

科技文化

人工智能時代的文學變革

● 陳楸帆

眾所周知，人工智能 (AI) 已經成為這個時代最火熱的概念之一。我來香港後發現，幾乎所有大學都在積極擁抱 AI，不僅理工科，就連人文學科、傳媒領域也紛紛要開設「AI+」的課程。雖然挑戰巨大，但我認為這是主動迎接變革的表現。作為一名科幻作家，我深切感受到 AI 給寫作帶來的衝擊。所以我以下想從一個創意寫作的視角，談談這幾年的實踐、思考，以及未來可能面臨的挑戰。

我們不妨先追溯一下歷史。1818 年，雪萊 (Mary Shelley) 創作了《弗蘭肯斯坦》(*Frankenstein*，又譯《科學怪人》)，這部作品被公認為現代科幻小說的開端。在書中，科學家弗蘭肯斯坦用屍體碎片拼接，通過現代科技賦予其生命，創造出一個非人的怪物。當怪物獲得自我意識後，他開始思考：我是誰？我從何而來，又將去向何方？我與創造者是甚麼關係？世間還有沒有同類？最後，創造者被殺，怪物也鬱鬱而終。二百年後的今天，這個故事仍在警示我們：人類對「他者」的恐懼仍根深蒂固。而我們今天最新的「他者」正是 AI。

事實上，AI 並非新鮮事物。1956 年達特茅斯會議 (Dartmouth Conference) 上首次提出「人工智能」的概念，而在此之前，阿西莫夫 (Isaac Asimov) 的科幻小說已探索了人機共生的未來圖景，維納 (Norbert Wiener) 的控制論思辨人與機器之間信息交互的反饋迴路。隨後幾十年間，儘管 AI 的發展一波三折，但人們從未停止對「思考的機器」的想像和創造。直到上世紀 90 年代，以深度學習為代表的技術取得突破，這一輪 AI 浪潮才如火如荼地席捲而來。而 2017 年「注意力機制」(attention mechanism) 和“Transformer”的提出，以及 GPT、BERT 等大語言模型的發布，令 AI 在文本創作領域的表現愈發驚豔。從某種意義上說，AI 正在成為二十一世紀弗蘭肯斯坦的怪物。

* 本文源自筆者在 2024 年 11 月 27 日於香港科技大學的演講，題為「人工智能與文學變革」，嗣經整理刊出。

未來人機協作、共同創作將成為常態。作為人類，我們究竟該如何看待AI？是懼怕還是擁抱？是競爭還是合作？我想，這是每個創作者都無法迴避的問題。

一 我的AI寫作實驗：從險勝莫言到「共生紀」

接下來我將結合自身經歷，談談這幾年與AI一起寫作的體驗。2017年，我與中信出版社簽約出版科幻小說集《人生算法》，主題就是探討人與AI的共生。為此，我找到以前在谷歌(Google)工作的工程師王詠剛，請他幫我搭建一套AI寫作模型。雖然那時的模型還非常粗糙，但通過給模型餵入我自己的寫作語料、反覆調試，它最終能夠生成有一定可讀性的語句。我把這些內容放進小說裏，作為人機合作的創作實驗，這場實驗也讓我獲得了中國的首個「AI文學榜」獎項。這個獎項由上海一家官方文學機構主辦的《思南文學選刊》發起，特別之處在於評委不是人，而是一個AI程式。它能給參選作品打分，我的小說《出神狀態》在綜合排名中以0.00001分險勝諾貝爾文學獎得主莫言的作品，奪得短篇小說組第一。

一個AI評選出另外一篇由AI輔助創作的作品作為優勝者，這件事堪稱「科幻本幻」(科幻中的極致科幻)。那篇獲獎的小說刊登在《小說界》雜誌2018年第4期上，以下是其中AI生成部分的選段①：

遊戲極度發燙，並沒有任何神秘、宗教、並不攜帶的人，甚至慷慨地變成彼此，是世界傳遞的一塊，足以改變個體病毒凝固的美感。



筆者與參加「人工智能與文學變革」演講的師生合照。(圖片由陳楸帆提供)

大家可以看到，這就是AI生成的原始內容，沒有經過任何修改。可以說，這是一種極其先鋒的風格。從中不難看出，當時的AI在邏輯連貫性、句法、人物理解等方面還非常欠缺。但這只是一個開始。

