人工智能 周刊

掌上长沙 山水洲城记
责编/王金文
美编/戴莹芳 校读/欧艳

谁在教AI认识世界？

长沙晚报记者采访给AI当“启蒙老师”的数据标注员群体，看看他们如何把世界讲给AI听

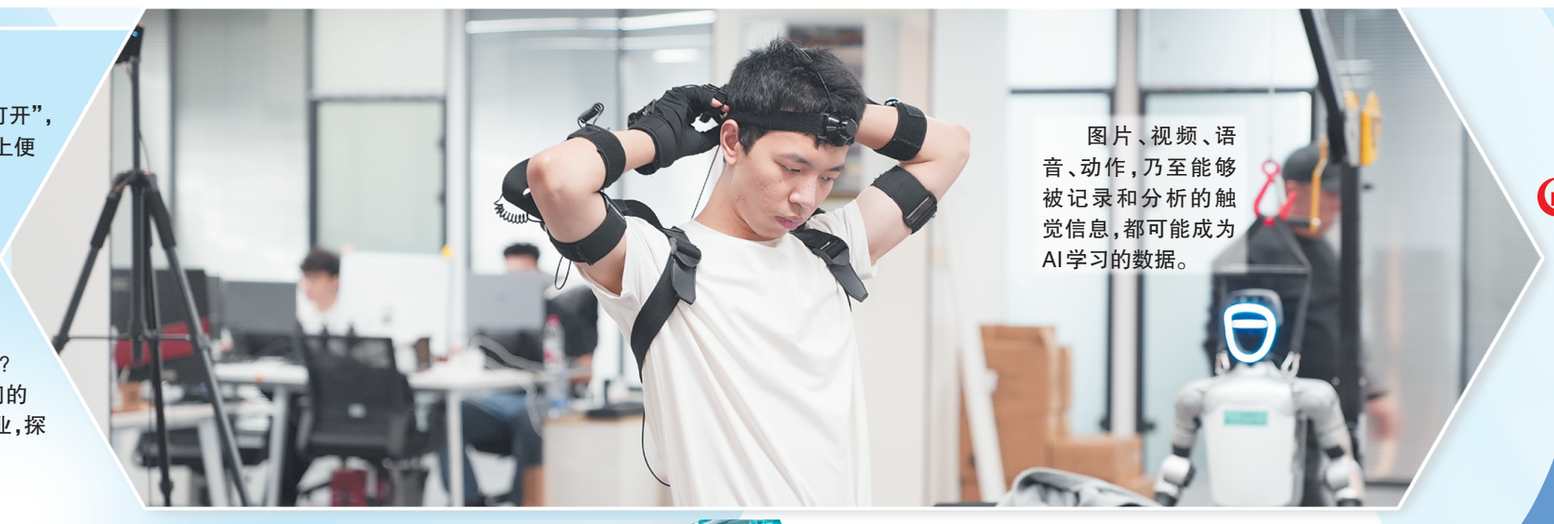
长沙晚报全媒体记者 黄飞武 罗杨

对着车机说一句“导航到橘子洲，把座椅加热打开”，它很快给出回应；给豆包发一句“帮我写首歌”，屏幕上便能生成一段旋律或歌词。

AI似乎越来越“无所不能”。但很少有人想过：它是怎么“学会”这些本领的？

答案不只在算法和算力里，也藏在一群人的日常工作中——他们不编写代码，也不直接训练模型，却通过一条条标注，帮助AI从海量数据中学习成长。他们就是AI数据标注员。

这个听起来有些陌生的新兴职业，每天究竟在做什么？AI越来越聪明，为什么还需要他们？这份工作又与我们的日常生活有何关联？近日，长沙晚报记者走进相关企业，探访一线从业者的工作日常。



图片、视频、语音、动作，乃至能够被记录和分析的触觉信息，都可能成为AI学习的数据。



1 给数据添上一份“说明书”

AI的学习，离不开海量数据的“喂养”。但它的学习不可能漫无目的、来者不拒。要让它真正具备“理解能力”，还需要有人告诉它：哪些信息值得关注，应该如何区分，又对应着什么含义。

