

「自駕模型需百萬級數據 背後是數十萬小時人工」

內企研發平台識別數據 AI訓練提速

經濟透視

「一段10秒的道路視頻，包含車輛、行人、紅綠燈、車道線、障礙物等數十類目標，標註員需逐幀識別、精準框選，完成一段視頻往往需要數小時；而訓練一個基礎自動駕駛模型，需要百萬級標註數據，背後是數十萬小時的人工投入。」寧波博登智能科技有限公司創始人兼CEO趙捷以自動駕駛場景為例，說明有多少智能，背後就有多少人工。

大公報記者 俞晞



博登智能位於安徽省馬鞍山市的具身機器人創新中心正在進行素材採集。

寧波博登智能科技創始人兼CEO趙捷。



業內人士指出，傳統標註模式的「三重困境」：一是效率低下，面對上游海量的數據需求，只能靠「堆人」來解決；二是質量參差，人工標註受疲勞、主觀判斷影響，誤差率普遍在10%左右，而AI模型對數據精度的要求近乎苛刻，「垃圾數據」只會訓練出「幻覺模型」；三是成本高企，伴隨着人力成本的逐年上漲，標註成本已然成為了AI模型研發不可忽視的沉重負擔。

「數據決定行業發展上限」

2019年從德國人工智能公司裸辭後回國創辦公司的趙捷，他要做的是真實世界AI訓練基礎設施生態的「構建者」：通過博登智能自主研發的「全自動化數據引擎」，先用雲端大模型對原始真實世界交互記錄進行70%至80%的自動化預處理；再由專家工程師對「邊緣場景」「疑難樣本」進行校驗、修正；最終形成高精度數據。

為了讓記者更直觀地感受「流水線」的便捷，趙捷推開了技術研發中心的大門，工程師們正在對BASE6.0平台進行測試。屏幕上，一段複雜的3D點雲數據被快速拆解，200餘個自動化標註小模型同步運行，精準識別車輛、行人、障礙物等「交通參與者」的三維坐標；這些坐標的下一站，是10餘個相關領域垂直大模型的實時校驗、修正，直至被送往最後的人工「關卡」。

「新一代平台的核心優勢，是『全模態、自動化、高精度、低成本』。」研發中心技術負責人向記者介紹說，BASE6.0版本融合智能體和多模態大模型，支持2D圖像、3D點雲、4D動態數據、文本、語音、視頻等所有主流AI數據類型，適配多種場景，「以前一個訓練數據工程團隊需要100人，現在靠平台+少量專家，20人就能完成同等工作量，而且質量更高、成本更低。」

傳統純人手操作易有偏差

面對傳統標註模式的「三重困境」，趙捷從來不認為數據工作就是簡單的重複勞動，真實世界的可執行經驗才是人工智能（AI）模型的靈魂素材，滿足自動駕駛、智慧政務、智能客服等多個行業的數據需求。

走進公司的具身智能實驗室，數百台顯示器同時跳動着像素級畫面：機械臂抓取水杯的動態軌跡、人形機器人跨越障礙的關節數據、廚房場景裏「視覺-觸覺-力反饋」的多模態交互，每幀畫面都在被精準拆解、標註、校準。「這裏每一秒產生的真實數據，都是訓練具身智能模型的核心素材……真實世界的訓練體系才是決定行業發展上限的關鍵變量。」

AI數據產業相關國家政策

2022年1月

國務院在《「十四五」數字經濟發展規劃》中首次明確提出培育壯大數據服務產業，將數據標註、清洗、脫敏等環節納入國家戰略

2025年1月

國家發展改革委等四部門聯合發布《關於促進數據標註產業高質量發展的實施意見》，明確提出到2027年產業年均複合增長率超過20%，培育一批科技型數據標註企業，建設特色鮮明的數據標註基地

