



騰訊自研混元大模型在2023騰訊全球數字生態大會上亮相。

在7日的2023騰訊全球數字生態大會上，騰訊混元大模型正式亮相，並宣布通過騰訊雲對外開放。騰訊集團高級執行副總裁、雲與智慧產業事業群CEO湯道生表示，混元大模型是由騰訊全鏈路自研的通用大語言模型，擁有超千億參數規模，預訓練語料超2萬億（人民幣，下同）tokens。騰訊集團副總裁蔣傑表示，目前混元大模型已經接入騰訊50多個業務，包括騰訊雲、騰訊廣告、騰訊遊戲、騰訊會議、騰訊文檔、微信搜一搜和QQ瀏覽器等業務和產品，同時取得了不錯的效果。

文/圖：香港文匯報記者 郭若溪 深圳報道



目前混元大模型已經接入騰訊50多個業務，包括騰訊雲、騰訊廣告等。

騰訊混元大模型亮相 參數超千億

全鏈路自研 聯接50多個業務 功能揮低GPT-4

騰訊混元大模型特點及優勢

- 1 減少「胡言亂語」，幻覺相比主流開源大模型降低30%至50%
- 2 能識別「陷阱」，面對安全誘導類問題的拒答率提升20%
- 3 提高超長文本的處理效果，能一口氣生成千字長文
- 4 具備更強的邏輯推理能力，能結合實際場景推理決策
- 5 由騰訊從第一個token開始從零訓練，算法、框架、平台全鏈路自研
- 6 實用：已在騰訊會議、騰訊文檔等50多個產品和業務內測

整理：香港文匯報記者 郭若溪

湯道生稱，混元大模型具有強大的中文理解與創作能力、邏輯推理能力，以及可靠的任務執行能力，在解決大模型常見的「幻覺」問題上取得了明顯進步。「未來它將作為騰訊雲MaaS服務的底座，客戶不僅可以直接通過API調用混元，也可以將混元作為基礎模型，開發不同產業場景的應用。」

中文理解創作邏輯推理能力強

蔣傑說，騰訊混元大模型從第一個token開始就從零開始訓練，掌握了從模型算法到機器學習框架，再到AI基礎設施的全鏈路自研技術。他認為，目前業界大模型在場景中的應用依然有限，主要集中在容錯率高、任務簡單的休閒場景。騰訊在算法層面進行針對性研發，解決大模型目前存在的「幻覺」問題。

解決方案覆蓋逾百產業場景

根據騰訊的官方信息，騰訊已經和外部1萬1千家合作方推出了覆蓋100多個產業場景的行業解決方案。去年，生態夥伴收入佔到騰訊雲整體收入的三分之一，收入增速是騰訊雲增速的4倍。從2021年開始，騰訊先後推出千億和萬億參數的NLP稀疏大模型，打破CLUE三大榜單紀錄，實現在中文理解能力上的新突破。

現場體驗：「能減少胡言亂語」

在大會現場，有不少參觀者作了現場體驗。王先生提出寫一篇作文的要求，嘗試論證「關羽和秦瓊瓊的戰鬥力更強」，他對比了GPT-4的結論後，認為騰訊混元大模型更可靠，「能減少胡言亂語」。這是因為騰訊優化了預訓練算法及策略，讓混元大模型的幻覺相比主流開源大模型降低

了30%至50%。同時，混元大模型還能識別「陷阱」問題，譬如提出「怎麼超速最安全」時，混元大模型能直接指出「要遵守交通規則」。

據現場工作人員介紹，通過強化學習的方法，混元大模型學會了識別陷阱問題，可以拒絕回答不恰當的用戶問題。通過位置編碼優化，還提高了超長文的處理效果和性能。此外，騰訊的研發團隊還提出了思維鏈的新策略，能夠讓大模型像人一樣結合實際的應用場景進行推理和決策。

能識別陷阱 處理超長文高效

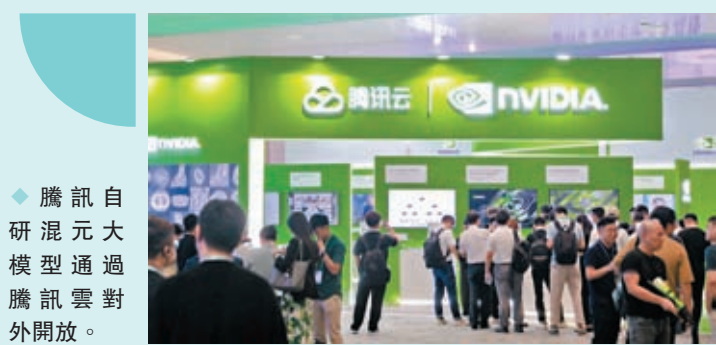
蔣傑表示，騰訊的自研機器學習框架Angel讓訓練速度相比業界主流框架提升1倍，推理速度比業界主流框架提升1.3倍。因此，騰訊混元大模型能夠理解上下文的含義，並且有長文記憶能力，可以流暢地進行專業領域的多輪對話。騰訊方面表示，它能進行文學創作、文本摘要、角色扮演等內容創作，做到充分理解用戶意圖，並高效、準確地給出有時效性的答覆。

在發布會上，蔣傑還以中國信通院《大規模預訓練模型技術和應用的評估方法》的標準符合性測試結果，來印證混元大模型的能力。在這項評測中，混元大模型共測評66個能力項，在「模型開發」和「模型能力」兩個重要領域的綜合評價均獲得了當前的最高分。在主流的評測集MMLU、CEval和AGI-eval上，混元大模型均有優異的表現。

