

# 生成式人工智能： 讀懂AIGC的奧秘

# 生成式人工智能： 讀懂 AIGC 的奧秘

丁磊 著

人從出生開始，就在不斷通過視覺、聽覺、嗅覺、味覺、觸覺等各種方式認識這個世界。我們通過不停地與外界接觸、學習，逐漸長大成人，再通過專業課程的學習在某些方面獲得一技之長從而立足於社會，並試圖改造世界。

「硅基」AI（人工智能）也按照類似的模式成長，但是相比於「碳基」人，它在速度方面極具優勢。AI 經歷了從最初的機器學習到神經網絡，再到 Transformer 模型的發展，2022 年底 ChatGPT 以及 2023 年初 GPT-4 橫空出世，引燃了公眾對生成式 AI 的關注，其中最讓人激動的就是 AI 大模型已經初步具備了人類的通識和邏輯能力——這恰恰是之前的 AI 所缺失的。此前，無論是 AlphaGo 還是 AlphaFold，最多只能稱作其各自領域的「專家」，而 ChatGPT 是通用的。

正如 OpenAI 首席科學家、ChatGPT 背後的技術大佬伊爾亞·蘇茨克維 (Ilya Sutskever) 所說，GPT（生成式預訓練模型）學習的是「世界模型」。他將互聯網文本稱作世界的映射，因

此，將海量互聯網文本作為學習語料的 GPT 學習到的就是整個世界。在我們認識世界的同時，GPT 模型也以驚人的算力，快速地獲取我們數年甚至數十年才能擁有的認知，即將成為一個接近成年人思維水平的「世界模型」。

不僅如此，已具備了「世界模型」能力的 GPT 還能夠生成「萬物」。當然，如蘇茨克維所說，這裡的萬物指的是世界萬物在數字空間的映射，包括文本、圖片、音頻、視頻、劇本、代碼、方案、設計圖等一切和我們生產、生活息息相關的事物。因為 GPT 模型在一定程度上可能已經具備了成年人的通識和邏輯，所以我們只需要拿特定專業領域的數據對其再做訓練（稱為「微調」），它就可以成為獨當一面的專業人才，可能成為藝術家、設計師、程序員、工程師或廣告優化師、供應鏈專家、客服人員等。這也許就是生成式 AI 或者說 AIGC（AI generated content，人工智能生成內容）帶給我們的核心價值。

在 AI 技術大爆炸的今天，生成式 AI 處在高速發展階段，技術和應用領域日新月異，因此我們非常有必要系統地了解生成式 AI。在這樣的背景下，本書將系統介紹生成式 AI 的原理與模型，同時也將對其在行業場景中的應用展開論述，將理論和實際相結合，讓大家從本源上了解 ChatGPT 里程碑式存在的意義。結合作者二十餘年 AI 領域研究與工作的經驗，本書會為讀者指明方向。尤其值得一提的是，本書既在理論上解釋了數字媒體即虛擬世界的生成式 AI，又探討了生成式 AI 如

何服務和賦能實體經濟。在當前的存量經濟時代，通過生成式 AI 重新定義生產力，助力行業更新發展，在存量裡促增長，具有尤為重要的意義。

如圖 0-1 所示，我們用圖表明本書所覆蓋的知識領域：X 軸是生成式大模型的維度，對應的是第二章「AIGC 的底層邏輯」，我們將了解「用甚麼去生成」；Y 軸是數字媒體形態的維度，對應的是第三章「功能分析：AIGC 能生成甚麼內容？」，我們將了解生成式 AI 能「生成甚麼」；Z 軸是行業職能的維度，對應的是第四章「商業落地：AIGC 的產業應用與前景」，

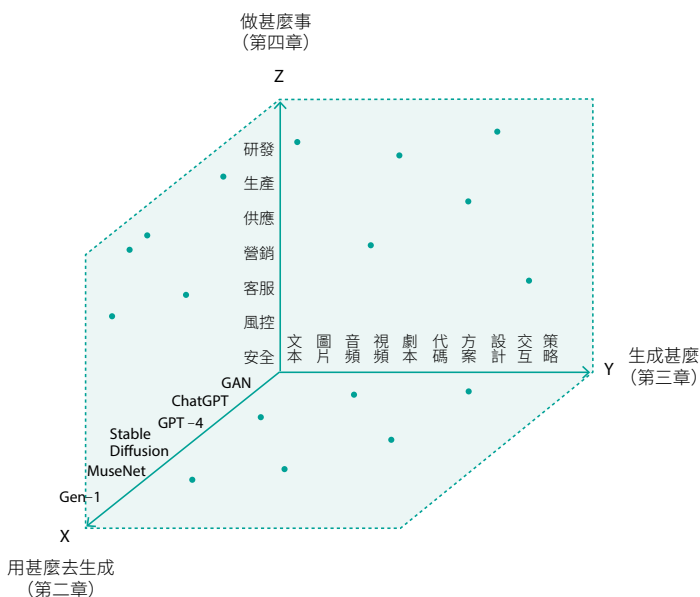


圖 0-1 本書所覆蓋的知識領域

我們將了解用生成式 AI 可以「做甚麼事」。三個軸所形成的空間裡的每個點都有其特定含義，例如：通過 GPT-4 模型生成代碼用在生產上，通過 Stable Diffusion 模型生成圖片用在營銷上。除了這三章，第一章會帶領讀者初識生成式 AI，第五章則探討生成式 AI 是否會取代大量的工作崗位，以及我們應該如何主動應對。