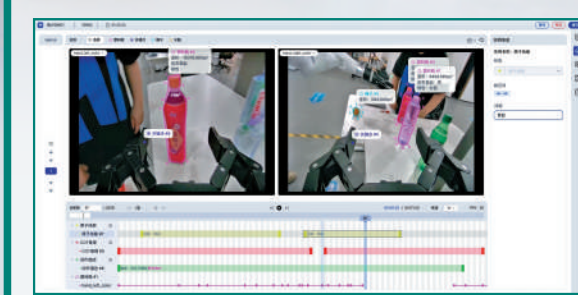
2024年4月

國家數據局等十七部門印發《「數據要素×」三年行動計劃（2024-2026年）》，強調健全數據採集、管理等標準體系，協同推進行業標準制定

2025年12月

國家數據局發布《關於加強數據科技創新的實施意見》，明確提出要加快攻關具身智能等技術創新發展的高質量數據集構建，謀劃布局一批數據領域部級重點實驗室

大公報記者 俞晞製表

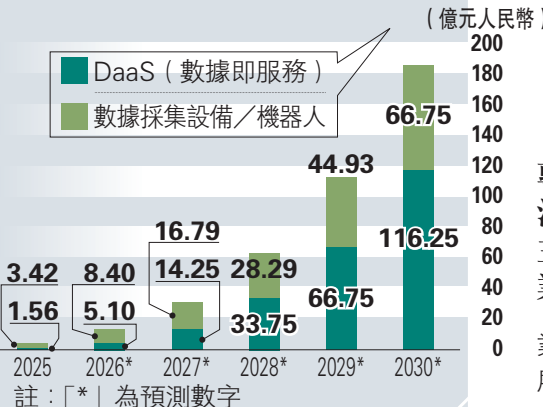


博登智能自研平台正在進行原子技能的標註。

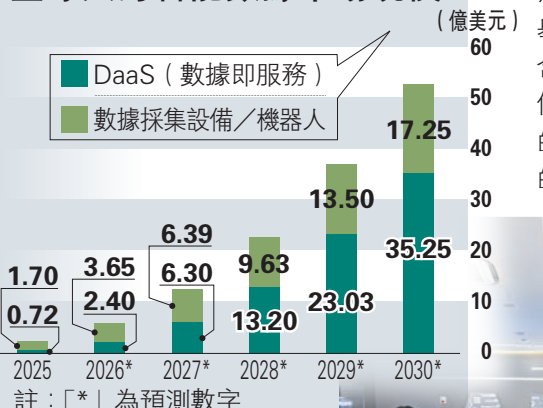


博登智能自研標註平台手部骨骼關鍵點標註。

中國具身智能數據市場規模



全球具身智能數據市場規模



具身智能領域將成為AI產業的下一個萬億級金礦。

具身智能發展迅速 數據分析炙手可熱

專注主業

「行業的共識是，具身智能將成為AI產業的下一個萬億級金礦，既然這樣，博登智能想成為那個「賣鏟子的人」。趙捷2023年將公司業務從自動駕駛、大模型數據服務，拓展至具身智能數據服務的決定。「傳統的AI訓練素材是靜態的、單模態的、事後標註的，而具身智能的訓練素材是動態的、多模態的、因果關聯的。」趙捷舉例說，一個「拿起杯子」的動作，包含了視覺、觸覺、運動學、音頻等多種信號，並且每一個信號之間都存在嚴格的時間因果關係，需要給它「餵」大量的數據才能習得。

對於從原來業務拓展至具身智能數據服務的轉變，趙捷表示，認真研究過中國具身智能企業的創始團隊後發現，超過半數都是曾經從事或參與過自動駕駛技術的研發。兩者在包括關節、傳感器等硬件，還是底層算法，都是通用的。所以目前新業務的很多技術人員就是從原本的兩個賽道「平移」過來的。他補充，在新的賽道上，博登智能沒有選擇跟風研發機器人本體，而是聚焦核心業務，自主研發了具身智能BRIC Robo數據採集平台與Blink數據管理平台，這樣「我們僅用了一年多時間，就將具身智能訓練素材的工程化效率最高提升700%、質檢效率提升300%、綜合交付準確率超99%。」

