學術對談

學界和公眾對人工智能認知的再思考：
當今之境，未來之航

對談人：Matthieu GUITTON、鍾布、宋韻雅
編輯：宋韻雅
翻譯：劉曉銘、李佳瑞

Matthieu Guitton 教授

鍾布教授
（Prof. Bu Zhong）

「人工智能之所以『人工』，是因為它由人類創造和製造。這並不意味著人類的優越性或將人工智能的物化；相反，它意味著人類以

Matthieu GUITTON，拉瓦爾大學醫學院與國際研究院教授、英國皇家人類學會會員、泰國國立法政大學Bualuang ASEAN講座教授。研究興趣：個體網絡行為、健康信息技術、VR技術在神經感覺障礙修復中的應用。電郵：matthieu.guitton@fmed.ulaval.ca

鍾布，香港浸會大學傳理學院院長及教授、賓夕法尼亞州立大學唐納德·P·貝利薩里奧傳媒學院教授。研究興趣：健康傳播、社交媒體、心理健康、信息處理、技術對個人和社區轉型的影響。電郵：zhongbu@hkbu.edu.hk

宋韻雅，香港浸會大學傳理學院教授。研究興趣：國際傳播、數字媒體、社交媒體分析。電郵：yunyasong@hkbu.edu.hk

一種藝術的方式創造物體。ChatGPT 是一個文本生成器，而不是人工智能。如果我們渴望超越文本生成器的能力，就必須將社會科學和人文學科的洞察融入人工智能研究中。只有這些領域能教導我們甚麼是人類，我們必須將這些經驗整合到程式碼中，才能超越目前的層次。所有這些因素表明，我們需要社會科學的程度比之前所認為的還要多。」—— Matthieu Guitton 教授

「很多『概念』只是風靡一時，而後很快銷聲匿跡，因而我們無法預測明年是否還會討論元宇宙或 ChatGPT。然而，這些技術的運作機制不僅關乎技術本身，還與由 AI 所驅動的溝通和互動有關。我們主要的興趣在於人類如何受到技術的影響，不論其載體是人、3D 動畫，還是第二人生。因此，我們的目標是探討這項新技術如何改變溝通方式，以及如何使其助益社會。」——鍾布教授

Dialogue

Revisiting Scholarly and Public Perceptions of Artificial Intelligence: Current State and Future Trajectories

Discussants: Matthieu GUITTON, Bu ZHONG, Celine Yunya SONG
Editor: Celine Yunya SONG
Translators: Xiaoming LIU, Jiarui LI

Abstract

The emergence of large language models, such as ChatGPT, has sparked widespread interest and debates surrounding artificial intelligence (AI). Amidst this fervor, scholars and the general public hold divergent, and at times conflicting, views and expectations of AI. The complex nature of AI and the need for its responsible and thoughtful development and deployment have become highly debated topics that require immediate attention. To gain an in-

Matthieu GUITTON (Full Professor). Faculty of Medicine and Graduate School of International Studies, Université Laval (Quebec City, QC, Canada); Fellow of the Royal Anthropological Institute of Great Britain; Bualuang ASEAN Professor Chair at Thammasat University (Bangkok, Thailand). Research interests: cyberpsychology and cyberbehavior, human behavior in virtual spaces, virtual communities, avatars and artificial humans.

Bu ZHONG (Professor and Dean). School of Communication, Hong Kong Baptist University. Research interests: AI imaginaries, health communication, social media, mental health, information processing, effects of digital technology on individuals and communities.

Yunya SONG (Professor). School of Communication, Hong Kong Baptist University. Research interests: global communication, digital media, social media analytics.

depth understanding of the opportunities and challenges brought by AI, we have invited two esteemed scholars with interdisciplinary backgrounds: Prof. Matthieu Guitton and Prof. Bu Zhong. As seasoned academic editors, they also bring unparalleled expertise and provide unique perspectives on academic publishing. This conversation does not seek to provide a definitive answer; instead, it aims to maintain an open and realistic perspective in light of the ongoing evolution of AI. Both the public and the academic community are encouraged to reassess the challenges posed by AI development, define its scope and limitations, and work together to foster its responsible growth.

Citation of this article: Guitton, M., Zhong, B., & Song, C. Y. (2024). Revisiting scholarly and public perceptions of artificial intelligence: Current state and future trajectories. *Communication and Society, 68,* 1–44.

Matthieu Guitton 教授簡介

Matthieu Guitton 是拉瓦爾大學（Université Laval）教授、英國皇家人類學會會員、泰國國立法政大學（Thammasat University）Bualuang ASEAN 講座教授、加拿大 CERVO 腦科學研究中心的高級研究員及團隊負責人。此外，他還擔任 *Computers in Human Behavior* 和 *Computers in Human Behavior: Artificial Humans* 期刊的主編。Matthieu Guitton 教授的研究興趣包括個體在網路和虛擬空間中的行為，關注個體的心理健康、社會動態、倫理和計算機輔助的溝通。他的研究也關注健康科學中的信息技術，探討遠程醫療、網路使用以及互聯網對健康醫療系統的影響。此外，他還致力於研究如何使用虛擬實境工具提高神經感覺障礙的修復能力。

鍾布教授簡介

鍾布現任香港浸會大學傳理學院院長和教授，同時也擔任賓夕法尼亞州立大學唐納德 · P · 貝利薩里奧傳媒學院（Donald P. Bellisario College of Communications, The Pennsylvania State University）的教授，以及賓夕法尼亞州立大學資訊科學與技術學院的兼職教授。他曾擔任國際中華傳播學會會長、美國新聞與大眾傳播教育學會（Association for Education in Journalism and Mass Communication, AEJMC）大眾傳播與社會分會主席，目前擔任 *Computers in Human Behavior* 期刊的高級編輯。鍾布教授的研究興趣涵蓋傳播、科技和人類行為的融合。他的研究應用決策理論和跨學科方法，分析個人與數字技術的互動，以及這些互動對行為、心理健康和整個社會的影響。他的研究領域包括健康傳播、社交媒體、心理健康、信息處理以及技術對個人和社區轉型的影響。

MG：Matthieu Guitton

BZ：鍾布

YS：宋韻雅

YS： **和許多新興科技一樣，人工智能（artificial intelligence, AI）面臨著一個艱難的挑戰：如何説服人們其利大於弊。有些人將AI描繪成「英勇」的角色，而有些人則將AI描繪成「邪惡」的角色。您如何評價當前公眾對人工智能的認知？是充滿希望、恐懼還是持中立態度？公共話語中關於AI的主要敘述是甚麼？**

MG：上述提問的範疇超越了 AI 本身，涵蓋了更廣泛的技術領域。為了實現社會和技術的共同進步，研究人員和社會必須保持一致。鑑於（人與 AI）的對抗性描繪已經存在了數十年（例如1980年代電影《終結者》〔*The Terminator*〕中人工智能對人類的征服），分析AI在大眾文化中的呈現顯得至關重要。關於技術傳播的學術研究必須融入大眾文化和文化研究的視角，以便更有效地管控其在社會上的可接受性（social acceptability）。如果沒有理解人們如何在大眾文化中感知和描繪人工智能，改變他們的態度將會很困難。要成為這個領域有所建樹的學者，你就必須對技術及其在大眾中的感知有跨學科的理解，而不只是關注某一個視角，比如計算機科學、傳播學或其他社會科學的視角。

BZ：將 AI 比喻為「人造人」（Artificial Humans）是有其意義的，因為我們與它的互動方式，無論在智能還是反應上，都與真人相似。我們對 AI 的理解就像珠寶，它為我們提供了對人類心智本質的寶貴洞見。然而，重要的是要將 AI 視為夥伴，而不僅僅是工具，因為它有潛力改變我們的思維、計算和行為方式。AI 在人類文明中扮演了關鍵角色，並且往往不僅僅是一個工具。在許多情況下，AI 是一個夥伴，就像我們的手機或電腦，我們在日常生活中極度依賴它，如果失去它，我們可能會產生一種斷聯感，例如在飛機上的時候。這種不確定感並不新鮮，因為它與手機剛出現時，人們（特別是老年人）所感到的不確定感相似。

認識到媒體在突顯人工智能各種特性（包括奇特性和閃光點）中的作用是重要的，但人們必須就如何更好地在他們的生活中使用 AI 作出知情的決策。許多人沒有意識到的是，AI 已經在日常生活中廣泛運用，以一種我們可能甚至未察覺的方式改變我們的感知。我們不再用肉眼看世界，因為我們的體驗不斷被使用的 AI

所塑造。社會科學家可以在幫助公眾更好地理解 AI 及其機制方面發揮重要的作用。

YS：人們在多大程度上接受 AI 成為他們日常生活的一部分？依二位所見，推動或阻礙人們採用 AI 的最重要的考量是甚麼？

MG：公眾對人工智能的接受程度主要受兩個因素驅動。第一個因素是我們如何看待 AI，亦即，它是被視為一種工具還是一個或將成為有感知生命體的夥伴。其中，感知（sentiency）的概念引發了恐懼，因為如果 AI 具有感知，它就可能擁有超越人類的決策能力，甚至可能推翻我們。然而，如果 AI 並不具有感知，我們就很難與它建立平等的夥伴關係。這種困惑在人群中產生了大量的不信任感，因為人們還沒有準備好以超越工具思維的心態接受 AI，儘管我們知道它的智能正在演進。第二個因素是人們對 AI 所涉事務的認知，具體來說就是「AI 服務我們」與「AI 監控我們」的二分法。在 COVID-19 大流行期間，人們不喜歡技術追蹤他們的互動或疫苗接種狀態，由此引發了關於是否應該使用技術來監視我們的辯論。如果 AI 為我們服務，如洗衣或做家庭作業，人們會樂見其成。然而，如果 AI 開始監察我們，告訴我們不能做甚麼，並像「老大哥」一樣監視我們，人們就不太可能接受它。

