大同市众梦云科技有限公司

数据产业人才培养就业

实施方案

2025年7月

一、项目背景

在数字经济成为全球经济增长新引擎的时代背景下，数据标注成为了推动技术创新的核心因素，国家发改委等四部门联合出台了《关于促进数据标注产业高质量发展的实施意见》等政策文件，明确提出加强标注人才队伍建设，深化产学研融合，鼓励行业联盟、高校、科研院所与企业建立长期合作机制，加大数据标注实践项目、继续教育和公共实训基地建设合作力度。

近年来，大同市积极做大数字产业集群，发展数字经济，相继出台了《数据产业发展三年行动计划（2024年—2026年）》、《关于加快人工智能领域人才就业若干措施》等系列文件政策，依托算力、地域、人力资源等方面的优势，数据标注产业集群发展初步成型，探索走出了一种“算力筑基、数据赋能”的资源型城市转型新路径。但规模化发展，人才需求激增，同时企业对员工培训专业性不足，导致人员流动性增大，人员流失比例偏大，专业性人才稀缺，尤其是专业知识要求较高的领域更是人才短缺，多种因素制约了数据标注行业的良性健康发展。

北京众梦未来集团作为产教融合领域的标杆企业，以“赋能产业升级，培育时代人才”为使命，形成了“总部引领、区域协同、多元合作”的发展格局。集团总部扎根北京，辐射全国，在多地建立专业人才培养基地与产业合作平台，深度整合政、产、学、研多方资源，推动教育链、人才链、产业链与创新链的有机衔接。目前，集团在江苏盐城、东台市、江西吉安市、宁夏吴忠市均有独立公司，目前山西大同是第五个独立公司(大同市众梦云科技有限公司）。

为助力高质量完成国家级数据标注基地建设任务，响应《国家产教融合建设试点实施方案》政策要求，破解数据标注人才短缺制约行业发展的难题，充分挖掘本地化数字人才就业，短时间解决数据标注企业用工需求，探索数字化人才培养就业新模式。按照大同市《关于加快人工智能领域人才就业若干措施》文件精神，大同市众梦云科技有限公司（以下简称“众梦云”）依托产业资源与职业教育优势，充分结合大同市数据产业发展实际，最大程度保障财政资金安全，制定本产教融合人才培训计划。