與自駕訓練邏輯相若

趙捷提到，主流大語言模型的預訓練數據量有數十萬億token，自動駕駛需要百億小時數據才能取得今天的效果。具身智能要從「想做」到「會做」，同樣需要大量的數據來學習訓練。但據不完全統計，全球範圍內研發端對高質量數據的需求量約為120萬小時，中間有着巨大的缺口。

「場景訓練與迭代學習 決定AI上限」

發展邊界

目前，博登智能已在寧波、馬鞍山、湖州等城市建設三大具身機器人創新中心，總佔地面積超3萬平米，覆蓋上百類生活化、工業化真實場景。「未來決定AI能力上限的，不僅取決於模型架構本身，更依託於真實場景訓練、物理化驗證與持續性迭代學習的完整體系。」在趙捷看來，當企業可以合法、合規地採集數據後，它的整個業務才能實現邏輯閉環。

趙捷指出，在自動駕駛的道路訓練素材裏，有一類是國家嚴格管控的，即導航數據。對於AI數據基礎設施企業而言，這是一條紅線。具身智能的數據幾乎都不是敏感數據，

在具身智能創新中心，技術員正在進行超場景數據的採集。



「普通消費者買台機器都能自己採集，企業能做的就更多了。」

讓機器理解現實世界任務

趙捷認為，過去十年，人工智能最大的挑戰是讓機器理解數字世界；未來十年，它將變成讓機器理解並執行現實世界任務。公司的目標是「讓機器人從

實訓基地「畢業」後，「能上崗、能幹活」。

「我們在基地裏部署了超500台多型號實體機器人，數千台自研Ego採集設備，它們在真實場景中反覆實操，通過多部位攝像頭採集數據，再經過審核、工程化清洗、結構化處理等標準化程序，可實現年產50萬小時真機訓練數據、百萬小時級Ego場景數據的穩定產能，搭建起國內領先的規模化真實場景訓練網絡。」

事實上，受益於新業務的拓展，博登智能今年的營收預期是去年的三倍有餘。就在記者採訪趙捷的一個多小時裏，他接到了三家具身智能企業的諮詢電話。

競爭轉向數據 科企加速布局

發展迅速

隨着具身智能的核心瓶頸逐漸指向數據，互聯網大廠的入場方式也從技術探索轉向更直接的基礎設施布局。「2018年到2026年，具身智能產業規模擴大將近5倍。」在京東集團技術委員會主席、京東雲總裁曹鵬看來，當前具身智能硬件進展迅速，但模型能力仍明顯滯後，「特別是在機器人大腦模型上面還有很大的短板，本質原因是缺少數據。」

「訓練具備泛化能力的具身模型至少需要千萬小時量級數據，而目前行業可獲得的數據規模仍停留在幾十萬小時。」基於這一判斷，京東於今年4月推出了穿戴式超高清採集終端JoyEgoCam，可在物流、零售、醫療、家庭等多種場景下「即戴即採」，並聲稱兩年內採集超過1000萬小時真實場景第一視角數據，疊加百

萬小時機器人本體數據。

向百萬小時級數據邁進

就在京東發布採集終端同一天，智元推出數據服務平台，小米披露機器人與工廠數據積累。這些動作背後，都指向同一趨勢——具身智能競爭正在進入數據基礎設施階段。「今年行業對數據需求正在快速上升，可以稱之為具身數據的元年。」曹鵬坦言，過去頭部企業擁有兩三萬條遙控數據已屬規模較大，而當前行業開始向百萬小時級數據邁進。

「我們最近剛剛完成了由鼎暉百孚、清新資本、魯信創投、深產投等多家知名機構聯合參投的數億元A+輪和A++輪系列融資。」趙捷向記者透露，「從博登智能的融資頻率上，也能窺見市場對於「真實世界AI數據基礎設施」這一新藍海的長遠發展潛力的看好。」