

大数据时代基层统计数据采集质量优化与标准化体系构建

王语嫣 朱迪 宋圣阳 庞静 孙旭阳
泰州学院

DOI:10.32629/acair.v3i4.17888

[摘要] 本文聚焦于大数据时代基层统计数据采集过程中面临的质量控制难题,围绕数据采集、处理、分析与应用全流程,提出构建以“标准化体系”为核心的“全流程智能质控+多源数据融合”框架。通过整合物联网设备、政务平台、社会感知等多源异构数据,结合联邦学习与可解释性人工智能技术,实现跨域数据清洗与质量核验,显著提升数据格式兼容率与语义一致性。建立涵盖实时性、完整性、一致性等多维度的动态质量评估指标体系,并在试点地区开展应用验证。结果表明,所构建体系在数据采集效率、质量可控性与决策支持能力方面均取得显著提升,为推进基层统计现代化提供了理论参考与实践路径。

[关键词] 数据; 基层统计; 数据采集; 质量优化; 标准化体系

中图分类号: S973.1+5 **文献标识码:** A

Optimization of Grassroots Statistical Data Collection Quality and Construction of Standardization System in the Big Data Era

Yuyan Wang Di Zhu Shengyang Song Jing Pang Xuyang Sun
Taizhou University

[Abstract] This paper addresses the quality control challenges in grassroots statistical data collection in the era of big data, focusing on the entire process from data acquisition to application. It proposes a framework centered on a “standardization system,” characterized by “full-process intelligent quality control + multi-source data integration.” By integrating multi-source heterogeneous data from IoT devices, government platforms, and social sensing, and employing federated learning and explainable AI technologies, cross-domain data cleaning and quality verification are achieved, significantly improving data format compatibility and semantic consistency. A dynamic quality evaluation index system covering multiple dimensions such as real-time performance, completeness, and consistency is established. Pilot applications demonstrate that the proposed system significantly enhances data collection efficiency, quality controllability, and decision-making support capabilities, providing both theoretical reference and practical pathways for advancing grassroots statistical modernization.

[Key words] big data; grassroots statistics; data collection; quality optimization; standardization system

引言

随着大数据、物联网、人工智能等新一代信息技术的迅猛发展,基层统计工作正面临数据源多样化、处理实时化、分析智能化等多重变革。传统统计数据采集模式在数据覆盖面、时效性、准确性等方面已难以适应现代治理需求,尤其在多源异构数据融合、质量控制与标准化管理方面存在明显短板。为应对上述挑战,本文立足于“技术—流程—管理”协同视角,探索构建覆盖数据采集、清洗、存储、分析与应用全流程的标准化体系。通过制定《多源数据接入规范》《AI模型评估指南》等标准文件,并引入“质量追溯码”等管理机制,实现从数据源头到分析应用的全链路质控。实践表明,该体系可有效提升数据权威性与决策

支撑能力,为推动基层统计向智能化、标准化转型提供系统性解决方案。

1 基层统计范式转型的技术驱动与形态演进

1.1 数据采集的全域化拓展与融合机制

大数据技术推动基层统计从传统的“抽样推断”向“全量感知”转型,构建起覆盖经济社会多维度场景的数据采集生态。在农业统计中,传统人工填报方式存在覆盖率低、偏差大等问题,部分县域粮食产量统计偏差率高达25%以上。通过部署物联网传感设备,如每50亩布设墒情监测点,实现土壤湿度、作物生长等多参数实时采集,使农业统计误差降至6%以内。在商业与社会统计领域,电商交易、物流信息、政务数据等多源信息的融合,大

幅提升了统计数据的时效性与颗粒度。某地级市通过整合市场监管、税务、人社等12个部门的数亿条数据,将企业活跃度统计细化至街道层级,为区域经济监测提供高精度画像。全域化采集模式不仅将统计覆盖范围从传统领域的70%拓展至新兴领域的90%以上,还通过交叉验证机制将数据可信度提升35%以上。

1.2 处理技术的智能化跃迁与流程再造

智能技术重构了基层统计数据处理流程,实现效率与精度的双重突破。传统人工录入模式下,数据错误率普遍高于7%,而基于OCR与自然语言处理的智能录入系统可将非结构化数据转换准确率提升至98%以上,大幅缩短预处理时间。在工业统计中,边缘计算节点实现对生产数据的实时清洗与异常识别,响应延迟低于50毫秒,有效提升数据信噪比。区块链技术的引入进一步强化数据溯源能力,通过分布式记账与时间戳机制,确保数据全程可追溯、不可篡改,显著提升数据可信度与验证效率。这些技术的综合应用使基层统计全流程处理周期从数周压缩至数天,为实时监测与动态决策奠定基础。

1.3 分析方法的预测性转向与模型创新

大数据推动统计分析方法从“事后描述”向“前瞻预警”转型,构建起多维度、高精度的预测体系。传统GDP核算模型依赖有限指标与线性方法,预测误差较大。某省统计局通过融合数百项宏观指标与数百万条微观数据,构建智能预测模型,将短期经济预测误差控制在7%以内。在民生统计领域,基于社交媒体评论的情感分析技术可提前识别民生诉求热点,准确率达88%以上,为公共服务优化提供数据依据。市场监管中,关联规则挖掘技术能够从企业注册、纳税、用工等多维数据中识别异常经营行为,支持度超过0.85,显著提升监管效率与精准度。预测性分析能力的构建,使基层统计从被动响应转向主动干预,决策支撑价值提升40%以上。

2 大数据时代基层统计数据采集质量优化策略

为系统提升基层统计数据采集质量,需从技术、管理、制度等多维度构建协同优化策略,具体包括以下方面:

(1) 推进数据采集工具的智能化升级。积极应用物联网、智能传感、移动互联等先进技术,实现数据采集的自动化、实时化与精准化。通过布设高精度传感器与边缘采集设备,实现对物理世界与社会经济活动的全域感知,最大限度减少人为干预与操作误差,从根本上提升数据源的准确性与完整性。

(2) 强化数据清洗与预处理的标准化流程。针对多源异构数据质量参差不齐的问题,建立统一的数据清洗规则与预处理标准,涵盖噪声过滤、异常值识别、格式转换、语义对齐等环节。通过引入联邦学习、知识图谱等技术,实现跨域数据的语义融合与质量提升,为后续分析提供高质量数据基础。

(3) 构建高效可靠的数据存储与管理体系统。采用分布式存储、云计算、数据湖等先进架构,支持海量数据的弹性存储与高效管理。建立数据分类、索引、备份与容灾机制,确保数据的完整性、可用性与安全性。同时,通过元数据管理实现数据资产的全生命周期可控。

(4) 深化数据分析与挖掘的价值释放。综合运用机器学习、数据挖掘、可视化分析等方法,从海量数据中提取隐含规律与知识洞察。通过构建预测模型、关联分析、异常检测等算法工具,提升统计分析的深度与广度,为政策制定和业务优化提供科学依据。

(5) 完善数据采集管理的制度保障。制定涵盖数据采集、传输、存储、使用等各环节的管理规范与操作指南,明确责任主体与工作流程。通过制度化、标准化管理,减少人为因素导致的质量漏洞,确保数据采集工作规范、有序、高效开展。

(6) 建立全过程数据质量控制机制。构建覆盖数据生命周期的质量监控体系,包括数据校验、审核、评估与反馈等环节。通过设立质量阈值与预警规则,实现数据质量的实时监测与动态干预,确保数据从采集到应用的全链条可控、可信。

(7) 促进数据共享与业务协同。打破部门与系统之间的数据壁垒,建立统一的数据共享平台与交换机制。通过数据互联互通,避免重复采集,提升数据利用效率,支撑跨领域、跨层级的协同分析与决策。

(8) 加强数据安全与隐私保护。建立健全数据安全管理体系,综合运用加密传输、访问控制、安全审计等技术手段,防止数据泄露、篡改与滥用。在推进数据开放共享的同时,切实保障个人隐私与公共利益。

3 案例分析:某地区基层经济统计改革实践

某地区在大数据背景下积极推进基层经济统计工作的改革与创新,致力于构建覆盖全面、响应及时、质量可控的现代统计体系。改革前,该地区主要依赖工商登记与税务报表等传统数据源,数据覆盖面窄、时效性差,难以全面准确反映区域经济发展的真实状况。

为突破这一瓶颈,当地政府统筹建设多源数据采集平台,整合工商、税务、电商平台、物流企业等多方数据资源。通过数据接口标准化与语义对齐,实现企业注册信息、纳税记录、交易数据、物流动态等多维信息的深度融合,极大丰富了统计数据的来源与维度。

在采集手段方面,大力推进物联网与智能传感技术的落地应用。在重点企业的生产车间部署环境监测与设备运行传感器,实时采集产量、能耗、工艺参数等关键数据,为企业生产经营状况的精准统计提供第一手资料。同时,引入移动采集终端与智能填报系统,提升数据采集的便捷性与实时性。

在数据分析环节,构建基于大数据与人工智能的智能分析平台。整合宏观经济政策、行业趋势、市场供需等多重因素,构建区域经济发展预测模型,支持短期趋势判断与中长期规划制定。通过数据可视化与交互式分析,为政府部门提供直观、深入的数据洞察,助力产业政策调整与资源优化配置。

在质量管理方面,建立全过程数据质量控制体系。制定从采集、清洗、存储到分析的全链路质量标准,明确各环节的质量要求与责任分工。定期开展数据质量评估与审计,及时发现并纠正数据偏差,确保统计结果的准确性与可靠性。

在安全保障方面,构建多层次的数据安全防护体系。采用数据加密、身份认证、访问日志记录等技术手段,严防数据泄露与非法使用。同时,建立数据分级分类管理制度,平衡数据开放与隐私保护的关系。

经过系统化改革,该地区基层统计数据的准确性、时效性与应用价值显著提升,统计数据在区域经济监测、政策评估与社会治理中发挥出越来越重要的作用,为区域高质量发展提供了有力支撑。

4 标准化体系构建的关键路径

要系统构建基层统计数据采集的标准化体系,需从技术标准、管理规范与制度保障三个层面协同推进,形成闭环式标准生态。

在技术标准层面,应制定《基层统计多源数据接入与接口规范》,明确物联网设备、政务平台、商业系统等15类主流数据源的传输协议、数据格式与安全要求,推动非结构化数据标准化转化率提升至85%以上。同时,出台《统计数据处理与质量控制技术指南》,规范数据清洗、标注、融合等环节的技术路径与质量阈值,如设定异常值识别准确率不低于90%、数据实时延迟不超过10分钟等具体指标。

在管理规范层面,需建立《基层统计全流程质量管理办法》,涵盖数据采集计划制定、现场操作规范、质量抽检机制、问题追溯与整改等全链条内容。创新推行“统计质量追溯码”制度,为每一条数据赋予唯一标识,实现从采集源头到分析应用的全生命周期溯源,将问题定位时间从平均3天缩短至2小时以内。

在制度保障层面,应构建跨部门协同治理机制,明确统计、工信、市场监管、农业农村等部门的职责分工与数据共享义务。建立统计质量考核与激励约束机制,将数据质量纳入地方政府及部门绩效考核体系,对标准化执行良好的单位给予政策倾斜,对数据造假、质量失控等行为实施问责。

通过技术、管理与制度的三维联动,可系统解决基层统计中长期存在的标准缺失、执行不力、协同不足等问题,为统计数据质量提升提供坚实基础。

5 结束语

大数据时代的到来为基层统计工作带来深刻变革,推动数据采集、处理与分析方式向智能化、标准化、协同化方向演进。本文系统探讨了基层统计数据采集质量优化与标准化体系构建

的路径,提出以“技术—流程—管理”协同为主线,构建覆盖数据全生命周期的质量控制与标准化框架。通过推进采集工具智能化、强化清洗预处理、构建高效存储体系、深化分析挖掘、完善制度保障、建立质控机制、促进数据共享、加强安全防护等多维策略,全面提升基层统计数据的质量和标准化水平。

未来,随着5G、人工智能、区块链等新技术的不断成熟与应用,基层统计工作将迎来更广阔的发展空间。我们应持续推动统计理论、方法、标准与技术的创新,加强跨学科、跨领域的协同研究,构建更加适应大数据时代要求的基层统计工作体系,为经济社会高质量发展提供更加准确、及时、全面的统计服务。

【参考文献】

- [1]邵培.提升统计数据质量以驱动经济高质量发展研究[J].中国集体经济,2025(2):21-24.
- [2]彭诗语.大数据背景下基层政府统计数据质量提升策略分析[J].中文科技期刊数据库(全文版)经济管理,2024(12):157-159.
- [3]于庆华.能源统计数据在“双碳”背景下的提升研究[J].知识经济,2024(28):136-138.
- [4]黄小平.经济统计数据收集与处理的质量控制研究[J].大众文摘,2024(20):6-8.
- [5]张岩岩,刘英华,梁士平.基层公共卫生统计数据质量改进研究[J].统计与管理,2022,37(7):122-128.
- [6]李媛媛.基层统计视角下统计数据质量的提升策略探讨[J].中文科技期刊数据库(全文版)经济管理,2022(7):264-267.
- [7]刘晓伟.大数据背景下提高基层统计数据质量的思考[J].中文科技期刊数据库(全文版)社会科学,2021(7):42.
- [8]李国.大数据背景下政府投资审计问题及应对措施研究[J].首席财务官,2024,20(15):193-195,202.
- [9]刘春霞.大数据时代下的信息处理与分析技术[J].微型计算机,2024(5):55-57.
- [10]张绍娟.大数据背景下建筑企业财务管理发展分析[J].财会学习,2022(1):34-36.

作者简介:

王语嫣(2004--),女,汉族,江苏省泰州人,本科,研究方向:信息统计与数据分析。