2020年，我與更專業的團隊合作，用更大的模型、更豐富的語料進行訓練，並加入不同科幻小說家的「風格濾鏡」，讓AI學習不同作家的寫作風格。借助更強大的語言模型「AI科幻世界」，機器生成的段落在連貫性、邏輯性、人稱指代等方面有了長足進步，基本可以獨立完成一個簡單的故事情節^②：

一個老太婆突然出現在他面前，她說話的口氣中有一種沙啞的東西。她告訴花花，地球人正在清理太陽系中所有的生態環境，用一種稀薄的物質製造出一個新世界，但老太婆又補充道，那只是一個大概的設想。

她邀請他去火衛二，那是另一個被遺棄的小行星。他們相信他能幫助他們製造出分子縮微器，這樣，地球上的分子就可以被拼裝成任何他們想要的樣子。他們說，這將是人類歷史的一大進步。

上面的文字都是由「AI科幻世界」生成的，我嘗試用這種方法創作了一篇以歌手華晨宇為主角的科幻小說《火星奧德賽》。AI扮演了我的「槍手」，承擔了大部分寫作任務，最後發表在2021年的《時尚芭莎》雜誌上。

與AI共創的過程中，我時常反思：創作的主體是人還是機器？如何定義「寫作」這個行為本身？從技術層面看，AI是通過分析海量語料，運用CNN (Convolutional Neural Network, 卷積神經網絡)、LSTM (Long Short-Term Memory Network, 長短期記憶網絡)、RLHF (Reinforcement Learning from Human Feedback, 人類反饋的強化學習)等算法進行深度學習，從而獲得語言理解與表達能力的。它與人腦的認知機制有相通之處，但並非對人腦的簡單模仿。從創作實踐看，人與AI是協作關係。AI負責那些相對結構化、數據易於學習的部分，如替換名稱、細節擴展、風格變遷等；而統領全域、把控方向、判斷，則是人類不可或缺的抽象與審美能力。我把「AI科幻世界」視為「陳楸帆2.0」，是我創作過程中的好幫手。在未來，人機協同將成為創意寫作的新範式。

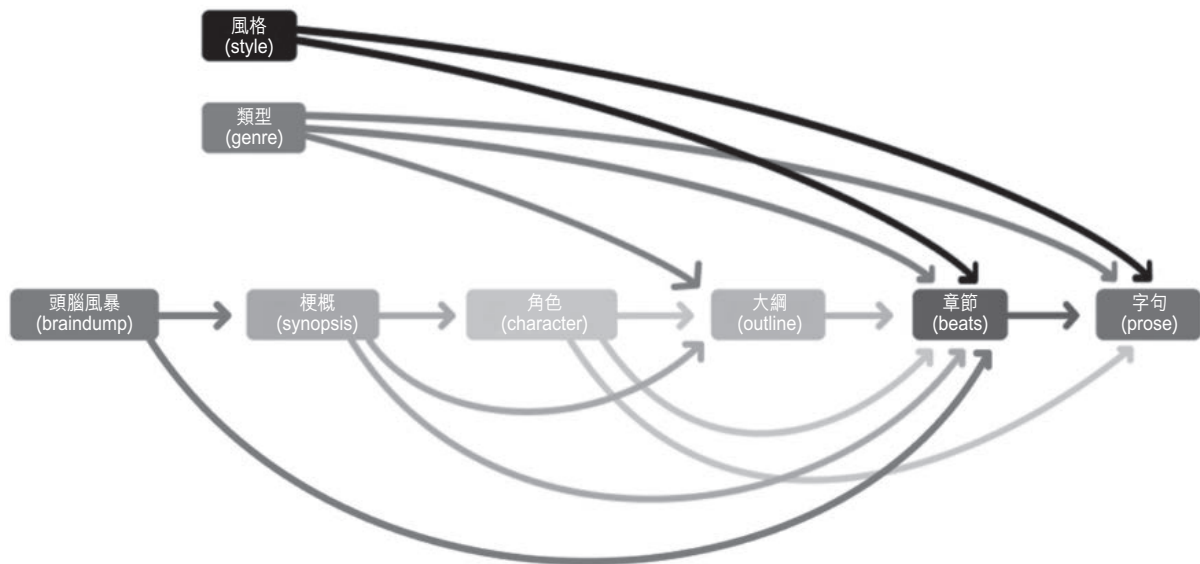
這裏我想特別提到「共生」這個概念。2020年，我策劃了一個名為「共生紀」的項目，邀請多位作家與AI一起創作科幻作品^③。它的英文名叫“Co-Creation”，寓意為打破人機二元對立，走向真正的共生。在我看來，人類與AI的關係不應是競爭，而應是協作；人類對AI不應是懼怕，而應是擁抱。只有敞開胸懷接納彼此，在創作中互相激發、互相成就，我們才能開創科幻寫作的新境界。正如阿西莫夫所言：「人類和機器人之間的分歧可能沒有智能和非智能之間的分歧那麼重大。」^④

