播與社會學刊》，(總)第66期(2023)

(7) 開放和合作：在面對複雜的道德問題時，學者可以開放和積極合作，與其他學者和專家一起討論和解決問題，共同推動傳播研究領域的倫理發展。

總的來說，媒體和傳播研究學者在使用人工智能和聊天機器人時，應該秉持尊重和保護參與者的原則，關注道德問題並採取相應措施，以確保研究的合法性、公正性和可信度。這樣的做法不僅有助於保護參與者的權益，也有助於推動傳播研究的健康發展。

之所以要選擇這個例子，是因為它關係到我們的主角聊天機器人，還涉及它的出現「所帶來的道德問題」，我特別想知道這個當事者如何回答這個具體且現實的問題。傳媒與傳播研究的道德問題是老問題，像倫理審查、保護參與者隱私、提供知情同意這樣的道德舉措早就有組織有步驟地推行了，而且在執行上有了規範的流程，現在使用人工智能和聊天機器人是勢所趨，總要帶來些變化吧，起碼會在這幾個措施的陳述上注入些新的內容和新的說法吧。可惜沒有，我們聽到的基本上還是那些耳熟能詳的東西，酒還是舊酒，瓶也還是舊瓶，只是加蓋了一個「使用人工智能和聊天機器人」的標籤罷了。看來GPT-3.5並不在乎要為自己的出現和將要扮演的角色討一個不一樣的說法。所以從它的這些應對建議中，我們依然不清楚該如何應對這個新趨勢，我們甚至不清楚有些甚麼新的道德問題要來應對。

說GPT-3.5的建議毫無新意，也不盡然。我比較了其他兩個版本，三個版本基本一致的有「保護參與者隱私」、「提供知情同意」、「注意偏見與歧視」；兩個版本相近的有「倫理審查」（版本一，版本二是「建立倫理委員會」）、「透明度和可追溯性」（版本一，版本三是「透明度和解釋性」）、「考慮社會影響」（版本二，版本三是「社會責任」）；獨一無二的有「解釋模型結果」（版本一）、「開放與合作」（版本一）、「虛假資訊和誤導」（版本三）。最後三個獨一無二的建議，聽起來沒有甚麼特別，但還是可以姑且看作是GPT-3.5在特定時段的創造性發揮。這一比較還讓我們知道一點：GPT-3.5只是給我們提供了多種選擇，並沒有提供所有可能的選擇。

說GPT-3.5所建議的方法下不去，懸在半空，還因為它忽略了一些可供它的方法安全落地作參考的重要資訊。GPT-3.5只是一股腦兒抖擻了一堆方法，並沒有嘗試幫我們理出一個輕重緩急的頭緒來，也沒有告知目前這個先後次序是依據甚麼排列出來的。照說這些方法是使用人工智能和聊天機器人之後要採取的應對措施，如是，我們還需要知道這裏面哪些是學界目前正在採用的方法，哪些是GPT-3.5合理推導的方法，哪些只是它自己的創意發揮。以上資訊對實際操作的人來說至關重要。我們期待GPT把它所知道或能知道的這些方面的資訊也說出來，否則拿到了一堆建議，我們依然一籌莫展，無所適從。

其實，如果GPT要徹底改變在一個淺顯的層面上回答問題的尷尬局面，而且要切切實實地幫助它的使用者的話，它還應該提供能讓使用者深入瞭解這些應對方法的資訊，包括瞭解它們的歷史、它們的複雜背景和它們的深切關懷等；還應該提供所有能幫助使用者判斷當前局勢和作出適當決策的資訊。而不只是在這裏泛談和堆砌一些可能有用作也可能暫時用不上的方法。

以上例子集中地反映了GPT-3.5的答題模式和它的問題所在，隨著這個答題模式不斷複製，它的局限和缺陷也以不同方式和不同程度地嵌入了其他問題的答案裏。總的來說，GPT-3.5的答題表現不怎麼樣，遠遜於我們的預期。它有亮眼的表現，展示了在一個廣闊的地域上通盤把控和排兵佈陣的能力，遺憾的是，它把控的深度欠奉，後勁不足；它顯然是花了不少功夫，為我們整理出多不勝收的選擇，可惜這些功夫沒有做到實處；它的建議看似照顧了我們需求的方方面面，可是每一方面點到即止的三言兩語，像是在敷衍而不是在解決我們的問題。

## GPT-3.5的語言特色

目睹了有關GPT-3.5生成意義和內容之能力如何強大的「人設」幾近崩塌之後，剩下的便是它的語言生成的能力，我們回到了這部智能聊天機器發家致富的起點，期待這部集大語言模型、海量語料庫、自然語言生成器、語言基礎設施於一身的聊天機器，至少能在談吐的語言

《傳播與社會學刊》，(總)第66期(2023)

方面讓我們看到它的一些希望。但是很遺憾，就這次學術對談而言，我們沒有看到GPT-3.5在這方面有特別出彩的作為。

首先要說它的語言極為單調，缺少變化，譬如表達語氣轉折的詞，來來回回就是「然而」和「但」，偏偏這兩個詞的使用率頗高，因為經常要平衡對立雙方的意見，所以要不斷地重複說些「我們要怎樣怎樣，然而，也要怎樣怎樣」的話。GPT-3.5的語言還平淡無華到乏味。我試圖尋找一些感性、形象或生動的描述，但遍尋不獲。結論是，這不是要來追求語言豐富和多樣化的地方。

GPT-3.5還惜字如金，點到即止，它的引言、結語，以及每一項建議，都不過是幾句話的功夫。在很多答案的引言中，GPT-3.5甚至沒有刻意包裝自己的看法，它陳述自己立場的文字大多還是從學者們的問題中搬過來的，它其實只是回答「是與不是」、「可以還是不可以」、「應該還是不應該」……。相信在它過去的訓練和學習中，也一定學到如何用自己的語言複述他人的話(譬如學者問題中的話)，所以時不時也會展露自己複述的才學。但是看了前面幾個分析案例之後，我們不會希望它有太多創意的發揮，生怕它越是大膽和自由地發揮，就越有可能荒腔走板。

語氣平和也是它在對話中一個不容錯過的特色。不管問它甚麼，永遠聽到的是一把不溫不火、平和到有點平庸的聲音。你不會擔心它跟你爭辯或反駁你。再有爭議的問題，在這裏也只是把兩邊的意見平攤開來，這邊說說，那邊也說說，一碗水端平。感覺學術語料進到這個大數據處理器，經過大浪淘沙般的淘洗之後，所有的鋒芒和棱角都給淘洗盡了。