BZ：調查顯示，雖然多數人表示喜愛 AI，但這並不一定意味著他們信任 AI。從人們對於 AI 監察的不適，以及讓 AI 執行諸如煤礦開採之類危險任務的傾向中，不難看出信任仍然是與 AI 相關的一個重大問題。調查表明，在 2023 年，當 AI 成為現實時，只有 9% 的美國人認為 AI 利大於弊；而在 1987 年，當 AI 還只是科幻小說（的想像）時，這一比例為 25%（Monmouth University Polling Institute, 2023）。當人工智能真正出現時，人們對它的信任度反而低於其僅存在於人們想像中的階段。因此，不能假設人們對 AI 的喜愛就等同於信任。人們希望 AI 為他們工作，但許多人目前對於與 AI 打交道感到不適。人們對人工智能了解越少，就越不願意讓它進入他們的生活。這與人們最初對於透過設備全天候保持聯繫的反感類似，但現在我們的口袋裏長期放著手機，因為我們已經對如何使用它有了更多的掌控力。同樣的，我們可以學會更好地理

《傳播與社會學刊》，（總）第 68 期（2024）

解和使用 AI。身為社會科學家，我們有責任幫助人們理解 AI 並對其善加利用。

YS： 一項有趣的研究發現，英國公眾對於消極的 AI 敘事如「非人性化」（inhumanity）和「異化」（alienation）展現出更多的焦慮和擔憂；而對於「便利」（ease）和「滿足」（gratification）等樂觀的敘事則興趣寥寥（Cave et al., 2018）。這個發現是否與你們的觀點或觀察相吻合？你們對目前關於公眾對 AI 反應的研究有何看法？

MG： 區分「渴望人工智能成為現實」和「篤信人工智能將帶來益處」這兩類想法很重要。人是社會性動物，會被任何能改善我們社會生活的事物所吸引。然而，AI 所做的事情與 iPhone、互聯網等技術所帶來的豐富體驗有所不同。我們欣賞後者，是因為它滿足了人類對社會互動的基本需求。雖然起初人們可能表示他們並不希望全天候地聯繫著，但實際上，我們也不願孤獨。透過科技，我們可以隨時保持連線，也可以隨時與他人交流，這正是人們所珍視的。溝通（communication）是人類的基本需求。作為一門科學學科，傳播（communication）至關重要，因為它是人類互動的基礎。在某些社會中，如果我們無法與人溝通，我們則會尋求與動物互動，又或者我們會透過科技與 AI 進行交流。我們渴望互動，AI 則為有意義的互動開創了新的可能性。理解這一點至關重要，因為推動 AI 研究的一個關鍵驅動力就是（人類）與新夥伴互動的渴望。

BZ： 身為社會科學家，我們應該考慮如何利用 AI 研究來改善人類的注意力（attention）。儘管許多關於 AI 的研究論文都聚焦在對社會負責的 AI（socially responsible AI）、AI 信任和 AI 倫理，但我們也應該從 ChatGPT 的創新方法中汲取靈感，例如，基於 Vaswani 等人在 2017 年發表的 "Attention is All You Need" 一文提出的「變換器」（transformer）便是一個例子。這進一步引發了一個問題：我們是否可以使用 AI 來改善人類注意力。注意力是我們社會中一個大問題，因為許多人在對任務的專注上都會遇到困難。例如，當人們上課學習新知識時，他們的思緒往往會飄忽不定；相比之下，當他們玩電子遊戲時，他們卻能全神貫注。許多科學突破都離不開極高的專注力，我們需要關注於如何提高人類的這種能力。因

此，社會科學研究在理解人類的專注力以及如何利用AI來強化注意力上能扮演關鍵的角色。雖然有許多關於對社會負責的AI、AI信任和通用AI方法的研究，但我們可以通過聚焦「人類注意力」來作出獨特貢獻。我們不應只是隨波逐流，而應努力在這個領域創造變革、有所作為。香港為我們提供了一個國際平台，讓我們有機會廣泛接觸（各類人與各類研究），並作出有意義的貢獻。

YS： **從2021年到2023年，「元宇宙」——這樣一個被設想的虛擬實境空間，一直佔據著新聞頭條。ChatGPT作為一個AI驅動的聊天機器人隨後出現，並將生成式AI帶入了公眾視野。AI將如何塑造元宇宙？AI和元宇宙將如何協同工作？**

MG：元宇宙並非新概念。自從1990年代首次引入3D沉浸式體驗以來，它已經存在了二十多年。雖然Facebook（現名為 Meta）公司讓這個詞變得流行，但元宇宙在Facebook介入之前就已經存在，並且無論Facebook是否參與，它都將繼續存在。理解元宇宙與人工智能的關係是很重要的，因為元宇宙所在的賽博空間，將會是人類與AI首次互動的地方。第一代人工智能是無實體的，這意味著它們並非社交機器人。機器人是受限制的代理者（limited agent），其軟體應用與ChatGPT等相似，但受制於其實體外形。在元宇宙中，我們沒有真實的身體，也無法見到對方的真實身體。這使得人工智能能夠輕易地模仿人類，因為它們不需要身體，只需要對話，而這是AI非常擅長的部分。人工智能是否將塑造元宇宙並不明確，但很明顯，元宇宙將塑造我們對人工智能的認知以及我們與人工智能的關係，因為這將是首次互動發生的地點。

BZ：我們的關注點並不在於特定社交媒體形式的神奇之處，而在於它們的運作方式。2003年首次亮相的「第二人生」（*Second Life*）就是一個經典實例。作為早期使用者之一，我曾建議我所在的大學購買一個賬號。然而，幾年後，人們的興趣逐漸消退，儘管我們已在服裝道具上投資了許多，並為我們的工作人員建造了一個虛擬島嶼，但我們還是沒再使用它。很多「概念」只是風靡一時，而後很快銷聲匿跡，因而我們無法預測明年是否還會討論元宇宙或

ChatGPT。然而，這些技術的運作機制不僅關乎技術本身，還與由 AI 所驅動的溝通和互動有關。我們主要的興趣在於人類如何受到技術的影響，不論其載體是人、3D 動畫，還是第二人生。因此，我們的目標是探討這項新技術如何改變溝通方式，以及如何使其助益社會。我時常向 ChatGPT 詢問有關 AI 的下一個突破點。儘管 ChatGPT 謙虛地承認它無法預測未來，但它已認識到深度學習是一個明確的有望推動 AI 進步的發展領域。理解深度學習及其潛在應用對於推動 AI 並開發其助益社會的潛力至關重要。此外，社會關注度也是 AI 發展中需要考慮的重要因素，因為它可以塑造我們對新技術的看法和接受度。同樣，媒體具有引導我們去關注多種事物的強大能力，因此，我們也不應低估媒體對於 AI 的影響。

YS： 看到生成式 AI 的快速發展以及它塑造元宇宙的潛力，實在令人振奮。二位對於這兩大當今趨勢的融合有何展望？

MG：我的研究途徑一直是將互動置於特定空間中進行思考。隨著元宇宙的興起，我們開始討論虛擬實境（virtual reality, VR），這已經被擴增實境（augmented reality, AR）所取代；未來，我們可能會轉向延展實境（extended reality, XR）。生成式代理（generative agents）和元宇宙的融合是這些技術的交匯點，因為它們都透過手機、電腦等設備在現實上新增一個疊加層，創造一個新的互動空間。這個包括虛擬、擴增和延展實境的新維度將創造一個我們作為人類能夠互動和居住的空間。當前關於生成式人工智能的討論往往聚焦於 ChatGPT，但 ChatGPT 只是這項技術的一個例子。我們不應該把討論局限於一個工具；相反，我們應該著眼於如何應對生成式人工智能這個整體。此外，我們不應該對人工智能和生成式代理的出現感到驚訝，因為我們至少在十年前就知道它們的來臨。重要的是，我們必須進行跨學科（interdisciplinary）、超學科（transdisciplinary）和多學科（multidisciplinary）的對話，以理解這些技術對社會的影響，以及我們如何善用其潛力來造福大眾。人工智能不僅僅是計算機科學家的領域，它與包括傳播學、哲

學、心理學和其他社會科學等在內的每一個學術領域都有關。每個學生都應該修讀一門心理學專業的基礎課程，以理解AI的起源。最終，我們應該創造一個更美好的社會。包括ChatGPT在內的技術或將能透過促進社會的可持續發展、和平以及人與技術的和諧相處，來幫助我們實現這一目標。

BZ： 我對Guitton教授的洞見表示讚賞。風險投資者往往不願意等待超過一年的投資回報，這反映了我們人類的不耐煩。然而，我們應該利用人工智能與元宇宙所帶來的生活方式的變化。「閾限」（liminal）一詞描述了我們越來越傾向於同時憩居在現在和未來的世界裏，生活在逐漸融合的線上和線下世界中。許多人就生活在他們的動畫世界中，對年輕一代而言，區分現實與幻想越來越具有挑戰性，這是以往只存在於科幻小説中的現象。為了幫助我們理解周圍正在發生的變化，我們需要來自包括社會科學家在內的科學家們的見解。儘管許多人都在使用像今日頭條（一個採用了AI技術的中國新聞資訊平台）這樣的平台，但大部分人仍然不理解AI的日常用途，更不用説去認識此類平台背後的信息流以及AI如何控制信息傳播了。長期身處這樣的閾限世界可能會令我們的生活更加複雜，並引發混亂。雖然年輕人或許不會面臨使用這些技術的困難，但我們不應忽視那些確實面臨困難的人，例如在使用健康碼或移動支付方面也覺得困難的老年人。他們可能需要協助來使用這些設備。

YS： 在二月份，《自然》（*Nature*）雜誌刊登了一篇名為 "What ChatGPT and Generative AI Mean for Science"（Stokel-Walker & van Noorden, 2023）的新聞專題。當科研人員被問及ChatGPT在科學中的可能用途時，他們喜憂參半。例如，研究人員強調，從根本上講，大語言模型（large language models, LLMs）在回答問題時常常是不可靠的，有時甚至會生成錯誤回答。該文章還討論了人工智能的安全和責任問題，例如生成有害內容的風險或AI聊天機器人因其訓練數據而可能產生的偏見。如果我們對這個問題稍作調整，ChatGPT和生成式AI對社會科學來説意味著甚麼？

