

武汉市数据局

关于组织开展武汉市 2026 年度高质量数据集建设任务专项申报工作的通知

各区（含开发区、东湖风景区）数据主管部门：

根据《武汉市促进人工智能产业发展若干政策措施》（武政规〔2025〕3号）相关要求，现开展我市 2026 年高质量数据集建设任务专项申报工作，有关事项通知如下。

一、建设目标

聚焦先进制造、生命健康、汽车制造与服务、光电子信息、数字创意、现代农业、现代商贸物流、工程设计、数据安全和网络安全、绿色环保、低空经济、智能机器人等 12 个专项领域（详见附件 1），建设一批产业亟需、模态丰富、质量过硬、富有行业特色的行业高质量数据集，丰富行业高质量数据集应用场景，研究建立行业高质量数据集联合共建、流通使用等创新机制，为武汉市推进人工智能创新高地建设夯实数据基础。

二、申报条件

（一）申报单位应为具有独立法人资格的组织机构（包含企、事业单位），具备在本市开展工作的必要基础，能够确保日常管理、信息联络与报送的及时准确。

（二）同一组织机构原则上不得同时申报多项建设任务专项。由多个法人单位共同进行联合攻关的项目，实行牵头单位负责制。鼓励

联合国国内数智经济头部企业合作申报。

(三) 申报单位应具有行业领先的数据汇聚能力、技术先进的数据治理能力,具有充足的资金投入、稳定的技术队伍,对申请的高质量数据集建设项目中运用的技术或产品,拥有明晰的知识产权。申报单位或联合体单位近三年无不良信用记录、财务会计制度健全,无重大违法行为。

(四) 申报单位对照建设任务专项要求填写申报材料,任务指标和完成时限原则上不得变更和调整。我局将根据《武汉市促进人工智能产业发展若干政策措施有关支持高质量数据集建设和数据产品利用条款资金管理办法(试行)》(以下简称《资金管理办法》)相关要求对申报项目进行遴选,及时向社会发布。

(五) 申报项目应为本市在建或拟建项目,原则上应于2026年10月31日前建成。

三、工作要求

(一) 有意向的申报单位应根据通知要求编制《高质量数据集建设任务专项申报书》(附件2),并准备相关佐证材料(包括单位资质证明、平台能力说明、数据集建设方案、合规承诺书等)。申报材料须真实、完整、规范,于申报截止日期前报送至所在区数据主管部门。纸质申报材料一式两份,胶装成册,电子扫描版(PDF格式)需与纸质版内容一致。

(二) 各区数据主管部门应结合本区实际,按照年度任务专项建设要求,组织指导辖区内符合条件的企事业单位做好申报工作。依据《资金管理办法》及本通知规定,对申报单位资格、申

报材料的完整性与合规性进行初审。通过初审的项目，请各区正式行文（附推荐表，附件3），将申报材料统一报送至市数据局。

（三）市数据局将根据各区组织申报情况，统筹推进项目评审及立项实施等后续工作。

（四）申报截止日期：2026年4月24日，逾期不予受理。

四、联系方式

联系人：张嵘钧 联系方式：65770039

邮箱：ssjjsjzlc@163.com

- 附件：1. 武汉市2026年高质量数据集建设任务专项清单
2. 武汉市2026年高质量数据集建设任务专项申报材料
3. 2026年高质量数据集建设任务专项申报单位推荐表