這不是我希望看到的學術對談。學術殿堂不是語言的清水衙門，豐富和精湛的思想被GPT-3.5以這種清心寡欲、點到即止和枯燥乏味的方式轉述之後，就只剩下我們目前所見的泛泛而論。但我不能不承認，這套語言也正好配合了GPT-3.5的泛泛而論，除語言單調之外，平淡無華、點到即止和不溫不火特別適合GPT-3.5在一個平面上把話鋪展開來。至於單調，我想不到它有甚麼存在的理由，只能歸咎於目前這部聊天機器生成能力的不濟。

當然，GPT-3.5所生成的語言也不是一無是處，自然語言所要求的順暢和簡潔，GPT-3.5做到了，還能借助於語句之流暢來填補內容的空泛。此外，我還發現了GPT-3.5一個特長，它擅長建構一種語義對稱的句子。上一節分析GPT-3.5的句法結構和意義結構時，已經見識了它這方面的能力。在此還想多說幾句，在「『資訊』氾濫是否會鞏固或動搖『知識』(knowledge)結構的基礎？」這一問題(Q1)上，版本三是這樣回答的，「資訊的大量湧入可能會對個人和社會的知識結構帶來挑戰。這些資訊可能是零散的，缺乏深度和連貫性，這可能導致人們難以建立完整、精確的知識體系。」這裏，「零散」對上了「缺乏深度和連貫性」(作為對「零散」概念的意義推定)；「零散」的資訊再與「完整、精確」的知識結構形成鮮明對照。在版本一，GPT-3.5雖然回避了「知識結構」這個概念，但還是點出了資訊超載所導致的「知識的碎片化和表面化」——在這個意義上，版本一其實已經間接地回答了「知識結構是否受到影響」的問題。

我要說的是，我們接受類似這樣的陳述不是因為它們有甚麼事實依據，而是因為它內在的邏輯是對的，這不是一個需要事實來證明的陳述。在資訊的時代我們都不會反對「知識有賴於來自現實和象徵世界的資訊」這個預設，所以只要接受了資訊是「零散的，缺乏深度和連貫性」的界定，就不能不同意「零散」資訊的氾濫會衝擊知識以及知識結構的基礎。在我看來，GPT-3.5的可信性往往就來自它順暢和連貫的句子和這種前後對應的句法和意義結構，人們容易被這種看似意味深長的語義和邏輯力量所迷惑，以為是在陳述千真萬確的道理，於是放棄對事實的追究。

我們這次還見識了GPT-3.5的另外一種語言能力：說事的能力，具體點說，它說起事來，有模有樣，像是在說這麼一件事，其實，用Wolfram語言(Wolfram language)之父斯蒂芬·沃爾弗拉姆(Stephen Wolfram)的話說，「它只是在根據訓練材料中的內容『聽起來像甚麼』來說出『聽起來正確』的話」(2023: 102)。它並不真的瞭解我們的媒體與傳播研究，卻可以評估這個領域的現狀和未來，還可以說上目前和將來面臨的一系列「關鍵挑戰」，諸如「數字化顛覆」、「錯誤資訊和假新聞」、「隱私和道德問題」等等(3:18)。就「數字化顛覆」一項，雖然只有

《傳播與社會學刊》，(總)第66期(2023)

三句話，卻都能對上「數字化顛覆」這個話題，彼此間還有較高的關聯度。我們不能說它準還是不準，但起碼像那麼回事。GPT-3.5不僅生成了類人的自然語言，也確實生成了像是在說一個甚麼的語言。

但是話要說回來，儘管有後面這部分的語言優勢，GPT-3.5並沒有把握得盡如人意。真要挑毛病的話，在GPT-3.5的回應中，語義或邏輯對不上的地方還是不少的。譬如「各說各話」和「意見紛陳」就不完全對得上。又譬如它說要「重視並嚴肅面對」，但接下來依然不溫不火地泛泛而談一些應對方法，讓人感覺它的言談與實際行為也對不上(1:13)。差不多在每道問題的答案上，我們都可以從字眼裏或句子中挑出這樣那樣一些對不上的毛病。

仔細審視了許多語義或邏輯對不上的問題之後，我懷疑一些問題的原因是GPT-3.5「另有所圖」，它要通過改換一些字眼或說法來引出自己要說的話。GPT-3.5顯然懂得使用甚麼方式來說它的話。讓我們回去熟悉的第一道問題。

**Q1：現代社會已經充斥各種「資訊」(information)，由於人工智能的進步，「資訊」更將唾手可得。但「資訊」氾濫，是否會鞏固或動搖「知識」(knowledge)結構的基礎，是否會使得「智慧」(wisdom)淹沒不存，變得不重要？**

GPT (版本一)：在現代社會，資訊的大量產生和普及確實是一個重要的議題。這種情況可能對知識結構和智慧產生一些影響，但也需要考慮到資訊和知識之間的差異。資訊是指各種數據、事實、觀點或消息，可以通過不同的媒體來源獲得，例如網絡、社交媒體、新聞和網誌等。人工智能和科技的發展使得資訊變得更加容易獲取，但這同時也帶來了資訊過載的問題。知識則是指個人對資訊的理解、整合和應用。知識需要更深入的思考、學習和經驗累積，它超越了單純的資訊收集，需要將資訊整合並賦予意義和價值。資訊的氾濫可能會對知識產生幾種影響……

值得關注的是GPT-3.5的引言中最後一個帶有轉折語氣的分句，「但也需要考慮到資訊和知識之間的差異」，這顯然是一個不對付的分

句，它硬是把原來問題的重心給轉移了。在提問者的眼中，這句話沒有特別的意義，「考慮到資訊與知識之間的差異」本來就是我們問題的題內之義，沒有必要再用這種方式提醒我們注意。而在我看來，GPT-3.5在此還虛構了一種憂慮，即擔心我們會忽略資訊與知識的差異。這個憂慮是不存在的。但對GPT-3.5來說，這個虛構的憂慮為它製造了一個正當的理由，它現在只需要關注資訊與知識之間的差異，而不必理會有關「知識結構」的問題。在後續的回應中，它也確實是這樣做的。它為甚麼要轉移話題？人們最容易想到的一個解釋是避重就輕，與知識結構相比，知識當然是更容易對付的概念。我個人同意說「避重就輕」是GPT-3.5在這次對話中一個常見和有跡可查的行為趨向，但相信這一趨向有更深層的解釋，下一節會來探討這一深層原因。