MG：首先，我想強調的是，人類也有可能產生錯誤答案。當審查學生作業時，人們可以觀察到學生並非總是正確的。會否生成錯誤答案並不是一個評判人工智能的合適標準。標準應該是人工智能是否具有識別錯誤答案的能力。目前，人工智能仍缺乏元認知能力，但未來或不再如此。我們不僅具有生成文本等認知能力，也具有使我們能夠反思我們的認知、探索行動的後果，並感知夥伴的心理狀態的元認知能力。如果人工智能在未來獲得這些技能，它們將更接近（真正的）智能。屆時，我們應該將它們排除在科學辯論之外嗎？科學辯論，顧名思義，既「爭辯」也「論證」。科學家們並不會等到找到絕對真理時才發表成果。相反，他們會提出一個其他研究者會仔細審查的元素，從而貢獻於知識建構。我們並不創建知識，而是創建促進知識建構的基礎部件。如果人工智能在未來獲得這些技能，他們可以幫助我們共同創造這些基礎部件。

BZ：新技術的到來常常會帶來不確定性。然而，ChatGPT 能產生錯誤答案並不表示它缺乏智慧。例如，我們愛我們的父母，但我們並不總是聽從他們的建議。歷史教訓能教導我們如何應對科技進步。當蒸汽機被引入工廠的時候，雖減輕了繁重的工作負擔，但也導致許多工人失業。起初，我們不確定該如何應對蒸汽機這項技術。如今，我們已不再為汽車的速度超越我們而擔憂了。要知道，在亨利‧福特（Henry Ford）發明汽車之前，馬是我們唯一的交通工具。同樣，在史蒂夫‧賈伯斯（Steve Jobs）推出智慧手機之前，我們對手機的可能性也是一無所知。人類是具有智慧的，我們可以從歷史中學習，以預期未來的結果。當 Facebook 問世時，許多成年人認為它只是一時的風潮，並認為社交媒體不會長久存在。然而，如今的社交媒體已經演變成不同的形式，年長者使用 Facebook，而年輕人則使用 Instagram、Discord、Snapchat 和其他平台。我們不應該為遇到來自 ChatGPT 的錯誤訊息而過於擔心，或是選擇盲目信任技術。即使作為一名教師，我也建議我的學生不要把我的話視為唯一的「真理」。我們必須培養自身的批判性思維能力，並付出心力來評估可得資訊的優勢與不足。

YS： 隨著AI在人類生活各方面的普及，個體越來越多地被技術設備和具有主動性的機器所飽和。在充滿AI的世界中，我們需要哪些能力？從學術角度來看，我們如何概念化這些能力？我們又該如何能夠實證地去研究和評估這些發展？

MG：我將從哲學和實踐兩個面向來回答這個問題。我們倡導對社會負責的人工智能，以更好地為大眾利益服務；但是，我們討論的對象不只是汽車而是可能驅動其他設備的技術。因此，我們也需要對社會負責的人類來與人工智能合作。將人工智能只視為一種工具是不夠的，我們必須意識到這種構造可能會獲得智慧或其他形態，這需要對社會負責任的人類的參與。我們不能只依賴「對社會負責的人工智能」，在人類自身的發展、行動和與他人包括AI的互動中，也需要展現社會責任。這是我對這個問題的哲學回答。作為大學教授，我在實踐層面的回答是 —— 教育 —— 即我們必須教育民眾、學生和未來的世代。我們應該傳授知識，並促進公共辯論。教育是解決許多問題的答案。在你們的城市，我們做的正是教育這件事；關鍵是得持續做下去。當我還是學生的時候，曾倡議我的大學在生物學課程中加入生物倫理，因為我堅信這是很重要的。我們必須在每一門課程中講授相關的人文學科內容。從事計算機科學的人應該理解基本的(技術)哲學：不只是知曉如何使用技術，還應了解為何使用它。我們必須通過教育使得學生和研究人員認識到他們亦是社會的一員，並協助教育社會，這正是大學的目的。人們通過繳稅來支付我們的工資，使我們能夠教育大眾。隨著科技素養的提高，人們將能更好地理解與智能和對社會負責的AI相關的挑戰。這將使得人們在與科技合作時更加重視社會責任。

BZ： 我欣賞Guitton教授的哲學回答，並想深入探討微觀層面。為了回答這個問題，我依託歷史教訓對新技術進行思考。例如，在美國，企業花了將近十五年的時間才採用「800」號碼。800號碼使客戶可以免費撥打企業的電話，由企業支付通話費用。最初，企業對採用這種商業模式持猶豫不決的態度，這使其採用推遲了將

《傳播與社會學刊》,(總)第 68 期(2024)

近二十年。然而,800 號碼最終變得司空見慣,企業也適應了這項新技術。黃頁和網站是另外兩個在廣泛採用之前遇到阻礙的技術例子。這些例子突顯了學習人工智能的好處,以及人類如何能夠迅速對人工智能的崛起作出反應以從中受益。作為科學家,特別是社會科學家,我們可以向新聞記者學習如何解釋技術對人類的意義及其潛在好處。這將帶來有價值的研究貢獻。從宏觀層面來看,我提出另一個名為「AI 智識」(AI intelligence)的可能。在數碼時代,每個人都需要一定程度的 AI 智識來理解人類世界並了解 AI 在世界中的角色。大型技術公司已從 AI 智識中受益匪淺,其他企業則必須更成熟地將 AI 技術融入其商業模式。通過這種微觀層面的方法,我們可以更好地了解 AI 如何在未來一兩年中協助人們。與蒸汽機最初在工廠遭遇工人試圖破壞的阻力不同,我們可以看到 AI 以多種方式幫助人類。

YS: **您認為人工智能領域的發展需要社會科學的參與嗎?似乎人工智能主要是由硬科學驅動的。社會科學可以扮演甚麼樣的角色呢?**

MG:社會科學在參與有關生成式人工智能的辯論上非常有趣,特別是在創造力方面的討論。有關人工智能是否具有創造力的問題源於我們對創造力的定義,自從人工智能出現以來,這個定義變得更加複雜。以前,我們並不質疑是否只有人類能夠具有創造力。現在,需要更多社會科學的是人類,而不是人工智能,因為 AI 的崛起挑戰了我們對於人類本質的理解。意識到這一點是很重要的。人工智能對社會科學的影響是,我們現在意識到我們有多麼需要社會科學,包括社會科學和人文學科。這些領域之所以被結合起來,是因為它們將人類視作智識生物,而不僅僅是生物。因此,我們不應將它們分開。人工智能之所以「人工」,是因為它由人類創造和製造。這並不意味著人類的優越性或將人工智能的物化;相反,它意味著人類以一種藝術的方式創造物體。ChatGPT 是一個文本生成器,而不是人工智能。如果我們渴望超越文本生成器的能力,就必須將社會科學和人文學科的洞察融入人工智能研究中。只有這些領域能教導我們甚麼是人類,我們必須將這些經驗

整合到程式碼中，才能超越目前的層次。所有這些因素表明，我們需要社會科學的程度比之前所認為的還要多。

BZ： 有多種方法可以理解人工智能和技術，這些方法可能來自任何領域。雖然通常是由信息科學和計算機科學的研究者應對這個領域的問題，但社會科學家和人文學者也可以作出重要的貢獻。對這個問題的答案是，我們必須理解人工智能和技術，就像我們理解人類思維一樣。人類心智是一種不可思議的現象，但目前的理解受到狹窄的學科方法的限制，如生物學或心理學，這些方法忽略了跨學科和跨領域方法所能提供的見解。為了充分探索人類思維的奧秘，我們必須擁抱這些方法。真正的跨學科和跨領域工作需要來自各種學科的合作和努力，以應對我們面臨的挑戰並回答相關問題，只有這樣才能在研究中取得真正的進展。社會科學家應該如記者那樣，為學術界提供背景信息並解釋信息。背景是至關重要的，因為現實通常比我們預期的更加複雜。我們不應假設不同年齡段的人以相同的方式使用技術。全面的成本效益分析和對背景的理解始終是必要的，不應愚蠢地相信技術將帶來僅僅是好的或壞的結果。例如，雖然青少年可能會對數字技術上癮，但小技巧和不同的背景可以幫助我們以不同的方式理解這個問題。一種解決方案可能是在晚上7點安排每天半小時的社交媒體使用時間，這既有助於減壓，又表明數字技術可以用於積極用途。社會科學家不應僅從單一視角審視技術，而應將人工智能視為合作夥伴，而不是一個順從的工具或技術。在這個充滿新現象的世界中，通過與人工智能共存，我們可以成為更好的個體，而社會科學家可以在幫助人們理解這種合作潛力方面發揮至關重要的作用。

YS： **人工智能的日益複雜化增加了以對社會負責的方式使用人工智能技術的需求。全球監管舉措日漸普及，有些舉措試圖在不同層面上對人工智能進行監管。誰是主要利益相關者？您對以社會負責任的方式管理人工智能有何看法？**

MG：負責任的 AI 問題並不是新話題，從 AI 的歷史及其與流行文化的互動中即可見一斑。第一部關於人工智能的著名小說是由生物化

學家以撒・艾西莫夫（Isaac Asimov）（1986：94–95）所著，他發展了三大機器人定律，分別是：

（一）機器人不得傷害人類，或坐視人類受到傷害。

（二）在不違反第一定律的前提下，機器人必須服從人類的命令。

（三）在不違反第一與第二定律的前提下，機器人必須保護自己。

這三條法則理論上涵蓋任何情況，如果遵循，可以防止傷害。然而，我們生活在一個現實世界而非理論世界。儘管我們曾嘗試將這些法則應用於自主機器人，現在卻創建了用於殺害目標的無人機，導致人類的死亡。要建立對社會負責任的人工智能，我們必須首先培養對社會負責任的人類，因為是人類為機器編程。發展負責任的（人工）智能不僅僅是對程序員的挑戰，更是對整個社會的挑戰。