希望任何一個不想在生成式 AI 時代落伍的人，在閱讀本書之後，都能理解生成式 AI 的底層邏輯和實際應用，也希望本書對他們的工作和生活有所助益。「萬物皆可生成」的時代已經來臨，理解 AI、訓練 AI、使用 AI，甚至和 AI 一起工作，對每個人來說或將無法避免。未來已來，讓我們一起出發！

## 第 1 章 初識生成式人工智能

縱觀 AI 產業版圖	003
聚焦 AIGC：內容皆可生成	018
生成式 AI 的核心價值	031
里程碑式的存在 —— ChatGPT	041

## 第 2 章 AIGC 的底層邏輯

生成式模型基礎	063
Transformer 和 ChatGPT 模型	077
Diffusion 模型	090

## 第 3 章 功能分析：AIGC 能生成甚麼內容？

生成文字：新聞、報告、代碼都可一鍵生成	105
描繪圖像：分辨率、清晰度、真實性與藝術性	115

音頻製作：精準還原、實時合成	130
影視創作：海量場景任你選	140
互動娛樂：遊戲中的生成式 AI 革命	151

## 第 4 章 商業落地：AIGC 的產業應用與前景

研發設計：設計能力樣樣俱全	167
生產製造：「L4 級別」的智能控制	183
供應鏈管理：庫存計劃可自動編程	196
市場營銷：營銷文案不再發愁	207
客戶服務：貼心服務打動客戶	218

## 第 5 章 主動還是被動？決勝 AIGC

展望未來：AIGC 是否是新一輪的技術革命？	233
智能並非萬能：AIGC 的優勢與瓶頸	243
我們的工作機會還在嗎？	255
後記	268
參考資料	271





[第 1 章]

# 初識生成式人工智能

現象級產品 ChatGPT 的橫空出世帶火了 AIGC 和它背後的生成式 AI，讓不少人對使用 AI 工具躍躍欲試。在本章，我們會一起進入 AI 的產業國度，從決策式 AI 躍遷至生成式 AI，對比這兩種人工智能模型的異同，深度挖掘它們的應用場景和商業潛力，同時還會探討「內容皆可生成」的生成式 AI 及其核心價值。在概覽生成式 AI 後，我們會把視角轉向具體工具，深入解析大眾已經熟知的 ChatGPT，了解這項「奇妙工具」背後的原理和發展歷程。若你對生成式 AI 一知半解，可以在本章的內容中初步認識它，也能了解到近期最值得關注的生成式 AI 工具。

# 縱觀 AI 產業版圖

如果要選出 2023 年最熱的幾個話題，ChatGPT 一定榜上有名。2023 年初，ChatGPT 席捲全球並成為流量熱點，人們都在前赴後繼地挖掘 ChatGPT 的各種潛能，探討其未來發展趨勢，甚至是與人類的關係。作為「人工智能家族」的大熱應用，以 ChatGPT 為首的各大人工智能應用開始被越來越多的人關注，也引發了人們的更多思考。

人工智能技術被稱為當代三大尖端技術之一，近年來在人們生活中的「存在感」也越來越強，這都是產業飛速發展的結果。想要清晰地了解以 ChatGPT 為代表的新興智能技術，完整地認識人工智能，我們可以先從其產業版圖的發展和現狀入手。

其實，人工智能的發展、傳播和被接受是經過了一段漫長的寒冬的。十多年前，它還只是一個不被人看好的小眾領域，但是現在，它卻已經成了街頭巷尾的熱點談資，幾乎任何事情都可以和人工智能聯繫在一起。短短十多年間，世界發生了天翻地覆的變化，新數據不斷湧現，各種問題層出不窮，直到現在，人工智能的春天才算是真的到來了，各個領域都急需人工

智能的幫助。這也是為甚麼人工智能的行業應用範圍如此廣闊，人工智能市場更是如一塊一望無際的遼闊土地，有待進一步開發。如圖 1-1，這是一份人工智能的行業應用版圖，不同的行業領域（零售、金融、醫療和教育等）與不同的職能方向（營銷、風控和安全等）共同構成了一個人工智能應用矩陣，對於每個行業中的相關職能，人工智能都可以找到應用場景，例如在零售行業的供應鏈、營銷、客服等方面以及金融行業的研發、營銷、客服、風控等方面都已經有人工智能落地實踐（圖中綠色表示）。但是，現在的人工智能只填充了廣闊的行業領域中的一部分，還有更多沒嘗試和拓展的行業以及職能中的應用場景。