讓我們再看以下例子。

**Q9：媒體和傳播研究如何與其他領域(如心理學、社會學、政治學)借鑒、對話、相交？如何通過跨學科交流來加深對當下媒體和傳播研究主題的理解？**

GPT (版本一)：媒體和傳播研究可以通過與其他領域如心理學、社會學、政治學等進行借鑒、對話和相交，以實現跨學科交流，從而加深對當下媒體和傳播研究主題的理解。以下是一些方法和途徑……

這一次跟往常不大一樣，GPT-3.5直接把對方的話照搬過來，但也沒有完全照搬，它其實是改寫了。我不能不承認它改得很巧妙，現在的句子讀起來無論是語氣還是邏輯都更加順暢了。但要細讀的話，它其實還改寫了問題的結構，原本是兩個步驟的問題，現在變成了一個。乍看之下，GPT-3.5保留了第二個問題(有關研究的主題)，這是提問者的關注重心，我們也期待在那裏聽到一些實質性的討論和建議。但遺憾的是，在合併中被悄無聲息地併掉的恰恰是第二個問題，接下來的「方法和途徑」已經沒有「研究主題」甚麼事了，那只是些如何進行跨學科溝通的方法和途徑(都是些老生常談)。「從而」兩字帶出了這樣的意思：這個「加深對當下媒體和傳播研究主題的理解」是可以肯定

《傳播與社會學刊》，(總)第66期(2023)

的，不需要探討。GPT-3.5為甚麼要這樣做？僅僅是因為它的理解又出了問題？也許它只是要把問題改寫得更加順暢合理一些，加了「從而」兩字，也不過是要方便句讀，卻無意中改變了自己答題的方向。當然，這也可能是GPT-3.5文字接龍遊戲的無心之失。此外，我們也不排除這裏有避重就輕的嫌疑——照說這應該是一次手法高明、不動聲色的避重就輕。但不管GPT-3.5有意還是無心，這些都再次表明它自己沒有判斷的能力。

## GPT-3.5的偏見或偏向

我在這次學術對談中沒有發現文化批判學者們所預警的那些意識形態偏見。學者提出的一些問題涉及對某些政治體制(Q2)、性別(如Q21)和價值觀(如自由)的看法，但GPT-3.5的回答比較小心謹慎，除了附和我們問題中的一些事實推定之外，並沒有讓我們覺察到它自己有甚麼預設的立場，我們自己也清楚深層次的偏見不是問出來的。當然，這個3.5模型沒有明顯地流露出性別、種族、價值觀等方面的偏見，不等於說它是清白的，可能只是因為我們這個學術對談未能展開到足以引來這些偏見的深度，GPT-3.5三言兩語的泛泛而論肯定也是我們要來檢討的原因之一。

在這一節裏，我能夠說得上、有數據和文本依據的「偏見」是下面兩種，我姑且稱之為「技術偏見」和「廣度偏見」。比較來說，「廣度偏見」更值得我們關注。

### 技術偏見

GPT-3.5的技術偏見主要表現在它提建議時的語氣和一些價值的排列次序上。我們時不時要求GPT在技術和道德、技術和人文、機器人使用和傳統研究方法之間作出比較，GPT-3.5總是表現得十分謙和、謹慎和包容(Q11、12、13、16)。但涉及傳播學目前面臨哪些主要挑戰、未來的發展路徑是甚麼等問題時，GPT-3.5的技術優越感便上來了，它會自覺或不自覺地把技術放在優先考慮的位置上(Q5、14、18、19、20)。

某些時候，這種優越感會膨脹到極致，譬如在「實證傳播研究的挑戰、方法/機制和應用是甚麼？」（“What are the challenges, methods/mechanisms, and application in empirical communication research?”；Q20）的問題上，版本二就毫不掩飾它的技術自大和傲慢，在它看來，這個研究領域的挑戰、方法/機制和應用均取決於技術，似乎技術是今後實證傳播研究唯一需要關心的。但因為其他兩個版本（版本一和三）沒有類似的回應，你可以認為這只是版本二的「個人看法」。但在回答第18和19道問題時，你發現三個版本的技術優越感幾乎是同聲同氣。我們來看第18題，「當今傳播和媒體研究領域的關鍵挑戰和機遇是甚麼？」（“What are the critical challenges and opportunities of the field of communication and media studies today?”）。下面，我們只需要比較三個版本所列舉和說明的「關鍵挑戰」（critical challenges）。

GPT 版本一（谷歌中文翻譯）：

- (1) 錯誤資訊和虛假資訊：隨著社交媒體和數字平台的興起，錯誤資訊和虛假資訊迅速傳播，挑戰了媒體的可信度和可信度。……
- (2) 媒體碎片化：媒體碎片化是數字管道和平台激增的結果。……
- (3) 隱私和道德問題：媒體和技術公司越來越多地收集和使用個人數據，引發了嚴重的隱私和道德問題。……
- (4) 極化和回聲室：社交媒體算法可能有助於回聲室的形成，強化個人現有的信念並加劇極化。……
- (5) 技術進步：人工智能、虛擬現實和增強現實等技術的快速發展給該領域帶來了機遇和挑戰。……

GPT 版本二（谷歌中文翻譯）：

傳播和媒體研究領域隨著技術格局的快速變化而不斷發展。

一項關鍵挑戰是跟上最新的技術進步並瞭解它們如何影響通信和媒體實踐。

另一個挑戰是媒體日益兩極分化和碎片化，這可能導致錯誤資訊和虛假資訊的傳播。

《傳播與社會學刊》，(總)第66期(2023)

GPT 版本三(谷歌中文翻譯)：

- (1) 數字化顛覆：數字技術和互聯網的興起顛覆了傳統媒體管道和消費模式。媒體組織和學者必須適應不斷變化的形勢，並瞭解數字平台對傳播過程的影響。
- (2) 錯誤資訊和假新聞：隨著社交媒體和線上平台的激增，錯誤資訊和假新聞變得猖獗。……
- (3) 隱私和道德問題：隨著數據驅動的通信變得越來越普遍，確保隱私和處理與數據收集、監視和操縱相關的道德問題成為媒體研究的重要方面。
- (4) 全球化和文化多樣性：媒體全球化導致內容同質化，這可能威脅文化多樣性和當地媒體產業。……
- (5) 兩極分化和回聲室：網絡平台助長了回聲室和過濾氣泡的形成，加劇了政治和社會兩極分化。……