二、项目目标

一年内完成2000名数据标注及相关领域技能型人才的闭环培养，打造“本地培训—本地实习—本地就业”的可持续发展链条。产教园区引入数据类企业5家，年产值达到1亿元。

三、具体措施

（一）招生与选拔

**目标群体：**面向本地高校应届生、职校学生及待业青年，优先选择大数据、计算机、人工智能等相关专业。

**招生规模：**分3批次完成2000人招募（每批500人），第一批500人，第二批500人，第三批1000人，确保全年指标达成。

（二）实训阶段（3个月）

**理论课程：**

1.基础模块：数据标注规范、AI算法原理、工具操作（如Label Studio、CVAT）。

2.进阶模块：自动驾驶标注、医疗影像标注、自然语言处理标注等垂直领域技能。

3.职业素养：团队协作、项目管理、行业法规（如《数据安全法》）。

4.证书赋能：联合行业协会组织“人工智能训练师（中级）”“数据标注工程师”等资格认证考试，通过率目标≥85%。

 **实操课程：**

1.基础技能实操工具熟练度及训练标注工具实战

（1）图像标注:LabelImg、CVAT 工具进行矩形框标注(如车辆、人脸)、多边形标注(医疗影像病灶区)、语义分割(遥感图像)。

（2）文本标注:BRAT 工具完成实体识别(人名、地点)、情感分类、意图标注(客服对话)。

（3）音频标注:Praat 工具完成语音转写、声纹分类、情绪标签标注。

（4）数据清洗实战:处理模糊图像、噪声音频、缺失文本数据,学习数据去重与标准化流程。

2.质量控制训练

（1）分组互检:学员交叉审核标注结果,学习标注一致性评估(Kappa 系数计算)。

（2）错误案例库分析:针对常见标注错误(如边界模糊、标签错配)进行修正演练。

3.垂直领域专项训练

（1）自动驾驶标注

3D点云标注:使用 CloudCompare 标注障碍物、车道线,学习坐标系转换与 LiDAR 数据处理。

多传感器融合标注:同步标注摄像头图像与雷达数据,掌握时间戳对齐技术。

（2）医疗影像标注

DICOM 文件处理:标注 CT/MRI 中的肿瘤区域、骨骼结构,学习医学影像隐私脱敏规范。

病理切片标注:使用 QuPath 工具标注细胞核分级(如乳腺癌分级)。

（3）自然语言处理(NLP)标注

知识图谱构建:实体关系标注(如“药物治疗疾病”三元组提取)。

多语言标注:针对跨境电商场景的跨语种意图识别标注。

4.项目制实战

（1）企业真实项目模拟

分组承接虚拟企业订单(如标注 10 万张零售货架图像用于 AI 巡检系统),完整演练。

需求分析:理解客户标注规范文档(如标注粒度、标签体系)。

任务拆分:制定标注计划与分工,使用项目管理工具(如Jira)。

标注实施:按企业标准完成全流程,每日提交进度报告。

验收交付:生成标注报告,包括准确率、覆盖率等指标。

（2）标注项目管理实训

角色扮演:学员轮岗担任质检员、项目经理,学习标注成本核算(如工时/千张)、团队协作与风险管控。

（三）就业闭环

**定向输送**：实习考核合格者直接入职合作企业，签订1年以上劳动合同，为职工缴纳五险。

**就地就业**：依托汇智创新城标注1500人职场、经开区标注产业园区（拟改建京东科技西侧现有园区），可吸纳就业3000人。

# 四、资金拨付及考核验收

（一）一次性兑现30%人才培养金人民币壹佰伍拾万元整（￥1500000），此后根据实际就业人数，根据相关印证材料，每月兑现1500元/人，直至该批人员对应的剩余70%资金（7000元/人）兑现完毕。

（二）培训层面验收：每次开班时，及时完成每批培训学员信息系统录入工作（大同市数字产业公共服务平台）；学员培训期间有签到签退记录、教学录像等相关资料，每批学员教学培养完成后，及时提供培训资料和培训合同，包括但不限于学员名单、培训计划、培训实施情况的证明文件、培训合同（或实习就业合同）、培训合格证书等相关证明材料。就业层面验收：大同市数据标注企业职场就业，按月支付银行工资流水证明。

（三）由市数据局或其委托第三方对人才培养实习就业情况按照原协议规定的标准、流程进行分批验收，每批（500人以上）验收结果要形成综合评估报告，并根据实际验收情况公布验收报告。通过培训层面验收的，则市数据局首次支付的3000元/人培养经费归众梦云所有；根据就业人员稳岗月数兑现培养经费，每月兑现1500元/人，直至该批人员对应的剩余70%资金（7000元/人）兑现完毕（相应支付比例可以根据项目实际运营情况进行适当调整）。

（四）众梦云收到的每笔人才培养金，在收到资金后需进行公示，公示期不少于七天，并严格审批流程、追踪资金流向，确保资金专款专用，项目实施后配合审计部门监督审查。

（五）本项目周期结束后，达到每年2000人数据标注人才培养和就业目标，市数据局审核后公布整体审核验收报告。若未能完成目标任务验收，众梦云应在30天内完成整改并重新提交验收申请。若整改后依然未能完成提供每年2000名学员的目标任务，市数据局按照实际完成人数支付对应比例的资金。若已支付资金超出实际完成人数对应资金的，超出部分全额返还市数据局。

（六）市数据局应当对学员的培训及就业情况进行不定期不限次数的抽检调查，抽检结果将作为考核评估的重要依据；对已就业人员，需持续对其就业情况进行跟踪，阶段性形成人才就业情况报告，作为今后是否延续本合作模式的参考依据。