

3C 科技

在近兩年的時間裏，AI大模型技術的發展可謂飛猛進，它已經成為科技界大熱的話題。ChatGPT的橫空出世，不僅在行業內引起了轟動，也讓我們普通人見證了人工智能的巨大潛力。

目前，在全球科技界，人和資本都在爭相湧入AI大模型的領域。AI大模型的應用已經滲透到小說等內容創作、學術研究、音樂製作，乃至於生成電影片段等各個領域，為我們的生活帶來了前所未有的可能性。

姚剛

我們也不斷聽到AI大模型企業獲得巨額資本注入的消息。這一切似乎都在暗示，AI大模型技術已經成熟，它就像移動互聯網一樣，成為了我們生活中不可或缺的一部分。

儘管有人認為ChatGPT的出現標

誌着人類進入第四次工業革命的新時代，它就像歷史上的「火」和「電」一樣，將徹底改變我們的生活方式，但事實上，我們對當前的AI大模型的理解還存在許多誤區。我們需要以更加理性的態度來看待這項技術，避免盲目跟風。

AI大模型

話你知

AI大模型是具有大規模參數和計算資源的機器學習模型，在訓練過程中需要大量的數據和計算能力，通常擁有數十億乃至數千億個參數，其設計目的在於提高模型的表達能力和預測性能，在處理複雜任務時能更好地捕捉數據中的模式和規律。

誤解四 人人都會用 AI大模型

目前，普通人能夠接觸到的主流AI大模型普遍都採用了和AI聊天的交互方式。這種方式直觀易懂，幾乎不需要額外學習成本。可以說，只要會打字，人人都知道如何跟AI大模型進行交流。然而，這種認知忽略了一個重要的事實：雖然與大模型交流容易，但要充分利用它們並獲取精準的結果卻是一項挑戰。

例如，簡單地說「請給我畫一朵花」，與詳細描述「你是一個油畫畫家，請參考梵高的畫風，畫一朵盛開的向日葵，背景是一片草原」，這兩條提示詞的差異將導致AI輸出截然不同的作品。前者可能只是一幅簡單的花的圖案，而後者則更可能產生一幅風格接近梵高、情感豐富的藝術作品。

因此，要想用好AI大模型，首先需要使用者能夠明確自己的需求，並且能夠從多個角度詳細描述這些需求。這與我們通常在搜索引擎上使用一兩個關鍵詞查找信息的習慣不同。在搜索引擎中，我們依賴算法來從海量數據中篩選

出最相關的結果。而在與AI大模型交互時，我們需要更主動地提供信息，幫助模型更準確地理解我們的意圖。

此外，用好大模型還需要用戶具備一定的技術背景和創造力。例如，在藝術創作、數據分析或編程等領域，用戶需要對所使用的工具和方法有一定的了解，才能有效地指導AI大模型生成符合預期的輸出。這就需要用戶不僅能夠提出問題，還要能夠提出正確的問題，這在很大程度上取決於用戶對大模型能力的理解和對任務需求的清晰表達。



▲GPT商店為用戶提供定製聊天機器人。

釐清誤解

AI大模型

非萬能

誤解一 AI大模型 可以信賴

雖然，已經有很多人在工作和學習中開始使用AI大模型幫助自己檢索資料、分析數據和創作內容。然而，細審視這些模型輸出的結果，我們有時會發現它們雖然看起來嚴肅認真，但實際上卻包含着AI編造的信息。如果未經核實就輕信這些內容，我們可能會被誤導，從而接受錯誤的知識或做出錯誤的決策。



▲「大模型型幻覺」是指大模型在面對某些輸入時，產生不準確、不完整或誤導性的輸出。

這種現象在業界被稱為「AI大模型幻覺」。在當前的AI大模型中，幻覺是一個棘手的問題。它的產生有多方面的原因：一方面，可能是因為模型訓練所用的數據不精確或不全面，使得模型在