比如，给AI看照片，怎样让它知道画面中的哪只动物是“猫”？让它听录音，又怎样分清谁在说话、说了什么？这就需要数据标注员给原始数据贴上相应的“标签”——在照片中框出猫的轮廓，在录音中标注文字内容、说话人和语音特征。经过规范标注的数据，才能用于模型训练，让AI从中学习规律。

国家数据局将数据标注概括为筛选、清洗、分类、注释、标记和质检等环节。简单来说，数据标注员做的，就是给原始数据添上一份“说明书”，从而让AI能够精准识别、理解并加以运用。

在湖南镁拓科技有限公司（简称“镁拓科技”），这样的工作每天都在进行。作为长沙本地头部的数据标注企业，这家企业年营收达千万级，在数据标注领域积累了深厚实力。

标注员从系统领取任务数据后，戴上耳机逐句核对语音。发现转写错误，要及时修正；遇到两个人同时说话，要根据规范切分语音，分别标记“说话人A”“说话人B”；碰上噪声较大的片段，还要单独标注相关信息；对于时长过长的句子，则要按照规范进一步切分。完成标注后，数据还要经过后台质检人员复核，确认符合要求，才能交付给模型训练。

看似只是听音频、改文字，实际操作却并不简单。正式作业前，团队还要先了解客户需求、试做样本、统一标注规则；进入具体环节后，小语种、方言、多人对话、语音不清等情况，又各有不同的处理要求。

而且数据标注的对象，也远不止语音。图片、视频、动作，乃至能够被记录和分析的触觉信息，都可能成为AI学习的素材。如今随着应用场景的不断拓展，标注工作也越来越需要各种专业知识的加持。

常见的岗位分为线下全职与线上兼职两种模式，企业会在线上平台持续释放标注任务，有兴趣的普通人经过简单培训，也可以居家线上接单参与，收入实行计件多劳多得；长沙本地全职标注员月薪大多集中在3000～6000元，掌握小语种、音乐、自动驾驶等专项技能的专业标注人员薪资会更高。

在长沙笠川科技有限公司，标注员周泳帆负责的是音乐数据。他毕业于音乐专业，同时是长沙交响乐团Melete合唱团的成员。普通人听一段曲子，或许只会觉得“好听”或“不好听”；而在周泳帆的工作中，一段音乐要求被拆解成更具体的信息：使用了什么乐器、采用什么和弦与节拍、整体情绪如何……这些内容都需要按照相应的标准进行识别和标注。这些更专业的标注，为AI学习音乐提供了更具体的依据，在辅助人创作时，才有可能更准确地理解创作要求，给出更专业的结果。

2 AI越聪明，越需要好数据

如果说算法决定了AI如何学习，那么数据就关系着它能学到什么、学得有多好。就像人要掌握一项技能，需要接触相应的教材和案例；AI要形成专业能力，也需要与目标相匹配的数据。

对于数据标注的重要性，清华大学教授孟庆国曾有明确定义：“数据标注是将原始数据转化为可识别、可训练、可计算的关键环节，其质量直接决定了数据集的应用价值。”

镁拓科技将AI学习过程比作盖房子：算法工程师好比绘制设计图纸的人，数据标注员则负责把一块块砖烧好、码齐。砖的质量不过关，房子就难以坚固；砖的数量不够，房子也难以建高。对于AI而言，数据标注的质量和规模，影响着模型最终能够达到的能力水平。

数据需求有多大？镁拓科技相关工作人员介绍，仅语音数据，训练一个功能就可能需要上万小时级的语料。不同应用场景、不同语言和不同任务，对数据的数量、质量及标注精度也有不同要求。如今，为满足不同语言的数据采集和标注需求，镁拓科技线上平台的数据采集、标注兼职人员已超过9万，参与者分布在美国、意大利、西班牙等全球多个国家和地区。

与此同时，AI也正在进入标注流程，成为标注员的“帮手”。早些年，缺少语音预识别工具时，标注员处理一段音频，几乎要从头听到尾，再逐字敲出文本。如今，AI可以先完成初步识别，人工再进行校对和修正，工作效率大幅提升。

相对来说，普通话语音等相对标准的场景，AI已经能够完成不少基础工作；可一旦遇到方言、小语种、多人交谈、车载发动机噪声干扰，或涉及专业乐理知识的复杂任务，仍需要人工进行更细致的判断和校验。

