

人工智能数据标注与共同富裕

刘贺洋^{*} 梁耀丹 闫谨 吴梦欣悦

北京物资学院, 中国·北京 101149

摘要: 我们将通过调查问卷和整理相关材料等方法, 探讨人工智能数据标注与共同富裕间的关系及影响。人工智能数据的标注是对数据进行筛选、清洗、分类、标记等各种加工处理, 产生高质量的数据集, 为人工智能发展提供优质“原料”, 有力支撑人工智能创新发展。从而使得数据标注的需求在中小城市迅速涌现, 扩大了数据标注的行业规模, 促进了地方经济发展; 为当地居民提供了大量就业机会, 增加了他们的收入。同时带动数字基础设施建设, 一定程度上缩小区域、城乡间的数字差距, 使各地区人群都能从数字经济发展中受益, 促进共同富裕。

关键词: 人工智能; 数据标注; 共同富裕; 就业促进; 区域经济; 数字鸿沟

Artificial intelligence data annotation and common prosperity

Liu Heyang^{*}, Liang Yaodan, Yan Jin, Wu Mengxinyue

Beijing University of Civil Engineering and Architecture, China Beijing 101149

Abstract: We will explore the relationship and impact between artificial intelligence (AI) data annotation and common prosperity through methods such as survey questionnaires and sorting relevant materials. AI data annotation involves various processing and handling of data, such as screening, cleaning, classification, and labeling, to produce high-quality datasets. These datasets provide high-quality "raw materials" for AI development and strongly support the innovative development of AI. As a result, the demand for data annotation has rapidly emerged in small and medium-sized cities, expanding the industry scale of data annotation and promoting local economic development. It provides a large number of employment opportunities for local residents, increasing their income. At the same time, it drives the construction of digital infrastructure, narrows the digital gap between regions and urban-rural areas to a certain extent, enabling people from all regions to benefit from the development of the digital economy and promoting common prosperity.

Keywords: Artificial intelligence; Data annotation; Common prosperity; Employment promotion; Regional economy; Digital divide

0 引言

研究背景:

共同富裕是中国式现代化的本质要求。《浙江高质量发展建设共同富裕示范区实施方案(2021—2025年)》^[1] 强调以数字经济推动城乡区域协同发展, 并将数字产业作为缩小差距的重要抓手。人工智能技术的迅猛发展使高质量训练数据成为关键瓶颈, 数据标注产业通过对原始数据进行加工处理, 为 AI 模型提供基础支撑, 广泛应用于医疗、自动驾驶、金融等领域。

2023 年, 国家出台《关于促进数据标注产业高质量发展的实施意见》^[2], 首次将数据标注提升至国家战略层面, 推动其规范化、规模化发展。据统计, 2023 年中国数据标注行业规模达 60.8 亿元, 预计 2025 年将突破 120 亿元。该产业呈现明显“下沉趋势”, 80% 的企业分布于四五线

城市及县域, 创造超 50 万个就业岗位, 其中超 60% 面向低技能劳动力。典型案例如云南蒙自、贵州“AI 标注村”等, 表明该产业在促进增收、弥合数字鸿沟和推动区域协调方面发挥重要作用。然而, 当前关于其与共同富裕的系统性研究仍不足, 技术及制度瓶颈亟待突破, 需深入探索其内在机制与发展路径。

研究意义:

从理论方面看, 本研究聚焦数据标注产业与共同富裕的关系, 突破以往对传统产业的路径依赖, 构建“技术—就业—公平”分析框架, 丰富数字经济背景下共同富裕的理论体系。通过跨学科视角解析产业价值创造机制, 填补相关研究空白, 为后续学术探讨提供基础。从实践方面看, 通过实地调研与案例剖析, 总结数据标注产业推动共同富裕的实践经验, 识别关键瓶颈。研究成果可为地方政府制

定产业政策、企业优化布局及社会组织参与民生改善提供参考，推动产业高质量发展与共同富裕目标协同实现。

研究方法与数据来源：

本研究采用多元化的研究方法，系统整合文献研究与案例分析，以全面、深入地探讨产业赋能共同富裕的实践机制与路径。首先，通过文献研究法，对国内外相关政策文件、学术成果及行业报告进行系统梳理与综述，把握当前理论前沿与实践动态，构建扎实的理论基础和分析框架。在此基础上，运用案例分析法，选取贵州、河南、云南等地的典型实践模式作为研究对象，通过深入剖析其具体做法、成效与问题，总结不同区域背景下产业赋能推动共同富裕的实践路径与共性经验，为研究结论提供扎实的实证支撑。

1 理论基础与核心概念界定

1.1 核心概念界定

1.1.1 人工智能数据标注产业

指对人工智能训练所需数据进行筛选、清洗、分类、标记等加工处理，生产高质量数据集的产业形态，涵盖数据采集、预处理、标注执行、质量审核等环节，具有劳动密集与技术辅助相结合的特征，广泛服务于计算机视觉、智能语音、自动驾驶等 AI 应用领域。