生成某些特定信息時缺少必要的背景知識。另一方面，大模型的訓練重點在於生成連貫、符合語言規範的文本，而非保證內容的真實性和準確性。此外，AI大模型缺乏人類的「常識」，無法像人類那樣依靠常識來判斷信息的合理性，因此在處理超出其訓練數據範圍的知識時，可能會產生一些反常識的內容。

誤解二 AI大模型 是全能的

我們在看到大模型寫詩、生成圖片甚至生成高質量視頻這樣的精彩應用場景之後，往往會認為AI大模型無所不能。而如果讓AI大模型做精準量化的任務時，其輸出的結果的準確性就大打折扣了。

在今年6月，上海人工智能實驗室研發推出的大模型評測體系——OpenCompass對7個主流AI大模型進行了高考試卷語、數、外的測試，從結果來看，大模型的語文、英語考試水平普遍不錯，但數學不及格。由此可見，在數學方面大模型還有很大的提升空間。數學能力決定其複雜推理相關能力，複雜推理能力是大模型在科研、金融、工業等要求可靠結果的場景落地時需要的關鍵能力。



▲盤古CV大模型。

誤解三 AI大模型 很賺錢

AI大模型的炙手可熱使得國內外的大型企業紛紛投身於這場技術革命，引發了一場「百模大戰」的競爭局面。然而，AI大模型的訓練不僅需要大量的算力支持，還需要依賴高質量的數據資源。目前，全球算力市場主要被海外巨頭所壟斷，這使得企業在獲取穩定算力方面面臨諸多挑戰。同時，構建AI大模型的基礎設施需要巨額投資，通常需要數億甚至數十億人民幣。



▲百度文心一言（ERNIE Bot）。

那麼，我們不禁要問，AI大模型真的很賺錢嗎？AI大模型公司的盈利模式主要分為兩種：面向企業客戶的To B模式和面向個人用戶的To C模式。

在To B模式中，大模型廠商通過向企業客戶提供API調用服務來收取費用，這些服務可能包括按使用時長、按調用量計費，或是提供包含硬件的一站式解決方案。例如，OpenAI就推出了30款左右的大模型供用戶調用，涵蓋了文本、圖像、音頻等多個領域。

另一方面，大模型廠商也在探索To C模式的變現途徑。對於大多數AI公司而言，大模型的大規模商業化最終需要依賴於個人用戶市場。OpenAI的ChatGPT聊天機器人就是一個成功的To C應用案例，它已經成為一個現象級應用，這是目前國內廠商尚

未達到的水平。儘管AI大模型的商業潛力巨大，但在當前階段，它們並不如預期那樣賺錢。以百度為例，2024年第一季度，其雲業務收入中只有6.9%來自大模型及生成式AI相關服務，約為3.24億元人民幣。這些收入相對於百度的資本投入，還遠遠無法收回成本。

而為了爭奪市場份額，國內外各大AI大模型廠商，都相繼推出了新模型並宣布降價，甚至提供免費版本。大模型服務屬於典型的規模經濟，隨著使用人數的增加，企業的研發和算力成本得以分攤。此外，隨着技術的進步，算力成本預計將持續下降，這也為大模型廠商提供了降價的空間。降價策略反映了業界對AI大模型未來前景的樂觀預期。

誤解五 AI大模型可以 持續交流

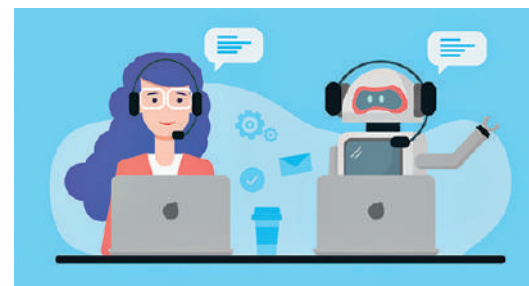
許多人認為，既然這些大模型能夠理解和生成自然語言，那麼它們應該能夠像人類一樣進行持續的、無限制的交流。然而，這種看法忽略了AI大模型在技術和功能上的關鍵限制。

