

端傳媒11周年 大陸 人工智能

教會機器認識蘋果之後，縣城「寶媽」被 AI 拋棄

對數據公司來講，被家庭責任留在本地的女性曾是一羣理想的工人。但當 AI 變得更強，它沒有繼續停留在縣城。



永和愛豆科技曾經的項目辦公室。圖：作者提供

在中國人工智能產業擴張的早期，一部分最基礎的數據標註工作，曾被有意識地送往中西部欠發達縣域。

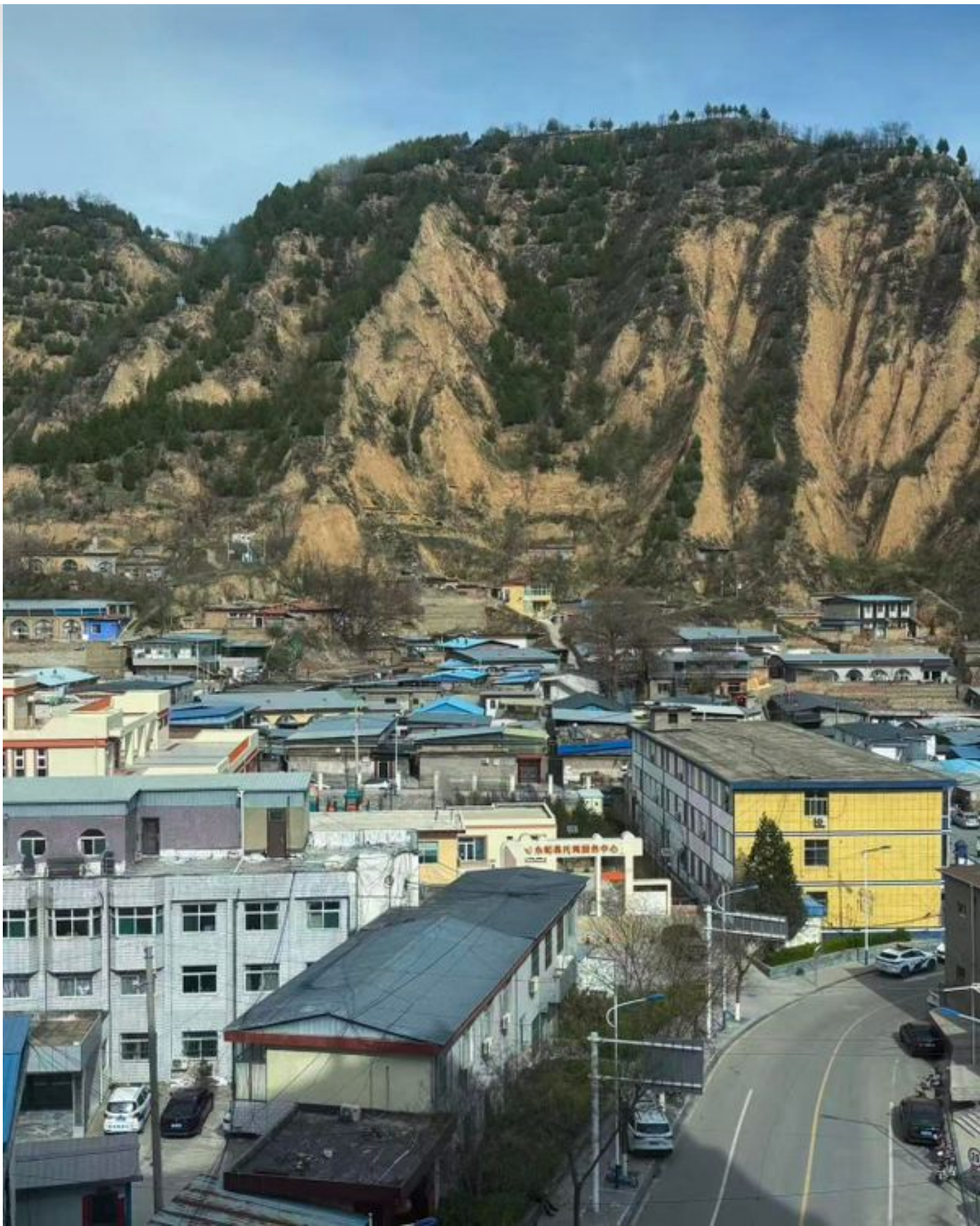
2019 年前後，數據標註進入數字扶貧和女性就業的敘事。一批由科技公司、公益基金會和地方政府共同推動的標註基地，陸續落地陝西、山西、甘肅、貴州、寧夏等地。截至 2023 年底，這一計劃已建立 17 個縣域數字就業中心，累計就業人數約 5800 人，其中近七成女性。

