

我国公共数据资源高质量供给任务举措研究

我国数据基础制度逐步完善，数据供给基础能力快速提升，供给利用模式多样化发展，数字中国建设已迈上新台阶。然而，我国仍存在公共数据底数不全、治理能力不强、供给动力不足等问题。针对这些问题，文章提出了具有针对性和可操作性的政策建议，旨在为我国下一阶段实现公共数据资源高质量供给和高效利用提供支持。

文 | **宋志刚** 数字中国研究院（福建） **张慧** 北京国脉互联信息顾问有限公司
姜煜坤 数研院（福建）信息产业发展有限公司

一、引言

本文通过系统分析当前公共数据资源供给的现状与问题，借鉴国内外先进经验，围绕公共数据底数管理、数据治理、供给激励和赋能新质生产力等方面，提出具有针对性和可操作性的政策建议，以期为我国下一阶段实现公共数据资源高质量供给和高效利用、降低社会用数成本和用数门槛提供支撑。

二、公共数据资源供给发展现状

目前，我国公共数据资源供给利用基础制度逐步完善，数据资源体系已初步成型，基础设施支撑能力快速提升，供给利用模式不断创新发展，数据资源价

值已得到初步释放。

（一）公共数据基础制度逐步完善

2019年以来，国家出台了《关于构建更加完善的要素市场化配置体制机制的意见》《关于加快建设全国统一大市场的意见》《关于构建数据基础制度更好发挥数据要素作用的意见》《关于加快公共数据资源开发利用的意见》等一系列文件，以《公共数据资源登记管理暂行办法》为代表的一批政策制度也已进入公开征求意见阶段，顶层设计不断加强，公共数据供给、流通、利用工作统筹推进持续发力。

（二）公共数据供给基础能力快速提升



赛迪网官方微信



数字经济官方微信

我国数据资源一体化目标框架已经初步形成，围绕数据资源全生命周期的治理标准、技术标准、安全标准也在加速完善。公共数据资源体系格局不断健全，全国一体化数据中心协同创新体系、东数西算加速布局，全国一体化政务大数据体系建设加速，截至2023年年底，全国一体化政务数据枢纽接入53个国家部门、31个省（区、市）数据，挂接资源达2.06万个，累计调用服务5361.35亿次，公共数据基础供给能力显著提升。

（三）公共数据供给利用模式多样化发展

我国公共数据资源的供给利用模式呈现出多样化的发展趋势。全国范围内已建设226个省市级公共数据开放平台，覆盖市监、工商、交通等多个领域。多个国家部委及省级数据管理部门开始探索公共数据授权运营机制，分别采取统一授权、分场景授权和分级授权等多种授权运营模式，满足社会对公共数据资源的应用需求。行业主管部门积极推进行业数据开发利用，在司法数据利用、社保数据运营等方面展开探索。政企数据融合也不断加深，有18.6%的平台企业和51%的中央企业在数据开发利用过程中应用到工商、气象、交通、地理等公共数据。

三、公共数据资源供给存在的问题及挑战

当前，公共数据资源供给主要存在数据底数不全、治理能力不强、供给动力不足以及新范式不清晰等问题，影响了

数据资源释放价值和技术创新。

（一）公共数据底数不全

目前部分公共数据资源“有数据未登记”“已登记未汇聚”“汇聚不及时”等问题较为突出。一方面，全国31个省市均有发布公共数据资源目录，但目录清单下公共数据项未完成全量归集和汇聚，尤其是水电气等高频数据的归集不能完全满足应用场景建设需求，铁路、民航等公共服务运营单位尚未实现常态化数据归集共享，同时部分已归集的公共数据通过人工报送，数据完整性、时效性难以保障。另一方面，《公共数据资源登记管理暂行办法》尚未正式出台，各省市对“应编尽编”“全量登记”界定不统一、统计口径不一致。此外，虽然浙江、广东、海南等地在“十四五”期间基本完成公共数据普查和全量编目，但仍有大部分省市公共数据资源目录有待补充、完善和动态更新。

（二）数据治理能力不强

数据治理存在的堵点卡点造成数据质量不高等问题，直接制约了公共数据高质量供给。一是“一数一源”作用发挥不足，各地公共数据采集规范和标准不健全，造成了数据资源重复、多余，数据采集质量不高，数据清洗比对工作大幅提高了人力物力等成本。二是公共数据数源单位的数据供给方式单一，仍以库表交换、API接口交换等方式为主，难以应对各种复杂的数据对接汇聚。三是中西部地区受制于人才、技术和资金等客观因素，数据基础能力薄弱，区域高质量数据供给“东强西弱，南强北弱”