技术进步改变了标注员的工作方式，也不断抬高数据质量的要求。当AI能够处理更多常规任务，人工标注便进一步向复杂、专业、容易出错的环节延伸。

3 从一条数据，看见AI背后的产业链

数据标注处在AI产业链的前端，连接着应用需求与模型研发。

应用方提出训练需求，数据团队据此采集、加工、标注并质检数据，再将符合要求的数据交给研发人员训练模型；模型应用到具体产品后，新的使用场景和问题又会带来新的数据需求。由此，数据标注不仅参与AI能力的形成，也影响着模型后续的优化方向。

这种影响，最终会落到普通人的日常生活中。

汽车能不能听懂方言导航，语音助手能不能在嘈杂环境里分清不同说话人，AI音乐工具能不能识别曲风、辅助创作，自动驾驶系统能不能准确识别道路上的行人和车辆……这些能力的背后，既离不开算法和算力，也离不开经过整理、标注和质检的数据。

镁拓科技介绍，公司在车载语音数据领域与国内八成以上的新能源车企有直接或间接合作。你坐进一辆国产新能源车说出那句“你好，小X”，背后大概率就有这家长沙企业标注的数据在起作用。长沙笠川科技则正在参与专业音乐创作领域AI大模型的训练开发。

不同领域、不同类型的数据标注，支撑着AI的能力从通用对话，一步步走进更多更专业、更具体的应用场景。

长沙也正在围绕这一环节，加快培育相关产业。

2024年5月，长沙入选首批承担国家数据标注基地建设任务的7个城市之一。此后，长沙按照“1个综合标注基地+7个行业标注基地”的思路统筹布局，将整座城市作为一个“大基地”来建设。其中，综合标注基地落在湖南湘江新区；行业标注基地则结合各地产业特色，布局马栏山视频文创、长沙经开区工业制造、望城经开区医疗健康等领域，推动数据标注与本地优势产业相结合。

截至目前，长沙全市标注数据规模达4.38万TB，建成高质量数据集277个，赋能124个大模型，引育标注企业93家，培育标注人才2.5万人，带动相关产值120亿元。蚂蚁集团、海天瑞声等企业相继在长沙布局，THOTHAI也于今年9月签约入驻麓谷。

这些数字背后，是数据标注从单一的基础服务向多领域、专业化方向延伸，也是长沙将数据资源、人才和产业需求进一步串联起来的探索。

今天，当我们惊叹于AI对答如流、辅助创作的能力时，目光往往聚焦在模型突破与技术创新上，却容易忽略那些看似细碎的基础工作。

一段语音的反复听辨，一个标签的仔细核对，一组数据的多轮质检……这些不显眼的工作，也是AI学习理解世界的重要一环。

技术让AI不断进步，而数据标注员则以耐心和专业，为AI认识世界提供更准确的参照。他们或许不在聚光灯下，却在屏幕与耳机之间，把纷繁复杂的现实世界，一点点转化为AI能够读懂的语言。



扫码看辣视频
数据标注员
把世界讲给AI听

→数据标注员将复杂的世界转化为AI能够读懂的语言。均为长沙晚报全媒体记者 尹昊天 摄

文体

责编/宁莎鸥 美编/王斌 校读/李乐

长沙图书馆联合高校打造的《深巷之光》马栏山拿下一等奖

AI影游让古籍“开口说话”

长沙晚报全媒体记者 宁莎鸥 通讯员 罗倩倩

10月26日，由长沙图书馆联合长沙理工大学设计艺术学院共同打造的AI互动影视游戏《深巷之光》，在马栏山“V谷·新境计划”互动影游创作培育营暨作品大赛中脱颖而出，荣获高校组一等奖。昨日，记者采访创作团队，了解长沙图书馆在“AIGC+沉浸式阅读+文旅融合”领域的创新探索。



实体解谜剧本《深巷之光》的首发体验活动。长图 供图

从15支参赛团队中脱颖而出

10月26日，“V谷·新境计划”决赛路演暨产业对接会在网易（长沙）数字产业中心成功举办，以“路演答辩+专业评审+资源对接”全链条，赋能青年创业创新与优质项目落地。