武汉市 2026 年高质量数据集建设任务专项清单

序号	建设任务专项名称	建设场景	任务专项内容
1	先进制造业高质量数据集	工业制造、科学研究、绿色低碳、智慧能源等	依托研发设计图纸、生产工艺参数、设备运行日志、质量检测数据、前沿技术创新等数据资源，建设多模态高质量数据集，打造智能维护、产品质量检测、能耗优化管理等典型应用场景。
2	生命健康行业高质量数据集	医疗卫生、科学研究等	依托临床标注数据、医学影像、电子病历、公共卫生监测数据、基因测序数据等数据资源，建设多模态高质量数据集，打造智能辅助诊断、流行病预测预警、个性化健康管理、药物研发辅助等典型应用场景。
3	汽车制造与服务行业高质量数据集	工业制造、交通运输、金融服务、科学研究、智慧能源等	依托车辆运行轨迹数据、生产制造工艺数据、供应链协同信息、用户驾驶行为数据、充电网络运行数据等数据资源，建设多模态高质量数据集，打造自动驾驶仿真训练、车辆故障预测预警、智能充电调度、出行服务等典型应用场景。
4	光子信息产业高质量数据集	工业制造、科学研究、智慧能源、绿色低碳等	依托光电器件测试数据、生产工艺参数、产品缺陷图像、光通信网络日志、行业应用场景数据等数据资源，建设多模态高质量数据集，打造生产过程智能控制、产品缺陷自动检测、性能优化仿真、行业应用解决方案等典型应用场景。
5	数字创意行业高质量数据集	文化旅游、教育教学、电子商务、数字创意、科学研究等	依托景区文创图像素材、视频资源、音频文件、用户交互数据、教育教学等数据资源，建设多模态高质量数据集，打造 AI 辅助内容生成、沉浸式体验构建、智能推荐系统、虚拟仿真教学等典型应用场景。

任务专项内容		
序号	建设任务专项名称	建设场景
6	现代农业行业高质量数据集	农业农村、气象服务、绿色碳、金融服务、电子商务、自然资源、科学研究等
7	现代商贸物流和金融行业高质量数据集	交通运输、金融服务、电子商务、人力资源等
8	工程设计行业高质量数据集	住房建设、应急管理、自然资源、交通运输、城市治理等
9	网络安全与网络安全行业高质量数据集	公共安全、城市治理、教育学、科学研究等
10	绿色环保行业高质量数据集	绿色低碳、气象服务、应急管理、自然资源、智慧能源等
11	低空经济行业高质量数据集	交通运输、应急管理、气象服务、自然资源、工业制造等
12	智能机器人行业高质量数据集	工业制造、科学研究、城市治理、应急管理、医疗卫生等

任务专项内容

依托气象观测数据、遥感影像、农作物生长数据、农产品市场信息等数据资源，建设多模态高质量数据集，打造作物生长监测预警、病虫害智能识别、农产品价格预测、农业保险精准定价等典型应用场景。

依托物流轨迹数据、金融交易流水、供应链订单信息、人才画像数据、市场供需信息等数据资源，建设多模态高质量数据集，打造智能路径优化、供应链金融风控、人才精准匹配、市场趋势预测等典型应用场景。

依托工程设计图纸、政策标准文档、地理空间信息、环境监测数据、水文灾害数据等数据资源，建设多模态高质量数据集，打造智能建筑方案生成、合规性自动审查、环境影响评估、城市数智化更新等典型应用场景。

依托网络安全公开报告、漏洞库信息、安全事件案例、网络流量特征样本、数据脱敏样本、合规政策文档等数据资源，建设多模态高质量数据集，打造网络威胁情报分析、安全态势感知、数据分类分级、隐私合规检查、数据交易合规评估等典型应用场景。

依托环境监测数据（大气、水质、土壤）、气象数据、碳排放核算数据、新材料及能源数据、水文灾害数据等数据资源，建设多模态高质量数据集，打造环境质量预测预警、碳排放智能核算、新能源功率预测、新材料性能模拟与应用等典型应用场景。

依托无人机飞行轨迹数据、空域管理信息、遥感影像、气象观测数据、地理空间数据等数据资源，建设多模态高质量数据集，打造低空航线智能规划、飞行安全实时监控、空中识别、低空物流调度等典型应用场景。

依托机器人运动控制数据、环境感知信息、人机交互记录、任务执行日志、高精度空间建模数据等数据资源，建设多模态高质量数据集，打造自主导航避障、多机协同作业、智能人机交互、特种作业执行等典型应用场景。

