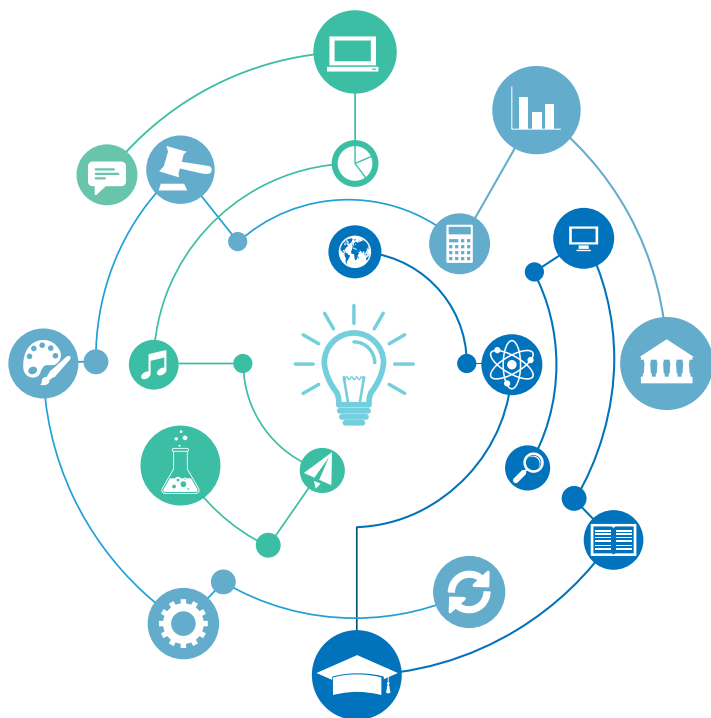


政务数据开发利用研究报告

2021版



全国信息技术标准化技术委员会大数据标准工作组
中国电子技术标准化研究院、国家信息中心

二〇二一年十月

政务数据开发利用研究报告

2021版

全国信息技术标准化技术委员会大数据标准工作组
中国电子技术标准化研究院、国家信息中心

二〇二一年十月

《政务数据开发利用研究报告》（2021版）

■ 专家指导组

梅 宏 周 民 孙文龙 杜小勇 徐春学

■ 编写组

王晓冬 王皓磊 董 超 蔡军霞 范科峰 董 建 张 群 李 冰
尹 卓 王为中 韩晓露 高福信 符海芳 马祎霞 冀 辰 马 力
赵晓靓 曹 杨 程 序 冯念慈 邹丽华 李峰风 张永丽 胡嘉斌
魏 韬 张耀能 成 扬 相明科 何志林 顾 勤 李 正 黎俊茂
何旭珩 张海波 程菲菲 蒋 韬 谭 成 胡绍勇 谢 江 侯中洋
何立鹏 刘祎铭 曹幼林 陈 鹏 张自奇 张超超

■ 编写单位

国家信息中心

华为技术有限公司

东方国信科技股份有限公司

智慧神州（北京）科技有限公司

成都市大数据集团股份有限公司

北京东方金信科技股份有限公司

上海观安信息技术股份有限公司

中国电子技术标准化研究院

中电科大数据研究院有限公司

中国信息通信科技集团有限公司

山西奥斯迪信息工程有限公司

深圳市华傲数据技术有限公司

同盾科技有限公司

四川久远银海软件股份有限公司

版权声明：如需转载或引用，请注明出处。

目 录

一、政务数据开发利用的重大意义	1
1.1 赋能数字政府建设	1
1.2 支撑数字经济创新	2
1.3 推进智慧城市建设	3
二、政务数据开发利用基本概念	4
2.1 相关术语和定义	4
2.2 数据资源要求	5
三、政务数据开发利用现状和问题	8
3.1 各国大数据发展战略	8
3.2 国内政务数据开发利用现状	10
3.3 政务大数据相关政策和标准现状	14
3.4 政务数据开发利用目前存在的问题	20
四、政务数据治理	26
4.1 数据采集汇聚	26
4.2 数据共享交换	30
4.3 公共数据开放	31
4.4 数据开发利用	33
4.5 数据质量管理	34

五、政务数据开发利用参考架构	41
5.1 政务数据总体参考框架	41
5.2 政务数据分类管理	42
5.3 政务数据资产目录管理	44
5.4 政务数据中心框架	50
5.5 政务数据主题联接框架简化	59
5.6 政务数据共享服务框架	60
5.7 开发利用新技术	70
六、实践与探索	76
6.1 成都市公共数据运营服务平台实践案例	76
6.2 乐山市智慧城市大脑案例	79
6.3 重庆市政府大数据资源中心实践案例	83
6.4 山西综改示范区基于数据统筹项目管理体系案例	85
6.5 唐山市政务大数据开发利用实践案例	89
6.6 漳州市数据汇聚共享服务平台建设项目案例	92
6.7 贵阳市政务数据共享开放考核实践探索	96
6.8 成都市公民大数据平台案例	99

一、政务数据开发利用的重大意义

“十四五”时期，是我国全面建成小康社会、实现第一个百年奋斗目标之后，乘势而上开启全面建设社会主义现代化国家新征程、向第二个百年奋斗目标进军的第一个五年，也是新一轮科技革命和产业革命同我国深入实施创新驱动发展战略和实现高质量发展的历史交汇期。党的十九届四中全会将数据列为新型生产要素之一，充分彰显了数据资源的极端重要性。数据日益成为国家战略性资源，大量基础性、关键性的数据掌握在政府手中，这些数据是巨大的“创新财富”，在保障国家秘密、商业秘密和个人隐私的前提下，加强政务数据共享、加快公共数据开放，深入开展数据开发利用，有利于释放数据红利、激发创新活力、创造公共价值，有利于深入实施国家大数据战略、推动落实创新驱动发展战略，有利于加快数字经济发展、实现经济高质量发展。总体而言，政务数据开发利用将在以下几个方面发挥重要作用。

1.1 赋能数字政府建设

近年来，以大数据技术驱动的政务数据开发与利用，无论是在微观事件变化还是在宏观方向决策层面，都深刻地影响着政府治理的各领域、各环节。通过加大对政务数据的开发利用，结合当前5G、人工智能、区块链等新技术，政务工作将实现新的突破，政府治理创新迎来新的变化。

目前，我国已经建成了人口、法人、社会信用、空间地理等一批基础库，以及投资、价格、信用、公共资源交易、就业、社保等主题数据资源，亟待深度开发利用和有序开放。政务数据作为国家的重要数据，不但体量巨大，而且数据具有权威性和可靠性，具备开发利用价值。通过建立从过程可信到结果可信的数据协同生态，在安全管理的基础上实现外部大

数据资源、人工智能算法与政务数据的碰撞，能够充分激发政务数据应用场景，提升政府智能决策水平，同时为更大范围、更深层次部门协调应用的产生提供坚实的基础保障。

在公共服务方面，随着“放管服”改革的持续发展，“让信息多跑路，让群众少跑腿”的创新服务理念逐步落到实处，“最多跑一次”、“一网通办”、“一门式服务”等应用层出不穷。通过开发利用政务数据，政府公共服务能够为公众提供更丰富贴心的服务选择，带来更便捷舒适的服务体验。通过结合智慧管理、精细管理的理念对政务数据进行开发，建立智慧社区、智慧电网、智慧交通，能够有效提高公共管理的安全水平和效率，降低管理成本和失误率。

在行业监管方面，结合创新的监管方法，汇聚交通出行、餐饮就医、购物服务等多源数据与政务数据碰撞，加强监管数据共享和业务协同，建立“事前事中事后”全流程监管体系，能够为政府应用大数据实现精准的事中事后评估提供保障，并通过监管数据的持续积累，能够进一步通过机器学习和人工智能方法展开负面问题的预测预警，有效防范社会治理中的风险。

1.2 支撑数字经济创新

从发展趋势看，数字经济是全球未来的发展方向。“十四五”时期，科技革命和产业革命将持续深化，新一代信息技术同实体经济加速融合，数据作为新生产要素的创新引擎作用日益凸显，将不断为经济高质量发展注入新动能。我国数字经济发展尚处于初级阶段，数据资源要素的高效配置，是推动数字经济成长的关键所在。为贯彻落实十九届五中全会精神，坚持创新在我国现代化建设全局中的核心地位，把握数字经济战略机遇，政务数据开发利用在健全数据开放的制度规则体系，推进政务数据开放共享和开发利用，鼓励各类主体参与政务数据的采集、开发和应用，构建完

善数据要素基础性制度等方面起着重要作用，对推动实现我国经济社会高质量发展以及增强提升国际塑造力影响力引导力都具有重大现实意义。

1.3 推进智慧城市建设

在大数据发展的背景下，基于对数据挖掘和运用基础上的智慧城市建设是城市发展的必然趋势，是新时期实现城市科学发展、高效管理与公共服务更优化的重要抓手。智慧城市建设可以提升城市的运营管理水平、经济发展水平和市民服务水平，而政府掌握的数据正是实现城市智慧化的关键支撑，另外，公众积极主动参与数据创新应用开发也将有助于发挥城市运行相关数据的社会和经济价值。研究政务数据开发利用的机制，将为智慧城市建设提供充足的数据保障，形成数据驱动城市发展的良好态势，促进城市治理现代化水平的提高，助力政府数字化转型与服务质量的优化，推动城市治理体系与治理能力的现代化进程。

作为智慧城市建设的主要责任主体，政府拥有日常办公、城市运行中产生的大量数据，对这些数据的开放、利用是充分融合与调动城市各构成要素的关键一环，有利于智慧城市建设进程的高效推进。因此，政务数据的有效共享与开放，能够让其在不同主体之间进行科学流转，进而有效提取并整合相关领域的关联数据，使之产生新的价值数据，以此提高政务数据利用率与增值空间，促进以数据为驱动的城市科学决策体系的形成。

二、政务数据开发利用基本概念

2.1 相关术语和定义

2.1.1 政务数据

各级政务部门及其技术支撑单位在履行职责过程中依法采集、生成、存储、管理的各类数据资源。

注：根据可传播范围，政务数据一般包括可共享政务数据、可开放公共数据及不宜开放共享政务数据。

2.1.2 公共数据开放

政务部门在安全保密、公共利益导向前提下，面向公民、法人和其他组织以非排他形式提供政务数据的行为。

2.1.3 政务数据共享

各级政务部门因履行职责需要，使用其他政务部门的政务数据以及为其他政务部门提供政务数据的行为。

2.1.4 政务信息资源目录

通过对政务信息资源依据规范的元数据描述，按照一定的分类方法进行排序和编码的一组信息。

注：一般用以描述各个政务信息资源的特征，以便于对政务数据的检索、定位与获取。

2.1.5 数据共享交换平台

面向各级政务部门，维护政务信息资源目录、汇聚基础信息资源、管理共享交换需求，提供数据安全和运行监控保障、支撑跨部门、跨地区开展政务数据共享交换的关键信息基础设施。

2.1.6 公共数据开放平台

基于开放目录汇聚政务部门可开放政务数据，面向社会提供可机读、可下载数据服务的关键信息基础设施。

2.1.7 政务数据提供者

基于统一数据共享交换平台,利用各种技术向其他政府部门、事业单位、企业或公众提供政务数据的实体。

2.1.8 政务数据使用者

使用政务数据的实体。

注：包括政府部门、团体机构、企事业单位和个人。

2.2 数据资源要求

2.2.1 政务信息资源目录要求

政务信息资源目录要求如下：

1.政务信息资源目录是开展各政务部门之间数据共享交换及向社会开放公共数据的依据和导引；

2.政务信息资源目录编制工作应包括政务信息资源分类、元数据描述、代码规划和目录编制，以及相关工作的组织、流程、要求等方面的内容；

3.开放目录和共享目录独立存在时，同一政务信息资源的相同信息项信息在共享和开放目录中应统一；

4.政务信息资源目录中的政务信息资源应按照资源属性、涉密属性、共享属性、层级属性等属性进行分类；

5.政务信息资源目录元数据应包括核心元数据和扩展元数据；

6.政务信息资源目录核心元数据应包括信息资源分类、信息资源名

称、信息资源代码、信息资源提供方、信息资源提供方代码、信息资源摘要、信息资源格式、信息项信息、共享属性、开放属性、更新周期、发布日期、关联资源代码。

2.2.2 政务数据内容要求

（一）数据格式要求

- 1.政务数据提供者应提供可机读的电子格式及相关软件版本信息；
- 2.数据集应以开放的、非专属的格式提供，包括电子文件格式、电子表格格式、图形图像文件格式、流媒体文件格式、自描述格式等；
- 3.数据库类格式需明确具体的数据库表结构定义；
- 4.特殊行业领域数据应由数据提供方提出其特殊行业领域的通用格式；
- 5.无法按照要求的形式提供数据的，可通过安排查阅相关资料、提供复制件或者其他适当形式提供。

（二）数据质量要求

数据质量责任要求：

- 1.政务数据提供者应对提供的数据质量负责；
- 2.原始数据提供者应对原始数据质量负责；
- 3.数据加工者应对加工后的数据质量负责。

数据质量要求：

- 1.规范性：数据信息项定义应优先采用国家、行业相关数据标准，实施开放共享数据工程前应先统一数据标准；
- 2.完整性：对于应满足需求的数据记录不应有缺失、重复，数据信息应完整，对敏感数据脱敏应保证最小颗粒度数据的完整展现；
- 3.一致性：不同数据集中描述同一对象的同一度量值在信息含义上不可冲突；
- 4.准确性：数据记录的内容应真实反映实际情况且有效，不能存在异

常或错误，不能出现不可识别的内容。

数据质量需求、数据质量检查、数据质量分析、数据质量提升等数据能力要求应遵循GB/T 36073-2018的规定。

2.2.3 数据更新与完善要求

- 1.应根据情况变化对政务信息资源目录进行更新维护；
- 2.应建立政务信息资源更新机制，进行动态管理；
- 3.政务数据提供者应确保提供的数据信息完善，确保数据得到及时、持续更新；
- 4.开放共享的信息内容发生变化时，政务数据提供者应当及时报告本级政务信息资源主管部门，更新相应政务信息资源目录和内容，并通知该政务数据使用者；
- 5.政务数据使用者应当及时比对和更新所获取的政务信息资源，确保数据一致性。

三、政务数据开发利用现状和问题

3.1 各国大数据发展战略

世界各国政府对政务数据应用日益重视，2012年3月29日，奥巴马政府率先推出了《联邦大数据研究与开发计划》，2014年5月，发布了《大数据：抓住机遇、保存价值》白皮书。2016年5月，美国网络和信息技术研发计划（NITRD）的大数据高级领导小组（SSG）编制发布了《联邦大数据研究和发展战略计划》，这是对美国大数据战略的延续和落实，其目的是对联邦机构的大数据相关项目和投资进行指导，促进各联邦部门深化大数据分析利用。该计划提出了提高大数据处理能力，加强研发投入，健全基础设施，促进数据共享，注重隐私、安全、伦理，强化教育培训，促进大数据创新生态系统协同创新等七个方面的战略计划，强调了网络与信息技术开发机构的项目和研究资金将在这七个方面重点支持突破的预期成果，确保美国在研发领域继续发挥领导作用，通过研发来提高美国和世界解决紧迫社会和环境问题的能力等。该计划还介绍了美国联邦机构的全国火灾事故报告系统（NFIRS）、科学气候大数据可视化、能源科学网络（ESnet）、对抗视力衰退、全球综合地球观测系统（GEOSS）、大数据知识（BD2K）计划、材料基因组计划（MGI）等大数据创新工作。

欧盟致力于数据的流动和广泛使用，希望通过建立数据市场，确保欧洲在未来成为世界上最具吸引力、最安全、最具活力的数据敏捷型经济体。2020年2月19日，欧盟委员会公布了《欧盟数据战略》，提出了四大支柱性战略措施：一是构建跨部门治理框架，二是加强数据投入，三是提升数据素养，四是构建数据空间。以数字经济发展为主要视角，概述了其在数据方面的核心政策措施及未来五年的投资计划，以助力数字经济发展。

英国将人工智能上升为国家战略，以政策引领推动全社会积极参与人工智能发展。与此同时，英国特别注重政策的可实施性，确保人工智能能够在英国“落地”，取得实实在在的效益。英国上议院于2017年牵头成立专门委员会对英国人工智能发展中所面临的重要问题和挑战进行了全面和深入的调研，并提出了解决办法和政策建议，形成了《英国的人工智能发展报告——准备、意愿和能力》。报告认为，人工智能的发展涉及面较广，不仅需要做好数据的开放利用和保护，开发“可被理解的人工智能技术”，重视大数据、人工智能创新和人才培养，还要正确处理好大数据、人工智能在就业、医疗和道德等领域对社会造成的风险和挑战。

2020年7月17日，日本政府发布《统合创新战略2020》，阐述了日本开展科技创新的年度路线方针。战略由以下四个部分组成：积极应对新冠肺炎疫情，构建具有韧性的经济和社会结构；持续推进社会5.0建设；强化研究能力；重点推动基础技术等研究领域。落实2019年6月G20大阪峰会提出的“可信赖的信息自由流通”理念，整合不同领域产生的数据以创造新的价值。从振兴地方经济发展和便利民众生活的角度，加快建设智慧城市并与国际社会交流经验。

除各个国家的数据战略外，国际组织也十分强调数据在全球化发展中的重要性。2020年4月，世界银行呼吁各国政府、相关企业以及学术界共同合作，通过大数据等技术手段应对新冠肺炎疫情所带来的危机。在2020年7月召开的G20数字经济部长会议中，数据流动成为各国部长们重点讨论的议题之一，在新一轮的国际经贸规则中，跨境数据流通成为全球双边/多边贸易合作的重要议题。一方面，基于“共同理念”的全球数据同盟体系加速构建，形成了欧盟GDPR和APEC跨境隐私规则体系（CBPR）两大区域性的数据隐私与保护监管框架，众多国家以二者为蓝本，对本国的数据跨境与数据保护规则进行修订；另一方面，两大框架在国与国、区域与区域之间衍生诸多灵活性的解决方案。2019年，日韩分别启动与美欧