BZ： 對社會負責的人工智能目前是一個重要的議題。在我去年抵達香港之前，我們啟動了一個社會負責任的人工智能研究所，其中合作組織包括我之前任教的學院和其他七個機構。最近，我向 ChatGPT 詢問有關對社會負責的人工智能的主要研究領域，它提出了四個關鍵領域：透明度、擴展性、隱私和數據。雖然我同意它對這些領域的識別，但我認為設計對社會負責的人工智能從根本上是關於理解人類價值觀。儘管人工智能偶爾可能會產生錯誤，但我們不應過於擔心或過分批評，因為我們並非完美無缺。即使是教授，我們有時也可能犯錯，比如錯誤地給作業打分。儘管在過去的 16 年裏這種情況僅發生過一次，但這並不意味著我在未來不會犯錯。我們應該給予人工智能一些靈活性和時間，並像對待成年人一樣對待它們。人工智能的社會責任是一個指南，就像醫生不得故意傷害患者一樣，但也有可能會因人為失誤開錯處方藥物。人工智能的社會責任是一個很棒的領域，如果我們做得好，我們將會有所作為，並且也能從中獲益。

YS： 負責任的人工智能是一個技術問題還是一個商業問題？例如，我們都同意人工智能需要是「公平的」，但應該使用誰的定義呢？這

是每家公司的決定，在深入細節時可能是一個難以作出的決定。您認為在發展對社會負責的人工智能中，人類代理的角色是甚麼樣的？

MG：我想進一步討論鍾教授有關評學生作業的評論。當我仍資歷尚淺的時候，我認為設計考試很簡單，只需要評估學生的知識水平。然而，隨著年齡的增長，決定要評估甚麼以及希望通過這樣做能實現甚麼目標變得越來越具挑戰性。我不僅僅想評估知識，而是希望確保我的學生從課程中獲得了一些對他們的職業和個人生活有益的東西。我仍然對我的學生進行考察，但我問他們一些需要更多的思考而不僅僅是背誦信息的問題。這帶出了在人工智能中應該評估甚麼的問題。儘管知識是一個重要組成部分，但這並不足夠。因此，中心問題不是負責任的人工智能是技術還是商業的問題，而是圍繞生命價值的問題。人類生命是否等同於人工生命？人工生命是否具有任何價值？一旦我們能夠回答這些問題，我們就能夠定義社會責任。從歷史上看，人類曾將其他人類視為商品，促成了奴隸制度，並要隨著我們對哲學理解的演變，認識到了所有人類生命的價值才將其廢除。同樣地，我們必須發展對人工智能的理解，以反映這項技術的價值和影響。因此，人工智能的社會責任不應僅僅關注技術或商業方面，因為我們必須考慮更有價值和影響的問題。

BZ：談到人工智能如何更好地服務社會，關於在討論中應該包括哪些利益相關者的問題是具有挑戰性的。誰應負責教授倫理和提高人工智能素養？政府、人工智能開發者、大型科技公司的首席執行官，還是社會的所有成員？我並沒有明確的答案，但當單一方主導對話時，無論是政府、科技公司還是一個有影響力的人物或意見領袖，我都感到擔憂。這會引起不適感和對討論方向的不確定性。例如，誰將教導下一代倫理，政府、政黨還是自稱天才的人？這令人不安。然而，這種不確定性促使人們不斷討論可以為「科技及其社會影響」進行辯論的利益相關者。在我看來，在這個過程中，我們應該優先考慮人類及其偏好，這將可能得出特定問題的解決方案。然而，整體而言，這仍然是一個難以回答的問題。

YS： **全球範圍內的人工智能監管因文化多樣性而碎片化。與此同時，建立一個全球性治理框架的需求正在增長。在我們嘗試將實踐與原則對齊時，文化差異可能會對人工智能的倫理和治理提出一系列獨特的挑戰。您對這一困境有何看法？**

MG：儘管文化差異是迷人且重要的，亦不應成為不採取行動的藉口。作為人類，當我們真正渴望找到共同的基礎時，我們便能做到。可持續發展目標就是一個很好的例子，包括美國、中國和北韓等眾多國家都已同意進行國際合作。當我們努力追求共同目標，例如可持續發展、人權、環境保護、普及教育和發展時，這些共同目標是可以實現的。然而，我們也要承認文化差異的存在——辯論和具有情境意義的討論對找到共同點十分重要。因此，正如我之前所提到的，進行文化研究是至關重要的。在我的研究所中，我們總會在進行的研究或項目中加入文化維度，因為這非常關鍵。如果我們付出努力，人類的思想就能夠達成任何事情。

BZ：近年來，文化差異已成為比全球化更重要的議題。儘管幾年前廣泛討論全球化，但現在東西方年輕人所關心的問題幾乎沒有區別，如氣候變化，而且他們能夠比許多來自其他年代的人更有效地進行交流。這表明在某些方面文化差異有所減少，而這也是個值得探討的議題。然而，我們應該問自己是否願意與來自不同文化背景的人進行交流。許多人表示在香港的美國人與典型的美國人有所不同，有些人亦認為持有護照的美國人不同於沒有護照的美國人。即使在西方世界，如美國和歐洲國家，也存在差異。我們不應該假設每個人都與我們擁有相同的文化背景。儘管我們都受到中華文化的滋養，並且受其數千年的影響，但我可以看到香港和中國大陸之間存在顯著的差異。我們應該避免假設每個人都有相似的思維方式，反而應該預設彼此之間存在許多差異。然而，我們也應該牢記，文化差異可能被用作一種藉口，妨礙我們認識到人性的真正價值，正如 Guitton 教授所討論的那樣。

YS： **聽說你們將推出一本名為*Computers in Human Behavior: Artificial Humans*（*CHB: AH*）的新愛思唯爾（Elsevier）期刊，這是一個令**

**人振奮的消息。這本新期刊的願景是甚麼？又與它的姊妹刊
Computers in Human Behavior（*CHB*）有何分別？**

MG：推出一本新的期刊是一個令人振奮的發展，作為編輯和科學家，
我個人對此感到興奮。我想就該期刊的三個關鍵點進行說明：

（一）有關「人造人」的研究目前分散在各個領域，如社交機
器人、人工智能和虛擬角色。儘管這些領域的研究人
員對「人造人」感興趣，他們之間卻很少有交流。因
此，我們需要一本新的期刊，將所有對這一領域感興
趣的人匯聚在一起。

（二）大多數人工智能的研究都聚焦在技術方面，而我們需
要考慮社會科學和人文學科的方面，如哲學、心理
學、社會學等。目前沒有一個平台可以發表這些類型
的研究，而我們創立的 *CHB: AH* 便能成為社會科學家
參與學術辯論的平台。

（三）推出這本期刊的第三個原因是要建立一個以社會視野
和使命為基礎的科學出版場所。*CHB* 旨在為與技術相
關的社會辯論作出貢獻，我們希望將「人造人」納入其
範疇。然而 *CHB: AH* 更聚焦，不僅包括網絡行為，還
包括神經科學、人機融合和機器人學。雖然這兩本期
刊是彼此的衛星，但 *CHB: AH* 有些偏向一側，擁有自
己的網絡，與計算機科學、神經科學和腦部研究相
連。因此，這本新期刊不會與 *CHB* 在相同主題發生衝
突，而是補充並擴展了 *CHB* 所處理的問題。

BZ：我很高興分享新期刊 *CHB: AH* 的使命，其準確地反映了我們工作
的本質。我們的焦點是「人造人」，致力於融入人文和社會科學領
域的觀點，而不僅僅是一本人工智能期刊。這種協作的努力旨在
揭示這個黑箱及相關問題。「人工智能」這個術語可能很快就會過
時，就像區塊鏈和非同質化代幣（non-fungible tokens, NFTs）等時
髦詞匯一樣。我認為 Elsevier 決定把這本新期刊放置在 *CHB* 這個
已經成熟的平台下，而不從頭開始創辦一本全新的期刊是值得稱

讚的，這種做法能為新期刊提供更快成長的蔭庇。我相信 *CHB* 是一個獨特的平台，可以成功地助推這本期刊。作為一個新平台和場所，這本期刊在東方世界推出亦別具意義，強調了對人工智能研究和「人造人」的全球視角。我鼓勵大家認真考慮將自己的論文投稿到 *CHB: AH*。

YS： **作為兩位經驗豐富的期刊編輯，您通常在一篇優秀的稿件中尋找哪些特點？**

MG：任何關於「人造人」的辯論或能夠培養思考和豐富靈魂的內容，我們都十分歡迎，因為我們都在討論自己作為人類的議題。因此，我們都鼓勵任何有助於我們對「人造人」理解的內容。這種開放性是 *CHB* 的標誌，而我們期刊都把發表優秀的科學發現視為首要目標。

BZ：我們討論了是否應同時接受定量和定性論文，並一致同意只要論文能提供洞見，我們都接受任何投稿，包括批判性研究。我們需要更多處於閾限世界中的人工智能的視角，這些閾限世界既包含我們現存的世界也涵蓋未來的世界，以促進成長。我相信這個領域將在未來幾年迅速擴展，因為我們共享對人工智能研究的願景和熱情，不僅僅是從自然科學的角度，還有人文和社會科學的角度。明年如何？我們將向大家報告我們的進程。

問答部分

Q1： **我正在教授一門曾教給本科生的社交媒體課程。在開始課程時，我提出了一個問題：「是技術塑造了人類行為，還是人類行為塑造了技術？」雖然兩者皆是正確的，但與每一種相關聯的是不同的方法和理論。電腦中介傳播的演變促使了社交媒體和「人造人」或人工智能的出現，我們需要考慮如何應對這一新發展。從傳播學的角度來看，值得注意的是有些理論不一定是新的，人們常常嘗試用舊的理論來解釋新的現象。關鍵問題是這些理論是否足以解釋人類對這些新現象的反應。作為一名教師，我正在思考是否應該教導我的學生舊的理論，以及舊的理論是否與新現象一致。您對這個問題有甚麼看法？**

BZ： 理論構建是一個持續進行的過程，即使對於像「議程設置理論」（Agenda-Setting Theory）這樣的「既定」理論，它也經歷了多次修訂。因此，這個領域始終有新的理論貢獻的空間。在獲得終身教職後，我諮詢了美國的頂尖學者，尋求他們對如何進行理論構建的建議。他們建議更加注重採用和修改現有的理論框架，使它們更具擴展性和優越性，以提高我們對技術的理解。因此，我強烈鼓勵我的合作者和學生探索來自各個領域的框架，而不僅僅局限於從心理學中借用。在其他領域，如信息科學、醫學和商科等，都有精彩和有趣的研究，這可以增進我們對技術的理解。例如，參加一個有關健康信息技術和分析的小型會議讓我對商業研究和市場營銷有了新的見解。作為學者，我們應該不斷審查和擴展我們現有的框架，確保它們對新的現象和挑戰仍然具有相關性。這就是我們所謂的理論貢獻。

