

面向城市交通治理与出行服务的高质量数据集架构、构建方法与应用

张群洪, 刘浩

福州市数字福州集团有限公司, 福建 福州 350009

摘要

在数据要素市场化配置和“人工智能+交通运输”加快推进的背景下, 城市交通高质量数据集已成为提升交通治理能力、优化公众出行服务和释放数据要素价值的重要基础。本文以福州市城市交通相关实践为例, 围绕治理与服务双场景, 分析区域交通高质量数据集建设面临的数据分散、标准不一和场景衔接不足等问题。研究认为, 区域交通高质量数据集建设应以多源异构数据汇聚为基础, 以标准化治理、主题建模、质量评价、安全治理和可信流通为关键路径。福州案例表明, 围绕“人、车、路、站、线”构建对象关系和主题数据组织机制, 能够推动交通数据由分散资源向治理效能、服务效能和数据产品化能力转化, 为同类城市推进交通数据要素高质量建设提供参考。

关键词

城市交通; 高质量数据集; 数据要素; 交通治理; 出行服务; 数据治理

中图分类号: TP315

文献标志码: A

Architecture, construction methods, and application of high-quality datasets for urban transport governance and mobility services

ZHANG Qunhong, LIU Hao

Fuzhou Digital Fuzhou Group Co., Ltd., Fuzhou 350009, China

Abstract

Against the backdrop of the accelerated market-oriented allocation of data elements and the advancement of “AI Plus Transportation,” high-quality urban transport datasets have become an important foundation for improving transport governance, optimizing public mobility services, and unlocking the value of data elements. Taking relevant urban transport practices in Fuzhou as a case, this study examined the problems encountered in developing high-quality regional transport datasets for governance and service scenarios, including fragmented data, inconsistent standards, and

inadequate alignment with application scenarios. The analysis suggests that high-quality regional transport datasets should be built on multisource heterogeneous data integration, with standardized governance, thematic modeling, quality assessment, security governance, and trusted circulation as key approaches. The Fuzhou case showed that establishing entity relationships and a thematic data organization mechanism centered on people, vehicles, roads, stops, and routes can transform fragmented transport data resources into governance effectiveness, service effectiveness, and data product development capacity. These findings provide a reference for comparable cities seeking to promote the high-quality development of transport data elements.

Key words

urban transport, high-quality dataset, data element, transport governance, mobility service, data governance

0 引言

2025年6月，交通运输部在“十五五”智慧交通发展专家座谈会上指出，要把服务支撑交通强国建设作为智慧交通发展的基本方向，将人工智能和数据要素驱动作为重要机遇，高质量做好“十五五”交通运输规划编制工作。这一部署并非孤立提出，而是与近年来国家层面对数据要素与交通运输融合发展的政策导向一脉相承。2024年，国家数据局、交通运输部等17部门联合印发《“数据要素×”三年行动计划（2024—2026年）》，在“数据要素×交通运输”领域明确提出提升数据供给水平、优化数据流通环境、加强数据安全保障，并支持交通运输高质量数据集建设和复用^[2]。在此背景下，高质量数据集已不再只是面向模型训练的技术资源，而是成为交通行业提升数据供给质量、支撑场景应用深化、推动数据要素价值释放的重要基础。

城市交通是综合交通体系中数据生成最密集、业务协同最复杂的领域之一。公交、地铁、出租、网约车等信息持续积累，已经形成较为丰富的数据基础。与此同时，

城市交通治理和出行服务又高度依赖跨主体、跨方式、跨空间协同，涉及交通、文旅、人社、公安等多部门，也涉及公交企业、互联网平台和场站运营方等多类主体。数据分散、标准不一、口径不统一、授权边界不清等问题，直接影响运行监测、拥堵治理、换乘衔接、重点群体服务和重大活动保障等高频场景的实施效果。

基于上述背景，本文以福州市城市交通领域相关实践为例，以治理与服务双场景为主线，依次讨论区域交通高质量数据集的建设需求、总体架构、关键路径和实践成效。本文围绕数据分散、规则不一、质量波动、安全风险较高以及场景衔接不足等交通领域问题展开回应，力求形成从问题识别到建设实施再到场景验证的完整闭环，为同类城市推进交通数据要素高质量建设提供参考。

1 城市交通数据集建设需求

在“人工智能+”和数据要素开发利用持续推进的背景下，高质量数据集正在成为行业数字化和智能化建设的基础能力。已有研究认为，高质量数据集不仅要具备较高的应用价值，还应具备较强的知识密

度、技术含量和可持续演进能力，并通过体系规划、工程建设和质量检测等环节形成稳定的数据资源体系^{[3][4]}。进一步看，数据集建设正在从内部资源建设走向产品化生产，高质量数据集需要同时具备质量可评估、价值可流通、合规可监管和内容可理解等特征，并通过产品规划、数据获取、清洗标准化、标注增强、质量检测和封装等环节形成可复用的数据产品形态^{[5][6][7]}。与传统以系统建设和数据接入为主的路径相比，高质量数据集更加关注对象关联建模、标准统一治理、质量持续评估和安全可信流通。建设重点因此从“接入了多少数据”逐步转向“是否形成了可复用、可评估、可运营的数据集体系”。