行業 職能	零售	金融	醫療	教育	製造	能源
研發						
生產						
供應鏈						
營銷						
客服						
風控						
安全						

圖 1-1 人工智能的行業應用版圖

從產業的視角來看，人工智能包括基礎層、技術層和應用層。其中，基礎層是人工智能產業的基礎，為人工智能提供數據及算力支撐；技術層是人工智能產業的核心，主要包括各類模型和算法的研發和升級；應用層則是人工智能面向特定場景需求而形成的軟硬件產品或解決方案。那麼，人工智能的產業規模發展至何種程度了呢？英國德勤 (Deloitte) 的報告中預測，世界的人工智能產業規模會從 2017 年的 6900 億美元增長至 2025 年的 64000 億美元，2017–2025 年的複合增長率將達到 32.10%，整體呈現出飛速攀升的趨勢。另外，人工智能近幾年成了各個行業在進行投資的熱門選擇。人工智能完全稱得上是風頭正勁，受萬人追捧，為經濟帶來了十分顯著的增量。

在產業應用上，人工智能發展到今天，我們能看到其在各個行業都有用武之地：製造業、零售業、金融業、醫療衛生行業……它在一定程度上改變了組織的運轉方式，使其可以更快更好地解決遇到的問題，並壓低各類成本。站在消費者的角度，人工智能的出現也為廣大的用戶群體帶來了更多的選擇。總的來說，人工智能可以看作一塊已開始被打磨的原石，露出了它璀璨的一角，它在推動世界經濟發展的同時，也將深層次地改變人類的生活。為了進一步了解 AI 產業版圖，下面我們從兩種不同的 AI——決策式 AI 和生成式 AI 談起。

## 決策式 AI 和生成式 AI

人工智能可從不同的維度進行劃分。如果按其模型來劃分（人工智能是由模型支撐的）可以分為決策式 AI 和生成式 AI。

決策式 AI（也被稱作判別式 AI）學習數據中的條件概率分佈，即一個樣本歸屬於特定類別的概率，再對新的場景進行判斷、分析和預測。決策式 AI 有幾個主要的應用領域：人臉識別、推薦系統、風控系統、其他智能決策系統、機器人、自動駕駛。例如在人臉識別領域，決策式 AI 對實時獲取的人臉圖像進行特徵信息提取，再與人臉庫中的特徵數據匹配，從而實現人臉識別。再例如，決策式 AI 可以通過學習電商平台上海量用戶的消費行為數據，制訂最合適的推薦方案，盡可能提升平台交易量。

生成式 AI 則學習數據中的聯合概率分佈，即數據中多個變量組成的向量的概率分佈，對已有的數據進行總結歸納，並在此基礎上使用深度學習技術等，創作模仿式、縫合式的內容，相當於自動生成全新的內容。生成式 AI 可生成的內容形式十分多樣，包括文本、圖片、音頻和視頻等。例如，我們輸入一段小說情節的簡單描述，生成式 AI 便可以幫我們生成一篇完整的小說內容；再例如，生成式 AI 可以生成人物照片，而照片中的人物在現實世界中是完全不存在的。如圖 1-2，它展示的是國外一個網站生成的「不存在的人」的照片。



圖 1-2 「不存在」的人

圖片來源：<https://generated.photos/faces>

總的來說，不管是哪種類型的模型，它的基礎邏輯是一致的：AI 模型從本質上來說是一個函數，要想找到函數準確的表達式，只靠邏輯是難以推導的，這個函數其實是被訓練出來的。我們通過餵給機器已有的數據，讓機器從數據中尋找最符合數據規律的函數。所以當有新的數據需要進行預測或生成時，機器就能夠通過這個函數，預測或生成新數據所對應的結果。

決策式 AI 和生成式 AI 作為 AI 模型的兩個主要分支，顧名思義，在諸多方面都有相異之處。

從宏觀角度來看，決策式 AI 是一種用於決策的技術，它利用機器學習、深度學習和計算機視覺等技術來處理專業領域

的問題，並幫助企業和組織優化決策。而生成式 AI 則是一種用於自動生成新內容的 AI 技術，它可以使用語言模型、圖像模型和深度學習等技術，自動生成新的文本、圖片、音頻和視頻內容。因此，決策式 AI 可以說是在對人類的決策過程進行模仿，但生成式 AI 就聚焦在創作新內容上。