附件 2

武汉市 2026 年高质量数据集建设任务专项 申报材料

申报场景：_____

申报项目名称：_____

申报单位：（加盖单位公章）

推荐单位：_____

申报日期： 年 月 日

填报须知

一、申报单位应仔细阅读《武汉市 2026 年高质量数据集建设任务专项清单》的有关说明，如实、详细地填写每一部分内容。

二、除另有说明外，申报表中栏目不得空缺。

三、申报主体所申报的项目需拥有知识产权，对报送的全部资料真实性负责，对能否按计划完成重点任务作出有效承诺，并签署企业承诺声明（见“申报项目承诺书”模板）。

2026 年武汉市高质量数据集建设任务申报表

一、申报单位基本信息				
单位名称 (全称)		统一社会信用代码		
法定代表人		注册资本 (万元)		
单位负责人		联系电话		
单位性质	<input type="checkbox"/> 国有企业 <input type="checkbox"/> 民营企业 <input type="checkbox"/> 外资企业 <input type="checkbox"/> 事业单位 其他(请注明):			
单位地址				
单位代码				
整体业务收入 (万元)	指上一个财年(提供证明材料)	研发投入 (万元)	指上一个财年(提供证明材料)	
联系人	姓名		手机	
	职务		邮箱	
联合单位				
二、项目基本情况				
专项领域	<input type="checkbox"/> 先进制造 <input type="checkbox"/> 生命健康 <input type="checkbox"/> 汽车制造与服务 <input type="checkbox"/> 光电子信息 <input type="checkbox"/> 数字创意 <input type="checkbox"/> 现代农业 <input type="checkbox"/> 现代商贸物流与金融 <input type="checkbox"/> 工程设计 <input type="checkbox"/> 绿色环保 <input type="checkbox"/> 数据安全和网络安全 <input type="checkbox"/> 低空经济 <input type="checkbox"/> 智能机器人			
项目名称				
建设周期	202X 年 XX 月-XX 年 XX 月	建设类别	<input type="checkbox"/> 在建 <input type="checkbox"/> 拟建	
总投资(万元)		资金来源		
建设背景与意义(300字以内)	<p>要求: 阐述目标领域内数据资源的现状、核心瓶颈(如稀缺性、质量低、标准化不足等)及其对技术或产业发展的制约,说明本数据集建设的紧迫性与广泛价值。</p> <p>示例: 当前工业质检环节存在三大痛点: 缺陷样本采集难,年人工采集费超 X 万元; 标注标准不一,模型复用率不足 X%; 长尾缺陷漏检率高达 X%,年质量损失 X 万元。本项目响应“数据要素×工业制造”政策,旨在构建覆盖多类缺陷、标注规范的工业视觉数据集,为产线智能化升级提供核心</p>			

	数据支撑。
建设内容 (300字以内)	<p>要求：严格参照全国数据标准化技术委员会（SAC/TC609）发布的 TC609-5 系列技术文件（特别是《高质量数据集 建设指南》与《高质量数据集 质量评测规范》），应遵循“需求驱动、规划先行、过程可控、迭代优化”的原则，全生命周期包括六个核心阶段：数据需求、数据规划、数据采集、数据预处理、数据标注、模型验证。</p> <p>示例： 工业缺陷检测高质量数据集</p> <p>采集：通过工业相机在 X 条产线部署实时拍摄，同步采集 OK/NG 品图像，年采集量 X 万张；引入主动学习算法自动筛选难例样本，减少无效采集 X%。</p> <p>加工：标注团队依据《工业产品表面缺陷自动检测系统技术要求》进行像素级分割标注，采用预标注模型辅助，人工复核量减少 X%；缺陷类别涵盖划痕、脏污、毛刺等 X 类，标注一致性达 X%。</p> <p>管理：依托高质量数据集平台，记录采集参数、标注版本、模型迭代信息；编写数据说明文档，含缺陷分布、光照条件、使用限制等。</p> <p>标准化：遵循《高质量数据集 质量评测规范》，经测试完整性 X%、标注准确率 X%、多样性覆盖 X 类场景，形成高质量数据集规模达 XTB，支撑 X 个模型及智能体训练。</p>
预期成果 (300字以内)	<p>要求：聚焦数据集产生的直接价值与可量化的经济社会效益，表述具体、明确，尽可能避免描述性介绍。</p> <p>示例： 经济效益：年节约采集成本 X 万元，人工标注成本降低 X%；模型跨产线复用周期从 X 月缩至 X 周，研发效率提升 X%；数据服务年收入 X 万元，项目 ROI 约 X%。</p> <p>社会效益：在 X 家制造企业试点应用，覆盖 X 条产线，年减少不良品流出 X 万件，节约质量损失 X 万元；带动供应链质检效率提升 X%。</p> <p>行业带动：牵头（或参与）制定《工业缺陷数据集标注规范》团标（或行标），举办 X 场技术培训或对接会（覆盖 X 人次、X 家企业），开源标注工具、开源数据集获 X 家企业采用。</p>