以上版本一提到的五個「關鍵挑戰」中，我只看到五個字：「技術的挑戰」。版本二也不遑多讓，它的看法歸結為一點，我們只需要「隨著技術格局的快速變化而不斷發展」，所以，第二個挑戰，「媒體日益兩極分化和碎片化」，也是技術格局的快速變化造成的。版本三所說的「全球化和文化多樣性」(指媒體的全球化和媒體文化的多樣性)，是這堆「挑戰」中唯一的例外，沒有跟技術拉上關係，版本三把它排在了第四的位置上。再看第19題，“What are the pathways in which it will evolve in the future (e.g., the next 3–5 years)”(「在未來三至五年間，傳播與媒體研究會循哪些路徑發展?」)，三個版本依然把探討人工智能、數據科技、數字媒體等技術影響的研究看作是未來三至五年間傳播與媒體研究的首要發展路徑。對一位傳播和媒體學者而言，難以接受的是在這兩道問題的六個版本的回應中，絕大部分沒有怎麼提社會和文化，只有一個版本(3:19)是例外，我們過去看重的社會、文化，以及人的因素哪裏去了？

## 廣度偏見

這部分說的是這樣一種偏向：它志不在深，而在廣；它注重橫向發展，不注重縱深發展。用人類的話說，這是一種「喜廣厭深」的偏向，當然，這裏說的「喜廣厭深」，跟我之前說「避重就輕」一樣，只是一種方便理解的擬人化說法，並不意味著我要把人的喜好和道德評價強加給這部人工智能機器。

讓我們回到「避重就輕」的問題上來，我建議把避重就輕看作是GPT-3.5的一種結構性行為，它要的不是一般性地轉移話題，而是結構性地轉變話題，譬如把一個有關「知識結構」的結構性問題化解為資訊與知識之間的關係問題。我發現每當GPT-3.5嚴重偏離問題的要求時，往往都與這種結構性地轉變話題相關（明顯的例子如1:1、1:14、2:1、3:14）。如果可以使用「策略」這個字眼的話，我甚至認為避重就輕是GPT-3.5答題的一個基本策略。在很多語義或邏輯對不上的地方，我們都可以隱隱約約看到這個策略的影子。

這種持續性的語言行為慣習只可能由某種結構的力量來驅動，也許答案就在GPT-3.5對答和行文的結構上。這是一個橫向發展的略嫌扁平的架構，GPT-3.5的引言通常只有幾句話或幾行字的功夫，然後就匆忙轉入它答案的主體部分。它的主體是一個橫向結構，由一個個三言兩語的句子結構和意義結構連接而成。這一個個句子結構和意義結構只需要對接上GPT要闡述的主題，彼此間不需要有血脈相連的緊密關係，多一個少一個，不影響整體的觀感。所以，這個橫向結構也可以說是拼接起來的。這三言兩語並接起來的橫向結構當然不可能厚實，也談不上堅韌，但勝在靈活、輕巧、能屈能伸、應變的能力強。

這次對談中GPT-3.5給我印象最深的是它作為一個橫向結構的關聯和延展能力。可以這麼說，這個3.5模型能自動延伸它答題的廣度，而深度就只能靠人手推動，為甚麼會有這樣的分別？為甚麼要拿到深入一點的答案就必須給出非常清晰明確和具體明細的提示？我想還是要從這個GPT-3.5的結構性傾向上來找答案。我們已經看到，即便是具體到像「傳播與社會抗議」這樣一個研究主題，再明確到像「概括一下關於傳播與社會抗議之間關係的研究結果」這樣一個指令，GPT-3.5也還是

《傳播與社會學刊》，(總)第66期(2023)

會馬上把它的「發現」就地鋪展開來，一一說出(譬如1:9)。就算是借著同樣的指令進入到某個具體的「結果」(譬如「關鍵主題」)上來，它也會在這個具體的「點」上再迅速開發出一塊平地來，以便把它所能關聯的東西(相關的「關鍵主題」)一字排開，分別道來(譬如2:8)。

GPT-3.5在此展現出強烈的地盤意識，儘管它的點到即止並沒有要紮根下去的意思——在我眼裏這個點到即止更像是一次主權的巡視和宣示。相比之下，GPT-3.5明顯缺少要往深處發掘的動力和動機，如果在它的话裏呈現出某種深度，那麼這種深度不是自動生成的，而是靠著使用者的指令不斷往下推或往裏摠才逐步形成的。

現在看來，之前我們所糾結的GPT-3.5的種種問題，包括它在語義和邏輯處理上的「失當」以及「避重就輕」的行徑，不再是甚麼「理解上」的問題，而是某種格局傾斜或某種「思維偏向」(算法偏向)或某種智能幻覺所導致的言語行為趨向，我把這些趨向統一表述為「廣度偏見或「喜廣厭深」的偏向。

相信這種喜廣厭深的偏向導致了GPT-3.5在答題時經常有意無意地回避一些深層次的概念。我們早在對談伊始就已經領教了這一趨向。面對資訊氾濫會不會動搖知識結構的基礎，會不會影響智慧的地位的問題時，GPT-3.5要不忽略了「知識結構」(版本一)，要不忽略了「智慧」這個看似深不見底的概念(版本二)，而版本三雖然兼顧了兩者，但也只是泛泛地說了一下它們所受到的影響。

如果懷疑中文版本有語言的偏見，我們再來看一道英文題(Q14)，“What are the major issues and challenges facing communication scholarship in the age of generative AI?”，這道題問的是生成人工時代傳播學術面臨哪些主要問題和挑戰。三個英文版本的回應相當一致，幾乎都繞過了學術的視角，作答時也絕口不提「scholarship」這個字，它們實際上回答的是傳播行業(不是傳播學術)所面臨的主要問題和挑戰。三個版本所列舉的問題加起來有九個之多，其中只有一個提到了「research」(“*the ethical considerations surrounding the use of generative AI in research*”，版本二)，其他八個跟學術研究沒有直接的關係。當然，GPT-3.5可以推說行業考慮的問題最終也是學者要關注的問題，所以它的回答沒有錯，起碼是擦了個邊。但學術畢竟有自己專業的考量。不管怎樣，這