而從微觀上看，這兩類技術的區別就更加明晰了，我們就從技術路徑、成熟程度、應用方向這三個角度來挖掘其深層次的不同（表 1-1）。

表 1-1 決策式 AI 和生成式 AI 的對比

	決策式 AI	生成式 AI
技術路徑	將數據分類打標籤，從而區分不同類別的數據，例如區分貓和狗的圖片	分析歸納已有數據後生成新的內容，例如生成逼真的狗的圖片
成熟程度	底層技術相對成熟，在各領域有廣泛的商業應用	2014 年開始迅速發展，近期呈指數級爆發，並且出現多個現象級應用
應用方向	人臉識別、推薦系統、風控系統、機器人、自動駕駛等	內容創作、人機交互、產品設計等

從技術路徑來看，決策式 AI 的主要工作是對已有數據「打標籤」，對不同類別的數據做區別，最簡單的例子如區分貓和狗、草莓和蘋果等，幹的主要是「判斷是不是」和「區分是這個還是那個」的活兒。生成式 AI 就不一樣了，它會在歸納分析已有的數據後，再「創作」出新的內容，如在看了很多狗



的圖片後，生成式 AI 再創作出一隻新的狗的圖片，實現「舉一反三」。

從成熟程度看，決策式 AI 的應用更為成熟，已然在互聯網、零售、金融、製造等行業展開應用，極大地提升了企業的工作效率。而生成式 AI 的「年歲更小」，2014 年至今發展迅猛，堪稱指數級的爆發，已在文本和圖片生成等應用內落地。

從應用方向來看，決策式 AI 在人臉識別、推薦系統、風控系統、機器人、自動駕駛中都已經有成熟的應用，非常貼合日常生活。生成式 AI 則在內容創作、人機交互、產品設計等領域展現出巨大潛力。

我們來舉一些生活中的例子，以更深入地了解兩者在日常生活中的應用。喜歡購物的讀者都知道，你在購買某一類產品後，購物平台會自動給你呈現諸多同類或相關商品。這件事的背後就是，電商平台會根據用戶常看的商品，剖析用戶和商品的關聯，從而有針對性地為用戶推薦內容，而這項功能就應用了決策式 AI 技術。從 2003 年開始，亞馬遜就將此技術應用到了電商領域，推薦的商品精準地匹配用戶需求，可以極大地降低用戶的搜索次數，並因此增加產品的銷售額。由此你可能會發現，平台似乎比你更清楚你需要甚麼，自然而然，自己的消費額也跟着上去了。平台也憑藉這個功能，讓更多用戶心甘情願地掏了腰包，來獲取更廣闊的商業價值。

根據行業數據統計，在亞馬遜的收入中約有 40% 來自個

性化推薦系統，而推薦系統每年能給網飛 (Netflix) 帶來 10 億美元以上的產值。除了電商平台，新聞、音樂、視頻等平台，也會利用個性化推薦系統為用戶推薦內容，在剖析用戶的長期興趣和短期興趣後，將精細化內容推送給用戶，並可以通過對用戶的停留、觀看時間、點讚、收藏等行為特徵的實時分析，精準刻畫出用戶畫像，減少人工運營的介入，顯著提升用戶黏性，這已將人工智能的價值凸顯無遺。在自動駕駛領域，AI 可進行智能分析、識別路況，滲透率穩步提升。自動駕駛汽車可以藉助決策式 AI 技術，分析判別各種路況，對多種物體進行識別與跟蹤，提升行車安全。無須人工干預的自動駕駛汽車雖然現在並不成熟，但隨着技術的迭代升級，有望獲得更大的市場潛力。

對於生成式 AI，ChatGPT 的出現讓我們對其有了衝擊式的關注和理解。因生成式 AI 功能強大、應用範圍廣泛，文字、圖片、音視頻內容相關的從業者在面對「強大助手」上線時，也會感覺到焦慮，恐被其取代。從可能性來講，它可以進行文字生成語音、圖像智能編輯、視頻智能剪輯、文字續寫或糾錯等十分多樣的工作，讓大家擺脫機械勞動，把時間花在創意性工作上，給文字作者、翻譯人員、插畫師、視頻剪輯師等帶來極大的支持。不僅如此，生成式 AI 還能勝任部分由設計師、程序員甚至專業工程師從事的設計與編程類工作，在提升工作效率的同時讓這些專業人士更能發揮所長，減少在初級工作上

的時間投入。與此同時，生成式 AI 對於從業人員的素質和技能，也提出了新的要求。總的來說，決策式 AI 和生成式 AI 均可以幫助用戶推進部分工作，如事件決策、創作內容等。可以說，人工智能的合理利用有助於提升客戶體驗，幫助企業降本增效，並抓住新的商業機會。

如前文所述，數據和模型分屬人工智能產業的基礎層和技術層，無論是決策式 AI 還是生成式 AI 的應用都離不開數據和模型，下面我們進一步了解「大數據」和「大模型」是如何重塑人工智能版圖的。