之间的推动跨境数据流动的双边协定，并与欧盟达成充分性保护互认协议。2020年3月，澳大利亚信息专员办公室（OAIC）与新加坡个人数据保护委员会（PDPC）签订关于跨境数据流动的谅解备忘录，加强数据治理方面的合作，促进两国之间的经济一体化；2020年6月，英国宣布脱欧后的未来科技贸易战略，允许英国和某些亚太国家间的数据自由流动，并希望与日本等国达成比其作为欧盟成员国时期更进一步的数据协议。

3.2 国内政务数据开发利用现状

近年来，我国大数据产业蓬勃发展，融合应用不断深化，数字经济量质提升，对经济社会的创新驱动、融合带动作用显著增强。主要体现在以下几个方面：

3.2.1 大数据产业发展政策环境日益完善

自2015年国务院发布《促进大数据发展行动纲要》以来，各地陆续出台了促进大数据产业发展的多项文件，包括规划、行动计划和指导意见等。据统计，截至2020年底，全国共有30个省份出台了56份政务数据开放的相关政策文件（详细内容请见本章第三节）。

在颁布的这些文件中，可以看出各地除了注重推进大数据产业以外，也在更多的关注政务服务和产业数字化等方面的内容，体现了大数据与行业应用的结合，同时还加入了人工智能、数字经济等内容，将新一代信息技术应用纳入考量的范畴。

在确定数据共享开放内容方面，各地坚持需求导向、目标导向，征求行业协会、相关企业、社会公众和行业主管部门的意见建议，形成数据开放目录清单，并动态调整。在共享开放数据范围方面，由政务数据扩展至企事业单位所涉及的公共数据资源；在共享开放数据成果形式方面，提供包括服务应用、数据可视化、研究成果、创新方案等数据共享开放形式。

3.2.2 各地大数据主管机构陆续成立

近年来，各省市陆续成立了大数据局等相关机构，对当地大数据发展进行统一管理。这些机构中大多数都包含制订地方大数据战略规划的功能，但在产业发展政策制订、数据资源整合、数据资源开放共享、电子政务系统建设、信息安全、政府网站建设等方面的职能则并非所有大数据主管机构都具备。

目前国内政务数据的开发利用主要基于政务数据开放平台，政务数据开放平台的管理机构主要是大数据管理机构。例如，北京市政府大数据管理机构为北京市经济和信息化局（北京大数据管理局），上海市政府大数据管理机构为上海市大数据管理中心，天津市政府大数据管理机构为天津市大数据管理中心，江苏省政府大数据管理机构为江苏省大数据管理中心，四川省政府大数据管理机构为四川省大数据中心，贵州省、浙江省、内蒙古自治区等专门设大数据发展管理局管理政府大数据的开发和利用，福建省、海南省、河南省政府大数据开发和利用的管理机构为专门的大数据管理局，其他地方政务数据开发利用的管理机构也都在政务数据公开主管部门。部分地区政务数据管理机构见下表：

表3-1 部分省级政务大数据机构

省级行政区	设立时间	机构名称	隶属机构
北京市	2018年	北京市经济和信息化局（市大数据管理局）	北京市政府
上海市	2018年	上海市大数据中心	上海市政府办公厅
天津市	2018年	天津市大数据管理中心	天津市委网信办
重庆市	2018年	重庆市大数据应用发展管理局	重庆市政府
广东省	2014年	广东省大数据管理局	广东省经信委
	2018年	广东省政务服务数据管理局	广东省政府办公厅
贵州省	2015年	贵州省大数据发展管理局	贵州省人民政府
内蒙古自治区	2017年	内蒙古自治区大数据发展管理局	内蒙古自治区政府

省级行政区	设立时间	机构名称	隶属机构
江西省	2017年	江西省信息中心（省大数据中心）	江西省发改委
江苏省	2017年	江苏省大数据管理中心	江苏省政务服务管理办公室
吉林省	2018年	吉林省政务服务和数字化建设局	吉林省人民政府
陕西省	2017年/2018年	陕西省工信厅（省政务数据服务局）和陕西省大数据管理与服务中心	陕西省工信厅
广西壮族自治区	2018年	广西壮族自治区大数据发展局	广西壮族自治区政府
浙江省	2018年	浙江省大数据发展管理局	浙江省政府办公厅
安徽省	2018年	安徽省数据资源管理局（省政务服务管理局）	安徽省人民政府
福建省	2018年	数字福建建设领导小组办公室（省大数据管理局）	福建省发改委
山东省	2018年	山东省大数据局	山东省政府
河南省	2018年	河南省大数据管理局	河南省办公厅
黑龙江省	2019年	黑龙江省政务大数据中心	黑龙江省营商环境建设监督局
湖北省	2019年	湖北省大数据中心	湖北省政务管理办公室
海南省	2019年	海南省大数据管理局	海南省人民政府
四川省	2019年	四川省大数据中心	四川省人民政府

3.2.3 开放平台建设日趋完善

政务数据开发和利用主要依靠数据开放平台，在我国，东部沿海发达地区和部分中部城市率先建设了政务数据开放平台，其他中西部地区政府也都在加快建设数据开放平台。在2018年1月上旬，中央网信办、发展改革委、工业和信息化部联合印发《公共信息资源开放试点工作方案》之后，多个省市自治区结合当地特色，启动省市平台互通和省直部门接入工作，并实现与国家平台对接。目前我国地方政务数据开放的整体情况非常

迅速，各地政务数据开放直接相关的法规政策和标准规范实现了从无到有、逐渐落地的发展态势，这些政策驱动各地政府部门数据开放加速，地方政务数据开放平台数量和开放的有效数据集呈现爆发式增长。2012-2017年我国地方政府上线的数据开放平台的数量年均复合增长率为46.1%，2017-2020年达到92.2%。截至2020年底，我国已有142个省级、副省级和地级政府上线了数据开放平台，与2019年下半年相比，新增了4个省级平台和36个地级(含副省级)平台，平台总数增长近4成。

3.2.4 开放数据集增长明显

对政务数据的开发利用通常以下载或API接口的方式提供数据集，数据集总量统计的是平台上可通过下载或API接口获取的有效数据集总数。从有效数据集总量来看，全国各地政府开放的有效数据集总数从2017年的8398个，增长至2020年的98558个，四年间增长超过10倍。部分省市共享开放平台数据统计见下表。

表3-2 部分地区政务数据开放平台数据统计

地区	政务数据开放平台	部门	数据集 (个)	数据项 (个)	数据量 (亿条)	API (个)
北京	北京市政务数据资源网	103	10,198	444,332	57.18	-
上海	上海市公共数据开放平台	51	5,204	44,994	9.91	2,396
天津	天津市信息资源统一开放平台	50	1,058	-	-	501
贵州	贵州省政务数据开放平台	83	3744	-	0.149	381
浙江	浙江省公共数据开放平台	-	18,493	92,193	51.96	9484
山东	山东公共数据开放网	51	-	154,148	28.3	66000
广东	“开放广东”政务数据统一开放平台	51	5,103	-	1.515	-
重庆	重庆市公共数据资源开放系统	48	800	-	-	-
福建	福建省公共信息资源统一开放平台	36	2,205	-	8	563

3.3 政务大数据相关政策和标准现状

3.3.1 政务大数据相关政策现状

（一）国家层面相关政策

为深化大数据的应用和发展，我国政府相继出台了一系列政策。相关政策的出台不仅推动了政务大数据的发展，也对政务大数据的建设过程具有一定的指导、规范和保障作用。

从国内立法政策来看，在早期阶段，我国政务数据开放主要是以政府信息公开为主。2007年国务院制定并发布了《中华人民共和国政府信息公开条例》（以下简称《条例》），《条例》规定了政府信息公开的基本原则、公开的范围、公开的方式和程序以及监督和保障措施，初步奠定了我国政府信息公开的基础。同时，为正确审理政府信息公开行政案件，最高人民法院于2011年发布了《关于审理政府信息公开行政案件若干问题的规定》的司法解释，进一步明确了政府信息公开中的相关操作问题。

2015年以来，随着大数据技术发展和产业推动，我国也从政府信息公开转向强调政务数据开放。在党中央、国务院发布的多个文件都对推进政务数据开放进行了不同程度的部署，包括：《关于促进云计算创新发展培育信息产业新业态的意见》《关于大力推进大众创业万众创新若干政策措施的意见》《国务院办公厅关于运用大数据加强对市场主体服务和监管的若干意见》《国务院关于积极推进“互联网+”行动的指导意见》《促进大数据发展行动纲要》《关于加快构建大众创业万众创新支撑平台的指导意见》《关于全面推进政务公开工作的意见》《2016年推进简政放权放管结合优化服务改革工作要点》以及《国家信息化发展战略纲要》等。其中，国务院《促进大数据发展行动纲要》明确提出2017年底前形成跨部门数据资源共享共用格局，在2018年底前建成国家政务数据统一开放平台；《国家信息化发展战略纲要》要求建立公共信息资源开放目录，构建统一规范、互联互通、安全可控的国家数据开放体系，积极稳妥推进公共信息

资源开放共享。

2016年9月，国务院发布《关于加快推进“互联网+政务服务”工作的指导意见》（国发〔2016〕55号），重点提出“推进政务信息共享、加快新型智慧城市建设、建立健全制度标准规范、完善网络基础设施、加强网络和信息安全保护”。

2016年9月，国务院发布《关于政务信息资源共享管理暂行办法》（国发〔2016〕51号），要求加快推动政务信息系统互联和公共数据共享。

2016年12月，国务院办公厅发布《关于印发“互联网+政务服务”技术体系建设指南的通知》（国办函〔2016〕108号），通知明确了政务服务信息的汇聚发布与展示、政务服务事项的一体化办理、互联互通与信息共享等相关要求。

2017年5月，国务院印发《政务信息系统整合共享实施方案》（国办发〔2017〕39号），强调加快消除“僵尸”信息系统；推进整合，加快部门内部信息系统整合共享。

2017年7月，国家发改委、中央网信办联合印发《政务信息资源目录编制指南试行》（发改高技〔2017〕1272号），指导加快建立政务数据资源目录体系，推进政务数据资源的国家统筹管理。

2018年7月，国务院发布《关于加快推进全国一体化在线政务服务平台建设的指导意见》（国发〔2018〕27号），要求深入推进“互联网+政务服务”，加快建设地方和部门政务服务平台。

2020年3月发布的《中共中央、国务院关于构建更加完善的要素市场化配置体制机制的意见》，要求推进政务数据共享开放、提升社会数据资源价值、加强数据资源整合和安全保护，同时提出要探索建立统一规范的数据管理制度、提高数据质量和规范性、丰富数据产品。

（二）地方相关政策

围绕政务大数据发展和应用需求，各省市也纷纷出台相关数据政策和规划，积极推动当地数字经济产业发展。部分省市出台的政务数据相关政策规划见下表。

表3-3 部分省市政务数据相关政策规划

发布省市	文件名称	发布时间
北京市	北京市政务数据开放服务指南（试行）	2020年09月
天津市	《天津市促进大数据发展应用条例》	2019年01月
重庆市	《重庆市以大数据智能化引领的创新驱动发展战略行动计划（2018-2020年）》	2018年03月
广东省	《广东省政务数据资源共享管理办法（试行）》	2018年11月
山东省	《数字山东发展规划（2018-2022年）》	2019年02月
	《数字山东2019行动方案》	2019年03月
山西省	《山西省政务数据管理与应用办法》	2020年12月
广西壮族自治区	《广西数字社会建设三年行动计划（2018—2020年）》	2018年08月
湖南省	《湖南省大数据产业发展三年行动计划（2019-2021年）》	2019年01月
湖北省	《湖北省政务数据资源应用与管理办法》	2021年01月
河南省	《河南省政务数据管理办法（征求意见稿）》	2020年12月
贵州省	《贵州省大数据战略行动2019年工作要点》	2019年03月
四川省	《四川省促进大数据发展工作方案的通知》	2018年01月
安徽省	《安徽省政务数据资源管理办法》	2021年01月
深圳市	《深圳经济特区数据条例（征求意见稿）》	2021年06月

3.3.2 政务大数据标准化现状

（一）国家标准化现状

标准化是促进政务大数据相关技术发展和开发应用的重要基础和必然途径，为推动和规范我国政务数据的汇聚、整合和开发利用，在国家层面，依托全国信息技术标准化技术委员会，先后成立了电子政务标准工作

组、公共数据开发利用专题组、大数据安全标准特别工作组。

电子政务标准工作组于2020年组织制定了《国家电子政务标准体系建设指南》，该指南致力于推动电子政务标准化发展，优化电子政务标准体系，完善电子政务标准，突破关键技术标准，强化标准实施与监督，增强电子标准化服务能力，为推进国家电子政务健康协调可持续发展、推进国家治理体系和治理能力现代化，提供有力支撑。

公共数据开发利用专题组是全国信标委大数据标准工作组下设的重点专题组之一，负责政务大数据方向的标准化工作，目前，专题组已研制并发布了3项政务数据开放共享方面的国家标准。

大数据安全标准特别工作组由全国信安标委设置，主要聚集于大数据安全领域的研究工作，目前，已开展多项大数据安全的国家标准研制工作。

截止目前，政务数据领域研制的主要国家标准如下表所示。

表3-4 政务数据领域部分国家标准

序号	内容方向	标准号 / 计划号	标准名称
1	数据元	GB/T 19488.1-2004	电子政务数据元 第1部分：设计和管理规范
2		GB/T 19488.2-2008	电子政务数据元 第2部分：公共数据目录
3		GB/T 32617-2016	政务服务中心信息公开数据规范
4		GB/T 39046-2020	政务服务平台基础数据规范
5	信息资源目录	GB/T 21063.1-2007	政务信息资源目录体系 第1部分：总体框架
6		GB/T 21063.2-2007	政务信息资源目录体系 第2部分：技术要求
7		GB/T 21063.3-2007	政务信息资源目录体系 第3部分：核心数据
8		GB/T 21063.4-2007	政务信息资源目录体系 第4部分：政务信息资源分类
9		GB/T 21063.6-2007	政务信息资源目录体系 第6部分：技术管理要求

序号	内容方向	标准号 / 计划号	标准名称
10	开放共享	GB/T 21062.1-2007	政务信息资源交换体系 第1部分： 总体框架
11		GB/T 21062.2-2007	政务信息资源交换体系 第2部分： 技术要求
12		GB/T 21062.3-2007	政务信息资源交换体系 第3部分： 数据接口规范
13		GB/T 21062.4-2007	政务信息资源交换体系 第4部分： 技术管理要求
14		GB/T 38664.1-2020	信息技术 大数据 政务数据开放 共享 第1部分：总则
15		GB/T 38664.2-2020	信息技术 大数据 政务数据开放 共享 第2部分：基本要求
16		GB/T 38664.3-2020	信息技术 大数据 政务数据开放 共享 第3部分：开放程度评价
17		GB/T 33780.5-2021	基于云计算的电子政务公共平台 技术规范 第5部分：信息资源开 放共享系统架构
18	安全管理	GB/Z 24294.1-2018	信息安全技术 基于互联网电子政 务信息安全实施指南 第1部分： 总则
19		GB/Z 24294.2-2018	信息安全技术 基于互联网电子政 务信息安全实施指南 第2部分： 接入控制与安全交换
20		GB/Z 24294.3-2018	信息安全技术 基于互联网电子政 务信息安全实施指南 第3部分： 身份认证与授权管理
21		GB/Z 24294.4-2018	信息安全技术 基于互联网电子政 务信息安全实施指南 第4部分： 终端安全防护
22		GB/T 34080.2-2017	基于云计算的电子政务公共平台 安全规范 第2部分：信息资源安 全
23		GB/T 39477-2020	信息安全技术 政务信息共享 数据 安全技术要求
24		GB/T 38667-2020	信息技术 大数据 数据分类指南

（二）地方标准化现状

在相关政策和需求促进下，全国各地相继成立地方性大数据、数字政府等专业标准化技术委员会，逐步规范当地政务大数据标准化工作，继而服务当地数字经济产业发展。部分省级大数据以及数字政府标准化技术委员会如下表所示。

表3-5 部分省级大数据&数字政府标准化技术委员会

序号	标准化技术委员会名称	成立时间
1	贵州省大数据标准化技术委员会	2017年
2	广东省大数据标准化技术委员会	2017年
3	内蒙古自治区云计算与大数据标准化技术委员会	2017年
4	山东省大数据标准化技术委员会	2018年
5	山西省网络安全和大数据信息技术标准化技术委员会	2019年
6	上海市公共数据标准化技术委员会	2020年
7	广东省数字政府标准化技术委员会	2020年

依托省级专业标准化技术委员，各地研制、发布了一批政务大数据相关地方标准，内容涵盖了政务数据平台、数据清洗加工、数据共享开放、数据安全等多个方面。政务数据领域部分省市地方标准如下表所示。

表3-6 政务领域部分省市地方标准

序号	标准号	标准名称	发布省市
1	DB52/T 1540.2-2021	政务数据 第2部分：元数据管理规范	贵州省
2	DB52/T 1541.3-2020	政务数据平台 第3部分：数据存储规范	
3	DB52/T 1541.2-2020	政务数据平台 第2部分：数据归集规范	
4	DB52/T 1540.3-2020	政务数据 第3部分：数据清洗加工规范	
5	DB52/T 1541.4-2021	政务数据平台 第4部分：接口规范	
6	DB52/T 1541.6-2021	政务数据平台 第6部分：面向全网搜索应用的数据处理规范	

序号	标准号	标准名称	发布省市
7	DB35/T 1777-2018	政务数据汇聚 数据集的规范化描述	福建省
8	DB35/T 1967-2021	政务数据汇聚 人口数据描述	
9	DB35/T 1966-2021	政务数据汇聚 企业法人数据描述	
10	DB15/T 1591-2019	政务数据生命周期管理规范	内蒙古自治区
11	DB15/T 2104-2021	政务数据开放共享 元数据	
12	DB37/T 3520-2019	政务数据资源体系 总体架构	山东省
13	DB36/T 1098-2018	政务数据开放平台技术规范	江西省
14	DB36/T 1179-2019	政务数据共享技术规范	
15	DB51/T 2693-2020	地理空间政务数据资源目录规范	四川省
16	DB44/T 2111-2018	电子政务数据资源开放数据管理规范	广东省
17	DB44/T 2110-2018	电子政务数据资源开放数据技术规范	
18	DB3301/T 0169-2018	智慧电子政务数据资源共享管理规范	杭州市
19	DB3301/T 0276-2018	政务数据共享安全管理规范	
20	DB3301/T 0322.4-2020	数据资源管理 第4部分：政务数据共享流程	
21	DB3301/T 0322.3-2020	数据资源管理 第3部分：政务数据分类分级	
22	DB3301/T 0322.2-2020	数据资源管理 第2部分：政务数据安全责任	
23	DB3301/T 0322.1-2020	数据资源管理 第1部分：政务数据安全监管	