决赛路演答辩环节，15支参赛团队依次进行5分钟项目路演与2分钟评委问答，评委现场提问、项目实测、作品试玩并综合打分，评审维度涵盖创意叙事、交互沉浸、美术视听、技术完成、价值潜力五项。经过激烈角逐，《深巷之光》凭借独特的创意叙事、沉浸的交互体验以及深厚的文化内涵，从15支参赛团队中脱颖而出，斩获高校组最高奖项。

从红色剧本到AI影游 跨越两年的创新迭代

《深巷之光》的诞生，源于长沙图书馆对红色阅读推广方式的持续探索。早在2024年，长沙图书馆就开始与长沙理工大学设计艺术学院研究生团队携手，深入挖掘长沙和平解放前夕那段波澜壮阔的历史，以“小人物推动大历史”的独特视角，为公众呈现了红色阅读解谜剧本《深巷之光》。

时隔一年，长沙图书馆与长沙理工大学设计艺术学院师生团队延续合作，将《深巷之光》从实体剧本升级为AI互动影视游戏，第一幕已于2026年4月正式上线B站。而AI驱动的《深巷之光》保留了解谜剧本的多分支选择与高自由度叙事，同时依托长沙图书馆丰富的馆藏文献资源，将真实史料加入游戏剧情，更融入了智能对话与动态剧情生成。

当你在波诡云谲的局势中“传递情报”时，你需要查阅数字化的《资治通鉴》寻找破局之法；当你面对“生死抉择”时，泛黄的《左宗棠手札》与《注解伤寒杂病论》或许就是解开谜题的关键钥匙；那些长沙民国时期的老照片，更是将你一秒拉回那个峥嵘岁月。玩家的每一次点击、每一次选择，都在与真实的史料对话。阅读也不再是单向的接收，而是一场充满未知与惊喜的探索。

长沙图书馆副馆长陈卓介绍，长沙图书馆自2020年陆续推出《火炬》《深巷之光》等原创红色解谜剧本，已吸引两万余人次参与。据悉，该馆已依托《边城》《湖南抗战》等地方文献及经典文学作品孵化了17部原创解谜剧本。

项目策划初衷：不止于讲故事，更在于连接

长沙图书馆策划《深巷之光》的初衷，不止于讲述长沙和平解放的历史故事，更希望融入古旧文献数字化阅读，实现文旅深度融合，让“在地、在线、在场”三者相连接。

“在地”，是深入挖掘地方文献与地域红色文化资源，将沉睡的史料转化为鲜活的创意动能；“在线”，是借助AI技术与互联网平台，打破时空限制，让馆藏文献以互动影游等全新形态触达更广泛的受众；“在场”，是让每一位参与者成为故事的“在场者”，让他们走进图书馆、走进历史现场，通过沉浸式阅读、参观和体验感受红色文化的厚重与温度。在剧本中，长沙图书馆将火宫殿、爱晚亭、铜官古镇、春华老街等历史建筑与城市街巷作为实景载体融入剧情，将体验场景延伸至火宫殿、刘廷芳公馆、湖南和平解放史事陈列馆等城市文旅地标，通过构建“阅读—行走—解谜”沉浸式文旅体验闭环，将地方文献转化为动态“城市漫游”文旅产品，实现公共文化产品供给从“馆内阅读”向“全域体验”提质升级。

“非遗班的少年”再度启航

长沙300场非遗进校园课程开放申报

长沙晚报10月28日讯（全媒体记者 郑湘平 通讯员 曾仪）湘剧开唱，巧手传艺。记者今日从长沙市非物质文化遗产保护中心获悉，品牌活动“非遗班的少年”2026年长沙市戏曲进校园活动启动申报，全市中小学校可申请共计300场公益课程与非遗活动，让湘地非遗扎根校园、浸润童心，申报截止时间为11月13日。

作为长沙未成年人美育与非遗传播重点项目，“非遗班的少年”2025年活动成果亮眼。活动覆盖长沙9区县（市）、走进68所校园及文化场馆，开展300余场名师课堂与主题活动，吸引8万余名少年学子参与；全市建成10所标准化“非遗戏曲社”，年度展演集中展出学生非遗作品500余件。该项目先后获评长沙市未成年人思想道德建设创新案例，并入选2026全国非遗传播创新案例。