## 從大數據到大模型

無論是決策式 AI 還是生成式 AI，以其現在的功能和潛力，都能為人類做很多工作，未來甚至有點萬能，那麼這麼萬能的技術，是怎麼被「訓練」出來的呢？這就要說到大數據了，決策式 AI 和生成式 AI，其實都離不開用大量數據來訓練模型。

對於大數據，大眾已經比較熟悉。顧名思義，大數據指的是海量的數據，但大數據並沒有看上去這麼簡單，它還有多樣性和高速增長的特性。圖 1-3 展示了從 2017 年到 2025 年全球數據總量的增長趨勢及預測情況。收集、存儲、處理和分析各種形式和來源的大數據，可以幫助企業和組織迅速獲得有價值的信息，並做出正確的決策，它還可以用於商業活動的改

善，如此能提升工作效率，降低工作成本，並推動企業實現更大的增長。就如人類通過經歷各類事件來積累經驗一般，在人工智能領域，我們通過大量的數據來訓練模型。

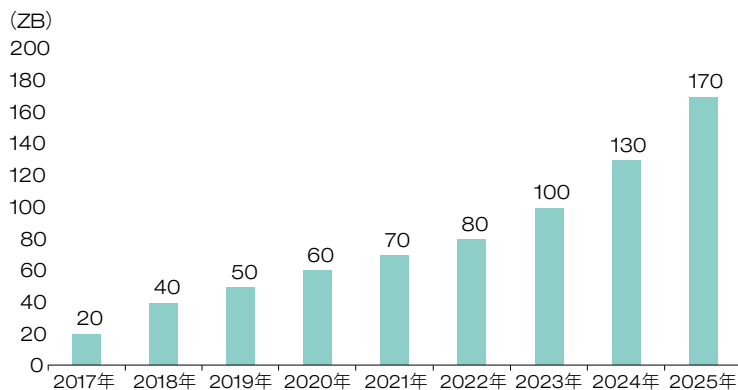


圖 1-3 2017-2025 年全球數據總量增長趨勢及預測情況

數據來源：國際數據公司發佈的白皮書《數據時代 2025》

而隨着深度學習的落地和發展，模型本身所需的存儲空間在近年有了顯著增長，最初的 GPT 就有 1.17 億個參數，ChatGPT 有 1750 億個參數，最新的 GPT-4 參數數量更多，有報道稱可能達到 1T（即 10000 億），但 OpenAI 公司其實並沒有公佈具體的參數數量，這些擁有海量參數的模型都被稱為「大模型」。如圖 1-4，它展示了大模型參數數量變化趨勢。這裡我們提到了深度學習，這是一種受人腦的生物神經網絡機制啟發，並模仿人腦來解釋、處理數據的機器學習技術，它能

自動對數據進行特徵提取、識別、決策和生成。你可能覺得這個詞有點耳熟，其實它大規模地應用於自然語言處理（NLP）、計算機視覺、機器翻譯等領域。深度學習的出現，為很多領域的工作帶來了前所未有的精度和效率。人工智能行業也因深度學習收穫了前所未有的發展速度，整個人工智能領域的發展都曾被它帶動。

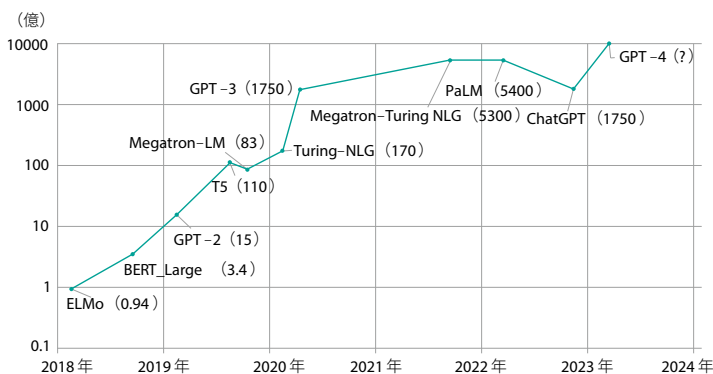


圖 1-4 大模型參數數量變化趨勢

大模型能分析處理海量的數據，在解決問題上取得更好的效果。本書的「主角」生成式 AI 就是大模型的產物。近年來，大模型在越來越多的行業和消費類應用中嶄露頭角，原因主要是它能夠迅速有效地處理海量的數據，幫助個人和企業提升效率。大模型與人工智能技術相輔相成，隨着人工智能技術的發展，大模型也會持續發展進步。另外，生活中日益普及的 5G