（注：此表由任务项目建设单位填写后，报送所属区数据主管部门汇总）

武汉市 2026 年高质量数据集建设报告 编制指南

一、申报依据

1. 建设背景及意义

阐述目标领域内数据资源的现状、核心瓶颈（如稀缺性、质量低、标准化不足等）及其对技术或产业发展的制约，说明本数据集建设的紧迫性与广泛价值。

2. 已有技术积累和应用条件

针对建设任务专项，项目单位情况已有的技术条件和相关应用成果等。

二、任务要求

该板块为项目书主体内容，各建设任务专项申报单位围绕任务要求进行编制，确保项目满足行业高质量数据集内容建设、技术标准建设、设施化应用、生态体系支撑、数据安全保障、经济社会效益突出等相关要求。

（一）数据集内容建设

以高价值场景需求为牵引，全面梳理自有数据，并充分依托外部数据，建设一批产业亟需、规模庞大、模态丰富、质量过硬、行业特色的数据集，形成可持续迭代的行业数据集供给体系。建成的高质量数据集应按需及时更新，确保时效性和适用性，有效提升人工智能模型性能。

（二）技术标准建设

验证全国数标委研制的《高质量数据集建设指南》《高质量数据集格式要求》《高质量数据集分类指南》等相关国家标准和技术规范。按照《高质量数据集质量评测规范》对所建数据集进行质量评估，建成的行业高质量数据集说明文档完整，包括数据集的模态类型、数据规模、应用场景、建设过程等。

（三）价值化探索

支持高质量数据集在国家数据流通利用基础设施上登记、上架、交易、交付。建成的行业高质量数据集应服务不少于3个应用场景，服务人工智能模型数量不少于3个。

（四）生态体系支撑

以打造互利共赢、协同创新、开放共享的数据生态为导向，积极推动跨机构、跨层级、跨领域的行业数据资源整合，探索建立高质量数据集联合研发与共建共享机制，围绕高价值应用场景牵引产业链协同突破。鼓励建设单位在符合法律法规的前提下，通过技术培训、资源对接等形式培育产业生态，将高质量数据集向开源社区开放，助力行业人工智能技术成果的广泛落地与普惠应用。

（五）数据安全保障

构建覆盖数据集全生命周期的安全防护体系，严格落实网络安全法、个人信息保护法等法规要求，对数据进行分级分类管理，实施严格的访问控制，并采用匿名化等技术手段对个人隐私等敏感信息进行深度处理。建立内容审核机制，及时过滤有害、虚假、歧视性及违背公序良俗的信息，确保数据集内容符合社会主义核

心价值观。完善版本管理流程，详细记录迭代内容、责任人及时戳，实现全过程可追溯。强化数据存储、传输及边界防护能力，定期开展安全审计、漏洞扫描与备份恢复演练，保障数据持续可用、完整且安全。

（六）经济社会效益

所建高质量数据集应紧密服务国家和区域发展战略，切实赋能实体经济与社会治理，确保取得显著且可验证的综合效益。项目应通过数据赋能，在产业增收节支、效率提升、模式创新等方面产生明确的经济价值；同时应具备较强的产业带动性，助力培育新业态与产业链协同；还应积极回应民生与服务需求，在公共服务、安全保障、绿色发展等领域彰显社会价值。项目申报时需明确预期效益的具体指标与测算依据。

三、任务核心指标

建设任务专项申报单位请根据本单位实际情况填报行业高质量数据集建设任务专项指标表。

表 1 高质量数据集建设任务专项指标表

序号	一级指标	二级指标	指标要求/响应情况		备注
			类型	规模	
1	数据集建设	数据集类别及规模			
		数据集质量	<input type="checkbox"/> 预训练集		≥ 600 亿 tokens
			<input type="checkbox"/> 指令数据		≥ 5 万条
			<input type="checkbox"/> 偏好数据		≥ 4 万组
		<input type="checkbox"/> 基准测试数据		≥ 5000 条	