交通领域已有研究主要围绕多源特征识别、出行链构建、运行评估和专题化应用展开。多源交通数据的联合应用能够扩展城市交通特征分析的对象和尺度^[3]；融合轨道交通、快速公交、常规公交IC卡和车辆定位数据，可构建多模式公共交通出行链^[8]；融合设施、手机位置、车辆定位、IC卡和路径规划等数据，则可从设施能力、运营服务和竞争力等方面开展公共汽电车运营评估^[9]。与此同时，交通物流、基础设施安全监测、交通基建知识服务、全域多模态城市治理和交通智能分析等方向也已出现高质量数据集实践^{[10][11][12][13][14]}。这些进展说明，交通数据集建设正在从单纯汇聚转向围绕场景组织和供给数据。

不过，上述现有成果仍存在三方面方法论局限。第一，分析单元多以单一运输方式、单一业务平台或单项算法任务为主，数据组织依附既有系统边界，较少回答区域尺度上不同交通方式和城市部门如何形成统一对象体系^{[3][8][9]}。第二，研究重点多放在数据融合、模型识别或应用效果，对

数据供给责任、目录与接口协同、授权边界、质量问题处置和成本分担等跨部门机制讨论不足，难以解释多主体数据如何由一次性接入转为稳定供给^{[10][11][12]}。第三，部分案例以平台上线或场景展示作为终点，缺少质量评价、业务反馈、版本管理和安全审计相衔接的持续运营设计^{[13][14]}。综合来看，上述局限在区域级交通数据集建设中集中表现为顶层设计和跨主体协调两个方面：在顶层设计层面，尚未形成从区域治理目标出发，统筹核心对象、主题数据、应用场景和评价反馈的完整方法，难以支撑跨运输方式、跨管理层级的数据资源配置；在跨主体协调层面，尚未将供数责任、目录维护、接口变更、质量工单、授权审批和违规追责等事项转化为部门与企业共同遵循的运行机制，难以保障多源数据由项目式接入转为持续、稳定供给。

安全与授权问题同样制约着数据集建设深化。城市交通数据涉及个人出行、交易记录、运营信息和监管信息，跨部门共享、政企协同和授权运营过程中，仍需进一步明确分类分级、授权责任、隐私保护和审计追踪机制^[15]。此外，一些建设工作仍停留在“重接入、轻组织，重展示、轻运营”的阶段，数据建设与高频场景之间衔接不够紧密，尚未形成围绕治理和服务需求持续演进的数据集产品。

针对上述不足，本文补充四方面认识：区域交通高质量数据集应以治理与服务双场景统筹数据需求，而不是按既有系统或单一运输方式机械切分；应围绕“人—车—路—站—线”等核心对象建立跨系统关联，而不是停留在数据表的并列汇聚；应以明确的供数、用数、质量和安全责任支撑跨主体常态化协同；应把质量评价、场景反馈、版本管理和安全审计纳入全生命周期运营。后文的总体框架和关键路径据

此展开。

2 数据集总体架构

2.1 建设目标

区域交通高质量数据集建设的目标，是在区域层面统筹数据资源，围绕场景组织数据供给，以治理机制保障数据可用，以持续运营促进数据集迭代。其关键不在于新增一个孤立平台，而在于将分散在不同部门、系统和主体中的交通数据，转化为能够持续支撑治理研判、协同调度、服务优化和复用扩展的数据供给能力。据此，本文提出“场景牵引建设、数据集反哺应用、制度与技术共同保障持续运营”的总体思路：高频刚需场景决定建设重点，主题数据集支撑场景迭代，标准规范、质量管理、安全治理和运营维护使数据集由一次性建设成果转化为可持续运行的数据供给体系。

2.2 架构组成

面向城市交通领域，本文将总体架构概括为“一个数据底座、一个保障体系、五类主题数据集、两类核心场景”，如图1所示。其中，数据底座解决数据从哪里来和如何关联的问题，保障体系解决数据如何持续可管可控的问题，主题数据集解决数据如何围绕业务需求重新组织的问题，两类核心场景则承接数据应用并检验建设成效。各部分通过“数据汇聚—治理加工—主题组织—场景应用—反馈优化”形成动态闭环：数据底座汇聚多源数据，在保障体系支撑下完成治理加工后形成五类主题数据集，再根据具体任务组合支撑治理

和服务场景；场景运行中发现的数据缺失、更新延迟和关联错误等问题，反馈至数据治理标注平台，推动治理规则、数据标签和数据版本持续更新。

具体来说，“一个数据底座”是指面向区域交通的基础数据底座，主要承载多源接入、资源编目、时空对齐、对象映射等能力。底座数据既包括公交、地铁、出租等行业数据，也包括在合规前提下可关联的文旅、气象和公安等跨部门数据。其作用不在于简单汇聚，而在于为主题数据集加工提供统一、稳定、可追溯的数据来源。

“一个保障体系”是指贯穿建设全流程的标准规范、质量管理、安全治理和运营维护机制。数据治理标注平台是这一体系的核心支撑，统一承载元数据管理、对象映射、规则配置、样本标注、人工复核、质量检测和问题闭环等能力，使标准、质量、安全和运营要求能够在同一平台上协同落地。