1.1.2 共同富裕

指在高质量发展基础上实现的全体人民物质富足与精神富裕的统一，核心是通过缩小城乡、区域、群体间的发展差距，实现社会公平与包容发展，具体体现为就业充分、收入均衡、公共服务均等化等维度。

1.2 理论基础

1.2.1 数字红利分配理论

该理论认为数字技术的发展产生了新型生产力红利，但其分配具有不均衡性，易形成“数字鸿沟”。数据标注产业通过将 AI 技术发展的中间环节下沉至欠发达地区，为弱势群体提供参与数字经济的机会，实现数字红利的普惠性分配。

1.2.2 劳动力市场分割理论

传统劳动力市场存在城乡、技能、地域等分割现象，导致弱势群体就业机会匮乏。数据标注产业以低技能门槛打破传统市场分割，构建跨地域的弹性就业市场，为农村剩余劳动力、城市低技能劳动者等群体开辟新的就业通道。

1.2.3 产业集聚理论

产业集聚通过规模效应与溢出效应促进区域经济发展。数据标注产业在四五线城市及农村地区的集聚，不仅

直接带动就业，更推动数字基础设施建设与相关产业发展，形成区域经济增长的新极点。

2 数据标注产业驱动共同富裕的现实基础与实现路径

2.1 现实基础：产业发展与政策支持的双重保障

随着 AI 技术在各领域的普及应用，数据标注需求从互联网行业向医疗、金融、自动驾驶等垂直领域延伸，产业规模持续扩张且需求日益多元。在政策层面，体系不断完善且形成协同发力的格局，国家层面出台《关于促进数据标注产业高质量发展的实施意见》《“十四五”数字经济发展规划》^[1]等政策，明确释放公共数据需求、加大财税支持、加强人才培养等举措，地方层面也积极响应，云南推进“数字乡村”标注项目^{[4][7]}覆盖，贵州建立标注技能认证体系，中央与地方协同推进产业发展。同时，数据标注产业的就业吸纳能力突出且包容性强，该工作对学历、技能要求较低，经过短期培训即可上岗，能够适配农村剩余劳动力、城市低技能劳动者、转行者等群体。

2.2 实现路径

产业赋能共同富裕的实现路径，首先体现在创造普惠就业、扩大收入基础。通过大量低门槛“数字岗位”，产业快速吸纳不同技能劳动力，将规模增长转化为广泛可持续的就业机会。如飞数信息（安徽）有 2000 余名专业标注员及 50 万众包劳动力；京数云（辽宁）作为残健融合就业基地，为特殊群体提供包容性岗位，彰显社会包容性。

其次，优化区域布局、促进协调发展是重要路径。产业数字化打破地理限制，推动数字产业链从沿海向中西部、东北扩散，实现“数字产业西迁”与“就业本地化”，缩小区域差距。如多鲸科技与空港数据（河南）成为区域性数据服务集聚中心；恒达地理（河北）、数字天码（重庆）成功拓展全国市场，体现良好区域协同。

进一步提升人力资本、实现技能增值，是关键机制。通过产教融合与真实项目实践，普通劳动者可成长为专业“AI 训练师”，实现就业规模扩张到劳动质量升级的转变。如多鲸科技（河南）年培养超 2000 名 AI 训练师；百度智能云通过众包生态建设，提升合作企业人员数字技能。

最后，驱动传统产业升级、共享发展红利，是可持续共同富裕的重要路径。高质量数据服务赋能自动驾驶、金融、医疗等行业智能化转型，提升全社会生产效率，做大经济蛋糕。典型案例有数族科技（江苏）为金融风控降本增效，本原智数、澳鹏中国以高精度数据标注推动高附加值产业发展。

这些路径的规模化与可持续性，得益于大模型自动化工具、人机协同和智能质检等关键技术支撑。这些技术在提升数据服务效率与质量的同时，降低行业参与门槛，为产业赋能共同富裕提供坚实保障。

3 数据标注产业驱动共同富裕面临的困境

3.1 技术迭代与人才短缺的双重压力

3.1.1 AI 技术对传统标注的替代风险

随 AI Agent 等技术发展，自动标注在简单重复任务（如图像分类、文本关键词提取）占比 2024 年达 45%，2026 年预计突破 60%，仅靠基础标注技能的岗位将逐渐压缩。

3.1.2 高端复合型人才严重短缺

产业向“专业化、场景化、高质量”转型，医疗标注需基础医学知识、金融标注需懂信贷流程，但既懂 AI 又熟悉垂直行业的人才稀缺。部分高校人才培养与产业需求脱节，叠加发达地区人才虹吸效应，中西部及基层企业难引留高端人才。