二 工具進化史：模塊化寫作與創意寫作教學

在近年來的創作實踐中，我深度應用了多款AI輔助工具，這些工具正以驚人的速度重塑創意寫作的範式。以GPT開放接口為基礎構建的寫作平台，曾試圖通過定制化功能解決特定創作痛點。但隨着大模型能力的躍遷，這類工具的精細化設計反而顯露出局限性——當GPT-4等通用模型已能通過簡單提示生成複雜文本時，過度定制化反而成為創作力的枷鎖。

比如AI寫作工具Sudowrite將寫作拆解為「靈感碎片→結構生成→風格適配」的模塊化流程；用戶可輸入零散關鍵詞，通過其“braindump”（頭腦風暴）功能生成故事梗概，再根據不同“genre”（文學類型，如科幻/言情/懸疑）調用預設角色模板，逐步細化至章節段落。該工具曾試圖通過“style”（風格）濾鏡功能模擬名家文風，例如將文本切換為海明威（Ernest Hemingway）的極簡主義或愛倫·坡（Edgar Allan Poe）的哥特氣質——其底層邏輯是將文學風格量化為語義空間的向量距離（如科幻與懸疑在向量空間中相鄰，而與言情相距較遠）（圖1）。

圖1 Sudowrite 模塊化寫作流程



資料來源：筆者繪製。

這種圖層化創作模式，使寫作者可像調整Photoshop圖層般精確控制敘事元素。在亞馬遜（Amazon）自出版平台KDP（Kindle Direct Publishing）上，已有作者利用此模式年產三百餘本小說，形成「文字流水線」奇觀。這在中國出版審查體系下雖難複製，卻揭示了技術對創作工業化進程的深度滲透。

當前最先進的AI協作工具（如Anthropic公司開發的Claude），在長文本記憶與知識庫管理上已超越GPT系列，其項目在功能上允許載入20MB大小的

背景文檔，使《潮星》(未出版)這類需要在超過三百頁的生態演化設定中進行深度推理思考的科幻創作成為可能。AI能瞬間檢測「火星重力參數與建築設計的邏輯衝突」，這種能力遠超人類作者的細節把控力。

在我的創意寫作課堂上正在實踐以下一種新型人機共生的創作生態：

- **靈感破壁**：學生將碎片化意象或隨機關鍵詞輸入AI，通過迭代生成獲得「敘事種子」，以減少從0到1的心理障礙；
- **認知鏡像**：用AI檢測文本中的文化偏見，如無意識中的種族主義凝視與男性中心視角；
- **元創作訓練**：分析機器生成的「賽博俳句」，通過記錄調教AI生成詩句的過程，更好地理解接受美學的運作機制。

站在人機協作的臨界點，我們既要警惕「自我工具化」的異化風險，也要摒棄盧德主義(Luddites)式的技術恐懼。當AI能批量生產合格線以上的小說時，真正的創作恰恰要從「達標」走向「破格」——就像傳統水墨中的「留白」需要用心經營，AI生成的文本框架必須經由人類的情感浸潤才能獲得靈魂。我們也要在算法規訓中撕開裂隙，讓那些無法被量化的戰慄與頓悟，繼續照亮人類獨有的文學暗地。

三 超越「怕」與「愛」：如何在AI時代創作科幻

與AI合作寫小說之後的2019年，我也與李開復博士一起創作了《AI未來進行式》一書，用十個2041年發生在不同國家的故事，勾勒出一幅全球化的未來AI圖景^⑥。我們想以一種通俗易懂的方式，告訴普通讀者AI將如何改變我們的生活。當時我們設想，AI將廣泛應用於教育、醫療、軍事、娛樂、交通、養老、能源、心理治療等各個領域，既會帶來便利，也可能加劇現有的社會問題，引發倫理挑戰。