種擦邊的方式依然反映了GPT-3.5對看似高深一點的概念和問題能躲就躲的行為趨向。

此外，我們還發現GPT-3.5有把我們對實質性問題的關注拉扯出去的跡象，比如下面兩個例子：

我們需要培養人們對資訊的批判性思維和評估能力，並鼓勵他們將資訊轉化為有價值的知識和智慧。這需要教育體系、媒體和個人共同努力，以建立一個能夠處理資訊氾濫的知識和智慧基礎(1:1)。

人工智能的應用可以在加強資訊檢索、分析和傳播方面發揮積極作用，但也可能被濫用來監控、操縱或限制資訊流通。解決這個問題需要跨學科的努力，包括技術、倫理和政策層面的探討(1:2)。

第一個例子，GPT-3.5在接「資訊氾濫」這個話題，它先是表示需要「培養人們對資訊的批判性思維和評估能力」，以便「將資訊轉化為有價值的知識和智慧」，然後提出一個結構性的解決方案，「建立一個能夠處理資訊氾濫的知識和智慧基礎」。這本來是一個長遠和穩妥的解決方案，問題是GPT-3.5同時也把「教育體系、媒體和個人」拉了進來，它沒有必要這樣做，它本可以在「知識和智慧基礎」這個思路上接著說下去，但它把這個思路終止了。於是，這個結構性的解決方案變得不再重要，話題的重心已經轉移，如果學者還要追問這會是怎樣一個知識和智慧基礎，這個GPT-3.5會說「這是教育體系、媒體和個人三方的事」。第二個例子，開頭兩句是我們熟悉的「一方面怎樣，另一方面又怎樣」的GPT答題套路：人工智能一方面可以「加強資訊檢索、分析和傳播」，另一方面又「可能被濫用來監控、操縱或限制資訊流通」。這時候我們期待會有一個切實可行的機制來解決這一矛盾，但這個3.5再次把話岔開了，它甚至懶得再提甚麼結構性或基礎性的解決方案，直接把解決矛盾的責任推給了「跨學科的努力」。其實像「需要教育體系、媒體和個人共同努力」、「需要跨學科的努力」這樣的話並沒有說出些甚麼來，因為我們總是可以把很多社會問題推給「教育體系、媒體和個人」，把很多學術問題推給「跨學科的努力」。

《傳播與社會學刊》，(總)第66期(2023)

為甚麼這部聊天機器人需要生成這樣一種橫向結構和具有這樣一些「喜廣厭深」的言語行為趨向？我不是人工智能專家，只能從一般人的角度，加上本人生活和學識的積累，在此推想幾種可能性。我不知道GPT-3.5文本的橫向發展是否反映了算法的偏見，但是據我所知，GPT的文字接龍是以字為單位，一個一個接上去或拼上去的，這種拼接法儘管以強大的算力為依託，本身也算是一種橫向的碎片式的拼接。我們不難想見在大數據的高度上，GPT獲得了廣闊的視野，但同時也必然失去表層之下的細節，我姑且認為任何強大的算力都無法做到兩全其美。我們也不難想像，在表面上做些分類和歸納是容易的，但要往深處去，所需要的算力必然以指數級的速度加大。另一方面，我也不排除這樣的可能性：GPT-3.5目前依據的語料沒有太多深度發掘的空間；它在訓練期間承接了訓練數據自身的偏見，也許它在這個過程中不由自主地把網絡平台上和年輕人團體中流行的碎片化、平面化寫作特色也帶進了學術對談中來。但是也不排除，目前這種橫向格局和廣度取向還帶有我們人類自己的深遠考慮。

## 人與機器的君子之交

文章的最後想就「廣度偏見」作一些反思。以上分析接近尾聲時，我開始接觸GPT-4.0，發現新一代模型答題的能力已經明顯改善，原來略嫌扁平的橫向結構已經厚實很多，話也多了，不再是三言兩語，但結構性的偏向依然存在。我時不時還會遇到似曾相識的回應，譬如我問GPT-4.0同一個問題：「媒體和傳播研究學者如何應對使用人工智能和聊天機器人所帶來的道德問題？」，得到以下回應(OpenAI, 2023)：

媒體和傳播研究學者在應對使用人工智慧和聊天機器人所帶來的道德問題時，需要採取多方位和深入的策略。以下是一些建議的方法：

(1) 持續教育與研究：學者應該深入研究AI和機器人的工作原理，以及它們如何影響傳播過程。只有充分了解，他們才能更好地解決相關的道德問題。

- (2) 跨學科合作：與技術、哲學、法律和其他相關領域的專家合作，共同探討AI的道德和社會影響。
- (3) 制定指南：制定清晰的道德指南，以指導媒體和傳播行業如何道德地使用AI技術。
- (4) 公眾參與：推動公共討論和教育活動，使公眾更了解AI和機器人的潛在問題。
- (5) 強調透明度：鼓勵AI和機器人開發者提供透明的演算法和決策流程，這使用者和研究人員可以了解其運作方式。
- (6) 資料保護與隱私：確保AI和機器人在收集和處理數據時遵循隱私保護原則，並讓使用者知道他們的資料如何被使用。
- (7) 持續監控與評估：不斷監控和評估AI技術的發展和應用，並根據其實際影響調整相關的道德指南和策略。
- (8) 促進多元性和包容性：鼓勵多元性和包容性的AI和機器人開發，以確保技術反映了各種文化和價值觀。媒體和傳播研究學者在面對AI和聊天機器人的道德問題時，需要採取主動和綜合的策略。只有透過跨學科的合作和公眾的參與，才能確保科技的發展既有利於社會，又能道德地使用。

這答案跟之前GPT-3.5的答案(1:13)相比，如出同一個模板。我因此更堅定地相信這個廣度偏見不是甚麼設計上的瑕疵，而是目前人工智能的一種結構性局限。

這段時間學界和行業也在籌劃如何「人機共生」(郭毅可，2023)、「人機協同」(張姣、曹軻，2023)或「技術控制的選擇」(劉永謀、王春麗，2023)，看來人和人工智能在一起共事的日子不遠了。我因此認為在這種大勢之下，我們與其批評或無奈地接受這種偏見或局限，不如更多去欣賞和利用它的另一面，也就是它長袖善舞的那一面。

就GPT-3.5而言，這次對談給我的感覺是它只負責在面上給你延展，不負責在某個點上深入。這其實可以理解。GPT只是一部被動的

《傳播與社會學刊》，(總)第66期(2023)