這些項目找到了一羣特殊的勞動力，「寶媽」。對許多被家庭責任留在本地的女性，AI 標註是一份少見的工作——可以坐在辦公室，週末雙休，按件計酬。

對數據公司來說，這也是一羣「理想的工人」。早期標註技術門檻不高，標註公司需要的是大量穩定、能夠長時間坐在電腦前重複操作的人。年輕人會離開縣城，男性更多外

×





中國永和縣城。圖：作者提供

教機器人認識蘋果

數據標註員崗位在永和的第一次招聘在 2020 年，由一家名爲「愛豆科技」的數據標註企業發起。這是一家縣屬國有企業，由國家衛健委、永和縣政府、螞蟻集團等共同扶持成立。它承接科技公司的數據訂單，再在當地招募和培訓標註員。

×



[點擊吃蛋糕→](#)

間生意不好，很快不再經營。

同一時期，中國的人工智能、自動駕駛等產業快速發展。2020 年 2 月，數據標註員作爲「人工智能訓練師」的一個工種，被正式納入國家職業分類目錄。人工智能產業對訓練數據、標註人員的需求相應增長。

迎着這股風潮，一條數據標註流水線也出現在永和縣城：主幹道邊的一棟四層小樓裏，十幾名標註員各自坐在電腦前，盯着屏幕，反覆移動和點擊鼠標。文字、圖片、音頻和視頻等通過網絡傳送到她們面前，再由她們一點點轉換成機器可以使用的數據。

和許多其他基地一樣，在永和，這條流水線的運作和本地女性深度綁定在一起。招聘雖沒有限定性別，但首批入職永和愛豆科技的 19 名員工中，長期留下來的幾乎都是像李霞一樣，需要兼顧家庭的本地女性。

她們經歷並不完全相同，有人此前很少有正式工作的機會，也有人已經在外地或縣城工作多年。她們大多有孩子，年齡從幼兒園到中學不等，因而常被稱爲「寶媽」。

在當地，男性更傾向於進入天然氣行業，雖然工作有一定危險，但收入也會更高；年輕人則更多選擇進城。而真正能長期坐在辦公室裏不斷做標註的，正是那些孩子還小、自己也走不開的女性。後來，基地一度只剩下一名男員工，他在幾乎全是女性的工作環境裏也沒有待太久。

在許多「寶媽」看來，數據標註首先是一份比飯店、早餐店更體面的工作。辦公室夏天有空調、冬天有暖氣，朝九晚五，也有周末和法定節假日。「不像飯店裏一樣又髒又累，」李霞說，「孩子放假我也能休息。」

黃土高原上的永和有着典型的西北風貌，少雨乾旱，地質破碎，約 1212 平方公里的土地被兩千五百多條溝壑切開。常住人口不足五萬人，曾是山西省的深度貧困縣之一。對於本地女性，大部分工作機會集中在飯店、超市、服裝店等服務行業。「能拿到手的錢也就 2000（元）左右。」李霞表示。