3.4 政务数据开发利用目前存在的问题

3.4.1 权责问题

政务数据的权责问题主要包括权属和责任两个层面的问题。当前现行法律法规针对数据权属缺乏明确规定，首先，我国尚未出台专门适用于规范数据权利的法律法规，只有部分相关条款散落于《民法典》《刑法》《网络安全法》中，进行了原则性、模糊性的规定。但是，从所调整的数

据社会关系（围绕数据生成、存储、应用与监管过程中所形成的社会关系）来看，现有的法律法规并不完全适用。

其次，当从数据产权上看，数据的特殊性决定其与我国现行产权制度下的客体不完全相符，数据财产属性和数据权利属性不明确，因此以物权、债权和知识产权为代表的财产权制度就难以直接用于数据；关于政务数据的权利归属，国内多个省市展开了积极的探索。比如《福建省政务数据管理办法》中规定“政务数据资源属于国家所有，纳入国有资产管理”；《重庆市政务数据资源管理暂行办法》也认为“政务数据资源属于国家所有”；《长沙市政务数据资源管理暂行办法》则具体到“长沙市各级政务部门根据法定职责依法履职产生和采集的数据所有权归长沙市人民政府所有”，但更多的省份都暂时回避了政务数据的权属问题。

再次，数据责任方面，国内各省、市都通过专门的政务数据管理办法，进行了政务数据开发开放共享，或开发利用的责任划分。这种落实到部门和具体人员的责任和考核机制，一方面加强了数据提供部门的责任心，强化了对数据使用环节的监督管理，但另一方面，在实际工作中也往往会造成各部门行事的谨小慎微，尤其对于某些存有高敏感信息，而又对其他部门数据需求较小的部门而言，这种责任机制，往往容易成为其不提供数据的借口。随着政务数据相关业务的不断拓展，所涉及的利害关系也将更加的突出。政务数据的责任体系，仍需要在工作实践中逐步完善，以实现各部门间权责的均衡，促进政务数据的开发利用。

政务数据开发利用中的权责问题，主要表现在四个方面：一是在项目全生命周期各阶段，各方干系方关于政务数据生产流通的权责不清晰；二是在政务数据资产管理体系中，数据资产开发、管理、运维等各干系方权责边界不清晰；三是政务数据流通及开发利用等相干系方安全责任边界不清晰；四是流通及开发利用安全审计责任主体缺失。

当前，政务数据的权属不清、权责不均衡等问题，一方面让政务数据

的汇聚工作困难重重，另一方面也使得政务数据的二次开发利用面临较大的法律风险。政务数据权责不明确已经成为阻碍政务数据流通和开发利用的核心问题。明确政务数据的权利归属与责任范围，已经是政务数据汇聚和开发利用的迫切需求。

3.4.2 机制体制问题

2016年，国务院印发《政务信息资源共享管理暂行办法》以来，各地也先后出台了地方政务数据相关管理办法，并基本建立了由当地政府统筹协调、数据资源主管部门牵头实施，各政务部门，以及第三方单位等广泛参与的，以当地政务大数据开放共享平台为数据汇聚载体的，跨层级、跨部门协同参与机制，对政务数据进行汇聚、开放共享，以及开发利用。在实施过程中，还存在以下几个方面的问题：

一是针对公共数据开发利用环节考核规定及实施细则缺失。在当前地方政府出台的关于政务数据共享开放考核相关文件中，针对政务数据的共享开放考核均有明确规定，但部分地区针对政务数据开发利用环节的考核尚未纳入考核管理范畴。公共数据开发利用考核管理办法及实施细则的缺位一定程度上制约了公共数据的开发利用进程。

二是考核工作实施难客观存在。在公共数据共享开发及开发利用管理办法出台后，部分地区配套的考核指标体系正处于设计阶段，一定程度上制约了考核工作的开展。此外，针对已出台公共数据开发利用考核管理办法的地区，在考核工作推进过程中，考核标准单一、考核过程中主观因素过多等导致具体的考核工作落实不到位等问题依然存在。

三是纵向的数据共享回流机制不畅通。当前上下级的政府部门从采集、流动到共享，往往是单线的，即下级政府负有采集流转义务，却难拥有共享权利。尤其是基层政府，通常基于上级垂管系统填报数据，本地却没有数据留存管理权限。回流机制不畅通，本地或上级数据难获取，导致基层业务应用难展开。

四是尚未建立相关管理责任体系。其中主要包括政务数据集中化管理的责任主体以及管理体系，政务数据资产管理体系和政务数据服务管理体系，以及政务数据安全审计的责任主体以及管理体系。

3.4.3 数据质量问题

政务数据的质量是影响政务数据开发效率、利用深度和应用成效等的关键因素之一。当前，国内各省市政务数据基本上都是由各政务部门自行控制、归集，再汇聚到当地政务数据管理平台进行统一存储、管理和开发利用。但由于目前各地信息化基础条件各异，对政务数据管理工作重视程度不一，在政务数据采集、数据标注、清洗、脱敏脱密等技术处理过程中缺乏统一的规划和标准，加之数据资源牵头主管部门对各政务部门政务数据管理工作的考核机制不健全、约束力不强等原因，导致事政务部门间数据质量参差不齐，并严重制约了政务数据开放、共享，和开发利用的效果。这些质量问题可归纳为六类：一是数据碎片化问题，如某地建设的事件库中，同一个事件，表与表之间较难找到关联性；二是数据重复性问题，同一数据的源数据重复，如疫情自主申报时同一人多次申报，导致相同数据多次推送；三是数据完整性问题，可能行缺失(如企业社保表中统一社会信用代码不全)，可能属性列缺失(如缺失了某些业务字段)；四是数据一致性问题，如医保业务中，病种的分类，不同医院、不同区域、不同医药企业的分类不完全一致，带来统计的数据混乱；五是数据冲突问题，如人口基础库中的人员联系方式与其他部门提供的有冲突，法人基础库中的企业注册地址与信用网中的实时地址有冲突等；六是数据时效性问题，如人口、法人、房屋等基础库中的数据无法做到实时更新，从而不能很好地作为基础数据进行各类业务支撑。

3.4.4 供需衔接问题

政务数据的“交易”，主要通过各政务数据平台或大数据交易平台进

行。通常情况下，政务数据平台直接对接各政务单位，汇聚政务数据，并作为需求对接平台，直接与数据需求方进行需求对接。大数据交易平台则普遍与政务数据汇聚平台的对接而获得政务数据供给，并作为数据和业务的对接和交易平台，与数据需求进行数据需求和业务需求的对接。这种以平台为依托和媒介的供需衔接模式，在保护政务数据安全、规范政务数据开发应用方面起到了积极的作用；但另一方面，数据供需双方缺乏直接、高效的沟通渠道，也经常导致实际业务中供需双方对对方的数据资源与数据需求互不了解，或者由于非专业人员业务理解偏差、部门间沟通衔接不畅等原因导致数据供需匹配不到位，从而大量的政务数据价值得不到有效释放和众多数据需求得不到满足的情况同时并存。

数据供需衔接问题主要分为以下几个层面，一是面向数据流通的数据属性定义不完善；二是区域数据资源台帐建设滞后；三是数据发布与检索机制不健全；四是数据请求、审批管理机制不配套；五是数据使用范围审计缺少有效的方式方法；六是数据预处理与数据需求脱节；七是缺少有效的数据需求培养或挖掘机制；八是缺少数据供需关系监控与调节机制等。

3.4.5 标准规范问题

在大数据时代，数据量井喷的同时也在产生巨量的噪声数据，要使数据成为“要素”，再转变为“资源”，就需要对数据进行标准化处理，以节约数据相关方收集、沟通和信任的成本，促进数据要素的商业交易和流通使用。当前我国数据开放共享还存在诸多不足，跨部门、跨地区、跨层级数据标准不一、数据理解难、机器可读性差、语义分歧等问题未能解决。譬如缺乏数据交易标准，致使种类多、格式多的大数据难以成为大规模交易的产品；缺乏数据质量标准，不利于区分大数据时代下的大数据与普通数据，并判断其是否能产生实际应用价值；缺乏行业数据应用标准，限制了大数据驱动行业的高速发展等问题。

政务数据标准不统一主要体现在以下几个方面：

1.数据范围不统一。比如在国务院2016年印发的《政务信息资源共享管理暂行办法》中的提法是政务信息资源，而北京、四川、成都等地在相关管理办法中，则是从早期的政务信息资源转为了近年的公共数据。由于各地对政务数据的提法不同，定义各异、范围不一，因此导致各地政务数据汇聚的内容存在一定的差异，给全国统筹开发利用造成了一定的困难。

2.数据格式缺乏统一标准规范。由于数据涉及范围涉及多个领域，数据类型包括多种形式，缺乏对每个类别数据和每种格式数据的详细规范要求，在实际执行中也因为数据体量大、复杂度高、人力不足、技术有限等各种原因，使得无法严格按标准执行，从而导致很多地方只是完成了数据的汇聚，但数据无法直接开发利用，还需要大量的数据预处理工作。

3.技术标准不统一。主要包括以下几方面：一是软硬件产品的技术标准；二是分层建设中的层间技术标准；三是数据生产流通的数据格式标准；四是数据流通审计技术标准。

4.管理标准不规范。基层数据专题库的建设没有形成有效的管理规范，如人口库、法人库等；操作层面的管理标准不规范，没有形成统一操作管理流程；围绕数据资产建设、数据共享开放的绩效考核标准不统一。

四、政务数据治理

政务数据治理的目标是通过构建公共数据端到端全生命周期的数据管理体系，形成统一的数据采集汇聚、共享开放、开发利用等过程的决策机制、流程和规则；基于高质量的数据，在确保安全合规的前提下，实现跨部门、跨业务、跨系统的数据流通。通过构建统一的数据开发利用技术体系，提升政务数据管理、分析及服务能力，支撑政务大数据应用场景的快速设计和开发落地，最大化释放政务数据的价值。数据治理总体流程框架如下：

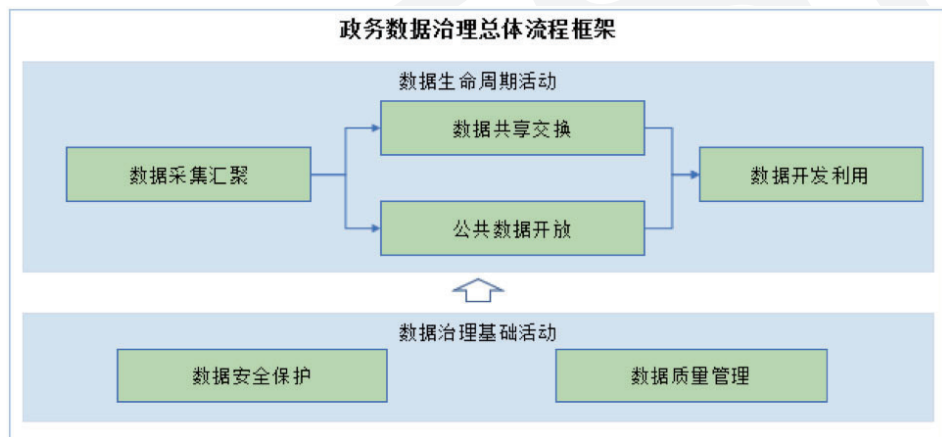


图4-1 政务数据治理总体流程框架

以下重点对数据采集汇聚、数据共享交换、公共数据开放、数据开发利用、数据质量管理等关键环节所涉及的技术流程、规则和技术进行阐述。

4.1 数据采集汇聚

数据采集汇聚是数据创建或获取的阶段，是数据生命周期中的关键点之一。

4.1.1 数据采集汇聚流程

数据采集汇聚阶段的工作从技术流程上可分为数据调研、规划设计、采集汇聚实施三个过程。

（一）数据调研

通过需求调研了解数据应用的场景和要求，通过现状调研了解现有数据基础和软硬件存储环境，为数据规划设计提供参考。

1.数据需求调研：根据建设目标及要求，与数据使用方进行数据需求的对接，明确目标数据建设范围和要求，由数据使用方提出数据需求清单，详细到信息项或维度值级别，对于需要进行二次加工的指标或信息项，明确业务规则和口径，同时，应明确目标数据所需遵循的标准及质量要求。

2.数据现状调研：根据数据使用方的数据需求，与外部数据提供方进行数据资源现状的对接，获取提供方的数据资源目录清单、数据字典说明（包括中文名称、字段类型、长度、单位、取值约束及代码表等）、样例数据，明确数据接入的方式、数据来源部门、数据安全等级等相关信息，对用户明确的数据需求了解其来源，有计划地进行沟通并协调数据（样例数据）。

（二）数据规划设计

建立正确的数据流和数据结构才能保证最好的结果，应根据应用场景和数据时效性需求，结合数据获取方式，选择合适的数据建模方式（如：实体建模、维度建模等）设计满足数据需求的数据结构、数据处理及更新流程，同时定义数据约束（如：是否允许空值、是否允许重复、数据之间的逻辑关系、编码规则、值域范围等。）

（三）采集汇聚实施

通过录入、迁移、同步、扫描、爬取等方式获取所需数据的过程。同时可能会根据目标数据物理存储要求将多源异构数据集成转换为统一的数

据格式。

4.1.2 数据采集方式

数据采集是指原始数据生成的过程，主要包括人工填报和系统采集两种方式。

（一）人工填报

人工填报是指在业务系统（PC端、移动端）表单中或文件系统（Excel、Word等）中录入数据的过程，是主要的数据采集方式。

（二）系统采集

系统采集是指由机器设备承载的系统自动识别、记录生成数据的过程。一种是利用扫描枪、扫描仪、摄像机、航拍仪以及其他各类物联网传感器设备进行的数据采集。另一种是在无相关方配合、需主动从网页或业务系统展示端获取数据情况下，采用深网数据爬取技术，通过对业务应用层面的语义解析实现数据的获取。

4.1.3 数据汇聚方式

数据汇聚是指将各部门已采集的原始数据向统一数据中心平台集中的过程，政务数据汇聚通常分为：物理汇聚和逻辑汇聚两种方式。

（一）物理汇聚

政务数据一般都是由业务部门依职权采集产生，数据分散存储在各种异构数据库和存储环境中。物理汇聚方式是指将原业务系统或存储环境中的数据，通过物理迁移的方式集中搬迁至中心数据存储环境下，经统一整合发布后，提供给数据需求部门进行应用。

（二）逻辑汇聚

逻辑汇聚通常也称为分布式汇聚。在数据共享应用前，数据依然是分散存储在各个部门系统或平台上，使用部门在使用数据的时候，再通过调用数据服务接口的方式获取和使用数据。

政务数据汇聚通常是两种方式结合，基于完全共享和有条件共享数据进行的跨领域批量数据关联融合计算的场景，如基础数据库和主题数据库建设，宜采用物理汇聚方式；基于不共享数据进行的身份核验、校验比对等场景，采用逻辑汇聚方式。

目前政务数据汇聚无论采用哪种方式，基本都是通过基于目录的共享交换体系来实现常态化数据更新。技术实现方法上根据具体场景分为数据推送和数据抽取两种方案。

4.1.4 数据采集汇聚治理要点

大部分数据问题源于数据产生的环节，数据采集汇聚阶段需加强数据落标管控和数据更新机制的建设，做好源头管控。

（一）数据落标管控

数据落标管控主要应用于数据采集过程。在具体执行数据采集工作之前建立标准化的数据模型，确定采集范围，将用于数据采集的数据模型与相关国家和行业标准进行标准比对，明确每类数据的定义、规则、关系，确保相关方对数据采集的内容、定义达成一致理解；明确数据采集的责任主体和数据采集的标准，最大限度的避免数据重复采集。

同时，应将数据采集的标准、规则嵌入数据采集系统，第一时间发现异常数据并及时修正。

（二）数据更新机制

数据更新机制主要应用于数据汇聚过程。数据汇聚过程中需要与源端系统确定更新周期（年、季、月、日、实时等）、更新时间（具体的时间点）、更新模式（全量、增量）、更新标识（基于哪个字段值或信息触发数据获取的动作）、重传机制（当已汇聚数据发生变更时如何处理）。

4.2 数据共享交换

数据共享交换是保障政务数据跨部门流通（汇聚和使用）的一种机制，也是实现数据供需对接的一种方式。

4.2.1 共享交换流程

（一）明确数据责任和需求

通过信息资源目录体系的建设，依据部门权责和跨部门数据需求目录，摸清数据资产情况，落实各部门数据需求来源，形成数据责任清单。

（二）目录注册发布及资源挂载

依托共享交换平台，各部门进行资源编目注册，经审核后统一发布，并完成目录和实际数据资源的挂载，实现目录和数据的映射。

（三）数据使用申请

目录发布后，数据使用部门通过共享交换平台查看所需数据资源目录，根据目录的共享属性和共享方式按照数据使用申请流程获取相应的数据。

4.2.2 共享交换治理要点

政务信息资源目录是数据共享交换的核心抓手，需遵照《政务信息资源共享管理暂行办法》（国发〔2016〕51号）和《政务信息资源目录编制指南（试行）》，建立高质量的政务信息资源目录。

各部门编制信息资源目录的过程中，应严格遵照有关规定，对数据资源进行分类、分级，确保目录元数据的准确性、完整性和规范性，并建立起完整的目录上下文信息（与职责和系统的关系），在目录发布后确保目录更新的及时性。技术实现上可考虑结合区块链技术，对资源目录（职责目录）进行分布式管理。

4.2.3 共享交换方式

政务数据共享交换主要有以下几种方式：库表交换、文件交换、服务接口、日志同步、消息队列。

（一）库表交换

部门或系统之间通过中间数据库（前置库）进行结构化数据的抽取、转换和加载，如oracle、mysql、mpp、hive等。

（二）文件交换

部门或系统之间通过共享文件夹（前置库）进行非结构化数据的共享交换，如txt、word、pdf等文件。文件交换过程中一般需要采用断点续传技术来提高数据交换速度。

（三）服务接口

部门或系统之间通过API接口服务的方式进行服务接口的对接，如webservice、微服务等服务接口。

（四）日志同步

业务数据库无法通过前置库采集，且需要实时采集的情况下，可以通过解析数据库日志的方式，实现数据库的增量数据捕获。该方式不需要在数据库中创建任何的触发器、临时表等对象。

