

三邊互動 三邊互動 三邊互動

當今世界局勢複雜多變，在不確定的時代中，大學發揮的社會功能及其扮演的傳統教育角色，在人工智能興起的背景下，需要如何作出改變，以應對前所未有的挑戰？本刊將於新學年伊始（10月號）刊出專輯，與讀者分享有關「大學之道」的一些思考。

——編者

透明與不透明的博弈與和解

人工智能（AI）出現後，知識生產主體發生了變化，人類與AI「平分天下」，甚至AI知識產出速度與佔比大幅超越人類，形成「一統天下」的局面。知識生產主體難分彼此，擾亂了受眾對生產源頭的知情權。然而在AI發展的現階段，人類在體驗了AI知識生產帶來的震驚之後，出於私欲本能而迴避、隱瞞、篡改AI貢獻的部分，以達到對AI知識的佔有目的，這一現象及其後果正是孫皓琛在〈輸出透明度：AI創造力的法律監管路徑〉（《二十一世紀》2026年6月號）一文所作的分析。隱瞞行為和技術，使知識的作者屬性變得撲朔迷

離，不但給鑒別和使用帶來了難度，且造成了文章作者所說的多層次、多維度的實質性危害。乃至對AI自身的發展來說，也由於開發者給AI不斷循環投餵「近親繁殖」的內容，有可能導致模型崩潰，AI演進發生了「優生學」意義上的退化。

那麼如何克服AI隱瞞，避開AI發展的陷阱？文章首先給出一個創造意圖、情感表達和創造過程控制的指標，表面上看，這可讓人類和AI創造力兩者從此涇渭分明。然而，事實正如文中的擔憂：「僅憑對輸出結果的觀察，我們已無法可靠地區分二者。」問題的癥結正在於此，受眾站在輸出端，無法了解知識生產過程「暗箱」內發生的活動。就像看見一張照片，不代表看見照相機內部不透明的機械、光學和化學製像過程。

在此，作者給出的方案是監管關口前移，好比站到照相機背後——攝影師的位置來觀察。問題是，製像裝置內部發生的活動依然是「不透明」的。然而悖論在於，要想得到照片就必須保證製像過程在「暗箱」內封閉地進行，否則成像、曝光這些技術運作過程就

無法實現。縱使通過法律介入，即以「輸出透明度」（output transparency）來規範AI的運用，也因技術問題而難以實現。

作者給出的「透明度」即強制嵌入標識。好比養殖場給雞蛋噴碼，帶上可溯源的出生證明，讓消費者通過體驗信息透明而建立購買信任。然而，作者也在擔心，標識（噴碼）只能控制科技公司（大型養殖場大宗供應超市的雞蛋），而無法控制用戶（農家土雞蛋消費者）。此外，這一方案在理論上可行，但在實際操作層面依然存在監管空白——比如對規模較大的養殖場可以監管，而對零散農家的土雞蛋則難以監管（監管成本過高）。

算法與數據透明度看似解決了「輸入性不透明」問題，但是AI隱瞞依然會發生在輸出端。而對「輸入—輸出」全過程強制標識，又會限制「人工+智能」的合作及人類的創造性空間。文章寄希望於作者說明義務，但這是一個軟指標，而非法律監管路徑。人類在「人工+智能」合作成果中分一杯羹的貪念，難以靠作者說明義務的道德約束來制止。無論是強制標識無法做到全覆蓋的漏洞，還是難以落地的作者說明義務，均構成現階段對AI輸出內容在法律監管與道德約束層面的困境。透明化的監管預期與不透明的「反動力量」，也許會在一個相當長的時期內博弈、共存。

作者對人類與AI創造力進行區分，認為人工與智能各自界限分明。然而，一旦人工

開始應用智能——兩種智能交匯之後，便很難分清彼此，難以稱量各自在輸出內容中所佔知識「版權」的比重。長時段地看，人類對AI有足夠應用經驗也有足夠認識之後，將走出當前的高度黏着期，迎來「人類的歸人類，AI的歸AI」的認知自覺階段。彼時人類將更加自覺地回到對生命情感、經驗、想像力這些AI不可替代的資源的開發利用。