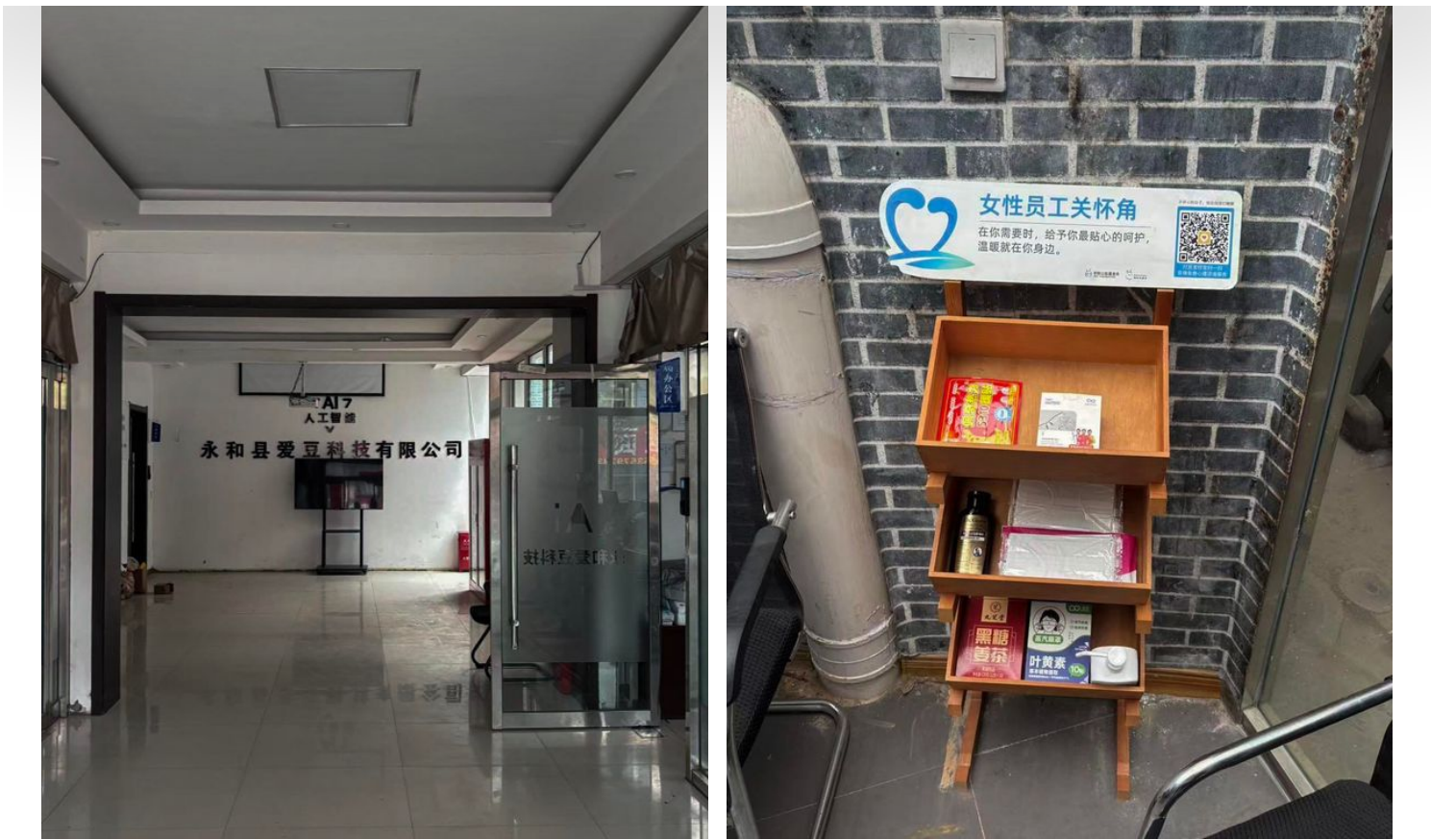
相比之下，訂單充足時，數據標註普通員工月收入有兩三千元；熟練員工如果願意加班，可以拿到四千元以上。

2021 年至 2023 年前後，是許多永和的標註女工們心中的「黃金階段」。

那幾年任務多，計件單價也高，有人月收入接近 8000 元，個別月份甚至可以過萬元。黃丹丹說，當時的標註任務簡單，一人手裏可能同時有三四個任務同時進行。工資按件計算，因此，做得快，收入就會明顯高過縣城裏許多的服務業崗位。最熱鬧的時候，基

×





(左)永和☐豆科技。(右)永和☐豆科技辦公室一角。圖：作者提供

「寶媽」的數據流水線

多數的基礎項目並不難上手。簽訂保密協議後，女工們通常接受兩三個小時的培訓就可以開工。不過，具體操作中真正需要適應的，是☐配合 AI 產業快速變化的需求而制定的嚴苛規則和生產秩序。

在李霞參與的一個項目裏，員工被分成幾個小組，每組有一個做事快的人被選☐小組長。組長會將每個組員的標註量、正確率記錄在表格裏，上傳給項目的管理方。

數量和準確率都被列入考覈。平常標註工作遵循多勞多得的計算方式，但臨近任務交付日期時，總量會進一步拆分☐個人績效指標，有時單天任務量就能達到三百或五百條。業務方還會通過測試標準流程和錄屏的方式，測算一名員工一個小時應該完成多少，制定合格標準。