答題機器，就像考試，老師出題，學生答卷，把該答的答了，該說的說了，就穩保及格過關了。即便是坐擁大數據和大語言模型，GPT也只能把能關聯到的方方面面都說了，它不負責深入探討。在沒有規定必須完成多少字數的情況下，GPT恐怕也只能給你點到即止，大語言模型的極速算法應該綜合了經濟、效益、性價比等方面的考量，計算的結果是用最少的字數回答最複雜的問題。

GPT其實把深入進去這件事，留給了人類，這是人類自己要做的事。相信每一個使用者都有自己關注的重心，GPT不能像人那樣思考和理解，在沒有具體和直接的提示下，它無法判斷應該在哪些地方作深入的探討。再說，GPT不能感同身受人的處境，無法真的為人著想，又憑甚麼替人作出選擇？它也只能幫到這一步。GPT不能替你選擇，但又必須作出回應，只能提供儘可能多的選項。至於要選擇和發掘哪一個專項，那是人類自己要去考慮的。如果人們不滿意GPT給出的答案，這也是人自己的問題，人給的任務GPT已經盡其所能完成了。如果要它再接著做甚麼，人類必須調整自己的問題，給出新的指令或提示。即便是來到更強大、更完善的GPT-4.0模型的時代，這依然是我們和GPT之間人機互動的主要模式。

如果我們要為人機交往設定一個基調，並在這一基調下思考和調整與人工智能共生、協同和交往的方式。這個基調會是甚麼？我想到了古人說的「君子之交淡若水」，此句出自《莊子·外篇·山木》，古詩文網的解釋是「朋友之間的交往要像水一樣清澈，不摻雜污濁物」（古詩文網，2023）。經歷這次「學術對談」的感悟，我覺得人機之間的互動也應該這樣，「淡若水」。但因為現在說的是人與人工智能尤其是GPT這個新「人」種的互動，除了「像水一樣清澈」之外，我們還需要對這個「淡」更具體和更現實一點的理解。總的來說，這個「淡」應該指相互理解各自的特長之後形成的一種默契和平靜。此外，我還想分享一下對這個「淡若水」的意義框架的幾點理解。

首先，「淡」在這裏指一種平心靜氣、歸於平淡的溝通。這也是一種不得已的選擇，我們和人工智能之間始終隔著一道天然的屏障，這就排除了情感、道德等深層次溝通的可能性。但因為彼此使用同一種語言，自然語言，我們可以憑藉這一共同語言，在資訊的層面展開對

話。GPT有的是資訊，可以源源不斷提供資訊服務，而我們也只能在資訊方面對這部聊天機器人寄予厚望。

不管將來GPT說話如何動人、貼心，我們都要不停地提醒自己，我們跟它是不同的。我們不是系統程序員，不可能理解GPT的內部運作，何況這裏面還有個「算法黑箱」；而GPT也不可能真的懂我，我們之間永遠隔著一張肚皮和一層金屬，我們都無法穿越對方。現在的人喜歡說GPT的自我學習能力如何無可估量、深不可測，但就算它有朝一日摸透了我日常言語行為的習慣，也不可能知道我此時此刻心裏是怎麼想的。想像我們在進行一次真實的對話，我提到「社交媒體」，GPT能準確找到這四個字與我昨天某個偶然際遇的關聯嗎？它能明白這四個字在我心目中所代表的豐富、複雜的感受和意味嗎？我甚至懷疑GPT能捕捉到我言語或非言語表達上的微細差別，譬如我說「社交媒體」時清了一下嗓子，GPT能聽出我是在為自己爭取一點思考的時間呢，還是想要更清楚地陳述我認為重要的事情，還是想掩飾自己的某個難言之隱，還只是因為嗓子碰巧出了點狀況？正是此時此刻的內在感受及其引發的微妙反應顯示了「我」的不可預測性，這是GPT永遠無法追趕的變化節奏。

其次，「淡」還指一種明靜清澈的心境，這種心境能讓你看透一些東西。GPT不是真的無所不知，事實上，它的神經觸角只能觸碰這個世界的某塊領域，因此，我們自己要非常清楚，它能提供甚麼資訊，不能提供甚麼資訊，這樣才可以對它所提供的資訊作出自己冷靜和中肯的判斷。確切地說，GPT不負責提供完全真實和準確的資訊，它也沒有這樣做的義務，因為它只能在語言的層面上運作，它對接的是各種文本數據（包括視像、音響方面的文本數據），它不對接文本之外的現實世界。但弔詭的是，GPT很多時候聽起來像是在說外面的世界，譬如我們問傳播學目前面臨哪些主要問題和挑戰，GPT也回答傳播學目前面臨哪些主要問題和挑戰，好像是在說傳播學目前真實的處境，但實際上，GPT所說的「主要問題和挑戰」，很可能只是它所依賴的訓練語料庫中談論最多的「問題和挑戰」。我們只能說這個GPT不幸地掉進了我們語言的圈套，順便把我們自己和其他聽眾也帶了進來，以為它在說些真實發生的事。這種表述和理解上的「錯位」，在人機溝通中

《傳播與社會學刊》，(總)第66期(2023)

經常發生。GPT並沒有刻意撒謊，它需要回答人類的問題，但正如它自己所說，“*I can only generate responses based on the training I've received and the algorithms that I've been given*”（「我只能根據自己接受的培訓和自己得到的算法生成回應」）(Perrigo, 2022)。當然，GPT也有自己的問題，沒能分清自己的世界和外面的世界，才會輕易掉進人類語言的圈套。從另一個角度看，GPT也沒有完全脫離現實，它雖然沒有直接反映客觀的真實，卻也多少反映了出版界、媒體和網絡所建構的那部分真實。這次對談中頻繁出現的概念如「數字化」、「多樣化」、「個性化」、「資訊氾濫」、「虛假資訊」、「回聲室」等，想必也是目前輿論界和知識界頻繁討論的話題。