網絡和彈性計算等基礎設施，也會給大模型的發展創造更多可能性，使其成為不可或缺的內容生成工具。

## 生成式 AI 市場格局

2021 年，高德納諮詢公司 (Gartner) 就曾預測，至 2023 年將有 20% 的內容被生成式 AI 創建，至 2025 年生成式 AI 產生的數據將佔所有數據的 10% (2021 年不到 1%)。2022 年 9 月，紅杉資本官網發佈的文章《生成式 AI：充滿創造力的新世界》預測，生成式 AI 將產生數萬億美元的經濟價值。據預測，2025 年，國內生成式 AI 應用規模有望突破 2000 億元，國內傳媒領域應用空間超 1000 億元。而且，生成式 AI「八面玲瓏」，它的應用場景十分廣泛，目前不僅應用於文本、圖片、音視頻、遊戲等數字媒體，還可以應用於製造業、建築業等實體行業。

在文本生成方面，生成式 AI 可以通過語言模型、神經網絡和深度學習技術，快速創建大量有助於改善客戶體驗的內容，如新聞資訊、劇本、營銷文本、智能客服等。其中作為經典應用的 AI 生成營銷文本、智能客服等都已在許多行業廣泛地應用；AI 生成新聞資訊和劇本等功能大家也可以期待一下，或許以後結合了 ChatGPT 等突破性的模型，文字性工作真的能依靠它變得輕鬆不少。

在圖片生成方面，生成式 AI 可以通過計算機視覺來分析圖片，生成營銷素材、設計方案和藝術作品等，幫助節省人力

成本和時間。另外，生成式 AI 還能在音頻生成、視頻生成和跨模態生成領域大展拳腳。

在音頻生成方面，生成式 AI 可以幫助使用者更好地分析、編輯和生成音頻文件，從而幫助創作出優秀的音頻作品。例如，克隆真人的語音、文本生成特定語音、作曲編曲等，生成式 AI 都能代替人類去做，並均已經廣泛應用於市場。

視頻生成也是生成式 AI 的重要應用，它可以幫助使用者生成高質量的視頻，如檢測和刪除特定片段、跟蹤剪輯、生成特效、合成視頻等。另外，大火的 AI 數字人也是它的「拿手絕活」。在李安執導的《雙子殺手》中，工作人員就用 AI 創造了一個數字人物小克。韋史密夫 (Will Smith) 在數字技術的幫助下同時出演了 50 歲特工亨利和 23 歲特工小克，該片實現了真實明星「年輕版」的數字化製作。

在跨模態生成中，生成式 AI 能夠根據文字生成創意圖片、根據圖片生成視頻、根據文字生成視頻，或根據圖片或視頻生成文字。對想像力豐富的朋友，或者影視行業從業者來說，這稱得上是「工作神器」。圖 1-5 就是一個根據文字「panda in a space suit」(穿着宇航服的熊貓) 生成圖片的例子。在遊戲方面，生成式 AI 可以用於遊戲開發，實現自動化的遊戲設計，同時能夠實現更好的遊戲體驗，如人工智能 NPC (非玩家控制角色) 等，說不定以後你玩的遊戲就有人工智能的深度參與。



圖 1-5 根據「panda in a space suit」生成的圖片

生成式 AI 不光在這些數字經濟領域廣泛應用，在實體領域的潛力也非常大，如在建築業等巨型垂直實體領域中，生成式 AI 所生成內容就不再僅局限於圖片和文字，而是進入了信息形式更為豐富的 3D（三維）設計領域。例如構建數字建築模型時，生成式 AI 能幫助建築師們產出 3D 建築模型，讓他們更好地理解項目。建築師們能夠使用 AI 圖像生成應用來豐富建築設計方案的細節，假如建築師們向應用中輸入較為初級的建築設計方案，AI 就能夠在初級設計的基礎上，繼續產出較為細化的設計方案，以此來深化設計。建築師們還可以隨手繪製一個潦草的建築場景線圖，讓人工智能來生成對應的建築實景圖。我們可以想像，隨着手繪信息的增加，生成式 AI 輸出的實景圖也越來越穩定。圖 1-6 所示的就是利用 AI 圖像生成工具生成的建築設計圖。





圖 1-6 由 AI 圖像生成工具生成的建築設計圖

圖片來源：<https://stability.ai/blog/stablediffusion2-1-release7-dec-2022>

技術的浪潮層疊翻湧，人工智能已成為人類社會衝向未來世界的戰艦，產業前景十分廣闊。生成式 AI 更是一個突破性的產業發展方向，它不僅能給數字媒體和虛擬空間帶來價值，還能促進實體行業的發展，在提升行業效率的同時優化原有的流程，創造出新的價值增長點，可以說是實體行業升級不可多得的機遇。