一度，員工還可以在平臺上看到同一項目裏其他人的進度。「別人都做三千個了，我才做兩千個，我得加把勁。」李霞時不時產生這樣的壓力。但速度也並非越快越好，做得慢會被詢問效率，明顯快於其他人，也可能被檢查是否因☐趕量影響了質量。

準確率則是另一套標準，根據項目不同，通常要求達到八成或九成。未達標的員工需要說明原因，再由組長提交給項目方。李霞記得，每週二，工作羣裏就會出現不合格者的

×



近兩年出現的與大模型相關的任務，則比遊戲項目還要複雜。

「一直在培訓、一直在學習，還要一直記筆記。」李霞說。大模型任務對理解能力的要求更高。比如，面對一張古代人物圖片，標註員不能只是框出人物，還要描述左邊的人穿什麼衣服、右邊是什麼角色，再按照規則補充儘可能多的細節。

在衆多項目中，最特殊的一種是「安全項目」。☑️提高互聯網平臺內容審覈效率，AI需要學習識別垃圾信息，而在此之前，需要有人先將它們識別出來。

2022 年到 2024 年前後，永和愛豆科技大約有 50 名員工都參與過大安全項目。很長一段時間裏，這也是公司最大的業務來源之一。

這要求女工長時間直面色情、暴力和血腥畫面。即使培訓時已經提前「打了預防針」，剛開始接觸時，很多人仍會產生明顯的不適。李霞參與這一項目的時間長達三年，她回憶，起初大家都會感到尷尬、噁心。遇到特別恐怖的圖片，有人會先跳過去，請膽子更大的同事處理。另一家同樣承接該項目的數據標註基地，甚至出現了員工直接在工位上嘔吐的情況。

由於內容的特殊性，大安全項目對數據保密性要求很高，因此也要設立特別的工作場地。☑️達到項目方要求，永和愛豆科技曾在二樓辦公室加設攝像頭，便於進行監控。辦公室門口也設立了更高的木板，防止外來人員隨意進入。參與這一項目的標註女工要在工作期間，把手機放在辦公室外的櫃子中。

安全項目帶來穩定的訂單，工作量也更大。一批任務可能包含三四十萬條數據，需要連續處理一兩個月，每人每天需要查看數千張圖片。永和愛豆科技經理張麗表示，公司會提前告訴員工，如果實在覺得噁心，可以暫時換到別的項目緩解一下。

但大部分人都選擇適應。「你不做，就沒有任務。」張麗說，「所以大家都做。最後也就習慣了。」

在她看來，女工們更看重回報。「工作肯定是有苦有累的，最重要的是能掙到錢。」剛入職時，很多沒有工作經驗的縣城寶媽完全不適應公司管理，連考勤和請假都需要適應。但久而久之也接受了。「大家大部分都是高中畢業，還有初中都沒有畢業的，在其他地方能拿這麼高工資嗎？根本拿不到。」

在這份和未來智能直接掛鉤的工作裏，流水線依然遵循傳統工廠的計件邏輯。系統嚴謹地衡量着工作量、正確率、速度、任務類型等種種指標，它們被不斷加權，最終彙總成工資裏的數字。

×





永和愛豆科技所處大樓。圖：作者提供

遷移的工作，無法遷移的母親

在永和之前，AI 標註已經出現在黃土高原的另一些縣城裏。

2019 年前後，數據標註被放進「數字扶貧」和「縣域就業」的框架中，縣城開始接入了遠方的人工智能產業鏈。這是一個無需大型廠房，在電腦上就可以操作的工作。

陝西清澗是較早的樣本之一。2019 年 10 月，「AI 豆計劃」以社會企業模式落地陝西清澗：由國家衛健委、縣政府等多方支持，在本地建立公司，並以國企方式運行。

光明網曾將這一模式描述為科技企業、公益基金會與地方脫貧和就業需求結合的結果。這些項目也被視為「數字扶貧」和「鄉村女性就業」的成功樣板，受到多家官方媒體報道，在各省多地複製。