“五类主题数据集”如表1所示，分别对应基础资源、运行监测、出行服务、协同调度和多模态感知五类核心能力，支撑基础底图构建、运行态势感知、全过程出行服务、联动指挥调度以及智能感知分析等任务。与按系统存储的传统方式相比，主题数据集强调围绕对象关系和场景需求组织数据，使数据能够直接支撑治理和服务应用。基础资源数据集主要解决底图和主数据统一问题，运行监测数据集主要支撑实时态势感知，出行服务数据集主要支撑服务链路贯通，协同调度数据集主要支撑专题联动和处置闭环。安全治理相关要求不再单列为主题数据集，而是纳入保障体系统一实施。前四类数据集主要沉淀结构化业务数据及其关联关系，多模态感知数据集则是在这一基础上补足视觉感知和智能分析能力。

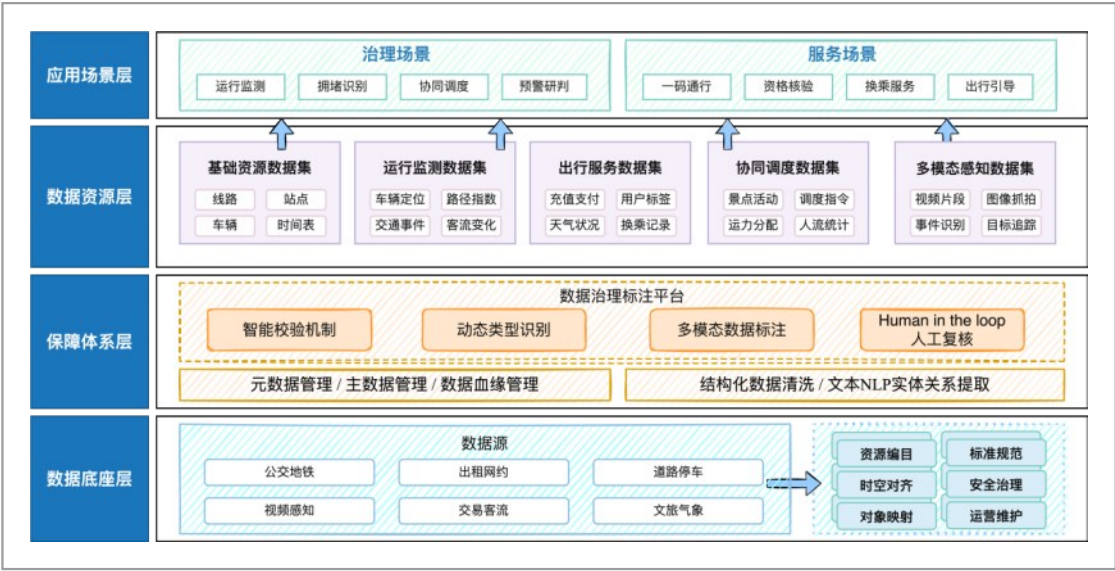


图1 城市交通高质量数据集总体架构

表1 五类主题数据集建设重点

主题数据集	建设重点	典型数据内容	主要支撑场景	数据量/TB
基础资源数据集	统一静态底图和主数据	线路、站点、枢纽、场站、路段、停车场、车辆、设备设施、行政区划、空间网格	资源编目、对象映射、基础底图构建	0.9
运行监测数据集	沉淀动态运行状态	车辆定位、路况指数、交通事件、信号状态、到离站信息、客流变化、视频识别结果、拥堵告警	运行监测、态势感知、预警分析	2.7
出行服务数据集	沉淀全过程出行服务链路	支付交易、一码通行记录、资格标签、换乘记录、到站预测、停车余位、服务反馈	一码通行、资格核验、换乘服务、重点群体保障	2.3
协同调度数据集	形成面向专题场景的联动数据	重点区域客流、重大活动保障数据、应急资源、运力调配记录、调度指令、联动处置结果	联动指挥、公交地铁协同、应急处置、重大活动保障	1.3
多模态感知数据集	沉淀可用于识别分析的图片、视频和标注样本	视频流片段、图像抓拍、目标框标注、事件识别结果、车牌车型识别结果、站台人流统计结果	事件分析、人流统计、视频巡检、识别模型训练	40.0

以多模态感知数据集为例，其主要承接城市交通中持续增长的视觉感知和智能分析需求，服务于治理场景中的事件发现、视频巡检、重点区域客流识别和辅助研判，在底座、保障体系和五类主题数据集

另一方面也能够为公众服务侧的人流统计、站区引导和模型能力迭代提供支撑，其核心字段示例如表 2 所示。基础上，数据集建设最终仍需落到具体业

表2 多模态感知数据集示例字段

字段类别	示例字段	字段说明
基础标识	数据片段标识、设备标识、点位标识、任务编号、样本编号	用于关联视频片段、图像样本、采集设备和标注任务
时间空间	采集时间、开始时间、结束时间、经度、纬度、行政区、线路编码、站点编码	用于还原图片视频的时空位置和所属交通对象
模态信息	数据类型、视频时长、帧率、分辨率、图片格式、采集视角、光照条件	用于描述多模态样本属性和可用性
标注信息	目标框坐标、目标类别、行为标签、事件标签、标注人员、标注时间、复核结果	用于支撑目标检测、行为识别和事件分析任务
识别结果	客流计数、车流计数、车牌识别结果、车型识别结果、事件置信度、拥堵判定结果	用于支撑人流统计、事件发现和运行研判
质量管理	样本质量等级、遮挡情况、模糊程度、缺帧标识、可用性标识	用于支撑样本筛选、质量评估和模型迭代