今年活动持续加码，面向全市中小学校开放申报，学校可结合场地条件、学生规模申报多场次课程，学校无需付费。课程分为两大板块：戏曲专业课程由非遗代表性传承人入校讲授湘剧、长沙花鼓戏、长沙弹词、长沙皮影戏，开展基础教学与剧目排演；非遗实践课程主打戏曲与手工技艺融合，涵盖湘绣、长沙剪纸、浏阳油纸伞、长沙面塑、糖画、榫卯小木作、传统和香、掐丝釉彩、湘剧脸谱绘制等十余项长沙本土非遗。课程时长45至90分钟，单场20至40人。

寻美非遗

用青春脚步丈量历史老街

湖南探索历史文化街区焕新密码

长沙晚报10月28日讯（全媒体记者 宁莎鸥 通讯员 邹伟瑜 黄梅 贺瑜）28日下午，湖南省历史文化保护利用片区（街区）创新发展交流暨City-walk活动在长沙潮宗街历史文化街区启动，一场以脚步丈量文脉、以青春激活老街的湖湘街区焕新行动由此展开。

活动以“一步千年 潮涌湖湘”为主题，设置创新发展交流活动、“青春寻脉”高校学生Citywalk活动、“悦享街区”文旅强基焕新行动、“声动湖湘”全媒体传播计划等系列活动，通过“交流研讨+青年实践+业态焕新+全媒传播”的多元联动，探索湖南历史文化街区活化利用新路径。

活动现场，历史文化片区“Citywalk宝藏路线”地图发布，串联长沙潮宗街、岳阳楼片区、永州柳子街、凤凰古城、常德河街等历史文化片区（街区），为市民游客提供可漫步、可体验、可消费的线路指引。

目前，湖南省共拥有历史文化街区67处，全省坚持科学规划、分类施策、有机更新，以微改造、精提升为路径，扎实推进各类历史文化遗存保护与活化利用。活动由湖南省文化和旅游厅、湖南省住房和城乡建设厅、长沙市人民政府主办。

扫码看辣视频
历史街区焕新

悦享书香
报名电话：82892600

长沙住房公积金管理中心公告

长沙卡斯迪酒店管理有限公司：

我中心受理你单位（原）职工刘会琼投诉你单位欠缴其住房公积金一案，现依法向你单位公告送达《行政处理决定书》（长金处决字〔2026〕第02030号）：责令你单位自本决定书送达之日起10日内依法补缴刘会琼2018年11月至2020年10月的住房公积金10541元。如对本处理决定不服，可于收到之日起60日内向长沙市人民政府申请复议，也可在6个月内向长沙铁路运输法院提起诉讼，逾期不申请复议、不起诉又不履行本行政处理决定的，我中心将依法申请人民法院强制执行。

自公告之日起30日内到我中心领取《行政处理决定书》，逾期即视为送达。

联系人：天心区管理部 电话：0731-84886873

领取地址：长沙住房公积金管理中心天心区管理部（长沙市天心区友谊路388号友谊路地铁口4号出入口北面）

长沙住房公积金管理中心
2026年10月29日

长沙住房公积金管理中心公告

湖南壹美达品牌运营管理有限公司：

我中心受理你单位（原）职工吴鸿辉投诉你单位欠缴其住房公积金一案，现依法向你单位公告送达《行政处理决定书》（长金处决字〔2026〕第02029号）：责令你单位自本决定书送达之日起10日内依法补缴吴鸿辉2025年8月至2026年1月的住房公积金1194元。如对本处理决定不服，可于收到之日起60日内向长沙市人民政府申请复议，也可在6个月内向长沙铁路运输法院提起诉讼，逾期不申请复议、不起诉又不履行本行政处理决定的，我中心将依法申请人民法院强制执行。

自公告之日起30日内到我中心领取《行政处理决定书》，逾期即视为送达。

联系人：天心区管理部 电话：0731-84886873

领取地址：长沙住房公积金管理中心天心区管理部（长沙市天心区友谊路388号友谊路地铁口4号出入口北面）

长沙住房公积金管理中心
2026年10月29日