永和基地就是從陝西清澗孵化出來的分支。這個由本地政府和螞蟻基金會牽頭建立的扶貧項目，一方面為縣城創造就業，另一方面也為人工智能產業提供了數據標註的勞動力。

永和基地的訂單大多都來自螞蟻集團內部，或由螞蟻集團作為中間樞紐，從高德、阿里雲、蜂鳥配送、網商銀行等企業引入。據報道，經標註的數據被用於各類不同的自動化

×



[點擊吃蛋糕→](#)

早期許多任務需要人工從頭做起，報酬也更高。多家媒體曾報道，最基礎的 2D 拉框，2017 年前後，單價能達到一毛多，甚至有公司開出五毛的價格。手速快的人一天干十幾個小時就能掙到五六百元。

但隨着機器預標註能力提高，簡單標註任務的價格一路下滑。到 2023 年前後，單價已經被壓到幾分錢。給自動售貨機打標的項目，過去，標註員需要從零開始，在售貨機圖片裏把一件件商品框出來並標明類別。如今，AI 已經可以自動進行第一輪「預標註」，標註員只需要再次覈對是否正確。

實際上，AI 標註產業的市場並沒有縮小，反而在擴大。數據機構艾瑞諮詢測算，2023 年中國 AI 基礎數據服務市場規模 45 億元，到 2028 年預計將增長到 170 億元。但新增需求正流向更復雜的項目管理，並對工人的行業知識和專業程度提出要求。2023 年，需求方自建團隊已經佔據主要市場，中小服務商只剩下 5.7% 的份額。

最先被吞噬的正是曾經最適合縣城寶媽的工作。

黃丹丹提到，大模型任務從 2023 年前後開始出現，後來難度越來越高，普通任務也許大專可以做，但稍微複雜一點就要求本科。與早期視覺標註的拉框、分類不同，大模型任務更依賴人的理解，要求標註員理解文本，判斷回答是否合理，或者把圖片寫成完整描述。一些項目還需要醫學、英語、漢語言文學等專業背景。

事實上，永和愛豆科技也曾短暫承接過自動駕駛 3D 標註訂單，這和大模型文本任務並不是一類業務，卻同樣困難——員工要在黑底綠色的點雲中辨認車輛、錐桶等物體。李霞說，這些圖讓她們「看得暈」，「真的看不懂」。

這讓數據標註不再適合永和。對於一個人口較少、年輕人容易外流的縣城來說，承接這些訂單的難度很高。「如果一個項目來了，要求招 20 名大學畢業生，我們就招不到，小縣城就這麼多人。」張麗說。

也是在 2023 年前後，永和基地的大模型標註組由於找不到合適的員工被解散，公司保留下來的訂單主要來自於螞蟻集團旗下業務、圖片審覈項目和一些臨時任務。今年，這家公司接到的一個臨時任務是永和方言的採集。一些看到招聘信息的寶媽來基地參加培訓，回家用手機朗讀文本，完成採集任務。但這樣的任務只招兼職，已經不足以支撐起過去辦公室裏一百多人的員工數量。



×



永和縣城的招聘廣告。圖：作者提供

相比之下，靠近西安、銅川的基地可以從城市和高校吸引更多年輕人，也更容易通過校企合作補充人才。標註產業在遷移，寶媽們卻很難隨之移動。

直到現在，許多永和人對於 AI 也沒有太多概念。至於數據標註，一些人模糊地知道有人在做這件事，但具體是什麼，他們並不清楚。永和距離山西省會太原市接近 300 公里，至今還沒有高鐵站。

AI 標註讓縣城女性和未來產業有了一次短暫的連接，卻沒有真正改變她們在地方就業結構中的位置。她們一些人重新開始做零散的兼職，也有人離開基地去別處工作，「還是服務業。」黃丹丹說。

人工智能產業繼續擴張，但它帶來的訂單和工作機會，沒有停留在縣城。如今，永和縣城的愛豆科技辦公室裏，只零散坐着幾名員工。曾聚集起一百多名女性的數據流水線已經安靜下來。

今年，女工李霞發現兒子開始和豆包對話。看到圖片提取文字的功能時，她想起自己過去做過的 OCR 標註任務，問了一句：「這是不是我們那會兒訓練出來的？」

(文內受訪者均匿名)



×