

山西省教育厅

晋教科信函〔2026〕52号

山西省教育厅关于组织 征集行业高质量数据集工作的通知

各高等学校：

按照《山西省数据局关于公开征集行业高质量数据集工作的通知》要求，现组织开展高质量数据集征集工作。请各校对照征集要求，认真填写《山西省高质量数据集申报书》和《高质量数据集汇总表》（详见附件），于3月25日前将电子版（WORD版和PDF盖章扫描版）发送至省教育厅科学技术与信息化处 kjc@sxedc.com。

附件：山西省数据局关于公开征集行业高质量数据集工作的通知



2026年3月18日

山西省数据局

关于公开征集行业高质量数据集工作的通知

省有关单位，各市数据局、山西转型综改示范区管委会：

为深入贯彻落实国家“人工智能+”战略部署，加快培育一批行业高质量数据集，强化数据要素高效供给，树立行业应用标杆，充分发挥示范引领作用，赋能我省数字经济高质量发展，现组织开展行业高质量数据集征集工作。有关事项通知如下：

一、征集要求

(一) 征集领域。主要面向但不限于科学研究、工业制造、农业农村、智慧能源、交通运输、金融服务、医疗卫生、教育教学、商务领域、人力资源、文化旅游、应急管理、气象服务、绿色低碳、公共安全、城市治理等**重点领域**，以及低空经济、智慧驾驶、具身智能、生物制造等**创新领域**的高质量数据集。

(二) 申报主体。在山西省内注册的独立法人单位，包括政府部门、企事业单位、科研院所等。申报单位近三年经营状况良好，在信用、质量、安全等方面无不良记录。涉及多家单位的，可结合实际情况联合申报。

（三）数据集质量要求。行业高质量数据集是指经过系统化的采集、汇聚、清洗、标注等全流程数据处理，可直接用于开发和训练人工智能模型，并能有效提升模型性能的行业数据集合。根据数据所蕴含知识类型，可分为行业通识数据集和行业专识数据集。数据集可包含结构化数据、文本、视频、音频、图形图像等多种模态，数据规模原则上应达到 TB 级。数据集应为已建成或在建项目，具备实际业务应用场景，并在技术先进性、模式创新性、应用示范性以及动态更新机制等方面表现突出。

二、组织实施

请省有关单位、各市数据局、山西转型综改示范区管委会（以下简称“推荐单位”）认真组织申报单位对照本通知要求进行填报。

（一）申报材料

1. 《山西省高质量数据集申报书》（附件 1）

申报书由基本信息表和申报材料正文两部分组成，申报单位须对所填内容的真实性负责，并加盖单位公章。

2. 《XXX 行业高质量数据集汇总表》（附件 2）

各推荐单位在对申报材料进行初审后，按附件 2 格式汇总，形成本单位统一的推荐清单。

（二）支持措施

对入选的具有示范带动作用的高质量数据集，省数据局将在

供需对接、资源协调、场景应用等方面给予支持，并择优推荐至国家数据局。同时，推动其入驻国家高质量数据集管理服务平台，开展数据资源登记，试点数据产权登记，支持在数据交易中心上架，多渠道实现推广应用与价值释放。

三、报送要求

申报单位应于2026年3月20日（星期五）前将附件1报送至所属推荐单位。各推荐单位经初审汇总形成附件2，并于2026年3月23日（星期一）前将全部附件1与附件2的电子版统一打包发送至指定邮箱：szkjhjcssjsc@163.com。

联系人：冯丹妮 0351-7676117

- 附件：1. 山西省高质量数据集申报书
2. XXX行业高质量数据集汇总表



附件 1

山西省高质量数据集申报书

(模板)