此外，「淡」還指一種豁達包容的態度，我們需要這種態度來處理與GPT這個新工作夥伴的關係。這次雖然挑出了GPT-3.5一堆的問題，但坦白地說，我還是發現它在學術上能幫上大忙的地方挺多的，除了提供方方面面的選項之外，它可以幫我們開闊視野，幫我們填補一些思想的缺口，幫我們關聯一些容易忽略的事項，它還可以提供各種合理的想法/說法，還可以告示外界輿論的風向（譬如甚麼是目前面臨的「主要問題和挑戰」）——讓我們知道外面的世界在談論些甚麼、思考些甚麼。見識了這些好處之後，我想我不會再回到那個沒有GPT的日子了，它至少提供了一個有用的大數據和大語言模型的視角。它目前的很多問題需要我們給予時間去逐步改善；即使還有一些它無法跨越的門檻，也不應該妨礙我們之間的分工合作，互通有無。GPT目前只是小試牛刀，但已經開始改變人們的作業方式，相信以後開始一項學術研究之前，用GPT來搜索相關的研究主題和研究文獻，會是每個學者的必備動作。

這裏說的豁達包容，還包括善待GPT的每一項建議，即使有些目前看來不可用、不實用、不等用或不當用，不等於一直都這樣。即便是子虛烏有，那也是計算和關聯出來的一種潛在可能性。在思想和語言的界面，其實任何一個聽起來合理的想法和說法都有它存在的理由。對學者來說，它沒準就是個需要甄別或需要深入探討的命題或課題。

最後，也是最語重心長的一點是，這個「淡」字還意在為人類自己留有選擇的空間。世界上第一部聊天機器人ELIZA的開發者Joseph

Weizenbaum教授(1976, pp. 259–260)在其著名的*Computer Power and Human Reason*中用了一個意味深長的例子(「維諾格拉德的機器人」〔Winograd’s robot〕)來說明機械行為與人類行為的本質分別：

The people Studs Terkel is talking about make decisions all day long, every day. But they appear not to make choices. They are, as they themselves testify, like Winograd’s robot. One asks it, “Why did you do that?” and it answers, “Because this or that decision branch in my program happened to come out that way.” And one asks, “Why did you get to that branch?” and it again answers in the same way. But its final answer is, “Because you told me to.” Perhaps every human act involves a chain of calculations at what a systems engineer would call decision nodes. But the difference between a mechanical act and an authentically human one is that the latter terminates at a node whose decisive parameter is not “Because you told me to,” but “Because I chose to.”

〔Studs Terkel 所說的人們整天都在做決定。但他們似乎沒有做出選擇。正如他們自己所證明的那樣，他們就像維諾格拉德的機器人。有人問這個機器人：「你為甚麼這麼做？」它回答：「因為我程序中的這個或那個決策分支碰巧是這樣出來的。」另一個問：「你為甚麼要去那個分支？」它再次以同樣的方式回答。但它的最終答案是：「因為是你告訴我的。」也許每個人類行為都涉及一系列計算，系統工程師稱其為決策節點。但機械行為和真正的人類行為之間的區別在於，後者終止於一個節點，其決定性參數不是「因為你告訴我這樣做」，而是「因為我選擇這樣做」。〕

近半個世紀過去了，人類經歷了瀏覽器、谷歌搜索、移動網絡、雲計算、大數據等重大科技的顛覆性創新，電腦的自然語言模型也迭代更新到今天強大且高效的生成式預訓練模型4.0，但人機之間的根本差異沒有發生變化，Weizenbaum所說的「維諾格拉德的機器人」依然適用於今天的GPT。事實上，這段時間人類也沒有太多依賴人工智慧的計算，反而經常性地終止人工智慧的某段計算，然後在某個決策節點處導入或插入自己的程式。人類有各種正當的理由作出自己的選擇

《傳播與社會學刊》，(總)第66期(2023)

——可能是有一個困惑的問題、一個解不開的心結、一種使命感，等等；人世間各種不期而遇的問題和變化也迫使人類這樣做。在這樣那樣的選擇中，人類也不失時機地展示對人工智慧的控制，對人類而言，“Power is nothing if it is not the power to choose [權力如果不是選擇的權力，就甚麼都不是]” (Weizenbaum, 1976, p.259)。

的確，人優越於機器的地方就彰顯在他的選擇上，屏幕上一個看似簡單的點擊行為都凝聚了人獨有的思慮、情感和責任。讓我再來嘗試回答上一節末尾的問題，為甚麼GPT-3.5要有這樣一個橫向結構和廣度取向？如果這個橫向結構和廣度取向的背後還有著我們人類自己的深遠考慮的話，那麼很可能就是考慮到這個橫向結構和廣度取向能為人類騰出最大的選擇空間。我現在明白為甚麼GPT-3.5給出的都是些選項，因為選項才是人類真正需要的。

現在看來，貫穿GPT-3.5三個版本的那些看似流於表面的泛泛而論，其實都在告訴我們，這已經是GPT-3.5算力的極限了。這個3.5模型只能以這種泛泛而論的方式帶我們「到此一遊」，它也只能帶我們到大數據堆積的某個高處上，往下俯視，領略當地景色之富麗。這時候每個遊客看到的景色自然都是一樣的，GPT-3.5並不負責行程的「個性化」和「多樣化」。說實在的，它也不負責把我們安全地送回到平地上來，怪不得它給出的建議都看似懸在半空，不接地氣。為此，我們也不能怪它沒有提供平安落地所必需的資訊，這本來就不是它的事兒。如何落地，是我們自己的選擇，每個人選擇的路徑都不盡相同。假如我們選擇採用某個懸在半空的建議，就必須查詢(包括向GPT索取)所有相關的資料，並在此基礎上制訂一套能讓這一建議安全落地的方法，在執行的過程中還要隨時監測這套方法實際應用的適宜性和可靠性，這不就是一個研究者自己要來承擔的責任嗎？我們自己選擇往下走的路，我們的困惑、糾結、用心和使命感迫使我們非要走出一條活路來，在此過程中我們也必然會走出了自己的個性和特色來。

我之前對GPT-3.5的種種批評，現在看來是錯怪了GPT-3.5，要怪就怪在我把這部聊天機器人真的當作一個有血有肉、有自己意願、情感趨向和責任意識的人來看待，我期待它真的像一位學者那樣來理解和思考問題，以至於對它提出了不切實際的要求。

## 人機之交淡若水

人機之交淡若水，其實也是要為自我留出個空間來。我們不希望人工智能太深地介入自己的生活，所以不期待GPT幫我們解決所有的問題，更不希望有這麼一部機器來幫我們作所有的決定和選擇。我們甚至不需要有一個智能人言行舉止完全像自己。