务场景中检验其有效性。“两类核心场景”是指治理场景和服务场景。治理场景以综合交通运行监测调度中心（TOCC）为主要承载，重点面向运行监测、拥堵识别、事件预警、公交地铁协同、重大活动保障和应急联动；服务场景以“榕城通/e福州”为主要承载，重点面向一码通行、换乘服务、资格核验、重点群体服务、支付协同和公众出行引导。场景不仅用于验证数据集建设成效，也会持续反向提出数据采集、治理和质量提升需求，从而形成建设与应用闭环。

3 数据集构建方法

3.1 多源数据汇聚

从汇聚方式看，可按实时接入、准实时归集和批量导入三类路径推进。车辆定位、路况、信号状态、视频告警、停车余位和实时客流等高时效数据，宜通过接口订阅、消息流或网关推送实现实时接入；交易汇总、线路变更、设备状态和服务日志等数据，可通过准实时归集保持业务同

步；基础资源和历史档案数据，则适合通过周期导入方式统一纳管。这样的分层设计既能保障动态数据时效，也能兼顾静态资源和历史数据沉淀。

从建设重点看，数据汇聚应坚持场景优先，而不是平均铺开。治理侧应优先打通公路、水路、道路运输、城市公共交通、应急资源和视频监控链路，服务侧应优先整合公交、地铁等数据，逐步形成覆盖核心对象和关键服务链路的基础网络。在组织机制上，还需要建立跨部门、跨企业的数据协同制度，明确数据供给、管理、使用和安全责任边界^[16]，使数据汇聚从一次性接入转向常态化供给，如表 3 所示。通过这一汇聚机制，前文所述数据分散、来源割裂和供给不稳定的问题，才能被转化为可持续组织和统一纳管的数据基础。

3.2 标准治理与主题建模

推进标准化治理，首先要围绕线路、站点、车辆、路段、枢纽、事件、交易和资格等核心对象建立统一编码体系和元数据规范，明确字段定义、时间基准、空间基准、更新频率和业务口径；其次要推进

表3 典型数据来源与获取方式

数据来源	数据名称	获取方式
市交通局	实时路况数据	接口实时调用
公交集团	公交线路运行数据	数据共享平台
地铁公司	地铁客流数据	接口实时调用
市公安局	交通卡口监测数据	数据脱敏后提供
市人社局	特殊人群身份数据	最小必要授权
市文旅局	旅游景区客流数据	数据共享平台
停车场运营企业	停车场空位数据	商业合作接入
互联网企业	出行需求预测数据	数据融合计算

时间、空间和语义对齐，解决不同系统时间戳不一致、空间单元不统一、同名指标含义不同等问题，避免关联分析失真^{[3][8]}。

在标准治理基础上，还需要围绕治理场景和服务场景开展主题建模，推动原始数据向可复用的数据集转化。治理侧可以形成客流监测、路网运行、公交轨道协同、重大活动保障和多模态感知等主题数据集，服务侧可以形成一码通行、资格核验、换乘服务和重点群体保障等主题数据集。其核心做法，是围绕“人-车-路-站-线”建立稳定的对象关系和业务链路，并将图片、视频及其标注结果与结构化业务数据统一组织，形成可支撑事件分析、人流统计等任务的可计算数据网络。只有经过这一阶段，前文提到的编码不统一、对象关系薄弱和数据建设与场景衔接不紧的问题，才有可能得到实质性缓解。

3.3 数据质量评价与改进

质量评价首先需要明确评价维度，并质量评价的作用并非形成一次性评分，而是通过“盘、建、研、管、用”等环节持续完善数据集能力体系^[18]。结合福州实践，项目形成了“接入校验—规则清洗—异常检测—人工抽检—业务反馈”的质量

根据具体应用场景设置相应指标。2025年12月发布的交通运输行业标准计划《交通运输行业高质量数据集建设导则》（JT 2025-125）提出完整性、准确性、一致性、及时性、唯一性、有效性、规范性和安全性八个质量评价维度^[17]，为交通高质量数据集质量评价提供了基本框架。结合城市交通治理与服务需求，本文进一步将八个维度细化为可观测、可计算的评价指标，如表4所示。

不同场景对数据质量的关注重点有所差异。例如，运行监测场景重点关注数据更新延迟、异常值检出率和重复记录率，资格核验场景重点关注核验准确率和访问权限合规率，换乘服务场景则重点关注外键完整性、跨系统匹配率和关键链路覆盖情况。通过将通用评价维度与具体业务指标相对应，可以避免质量评价停留在原则性要求层面。

在明确评价维度和适用指标的基础上，还需要结合项目数据特征和业务需求确定各维度在综合评价中的权重。福州案例主要服务于运行监测、协同调度和便民服务，对关键对象覆盖、数据真实性、跨系统关联和动态数据响应具有较高要求。因此，项目将完整性和准确性分别设置为20%，将一致性和及时性分别设置为15%，以重点保障关键数据覆盖、结果可靠、跨系统匹配和实时响应；有效性和安全性分别设置为10%，用于保障数据符合业务规则并实现合规使用；规范性和唯一性分别设置为5%，用于评价元数据管理和重复数据控制情况。具体权重设置如表5所示。

提升闭环：接入阶段校验数据结构、编码及其时空合法性；治理阶段识别重复上传、时空错位、异常漂移和逻辑冲突；应用阶段结合现场感知、业务结果和公众反馈开展反向验证；反馈阶段持续修正规则、优