（五）消息队列

数据应用端对数据源有较高实时性要求的情况下，可以采用消息队列传输通道，由源端系统将采集的数据先存入高吞吐量的消息传输通道中，如Kafka消息队列，可供应用端的业务实时消费。

4.3 公共数据开放

公共数据资源的开放利用，是深化“放管服”改革的新要求，也推进创新发展、高质量发展的必要途径。政务数据开放是指各级人民政府遵守《中华人民共和国保守国家秘密法》《中华人民共和国政府信息公开条例》，以及当地政务信息相关管理办法等的规定，根据经济社会发展需要，遵循合法有序的原则，经由当地政务信息开放平台统一为公民、法人和其他组织提供政务数据资源开放的服务。

在“开放是常态，不开放是例外”的普遍要求下，政务数据开放分为无条件开放、有条件开放和不予开放三类。可以提供给所有公民、法人和其他组织使用的公共数据属于无条件开放类；可以部分提供或者需要按照特定条件提供给公民、法人和其他组织的公共数据属于有条件开放类；涉及国家秘密、商业秘密、个人隐私和国家安全，以及其他不宜提供给公民、法人和其他组织的公共数据属于不予开放类。涉及国家、单位、团体敏感信息和个人隐私信息的公共数据经过脱敏、脱密等技术处理后符合开放条件的，应当开放。

政务数据开放数据全生命周期包括数据生成、数据传输、数据清洗加工、数据存储、数据服务和数据销毁共六个环节。

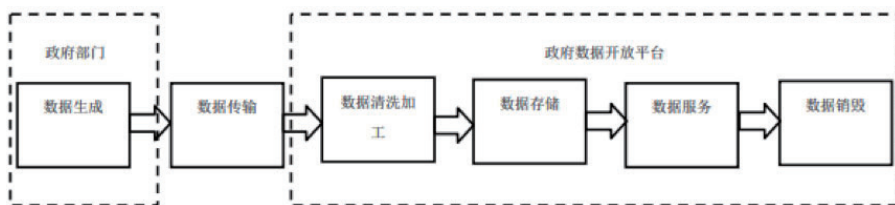


图4-2 政务数据开放数据全生命周期图

为充分推动当地政务数据开放工作，上海、浙江、天津等省级地方和威海、连云港、福州、哈尔滨、青岛等副省级和地级地方出台了专门针对政务数据开放的地方政府规章和规范性文件；上海、广东、浙江等省还专门制定了针对政务数据开放的标准规范。自2012年上海推出我国大陆地区第一个开放数据平台以来，截至2020年10月，我国已有142个省级、副省级和地级政府上线了数据开放平台，各地方政府开放的有效数据集总数达到98558个。我国政府开放数据的社会化利用生态体系愈发成熟。

4.4 数据开发利用

公共数据开发利用是深入实施大数据战略，推动落实创新驱动发展战略的重要举措之一。国家《促进大数据发展行动纲要》发布以来，各部委及国内多地都积极开展了公共数据开发利用的探索。比如在行业探索方面，环境保护部编制了《生态环境大数据建设总体方案》，以改善环境质量为核心，“通过生态环境大数据发展和应用，推进环境管理转型，提升生态环境治理能力，为实现生态环境质量总体改善目标提供有力支撑”；交通运输部印发《推进综合交通运输大数据发展行动纲要（2020-2025年）》，“以数据资源赋能交通发展为切入点”，聚焦基础支撑、共享开放、创新应用、安全保障、管理改革等重点环节，实施综合交通运输大数据发展“五大行动”，推动大数据与综合交通运输深度融合。

在地方探索中，各地从不同的侧重点，从不同的思路，积极探索公共数据开发利用新路径。比如，北京市在《北京市公共数据管理办法》中明确了市经济信息化部门负责组织、协调、指导和监督全市公共数据管理工作，提出了开放竞赛、补助奖励、应用培训等形式，以推动社会主体对开放数据的创新应用和价值挖掘。重庆市通过《重庆市政务数据资源管理暂行办法》，要求各政务部门依托共享系统和开放系统，开展一体化在线政务服务，“提高政府决策能力、管理能力和服务能力，推进治理体系和治理能力现代化”，并“鼓励公民、法人和其他组织利用政务数据资源创新产品、技术和服务，发挥政务数据资源的经济价值和社会效益，推动全社会创新创业和经济发展方式转变”。成都市在2018年通过《成都市公共数据管理应用规定》，对公共数据采集、共享、开放等方面做出了具体规定，有效推动全市公共数据共享和开放。2020年，根据新形势、新要求，成都市又创新性出台《成都市公共数据运营服务管理办法》，明确了开展公共数据“需求驱动”的“市场化运营”原则，对政务数据运营的内容、

责任单位部门及其职责分工、服务方式、监管和安全管理等都均做了具体规定，并提出了建立数据运营平台各环节“透明化、可记录、可审计、可追溯的全过程管理机制”，极大地提升了成都市公共数据增值利用效率。

为进一步拓展公共数据应用场景，上海市自2015年起，每年开展一次开放数据应用创新大赛（SODA），大赛通过参赛团队提供大规模、多维度的公共数据和行业数据开放样本，“以赛促用”，激发全社会对开放数据开发利用的热情，打造开放数据价值实现的示范案例。扬州市也于2021年开始，举办扬州公共数据开发应用创新大赛，进一步加强数字政府建设，提升公共服务、社会治理等数字化智能化水平，促进公共数据开发利用，优化提升营商环境。北京市和福州市通过建设公共数据创新基地，通过为特定对象，开放某些特定类别的公共数据等方式，探索公共数据开发利用新模式。成都市通过数据核验、数据补全、可信应用云、分析报告等服务方式，积极探索在公共数据权属不清环境下的公共数据应用新模式。

在多方安全计算、隐私计算、联邦学习等技术支撑下，我国公共数据开发利用在交通、金融、旅游、医疗、招商、工商、税务、综合治理等多个领域的创新性应用案例层出不穷，数据产品与服务形式包括了人工智能数据服务、API定制服务、可视算法、模型组件、行业解决方案等各个方面。随着政企数据融合不断推进，公共数据的开发利用在提升政府服务能力、提高社会生产力、建立数据驱动型增长新模式、培育创新经济、建设创新型国家，推动经济高质量发展等过程中，起着越来越重要的作用，并将助力实体经济突破发展边界，实现质的飞跃。

4.5 数据质量管理

高质量的数据才能创造业务价值，数据质量管理是贯穿数据开发利用全生命周期的一项基础核心的数据管理活动。当前政务数据资源普遍存在不准确、不完整、缺少上下文信息等质量问题，造成数据共享后难以理

解、开发和应用，无法达到辅助政府重大事项决策的作用。数据质量管理是当前各地政务资源统筹管理单位面临的难题之一。

在数据质量管理领域，可参考的标准和框架主要是ISO 8000国际标准和2018年发布的GB/T 36073-2018《数据管理能力成熟度评估模型》（DCMM）和GB/T 36344-2018《信息技术 数据质量评价指标》国家标准。ISO 8000包含了一般原则、主数据质量、交易数据质量和产品数据质量的标准。通过ISO 8000可以保证组织在满足决策需求和数据治理的基础上，在产品或服务的全生命周期内高质量的存储、共享、交换和利用数据。DCMM标准中给出了数据管理能力成熟度评估模型及相应的成熟度等级，定义了数据质量能力域中的数据质量需求、数据质量检查、数据质量分析及数据质量提升等4个能力项。GB/T 36344-2018从规范性、完整性、准确性、一致性、时效性和可访问性等6个维度规定了数据质量评价指标的框架。

参考以上国际和国家标准，实践中结合PDCA（计划、执行、检查和处理）质量持续改进和闭环方法，将数据质量管理过程分为原始数据剖析、数据标准建设、数据质量评估和问题数据清洗4个阶段的工作。其中：数据标准建设是核心，为数据质量评估和问题数据清洗提供依据，原始数据剖析为数据标准建设提供良好实践，数据质量评估发现问题数据，为问题数据清洗提供改进的方向。

4.5.1 原始数据剖析

原始数据剖析是通过一些活动的处理，熟悉数据集，发现数据特征，识别数据的质量问题，对问题数据进行分类，并尝试发现产生数据质量问题的原因的过程。包括以下工作过程：

（一）进行数据描述

使用描述性统计方法描述数据“总的”或者“表面的”特征。从数据字段、数据表、数据集三个维度分别进行描述。

1.数据字段描述：包括数据记录数、唯一记录数、唯一率、重复记录数、重复率、缺失记录数、缺失率、零值记录数、零值率、最大列长度、最小列长度、平均列长度等指标。

2.数据表描述：包括数据字段数、数据记录数、总存储空间大小、缺失数据元素数、包含缺失值的字段数、全部缺失的字段数、重复记录数、重复率等指标。

3.数据集描述：包括数据表数、数据字段数、数据记录数、总存储空间大小、数据表平均存储空间大小、缺失数据元素数、包含缺失值的字段数、全部缺失的字段数等指标。

（二）进行数据探查

通过查询、可视化和报告的方式对数据特征进行分析。包括：关键属性的分布，一对或几个属性之间的关系，分析结果的汇总，数据集子集的进一步探查。分为数据字段元数据信息探查、数据值分布探查、数据字段表示形式探查、数据表探查4个方面。

1.数据字段的元数据信息探查包括：分析数据字段的命名及命名规则；分析数据字段的属性信息，包括数据类型、数据长度、精度等；分析数据字段值的变量类型；分析数据字段的数据类型；分析数据字段的数据元表示；分析数据字段的真实值是否符合定义的数据类型。

2.数据值分布探查包括：分析数据表中的主键是否唯一；分析数据字段的值是否满足指定阈值，分析数据字段值的统计量；分析各个维度值在总体数据中的分布情况；分析数据字段中是否包含缺失数据；分析数据字段中是否包含约定的异常数据；分析数据字段中是否包含特定的字符。

3.数据字段表示形式探查包括：从业务角度规定数据项值的格式要求，包括数据值的类型，允许的最大和最小字符长度，数据项值的表示格式等；分析编号类的数据字段是否符合特定的编码规范；分析枚举类的数据字段的唯一值列表是否符合特定的代码规范；分析名称类的数据字段是

否符合特定的规则；分析日期时间类的数据字段是否符合特定的规则；分析数字类数据字段是否符合特定的规则；分析文本类数据字段是否符合特定的规则。

4.数据表探查包括：分析表中的外键是否都能在关联表中找到；分析数据字段之间，或数据表之间是否满足指定的业务规则；分析业务时间和更新时间，分析数据表中数据接入和更新的规则；分析数据表是否满足业务需求。

（三）剖析结果分析

对数据描述和数据探查的结果进行具体问题具体分析。以便更准确地理解和把握业务含义。站在数据表和数据字段两个维度进行分析，同时对具体的数据问题进行进一步数据特征分析，并尝试发现数据问题背后的原因。

（四）评估业务影响

按业务重要性程度和数据问题的严重性程度对原始数据进行业务影响分级。更进一步地确定导致数据问题的根本原因并制定一个行动方案。

4.5.2 数据标准建设

数据标准建设是通过一系列方法，形成可落地的数据标准的过程。包括以下工作过程：

（一）确定标准的范围

在充分理解标准建设的需求和数据剖析结果的基础上，明确标准约束的对象、内容范围和标准的适用范围。

1.按照数据组织的单元，数据标准主要包括数据元标准和数据集标准，数据元包括基础类数据元和指标类数据元。

2.按照对数据具体特性的描述，数据标准还包括业务术语标准、代码标准、编码标准。

（二）编制数据标准

根据实施数据质量管理的对象范围，形成数据标准草案，研讨、改进。

（三）发布数据标准

对已提交的数据标准进行技术审查和交叉职能审查，根据审查意见进行修改，并批准、发布数据标准。

4.5.3 问题数据清洗

问题数据清洗阶段是对问题数据进行清洗，将其转化为满足数据质量要求或应用要求的数据的过程。包括以下工作过程：

（一）数据清洗规则设计

包括具体的清洗策略、规则。数据清洗规则从执行策略上分为：补全修正、数据过滤、数据容错三类。补全修正是对问题数据进行修正或替换（转换），数据过滤是将问题数据与标准数据进行分离，数据容错是因业务和管理原因无法对问题数据进行修正又无法分离的的一种处理机制。具体可采用定性分析的方法，将数据的重要性程度和数据问题的严重程度结合起来，形成策略矩阵，依据数据在策略矩阵中所处的区域选择具体的处理方案。具体清洗规则需结合数据应用的场景设计，确保数据不失真。

（二）数据清洗规则审核

由业务和技术负责人分别对规则进行审核，通过之后规则方可生效。

（三）数据清洗任务执行

由技术人员通过相关工具执行清洗规则。

4.5.4 数据质量评估

数据质量评估阶段是通过数据校验方法，发现数据问题及对数据质量进行分析评价的过程。包括以下工作过程：

（一）数据质量评价指标体系设计

包括评价对象、评价维度、指标计算方法、权重设计、评价结果的表示等。通过可度量的数据质量评价指标来明确数据质量管理的目标和

重点。

- 1.评价对象：包括部门、数据集或数据表。
- 2.评价维度：参考GB / T 36344-2018《信息技术 数据质量评价指标》，可划分为规范性、完整性、准确性、一致性、时效性、可访问性。
- 3.指标计算方法：一般按占比或按是否满足条件设置固定的分值。
- 4.权重设计：根据业务影响分析确定的数据重要性级别，设置相应的权重。
- 5.评价结果的表示:定量或定性的评价。

（二）数据质量检查规则设计

参考业务需求、数据质量评价指标体系和数据标准，设计数据质量检查规则。规则的设计以满足数据应用需求为准则，并不是设计的越全面越好。数据质量检查规则分为技术规则和业务规则。

1.常见的技术规则包括：空值、重复、格式（如编号规则、异常字符等）、值域（数值值域、字符值域、日期值域）、参照完整性（待检数据是否全部包含在参照数据列中）、波动性（不同期指标数据的波动）、逻辑一致性（不同列之间的逻辑关系、明细记录汇总和汇总表一致性）等。实践过程中一般通过规则引擎实现技术规则的封装。

2.业务规则是技术规则的具体实例化结果和组合。例如：针对某一数据表中的公民身份号码设置的公民身份号码检查规则可由：空值、重复、格式检查3种技术规则组合而成。

（三）数据质量检查规则审核

由业务和技术负责人分别对规则进行审核，通过之后规则方可生效。

（四）数据质量检查任务执行

由技术人员通过相关工具执行数据质量稽核任务。

（五）数据质量检查结果分析

对稽核任务执行的结果进行分析，对问题数据进行分类及原因追溯。



（六）输出数据质量报告

对数据质量评估结果进行汇总，形成数据质量报告，并对问题数据进行反馈。将检查结果以工单形式反馈给责任部门，便于对问题数据处理过程及结果的跟踪。同时，通过数据质量报告，数据资源统筹管理部门可以直观的了解数据质量提升的过程和效率，及时发现数据质量管理的堵点，从而更好地推动数据质量提升工作。

五、政务数据开发利用参考架构

5.1 政务数据总体参考框架

政务数据总体参考框架如下图所示：

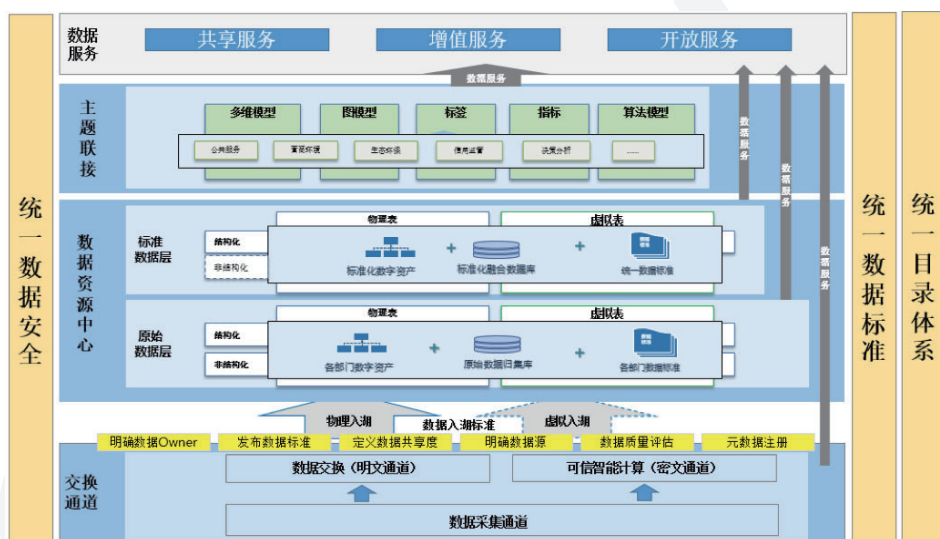


图5-1 政务数据总体参考框架

政务数据总体参考框架包括共享交换通道、数据资源中心、主题联接和数据服务。

共享交换是实现政务数据资源整合共享和服务的核心，也是基础性平台。基于共享交换实现公共数据资源归集整合，开展数据归集共享工作。

数据资源中心分为原始数据层和标准数据层两个部分，是多源数据汇集的数据集合，既有原始层贴源数据集合，也有数据管理部门非业务规则治理后的统一标准数据及其关联关系，为后续数据在业务层面的加工、分析做好数据准备。

数据主题联接是对数据资源中心的数据按业务流/事件、对象/主体进

行联接和规则计算等处理，形成面向数据消费的主题数据，具有多角度、多层次、多粒度等特征，支撑业务分析、决策与执行。基于不同的数据消费诉求，主要有多维模型、图模型、指标、标签、算法模型等五种数据联接方式。