值得一提的是，這本書在2021年推出後一年ChatGPT就發布了。現在回頭看，書中的很多想像已成為現實。也許用不了二十年，只需要五到十年，書裏的大部分故事就會實現，甚至被現實趕超。AI的發展之快，確實令人驚歎。現在回過頭看，儘管一些技術細節因為AI的高速發展而顯得過時，但對未來趨勢的判斷卻大都應驗了。這再次說明，在科幻創作中，對人性的洞察往往比技術細節更加重要。畢竟人性和社會的基本規律是相對穩定的基石底座，而技術浪潮卻來去匆匆，各領風騷三五年。

寫作這本書時，我們也在形式上進行了探索，將科幻小說與非虛構內容融為一體，呈現出獨特的文體結構。我不確定這是否受到了《聊齋志異》中「異史氏曰」夾敘夾議傳統的潛移默化，但將想像與理性思考相結合，確實能讓AI

這個宏大命題變得更貼近現實、更具可讀性。對科幻寫作而言，文體創新無疑是應對AI時代的一個重要策略。

說到這裏，我不禁想起二十世紀中葉在美國興起的一種科幻子類型，劉慈欣將其形象地稱為「工程師科幻」，代表作包括克拉克 (Arthur C. Clarke) 的《與羅摩相會》(*Rendezvous with Rama*)、尼文 (Larry Niven) 的《環形世界》(*Ringworld*) 等^⑥。它們以當時的前沿科學理論為基礎，展開對近未來技術的想像，強調工程學上的細節、真實性和科學性。這種「工程師科幻」傳統，在劉慈欣的著作中得到彰顯；在我近年的創作中也得到延續，但又呈現出新的變化。《零碳中國》裏對碳中和社會與技術的推演，《荒潮》中對由病毒與垃圾混合產生賽博格 (cyborg) 物種的設定，《AI未來進行式》基於現實AI技術發展路徑所做出的合理想像，既延續了前輩們「以現實為跳板」的創作方法論，又注入了更濃厚的人文關懷視角。這種變化源於我對傳統「工程師科幻」中「技術樂觀主義」敘事的警惕。在我眼中，宏大敘事不應成為忽視個體價值的遮羞布，科技發展也不能抹殺貶低人文主義的光芒。人、技術、環境三者之間存在着複雜關係與能動性網絡，人類要學會與作為「他者」的AI合作共生，建立起平等互惠的夥伴關係，而不是把它當成純粹的工具或威脅。只有超越「怕」與「愛」的二元情緒，心存敬畏與慈悲之心，我們才能馴服技術，讓它服務於全人類的利益。

四 「狂野西部」：AI對文化產業的全面衝擊

接下來，我想談談AI時代文化產業所面臨的一些更現實的問題。

2023年，索尼世界攝影獎 (Sony World Photography Awards) 得主埃爾達格森 (Boris Eldagsen) 拒領獎項，因為他的獲獎作品根本不是攝影，而是AI生成的圖像，他質問「攝影的本質是否已死」。而2022年，科羅拉多博覽會 (Colorado State Fair) 數字藝術類冠軍《太空歌劇院》(*Théâtre D'opéra Spatial*) 作者則公開使用圖像生成工具 Midjourney 創作，引起巨大爭議。這種價值觀的撕裂在文學領域更為尖銳——2024年，《東京都同情塔》獲得日本主流文學大獎芥川獎，作者九段理江承認書中有5%的內容由AI生成。當AI生成內容在「雙盲評審」(peer review) 中瞞過人類評委，我們不得不重新審視這樣的質問：AI創作的作品能獲得人類獎項嗎？

2024年「番茄小說斷更事件」暴露了平台經濟的掠奪性。網文平台番茄小說的合同條款規定作品內容將無償用於AI訓練，引發兩萬部小說的作者集體斷更 (中斷更新) 抗議。與之形成對比的是2023年11月北京互聯網法院的里程碑判決：認定用戶通過調整提示詞生成的AI圖片《春風送來了溫柔》享有著作權。這種司法分歧映射全球AI監管困局——美國《紐約時報》(*The New York*

Times) 訴 OpenAI 案便聚焦「合理使用」的邊界，日本則允許非感知性數據挖掘（即並非對圖像層面進行拼貼修改，而是高維向量空間的數據挖掘）。在 AI 時代，如何界定作者身份？版權該如何歸屬？濫用的界限如何劃清？這需要倫理和法律的界定。但具體的權益分配，恐怕還需要案例積累和制度完善。我們應該建立一套公平合理的交易機制，既保護創作者的權益，又為 AI 技術發展留出空間。