现象明显。四是受制于转型成本高、信心不足、回报低于预期等顾虑和问题，行业数据供给“技术型行业强、服务型行业弱”现象较为突出。

（三）数据供给动力不足

由全责边界不清、激励机制错配、数据供给保守等原因导致的数据供给动力不足问题，是制约公共数据供给能力构建提升的又一突出矛盾。一是公共数据提供部门、使用部门、管理部门等不同主体，尚未明确划分供数、治数、用数、管数等各环节涉及部门的权责边界，各方为降低安全管理边界成本，导致数据安全底线演变成了数据发展红线。二是目前北京、广东、上海等地已发布的激励措施主要集中在数据商层面，但数源部门对外供数过程中涉及的资金投入和人力投入等成本没有相对应的补偿机制。三是税务、海关等国家垂直部门数据供给保守，业务系统仍然孤岛林立，以查询为主的“假共享、无流通、弱应用”现象较为普遍；供电、不动产、婚姻金融等场景所需核心数据，常以数据安全为由，拒绝对外提供数据产品和服务。

（四）数据新范式不清晰

当前我国推动经济高质量发展、促进新质生产力方面作出了战略部署，但是现有的数据供给无法与日益提升的新质生产力发展需求相匹配。一是面向新质生产力的公共数据供给存量不平衡和增量不充分的问题较为突出，新能源、新材料、高端装备等关键领域的重点科研数据游离于公共数据平台之外，且高质量科研数据供给存量不均衡。二是人工

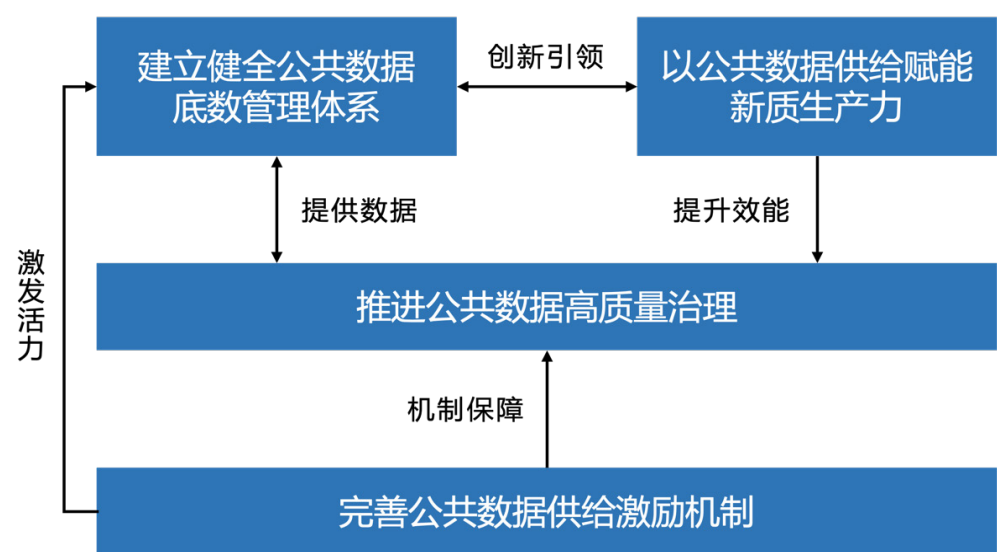
智能领域的高质量数据供给面临技术瓶颈，发展面临着高质量数据匮乏的问题，其中数据标签是人工智能应用的最大障碍。目前人工智能所需的数据标注和数据训练等专业数据服务尚处于起步阶段，存在人工智能所需的原始数据不充分、数据标准不统一、数据质量不达标等现实问题。

四、公共数据资源高质量供给举措建议

基于公共数据资源发展现存问题，建议通过四大任务举措，实现数据资源高质量供给。一是建立健全公共数据底数管理体系，摸清数据资源底数，为公共数据基础制定、扩大公共数据供给奠定基础；二是推进公共数据高质量治理，强化源头治理、降低区域发展不平衡，提升公共数据整体供给质量；三是完善公共数据供给激励机制，以需求为导向，不断提升数源部门的供数积极性；四是以公共数据供给赋能新质生产力，结合最新技术发展趋势，打造数据增量，服务新质生产力发展。

（一）建立健全公共数据底数管理体系

一是进一步完善公共数据资源普查工作机制。以全国公共数据资源“一本账”管理为目标，深化开展全国数据资源调查工作，健全全国统一公共数据资源普查工作机制、完善调查范围、细化工作方法，推动解决当前公共数据台账不全、数据基准性不高、数据碎片化等问题，强化普查成果和公共数据共享开放与授