2	技术标准建设	数据采集、处理、管理、标注、合成、质检、服务运营等环节自动化工具数量	单位：个 (包括：明确自动化工具类型)	≥ 20 个
		平台数据采集、汇聚、清洗、标注、生成、评测、封装等核心功能模块数量	单位：个 (包括：明确标准名称)	≥ 5 个
3	价值化探索	赋能应用场景数量	单位：个 (明确具体场景)	≥ 3
		服务人工智能模型及智能体数量	单位：个	≥ 3
		服务市场主体数量	单位：个	≥ 5
4	生态体系支撑	参与项目相关领域标准制定	单位：个	
		通过技术培训、对接会等形式培育产业生态	单位：场次	
5	经济社会效益	分经济效益、社会效益两方面，提出具体、明确、可量化的指标。		
6	数据安全	需建立覆盖行业高质量数据集建设全生命周期的安全管控体系，设备国产化率不低于 80%。		

四、项目实施

1. 项目负责人及团队成员履历及能力评价
2. 项目实施的可行性分析
3. 项目实施计划（项目进度计划及关键里程碑节点）
按月列出计划进度和关键的、必须实现的节点目标。
4. 投资构成

五、项目验收

1. 验收内容（请根据“高质量数据集建设任务专项指标表”及单位自身实际情况，在表 1 提出的共性指标外，结合项目特点，明确用于验收的其他关键指标内容）；

2. 验收方式（请根据《资金管理办法》，明确拟以何种方式对“高质量数据集建设任务专项指标表”内容进行验收）。

六、经费预算

项目实际投入经费预算表

科目名称	所需经费	具体内容及计算依据
经费预算（合计）		
1. 软件开发费		
2. 硬件设备费		
3. 能源材料费		
4. 测试加工费		
5. 劳务费		
6. 专家咨询费		
7. 其他费用		

注：计算依据可另附说明

申报单位相关证明材料

1. 主体资格与合规类：承诺书、营业执照（事业单位法人证书）、申报单位近3个月的纳税证明或无欠税证明、上一年度财务审计报告、信用中国信用记录，其法定代表人无重大违法记录声明函（加盖公章）。

2. 创新成效与成果类：数据服务与应用成效证明（如采购合同、用户应用证明、应用案例、成果报告）、与数据集建设相关的技术成果与创新水平证明（如专利、软件著作权、标准制定文件、论文检索报告、科技奖励、竞赛奖励等）。

3. 建设投入与保障类：项目核准或备案文件、项目投入证明（已开工项目提供相应采购合同及已完成投资的财务凭证，未开工的提供自筹资金证明）等。

4. 其他有助于证明项目质量的材料：如项目人员资质、成果鉴定等，由申报单位根据实际情况选择提供。

项目单位的申报材料应装订成册，封面加盖单位公章，侧面加盖骑缝章。

申报项目承诺书

根据《关于征集武汉市 2026 年度高质量数据集建设任务专项的通知》有关要求，我单位提交了项目参评。

现就有关情况承诺如下：

1. 我单位对所报送的全部资料真实性负责，保证所报送的项目拥有知识产权，符合国家有关法律法规及相关产业政策要求。
2. 我单位所报送的项目符合国家保密规定，未涉及国家秘密、个人隐私和其他敏感信息。
3. 相关材料中的文字和图片已经我单位审核，确认无误。
4. 支持基于项目建成的高质量数据集在国家数据流通利用基础设施上登记、上架、交易、交付。

我单位将根据《关于组织开展武汉市 2026 年度高质量数据集建设任务专项申报工作的通知》要求，增强大局意识，切实承担主体责任，若申报成功，将在项目实施期间认真组织、重点推进、加强保障，全力完成高质量数据集建设，力求 2026 年 X 月前达到或超过预期目标。

我单位对违反上述承诺导致的后果承担全部法律责任。

联系人：

联系电话：

法定代表人：（签字）

公司（企业盖章）

2026年 月 日

附件 3

2026 年高质量数据集建设任务专项申报单位推荐表

推荐单位（盖章）：

序号	单位名称	专项领域	申报项目	主要内容	项目联系人	联系方式
1						
2						
3						

- 注：1. 本表由各区（开发区、风景区）数据主管部门填报；
2. 推荐单位按优先次序排名；
3. 专项领域是指武汉市 2026 年高质量数据集建设任务专项清单中（附件 1）涉及的领域。