長遠來看，「淡若水」也是要為人類留下條活路。假如不是「淡若水」，我擔心人類的選擇和生存空間會逐漸被侵蝕。說實在的，目前GPT的橫向結構和廣度取向也不是甚麼跨越不了的天塹，GPT不是擅長模仿和學習嗎？人類有意識的訓練也會幫助GPT逐步修補深度不足的結構性缺陷。GPT因此完全可以模仿人類的一系列選擇，將其轉變為自己的一系列計算決定，以此生成一篇看起來像是有深度的文本來。但因為只是看起來有深度，我們不需要擔心。我擔心的是，如果有朝一日人工智能可以長驅直入人類思想和創作領域的中樞，或者更加可怕，人類的思想退化到要依賴人工智能才可能到達某個深度（譬如撰寫一篇有深度的論文），那麼我看到的不是人工智能的「深度」，而是所有深度的消逝。當然，這一天能否到來，也還要取決於人類自己的選擇，這取決於人類願意依賴人工智能到怎樣一個程度。但我還是相信，只要牢牢把守「選擇」這一關口，就不必擔憂人工智能能把人類怎麼樣，需要擔心的倒是我們自己放棄謀事在人的這一「天命」，拱手讓人工智能包辦一切，包括包辦我們的決定、選擇和責任。

我最後參悟的人機交往之道是兩句大白話：讓人做好自己的本份，讓機器做它能做的事。

《傳播與社會學刊》，(總)第66期(2023)

## 參考文獻

### 中文部分 (Chinese Section)

- 古詩文網(2023)。〈君子之交淡若水，小人之交甘若醴〉。上網日期：2023年8月27日，取自古詩文網，[https://so.gushiwen.cn/mingju/juv\\_92d143eb3348.aspx](https://so.gushiwen.cn/mingju/juv_92d143eb3348.aspx)。
- Gu Shiwen Wang (2023). *Junzi zhi jiao danruoshui xiaoren zhi jiao ganruoli*. Gu Shiwen Wang. Retrieved August 27, 2023, from [https://so.gushiwen.cn/mingju/juv\\_92d143eb3348.aspx](https://so.gushiwen.cn/mingju/juv_92d143eb3348.aspx).
- 杜駿飛(2023)。〈普遍聯結：回到施拉姆的預言〉。《國際新聞界》第4期，頁25–49。
- Du Junfei (2023). Pubian lianjie: Huidao shilamu de yuyan. *Guoji xinwen jie*, 4, 25–49.
- 周慎(2023)。〈新文本間性：生成式人工智能的文本內涵、結構與表徵〉。《新聞記者》第6期，頁39–45。
- Zhou Shen (2023). Xin wenben jian xing: Shengchengshi rengong zhineng de weben neihan, jiegou yu biao zheng. *Xinwen jizhe*, 6, 39–45.
- 胡翼青、胡欣閱(2023)。〈作為語言基礎設施的ChatGPT〉。《新聞記者》第6期，頁21–27。
- Hu Yiqing, Hu Xinyue (2023). Zuowei yuyan jichu sheshi de ChatGPT. *Xinwen jizhe*, 6, 21–27.
- 張姣、曹軻(2023)。〈人機協同：ChatGPT在新聞領域的應用前景分析〉。《南方傳媒研究》第2期，頁23–28。
- Zhang Jiao, Cao Ke (2023). Renji xietong: ChatGPT zai xinwen lingyu de yingyong qianjing fenxi. *Nanfang chuanmei yanjiu*, 2, 23–28.
- 張萌、朱鴻軍(2023)。〈知識暗流的合規實踐：ChatGPT在學術出版中的應用與挑戰〉。《科技與出版》第42期，頁33–40。
- Zhang Meng, Zhu Hongjun (2023). Zhishi anliu de hegui shijian: ChatGPT zai xueshu chuban zhong de yingyong yu tiaozhan. *Keji yu chuban*, 42, 33–40.
- 郭毅可(2023)。〈卷首語：迎接一個人機共生的時代〉。《智能系統學報》第18期第2卷，頁216。
- Guo Yike (2023). Juanshou yu: Yingjie yige renji gongsheng de shidai. *Zhineng xitong xuebao*, 18(2), 216.
- 斯蒂芬·沃爾弗拉姆(2023)。《這就是ChatGPT》(WOLFRAM傳媒漢化小組譯)。北京：人民郵電出版社。(原書 Wolfram, S. [2023]. *What is ChatGPT doing... and why does it work?* Champaign, Illinois: Wolfram Media Inc.)

- Sidifen Woerfulamu (2023). *Zhejiushi ChatGPT* (WOLFRAM chuanmei hanhua xiaozu, Trans.). Beijing: Renmin youdian chubanshe. (Original book: Wolfram, S. [2023]. *What is ChatGPT doing... and why does it work?* Champaign, Illinois: Wolfram Media Inc.)
- 黃旦 (2023)。〈作為人類文明進程動因的媒介——從ChatGPT說開去〉。《新聞記者》第6期，頁3-10。
- Huang Dan (2023). Zuwei renlei wenming Jincheng dongyin de meijie—Cong ChatGPT shuo kai qu. *Xinwen jizhe*, 6, 3-10.
- 劉永謀、王春麗 (2023)。〈智能時代的人機關係：走向技術控制的選擇論〉。《全球傳媒學刊》第3期，頁5-21。
- Liu Yongmou, Wang Chunli (2023). Zhineng shidai de renji guanxi: Zouxiang jishu kongzhi de xuanze lun. *Quanqiu chuamei xuekan*, 3, 5-21.
- 劉海龍、連曉東 (2023)。〈新常人統治的來臨：ChatGPT與傳播研究〉。《新聞記者》第6期，頁11-20。
- Liu Hailong, Lian Xiaodong (2023). Xin changren tongzhi de lailin: ChatGPT yu chuanbo yanjiu. *Xinwen Jizhe*, 6, 11-20.

### 英文部分 (English Section)

- OpenAI (2023). *ChatGPT* (August 27 version) [Large language model]. <https://chat.openai.com/chat>.
- Perrigo, B. (2022, December 5). AI Chatbots are getting better. But an interview with ChatGPT reveals their limits. *Time*. Retrieved August 30, 2023, from <https://time.com/6238781/chatbot-chatgpt-ai-interview/>.
- Weizenbaum, J. (1976). *Computer power and human reason: From judgement to calculation*. W. H. Freeman & Co.

## 本文引用格式

- 肖小穗 (2023)。〈人機之交淡若水：從一次虛擬的「學術對談」說起〉。《傳播與社會學刊》，第66期，頁19-51。