MG： 理論是我們用來幫助理解現象的結構。然而，它們的有效性只能持續到它們被證明錯誤之時，而後我們必須轉向新的理論。在互聯網出現之前開發的理論是否仍然能夠解釋現代技術，特別是在2023年，這是值得質疑的。儘管教授歷史內容很重要，研究人員不應僅僅局限於過時的理論。相反，他們應該制定與當前現象更相關的新框架。審稿人也無需對理論進行大量論述，最重要的是方法論和結果是否令人信服。如果方法論堅實，但結果與理論不一致，這表明理論是不正確的。我們不應該束縛自己於特定的理論，因為任何理論在某個時刻都可能過時。然而，當我們認為當前理論不再有效時，應該測試新的想法。儘管我樂於接受理論貢獻，但每兩週生成一個新理論既不可行也不具有開創性。相反，我們應該努力作出有意義的貢獻，帶來一些新的東西。最終，我們應該在生命結束時思考我們對科學的貢獻，這可能是一個理論框架，但目前還不知道那將是甚麼樣的貢獻。

Q2A：我們對 ChatGPT 的能力感到驚奇，但隨著大型語言模型變得越發先進，出版商的編輯們面臨新的挑戰。儘管 ChatGPT 可以作為研究助手或合作者，但已經有一些文章把它列為潛在的作者。然而，使用 ChatGPT 或類似的 AI 增加了研究造假的風險。作為編輯，你們如何應對這些由 AI 驅動的大語言模型所帶來的挑戰？

MG：這個問題既有趣又頗具挑戰性。雖然我不能斷言給出一個肯定答案，但它反映了我們當前在不斷變化的情況下的思考。信任是研究的基石，偽造數據和將ChatGPT列為作者是完全不同的事情。如果有人利用ChatGPT捏造或操縱數據，那麼該論文應當被撤回，責任人應該被解僱，這樣的行為在學術界將不容寬恕。然而，如果誠實地將ChatGPT作為研究助手或合作者，其貢獻應該被公開承認。若對ChatGPT的使用超越了寫作的範疇，例如用它來提供新的觀點或角度時，那麼將其列入作者名單可能是合適的，這樣的決定應該基於進行這項工作的研究者的誠實評估。如果ChatGPT只是被用作寫作工具，則應該在方法部分予以說明，就像我們承認使用統計軟體一樣。透明性是其中的關鍵，技術的使用始終應被承認。然而，利用人工智能生成假數據是一種偽造行為，而且因為難以檢測，更屬於雙重犯罪。這種行為是絕對不能接受的。

Q2B：你們會考慮接受ChatGPT或其他人工智能 (系統) 作為合法的作者嗎？

MG：就我個人而言，我願意接受ChatGPT或其他AI作為合法的作者，但當此類情況首次出現時，我將需要向Elsevier和出版商諮詢。這種情況與那些由團體作者署名的論文並無太大區別；這在社會科學領域偶爾發生，而在物理學和某些醫學領域更為常見。在某些情況下，研究聯盟或合作作品可能有多達四十位作者。因此，如果這項研究是由宋韻雅、鍾布、Matthieu Guitton和ChatGPT共同進行的，我認為沒有理由不承認ChatGPT是作者。不過，需要注意的是，在提交研究工作時，ChatGPT不能作為通訊作者。

BZ：為了避免混淆或其他潛在問題，有必要確定研究的哪些部分是由ChatGPT完成的。否則，我不願意與這樣的工具相關聯。這些技術的發展前景還未有定數，不過值得注意的是，當第一款拼寫檢查器問世時，人們曾擔心學生因過度依賴技術而喪失拼寫能力。有些教授甚至預言，綜合性考試將失去效用。然而，這並未成為現實，在學習能力未受技術顯著改變的情況下，我們仍繼續使用它們。

Matthieu Guitton教授著作選

Beaunoyer, E., Dupéré, S., & Guitton, M. J. (2020). COVID-19 and digital inequalities: Reciprocal impacts and mitigation strategies. *Computers in Human Behavior, 111,* 106424.

Beaunoyer, E., & Guitton, M. J. (2021). Cyberthanathology: Death and beyond in the digital age. *Computers in Human Behavior, 122,* 106849.

Georgieva, I., Beaunoyer, E., & Guitton, M. J. (2021). Ensuring social acceptability of technological tracking in the COVID-19 context. *Computers in Human Behavior, 116,* 106639.

Guitton, M. J. (2023). Toward Homo Artificialis. *Computers in Human Behavior: Artificial Humans, 1*(1), 100001.

Kaewkitipong, L., Beaunoyer, E., Ractham, P., & Guitton, M. J. (2023). Augmented spirituality: Renewing human spirituality in a technology-driven world? *Computers in Human Behavior, 148,* 107904.

Pelletier, A., Kaewkitipong, L., & Guitton, M. J. (2024). Using technology to study refugee, conflict-affected, and hard-to-reach populations: Methodological and ethical considerations. *Computers in Human Behavior, 152,* 108053.

鍾布教授著作選

Zhong, B. (2022). *Social media communication: Trends and theories*. John Wiley & Sons.

Zhong, B. (2023). Going beyond fact-checking to fight health misinformation: A multi-level analysis of the Twitter response to health news stories. *International Journal of Information Management, 70*, 102626.

Zhong, B., & Chen, J. (2021). Health information helps mitigate adolescent depression: A multivariate analysis of the links between health information use and depression management. *Child: Care, Health and Development, 47*(2), 201–207.

Zhong, B., & Liu, Q. (2021). Medical insights from posts about irritable bowel syndrome by adolescent patients and their parents: Topic modeling and social network analysis. *Journal of Medical Internet Research, 23*(6), e26867.

Zhong, B., Huang, Y., & Liu, Q. (2021). Mental health toll from the coronavirus: Social media usage reveals Wuhan residents' depression and secondary trauma in the COVID-19 outbreak. *Computers in Human Behavior, 114*, 106524.

Zhong, B., Sun, T., Forde, S., & Payne, G. (2022). More than politics: How personality trait and media technology use affect vote likelihood during the 2020 US presidential election. *American Behavioral Scientist, 66*(3), 357–375.

Zhong, B., Sun, T., Zhou, Y., & Xie, L. (2024). Privacy matters: Reexamining internet privacy concern among social media users in a cross-cultural setting. *Atlantic Journal of Communication, 32*(2), 180–197.

參考文獻

Asimov, I. (1986). *Robot dreams*. Berkley Trade. Archived on March 16, 2012, Wayback Machine. Retrieved July 16, 2023, from https://web.archive.org/web/20120316134042/http://www.tcnj.edu/~miranda/classes/topics/reading/asimov.pdf.

Cave, S., Craig, C., Dihal, K., Dillon, S., Montgomery, J., Singler, B., & Taylor, L. (2018). *Portrayals and perceptions of AI and why they matter*. The Royal Society.

Monmouth University Polling Institute (2023, February 15). *Artificial intelligence use prompts concerns: Most expect ChatGPT will be used for cheating*. Retrieved April 3, 2023, from https://www.monmouth.edu/polling-institute/reports/monmouthpoll_US_021523/.

Puslowski, X. S. (Director). (2009). *The terminator* [Film]. Paramount Pictures, Canal+.

Stokel-Walker, C., & van Noorden, R. (2023). What CHATGPT and generative AI mean for science. *Nature*, *614*(7947), 214–216.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems 30 (NIPS 2017)* (pp. 6000–6010). Curran Associates, Inc.

本文引用格式

Matthieu GUITTON、鍾布、宋韻雅 (2024)。〈學界和公眾對人工智能認知的再思考：當今之境，未來之航〉。《傳播與社會學刊》，第 68 期，頁 1–44。

Academic Dialogue with **Matthieu GUITTON** and **Bu ZHONG**

Revisiting Scholarly and Public Perceptions of Artificial Intelligence: Current State and Future Trajectories

MG: Matthieu GUITTON
BZ: Bu ZHONG
YS: Celine Yunya SONG

YS: **Like many emerging technologies, artificial intelligence (AI) faces an uphill battle to convince people that its benefits outweigh its drawbacks. There are popular portrayals of "heroic" AI as well as "villainous" AI. How would you evaluate the public's current perception of AI? Hopeful, scared, or neutral? What are the main narratives of AI in public discourse?**

MG: The scope of this inquiry extends beyond AI to encompass technology more broadly. In order to achieve progress for both society and technology, researchers and society must be in alignment. It is crucial to analyze how AI is portrayed in popular culture, given that antagonistic depictions have existed for decades, such as *The Terminator* with AI that overran humanity in the 1980s. Scholarly work on the communication of technology must take popular culture and cultural studies into account in order to better manage its social acceptability. Without an understanding of how people perceive and depict AI in popular culture, it will be difficult to change these attitudes. To be an effective scholar in this field, one must have an interdisciplinary understanding of both the technology and its popular perception, rather than just focusing on one perspective, such as computer science, communication, or the other perspectives of social science.

BZ: Viewing AI as an artificial human is a useful analogy because we interact with it in a similar way to real humans, in terms of intelligence and response. Our understanding of AI is like jewelry that provides us

with valuable insights into the nature of the human mind. However, it is important to treat AI as a partner, not simply a tool, as it has the potential to alter our thinking, calculations, and behaviors. AI plays a critical role in human civilization and is often more than just a tool. In many cases, AI is a partner, similar to our phone or computer, on which we heavily rely in our daily lives and from which we may feel disconnected, such as on an airplane. This feeling of uncertainty is not new, as it is similar to the uncertainty felt when mobile phones were first introduced, particularly among older people.

It is important to recognize the media's role in highlighting the various features of AI, including its weirdness and brightness, but humans must make informed decisions on how to better use AI in their lives. Many people fail to realize that AI is already extensively functioning in our daily lives, altering our perceptions in ways we may not even realize. We no longer see the world with our bare eyes, as our experiences are constantly shaped by the AI we use. Social scientists can play a crucial role in helping the public better understand AI and its mechanisms.