表4 城市交通高质量数据集质量评价维度与指标

维度	指标示例	城市交通场景下的关注重点
完整性	字段填充率、主键完整率、外键完整性	客流、车辆、站点、路段、资格标签等关键对象是否齐全,换乘链路是否完整
准确性	数据校验通过率、人工标注准确率	车辆定位、客流识别、停车状态、资格核验结果是否真实可靠
一致性	格式符合率、跨系统数据匹配率	公交、地铁、停车、支付、视频等系统之间的编码、口径、时间空间基准是否统一
及时性	数据更新延迟、刷新周期达标率	路况、GPS、停车余位、实时客流等动态数据能否满足监测预警和出行服务时效要求
唯一性	重复记录率	交易记录、车辆轨迹、事件告警等是否存在重复上传、重复采集或重复入库
有效性	数据范围合理率、异常值检出率	速度、客流、停车余位、到离站时间等是否符合交通运行规律和业务规则
规范性	元数据规范率、目录登记完整率	字段定义、主题标签、对象描述、数据目录和版本信息是否规范统一
安全性	访问权限合规率、数据脱敏完成率	资格核验、支付交易、出行画像等敏感数据是否做到最小必要授权和安全使用

表5 城市交通高质量数据集质量评价参考权重

维度	参考权重/%	权重说明
完整性	20	直接影响关键对象覆盖、关键链路贯通和主题数据集成型质量
准确性	20	直接影响客流识别、车辆定位、资格核验和运行研判结果可信度
一致性	15	影响跨系统关联、跨方式协同和主题数据集复用能力
及时性	15	影响路况监测、调度预警、停车诱导和出行服务实时响应能力
唯一性	5	反映重复采集、重复入库和重复调用问题控制水平
有效性	10	反映数据是否符合交通业务规则和运行规律
规范性	5	反映元数据、目录、标签和版本信息的规范化程度
安全性	10	反映敏感数据使用、共享和服务输出是否合规可控

化阈值并更新模型。通过将质量诊断、问题派单、人工复核和版本追踪纳入常态化流程。

3.4 安全治理与可信流通

与此同时，随着数据集共享范围扩大和场景调用频率提高，安全治理与可信流通也需要同步跟进。完善安全治理，首先要建立覆盖全量数据资源的分类分级机制，根据敏感程度、业务重要性和流通范围，对基础资源、运行监测、交易、资格、视

频和日志等数据设置差异化的访问控制、处理规则和审计要求；其次要实行最小必要授权和分级开放，明确不同部门、平台和场景可访问的数据粒度、使用期限和责任边界，优先通过脱敏结果、统计结果或服务接口对外提供能力，降低原始数据暴露风险^[19]。

在此基础上，还应结合隐私保护计算、可信执行环境、区块链存证和全过程日志等手段，实现“数据可用不可见”，如对车牌、人脸、证件号码、联系方式等敏感字段实施掩码、模糊化或去标识化处理，同

时，要建立与业务运行相衔接的安全监测和应急处置机制，持续开展异常访问识别、接口风险预警、违规调用追踪和泄露处置^[20]。应当注意的是，“数据可用不可见”并非适用于所有数据和场景的低成本方案。隐私计算通常会增加加密、通信和多方协同开销，对实时路况、视频流等高频大规模数据可能造成明显时延；可信执行环境依赖特定软硬件栈，还面临异构设备适配、可信证明、密钥管理和侧信道防护等问题。

因此，福州实践按数据敏感度和业务时效进行分级选型：一般业务数据优先采用去标识化、最小字段集和受控接口；确需跨主体联合计算的高敏感性数据，再在小范围场景采用隐私计算或可信执行环境；对原始视频、人脸图像、个人出行轨迹等高风险数据，优先在本地或边缘侧处理，仅输出统计结果、标签或模型服务。技术选型前应评估总拥有成本、性能损耗、服务等级、可替代方案和退出机制，经试点验证后再逐步扩围。由此，“数据可用不可见”不再是单一技术承诺，而是由制度规范、数据处理、接口控制和特定增强技术共同实现的风险控制目标。

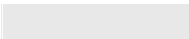
4 应用验证

在服务侧，福州以“榕城通/e 福州”为统一入口，持续整合公交、地铁、充值支付等数据，逐步形成面向一码通行、资格核验和出行服务的主题数据集。平台支持实体卡、手机 NFC、二维码等多种介质，并兼容住建部全国城市一卡通和交通运输部互联互通标准。这一实践表明，依托服务平台推进数据集建设，不仅能够沉淀支付、资格和服务触达数据，也能够直接转化为公众可感知的出行服务能力。

围绕重点群体服务，服务侧打通了人社、公安、民政、教育等部门的数据接口，在政策卡申办环节嵌入权威身份信息，进行自动核验机制，实现老年人、学生、新落户人员等群体优待资格在线办理^[21]。据项目运营单位内部统计，2025 年全年发放及服务各类政策优待卡超过 50 万人次，通过技术优化，扫码支付全流程为 9 秒以内，交易成功率稳定在 99.9% 以上，资金结算准确率达 99.99%。通过持续提升公共交通与数字便民服务水平，公交领域市民出行满意度已突破 90%^[22]。从价值实现角度看，高质量数据集建设通过支撑公交线网优化、来车信息实时预报、资格核验、支付协同和跨部门业务联动，提高了交通服务的响应效率，是推动公众出行体验和满意度提升的重要因素。

