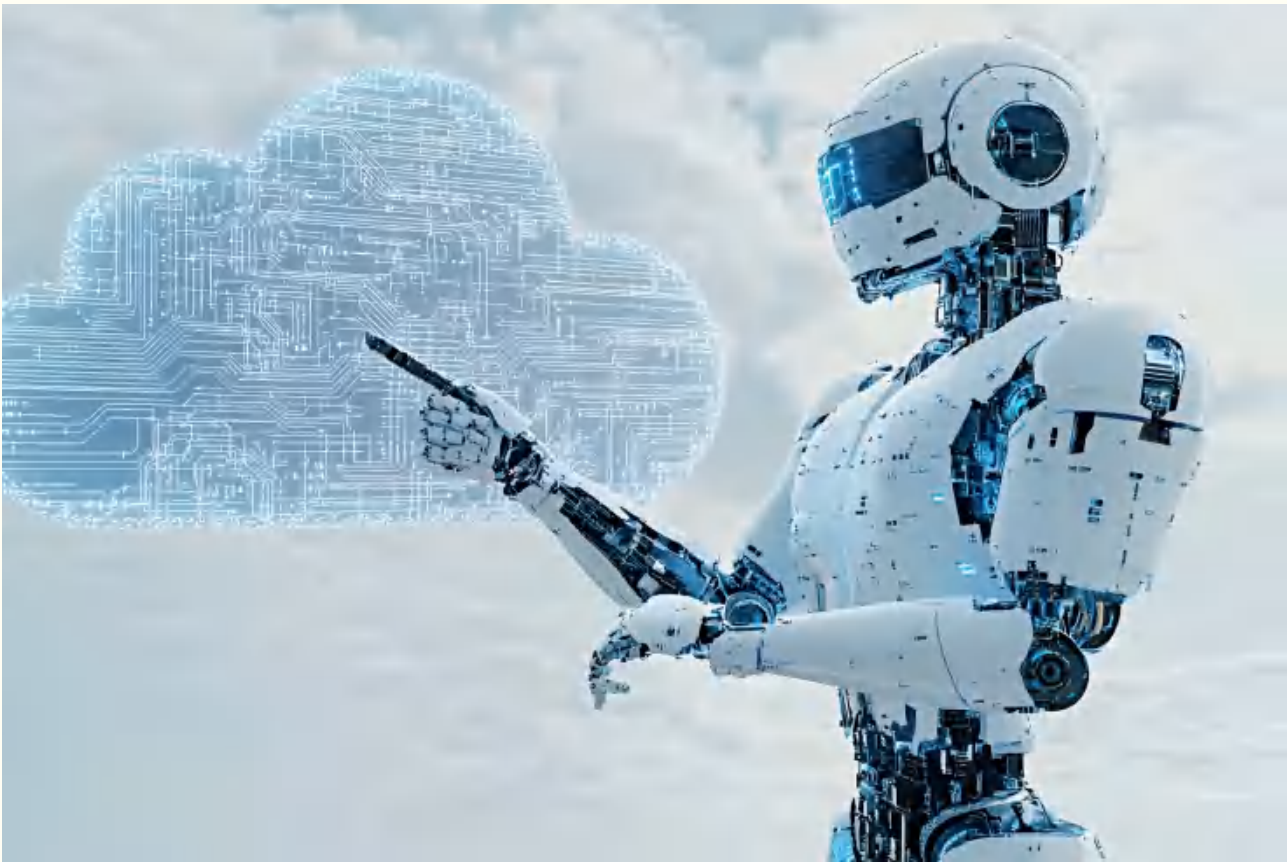


强化数据标注基地引领作用 带动数据标注产业高质量发展

文 ◆ 中国电子信息产业发展研究院院长 张 立

数据标注产业是对数据进行筛选、清洗、分类、注释、标记和质量检验等加工处理的新兴产业，对于提升数据供给质量，推动人工智能创新发展具有重要支撑作用。近日，国家发展改革

委、国家数据局、财政部、人力资源社会保障部联合印发《关于促进数据标注产业高质量发展的实施意见》（以下简称《实施意见》），明确提出要从深化需求牵引、增强创新驱动、培育繁荣生态、优化支撑体系等方面多管齐下、协同发力，同时强调要鼓励和支持数据标注基地先行先试，打造一批科技水平高、资源集聚强、辐射带动广的典型样板，为我国数据标注产业高质量发展指明发展方向和推进路径。



当前，在国家数据局的指导下，有关地方正在积极推动建设数据标注基地，推动产业集聚发展。下一步，有关地方应发挥数据标注基地的示范引领作用，从供需两端双向发力，构建高水平供给体系，全方位促进需求增长，推动数据标注产业高质量发展。

一、数据标注产业对推进“数据要素×”和“人工智能+”行动具有重要意义

发展数据标注产业是数据要素价值释放的前提条件。当前我国诸多领域数据要素开发利用仍处于摸索阶段，面临数据加工能力不足等问题。数据标注是数据加工的核心环节，发展数据标注产业有助于提升数据供给质量、提高数据开发利用水平、释放数据要素价值。发展数据标注产业是发挥我国海量数据规模和应用场景优势、推动数据要素在不同领域形成乘数效应的重要依托。

发展数据标注产业是推动人工智能创新发展的重要支撑。麻省理工学院等研究机构指出，互联网公域高质量文本数据将在2026年“耗尽”，人工智能发展将遇到数据壁垒。在此情况下，通过数据标注技术手段打开私域数据供给空间就成为人工智能创新发展的关键。数据标注构建了数据资源与人工智能应用的连接桥梁，为人工智能持续创新发展提供高质量数据集。大力发展数据标注产业对于“人工智能+”行动的顺利实施至关重要。