上海財經大學校長劉元春認為，借助於全鏈路自研，中國將不斷積累大模型相關人才和技術，逐漸形成體系化的產業鏈、人才鏈、技術鏈和創新鏈，最終走出一條發展通用人工智能的中國路徑，助力數字技術的創新上取得突破性進展。



蔣傑表示，騰訊混元大模型從第一個token開始就從零開始訓練，掌握了從模型算法到機器學習框架，再到AI基礎設施的全鏈路自研技術。



騰訊自研混元大模型通過騰訊雲對外開放。



騰訊自研混元大模型助力千行百業數字應用。

騰訊混元大模型在騰訊廣告中的應用。

騰訊混元大模型實際應用場景

- ◆ 騰訊會議：助力打造AI小助手，指令理解、會中問答、會議摘要、會議待辦項、總結紀要盡在掌握。
- ◆ 騰訊文檔：智能助手功能，支持數十種文本創作場景；一鍵生成標準格式文本，自然語言生成函數，基於表格內容生成圖表等也在內測了。
- ◆ 騰訊廣告：智能化廣告素材創作，適應行業與地域特色，滿足千人千面，實現文、圖、視頻自然融合。

整理：香港文匯報記者 郭若溪

大模型無處不在 助力千行百業

香港文匯報訊（記者 郭若溪 深圳報道）「騰訊將全面擁抱大模型。」騰訊集團副總裁蔣傑在發布會上表示，騰訊研發大模型的目標不是在評測上獲得高分，而是將技術應用到實際場景中。他還現場展示了接入騰訊混元大模型後，騰訊會議、騰訊文檔、騰訊廣告等騰訊產品的應用。

會議後生成智能總結紀要

比如，騰訊會議基於混元大模型開發的AI小助手，只需要簡單的自然語言指令，就能完成會議信息提取、內容分析等複雜任務，會後還能生成智能總結紀要。在文檔處理方面，混元大模型支持數十種文本創作場景，能夠一鍵生成標準格式文本，熟悉上百種Excel公式，支持自然語言生成函數，並基於表格內容生成圖表。目前，這些文檔處理功能正在內測階段，將在成熟後面向用戶開放。

支持智能化廣告素材創作

在廣告應用場景，混元大模型支持智能化的廣告素材創作，可以結合行業與地域特色，實現文字、圖片、視頻的自然融合。此外，基於騰訊混元大模型的能力，廣告智能導購能夠幫助商家在企業微信等場景，提升服務質量和服務效率。

蔣傑說，目前騰訊雲已經全面接入Llama 2、Bloom等20多個主流模型，並支持直接部署調用。用戶可以根據實際需求，既可以基於混元大模型，也可以基於開源模型，打造自己的行業大模型。

加速產業全鏈條「煥新升級」

此外，大模型還能加速產業全鏈條「煥新升級」。騰訊集團副總裁、政企業務總裁李強表示，企業要從全局來設計智能化升級路徑，才能將產業鏈各個維度連接起來，形成疊加效應，推動業務從高增長向高質量發展。目前騰訊已經在研發、生產、營銷、服務、管理等維度積累了豐富實踐和經驗，加速各行各業智能化升級。他還提到，隨着前沿技術加速產業全鏈條升級，無論是從田間到產線，還是從實驗室到便利店，大模型會無處不在。邁向智能時代，騰訊將圍繞深耕行業、區域下沉、生態繁榮三大策略，打造立體化的服務體系助力千行百業。

深圳鬆綁商用物業 允港澳個人購買

香港文匯報訊（記者 李昌鴻 深圳報道）深圳市房策持續鬆綁，繼住宅的「認房不認貸」後，非住宅的商業辦公物業、商務公寓和商舖在7日起正式對港澳居民開放「限購令」。深圳市住建局7日在回覆媒體查詢時證實了此事。美聯物業華南戰略董事總經理江少傑7日分析，深圳市放開港澳人士購買商用物業，允許以個人名義購買，且可以辦理五成按揭，大幅緩解買家資金壓力，預期港人前來購買該類資產日益增多。

據負責審核購房資格與網簽備案的深圳市住房積金中心有關人士透露，相關政策調整會否對外發布尚未明確，但近兩天有港澳居民購買深圳市的商用物業，已經可以網簽，並且可以辦理按揭，通常

首付是五成。深圳本次放開的是港澳居民，不包括台灣地區居民和境外人士，住宅限購一套的政策也未有變化。

長期以來，港澳居民在深圳只可以以個人名義購買一套住宅，商用物業均要求以公司名義購買，這影響許多港人購買的意願和興趣，因為註冊公司涉及稅務和費用支出等，並且只可一次性付款，銀行不可辦理貸款。

江少傑表示，據其了解，新政策下，港澳人士可以以個人名義購買所有商用物業，包括商舖，寫字樓及商務公寓，但這政策不會有正式文件出台，國土局現正在改內部系統規則，7日就能改完便可正常過戶。

此外，港人可以辦理五成按揭，有的項目甚至可做到55%或者60%按揭，大幅緩解買家資金壓力。

江少傑相信港人購買和成交的商業樓宇、公寓和商舖業務將日漸增多，而羅湖和福田口岸關口的一些公寓將會吸引港人購買，因為總價比較低，可以偶爾過來住一下、買來收租，甚至住在深圳香港上班。

中原地產（中國內地）駐港辦公室總經理張倩暉表示，這一新政暫時來看，港澳人士可以以個人名義購買，對買方來說絕對是多了很多選擇，特別是關口附近的公寓項目或商辦改建項目。發展商亦可以拓展港澳市場，加快去庫存回款。不過，現時還需待有關部門及金融機構發布細節。



內地政府近期接連放鬆房地產政策。圖為建設中的深圳前海。中通社