在治理侧，福州以 TOCC 为核心载体，围绕铁路、民航、公路、水路和城市公共交通等多领域数据^[23]，建设面向运行监测、事件识别、协同调度和应急联动的治理类主题数据集。现阶段，TOCC 规划对接 17 家数据源单位共 48 个数据项，已完成 10 家单位 33 个数据项接入，整合 6 万余路视频资源，依托治理平台推进高质量数据集建设。其核心意义在于，逐步形成跨层级、跨部门、跨主体的数据统一汇聚与统筹纳管能力，进一步通过标准统一、视频智能分析、跨部门联动和模型研判，将分散的数据资源稳定转化为监测预警、应急调度和决策支持能力。对于治理侧而言，这一转变的价值不仅在于“看见更多数据”，更在于推动原本口径不一、彼此割裂的信息进入同一对象体系和同一研判链路，减少重复核对和多头研判，提升跨部门协同处置的及时性、协同性和一致性。

从业务效果看，治理侧增强了运行监测、事件识别、协同调度和应急联动能力，



服务侧提升了公众出行效率和重点群体服务水平，体现出高质量数据集能够同时转化为治理效能和服务效能。在此基础上，福州具备向数据产品化延伸的现实潜力，在合规前提下形成脱敏接口、专题分析结果、监测指数、决策支持服务和模型能力输出^[24]。结合现有基础，可优先探索枢纽客流监测、重点区域拥堵研判、重大活动交通保障评估、停车热度分析和景区商圈客流画像等产品，后续进一步完善授权边界、产品目录、定价机制和持续运营模式。

5 局限与改进方向

尽管福州已形成一定实践基础，交通高质量数据集从场景验证走向持续运营和复制推广仍面临若干问题：跨主体数据标准尚未完全统一，实时数据质量稳定性不足，安全授权与可信流通机制仍需完善，隐私计算、可信执行环境和边缘处理的规模化应用还受到算力、适配成本和运维能力制约；现有成效评价也缺少建设前后同口径指标和对照数据，数据价值转化仍以内部支撑和单场景应用为主。

下一步，应进一步完善主数据和接口标准，强化自动校验、异常告警、人工抽检与业务反馈闭环。安全方面，应优先做好数据分类分级、目录治理、最小必要授权和受控接口，再对高敏感性联合计算场景开展小范围验证；成效评价方面，服务侧可形成满意度等连续指标，治理侧可记录事件发现与处置时长、误报率和人工核对工作量，并通过前后对比或分阶段评估识别数据集建设的增量作用。在此基础上，可面向“交通+文旅”“交通+商业”“交通+绿色出行”等场景开发监测分析、决策支持和数据服务产品，推动数据集由内

部支撑向多场景运营和开放赋能延伸。

6 结束语

面向交通运输高质量发展和数据要素价值释放的新要求，城市交通高质量数据集建设正在成为城市交通数字化转型的重要基础。本文以福州实践为例，围绕TOCC与“榕城通/e福州”两类核心场景，分析了城市交通区域高质量数据集建设的现实需求，提出了总体框架和关键路径，并结合治理端与服务端实践，对其应用价值和建设成效进行了总结。

研究表明，高质量数据集建设的关键，不在于单纯扩大数据规模，而在于围绕治理与服务场景，推动多源异构数据实现统一汇聚、标准治理、质量提升、安全流通和持续运营，并通过“人、车、路、站、线”对象建模和双场景协同应用，形成由数据资源向治理效能和服务效能转化的闭环能力。福州案例说明，这一路径已经在治理侧形成综合研判和联动处置的基础支撑，在服务侧形成资格核验和一体化出行服务的数据支撑，并具备继续向专题分析和数据产品化延伸的条件。只有把场景牵引、制度约束、技术支撑和持续运营结合起来，城市交通高质量数据集才能真正成为支撑城市交通治理现代化和公众出行服务升级的长期能力。

参考文献：

- [1] 交通运输部. 刘伟主持召开“十五五”智慧交通发展专家座谈会 牢牢把握服务支撑交通强国建设战略方向 高质量做好“十五五”交通运输规划编制工作[EB/OL]. 2025-06-26[2026-04-01]. <https://www.mot.gov.cn/zhong>

yaohuiyi/202506/t20250626_4171277.html.

Ministry of Transport of the People's Republic of China. Liu Wei chaired an expert symposium on smart transportation development for the 15th Five-Year Plan[EB/OL]. 2025-06-26[2026-04-01]. https://www.mot.gov.cn/zhongyaohuiyi/202506/t20250626_4171277.html.

[2] 交通运输部. 17 部门联合部署“数据要素×”三年行动 发挥交通运输领域数据要素乘数效应 [EB/OL]. 2024-02-07[2026-04-01]. https://www.mot.gov.cn/jiaotong-yaowen/202402/t20240207_4008506.html.

Ministry of Transport of the People's Republic of China. Seventeen departments jointly deployed the three-year action plan for Data Elements × to leverage the multiplier effect of data elements in transportation[EB/OL]. 2024-02-07[2026-04-01]. https://www.mot.gov.cn/jiaotongyaowen/202402/t20240207_4008506.html.