3.2 区域与产业发展不平衡

数据标注产业虽呈“下沉”趋势，全国 80% 标注企业分布于四五线城市及县域，但区域协同不足，如四川除成都外，其他市州产业基础薄弱，未形成全省协同发展局面。同时产业生态与品牌竞争力待提升，多地标注企业“散、小、弱”，缺乏全国影响力与技术创新能力的龙头企业，业务同质化严重、集中于基础标注，利润有限；供需对接不畅，如四川庞大 AI 产业需求未充分匹配本地标注服务能力。

3.3 就业质量与职业发展困境

数据标注岗位薪酬水平不高（部分外包岗月薪几千到一万多，正式岗与大厂技术岗有差距），职业发展路径模糊，影响长期吸引力，尤其难留高素质人才。且多数标注员从事高重复、低创造性工作，技能提升慢。虽有泰安等市建“数字经济工匠学院”培养数字技能人才，但广泛的系统性职业培训体系与明确职业资格认证通道仍待完善，致从业人员易陷“流水线”困境，难获有效职业晋升。

3.4 数据要素化与制度保障的挑战

3.4.1 高质量数据供给与合规成本

AI 大模型训练需大规模、高质量标注数据，但产业面临数据合规安全挑战。在维护国家数据安全、保护个人信息与商业秘密前提下，企业需解决合法合规获取、标注及流通数据的问题，一定程度增加运营成本。

3.4.2 产业政策与标准体系需持续优化

虽国家出台《关于促进数据标注产业高质量发展的实施意见》，各地亦推出支持政策（如山西 15 条措施、贵州 10 项政策），但政策细化落地、行业标准统一（当前标注标准不统一）、构建覆盖数据质量、隐私保护、安全审计的全流程规范体系，仍任重道远。

4 优化路径与政策建议

4.1 推动产学研转化

建立“企业需求—高校研发—成果落地”的闭环，如清华大学与华为公司签署科技合作框架协议，共同推进“产学研用”协同创新。2019 年 11 月，华为技术有限公司智能数据与存储产品线与清华大学软件学院、大数据系统软件国家工程实验室正式成立华为—大数据系统软件国家工程实验室联合创新实验室，进行大数据与数据库领域的联合创新。

4.2 政府支持

《国家发展改革委等部门关于促进数据标注产业高质量发展的实施意见》^[5]表明以促进数据开发利用、赋能经济社会发展为主线，布局数字科技新赛道，构建产业国际竞争新优势。发展数据标注产业坚持有效市场和有为政府相结合，预计到 2027 年，数据标注产业专业化、智能化及科技创新能力显著提升，产业规模大幅跃升，年均复合增长率超过 20%，形成相对完善的数据标注产业生态，区域协同发展的新格局。

4.3 人才保障

建立“标注员职业技能等级认定”制度，与薪资、晋升挂钩，提升职业吸引力。《关于促进数据标注产业高质量发展的实施意见》^[6]，文件明确提出要依托行业组织、院校、社会培训评价组织开展数据标注相关职业技能等级认定，支持分层次建设数据标注人才库，强化产业人才支撑

5 结语

本研究探讨了数据标注产业与共同富裕的内在联系及实现路径。研究表明，数据标注产业低门槛、劳动密集、地域下沉的特点，有效创造大量就业岗位，提升低技能劳动者、农村地区及弱势群体收入，为缩小城乡差距、促进区域协调发展提供路径。典型案例显示，该产业不仅带动县域经济发展，还通过技能培训与数字基建，增强地区可持续发展能力。

然而，产业仍面临自动化技术替代、标注质量参差、职业发展通道不畅、区域发展不均衡等挑战。未来，需推

动政策落地与跨部门协同, 加强技术创新与人才体系建设, 促进数据标注从劳动密集型向知识密集型转型; 同时深化数据要素市场化改革, 探索标注数据资产化路径, 释放产业更大价值。

展望未来, 数据标注产业有望在政策引导与技术赋能下, 成为推动包容性增长和共同富裕的重要力量, 为县域经济注入新动能, 也为全球数字时代公平与发展问题提供“中国方案”。后续研究可关注其长效作用机制与产业生态优化路径。

参考文献:

[1] 国家发展改革委. 浙江高质量发展建设共同富裕示范区实施方案(2021—2025年)[Z]. 2021.

[2] 国务院. 关于促进数据标注产业高质量发展的实施意见[Z]. 2023.

[3] 国家发展改革委, 中央网信办: “十四五”数字经济发展规划[Z]. 2021.

[4] 国务院办公厅. 数字乡村发展行动计划(2022—2025年)[Z]. 2022.

[5] 国家发展改革委, 国家数据局, 财政部, 人力资源社会保障部. 国家发展改革委等部门关于促进数据标注产业高质量发展的实施意见[Z]. 2025.

[6] 国家数据局. 构建高素质人才队伍, 助力数据标注产业发展[Z]. 2023.

[7] 中共云南省委办公厅, 云南省人民政府办公厅. 关于加快推进数字乡村建设的实施意见[Z]. 2020.

基金项目: 本文系北京物资学院“人工智能数据标注与共同富裕”项目阶段性成果(编号: 2025010402072)。