

技術詞彙表

以下為本文件中使用的與我們及／或我們業務有關的若干詞彙。因此，這些詞彙及其涵義未必與這些詞彙的標準行業涵義或用法一致。

廣告發佈者	指	根據《中華人民共和國廣告法》，廣告發佈者是指為廣告主或者廣告主委託的廣告經營者發佈廣告的自然人、法人或者其他組織
AGI	指	通用人工智能，即在人工智能中體現人類通用認知能力，能夠在面對任何不熟悉而人類卻能夠處理的任務時找到執行任務的解決方案
AI或人工智能	指	人工智能，指研究、開發用於模擬、延伸和擴展人的智能的理論、方法、技術及應用系統的一門先進技術科學
AI agent	指	一種能夠感知其環境、作出決策並採取行動以實現特定目標的自主AI系統或軟件實體，通常具有推理及決策能力
AI模型	指	一種依賴訓練數據來識別模式並作出預測或決策的算法
AI產品	指	可擴展的AI工具或APP，將AI算法與硬件、軟件及雲基礎設施相結合，提供自動化、數據驅動的功能
原生AI	指	基本以AI為核心組件進行設計
AlexNet	指	一種由Alex Krizhevsky、Ilya Sutskever和Geoffrey Hinton提出的深度卷積神經網絡，在2012年ImageNet競賽中，圖像識別準確率遠遠領先其他競爭對手，標誌著深度學習時代的開啟

技術詞彙表

AIGC	指	人工智能生成內容
算法	指	通過執行一連串指定行動（專門由計算機執行）解決問題的程序或公式
API	指	APP編程接口（Application Programming Interface，簡稱API）是應用系統對外開放功能的調用接口，為用戶或開發人員提供嵌入自身軟件或硬件的接口服務
單個付費用戶平均收入	指	單個付費用戶平均收入按當期總訂單支付金額除以同期付費用戶數量計算。訂單支付金額來源於第三方應用商店的訂單記錄，而付費用戶數量是在此期間完成訂單支付的用戶總數，經跨支付渠道去重並剔除僅參與試用的用戶數
複合年增長率	指	複合年增長率，等於 $(\text{期末價值} / \text{期初價值})^{(1/\text{年數})} - 1$
每千次展示成本或 CPM	指	一種按每千次展示支付廣告費的定價模式
客戶關係管理系統		Customer Relationship Management，簡稱CRM，是利用信息科學技術，實現市場營銷、銷售、服務等活動自動化分析和管理的，使企業能更高效地為客戶提供滿意、周到的服務，以提高客戶滿意度、忠誠度為目的的一種信息管理系統
數據規模化定律	指	AI中的一種原理，表明隨著數據量、計算能力和模型大小的增加，模型性能會以可預測的方式得到改善，是三大AI規模化定律之一

技術詞彙表

深度學習	指	深度學習是機器學習的一種方法，通過建立模擬人腦進行分析學習的神經網絡，模仿人腦的機制來解釋複雜數據，例如圖像，聲音和文本等
DIKI	指	數據－信息－知識－智能，一種用於構建從原始數據到可操作智能的進展過程的內部研發框架，用於在數據驅動的決策中創造價值鏈
生成對抗網絡	指	一種通過兩個神經網絡對抗性訓練來提高預測準確性的機器學習模型
GPU	指	圖形處理器，一種用於快速操作和修改內存以加速圖像生成的專用電子電路
幻覺 (AI領域)	指	AI模型生成錯誤的、無意義的或與輸入數據無關的輸出的情況，通常是大語言模型面臨的一項挑戰
ICDAR	指	國際文檔分析與識別會議
ICFHR	指	國際手寫識別前沿會議
ICPR	指	國際模式識別大會
ImageNet	指	一個廣泛用於訓練及評估計算機視覺模型的圖像數據庫。2010年至2017年，ImageNet競賽每年舉行一次，使用ImageNet數據集的一個子集來衡量提交模型的性能
Inference-Time Training	指	一種在推理階段 (模型使用時) 進行的AI模型訓練方法，允許持續學習及自適應，而無需從頭開始重新訓練
知識圖譜	指	使用圖譜結構數據模型存儲及組織信息的知識庫

技術詞彙表

大語言模型(大模型)	指	由數十億參數的人工神經網絡構成，通過自監督學習或半監督學習在大量未標記文本上進行訓練的計算機化語言模型
機器學習	指	一種實現人工智能的方法，使計算機獲取新的知識或技能，重新組織已有的知識結構使之不斷改善自身的性能
MAU	指	月活躍用戶數，即於一個歷月內錄得設備活躍度(包括APP啟動、訪問、瀏覽及點擊等行為)的用戶總數。該指標包括手機APP及小程序月活躍用戶數
月活付費轉化率	指	指定期間內付費用戶與月活躍用戶數量相除得到的比率
毫秒級實時處理	指	系統在極低延遲(通常在毫秒級)內實時處理及分析數據的能力，從而實現即時決策
多模態	指	能夠同時處理及整合多類型數據(如文本、圖像、視頻)以增強理解及功能的技術或系統
兆瓦時	指	兆瓦時，一種能量單位，等於每小時產生1,000千瓦電力，用於衡量電力輸出
自然語言處理(NLP)	指	系統在工作過程中積累數據，並通過深度學習模型及學習技術從大量的原始、非結構化及未標記的文本和語音數據集中提取越來越準確的含義的過程
神經網絡處理器(NPU)	指	神經網絡處理器，一種專門用於加速機器學習算法的微處理器，通常通過預測模型操作
光學字符識別(OCR)	指	光學字符識別，將文本圖像轉換為機器編碼文本的電子或機械過程

技術詞彙表

付費用戶	指	於報告期間至少完成一筆成功交易的用戶，無論交易次數如何，每個期間每名用戶只計算一次
後訓練技術	指	用於在初始訓練後優化及改進AI模型的技術
預訓練技術	指	在對特定任務進行微調之前，在大型通用數據集上對AI模型進行初始訓練，以構建廣泛的知識基礎
問答	指	問題與答案
研發	指	研究與開發
強化學習	指	基於獎勵期望行為及／或懲罰不期望行為的機器學習訓練方法
SDK	指	軟件開發工具包，為特定的軟件包創建APP的軟件開發工具集
文本智能	指	專注於將文本數據的處理、理解和操作作為其智能核心組成部分的AI系統，通常與其他數據類型相結合
Transformer	指	一種深度學習模型，其特點是採用自注意力機制，對輸入數據的每一部分賦予不同的權重