首先，AI大模型在處理對話上下文時存在信息量的限制。例如，當你與某個AI大模型討論一個複雜的項目時，你不能期望它還能記得幾個月前你們交流的細節，除非這些細節在最近的對話中

有被提及。這種局限是由於模型的架構和訓練方式決定的，它們通常只關注最近的對話內容，以便更有效地處理信息。

其次，AI大模型的「短期記憶」特性意味着它們在多輪對話中的能力是有限的。這並不是說AI不能進行多輪對話，而是說它們可能無法在長時間的交流中保持一致性和準確性。例如，如果你與一個AI客服機器人反饋一個問題，它可能在對話的早期階段理解你的需求，但如果對話持續時間過長，它的回答就可能模糊或者跑偏。

此外，AI大模型在處理大量數據時也存在限制。用戶不能期望AI一次性處理數千頁的文檔或大量的數據輸入。這種限制是由於計算資源和算法效率的限制。



▲AI大模型在處理對話上下文時存在信息量的限制。

誤解六 AI大模型+硬件 是救星

在消費電子行業，大模型技術的應用曾備受期待，被許多人視為市場的潛在救星。然而，現實的發展並不如預期那般樂觀。

在AI終端創新的浪潮中，一類備受矚目的創新產品是基於大模型技術打造的新型硬件設備。

在MWC大會上，一家名為Humane的初創公司提出了「手機能做的，AI Pin都能做」的宣言，展示了其產品通過語音對話和激光顯示進行交互的能力。

另一款備



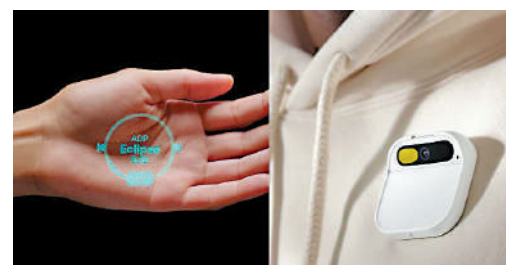
▲AI大模型硬件Rabbit R1。

受矚目的AI大模型硬件Rabbit R1宣稱其搭載了高性能的AI大模型LAM，能夠處理複雜的數據和學習任務，並配備了顯示屏幕，提供更加直觀的交互體驗。

對於那些對傳統智能手機感到審美疲勞的消費者來說，這些新型硬件無疑充滿了吸引力。但是，隨着這些產品逐漸進入市場並被用戶廣泛使用，一些問題也開始顯現。首先，像AI Pin和Rabbit R1在推廣中承諾的AI功能並沒有完全實現。其次，在用戶體驗方面，

這兩款AI硬件在實際應用中遇到了諸多挑戰，包括算力不足導致的延遲、系統不穩定以及缺少常用功能等問題。

隨着退貨率的上升和負面評價的增多，市場對這兩款硬件的熱情迅速冷卻。歸根結底，這些新興硬件雖然捕捉到了消費者對傳統智能手機交互體驗的不滿，但它們提供的產品解決方案還不夠成熟，無法讓消費者獲得更好的體驗以對這樣的產品產生依賴，從而難以轉化為實際的產品價值和持久的商業成功。



▲AI Pin擁有通過語音對話和激光顯示進行交互的能力。

結語

在探索AI大模型的潛力時，我們應當認識到，儘管AI大模型技術帶來了各種可能，但它仍然處於成長和發展的階段。AI大模型並非萬能，它也有自己的一些局限。我們應當以審慎的態度對待它輸出內容。同時，我們也應該意識

到，擁抱AI大模型不能一味地跟風，在大模型領域盲目的投入並不一定能創造更多價值。企業和AI創業者應當回歸理性，找到真正的市場切入點並發展合理的商業模式，這樣才能達到可持續發展的目標。