数据集名称：_____

申报单位（盖章）：_____

申报日期：_____年_____月_____日

基本信息表

申报单位：（盖章）

年 月 日

板块	字段	填写说明与示例
1. 基础信息	数据集名称	填写数据集完整、规范的名称，如“高速公路视频车流检测数据集”。
	所属领域	重点领域： <input type="checkbox"/> 科学研究 <input type="checkbox"/> 工业制造 <input type="checkbox"/> 农业农村 <input type="checkbox"/> 智慧能源 <input type="checkbox"/> 交通运输 <input type="checkbox"/> 金融服务 <input type="checkbox"/> 医疗卫生 <input type="checkbox"/> 教育教学 <input type="checkbox"/> 商务领域 <input type="checkbox"/> 人力资源 <input type="checkbox"/> 文化旅游 <input type="checkbox"/> 应急管理 <input type="checkbox"/> 气象服务 <input type="checkbox"/> 绿色低碳 <input type="checkbox"/> 公共安全 <input type="checkbox"/> 城市治理 创新领域： <input type="checkbox"/> 低空经济 <input type="checkbox"/> 智慧驾驶 <input type="checkbox"/> 具身智能 <input type="checkbox"/> 生物制造 <input type="checkbox"/> 其他：_____
	建设/牵头单位	填写数据集的主要建设单位或产权归属单位全称。
	当前状态	<input type="checkbox"/> 已建成 <input type="checkbox"/> 在建（预计完成时间：_____）
2. 数据内容	核心数据描述	用 1-2 句话简要说明数据的主要内容，例如：“包含全省主要路口 2019-2023 年车辆通行量、车型、速度等结构化数据。”
	数据规模与模态	规模：如“100TB”、“500 万条记录”等。 模态： <input type="checkbox"/> 文本 <input type="checkbox"/> 图像/视频 <input type="checkbox"/> 音频 <input type="checkbox"/> 时序数据 <input type="checkbox"/> 结构化表格 <input type="checkbox"/> 其他：_____
	数据来源与更新频率	来源： <input type="checkbox"/> 自营业务产生 <input type="checkbox"/> 采购/合作获取 <input type="checkbox"/> 公共数据 <input type="checkbox"/> 其他：_____ 更新频率： <input type="checkbox"/> 实时 <input type="checkbox"/> 每日 <input type="checkbox"/> 每月 <input type="checkbox"/> 季度 <input type="checkbox"/> 每年 <input type="checkbox"/> 不更新

板块	字段	填写说明与示例
3. 质量水平	加工处理情况	简要说明是否经过清洗、标注、脱敏等处理。示例：“已完成数据清洗和车牌、人脸信息脱敏，并对车辆类型进行了人工标注。”
	遵循的标准规范	可填写遵循的国家/行业/企业标准，如“参考国家《高质量数据集 格式要求》”。若不清楚可填“暂无”。
4. 应用情况	主要应用场景	说明数据集可用于解决什么问题，例如：“用于训练交通流量预测 AI 模型，以优化信号灯配时。”
	是否已用于 AI 训练	<input type="checkbox"/> 是（请注明模型类型/项目名称：_____） <input type="checkbox"/> 否
	已产生的成效或价值	可量化或定性描述，如：“在某试点路段将通行效率提升了约 15%。”或“支撑了 XX 智能诊断系统的研发。”
5. 合规与安全	合规安全说明	简要说明数据获取、处理的合规性及安全措施，例如：“数据经匿名化处理，已通过内部安全评审，不包含个人隐私信息。”
6. 联系信息	负责人及联系方式	姓名：_____ 电话：_____ 邮箱：_____

申报材料正文

(总字数建议不超过 3000 字，可附图表)

一、建设背景

简述数据集建设的背景与目的，说明拟解决的主要问题或支撑的业务需求。另介绍已有工作基础，以及建设、运营、数据提供等主要参与单位与分工。

二、数据集建设情况

全面描述数据集的基本情况，包括：数据集所属领域、当前状态；数据的主要内容、规模与模态；数据来源渠道及更新维护机制；数据加工处理情况；建设过程中参照的相关标准规范。

三、应用成效

说明数据集已赋能的具体业务场景，支撑了哪些人工智能模型或系统的研发训练，并介绍取得的实际成效，须用量化指标(如效率提升比例、模型性能改进幅度)体现，可辅以定性描述(如支撑某系统上线、优化某业务流程)。

四、创新点

阐述数据集在采集、处理、标注、应用等环节的技术创新或模式创新之处，突出与同类数据集的差异化优势。

五、合规与安全保障

说明数据集在获取、处理、使用过程中的合规性，包括数据授权、匿名化处理、隐私保护等情况；介绍在数据安全方面采取的主要技术措施和管理机制。

附件 2

XXX 行业高质量数据集汇总表

推荐单位：（盖章）

年 月 日

序号	数据集名称	所属领域	建设单位	数据集规模	数据集模态	服务模型名称	应用场景名称	联系人及联系方式	备注
1									
2									
3									
...									

（此件依申请公开）

山西省教育厅办公室

2026年3月19日印发