二、数据标注基地为地方产业发展提供新的增长引擎

为汇聚数据资源、提升数据质量、盘活数据要素价值，探索数据标注产业高质量发展的科学路径，国家数据局部署了成都、沈阳、合肥、长沙、海口、保定、大同等7个承担数据标注基地建设任务的城市，并支持上述城市在技术创新、行业赋能、生态培育、标准应用、人才就业、数据安全等方面先行先试。在技术创新方面，基地积极布局智能标注技术攻关。例如，沈阳市制定了《数据标注技术创新指导意见》，以“揭榜挂帅”形式支持建成人工智能医学图像分析系统，能够提供医学光镜图像、病理图像等数据智能标注，已标注10TB高质量数据集，并应用于有关大模型研发。在行业赋能方面，基地打造了一批行业高质量数据集。例如，成都市已储备金融、医疗、网络安全等领域的多模态高质量数据集5400TB，赋能训练30余个人工智能模型。长沙市天心经济

开发区已聚集地理、气象等8个领域共计700TB的数据资源，可支持智能交通、灾害预警、城市规划等人工智能应用。在生态培育方面，基地产业集群效应初步显现。例如，合肥市积极整合产业链资源，已与国内外上百家相关企业开展数据合作。海口市运营百度智能云人工智能基础数据产业基地，吸引大模型企业和数据标注企业等入驻，打造协同发展的产业生态。在标准应用方面，基地稳步开展标准体系建设。例如，成都市有关科研机构正在制定数据标注技术、数据加工分析能力评估、数据资源集成能力评估等地方和行业标准，为提升数据标注质量打好基础。在人才就业方面，基地推进产教融合多元培养成效显著。例如，大同市积极推进产教融合，培养数据标注人才，目前基地累计已培训就业数据标注相关人员约两万人次，其中本科及以上学历占比近三成。保定市举办首届人工智能基础数据标注职工职业技能大赛，旨在推动培养优秀数据标注人才，约93个团体400人参加。在数据安全方面，基地数据安全保障体系逐步建立。沈阳市、成都市、长沙市、大同市等从技术和制度两个维度强化数据标注全流程安全管理，例如搭建基于区

区块链技术的数据标注安全溯源机制、推动数据标注安全生产环境建设、开展数据合规认证、建立数据安全风险分级分类体系等。

三、提升四类供给，释放两类需求，做大做强做优数据标注基地

当前数据标注产业发展仍面临一些问题，从供给端看，面临智能标注技术供给不足、产业基础薄弱、标准体系不健全、高水平人才短缺等问题；从需求端看，面临数据标注市场需求规模低的问题。数据标注基地作为数据标注产业发展的示范标杆，应从供需两端双向发力，“提升四类供给，释放两类需求”，为解决上述问题提供指引和路径。

（一）供给端：多措并举，构建高水平数据标注供给体系

技术创新提升供给效率。《实施意见》强调技术创新对产业发展的推动作用，明确指出要加强数据标注领域关键技术攻关应用，加快科技成果转化和应用落地。数据标注基地应充分发挥高水平创新载体作用，加强前沿技术探索，推动数据标注技术融合创新，研发智能化处理工具和平台，以技术创新提高企业生产率，为市场提供更加精准、更

加多元、更加高效的数据标注服务。**生态培育提高供给规模。**《实施意见》指出，要着力壮大经营主体，培育数据标注领域龙头企业，加快中小企业发展步伐；要着力健全产业生态，推动数据标注产业上下游协同发展。数据标注基地应利用区位优势和资源禀赋，通过税收、资金、土地等优惠政策推动产业集聚，打造完善的产业链，促进技术、人才、资金等配置优化，形成数据标注产业集聚的规模效应，提升数据标注服务的供给规模。**标准完善保障供给质量。**《实施意见》指出，要围绕数据标注关键环节，建立标准体系框架，加快制定数据标注技术、质量、能力等方面的标准。数据标注基地应积极构建涵盖流程、安全管理、技术、质量检测、人员培训等多个方面的数据标注标准体系，保障数据标注产品服务的供给质量，形成数据标注基地品牌效应。**人才引育强化供给支撑。**《实施意见》明确提出，要加强数据标注人才队伍建设，通过引育高端专业人才、深化产学研融合、畅通人才发展通道等多种方式强化产业人才支撑。数据标注基地应牢牢把握人才引进和人才培养双管齐下的原则，着力构建高水平数据标注人才发展体系，利用优质人才资源助推数据标注应用场景不断深化，为数据标注产业高质量发展提供坚实支撑。

（二）需求端：需求驱动，推动数据标注产业蓬勃发展

释放公共数据标注需求。《实施意见》中提出，要深化人工智能在公共领域的应用，依法依规有序推动公共数据标注与开发利用，支持公共数据赋能实体经济发展。数据标注基地有关部门可结合实际，统筹安排数据标注服务采购费用，将标注后高质量公共数据开放或授权开发利用，释放公共数据标注需求。**挖掘企业数据标注需求。**《实施意见》明确指出，深挖企业生产管理全流程数据标注需求，加大企业数据开发利用。一方面，数据标注基地应引导企业通过数据标注开展数据资源管理，提升数据治理能力，助力企业乃至行业数字化转型，创造更多企业数据标注需求。另一方面，数据标注基地应通过开展供需对接、发放数据券等方式引导企业外包数据标注业务，释放更多企业数据标注需求。

展望“十五五”，在《实施意见》的指引下，数据标注产业的专业化、智能化及科技创新能力水平将显著提升，产业规模将大幅跃升，数据标注产业将成为地方新的经济增长点，数据标注产业将推进数据要素市场化配置改革，更好促进高质量发展。■

（文章来源：国家数据局公众号）