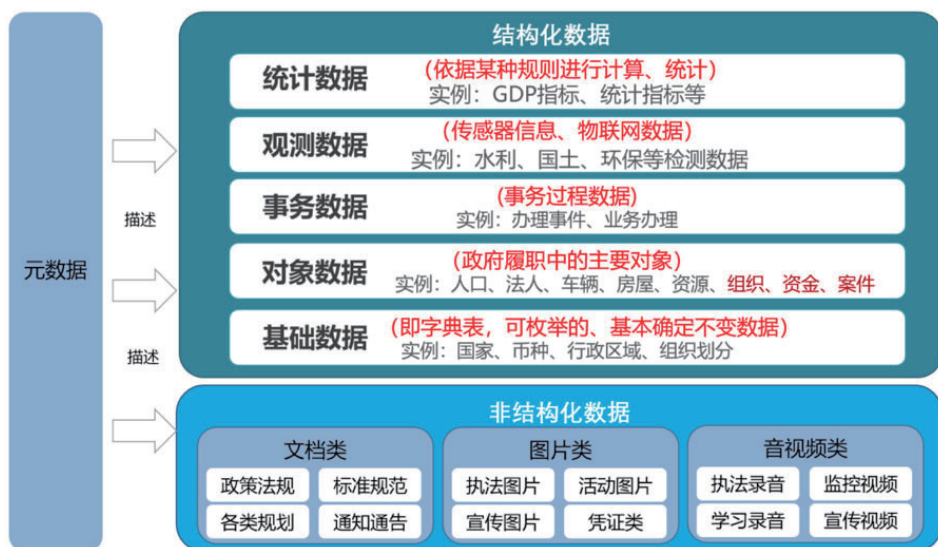
数据服务是以数据集、接口、自助在线分析等多种共享的方式向其他单位提供数据的服务过程，是推进数据跨部门、跨层级、跨地区汇聚融合和深度利用的重要手段。

软硬件大数据基础设施应支持国产主流技术和架构，不限于芯片、整机、操作系统、数据库、计算框架等形态；同时应不因关键软件、硬件、器件的扩容、维护、替换影响整体业务运行性能或导致业务停用，支持多种计算架构。

5.2 政务数据分类管理

政府部门在履职过程中会面临多种、复杂类型数据的管理问题。基于不同的目的，可以从多个角度上来对数据进行分类管理。对数据进行分类的目的，是为了针对不同特性的数据采取不同的管理策略，以期达成最大投入产出比。

业界常见的数据分类包含，结构化数据和非结构化数据、内部数据和外部数据、原始数据和衍生数据、明细数据和汇总数据等。在业界数据分类基础上，结合政府的数据管理分层需求，形成政务数据分类管理框架。



元数据，定义数据的数据，是有关机构所使用的物理数据、技术和业务流程、数据规则和约束以及数据的物理和逻辑结构的信息。其特征是描述性标签，描述了数据、相关概念、以及它们之间的联系。

基础数据，用结构化语言描述属性，用于分类或目录整编的数据。其特征是通常有一个有限的允许/可选值范围，可作为参考数据。属于静态数据，比较稳定，可以用作业务/IT的开关，职责/权限的划分或统计报表的维度。

对象数据，可以在政府内跨流程跨系统重复使用。具有唯一、准确、权威的数据源。其特征是业务事件的参与方，取值不受限于预先定义的数据范围。在业务发生之前就客观存在，比较稳定。

不同分类的数据，其重要程度及治理方法有所不同，在以上的数据分类中，基础数据、元数据是贯穿流程、跨部门的重要数据，需要由政府的统一数据管理部门和数据产生的部门协同统一管理。

几种数据的关联关系如下图所示：

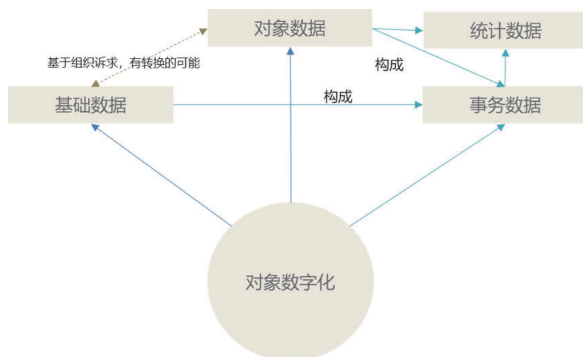


图5-3 数据关联关系

5.3 政务数据资产目录管理

数据资产目录是指对数据中有价值、可用于分析和应用的数据进行提炼形成的目录体系。编制数据资产目录可以给出业务场景和数据资源的关联关系，降低理解系统数据的门槛。

2015年3月，中共中央办公厅、国务院办公厅印发了《关于推行地方各级政府工作部门权力清单制度的指导意见》，要求将地方各级政府工作部门行使的各项行政职权及其依据、行使主体、运行流程、对应的责任等，以清单形式明确列示出来，向社会公布，接受社会监督。

2015年12月，国务院办公厅关于印发《国务院部门权力和责任清单编制试点方案的通知》（国办发〔2015〕92号），要求以清单形式列明试点部门的行政权责及其依据、行使主体、运行流程等，推进行政权责依法公开，强化行政权力监督和制约，防止出现权力真空和监管缺失，加快形成边界清晰、分工合理、权责一致、运转高效、依法保障的政府职能体系。

5.3.1 面向权责清单的数据资产梳理模型

政府权责清单很全面的覆盖了政府部门行使权力及履行义务过程中所涉及的所有内容，全面梳理了部门现有权责事项，将权力和责任事项进行结合，并参照行政许可、行政处罚、行政强制、行政征收、行政给付、

行政检查、行政确认、行政奖励、行政裁决和其他类别的分类方式，对行政权责事项进行全面梳理并逐项列明设定依据，是政府部门梳理“数据家底”的基础。理清政务数据资产应该建立面向权责清单的数据资产梳理模型。该梳理方法建立在中央编办、国务院法制办等相关审核单位已审核的权责清单基础上，充分利用已经梳理完成的工作成果，不仅大大缩减了重新梳理目录的工作量，同时增加了目录的权威性、全面性。

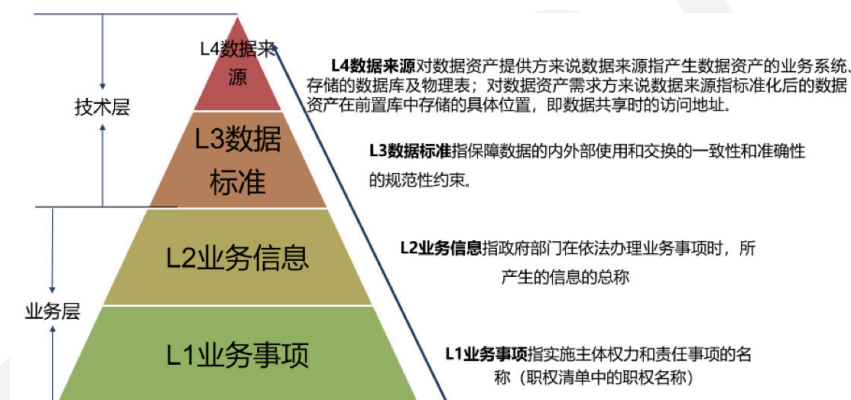


图5-4 面向权责清单的数据资产梳理模型

L1业务事项，指实施主体权力和责任事项的名称（职权清单中的职权名称）

L2业务信息，指政府部门在依法办理业务事项时，所产生的信息的总称。

L3数据标准，指保障数据的内外部使用和交换的一致性和准确性的规范性约束。

L4数据来源，对数据资产提供方来说数据来源指产生数据资产的业务系统、存储的数据库及物理表；对数据资产需求方来说数据来源指标准化后的数据资产在前置库中存储的具体位置，即数据共享时的访问地址。

L1-L2是业务层，便于部门从业务上理解数据资产所属的业务事项业务信息，了解数据资产的业务含义。数据资产L3-L4为技术层信息，包括

数据的表和字段，方便部门开展数据共享和使用数据。

5.3.2 政务数据资产梳理

（一）梳理业务事项

业务事项指实施主体权力和责任事项的名称，同职权清单中的职权名称。政府业务事项分为：行政许可、行政处罚、行政强制、行政征收、行政给付、行政检查、行政确认、行政奖励、行政裁决及其他类别。

在梳理政务数据资产时，政务数据资产的业务事项可直接复使用权责清单中的职权名称。如下图中，某市属局的业务事项为：市级社会团体年度检查、对非法开展社会团体活动的处罚、社会组织（社会团体、民办非企业单位、基金会）成立、变更、注销登记等。

表5-1 市属局业务事项

实施主体	职责	职权名称	类别
市属局	社会组织登记管理	市级社会团体年度检查	行政检查
		对非法开展社会团体活动的处罚	行政处罚
		对违规开展社会团体活动的处罚	行政处罚
		封存、收缴社会团体法人登记证书、印章和财务凭证	行政强制
		市级基金会年度检查	行政检查
		对非法开展基金会活动的处罚	行政处罚
		对违规开展基金会活动的处罚	行政处罚
		封存、收缴基金会登记证书、印章和财务凭证	行政强制
		市级民办非企业单位年度检查	行政检查
		对非法开展民办非企业单位活动的处罚	行政处罚
		对违规开展民办非企业单位活动的处罚	行政处罚
		封存、收缴民办非企业单位登记证书、印章和财务凭证	行政强制
		市社会组织评比表彰	行政奖励
		社会组织（社会团体、民办非企业单位、基金会）成立、变更、注销登记	行政许可

（二）确定业务信息

在梳理完业务事项后，需要对业务事项进行拆分或合并，这样才能更详细合理的梳理政府部门在依法办理业务事项时产生的所有信息。

业务信息是政府部门在依法办理业务事项时，该事项全生命周期可产生的信息。包括办理结果信息、相关的统计信息等，行政许可类的还需梳理出办理该业务服务对象的基本信息、所有行政许可类型名录信息清单。我们梳理确定的业务信息，一般即为我们的政务数据资产目录名称，然后根据业务信息，确定各数据项。

如办理“社会组织（社会团体、民办非企业单位、基金会）成立、变更、注销登记”业务事项，所产生的信息包括社会组织（社会团体、民办非企业单位、基金会）基本信息、社会组织（社会团体、民办非企业单位、基金会）成立、变更、注销办件结果信息、社会组织（社会团体、民办非企业单位、基金会）统计类信息、社会组织（社会团体、民办非企业单位、基金会）行政许可名录信息等。

如“对非法开展民办非企业单位活动的处罚”的业务信息包括对非法开展民办非企业单位活动的处罚信息、对非法开展民办非企业单位活动的处罚统计信息等。

（三）明确数据标准

数据标准是进行数据标准化的主要依据，构建一套完整的数据标准体系是开展数据标准管理工作的良好基础，有利于打通数据底层的互通性，提升数据的可用性。简述之，即数据标准是指保障数据的内外部使用和交换的一致性和准确性的规范性约束。因此，在建立业务信息的实际数据模型后，应制定本单位共享数据项的数据标准，以便于各部门数据需求对接，推动政务数据资产目录体系的统一共享应用。通过数据标准化，有利于拉通数据，有效提升业务效率和数据质量、促进数据共享。主要有以下作用：

提升业务效率：数据标准统一了业务语言，明确了业务规则，规范了业务处理过程，从而提升组织整体业务效率，满足管理决策对信息及时性的要求。

促进数据共享：数据标准统一了各类系统的数据定义，降低系统间集成的复杂度，提高系统间交换效率，并为管理分析系统提供一致的分析指标和分析维度定义。

提升数据质量：数据标准明确了数据填写及处理要求，规范数据源的格式，同时提供管控方面的保障，因此数据标准将直接提高数据质量。

政务数据资产的数据标准涉及各部门根据共享要求确定所共享的数据资产业务数据项，明确各数据项的数据项名称、编码、类型、长度、业务含义、质量规则、安全级别、域值范围等。对于有国标或者行标的，应该优先按照国标、行标执行，否则按照产生该数据的权威部门的标准执行。

如社会组织（社会团体、民办非企业单位、基金会）基本信息，根据业务信息梳理共享数据项后，每个字段应按照《政务服务中心信息公开数据规范》GB/T32617-2016中要求的数据标准进行共享。

结构化数据除应明确数据类型外，还应标明数据长度。包括：字符型C、数值型N、货币型Y、日期型D、日期时间型T、逻辑型L、备注型M、通用型G、双精度型B、整型I、浮点型F等。

（四）确定数据来源

在梳理完业务事项可产生的业务信息后，接下来的梳理工作需要找到产生这些业务信息的系统、这些业务信息的存储位置等，这时候，我们需要确定业务信息的数据来源。

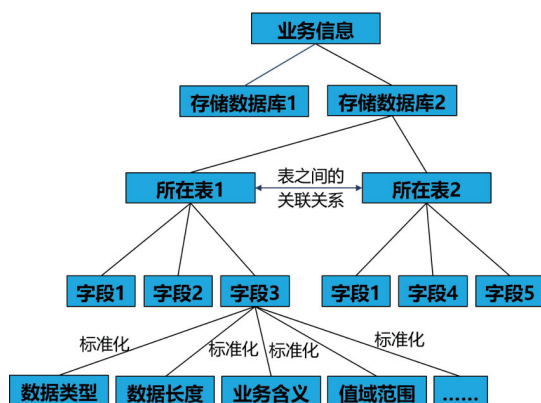


图5-5 业务数据来源

数据提供部门在确定数据来源时，应根据已确定的数据标准确定各数据项产生的业务系统、存储的物理表，对于在不同数据库、不同表存储的字段应进行合并、拆分、汇总整合，并确定共享数据项。对于有分类、标签、编码的，需整理出相应的编码规则和数据字典，并将整理出来的编码、数据字典按照相应数据标准进行标准化。如在梳理社会组织（社会团体、民办非企业单位、基金会）成立、变更、注销办件过程信息的数据来源时，应梳理登记注册审核状态信息，并按照相应标准对状态及对应状态码进行标准化，以便于数据共享时数据资产的统一提供、统一服务。

最后，库表数据资产应在本部门前置机中建库建表，并将标准化后的数据项汇总整合存储在已创建的数据表中；文件、文件夹类数据资产应在部门前置库中创建文件夹，并将标准化后的文件存储在文件夹中，以便向数据资产需求方提供统一的共享服务。当各信息项及信息项的数据类型、值域范围发生变化，则相应的共享目录也应进行相应的变更，以保证实际数据与数据资产目录的一致性，实现目录的动态联动更新。

（五）资产目录发布

1.数据资产目录编制

根据数据资产管理模型对政务数据资产元数据进行整理、分析、汇

总，确定数据资产分类、安全级别、提供方及相关属性，按照政务数据资产元数据要求，生成政务数据资产目录。

以社会组织（社会团体、民办非企业单位、基金会）基本信息为例，假如该部门主要共享社会组织名称、组织机构代码、营业执照、法定代表人、通信地址、联系人、联系电话这些信息项，则数据资产目录编制如下：

目录名称：社会组织（社会团体、民办非企业单位、基金会）基本信息。

目录分类：部门数据资产目录-省级地方-对象数据-其他组织。

表5-2 目录信息项

信息项名称	数据类型及格式	数据标准
单位名称	an..100	GB/T19488.2
组织机构代码	an9	GB 11714
营业执照	an..15	GB/T19488.2
法定代表人	a..30	GB/T19488.2
联系人姓名	a..30	
联系电话	an..18	
通信地址	an..400	

2.数据资产目录报送

为了保证数据资产的全面性、准确性和规范性，需对编制的的数据资产进行审核，各部门在对编制的政务数据资产进行复核、审查后，及时报送本级政务数据资产共享主管部门。

5.4 政务数据中心框架

政务数据资源中心是逻辑上各种原始数据的集合，包括原始层和标准层两个部分。

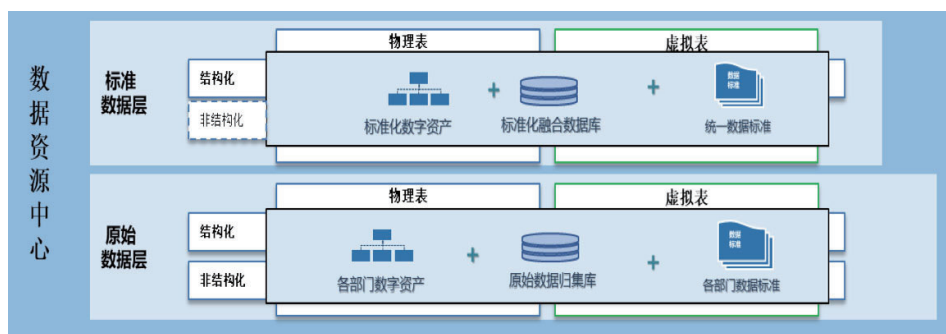


图5-6 政务数据资源中心

原始数据层是逻辑上各种原始数据的集合，除了“原始”这一特征外，还具有“海量”和“多样”（包含结构化、非结构化数据）的特征。数据资源中心保留数据的原格式，原则上不对数据进行清洗、加工，但对出现数据资产多源异构的场景需要整合处理，并进行数据资产注册。

标准数据层是在原始数据层的基础上，通过政务对象数据识别形成国家级统一数据标准，构建标准化融合数据库。这层的标准数据会进行通用规则的清洗、加工，在对象数据的识别和治理下，形成全局各类逻辑实体的关联关系，可以同时形成一个大宽表，并进行数据资产的注册，形成标准化数据资产。

5.4.1 基于入湖标准治理的原始层构建

（一）政务数据入湖标准

数据入湖是数据协同治理消费的基础，需要严格满足入湖的6项标准，包括明确数据所有者、发布数据标准、定义数据共享属性、明确数据源、数据质量评估、元数据注册。通过这6项标准保证入湖的数据都有明确的业务责任人，各项数据都可理解，同时都能在相应的信息安全保证下进行消费。

1.明确数据所有者

数据所有者由数据产生对应的流程所有者担任，是所辖数据端到端管

理的责任人，负责对入湖的数据定义数据标准和共享属性，承接数据消费中的数据质量问题，并制定数据管理工作路标，持续提升数据质量。

2.发布数据标准

入湖数据要有相应的业务数据标准。业务数据标准描述政府层面需共同遵守的“属性层”数据含义和业务规则，是政府层面对某个数据的共同理解，这些理解一旦明确并发布，就需要作为标准在企业内被共同遵守。