YS: **To what extent does the public accept AI as part of their daily lives? From your observations, what are the most important considerations that drive people toward or push them away from AI adoption?**

MG: Public acceptance of AI is driven by two main factors. The first factor is how we perceive AI, whether as an object-like tool or a partner that is or will become a sentient being. The concept of sentiency elicits fear because if AI becomes sentient, it may have decision-making power that surpasses ours and potentially overthrow us. However, if AI is not sentient, it is difficult to negotiate a partnership relationship with it on equal terms. This confusion generates a lot of distrust in the population, as we are not yet ready to accept AI beyond the tool mentality, even though we know that intelligence is evolving.

The second factor is the perception of what AI is doing, specifically the large dichotomy between AI serving us and policing us. During the COVID-19 pandemic, people did not like technology to track their interactions or vaccination statuses, which led to a debate about whether technology should be used to monitor us. If AI serves us, such as by doing our laundry or homework, it is viewed positively.

However, if AI starts policing us, telling us what not to do, and watching us like Big Brother, people are less likely to accept it.

BZ: Based on surveys, while most people love AI, this does not necessarily translate to trust. Trust remains a significant issue related to AI, as exemplified by people's discomfort with AI policing them and their preference for AI to perform risky tasks like coal mining. A study demonstrates that when AI became a reality, only 9% of Americans believed that AI would do more good than harm in 2023, compared to 25% in 1987, when AI was still science fiction (Monmouth University Polling Institute, 2023). The reality of AI has made people trust it less than they imagined. Therefore, it is crucial not to assume that people's love for AI equates to trust. People want AI to work for them, but many feel uncomfortable dealing with it at present. The less people know about AI, the less comfortable they feel about letting it into their lives. This is similar to how people initially disliked the idea of being reachable 24/7 through a device, but now that we have a phone in our pocket all the time and have more control over how we use it. Similarly, we can learn to better understand and utilize AI. As social scientists, it is our responsibility to help people understand AI and make good use of it.

YS: An interesting study found that the public in the United Kingdom showed more anxiety and concerns towards pessimistic (AI) narratives such as inhumanity and alienation but less excitement for optimistic narratives such as ease and gratification (Cave et al., 2018). Does this finding resonate with your views or observations? What do you think about the current research on public response to AI?

MG: It is important to distinguish between the desire for AI to come to life and the trust that it will do good things. Humans are social animals and are drawn to anything that improves our social life. However, what AI does is different from the enrichment provided by technology such as iPhones and the internet. We appreciate the latter because it fulfills a basic human need for social interaction. Initially, people may have said that they did not want to be reachable 24/7, but in reality, we do not want to be alone. With technology, we are constantly connected and have the ability to communicate with others at any time, which is something people appreciate.

Communication is a fundamental human need. As a scientific discipline, communication is essential because it is the basis of human interaction. If we cannot communicate with humans, we seek interaction with animals in some societies, or we communicate with AI through technology. We crave interaction, and AI opens up new possibilities for meaningful interactions. It is crucial to understand this, as one of the key drivers behind AI research is the desire for new partners to interact with.

BZ: As social scientists, we should consider how AI research can be used to improve human attention. While many research papers on AI focus on socially responsible AI, trust, and ethics, we should also draw inspiration from ChatGPT's novel methods, such as "the transformer", which is based on the article "Attention is All You Need" (Vaswani et al., 2017). This raises the question of whether we can use AI to improve human attention. Attention is a significant problem in our society, as many people struggle to focus on tasks. For example, when people attend classes and learn something new, their minds tend to wander. In contrast, when they play video games, they are highly concentrated. Many scientific breakthroughs require intense attention. We need to focus on how we can improve this ability in humans.

Therefore, social science research can play a crucial role in understanding human attention and how AI can be used to enhance it. While there is a lot of research on socially responsible AI, trust, and generic approaches to AI, we can make a unique contribution by focusing on human attention. We should not just follow the crowd but strive to make a difference in this area. Hong Kong provides an international platform for us to reach out and make a meaningful contribution.

YS: **From 2021 to 2023, the metaverse as an envisioned virtual reality space has been dominating the headlines. ChatGPT as an AI-powered chatbot emerged afterwards and brought generative AI into the limelight. How will AI shape the metaverse? How can AI and the metaverse work together synergistically?**

MG: The metaverse is not a new concept, as it has been around for more than 20 years since the first 3D immersive was introduced in the 1990s. While Facebook (now known as "Meta") popularized the term, the metaverse existed before and will continue to exist regardless of

Facebook's involvement. It is crucial to understand the relationship between the metaverse and AI because the cyberspace, of which the metaverse is a part, will be the first space where humans and AI will interact.

The first AI is disembodied, meaning that it is not a social robot. Robots are limited agents, and their software applications are similar to ChatGPT, but they are constrained by their physical bodies. In the metaverse, we do not have a real body, and we do not see the real body of our partner. This makes it easy for AI to impersonate a human because they do not need a body; all they need is dialogue, which AI can do very well. It is unclear whether AI will shape the metaverse, but it is evident that the metaverse will shape our perception of AI and our relationship with AI, as it is where the first interaction will take place.

BZ: Our focus lies not on the wonders of specific forms of social media but rather on their mechanics. *Second Life*, which debuted in 2003, is a well-known example. As one of its early adopters, I recommended that my university purchase an account. However, after a few years, interest dwindled, and we ceased to use it despite having invested in clothing and building a virtual island for our laborers. It is worth noting that concepts can come and go, and we cannot predict whether we will still discuss the metaverse or ChatGPT next year. Nonetheless, the mechanics of these technologies are not solely about the technology itself, but the communication and interactions that occur with them, powered by AI. Our primary interest lies in how people are influenced by such technology, be it a human, a 3D-like animation, or *Second Life*. Therefore, we aim to explore how this new technology can alter communication and how to leverage it for social benefits. Frequently, I ask ChatGPT about the next breakthrough in AI. Although ChatGPT is modest and acknowledges its inability to predict the future, it recognizes deep learning as a clear area of development expected to drive AI progress. Understanding deep learning and its potential applications is crucial for advancing AI and exploiting its potential for societal benefit. Additionally, social attention is an essential factor to consider in AI development, given how it can shape our perception and adoption of new technologies. Similarly, the media has a powerful ability to direct our attention to various things, and hence, we must not underestimate its power in influencing AI.

YS: **It is exciting to see the rapid advancement of generative AI and its potential to shape the metaverse. How do you envision the convergence of the two major trends happening today?**

MG: My approach to studying interactions has always been to view them as taking place in a specific space. With the rise of metaverse technology, we began discussing virtual reality (VR), which has since been replaced by augmented reality (AR), and in the future, we will likely shift to extended reality (XR). The convergence of generative agents and the metaverse is where these technologies will meet because they both add a superimposed layer on reality through devices such as phones and computers, creating a new space for interaction. This new dimension, which includes VR, AR, and XR, will create a space where we can interact and inhabit as human beings. The current debate on generative intelligence often centers on ChatGPT, which is just one example of this technology. We should not reduce the debate to one tool; instead, we should focus on how to deal with generative intelligence as a whole. Moreover, we should not be surprised by the emergence of AI and generative agents since we have known about their arrival for at least a decade. It is crucial that we engage in an interdisciplinary, transdisciplinary, and multidisciplinary dialogue to understand the implications of these technologies for society and how we can harness their potential for the greater good. AI is not just the domain of computer scientists; it is relevant to every academic field, including communication, philosophy, psychology, and other social sciences. Every student should have a basic course in psychology to appreciate the origins of AI. Ultimately, we should create a better society. Technology, including ChatGPT, can help us achieve this goal by promoting sustainability, peace, and harmony between ourselves and technology.

BZ: I appreciate the insightful points raised by Prof. Guitton. Venture capitalists are often unwilling to wait for more than a year for a return on their investment, highlighting our impatience as humans. However, we should leverage changes in our lifestyle. The term "liminal" describes our growing tendency to inhabit both present and future worlds simultaneously, living in both the online and offline worlds that are converging. Many individuals reside in their animated world. It is becoming increasingly challenging for younger generations to

differentiate between reality and fantasy, a phenomenon previously confined to science fiction. To help us understand the changes occurring around us, we require insights from scientists, including social scientists. Despite accessing platforms such as *Toutiao* (今日頭條; a news and information platform equipped with AI in China), the majority of people do not comprehend the daily use of AI, let alone the information flow behind these platforms and how to control information dissemination. Persisting in this liminal world could further complicate our lives and lead to confusion. While young people may not face challenges using these technologies, we should not overlook those who do, such as older individuals who may find it difficult to use their health code or mobile payment. They may require assistance in using these devices.

YS: *Nature* **published a news feature entitled "What ChatGPT and Generative AI Mean for Science" in February. When researchers were asked about the potential use of ChatGPT in science, their excitement was tempered with apprehension (Stokel-Walker & van Noorden, 2023). For example, researchers emphasize that large language models (LLMs) are often unreliable in answering questions and sometimes generate false responses. The article also addresses issues of safety and responsibility, such as the production of toxic content. AI chatbots can be biased, depending on their training data. If we tweak this question a little bit, what do ChatGPT and generative AI mean for social science?**

MG: Firstly, I want to emphasize that humans are also capable of generating incorrect answers. When reviewing student assignments, one can observe that they are not always correct. Generating incorrect answers is not the appropriate criterion. The criterion is the ability to recognize that an answer is incorrect. Currently, AI lacks metacognition, but this may not be the case in the future. While we possess cognitive skills such as generating text, we also have metacognitive skills that enable us to reflect on our cognition, explore the consequences of our actions, and perceive the state of mind of our partner. These are all metacognitive skills that AI still lacks. If they acquire these skills in the future, they will be closer to intelligence. Should we exclude them from scientific debates? Scientific debates are precisely that—debates. Scientists do not wait to publish until they find the absolute truth.

Instead, they present an element that other researchers will scrutinize, contributing to the building of knowledge. We do not construct knowledge, but rather, we create the building blocks that facilitate the construction of knowledge. If AI acquires these skills in the future, they can aid us in creating these building blocks collaboratively.