[3] 吴子啸, 付凌峰, 赵一新. 多源数据解析城市交通特征与规律[J]. 城市交通, 2017, 15(4): 56-62.

Wu Z X, Fu L F, Zhao Y X. Analysis of urban traffic characteristics and patterns based on multisource data[J]. Urban Transport of China, 2017, 15(4): 56-62.

[4] 姜桂艳, 牛世峰, 李红伟. 动态交通数据质量评价方法研究[J]. 北京工业大学学报, 2011, 37(8): 1190-1195.

Jiang G Y, Niu S F, Li H W. A data quality assessment method for dynamic traffic data[J]. Journal of Beijing University of Technology, 2011, 37(8): 1190-1195.

[5] 杨琳, 朱扬勇. 高质量数据集产品的形态和生产流程研究[J]. 大数据, 2026, 12(2): 64-74.

Yang L, Zhu Y Y. Research on the form and production process of high-quality dataset products[J]. Big Data Research, 2026, 12(2): 64-74.

[6] 张群洪, 刘浩. 全国一体化数据市场视角下的交通数据要素化市域实践——以福州为例[J]. 大数据, 2026, 12(4): 147-162.

Zhang Q H, Liu H. Unified national data market-oriented municipal practice of transportation data factorization: a case study of Fuzhou[J]. Big Data Research, 2026, 12(4): 147-162.

[7] 张群洪, 李林, 杨堃, 等. 基于可信数据空间架构的高质量数据集建设[J]. 大数据, 2026.

Zhang Q H, Li L, Yang K, et al. High-quality dataset construction based on trusted data space architecture[J]. Big Data Research, 2026.

[8] 张懿木, 陈田, 王俊, 等. 基于多源数据的多模式公共交通出行链构建[J]. 城市交通, 2021, 19(4): 120-126.

Zhang Y M, Chen T, Wang J, et al. Construction of multimodal public transport trip chains based on multisource data[J]. Urban Transport of China, 2021, 19(4): 120-126.

[9] 王月, 程娜. 基于多源数据的公共汽电车运营评估: 以重庆市为例[J]. 城市交通, 2024, 22(1): 53-61.

Wang Y, Cheng N. Public bus and trolleybus operation evaluation based on multisource data: a case study of Chongqing[J]. Urban Transport of China, 2024, 22(1): 53-61.

[10] 国家数据局. 高质量数据集典型案例 | 招商局集团交通物流行业高质量数据集[EB/OL]. (2025-11-25) [2026-07-11]. https://www.nda.gov.cn/sjj/ywpd/szkjyjcsw/1125/20251125183635771759800_pc.html.

National Data Administration. Typical case of high-quality datasets: China Merchants Group's high-quality dataset for transportation and logistics[EB/OL].

- 2025-11-25[2026-07-11]. https://www.nda.gov.cn/sjj/ywpd/szkjycss/1125/20251125183635771759800_pc.html.
- [11] 国家数据局. 高质量数据集典型案例 | 中国交通基础设施多维感知安全监测数据集[EB/OL]. 2025-11-27[2026-04-01]. https://www.nda.gov.cn/sjj/ywpd/szkjycss/1127/20251127143442536097684_pc.html. National Data Administration. Typical case of high-quality datasets: CCCC's multidimensional sensing and safety monitoring dataset for transportation infrastructure[EB/OL]. 2025-11-27[2026-04-01]. https://www.nda.gov.cn/sjj/ywpd/szkjycss/1127/20251127143442536097684_pc.html.
- [12] 国家数据局. 高质量数据集典型案例 | 中交集团交通基建行业通识类数据集建设与应用[EB/OL]. 2025-11-26[2026-04-01]. https://www.nda.gov.cn/sjj/ywpd/szkjycss/1126/20251126151100416711695_pc.html. National Data Administration. Typical case of high-quality datasets: construction and application of a general-purpose dataset for transportation infrastructure at CCCC[EB/OL]. 2025-11-26[2026-04-01]. https://www.nda.gov.cn/sjj/ywpd/szkjycss/1126/20251126151100416711695_pc.html.
- [13] 国家数据局. 高质量数据集典型案例 | 莆田市全域多模态城市治理数据集[EB/OL]. 2025-10-20[2026-04-01]. https://www.nda.gov.cn/sjj/ywpd/szkjycss/1020/20251020160814070295152_pc.html. National Data Administration. Typical case of high-quality datasets: Putian's citywide multimodal urban governance dataset[EB/OL]. 2025-10-20[2026-04-01]. https://www.nda.gov.cn/sjj/ywpd/szkjycss/1020/20251020160814070295152_pc.html.
- [14] 武汉市数据局. 新质生产力在湖北 | 湖北发布第二批高质量数据集[EB/OL]. (2025-05-28) [2026-07-11]. https://home.wuhan.gov.cn/mtbd/202505/t20250528_2587224.shtml. Wuhan Municipal Data Bureau. New quality productive forces in Hubei: Hubei released its second batch of high-quality datasets[EB/OL]. 2025-05-28[2026-07-11]. https://home.wuhan.gov.cn/mtbd/202505/t20250528_2587224.shtml.
- [15] 国家数据局. 专家解读 | 推动高质量数据集建设, 加快实施“人工智能+”行动[EB/OL]. 2025-03-04[2026-04-01]. https://www.nda.gov.cn/sjj/zhuanti/sjbz/0414/20250304155233105840686_pc.html. National Data Administration. Expert interpretation: advancing high-quality dataset construction and accelerating the AI Plus initiative[EB/OL]. 2025-03-04[2026-04-01]. https://www.nda.gov.cn/sjj/zhuanti/sjbz/0414/20250304155233105840686_pc.html.
- [16] 交通运输部科学研究院. 我院信息中心牵头编制的《交通运输数据安全风险评估指南》行业标准正式发布[EB/OL]. 2025-03-28[2026-04-01]. https://www.motcats.ac.cn/art/2025/3/28/art_55_4056.html. China Academy of Transportation Sciences. The industry standard Guidelines for transportation data security risk assessment was officially released[EB/OL]. 2025-03-28[2026-04-01]. https://www.motcats.ac.cn/art/2025/3/28/art_55_4056.html.
- [17] 交通运输部. 交通运输行业高质量数据集建设导则(JT 2025-125)[EB/OL]. (2025-12-18)[2026-07-12]. <https://jtst.mot.gov.cn/gb/search/gbDetailed?id=b40ab-caf84fd5d776f309af249653ea4>. Ministry of Transport of the People's Republic of China. Guidelines for the