# 聚焦 AIGC：內容皆可生成

當下，世人的目光被 ChatGPT、GPT-4 這些 AIGC 深深吸引。而在清楚地認識這些新事物之前，我們需要梳理一下它們的歷史脈絡，其實在數年硝煙瀰漫的「內容大戰」中，我們已經悄然經歷了多種內容形式的迭代：PGC (professional generated content)、UGC (user generated content) 和 AIUGC (artificially intelligent UGC)。PGC 即「專業生產內容」，主要指具備專業背景的內容生產者所創造的內容；UGC 則為「用戶生產內容」，其內容的源頭更偏大眾化，人人都可作為用戶進行內容生產；AIUGC 則為人工智能與 UGC 的結合，人工智能參與到了用戶創作內容的過程中。如今，在三度更迭之後，AIGC 正式來襲。與 PGC、UGC 和 AIUGC 不同的是，在 AIGC 的概念中，「無生命的」人工智能成了完全的內容源頭，「無生命主體」成了為人類創作內容的生產者。人工智能在人類社會的應用又取得了顛覆性的突破，透出了不同於以往的炫目光彩，吸引着人們不斷探索。如圖 1-7，從 PGC、UGC、AIUGC 到 AIGC，所對應的內容數量呈逐漸增加的趨勢。

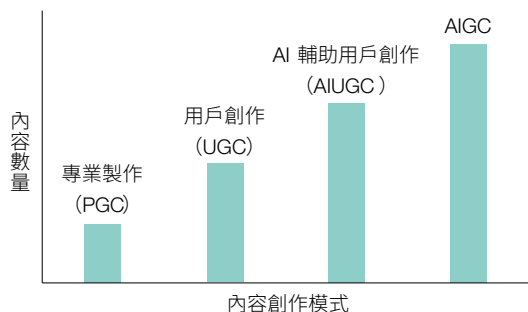


圖 1-7 內容創作的四個發展階段

從字面上看，AIGC 就是利用人工智能自動生成內容的生產方式，它可以在生成式 AI 模型、訓練數據等的基礎上，生成文本、圖片、音頻、視頻、代碼等多樣化內容，這種快速的內容生產方式給市場注入了令人興奮的新鮮血液。AIGC 的出現，使得各行各業都受益，使得人們的生活更加便捷。但在發展得如火如荼的同時，AIGC 又引發了我們對其更深層次的思考。

AIGC 開啟了新一輪的內容生產革命，它在多樣性、質量、效率三個方面推動了內容生產大踏步前進。AIGC 的出現，既可以滿足消費型內容亟待擴充的需求，也可以快速產出多樣化的內容形態，迎合多種細分場景，以 AI 作者的身份助力商業化浪潮的翻湧。或許我們現在正在看的某張圖片、某段視頻就是 AI 作者的「作品」，而我們卻不自知。

下面我們就圍繞 AIGC，對文本、圖片、視頻等不同的內

容形式展開論述，看看 AIGC 究竟是如何「長袖善舞」，在各個內容形式中發揮作用的。

## 文本生成

AIGC 生成文本目前主要被應用於新聞的撰寫、給定格式的撰寫、風格改寫以及聊天對話，GPT 是主流的文本生成模型之一。

GPT 的「學名」是生成式預訓練模型 (generative pre-training transformer)，這是一種用來分析和預測語言的人工智能模型，它可以幫助我們進行自然語言處理，例如機器翻譯、自動文摘和快速問答。GPT 的厲害之處是，它可以在文本中自動學習概念性內容，並自動預測下一段內容。也就是說，它可以根據上下文記住概念，並能夠在短時間內直接輸出相關內容。

GPT 背後的基礎模型是一種新型的機器學習技術，它可以幫助我們分析大量的自然語言數據。它背靠一個大型神經網絡，通過在已有文本庫中找到有關自然語言的規律來學習。GPT 無須人工設計特定的自然語言處理系統，可以根據已有文本，自動生成語法正確、內容相關的文本。有這樣一個「神器」，很多內容就可以藉助它的力量來完成了！

GPT 的發展目前經歷了 GPT-1、GPT-2、GPT-3、GPT-3.5 和 GPT-4 幾個階段。對於 GPT-1 模型，我們可以

這麼理解：先使用海量沒有進行標註的語料，預訓練出一個語言模型，而後對語言模型進行微調，使之應用於特定的語言任務中。GPT-2 則在 GPT-1 的基礎上進行了多任務的訓練，使用了更大的數據集，提升了語言處理能力。GPT-3 則在訓練的參數量、訓練數據和訓練費用上都高於前兩者，能完成更加複雜的任務。

OpenAI 推出的 ChatGPT 是 GPT-3.5 的延伸，這是一款聊天機器人程序，能通過學習和理解人類的語言與人類對話，還能實現視頻腳本撰寫、營銷文案寫作、文本翻譯、代碼編寫等功能。例如它在代碼理解和編寫方面的能力，就在程序員圈引起了廣泛的關注：它可以看懂你輸入的代碼片段，幫你解讀其中的含義，甚至可以根據你的要求幫你編寫一段完整的代碼。如此強大的能力，幾乎顛覆了人們的認知，並引發了諸多關於「AI 替代人類」的相關討論。