来源：数字中国研究院（福建）

图 1 公共数据高质量供给任务举措思路图

权运营安排联动管理。

二是公共数据分类施策管理。重点对政务数据和教育、医疗、水电煤气、交通通信、民航铁路等具有公共职能的公共企事业单位数据，以及科研等需求较高数据，制定公共数据分类分级配套管理标准和制度体系，逐步完善公共数据资源分级分类实施细则和操作指南，指导公共数据资源发挥数据倍增乘数效应。

三是加强公共数据资源登记管理。持续组织完善和更新公共数据目录，依托公共数据资源登记平台和可信可控的区块链底层技术体系，建立公共数据资源基础台账，做到动态更新和上链存证，推动公共数据资源全流程管理。

四是推进重点领域公共数据归集和提质扩容。梳理中央企业、地方国企掌握的公共数据，加大地理信息、卫星导航、

城乡融合、绿色低碳、气象等重点领域公共数据归集和汇聚力度。探索利用联邦学习、多方安全计算等新技术，实现电力等公共服务运营单位和税务、海关等国家派驻机构公共数据归集。

（二）推进公共数据高质量治理

一是深化“一数一源一标准”治理。建立公共数据多元数据治理机制，推动数据主管部门、数据中心、数源部门、数商机构等全方位、多渠道参与公共数据治理。推进信息系统按需升级改造，提升数源部门的数据管理能力。健全数据纠错管理机制，提升数据“发现—反馈—修正—校验”跨层级闭环管理能力。实现数据源头治理、综合治理、系统治理，强化公共数据闭环管理，从根本上解决数据供给的难点和堵点。

二是推进高频事项数据治理。建设

完善主题库、专题库，加强对生态环境、交通出行、教育文化、医疗卫生、社保就业、城建住房、社会救助、法律服务等领域应用支撑，着力提升数据资源利用效率。优化经济治理基础库，进一步整合统计、发改、经信、科技、财政、建设、税务、海关、市场监管等部门经济运行数据，为经济运行分析研判和辅助决策提供数据支撑。

三是加大中西部地区和传统产业数字化转型支持力度。结合中西部产业特色、经济优势和数字经济整体布局，研究制定公共数据汇聚共享、协同优化、复用增效、融合创新等相关政策。推动传统行业数据空间建设，建立产业链协同生态，为传统企业构建无歧义、无冗余、单一数据来源的生产力数据库，提供智能制造解决方案，帮助企业提升自动化、智能化水平。

（三）完善公共数据供给激励机制

一是厘清公共数据相关主体供数权责边界。制定公共数据供给体系管理规则制度体系和全国统一的公共数据资源体系管理规范，规范公共数据提供主体、使用主体和管理主体之间的权责关系，明确公共数据边界、范围和多元治理体系。

二是建立健全公共数据供给奖惩措施。构建公共数据供给效能评价体系，对供数积极、成效显著的数据源单位开展专项激励，对数据部门有数据不愿供、有数据不开放、有需求不配合等行为予以惩罚。建立信息化资金投入与部门数据供给挂钩机制，在政府投资信息化项目管理体系中，增加主动供数、供数积

极且供数有保障的奖励条款或措施。

三是推动央地数据回流。国家层面成立省部公共数据共享协调机制，统筹推进部委与地方之间公共数据安全高效应用，特别是加强垂管系统公共数据资源供给力度，推进垂管系统数据有序回流和应用。

（四）以公共数据供给赋能新质生产力

一是建设高质量公共数据集。建设安全合规的高质量通用数据集，为人工智能大模型预训练及行业大模型监督微调提供语料支撑。建设科学数据集，联合中国社会科学院、中国科学院等科研机构，深入挖掘各类科学数据和科技文献，抽取细粒度知识、融合多来源知识。打造公共性、公益性数据共同空间，提供大模型预训练、模型调优推理、安全评测以及模型管理等能力，降低用数成本，提高用数效率。

二是支持数据标注产业发展。支持数据服务企业围绕数据清洗、数据标注、数据分析等需求，强化数据服务供给能力，并给予配套奖励激励措施。统筹规划建设数据标注基地，围绕地域特色和行业优势，视情况增补新一轮国家数据标注基地。推动成立数据标注联盟，形成数据标注行业标准，建立人工智能产业数据资源清单，提升人工智能数据标注库规模和质量。

责任编辑：徐培炎 xupy@staff.ccidnet.com