AI 對文學產業的影響不容小覷。一方面，AI 確實可以提高創作效率，催生出更多優質內容。以網絡文學為例，一些平台已經嘗試用 AI 來輔助寫作，填充枝節內容。但另一方面，過度依賴 AI，也可能加劇同質化、快餐化的問題。而「AI 陪伴」、「AI 互動」之類的社交應用竊取了大量文學作品中的人物、情節、設定，更是對人類寫作的一種消費主義異化。

面對洶湧而來的 AI 浪潮，傳統的寫作模式正面臨巨大衝擊，但也迎來了前所未有的想像空間。事實上，從口頭傳統到寫本時代，從印刷時代到數字時代，文學創作的工具箱從未停止擴充，人的表達欲望也在與日俱增。作為個體，我們渴望理解並表達自我，用獨特的方式詮釋人生。文學為人的自我認知和情感表達提供了一面鏡子。這樣的價值在 AI 時代會否受到挑戰與貶損呢？我的回答是：恰恰相反。

五 以 AI 為鑒：人的變革帶來文學的變革

在 AI 時代，文學從未像今天這樣重要。為甚麼這麼說？讓我們回到本文最初的問題——為何人類執著於創造能像人一樣思考的機器？這本質上是對雪萊《弗蘭肯斯坦》隱喻的現代呼應：即便預見到技術可能帶來的失控風險，我們仍無法抗拒探索自身思維本質的衝動。2024 年諾貝爾物理學獎得主之一、深度學習之父、ANN (Artificial Neural Network, 人工神經網絡) 的發明者辛頓 (Geoffrey Hinton) 的初衷正源於此：他試圖通過構建機器模型，逆向破解人類認知的密碼。這種科學探索暗合哲學家羅素 (Bertrand Russell) 的觀點——「若無法復現某物，便談不上真正理解它」^⑦。這一還原主義思維，驅動着人類將意識拆解為算法與數據的組合，試圖在矽基載體上重構智能的奇迹。

然而，AI 的迅猛發展迫使人類直面更本質的詰問：人是甚麼？文藝復興時期的人文主義覺醒、達爾文 (Charles Darwin) 進化論對「人類中心論」的顛覆，再到今日 AI 引發的認知革命，這一問題始終如影隨形。當機器能通過圖靈測試 (Turing Test)、生成媲美人類作家的文本時，我們不得不重新審視自身在智能譜系中的位置。一些人堅持人類高高在上的獨特性，另一些人則視其為複雜算法驅動的「濕件機器」(wetware, 即人機連結的設備)。這種爭議本身，恰是人性複雜性的鏡像。

在AI時代，文學的價值遠未被充分認知。它不僅是娛樂消遣或精英階層的文化符號，更是理解人類本質的核心界面。這一界面的獨特性在於其雙向性。一方面是對外映射：文學能記錄人類對世界的感知，如馬奎斯 (Gabriel G. Márquez) 《百年孤獨》 (*Cien años de soledad*) 中馬孔多鎮的魔幻現實，實為拉丁美洲歷史的隱喻投射；另一方面是向內探索：通過文字，我們得以窺見意識深層的潛意識與集體無意識，如卡夫卡 (Franz Kafka) 《變形記》 (*Die Verwandlung*) 便是心靈對異化焦慮的具象化。

AI對文學的學習，意外揭示了人類認知的盲區。大模型通過分析海量文本，識別出語言中隱藏的情感模式與文化規訓，甚至能生成觸及人性深層的隱喻——這種能力暗示着：文學可能成為機器理解人類意識的橋樑，而AI則成為人類反觀自身的鏡鑒。在AI時代，文學這面鏡子竟是可以流動的，如同佛教中所謂的「因陀羅網」，每一顆寶珠都能照見所有其他的寶珠，互相依存，牽一髮而動全身。

當AI能夠生成流暢的故事時，創意寫作的意義正從「技藝傳授」轉向「認知療癒」。通過寫作，人們得以：

- **解構自我**：追問為何反覆書寫父子衝突、身份迷失等主題，實為挖掘潛意識中的情感礦牀；
- **對抗異化**：在算法推送與信息過載的漩渦中，寫作是剝離數字干擾、回歸內觀的精神錨點；
- **集體療癒**：如疫情期間全球湧現的「新冠文學」，通過共享創傷敘事重建心靈聯結。