而當人們還沉浸在 ChatGPT 帶來的無限遐想中時，就在 2023 年 3 月，OpenAI 推出了史上最強大的模型——GPT-4。它在文學、醫學、法律、數學、物理和程序設計等不同領域表現出很高的熟練程度，各方面能力已全面超越 ChatGPT。不僅如此，它還能夠將多個領域的概念和技能統一起來，並能夠理解一些複雜概念。OpenAI 在官網上演示了這樣一個示例：向 GPT-4 展示一張圖片（圖 1-8），並詢問圖中有甚麼有趣的地方。而 GPT-4 的回答相當精妙：這幅圖的有趣之處在於，

把一個大而過時的 VGA（視頻圖形陣列）接口插入一個小而現代化的智能手機充電端口，這是十分荒謬的。GPT-4 儼然擁有一個普通人的正常思維。



圖 1-8 一張「有趣」的圖片

圖片來源：<https://openai.com/research/gpt-4>

要想深刻了解 AI 技術的發展，我們就需要到推動主體——企業中去。主打 AI 文本生成的 Jasper 公司位於美國加利福尼亞州，通過其產品的文本生成功能，用戶可以輕鬆完成生成 Instagram（照片牆）標題，編寫 TikTok（抖音國際版）視頻腳本、廣告營銷文本、電子郵件內容等略顯燒腦的重複性工作。AI 文本生成功能一經推出，便給社交媒體、跨境電商、視頻製作等多個新興行業帶來了巨大的顛覆力量。

除了 Jasper 以外，OpenAI 更是近期談論 AI 時不可繞過

的熱門企業。OpenAI 是一家 AI 研究公司，成立於 2015 年，它旨在促進人工智能的安全可控發展。我們前文中提到的 GPT 這類卓越的自然語言處理模型，就是 OpenAI 首創推出的，這也使得 OpenAI 一躍成為 AI 行業的佼佼者。在自己進行技術創新之外，OpenAI 也通過與微軟等行業巨頭的合作，將 AI 的應用推向更高的層次，這也將為人類的日常生活帶來豐富的可能性。

由於 GPT 有基於英文語料庫且不開源的局限，國內的技術人員也在探索我們自有的自然語言處理模型。2020 年 11 月中旬，北京智源人工智能研究院和清華大學研究團隊就合作推出了中文預訓練模型——清源 CPM (Chinese Pretrained Models)，我們也有了自主研發的類似於 GPT 的模型。

## 圖片生成

你是否嘗試過用 AI 生成圖片呢？談到 AI 生成圖片，你第一時間又會想到哪個程序呢？你所使用的程序，很可能背後是由 Diffusion (擴散) 模型來進行技術支撐的。Diffusion 模型是一種新興的 AI 技術，它的靈感來源於物理學中的擴散現象：通過對圖片不斷加入噪聲來生成一張模糊的圖片，這個過程類似於墨水滴入水池的擴散過程；再通過深度神經網絡學習模糊的圖片並還原成原始圖片的逆擴散過程，實現生成圖片的功能。目前，Diffusion 模型在視覺藝術和設計相關領域

非常受歡迎。

Stability AI 是一家全球領先的 AI 研究型企業，致力於開發前沿的人工智能模型。2022 年，由該公司與另外兩家初創公司共同研發的 Stable Diffusion 模型發佈，可以真正實現「一秒出圖」，這個「一秒」不是誇張的代指，而是真正的事實。這就意味着你可以藉助 AI，實現自己瑰麗的夢境，復原宏大的想像，也可以為自己的小說配上極富幻想感的插圖，不論它們有多超現實，你都可以通過 AI 把它們呈現在大家的眼前，讓想像不再孤獨。

2022 年，AI 繪圖突然大熱，隨着 DALL·E 2、Stable Diffusion、Midjourney 等圖像生成領域現象級應用的紛紛興起，AI 繪畫就像一陣旋風，首先在國外引起了不小的風浪，社交平台上出現了大量的 AI 繪畫相關嘗試和討論。很快這場旋風就從國外颳到國內，引起了國內用戶的廣泛關注。這些應用到底有着怎樣驚奇的功能，而它們背後又有哪些企業在推動這場 AI 繪畫「旋風」呢？

首先我們把目光放到 Midjourney 身上（圖 1-9），這是由同名研究實驗室開發的 AI 繪畫工具。在 AI 繪畫領域，Midjourney 降低了藝術繪畫創作的門檻，用戶只需要輸入文字描述，計算機就會自動生成一張作品。Midjourney 採用了深度學習模型，能夠自動為用戶生成高質量的繪畫作品，包括素描、油畫等，讓用戶的使用更加方便。