3.认证数据源

通过认证数据源，能够确保数据从正确的数据源头入湖。认证数据源需遵循政务数据源管理的要求，数据源一般是指业务上首次正式发布某项数据的应用系统，并经过数据管理部门。认证过的数据源作为唯一数据源头被数据资源中心调用。当承载数据源的应用系统出现合并、分拆、下线情况时，需及时对数据源进行失效，并启动新数据源认证。

4.明确数据共享属性

定义数据共享属性是数据入湖的必要条件，为了确保数据资源中心中的数据能充分的共享，入湖的数据必须要定义数据共享属性。责任主体是数据所有者，数据管家有责任审视入湖数据共享属性的完整性。

5.数据质量评估

数据质量是数据消费结果的保证，数据入湖不需要对数据进行清洗，提升数据质量，但需要对数据质量进行评估，让数据消费的人员了解数据的质量情况，并了解消费该数据的质量风险。同时，数据所有者和数据管家可以根据数据质量评估的情况，推动源头数据质量的提升，满足数据质量的消费要求。

6.元数据注册

元数据注册是指将入湖数据的业务元数据和技术元数据进行关联，包括逻辑实体与物理表的对应关系，及业务属性和表字段的对应关系。

通过联接业务元数据和技术元数据的关系，能够支撑数据消费人员通

过业务语义快速的搜索到数据资源中心中的数据，降低数据资源中心中数据消费的门槛，能让更多的业务分析人员能理解和消费数据。

（二）政务数据入湖方式

数据入湖以逻辑数据实体为粒度进行入湖，逻辑数据实体在首次入湖时应该考虑信息的完整性，原则上一个逻辑数据实体的所有属性应该一次进湖，避免一个逻辑实体多次入湖，增加入湖工作量。

数据入湖的方式主要有物理入湖和虚拟入湖两种，根据数据消费的场景和需求，一个逻辑实体可以有不同的入湖方式。两种入湖方式相互协同，共同满足数据联接和用户数据消费需求，数据管家有责任根据消费场景的不同，提供相应方式的入湖数据。

物理入湖指将原始数据复制到数据资源中心中，包括批量处理、数据复制同步、消息和流集成等方式。

虚拟入湖指原始数据不在数据资源中心中进行物理存储，而是通过建立对应虚拟表的集成方式实现入湖，实时性强，一般面向小数据量应用，大批量的数据操作可能影响源系统。

数据入湖可以由数据资源中心主动从数据源通过PULL（拉）的方式入湖，也可以由数据源主动PUSH（推）的方式入湖。数据复制同步、数据虚拟化、以及传统数据仓库技术（ETL）批量集成都属于数据资源中心主动拉的方式。流集成、消息集成属于主动推送的方式。在特定批量集成场景下，数据会以CSV、XML等格式，通过文件传输协议（FTP）推送给数据资源中心。

表5-3 数据入湖方式

入湖方式	数据源	数据资源中心
PULL（拉）	被动：当被请求时提供数据	主动：决定何时获取数据
PUSH（推）	主动：按自己节奏提供数据	被动：响应接收数据

1.结构化数据入湖

结构化数据是指由二维表结构来逻辑表达和实现的数据，严格地遵循数据格式与长度规范，主要通过关系型数据库进行存储和管理。

触发结构化数据入湖的场景有两种：一是政务数据管理组织基于业务主动规划、统筹数据入湖；二是响应数据消费方需求实施数据入湖。结构化数据入湖过程包括：数据入湖需求分析及管理、检查数据入湖条件和评估入湖标准、实施数据入湖、注册元数据。

2.非结构化数据入湖

非结构化数据是指包括无格式文本、各类格式文档、图像、音频、视频等多样异构的格式文件，较之结构化数据更难对其标准化和理解。相较于结构化数据，非结构化数据管理不仅包括文件本身，也包括对文件的描述属性，也就是非结构化的元数据信息。

这些元数据信息包括文件对象的标题、格式等基本特征定义，还包括对数据内容的客观理解信息，如：标签、相似性检索、相似性连接等。这些元数据信息会便于用户对非结构化数据的搜索和消费。

非结构化数据入湖包括对非结构数据的基本特征元数据、文件解析内容、文件关系和原始文件入湖，其中基本特征元数据是非结构化数据入湖的必选内容，后面三项内容根据分析诉求可以选择入湖和延后入湖。

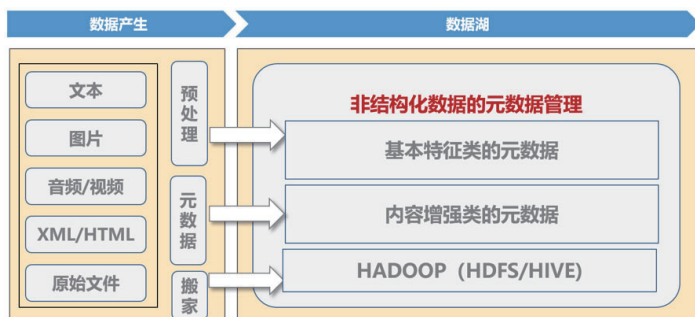


图5-7 非结构化数据入湖

（三）政务数据资源中心原始层的建设

政务数据原始层也是贴源数据层，是逻辑上各种原始数据的集合。各个部门会在数据资源中心形成各自的部门的数据资产和原始数据归集库。



图5-8 政务数据原始层

数据资源中心原始层保留数据的原格式，原则上不对数据进行清洗、加工。满足数据入湖标准的数据从中心前置节点或其它数据源完整的拷贝，保障大数据平台的数据和外部数据源数据的一致性。原始层数据在模型上采取贴源设计，偏向于维持源系统数据组织的原貌形态，基本上不会考虑信息要素的删减，保持与源系统数据同样的数据结构。数据按照抽取的频率和增全量进行存储，每个频率单独保存一个表文件，能满足后续其他层次的需求即可。归集贴源层存储上游源系统的明细数据，但不整合处理，尽量保持业务数据原貌。核心分为两个步骤：

第一步，将外部数据加载到数据资源中心的临时缓冲区，转换数据的形态，将转换后的数据纳入数据资源中心管理。

第二步，原始层保留源系统的全量历史数据，供后续数据深化加工使用，避免频繁访问各厅局委办的原始数据。

5.4.2 基于政务对象数据治理的标准层构建

（一）政务对象数据

政务对象数据是指有共享性的基础数据，可以在政府各部门之间、政府与社会个人、单位之间内被重复使用的。比如，自然人的对象数据是指每个自然人的基本特性描述，如证件号码、姓名等，在每个人的生命周期中基本不会发生变化。

人口基础信息库对象数据包括如下内容：

表5-4 人口基础库对象数据

表中文名	字段序号	字段中文名	数据类型
人口基本信息表	1	公民身份号码	字符型
人口基本信息表	2	性别代码	字符型
人口基本信息表	3	姓名	字符型

政务对象数据治理以对象数据生命周期管理为核心，基于对象数据生命周期各个环节建立对象数据标准，以数据标准为依据搭建数据质量管理体系。

（二）政务数据标准层的构建

政务数据标准层的构建是基于政务对象数据实现数据的标准化融合。建立一个统一、规范、标准化、涵盖各委办局信息系统数据，促进跨部门、跨层级数据互认与共享。



图5-9 政务数据标准层

对于结构化数据的标准化融合，遵循“一数一源”的原则对来自于多个部门的重复数据表和数据项在表结构层面进行归一和融合，同时，对于非结构化数据则进行文件解析和特征信息提取，并和上述结构化数据进行关联融合。

标准化数据库的设计和建设符合国家级统一数据标准和技术规范，符合国家相关管理规范要求。以人口基础信息资源库为例，依托于政务数据服务共享交换，以公安人口数据为基础，逐步融合卫计委、民政部、人社厅、教育厅等部门的相关信息资源，重点建设和整合公安户籍信息、人

口普查信息、公务员信息、社会保障信息和人事关系信息等信息资源，扩展健康、收入、婚姻、社保、救助、贫困、残疾、流动、死亡等信息，逐渐丰富人口基础信息资源条目，构建统一的、可共享的人口基础共享信息库。

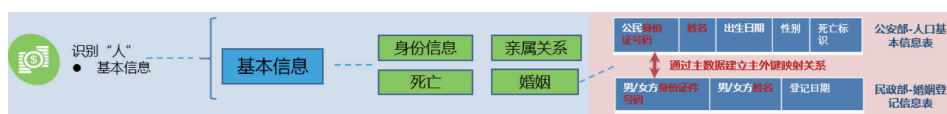


图5-10 人口基础信息资源库

通过对象数据识别来关联多个数据源，构建人口库，实现人口信息的交换、处理、存储、更新、服务为一体的运行、管理、共享体系。为政府相关部门的业务应用、跨部门业务应用和宏观决策等提供基础数据支持与服务。

1.数据来源。以公安核实的实有人口信息为基础，覆盖本级政府行政区人口的基本信息、扩充信息，并且根据需要可以及时追加基础信息。接入公安、卫计、教育、民政、人社、住建等单位扩展信息，建立以公民身份证号码为唯一代码的实有人口库，实现以人口档案管理为基本，延伸到人口健康、社保、教育等关联信息。人口基础信息资源库按照信息来源和信息归属划分，包括：劳动就业信息、社会保险信息、教育信息、卫生健康信息、民政信息、住房公积金信息、住房信息、税务信息、单位信息、申报信息等。根据不同信息资源类别，提供数据库表结构设计。

2.数据管理。数据管理根据建设人口基础数据库主题库所需的支撑模块和接口，包括用户管理、权限管理、用户组织机构管理、编目管理、消息中心、日志管理等功能。能够维护人口库的基本内容构成可细化到字段级，并可维护每个字段从哪个信息资源中的某个指标获取，如有多个来源，需要提前确定数据协商计算规则。

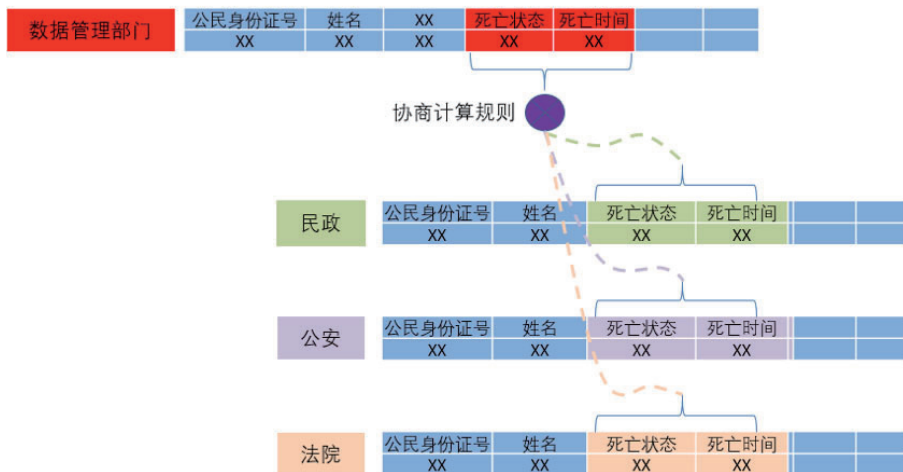


图5-11 协商计算规则

如上图，死亡数据在业务上有多个来源部门，如民政、公安、法院等，但数据管理部门在归集到自然人死亡相关数据时，需要组织多个数据源部门对数据最终状态的确定，制定协商计算规则，以计算规则确认数据最终状态。

3.数据集成。主要包括以下两方面内容：

数据清洗：照标准规范，针对数据本身属性错误的数据进行过滤，通过过滤规则，即可将这类错误数据过滤。通过数据清洗功能，为数据比对提供干净、准确的数据环境。

数据比对：将同一属性但来自不同数据源的数据，通过制定的比对规则进行比对。数据比对一致的，认定为准确；数据比对不一致的，认定为错误，反馈到相关部门进行核查。

4.数据更新机制。以人口库为例，人口库中的核心基础信息反映人口的基本属性，更新频率较低；扩展信息反映人口在不同阶段的状态属性，更新频率相对较大。基础信息更新要求：由牵头部门公安人工核对后增量更新。其他部门无权进行更新、如果发现基础信息有误的，提交给牵头

部门进行人工处理后增量更新。扩展信息更新要求需根据扩展信息设定的更新频率按谁主管谁负责增量更新。

5.5 政务数据主题联接框架简化

在当前政府数字化转型的背景下，政府部门对数据的消费需求已经不在局限于传统的报表分析，更多要支持政府用户的自助分析、实时分析，通过数据的关联关系支持业务的关联影响以及对目标对象做特征识别，进行特定业务范围圈定、差异化管理与决策等。目前，数据资源中心汇聚或关联大量的数据，虽然已能够满足用户不用到各个源系统取数，统一从数据资源中心中取数据的要求，数据资源中心中的数据结构都是同数据源、严格遵从三范式的，即使每个数据都有详细的定义和解释。然而，面对一个分析诉求，需要的数据全部分散在不同物理表中，数据零散，用户很难知道数据之间的关联关系。

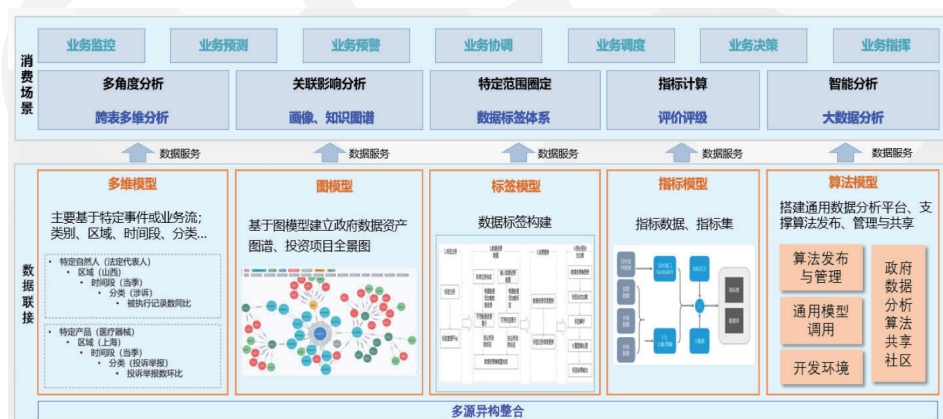


图 5-12 五类数据主题联接

1. 多维模型是面向业务多视角、多维度的分析，通过明确的业务关系，建立基于事实表、维表以及相互间联接关系，实现多维数据查询和分析。比如，在数据共享交换的场景，给定业务目标，数据管理部门可以构

建多维数据模型关联多个数据源，特定维度的数据源也可以有多个，这样就可以通过多维模型的牵引来实现业务目标的驱动多方数据源的共享与交换。在具体的业务场景也可以基于此抽象出各个分析维度，主要在于业务需求驱动下的事实表和维度的设计。

2.图模型是面向数据间的关联影响分析，通过建立数据对象以及数据实例之间的关系，帮助业务快速定位关联影响。比如，构建逻辑实体和物理表的图谱关联关系，可以应用于逻辑实体或者数据资产、服务资产的上下游关联关系，确定上游数据源或者逻辑实体对下游的变化影响。

3.标签是对特定业务范围的圈定。在业务场景的上下文背景中，运用抽象、归纳、推理等算法计算生成对目标对象特征的符号表示，是用户主观观察、认识和描述对象的一个角度。比如，给特定对象打上业务标签，是实体属性的一种属性补充和扩展。

4.指标是对业务结果、效率和质量的度量。依据明确的业务规则，通过数据计算，得到衡量目标总体特征的统计数值，能客观表征对象的某一业务活动中业务状况的数值指示器。比如，统计某个周期的标准化同比增长率和环比增长率。

5.算法模型是面向智能分析的场景，通过数学建模对现实世界的抽象、模拟和仿真，提过支撑业务判断和决策的高级分析方法。比如，预测跨行业的风险预警等级和告警等。

5.6 政务数据共享服务框架

政务数据共享主要面向两类对象，一类是人，主要是指业务人员；一类是软件，主要是政务业务系统。面向业务系统的数据共享服务包含数据集服务和接口服务。面向业务人员的数据共享服务包含页面检索服务和自助分析服务等。数据共享服务架构图如下所示：

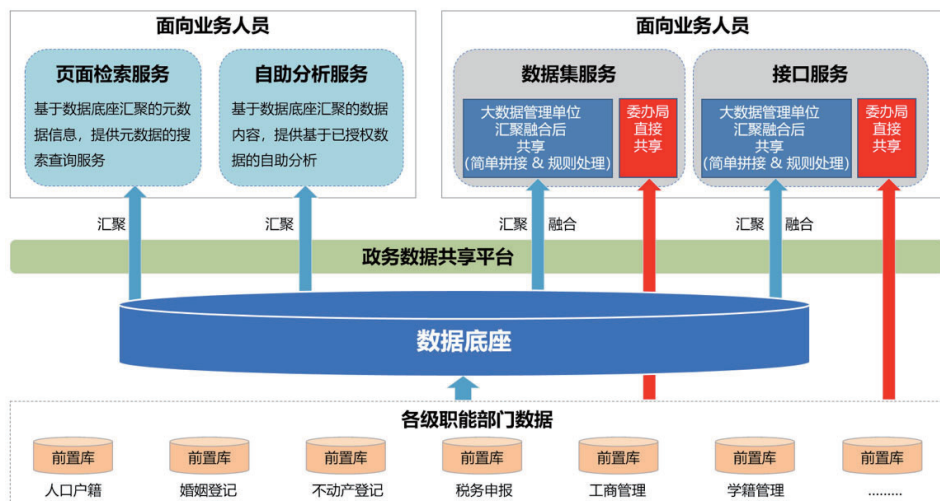


图5-13 数据共享服务架构图

5.6.1 数据共享服务的原则

数据共享服务以数据为核心，解决如何将数据方便、高效、安全地共享出去，降低数据获取难度，提升数据需求体验和效率。通过平台能力的建设，提供不同的数据共享服务形式，满足不同类型的数据共享服务需求。

数据共享服务的建设主要有以下原则：

1. 一致性原则。提供数据共享服务前，要确定每项数据的源头单位，由源头单位对数据的准确性、一致性负责。减少数据“搬家”，从而减少向下游二次传递所造成的数据不一致问题。