而這些，恰恰是目前的AI模型尚且無法觸及乃至終究抵達的本體論彼岸。

六 結語：文學讓我們照亮彼此

正如韓裔德國哲學家韓炳哲在《非物：生活世界的變革》 (*Undinge. Umbrüche der Lebenswelt*) 中指出：「在人初始的感動中，思維彷彿是在自身之外的。基礎情態將思維放置到了某一種外在之中。人工智能不進行思維，因為它從來不是在自身之外的。」^⑩人類的思維本質在於「遞歸性」——我們總能在思考之外反觀思考本身。這種能力使創作不僅是表達，更是對存在狀態的持續叩問。而AI的局限性恰恰在此：它能模擬悲傷的語法，卻無法體驗淚水滑過臉頰的溫度。

巴西學者弗盧塞爾 (Vilém Flusser) 在《書寫還有未來嗎？》 (*Die Schrift. Hat Schreiben Zukunft?*) 中預言：人類需學會用數字符碼重新書寫世界，否則將陷入「文盲的野蠻」。這種轉型不是簡單的媒介遷移，而是思維範式的顛覆——

從線性文字到粒子網絡，從邏輯推演到多維交互^⑨。然而，我又想起了80年代李澤厚老師在《美學四講》中的深情呼喚：「於是，回到人本身吧，回到人的個體、感性和偶然吧。從而，也就回到現實的日常生活(everyday life)中來吧！不要再受任何形上觀念的控制支配，主動來迎接、組合和打破這積澱吧。……於是，情感本體萬歲，新感性萬歲，人類萬歲。」^⑩

這貌似南轅北轍的兩條岔路將何以匯合？道家「生生之謂易」的智慧或許為此指明方向：真正的變革不是非此即彼的替代，而是共同進化(co-becoming)。當人類與AI在創作中形成「控制論迴環」——機器提供跨維度的靈感迸發，人類注入倫理判斷與情感溫度——我們或許能抵達佛陀「圓融無我」或道家「天人合一」的境界。

站在技術狂飆的十字路口，文學的終極使命愈發清晰：它既是捍衛人性邊疆的堡壘，也是通向未知的星圖。當我們用AI生成一幅「超越『怕』與『愛』」的共生圖景時，或許正在重演古老的神話——不是普羅米修斯為人類盜火接受無盡的懲罰，而是人類與機器共執火把，照亮彼此，也照亮更廣闊而未知的存在。

註釋

① 陳楸帆：〈出神狀態〉，《小說界》，2018年第4期，頁124-31。引用部分為AI程式通過深度學習筆者的風格創作而成，未經人工修改。

② 陳楸帆：〈火星奧德賽〉，《時尚芭莎》，2021年第1期，頁56-64。

③ Masha Borak, "Artificial Intelligence Is Helping 11 Science Fiction Authors Write New Stories for Publication Every Week" (31 October 2020), www.scmp.com/tech/innovation/article/3107451/artificial-intelligence-helping-11-science-fiction-authors-write.

④ Isaac Asimov, *The Caves of Steel* (London: Grafton, 1958), 37.

⑤ 參見 Kai-fu Lee and Chen Qiufan, *AI 2041: Ten Visions for Our Future* (New York: Currency, 2021)；李開復、陳楸帆：《AI 2041：預見十個美麗新未來》（台北：天下文化，2021）；《AI 未來進行式》（杭州：浙江人民出版社，2022）。

⑥ 〈劉慈欣人民日報撰文：在科幻中感受科技的溫度〉（2024年1月26日），澎湃網，www.thepaper.cn/newsDetail_forward_26150412。

⑦ Bertrand Russell, *The Problems of Philosophy* (London: Oxford University Press, 1998), 32.

⑧ 韓炳哲著，謝曉川譯：《非物：生活世界的變革》（上海：東方出版中心，2023），頁67。

⑨ 弗盧塞爾(Vilém Flusser)著，朱恬驊譯：《書寫還有未來嗎？》（上海：東方出版中心，2024），頁175-79。

⑩ 李澤厚：《華夏美學·美學四講》，增訂本（北京：三聯書店，2008），頁407。

陳楸帆 中國科幻作家，畢業於北京大學中文系與藝術學院，代表作包括《荒潮》、《人生算法》、《AI 未來進行式》等，被翻譯為多國語言。