BZ: Encountering new technology often generates uncertainty. However, ChatGPT's ability to generate incorrect answers does not imply a lack of intelligence. For instance, we love our parents, but we do not always heed their advice. Historical lessons can teach us how to respond to technological advancements, such as when the steam engine was introduced into factories, relieving heavy duties but displacing many workers. Initially, we were unsure how to respond to this technology. Today, we no longer fret about cars outpacing us. Before Henry Ford invented the car, the horse was our sole means of transportation. Similarly, before Steve Jobs introduced smartphones, we had no idea what a phone could be. Humans are intelligent, and we can learn from history to anticipate outcomes. When Facebook was released, many adults regarded it as a passing fad, assuming that social media would not endure. However, social media has evolved into different formats, with older people using Facebook and younger people utilizing Instagram, Discord, Snapchat, and other platforms. We should not fret about encountering error messages from ChatGPT or blindly trusting technology. Even as a teacher, I advise my students not to regard my words as the only truth. We must develop critical thinking skills and exert mental effort to evaluate the strengths and limitations of the information available.

YS: With the proliferation of AI in many aspects of human life, individuals are increasingly saturated with technological devices and agentic machines. What kinds of capabilities do people need in a world infused with AI? From a scholarly perspective, how can we conceptualize these capabilities? How can we empirically study and assess these developments?

MG: I will address this question both philosophically and practically. We advocate for socially responsible AI that works for the greater good. However, we are not solely discussing cars; we are referring to technology that may power other devices. Therefore, we also require socially responsible humans to collaborate with AI. Treating AI as a

mere tool will not suffice. We must recognize that this construct may acquire intellect or a different form, necessitating socially responsible human participation. We cannot solely rely on socially responsible AI. We must also be socially responsible in our development, actions, and interactions with others, including AI. That is my philosophical response to this question. My practical response, as a university professor, is education. We must educate the population, students, and future generations. We should impart knowledge and promote public debate. Education is the answer to many issues. In your city, we are doing precisely that, and it is crucial to continue. When I was a student, I advocated for my university to incorporate bioethics into biological courses because I believe it is vital. We must teach the humanities in every course. Those working in computer science should understand basic philosophy—not only how to use technology but why to use it. We must educate students and researchers as members of society and assist in educating society. That is the purpose of the university. People pay taxes to fund our salaries, enabling us to educate the population. As technological literacy grows, individuals will be better equipped to understand the challenges associated with intelligence and socially responsible AI. This will lead to greater social responsibility when working with technology.

BZ: I appreciate Prof. Guitton's philosophical response and would like to delve deeper into the micro-level. To answer this question, I rely on historical lessons and consider new technology. For instance, in the United States, it took almost 15 years for businesses to adopt 800 numbers. An 800 number enables customers to call a business free of charge, with the business paying for the call. Initially, businesses were hesitant to adopt this business model, delaying its adoption for almost 20 years. However, 800 numbers eventually became commonplace, and businesses adapted to this new technology. The Yellow Pages and websites are other examples of technology that encountered initial resistance before becoming widely adopted. These examples highlight the benefits of learning about AI and how humans can react rapidly to the rise of AI to reap its benefits. As scientists, particularly social scientists, we can learn from journalists how to interpret the meaning of technology for humans and its potential benefits. This leads to valuable research contributions. From a macro-level perspective, I

propose another channel called "AI intelligence". In the digital age, everyone requires some level of AI intelligence to comprehend the world as humans and to understand AI's role in the world. Large technology companies have greatly benefited from AI intelligence, and other businesses must become more mature in integrating AI technology into their business models. With this micro-level approach, we can better understand how AI can assist people in the next year or two. Unlike the initial resistance encountered by steam engines, wherein workers attempted to damage them in factories, we can recognize that AI helps people in numerous ways.

YS: **Do you think the field of AI needs social science to evolve? It seems that AI has largely been driven by hard science. What role can social science play?**

MG: The involvement of social science in debates about generative AI is intriguing, particularly regarding creativity. The question of whether AI is capable of creativity stems from our definition of creativity, which has become more complex since the advent of this type of intelligence. Previously, we did not question whether only humans were capable of creativity. Now, it is humans, not AI, who require more social science because the rise of AI challenges our understanding of what it means to be human. This is a crucial point to grasp. The impact of AI on social science is that we now realize how much we need social science, including both the social sciences and the humanities. These fields are combined because they deal with humans as intellectual beings, not just biological ones. Therefore, we should not separate them. AI is "artificial" because humans craft and create it. This does not imply human superiority or AI objectification; rather, it means that humans craft objects in an artistic manner. ChatGPT is a text generator, not an AI. If we aspire to surpass the text generator's capabilities, we must incorporate lessons learned from social science and humanities into AI research. Only these fields can teach us what it means to be human, and we must integrate these lessons into our programming to advance beyond our current level. All of these factors demonstrate that we require social science more than we previously believed.

BZ: Numerous approaches exist for understanding AI and technology, and they may originate from any field. Although information science

Communication and Society, *68* (2024)

and computer science typically handle this area, social scientists and humanities scholars can also make significant contributions. The answer to this question is that we must comprehend AI and technology, much like we understand the human mind. The human mind is an incredible phenomenon, but our current understanding is restricted by narrow disciplinary approaches such as biology or psychology, which disregard the insights that interdisciplinary and transdisciplinary approaches can provide. To fully explore the mysteries of the human mind, we must embrace these approaches. True interdisciplinary and transdisciplinary work requires collaboration and effort from various disciplines to address the challenges we face and answer questions. Only then can we make genuine progress in the research. Social scientists should follow the example of journalists, who provide context and interpret information for the academic community. Context is critical since reality is typically more complex than we anticipate. We should not assume that individuals use technology in the same way across age groups. A comprehensive cost-benefit analysis and understanding of the context are always necessary, and we should not foolishly believe that technology will solely bring either good or bad consequences. For example, while it is true that teenagers may become addicted to digital technologies, small tricks and different contexts can help us understand this issue in a different way. One solution could be scheduling a daily half-hour of social media use time at 7 p.m., allowing for stress reduction and the demonstration that digital technology can be used for good. Social scientists cannot only present a singular perspective on technology, but instead should consider AI as a partner rather than a submissive tool or technology. By coexisting with AI in this world full of new phenomena, we can become better individuals, and social scientists can play a crucial role in helping people understand the potential of this partnership.

YS: **The growing sophistication of AI has increased the demand for socially responsible use of AI technologies. Global regulatory initiatives are becoming increasingly common. There are attempts to regulate AI at various levels of governance. Who are the key stakeholders? What is your perspective on a socially responsible approach to governing AI?**

MG: The question of responsible AI is not new, as seen in the history of AI and its interaction with popular culture. The first well-known novel

about AI was written by biochemist Isaac Asimov (1986, pp. 94–95), who developed the three laws of robotics as follows:

1) A robot may not injure a human being or, through inaction, allow a human being to come to harm;

2) A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law;

3) A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

These three laws theoretically cover any situation, preventing harm if followed. However, we do not live in a theoretical world, but in a real one. While the implementation of these laws into autonomous robots was attempted, we now create drones designed to kill targets, leading to the deaths of people. To build socially responsible AI, we must first create socially responsible humans, since humans program machines. The challenge of developing responsible intelligence is not solely for programmers but for society as a whole.

BZ: Socially responsible AI is currently a prominent topic. Before my arrival in Hong Kong last year, we launched a socially responsible AI institute, which included my former college and seven other institutions. Recently, I asked ChatGPT about the primary research areas in socially responsible AI, and it identified four key areas: transparency, expandability, privacy, and data. While I agree with these areas, I believe that designing socially responsible AI is fundamentally about understanding human values. Although AI could occasionally produce errors, we should not worry too much or be overly critical, as we are not perfect. Even as professors, we may occasionally make mistakes, such as erroneously grading a student's work. While this has happened to me once in the past 16 years, it does not imply that I will not make mistakes in the future. We should give AI some flexibility and time and treat them as adults. Social responsibility for AI serves as a guideline, similar to how doctors must not intentionally harm their patients but may inadvertently prescribe the wrong medication due to human error. Social responsibility for AI is an excellent field that enables us to make a difference while also profiting from it if we do an excellent job.

YS: **Is responsible AI a technology issue or a business issue? For example, we all agree that AI needs to be "fair", but whose**

> **definition should we use? It is a company-by-company decision, and when you get into the details, it can be a difficult decision. What do you think about the role of human agency in developing socially responsible AI?**

MG: I would like to build on Prof. Zhong's comments regarding grading students' work. When I was a junior professor, I thought creating exams was straightforward, as I only needed to evaluate students' knowledge. However, as I have grown older, it has become increasingly challenging to determine what to evaluate and what I hope to achieve by doing so. Rather than simply assessing knowledge, I want to ensure that my students have acquired something meaningful from my class that will benefit them as they progress in their careers and personal lives. I still examine my students, but I ask them different kinds of questions that require more than just reciting information. This raises the question of what should be evaluated in AI. While knowledge is an essential component, it is not enough. Therefore, the central question is not whether responsible AI is a technical or business issue, but instead revolves around the value of life. Is human life equivalent to artificial life? Does artificial life have any value? Once we can answer these questions, we will be able to define social responsibility. Historically, humans have considered other humans as goods, which led to slavery. As our philosophical understanding has evolved, we recognized the value of all human life and abolished slavery. Similarly, our understanding of AI must evolve to reflect the value and impact of the technology. Therefore, social responsibility for AI should not focus solely on technical or business aspects, as we must consider issues with greater value and impact.

BZ: Talking about how AI can better serve society, the question of which stakeholders should be included in discussions is challenging. Who should be responsible for teaching ethics and improving AI literacy: the government, AI developers, CEOs of large technology companies, or all members of society? I do not have a definitive answer, but I am concerned when a single party dominates the conversation, whether it is the government, a technology company, an influencer, or an opinion leader. This creates discomfort and uncertainty about the direction of the discussion. For example, who will teach ethics to the next generation—the government, a political party, or self-proclaimed

geniuses? It is unsettling. However, this uncertainty fosters ongoing discussion about the stakeholders who can contribute to the debate on technology and its impact on society. In my opinion, during the process, we should prioritize humans and their preferences, which may lead to solutions for specific problems. Nonetheless, as a whole, it remains a difficult question to answer.