2. 黑盒原则。数据使用方不用关注技术细节，满足不同类型的数据共享服务需求。

3. 敏捷响应原则。数据共享服务一旦建设完成，并不需要按数据使用方重复构建集成通道，而是通过“订阅”该数据共享服务快速获取数据。

4. 自助使用原则。数据共享服务的提供者并不需要关心数据使用方怎么“消费”数据，避免了供应方持续开发却满足不了数据使用方灵活多变的数据使用诉求的问题。

5.可溯源原则。所有数据共享服务的使用都可管理，数据供应方能够准确、及时地了解“谁”使用了自己的数据，确保数据使用的合理。

5.6.2 数据集服务

（一）数据集服务概述

数据集服务是由服务提供方提供相对完整的数据集合或文件，数据以数据库表或文件的形式进行交换。数据集服务存在四种形式：

一是数据供应方直接通过传输系统将数据库表或文件传递到数据使用部门；

二是数据管理方直接将数据从不同的提供方传递到数据使用部门；

三是数据管理方对多个数据集进行简单拼接融合，但不改变数据后形成新的数据集，再传递到数据使用部门；

四是数据管理方对多个数据集进行融合，并按照业务规则对一个或多个字段的数据进行处理后形成新的数据集，再传递到数据使用部门。

对于业务规则是由多个数据集的数据供应方、数据集的需求单位、数据管理方共同协商确定，并固化在数据管理部门中，数据本身权责不变，谁的数据共享出去谁负责。

其中，前两种形式是当前政务数据共享中最常用的形式，第三种和第四种是共享服务提升演进的形式。随着业务需求的变化，简单的将数据提供方的数据传递给数据使用部门，已经不能满足业务诉求，需要数据管理方将相关业务数据汇聚在数据底座，并对数据进行融合处理，按照业务相关部门共同协商的取数规则，确定最终对外共享的、唯一的数据，由数据共享平台对外提供数据。

（二）数据集服务使用场景

主要使用场景1：对应形式一，如地方政府有关部门或人员直接访问统计局官网，获取第七次人口普查数据。

主要使用场景2：对应形式二，如数据管理方挂接卫健委等多部委的

疫情防控信息的目录，并没有存储多部委数据。数据使用方向数据管理方提出多部门数据集使用申请，经数据管理方处理、数据供应方审批与授权后，数据使用方从管理方获取多部门数据。

主要使用场景3：对应形式三，如数据管理方汇聚市场监管局、税务局、住建、自然资源、水利、林业和草原等部门数据，通过数据集的方式，批量的获取数据，生态环境部可利用数据进行追溯污染原因，结合空气质量检测数据，以达到动态感知，及时精准的获取空气治理效果。

主要使用场景4：对应形式四，如招商引资时，税务部门、土地部门、环保部门、人力部门甚至教育部门都会招商企业有不同程度的优惠政策。数据管理方综合处理各单位的优惠政策及适用条件，经数据整合、计算形成新的数据集，将此数据集共享给决策部门，用来辅助招商引资。

（三）数据集服务核心能力

根据支持数据共享服务的数据集有无关联处理，可将服务的能力分为无关联数据集服务能力、单一业务对象或资产的数据集服务能力和基于主题连接数据集服务能力。

1.无关联数据集服务能力

无关联数据集服务能力是指在数据进行结构化和标准化处理之后，数据质量尚可的情况下，业务对象之间无关联、主题之间无关联，将数据集直接对外提供服务。数据集可能直接从数据供应方到数据使用方，也可能由数据管理方传递给使用方。

2.基于数据湖的数据集服务能力

（1）将数据湖的同一个业务对象内的一个或多个资产封装为数据共享服务。在部分要求同一业务对象下多种数据的场景中，例如，政府投资项目管理时，业务对象是投资项目，该项目会包含若干个子项目及环评等项目，即可将子项目一、子项目二、子项目三的数据封装对外服务，形成可以支撑投资项目管理的数据集。

(2) 将数据湖内单个资产及其关联对象数据合并封装为数据共享服务。在部分需求、同一业务对象、多种数据且数据分属不同管理部门的场景下，需要向数据使用方提供相对完整的对象信息，以便于数据使用方自助分析。例如，政府投资项目管理场景下，可能需要项目审批数据、项目招投标数据等。由于分析服务面对的是具体的政府业务人员，而业务人员不可能读懂数据库中直接提取的项目ID等物理层主键或外键，并且没有必要让每个自助分析人员都重复进行共性数据联接。因此，可以在数据分析服务封装时，将必要的数据库联接在一起，比如将项目名称、项目编号和项目招投标数据合并封装为一个数据集服务。

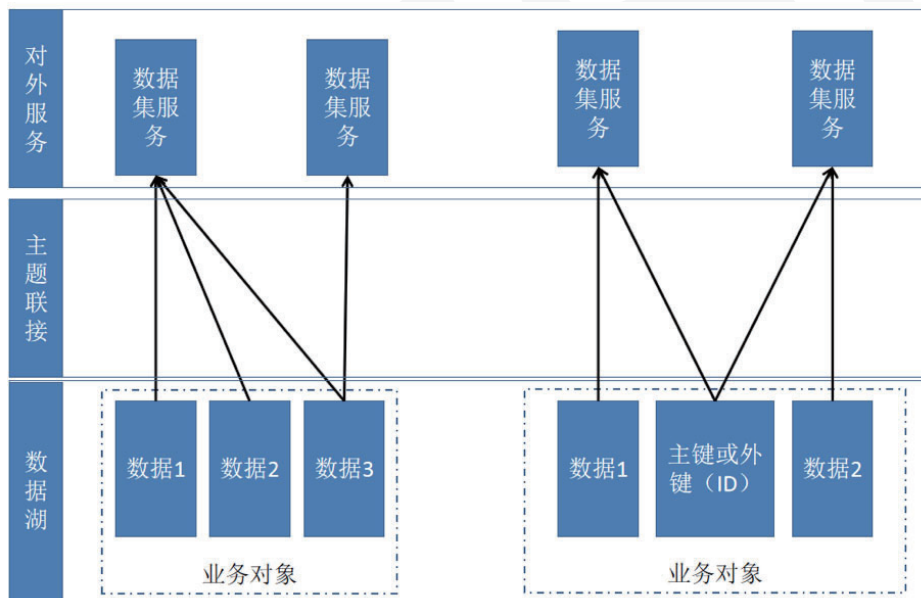


图5-14 关联相关数据

3. 基于主题联接数据集服务能力

(1) 将单个主题联接的数据资产封装为一个或多个数据共享服务。数据共享服务在面对不同数据使用方的不同需求时，可以适当地拆分为多个数据共享服务，以便更好地提供给数据使用方，减少冗余数据，提升数

据使用方体验。例如，教育场景下，面对不同的政务业务人员，需要提供学历、学籍等不同层面的信息以及幼儿教育、中小学、高等教育、职业教育等不同细分领域的的数据。比较恰当的方式是将多类需求分别封装为不同的数据共享服务，并确保这些数据共享服务的数据来源于同一个主题联接数据资产。

(2) 将由多个主题联接数据资产组成的多维模型整体封装为一个数据共享服务。在部分情况下，主题联接数据资产并不是以宽表（宽表是指在一张表格内，囊括所有需要的数据字段或属性）的形式落地，而是以多维模型的形式存在，此时可以将多维模型整体封装为一个数据集服务。例如，人社部的政务业务人员希望构建多角度分析的就业统计数据集时，就可能需要教育主题下的公民学历信息统计数据集、需要人口主题下的公民年龄或地域信息统计数据集、需要医疗主题下的公民健康情况统计数据集等。数据管理方将这些主题数据连接后，封装为一个新的数据集，满足数据使用方的就业统计需要。

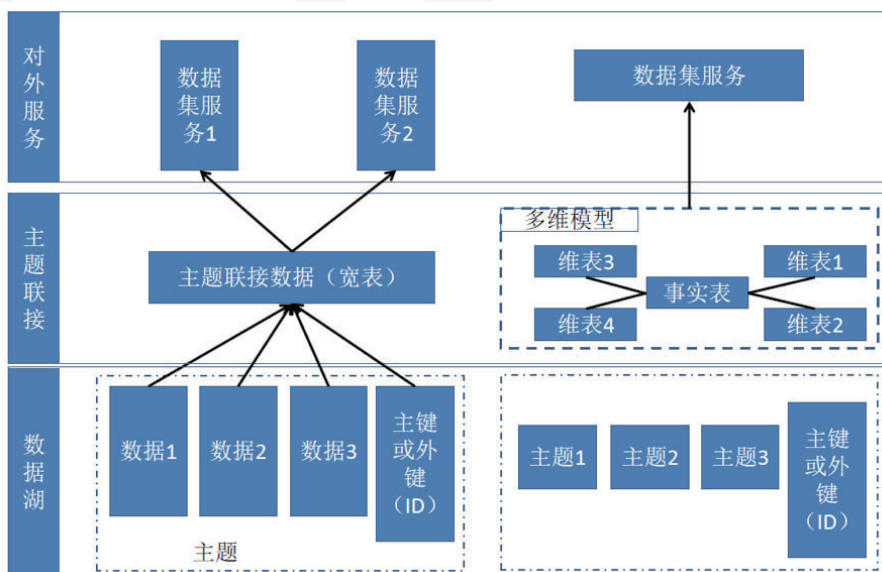


图5-15 多维数据模型封装

4.数据集融合管理能力

融合管理能力包括数据汇聚处理能力、数据检查能力、规范设计能力。

(1) 数据汇聚处理能力，该能力是指数据管理方将多部门多种类数据汇聚到自身后，需要从存储设施、网络网络、云、算力等多个层面构建硬件处理能力，和数据湖及主题关联能力。

(2) 数据汇聚条件检查能力，涉及服务需求识别、数据资产检查、数据资产质量要求，具体如下：

服务需求识别：包括汇聚数据需要有明确的数据应用需求；汇聚数据前需要充分分析数据应用需求，包括但不限于识别数据内容、数据需求频率、数据用途；汇聚数据前需要做服务复用性检查，判断需求所需的数据共享服务是否已存在，是否可以多场景复用。

数据资产检查：包括要服务化的数据需要进入数据底座，且数据资产必须满足标准（任命数据责任人，发布数据标准、有初始源或可信源、定义数据密级、数据质量满足要求、元数据注册）；封装成汇聚数据共享服务的数据，相关的汇聚逻辑，必须是由数据责任人提供。

数据资产质量要求：包括数据资产需要有主键、增量键字段或标识，能够唯一识别资产中的数据记录及增量数据；数据资产要有需求所需的所有维度字段；数据资产需要具有能够进行数据安全性控制的维度字段，以便于数据共享服务提供时满足数据安全控制要求。

(3) 数据融合管理属性设计能力规范，涉及明确数据融合责任人、明确数据融合安全策略、明确数据融合版本管理、明确数据融合SLA（服务标准）要求，具体如下：

明确数据融合责任人：数据融合的责任人取决于数据的责任人，跨业务域数据源封装的数据共享服务不改变原数据责任人的职责，大数据中心作为数据融合建设方，只承载汇聚逻辑的开发，汇聚逻辑由各数据责任人

联席确认，共同作为该数据融合的责任人。

明确数据融合安全策略：包括数据融合要有明确的密级定义；数据融合的密级取决于数据共享服务中包含密级要求最高的数据项，数据融合的授权审批遵从整体的数据安全策略；基于外部公开、内部公开的数据资产封装的数据融合原则上不需要进行安全性审批。

明确数据融合版本管理：针对基于API形式提供的数据融合，如果不提供向后兼容功能，则需要提供版本控制能力。

明确数据融合SLA（服务标准）要求：即数据融合必须有明确的可用性、响应性等SLA要求。

5.6.3 接口服务

（一）接口服务概述

接口服务是用API接口服务，根据数据量、计算规模的大小，实时或非实时地将数据推送给数据使用部门的服务。接口服务存在四种形式：

一是数据供应方直接通过接口将单条数据传递到数据使用部门；

二是数据供应方使用接头的方式，通过数据管理方挂接该接口，将单条数据传递到数据使用部门；

三是数据管理方通过接口形式，对多个部门的数据进行简单拼接融合，但不改变数据后形成新的接口，再传递到数据使用部门；

四是数据管理方通过接口对多个部门的数据进行融合，并按照业务规则对数据进行处理后形成新的最终对外共享的数据，再传递到数据使用部门。对于业务规则是由数据的多个数据供应方、数据的需求单位、数据管理方共同协商确定，并固化在共享平台中，对最终共享的数据权责不变，谁的数据共享出去谁负责。

（二）接口服务使用场景

主要使用场景1：对应形式一，如数据使用方通过国家卫健委官网，直接调用某一接口用于日常卫生健康相关业务办理。

主要使用场景2：对应形式二，数据提供方以接口的方式，将单一数据接口挂在数据共享交换的地方(如共享平台)。数据使用方需要获取某一具体数据值时，通过共享平台进行接口调用。如公安的人口数据、卫健委防疫数据、教育部的学历学籍数据、市场总局的企业数据均挂接于共享平台，当数据使用部门通过共享平台获取相应的数据权限后，即可进行调用，不用再于各数据提供方网站一家一家找，编写一家一家的接口调用程序。

主要使用场景3：对应形式三，当数据供应方的多个独立的数据接口无法满足数据需求方的要求时，需要对多个数据供应方的业务相关性强的数据进行简单拼接融合成一个新的数据接口，但不改变数据本身的值，然后再传递给数据使用部门。如数据管理方，汇聚扶贫主管部门的建档立卡贫困户信息、自然资源不动产登记信息、公安的户籍人口信息、死亡人员信息、车辆信息、残联的残疾人信息、市场监管局的工商注册信息等，通过身份证号等唯一标识字段进行关联，形成补贴发放人全景信息，再通过社保卡与人进行关联，将汇聚后的信息，通过封装成新的接口的方式传递给财政部门，以供其将补贴发放到农民手中。

主要使用场景4：对应形式四，数据使用部门所需的数据涉及到多个来源，并且需要对多个来源的相关业务数据汇聚在数据底座，按照业务规则进行处理后，确定最终对外共享的、唯一的数据，封装成新的接口后再传递给数据使用部门。如离婚状态信息，在民政部门、法院部门、公安部门等多个部门均存在，可能存在同一时间点的离婚状态的值是不同的，而数据使用部门（如人社部门）在需要离婚状态信息时，需要一个确定、唯一、有效的数据，在此场景下，需要三个数据供应方、一个数据使用部门、数据管理方一起协商取数规则，并固化在共享平台中执行，以便满足人社部门准确的数据需求。

（三）接口服务核心能力

1.常规的接口服务能力

资源提供方需在共享平台对接口进行注册、发布，核心能力包括服务注册、服务审核、服务撤销、服务变更、服务进度查看。数据需求方需在共享平台服务调用。

2.接口开发管理能力

接口开发包含数据API开发，函数API开发。

一是数据API开发是指通过编写SQL脚本的方式，将数据库提供的数据服务转换为API的能力。对数据提供方而言，对外开放其数据仓库的数据，使用更便捷的接口形态，提供业务增值能力，通过简单的操作，即可快速、低成本、低风险的开放数据或服务。对数据使用方而言，不再需要定制各种连接客户端，只需要使用一个简单的请求发送能力的客户端即可轻松驾驭各种数据仓库的能力。

二是函数API开发是指通过编写代码的方式，对多个API进行编排和适配，封装为一个新的API。对API数据使用方而言，通过将多个细颗粒度的API组合为一个特定场景功能的API，可以简化API数据使用方的开发难度，降低对API的学习成本，降低开发成本，缩短业务应用开发和上线时间。当API有所变化时，可以通过函数API进行适配，尽可能保持提供给API数据使用方的API不发生变化，避免因API变化而导致单个或多个API数据使用方的应用修改开发。

3.接口服务管理能力

数据接口服务包括管控平台、服务引擎、服务状态监控、可视化开发工具、传输代理、调度引擎、数据发布引擎、资源采集客户端等。

(1) 管控能力：数据共享服务统一管控，提供资源目录、数据开发、数据使用、服务引擎管理、调度计划、统计监控、服务权限配置等系列功能，对数据资源和服务资源统一注册管理等管控能力。

(2) 服务引擎：批量作业服务及文件传输服务引擎，负责批量作业模型解析、批量作业执行、文件传输服务；提供多协议、多数据源适配支

持；为服务运行提供高性能、高可靠的运行环境。

(3) 服务状态监控：提供日志解析及监控能力，对事前预警、事中告警、事后统计分析等功能提供后台支撑。

(4) 调度引擎：作业调度引擎，支持作业与作业流的串行及并行调度，提供日历、频度、事件等多种规则的调度，为作业和作业流运行提供多样化的调度方式。

(5) 数据发布引擎：基于相关架构提供数据服务发布及数据访问能力，支持单表、结果集等形式的实时接口服务发布。

(6) 资源采集能力：实时采集各引擎所在物理机的CPU、内存、磁盘、网络等指标，并进行展示。

4.接口服务调用监控能力

共享平台对接口服务的调用进行监控，包括调用方、调用次数、接口服务的数据来源、服务质量等内容。

5.数据融合处理能力

与数据集服务中的数据融合处理能力类似，包括数据汇聚处理能力、数据检查能力、规范设计能力。

5.7 开发利用新技术

5.7.1 人工智能

目前人工智能在政务领域的应用尚处于起步阶段，已经开始在一定程度上产生积极的带动效应。从政务数据开发利用的角度而言，人工智能将极大解决传统政府的效率缓慢、机构臃肿、部门协调和公共服务缺乏精准化等不足的问题，为构建高度柔性动态和为民服务的政府体系提供支撑。

人工智能将在一下几个方面有效提升政府服务能力：

一是有助于解决传统政府人力资源缺乏。传统政府不能无限制地扩大规模和增加人员，然而面对日益复杂的公共管理和服务诉求，政府越来越

面临着人力资政府创新源缺乏的问题。表现为大量的日常执法、决策、行政流程运转所需要的人力越来越多，为了解决这一困境，政府采用了购买公共服务、临时聘用等各种手段，然而其解决的效果依然较为有限。人工智能的应用可以极大缓解这一问题，在信息收集、行政流程、行政咨询应答等领域，可以大量替代传统人力投入，改善政府的人力资源局限。