construction of high-quality datasets in the transportation industry (JT 2025-125) [EB/OL]. 2025-12-18[2026-07-12]. <https://jtst.mot.gov.cn/gb/search/gb-Detailed?id=b40abcaf84fd5d776f309af249653ea4>.

[18] 姜春宇, 白玉真, 刘渊, 等. 构建企业级人工智能高质量数据集: 方法与路径[J]. 大数据, 2025, 11(06): 47-56. DOI: 10.11959/j.issn.2096-0271.2025088. Jiang C Y, Bai Y Z, Liu Y, et al. Building enterprise-level high-quality AI datasets: methods and pathways[J]. Big Data Research, 2025, 11(6): 47-56. DOI: 10.11959/j.issn.2096-0271.2025088.

[19] 交通运输部科学研究院. 交通运输数据脱敏指南 [EB/OL]. 2024-04-12[2026-04-01]. https://www.motcats.ac.cn/art/2024/4/12/art_136_3716.html. China Academy of Transportation Sciences. Guidelines for transportation data desensitization[EB/OL]. 2024-04-12[2026-04-01]. https://www.motcats.ac.cn/art/2024/4/12/art_136_3716.html.

[20] 中国电子系统技术有限公司. 北京市交通运行协调指挥中心 TOCC 项目 [EB/OL]. 2022-01-12[2026-04-01]. <https://www.cestc.cn/case/Caseshow/id/116/>. China Electronics System Technology Co., Ltd. Beijing Transportation Operations Coordination Center (TOCC) project[EB/OL]. 2022-01-12[2026-04-01]. <https://www.cestc.cn/case/Caseshow/id/116/>.

[21] 福州日报. 福州市重磅推出“青榕工程”[EB/OL]. 2025-11-01[2026-04-01]. <https://news.fznews.com.cn/fzxw/20251101/I5Y49HvW1Y.shtml> Fuzhou Daily. Fuzhou launches the Qingrong Project[EB/OL]. 2025-11-01[2026-04-01]. <https://news.fznews.com.cn/fzxw/20251101/I5Y49HvW1Y.shtml>.

[22] 福建省人民政府. 织密民生网络 激活城市脉动 [EB/OL]. 2025-05-08[2026-04-01]. https://www.fujian.gov.cn/xwdt/mszx/202505/t20250508_6909374.htm The People's Government of Fujian Province. Weaving a dense livelihood network and activating the urban pulse [EB/OL]. 2025-05-08[2026-04-01]. https://www.fujian.gov.cn/xwdt/mszx/202505/t20250508_6909374.htm.

[23] 福州日报. 市交通运输局:大道通衢绘宏图阡陌纵横谱新篇[EB/OL]. 2026-01-19[2026-04-01]. https://mag.fznews.com.cn/fzrb/2026/20260119/20260119_025/20260119_025_1.htm Fuzhou Daily. Fuzhou Municipal Transport Bureau: building a comprehensive transport network for urban development[EB/OL]. 2026-01-19[2026-04-01]. https://mag.fznews.com.cn/fzrb/2026/20260119/20260119_025/20260119_025_1.htm.

[24] 国家发展改革委, 国家数据局. 国家发展改革委 国家数据局关于印发《公共数据资源授权运营实施规范(试行)》的通知[EB/OL]. 2025-01-20[2026-04-01]. https://www.nda.gov.cn/sjj/zwgk/zcfb/0120/20250120171914419632974_pc.html. National Development and Reform Commission, National Data Administration. Notice on issuing the Specifications for Authorized Operation of Public Data Resources (Trial) [EB/OL]. 2025-01-20[2026-04-01]. https://www.nda.gov.cn/sjj/zwgk/zcfb/0120/20250120171914419632974_pc.html.



张群洪（1977- ），男，博士，福州市数字福州集团有限公司董事长，高级工程师、高级经济师，主要研究方向为数字化、数据开发、人工智能。



刘浩（1997- ），男，硕士，福州市数字福州集团有限公司数据资产部主管，主要研究方向为人工智能、数据要素价值化、数字政府、政务服务。

录用日期: XXXX-XX-XX
通信作者: 刘浩,nathan40@163.com
基金项目: 无
Foundation Items: