



提高数据质量的“三个关键点”

□ 韩红桂

数据是数字化、网络化、智能化的基础，是推动社会经济发展的新型生产要素。作为数字经济的重要组成部分，数据在推动社会生产力发展中的作用日趋显著，已深度融入生产、分配、流通、消费和社会服务管理等各个环节。中共中央、国务院印发的《数字中国建设整体布局规划》提出了“数据资源规模和质量加快提升，数据要素价值有效释放”的目标，强调要夯实数字中国建设基础，全面赋能经济社会发展。

数据质量是数据管理和数据分析的基础，决定了各行业的运营效率、服务品质以及决策能力等。高质量数据有助于优化业务流程、提升服务品质、提高决策精准度，对于推动信息化和智能化进程，实现数据驱动的创新发展具有深远意义。然而，在工业数据采集过程中，数据质量往往受多种因素影响：由于传感器故障或读数误差、操作人员失误、环境变化等原因，致使数据存在离群值和缺失值；由于环境干扰或设备老化等原因，造成数据包含噪声、干扰等信息；由于设备不同步、采集频率差异、数据延迟或传输错误等原因，导致数据存在时间不一致性等。低质量数据的产生不仅直接影响数据的可用性，同时会导致员工操作失误、决策偏差，企业运营效率降低和大量资源浪费等。

2022年12月印发的《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》，2024年9月印发的《中共中央办公厅 国务院办公厅关于加快公共数据资源开发利用的意见》，提出加强数据管理和利用的具体措施，包括建立健全数据质量管理体系、推动数据标准化、加强数据安全保护等。工信部发布的《“十四五”大数据产业发展规划》也明确了大数据产业的发展方向，强调优化数据资源配置，激发数据要素潜能，推动大数据产业高质量发展。国家系列文件的发布，为数据质量管理明确了发展方向与发展路径。

选择可靠有效的数据处理方法

系统化的数据质量提升流程是维护和提高数据质量的坚固防线。获取准确性高、一致性强的数据，

首先要有一套行之有效的数据处理方法，主要包括：异常数据检测、数据去噪以及数据同步对齐等。

异常数据检测是确保数据准确性的首要环节。及时、准确识别异常数据是数据分析中的关键一环。通过对异常数据的检测可以增强数据集的完整性与可靠性，为后续数据分析和模型建立提供坚实的数据基础，是维护整个决策链准确性的重要措施。异常数据通常分为偏差数据、离群值和缺失值等。其中，离群值的产生是由于测量误差或设备故障引起。比如，以特定的距离函数计算数据对象之间的距离来确定离群点，一般认为离其他数据更远的数据点可能是离群点。该方法可以简单直观地检测出异常数据，但需谨慎选择合适的距离度量标准。再比如，通过数据驱动方式学习数据中的模式或结构来区分正常数据或异常数据，该方法能够适应复杂的数据分布。异常数据还包括在监测过程中出现的数据缺失，通常通过直接空值法、统计分析法和时间序列法等检测来辨识异常数据。直接空值法通过简单操作就能够发现明显的缺失数据。统计分析法是通过分析数据的分布特征实现数据的缺失值检测，该方法有标准的流程，易于操作和复制。时间序列法是通过数据在时间维度上的连续性和趋势性的分析来检测数据的缺失值。

数据去噪处理是提升数据质量的关键环节。数据去噪处理是确保数据质量的核心步骤，能够有效提升数据集的纯净度和可信度，为后续数据分析和模型建立提供稳定的数据平台，是保障决策链精确性的重要基石。数据去噪的方法分为基于滤波器的数据去噪方法、基于频域分析的数据去噪方法、基于机器学习的数据去噪方法等。比如，通过建立特定的滤波器来滤除数据中的高频或低频噪声，保留有用的信号，确保数据的平滑性和准确性。还有通过将数据转换到频率域，利用信号在频率特征的差异来消除噪声。此外，通过训练模型来学习数据的内在模式和特征，能够处理更加复杂的数据结构和噪声类型。

数据同步对齐是提升数据质量的关键步骤。数据同步对齐对保障数据的质量具有决定性作用，提升数据集的一致性，为后续数据分析和模型建立搭建稳定的数据框架。动态时间规

整法通过计算两个时间序列之间最佳规整路径实现数据的非线性对齐，适用于不同速率的信号对齐以及不同长度的时间序列对齐。时序重采样法通过改变采样频率、填补缺失时间点或合并时间间隔等方式将原始时间序列数据转换为新的时间序列以实现数据同步，适用于解决由于采样不一致导致的数据对齐问题。

实施严格的数据验证与实时监控

为了确保数据的有效性和准确性，严格的数据验证是必不可少的环节。数据验证是保障数据源可靠性的重要措施，是数据质量提升的坚实支撑，主要包括格式验证、范围验证和逻辑验证等方法。格式验证旨在检查数据是否符合预定的格式要求，如日期格式、数值范围等，确保数据的标准化。范围验证是确保数据在预定义的合理区间内，防止数据因操作失误或采集误差等偏离正常范围。逻辑验证则基于业务逻辑关系检查数据之间的合理性，如数据的一致性和关联性。高效的数据验证机制能够识别并纠正数据集中的错误、异常和遗漏，提高数据集的整体品质，确保数据的有效性和准确性，为后续分析和决策提供可靠的数据保障。

实时数据决策的重要性在现代社会各个行业中日益凸显，这些决策都需要高效、可靠的实时数据监控。数据监控是维护数据实时性和连贯性的重要手段，是数据质量提升的必要保障，主要包括实时数据监控、阈值监控和趋势分析等方法。实时数据监控通过建立实时报警机制，迅速识别数据异常情况，确保潜在问题被及时发现并得到有效处理。阈值监控通过预设的合理范围对数据进行实时跟踪，一旦超出范围即发出警报。趋势分析通过挖掘历史数据，识别数据变化的长期趋势，为未来趋势预测提供依据。有效的数据监控机制能够促进数据的连续性和一致性，确保在快速变化的环境中实现精准的数据管理和快速响应，为各行各业的稳定运行和效率提升提供强有力的支撑。

综上所述，构建完善的数据质量提升流程是确保数据准确性、可靠性和应用价值的关键所在。异常数据检

测为提升数据质量奠定了基础，数据去噪和数据对齐则是提升数据质量的有效手段。三者有机结合，能够直观地增强数据的完整性、纯净度和规范性，确保数据集的高标准和可靠性。数据验证与数据监控虽然不能直接提升数据质量，但是能够确立衡量数据质量的关键准则，保障数据源的可靠性与连贯性，为数据质量提升提供支撑。上述环节相辅相成，共同构成了数据质量闭环管理，确保数据从采集到应用的全生命周期遵循高质量标准。

构建科学规范的数据质量管理体系

数据质量的提升过程是一项复杂的系统工程，需要精准匹配不同行业的实际需求，全面考虑特定环境下的设备类型、运行状态、业务流程等独特因素，采取针对性的方法与措施增强数据的准确性和可用性，提升整体的运营效率，优化资源配置，促进各行业的数字化转型。与此同时，根据变化的数据环境和业务需求，持续优化和迭代数据质量提升流程，确保数据价值最大化。

在数字化浪潮推动下，数据质量管理快速迈向智能化阶段，数据质量提升正面临严峻挑战：一方面，数据量呈爆炸性增长；另一方面，实时数据流的处理、数据精确度与一致性的维护等需求标准日益提高。因此，未来数据质量提升应深度聚焦数据处理智能技术开发，实现对大规模数据的深度挖掘、精准分析和快速响应；全力构建自适应数据验证系统，确保数据在动态变化的环境中自动调整，始终符合高质量标准；不断优化传感器技术，提升数据的采集效率和准确性；广泛推广边缘计算与分布式存储技术，增强数据处理的能力和灵活性，降低数据处理的延迟；精心制定和推广国际数据管理标准，促进数据质量的全球化统一；确立数据质量在促进各行业智能化和数字化转型中的核心作用，推动数据质量迈向更高层次。



数字现象

信息前沿

权，抓住网络文化传播的便捷性、互动性和互动性等特点，打通线上线下传播渠道，形成网上网下文化传播合力，以人民为中心推动中华优秀传统文化网络传播，满足人民群众日益增长的精神文化需求。加强顶层设计和制度建设，深化主流媒体国际传播机制改革创新，有效利用出海平台企业的竞争优势，努力保护出海企业的合法权益，引导商业文化网站和自媒体健康有序发展，形成具有国际影响力的新型文化集团和媒体集团，构建多渠道、立体式对外传播格局，从战略、资源、力量、战术等方面融合协同和统筹整合，促进中华网络文化国际传播力的提升。构建权责清晰、监管有力的新兴数字文化大市场，建立数字文化平台运营和数据监管标准体系，创新商业模式活跃数字文化市场，规范网络文化传播秩序，遏制违法有害信息网上传播，推进完善网络文化服务市场准入和退出机制，鼓励各类经营主体公平竞争，促进数字文化资源开放透明、规范有序流动，推动数字文化贸易繁荣发展。



在数据爆发的今天，大数据广泛应用于提高社会效率，推动着社会生产力发展。数据开放已经成为顺应时代发展的必然选择，但是，随之而来的数据安全和个人隐私也成为亟待解决的问题。当前，数据要素相关制度、技术、立法不完善，隐私泄露事件多发，影响数据价值的释放。《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》强调，要以保护个人信息为前提，促进数据合规高效流通使用。加强个人数据安全保护，需要推进制度规范监督的完善，充分释放数据要素价值，促进数据高效流通使用。

个人数据具有重要的社会经济价值，同时还关乎个人基本权利。然而，随着数据收集、存储、使用的规模不断扩大、场景的不断增多，个人数据面临的风险与日俱增。《中国互联网络发展状况统计报告》显示，约有四成网民遭遇信息泄露、网络诈骗、身份盗用等个人数据安全问题。个人数据面临的风险情境也愈加复杂。比如，个人信息的过度收集。社交媒体、移动应用和在线服务平台常以用户体验为名过度收集超出其本身功能必要范围的用户信息，在未经得用户同意的情况下对一系列个人数据打包授权，导致大量敏感数据泄露。再如，数据的不当使用。数据持有者出于经济利益考量，在未经得用户同意的情况下将收集到的个人数据用于其他商业用途或转售给第三方，数据被用于市场营销、行为分析或其他经济目的，加剧个人数据的泄露风险。还比如，网络诈骗。非法采集者通过“社会工程方法”，即通过操纵、影响或欺骗他人，以获取敏感信息，诱使个体做出不利的数据决策，敏感数据泄露导致身份盗用、财务损失及声誉受损等严重后果。最后，技术渗透。黑客攻击、恶意软件和数据挖掘等先进技术手段使得入侵并窃取个人数据变得愈加隐蔽，给个人数据安全带来更大挑战。

随着数据保护技术和隐私政策的不断健全，技术创新与社会制度的协调发展能够为个人数据保护提供可行方案，以安全治理机制健全数据基础制度，促进数据流通利用有序、可持续。

建立科学有序的风险预防机制。物联网设备、高清摄像系统、智能手机等技术的迭代升级使个人数据采集更具即时性、隐蔽性，云计算和区块链等技术更是打破了传统的数据生命周期。值得注意的是，数据一旦在生成源头失控，就有可能瞬间扩散并长期保存，增加了追溯和治理的难度。因此，应当加强对个人数据生成源头的治理，通过预防性策略降低数据不当使用和泄露的可能性。制定个人数据负面清单，明确规定禁止或限制采集的数据类型，尤其要对高度敏感的个人信息划定红线。约束个人数据生成场景，按照最小化和知情同意原则规范适当、适时的个人数据采集场景，对超高精度捕获技术和数据关联分析技术限定适用边界，防止个人数据采集“无孔不入”。提升个人数据安全素养，注重宣传引导和教育普及，推广个人数据保护的基本概念、法律规范和实践案例，提高公众对数据采集的敏感意识和应对能力。

建立动态连贯的过程监督机制。面对日益多发的隐蔽数据交易行为，应当加强对数据处理和流通过程的治理，常态化对数据生命周期进行合规性监管。进一步完善个人数据的可追溯管理，全面记录并分析数据处理的各个环节，实时监测数据流通活动，及时发现并精准识别潜在风险，防范数据使用中的不合规行为。开展个人数据的动态审计，通过多层次、多角度的数据审计体系掌握数据安全状况，感知个人数据安全态势，预测数据未来动向，减少数据泄露、滥用和其他安全隐患。引导公众参与个人数据风险评估，广泛吸取社会意见，制定评估标准，建立标准化评估流程，不断优化个人数据保护措施。

建立韧性可靠的应急响应机制。改进处置应对措施，为个人数据安全应急响应建章立制，界定个人数据安全的应急响应对象和应急处置流程，建立个人数据泄露的举报投诉机制，确保在个人数据危机事件发生时迅速隔离问题，提高数据处理器器的快速反应能力和资源调度能力。健全灵活有序、务实高效的个人数据安全应急响应机制，加强政府指导和行业协同，在危机发生后的最短时间内恢复数据系统的

健全个人数据保护的 安全治理机制

□ 王宇航

正常运行，减少对业务连续性和机构信任的负面影响，迭代升级并推广普及数据安全防护技术，弥补数据安全漏洞、更新数据安全基础设施，增强数据安全韧性，为个人数据流通利用保驾护航。

建立渐进均衡的维权申诉机制。数字业态下个人信息权益和企业数据加工过程相互交织，呈现出错综复杂的权属形态。应当适应行业、技术和社会发展阶段，统筹平衡个人信息权益保障和数据产业发展需求，搭建申诉反馈渠道，针对用户投诉的隐私泄露或数据滥用问题及时回应处理，链接便捷高效的司法途径和维权机构，确保公正透明的司法和行政响应。示范发布典型案例指引，以判例揭示个人权益保护的执行方式，明确行业合规要求，提供法律框架内的可操作性指引，指导企业在数据收集、处理和使用过程中更加审慎。采用分级分类管理手段，逐步完善个人数据保护措施，通过试点和阶段性法规引导行业自律规范发展，避免对数据行业过度限制性干预。

建立多元协同的责任共识机制。个人数据保护需要全社会共同参与，应当明确不同主体在数据治理中的角色，形成共同决策、共担责任、共建共治的合作共识。政府主导、引领合作，统筹制定政策、法规和监管框架，为个人数据的经营活动制定包括安全公平在内的管理规范，为个人数据使用及其安全提供一切必要条件和制度保障，确保整体社会利益最大化。平台主体要加强行业自律，承担起个人数据安全的行业责任，加强个人数据分类分层利用，综合运用多方安全计算、可信执行空间等技术手段提升数据安全水平。强化个人数据保护意识，承担个人数据保护责任，积极参与信息保护、社会规范和公共利益的社会利益最大化。不断优化个人隐私保护策略以适应新的安全挑战。



提升数字文化产品的供给与传播能力

□ 刘密霞

2024年10月28日，习近平总书记在中央政治局第十七次集体学习时强调，探索文化和科技融合的有效机制，实现文化建设和数字化赋能、数字化转型，把文化资源优势转化为文化发展优势。数字化、信息化已成为建设社会主义文化强国的重要时代特征，信息技术发展正从作为资源、技术、工具带来的物质满足转向满足人民群众的网络文化需求。先进文化在互联网上传播，使互联网成为推进数字文化产业发展和网络文明建设的重要阵地，提升数字文化产品的供给与传播能力是信息化条件下文化领域治理的重要内容。

在数字文化产品供给与传播过程中，发挥社会主义核心价值观引领，发挥主流媒体在发展壮大主流价值、主流舆论、主流文化方面的重要作用，引领网络文化朝着积极向上的方向发展，形成丰富多样、繁荣发达的网络文化，调动全党全国人民的凝聚力和向心力，构建充沛释放正能量和持续激发社会活力的网上文化生态，形成网络文化阵地与传播新格局。数字文化产品供给与传播能力是网络文化阵地建设和网络文化传播秩序构建的国家文化软实力的重要体现，关系网络文化产业的繁荣，文化资源的数字化建设、呈现与推广以及全媒体传播格局和数字文化大市场的构建。

促进网络文化产业繁荣，加强优质数字文化产品供给。建设网络强国

要有繁荣发展的网络文化和信息服务，当前网络文化产业发展与投资市场还不够丰富，开发、运营、服务、消费的链条尚不健全，乡村数字文化产品市场推广的资源和能力不足，内容治理和产业管理等面临的各方面的问题亟须解决。充分调动各方资源，以“融合创新”方式推动企业、行业组织等多元社会力量参与到网络文化建设中并提供服务，增加数字文化产品的市场供给，引导各类平台、网络文化企业和广大网民创作生产积极健康、向上向善的优质数字文化产品，激发文化与信息技术的融合创新，实现数字文化产品供给与需求的精准匹配，充分展示人类优秀文明成果。优化数字文化产业结构，培育具有影响力的数字文化品牌，通过网络直播、云展览、云旅游等方式推动乡村数字文化产业发展，形成数字文化产业规模，不断提升我国网络文化企业的全球影响力。加强网络文化产业管理，构建网络文化产业良好治理生态，通过文化产业基金、财经金融服务等方式推动文化金融服务模式创新，形成文化金融服务体系，提升综合服务能力，助力数字文化产业健康发展。

深入实施国家文化数字化战略，让全民共享中华文化数字化成果。中华文化全景呈现、中华文化数字化成果全民共享是国家文化数字化战略的主要目标，让全体人民共享数字化发展成果，与联合国2030年可持续发展

目标实现的原则“不让任何一个人掉队”不谋而合。加快推进文化资源数字化建设，推动中华优秀传统文化和当代文化精品数字化，以国家文化专网数字化基础设施为依托，搭建文化数据服务平台，推进文化产业数字化转型，提供丰富健康的数字文化服务。整合现有的公共文化数字资源平台，打通各层级公共文化数字资源，统筹城乡一体化发展，数字技术赋能乡村文化，关注乡村文化思维、发展能力、数据访问和连接等因素，促进乡村文化、旅游、产品与平台经济、共享经济深度融合。通过城乡优势互补，形成公共文化服务体系，满足群众公共文化服务新需求。建设中华文化数据库，创新数字文化业态和产品结构布局，开发文化产业数字化应用新场景，推动网络视听、数字文旅等新业态发展，通过资源整合、资源投入、资源分配以及资源监管等综合措施确保公平地向每个人提供数字信息和服务。

加强主流媒体的网络文化舆论引导，构建全媒体传播格局和数字文化大市场。主流媒体具有强大的传播力、引导力、影响力和公信力，促进优质网络文化产品生产传播，离不开主流媒体的网络文化舆论引导。面对新媒体的冲击和影响以及全媒体时代的挑战和机遇，传统主流媒体应牢牢把握网络文化舆论的话语权和主导