當然，法律與道德並非勢不兩立，柏拉圖(Plato)與亞里士多德(Aristotle)都在肯定城邦法律對公民德性的培養。在此，法律與道德殊途同歸，兩者都致力於構建「善」與「正義」的社群。而知識向來以「善」為目標，當人類的德性(必然包括AI素養)普遍提升之後，人類與AI之間爭奪知識生產所有權的戰爭方可停止，當人類與AI走向透明化的和解，彼時，正是實現作者所期待的「輸出透明度」的時刻。

張峻 蘭州
2026.6.22

社會、科技共同塑造下的加速主義

季衛東〈加速主義：數字時代的「華盛頓共識2.0」？〉(《二十一世紀》2026年6月號)一文從哲學與政治出發，聚焦美國現屆政府的政商關係，推測未來會更積極推動包括通用人工智能(AGI)的科技創新。文中明確指出，在規範科技與

加速創新之間，背後是人文主義與後人類主義之間的分別，而按照作者分類，加速主義(accelerationism)理念建基於後者。它大致可分為三大分支：主張科技解放人類的左翼，倡議科技企業家主政的右翼，以及支持AI驅動解決貧窮、戰爭和氣候問題的有效加速主義(effective accelerationism)。

文章強調2022年ChatGPT面世的背景下，美國現屆政府上台後大力推動改革，國家、科技與資本的合流無條件地促進科技創新。有別於上世紀新自由主義經濟學的所謂「華盛頓共識」(Washington Consensus)，作者提出值得關注的推測，指「華盛頓共識2.0」也許會由科技資本所驅動，出現效率至上的企業型國家、減少監管的自由市場、取代民意的演算法、後人類的「科技奇點」等特點。固然從控制論和後人類主義來看，宇宙演化、資本主義和AI都是一種必然、自我優化的力量，假若人類試圖減速將會徒勞無功。

加速主義的哲學思想，對於社會學者來說並不陌生。比如著有《社會加速：現代社會中時間結構的改變》(*Social Acceleration: A New Theory of Modernity*)的社會批判理論學者羅薩(Hartmut Rosa)認為，現代性與社會加速是如影隨形的關係；由於交通和通訊工具、社會變遷與人際關係、生活節奏三個結構性因素所形成的社會加速，使人無法自主、自控，產生出新型異化。羅薩也認為結構性因素令社會發展難以減速(deceleration)，而他提倡個人對現實世界有所觸動

並與之連結，建立人與事物的共鳴，以避免人的異化。

然而，社會學者提倡的社會科技取徑(socio-technical approach)既不認為科技發展是宿命，又不認為加速僅僅屬於社會與文化的現象，反而強調科技的進退是由制度、價值、物質與人造物持續互動共同塑造的。科技固然可以成為變革的催化劑，但它從來不是推動社會的唯一力量，也不應該形成不可逆轉的線性前進路徑。故而，科技是進是退，並非加速中的社會文化所能單獨決定的。

事實反映，社會和科技兩者往往難以分割。比如日本福島核事故之後，媒體批評大致可分為兩類：一類是指責科技，如反應堆的設計；另一類是指責社會，如規章欠奉。而針對核電技術的規範與放寬，則難以達成一致的國際共識。不同社會對福島核災的回應方法亦不盡相同，比如德國關閉國內的核電廠，而法國則仍舊倚重核能。社會與科技既彼此交互、共同塑造，而人在其中扮演的角色也不容忽視。

說到底，我們的未來，其實是由一連串的選擇而成，並不是特定的個人、團體或企業自上而下來決定的。我們不該只仰望「科技奇點」，未來具體展現在我們每一次的點閱、分享、買單等行為中，也體現在我們如何參與、塑造科技創新的集體決定中。在加速與減速的選擇之間，不僅在於向科技專家、資本家或政策制定者問責，更在於個人日常行為與集體決定。

張立健 新加坡
2026.7.6