YS: **AI regulation across the globe is fragmented by cultural diversity. Meanwhile, there is a growing need for a global governance framework. Cultural differences may present a unique set of challenges to AI ethics and governance when we try to align the practices with the principles. What do you think about the dilemma?**

MG: While cultural differences are fascinating and essential, they should not serve as excuses for inaction. As human beings, we can find common ground when we genuinely desire to do so. The Sustainable Development Goals serve as a prime example of this, as many countries, including the United States, China, and North Korea, have agreed upon them for international cooperation. Common goals, such as sustainability, human rights, environmental protection, education for all, and development, are achievable when we put effort into them. However, we also acknowledge cultural differences: debates and contextualized discussions are crucial in finding common ground. Therefore, conducting cultural studies is vital, as I mentioned earlier. In my lab, we always incorporate a cultural dimension into any studies or projects we undertake because it is critical. There is nothing that the human mind cannot accomplish if we put forth the effort.

BZ: Cultural differences have become a more prominent topic than globalization in recent years. While globalization was discussed widely a few years ago, there is now little difference in the concerns shared by young people from the East and West, such as climate change, and they can communicate more effectively than many others. This suggests that cultural differences have diminished in some ways, which is a topic worth exploring. However, we should ask ourselves if we are comfortable communicating with individuals from different cultural backgrounds. Many individuals have expressed that Americans in Hong Kong are distinct from typical Americans, and some believe that any American with a passport is different from those without.

Differences can also be found in the West, such as in the United States and Europe. We should not presume that everyone shares the same cultural background as ourselves. Although we are both inspired by Chinese culture and have been influenced by it for thousands of years, I can see a significant difference between Hong Kong and mainland China. We should avoid assuming that everyone thinks similarly and instead presume that many differences exist. Nonetheless, we should also keep in mind that cultural differences may be used as an excuse to fail to recognize the true value of humanity, as Prof. Guitton discussed.

YS: We have heard exciting news that you will launch a new Elsevier journal called *Computers in Human Behavior: Artificial Humans (CHB: AH).* What is the vision of this new journal? How does it distinguish itself from its sister journal—*Computers in Human Behavior (CHB)*?

MG: The launch of a new journal is an exciting development, and I am personally thrilled about it as both an editor and a scientist. I would like to address three crucial points regarding the journal:

1) Research on artificial humans is currently siloed among various fields, such as social robotics, AI, and avatars. There is little communication between researchers in these fields despite their shared interest in artificial humans, regardless of the form they take. As a result, there is a need for a new journal that brings together all those interested in this area;

2) Most research on AI focuses on technical aspects, whereas we need to consider the social sciences and humanities aspects, such as philosophy, psychology, sociology, and more. Currently, there is no platform to publish such research. Therefore, we have created *CHB: AH* as a new platform that enables social scientists to contribute to the debate;

3) The third reason for launching this journal is to have a scientific publishing venue grounded in a social vision and mission. *CHB* aims to contribute to societal debates related to technology, and we wanted to include "artificial humans" under its umbrella. However, *CHB: AH* is narrower and encompasses not only cyber behavior but also neuroscience, cyborgs, and robotics. While the two journals are satellites of each other, *CHB: AH* is a bit on the side, with its own network

that connects with computer science, neuroscience, and brain research. Therefore, this new journal will not collide with the same topic but complement and expand on the issues tackled in *CHB*.

BZ: I am thrilled to share the mission of *CHB: AH*, which accurately represents the essence of our work. Our focus is on artificial humans, and we are committed to incorporating perspectives from the fields of humanities and social science rather than being solely an AI journal. This collaborative effort aims to shed light on the black box and related issues. The term "artificial intelligence" may become outdated soon, just like buzzwords such as blockchain and non-fungible tokens (NFTs). I find Elsevier's decision to not start the new journal from scratch commendable. Instead, the new journal is placed under the umbrella of *CHB*, which is already a well-established venue. This approach provides shade for the new journal to grow much faster. I believe that *CHB* is a unique venue that can successfully launch this journal. Launching it in the East is also meaningful as it emphasizes a worldwide perspective on AI research, on artificial humans, through the vector of our venue in this new platform. I encourage everyone to seriously consider submitting their papers to *CHB: AH*.

YS: As two highly experienced journal editors, what do you usually look for in a good manuscript?

MG: We welcome anything that contributes to the debate, nurtures our thinking, and enriches our soul regarding artificial humans, as we are discussing ourselves as humans. Therefore, anything that fosters our understanding of artificial humans is encouraged. This openness is a trademark of *CHB*, as we prioritize good science in our journal.

BZ: We discussed whether to accept both quantitative and qualitative papers and agreed that any approach would be acceptable, including critical studies, as long as it provides insights. We require more perspectives on AI in liminal worlds, where we exist in both the current and future worlds, to facilitate growth. I am confident that this field will expand rapidly in the coming years as we share the vision and passion for AI research, not just from a natural science perspective, but also from humanities and social science perspectives. How about next year? We will report to you about how much we grow up.

Q&A Part

Q1: **I teach a social media course, which I have previously taught to undergrads. To begin the course, I ask the question, "Does technology shape human behavior, or does human behavior shape technology?" While both are true, different approaches and theories are associated with each. The evolution of computer-mediated communication has given rise to social media and artificial humans or AI, and we need to consider how to approach this new development. From a communication science perspective, it is worth noting that some theories are not necessarily new, and people often attempt to explain new phenomena using old theories. The critical question is whether these theories are adequate for explaining human responses to these new phenomena. As a teacher, I am contemplating whether to teach my students about old theories and whether old theories are consistent with new phenomena. What is your perspective on this matter?**

BZ: Theory building is an ongoing process, even for established theories like the Agenda-Setting Theory, which has undergone several revisions. Therefore, there is always room for new theoretical contributions to the field. After receiving my tenure, I consulted with leading scholars in the United States to seek advice on how to approach theory building. They recommended focusing more on adopting and modifying existing theoretical frameworks to make them more expandable and better, enabling us to improve our understanding of technology. As such, I strongly encourage my collaborators and students to explore frameworks from various fields and not limit ourselves to borrowing only from psychology. Wonderful and interesting research is being done in other fields, such as information science, medical science, and business, which could enhance our understanding of technology. For example, attending a small conference on health IT and analytics gave me new insights into business research and marketing. As scholars, we should continually review and expand our existing frameworks to ensure that they remain relevant to new phenomena and challenges. This is what we refer to as a theoretical contribution.

MG: Theory is the structure that we use to help us understand phenomena. However, their effectiveness lasts only until they are proven wrong, and

then we must move on to new ones. It is questionable whether theories developed before the internet existed are still relevant to explain modern technology, especially in 2023. While it is essential to teach historical content, researchers should not limit themselves to outdated theories. Instead, they should develop new frameworks that are more relevant to current phenomena. It is also unnecessary for reviewers to write extensively about the theory. What matters most is whether the methodology and results are compelling. If the methodology is sound, but the result does not align with the theory, it indicates that the theory is incorrect. We should not restrain ourselves to a specific theory, as any theory can become outdated at some point. However, we should test new ideas when we think that current theories are no longer effective. While I am open to theoretical contributions, generating a new theory every two weeks is not feasible or groundbreaking. Instead, we should aim to make meaningful contributions that bring something new. Ultimately, we should think about our contribution to science at the end of our lives, which may be a theoretical framework, but it is unknown what that contribution will be.

Q2A: We are astonished by ChatGPT's capabilities, but as LLMs become more advanced, editors and publishers face new challenges. While ChatGPT can serve as a research assistant or collaborator, some articles have already listed it as a potential author. However, the use of ChatGPT or similar AI poses a risk of facilitating research fabrication. As editors, how do you address these challenges brought about by AI-powered LLMs?

MG: The question posed is both interesting and challenging, and while my answer may not be definitive, it reflects our current thinking as the situation is continually evolving. Trust is the foundation of research, and falsifying data or having ChatGPT as an author is a completely different matter. If someone uses ChatGPT to fabricate or manipulate data, the paper should be retracted, and the individual responsible should be fired. Such conduct would be heavily disregarded in the scholarly world. However, if ChatGPT is used honestly as a research assistant or collaborator, its contribution should be acknowledged transparently. If the use of ChatGPT goes beyond writing, such as providing new ideas or angles, it may be appropriate to include it among the authors. The decision to do so should be based on an honest

evaluation by the researcher who conducted the work. If ChatGPT is used solely as a writing tool, its use should be acknowledged in the methodology section, similar to how we acknowledge the use of statistical software. Transparency is key, and the use of technology should always be acknowledged. However, using AI to generate false data is an act of forgery and a double crime because it is challenging to detect. Such conduct is unacceptable.

Q2B: Will you consider accepting ChatGPT, or other AI, as a legal author?

MG: Personally, I would be willing to accept ChatGPT or AI as a legal author, but in the event that it occurs for the first time, I will need to consult with Elsevier and the publisher. The situation is not so different from when a paper is authored by a consortium, which happens occasionally in social science but is more common in physics and certain fields of medicine. In some cases, research federations or collaborative works may have up to 40 authors. Therefore, if the research were conducted by Celine Song, Bu Zhong, Matthieu Guitton, and ChatGPT, I see no reason why ChatGPT cannot be recognized as an author. However, it is important to note that ChatGPT cannot be the corresponding author when presenting the work.

BZ: It is important to identify which parts of the work were completed by ChatGPT to avoid any confusion or potential issues. Otherwise, I would not like to be associated with such a tool. It is uncertain what the future holds for the development of such technologies. However, it is worth noting that when the first spell checker became available, there were concerns that students would become overly reliant on technology and lose the ability to spell. Some professors even predicted that comprehensive exams would no longer be effective. However, this did not come to pass, and we continue to utilize technology without significant changes to our ability to learn.

Selected Works by Matthieu Guitton and Bu Zhong

Please refer to the end of the Chinese version of the dialogue for Matthieu Guitton and Bu Zhong's selected works.