二是有助于提升传统政府的决策质量。传统政府基于人工的公共事务决策，由于收集数据有限，决策效果难以精确化，因此往往存在很大程度上的决策质量不高与不确定性的问题。决策质量的提升成为传统政府改进的最重要的领域。而人工智能可以全面提升更有效的决策信息支持，并根据需要自动生成相应的决策方案，供决策者选择，从而极大提升政府的决策质量。

三是有助于建立顺畅、高效民政双向沟通方式。公民交互渠道传统政府与公民的关系是相对刚性的，面向政府缺乏足够的人力和相对的渠道实现非常有效的公民联系，当公民与政府发生交互时，实际上也提供了大量的有效信息给政府，而传统的交互方式则很难收集和分析这些信息用以改进政务服务。可以通过人工智能建立公民的政务服务助手，实现随时随地对政务服务的咨询和协助解决，同时有效地将互动结果进行分析从而改进政务流程。

四是有助于构建更精准的公民、企业的个性化服务。面向众多的企业、公民如何提升服务质量，服务对象数量众多而差异化诉求又很大，实现个性化、精准化的服务，通过传统的人工方式基本无法实现。构建海量公民、企业的个性数据库，在大数据基础上的人工智能手段，有针对性地针对每一位公民建立完备的数据档案，并适时调配公共资源满足公民的需求。

5.7.2 区块链技术

区块链在促进数据共享、优化业务流程、降低运营成本、提升协同效率、建设可信体系等方面具有优势。目前，影响各级政务部门间数据共

享的主要问题包括：共享交换效率低、安全共享手段少、数据有效性差、目录信息不准确、权责不清晰、管控性弱等。具体而言，共享效率低：人工走平台的发文、申请、协调等方式获取数据，流程长、效率低；安全共享手段少：数据的可复制性导致敏感信息泄露，数据安全计算，可用不可得的技术手段少；数据有效性差：数据非实时归集容易出现无效数据，大数据平台对业务理解较弱，无法充分管理数据；目录信息不准确：大数据平台对目录信息的管理较弱，产生目录和实际数据不一致的情况；权责不清晰：数据的所有权和管理权分离，容易引起维护、安全等问题；管控性弱：政务数据管理部门对数据实际使用者、使用目的、使用频率等信息掌握不足。

构建基于区块链的政务数据开发架构，推动跨部门、跨区域政务服务一体化协同建设，可以有效推进解决上述问题，夯实政务数据流通、开放与共享能力等方面具有无法替代的现实价值。

5.7.3 可信计算

可信计算的主要思想是引入安全芯片,并从系统内核层到应用层都构建信任关系,建立相应的信任链,利用可信关系实现系统的自我保护、自我管理和自我恢复的机制。可信计算思想引入电子政务中的应用时,主要实现以下设计目标：建立相关机制,初步实现可信终端与可信服务器的双向认证。实现安全的数据共享,基本解决信息孤岛问题。形成基础数据资源库,实现部门数据集市,实现用户的统一认证、统一管理。提供可靠的公文流转平台,有效的缩短公务处理的周期,加快公文流转过程,保持公务处理渠道的畅通,实现无纸化办公、远程办公、异地办公、移动办公。通过信息网络平台,提高信息统计和分析的水平,为领导提供决策的辅助支持。提供完善的信息交流机制,方便各部室、中心和办公人员之间日常办公及业务信息的交流和管理,实现业务办理互联互通、信息数据共享。充分利用现有建设好的网络环境及已有的网络配置资源,充分地保护已有的

信息化投资。

5.7.4 可视化技术

可视化技术是基于人眼的视觉感知和人脑的智能认知，综合运用计算机图形学、计算机视觉、人机交互和认知心理学等领域的技术，将抽象的数据转化为直观可见的图形图像以增强人类感知、帮助人们更有效地完成任务的工具。

可视化技术在政务领域的主要应用包括：

1.自助数据建模

传统的基于数据仓库的数据建模用户只能用已经创建好高质量的数据仓库，而在运用数据可视化技术下，用户可以通过简单的可视化操作，将数据库中的数据表创建成待分析的数据模型，可以通过编写SQL语句建立一张虚拟的表创建模型和定义虚拟表之间进行关联，建立表与数据表和数据视图关联建模，把数据库中的数据视图作为一个虚拟的表创建数据模型。

2.可视化编排

政务大数据可视化平台设计是基于纯代码开发的，可以在大多数浏览器下进行数据可视化设计，设计器可以编排适合移动端显示的页面，并提供分辨率自适应功能，以确保编排的页面能够在一定范围内适应多种不同分辨率的显示终端。平台提供的可视化组件，包括各种统计图形、图表组件等，用户只需要通过简单的拖拽和参数配置，即可创造出专业的数据分析报表、仪表盘、数据可视化大屏以及适用于移动设备的图表页面。为了在进行数据分析的同时，也能编排出一个完整的数据呈现WEB系统。大数据可视化平台设计增加了独有的WEB交互控件，可以设计出多样化的，具备交互能力，可用于数据呈现、数据查询、数据联动、数据分析的全套WEB查询系统。

3.多维数据分析

政务大数据可视化平台能提供各种常见的联机分析处理操作，除基本的分组聚合外，还可以进行任意多维度分析。如拖拽式自由分析，在多维数据分析一体化工作台中，可以选择任意数据模型中的任意维度、度量进行自主拖拽分析。在进行多维分析时，支持自由创建维度间的层级关系，将具备层级的维度应用到各种图表上，即可实现自动多层钻取能力，逐层分析和查看各种数值情况。同时，在数据分析工作中，除了对普通数据的维度、度量进行直接分析和观察数据的整体效应外，也能分析平均数、最大值、最小值等常见需求。

4.支持集成框架和外部引用

数据可视化设计能够非常方便地与其他应用进行集成。提供专门的组件、提供二次开发调用接口，让开发者很容易就能进行二次开发。在平台集成和客户化功能方面，设计开发的报表可以灵活和第三方进行集成应用和跨平台部署。数据可视化平台发布的报表工程、页面，能通过API等接口方式嵌入到第三方应用系统。根据安全要求以及第三方系统特性，可同时支持免登录方式和权限认证方式。

5.7.5 数据图谱

数据图谱（知识图谱）是通过将应用数学、图形学、信息可视化技术、信息科学等学科的理论与方法与计量学引文分析、共现分析等方法结合，并利用可视化的图谱形象地展示学科的核心结构、发展历史、前沿领域以及整体知识架构达到多学科融合目的的现代理论。它把复杂的知识领域通过数据挖掘、信息处理、知识计量和图形绘制而显示出来，揭示知识领域的动态发展规律，为学科研究提供切实的、有价值的参考。

其在政务数据开发利用方面的应用主要涉及以下几个方面：

1.数据存储：借助图数据库，将数据湖中所有的数据进行深度融合和关联存储，实现从数据到信息、从信息到知识、从知识到情报的整个价值

链和智能推理研判。

2.数据检索：借助图库检索语句，能实现数据、关系、链路信息的快速检索。

3.多层关系探究：借助图算法、图形化的展示，可以探究数据之间多维度、的关联关系，包括N度的关联关系。同时借助图归类的算法、对数据进行划分、归类、管理。

4.快速建模：借助图算法、图运算功能，业务人员可以快速、方便的完成复杂的数据模型，解决特定业务场景库的建设；通过深度学习技术，可以做到AI建模、自助完成主题、场景库的建设。

5.时空数据分析：通过时间信息、地理信息的数据落图，实现空间化数据的图形化展示，将时间与地图轨迹进行关联。

六、实践与探索

各地区政府开展政务数据开发利用探索，形成极具地方特色的实践案例和典型应用。以部分实践为例，具体情况如下。

表6-1 地方政府政务大数据利用实践案例

序号	地区	案例名称
1	成都市	成都市公共数据运营服务平台
2	乐山市	乐山市智慧城市大脑案例
3	重庆市	重庆市政府大数据资源中心实践案例
4	山西省综改区	山西综改示范区基于数据统筹项目管理体系建设
5	唐山市	唐山市政务大数据开发利用实践案例
6	漳州市	漳州市数据汇聚共享服务平台建设项目案例
7	贵阳市	贵阳市政务数据共享开放考核方案实践探索
8	成都市	成都市公民大数据平台案例
9	贵阳市白云区	贵阳市白云区政务数据开发利用实践案例
10	深圳市罗湖区	深圳罗湖区大数据中心建设项目
11	苏州市吴江区	苏州吴江区大数据平台二期建设项目案例

6.1 成都市公共数据运营服务平台实践案例

6.1.1 案例背景现状

基于成都市公共数据运营开展和市场实际情况，成都市政府在确保数据隐私安全的前提下，以数据价值变现为目的，以数据产品和服务为抓手，提出了通过数据运营首先盘活数据要素市场中最权威、最完整的公共数据资源的发展思路，并于2017年明确了本地国资载体，在全国率先启动政务数据授权运营的路径探索，开创了以公共数据运营为核心的数据要素发展的“成都模式”。

6.1.2 案例内容

成都市公共数据运营服务平台作为成都市政务数据资产运营管理平台，通过构建可信应用云，在确保安全合规和数据隐私安全的前提下，融合安全多方计算、可信执行计算环境、联邦学习等隐私计算技术，在联合数据分析、隐私数据求交、联合建模等场景中开展应用。在保护原始数据不出域的同时，汇聚贯通政府、产业、社会第三方的数据资源，打破各机构间的既有数据壁垒，实现跨机构、跨行业、协同多方进行多维数据的融合。通过全新范式实现“不拿数据拿结果，不汇聚数据但汇聚价值”，完成数据流通向“价值”流通的升级，实现数据隐私保护与价值挖掘之间的平衡，提升数据价值挖掘下的隐私保护信任，向社会机构提供合法有序、契约化的政务数据开放服务，构建“数据可用不可见”的合作新模式，从而促进可信的数据流通，链接多维的数据价值，构建繁荣的数据生态。

6.1.3 案例成果/成效

在数据确权与定价尚无定论的背景下，成都创新模式对数据要素的发展具有重要的开创性意义。主要体现在以下几个方面。

1.发展思路创新

一方面该模式打破了数据流通中对数据资源确权、定价的依赖限制，形成数据价值流通理论，推动数据价值的挖掘、传递和变现，形成数据保值增值和开发利用可行路径；另一方面，在数据权属不明晰的情况下，通过本地国资公司运营，可以确保令行禁止、客观中立，也有效避免数据作为新型国有资产的流失风险。

2.管理机制创新

通过建立科学的公共数据运营服务管理工作机制，确保数据安全和隐私保护，有效保障公共数据运营工作的顺利开展。其中网络理政办负责指导、监督和协调推进政务数据授权运营服务工作，负责各部门、各有关单

位的数据汇聚，为数据运营提供资源保障；市政府各部门、有关单位（即数据提供单位）负责确认政务数据运营需求，基于风险评估结果对可以运营的数据进行授权确认；市大数据集团，负责建设维护并管理成都市公共数据运营服务平台，根据部门授权确认意见，采用市场化方式依法依规开展政务数据运营服务，并建立透明化、可记录、可审计、可追溯的全过程管理机制，形成政务数据使用的全程记录，再将日志数据提供给成都市政务信息资源共享平台，定期将运营服务、数据交互、数据使用情况汇报给市网络理政办、市新经济委和数据提供单位。市网络理政办不定期检查政务数据运营平台的数据使用日志和安全保护情况，加强数据运营过程中的数据安全保障。

3.技术创新

一方面，通过应用可信执行环境（TEE）、多方安全计算（MPC）、联邦学习、深度学习等技术手段，营造安全、可信、可计量的数据流通环境，建立数据安全防范、监测、预警和应急处置机制提升安全应急管理能力和数据流通审计机制提升数据安全监管能力，解决了当前“安全可靠”与“数据流通”之间互斥的难题；另一方面，基于成都超算中心的算力支撑，有效解决了“隐私安全与性能”之间的平衡问题。

4.理论创新

形成了一套兼顾数据价值变现、数据安全与群众获得感的数据运营方法论。一是开展了数据要素流通的分账机制研究，为数据要素跨区域跨层级流通，奠定了理论基础；二是开展了政务数据授权运营的实现机制与内在机理研究，以成都市为案例，探索政务数据授权运营新模式、新方法，其经验做法为全国各省市开展“政务数据授权试点”工作提供了实践指导。

6.2 乐山市智慧城市大脑案例

6.2.1 案例背景现状

依据四川省“一干多支”发展战略和乐山市委、市政府城市发展部署要求，项目由乐山市数字经济发展局牵头，市级部门密切配合，以心连心服务中心为基础，运用大数据、AI、5G等新技术，汇聚城市运行监测数据，建设市心连心“1+3+N”信息互联互通平台。平台以数据作为战略价值的核心，打造集城市管理中心、服务中心、应急指挥中心、运营资源整合中心于一体的乐山“新型智慧城市枢纽”，实现“一个城市，一个中心”，形成了具有鲜明特色的智慧城市“乐山方案”，并且围绕以下问题攻坚克难。

一是在当前社会经济飞速发展下，百姓需求、企业需求、政府要求与城市宜居环境平衡、经济产业稳定发展、政府精准决策等城市亟待解决的发展瓶颈相结合，解决技术手段跟不上社会发展的实际问题，实现各类应用场景快速搭建和拓展迭代，以场景应用驱动数据融通。

二是重点面向政府系统数据堰塞的情况，通过“数据大归集”将信息系统、应用数据、视频摄像、物联感知设备等，叠加到城市运行态势一张图上。解决数据因停留在“纸面上”、“文件里”、“硬件中”而无法有效形成数据效能的问题，以数据元素为基础形成城市态势感知体系，将“数据融通”赋能“协同治理”。

三是针对城市管理九龙治水的局面，平台将“单打独斗”的业务重构为“纵横联动”的场景，从城市治理、应急处置、社会诉求等业务实质出发，解决城市治理过程中一些问题“无人问津”、部分问题“多头治理”的情况，打通“城市大脑”与管理末端的“神经”，基本形成的智能城市治理运行体系。

6.2.2 案例内容

围绕城市治理和城市服务所面临的焦点、难点和堵点，根据新型智慧城市建设基础和目前所面临的形势，运用全新的信息化、智能化手段，汇聚城市数据，进行合理的分析和调度，改善城市运行环境，提升城市管理智慧化、服务人性化、应急快速化、决策科学化水平。为充分发挥在城市治理、城市决策和综合展示等维度的支撑性作用，构建平时模式、战时模式、会客模式三大工作平台，实现三个一：一个最优城市运营管理平台、一个最佳城市决策指挥中心、一个最靓城市综合展示窗口。

一个最优城市运营管理平台：全面接入综治、交通、城管、市场、生态、文旅、食品等部门数据，实现对民生诉求和城市运行状态的全源实时感知、态势预测、事件预警和决策支持，提高跨部门跨领域的协同指挥能力，形成平战结合的运行管理新模式，构建跨层级、跨区域、跨系统、跨部门、跨业务综合协同管理和服务平台，实现政府治理体系和治理能力现代化。

一个最佳城市决策指挥中心：通过对接入的城市信息资源进行整合与展示，实现城市运行的集中监控，城市应急的统一指挥和协同联动，为领导决策提供技术支持。透过实时各项城市信息的汇集与分析研判，为政府领导平时常规监管，战时应急指挥提供平台级服务。

一个最靓城市综合展示窗口：运用3D、GIS、AR以及可视化等技术，借助3D模型、综合大屏、手机移动端等载体，将汇聚整合的数据通过各种维度进行综合展示，提供宏观总体、中观区域以及微观细节的城市运行数据实时呈现，构建城市宣传会客厅。

平台总体架构采用设施共建、技术衔接、信息共享、应用协同、服务集成的思路设计，逻辑上分为感知层、基础层、数据层、支撑层、应用层、接入层6个层次和标准规范体系、信息安全体系、运维保障体系3个配套支撑体系。

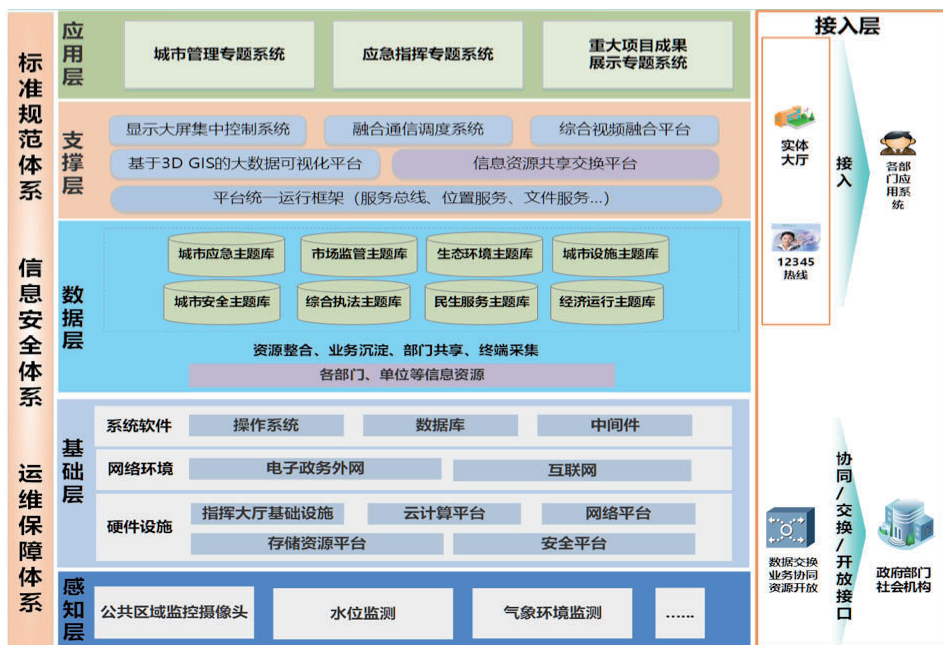


图6-1 乐山市智慧城市大脑总体架构

感知层：支持城市已建物联网系统的全面接入，包括人脸识别门禁、公共区域监控摄像头、车辆车牌识别、地下管线监测、消防通道地磁、气象环境及水位监测等。

基础层：支撑平台运行的基础，包括硬件设施、网络环境及系统软件。其中，硬件设施包括指挥大厅、配套机房、主机、存储、备份、网络、安全、终端等硬件设备及环境；网络环境依托电子政务外网、互联网、应急指挥通信网、视频专网实现安全接入。系统软件包括操作系统、数据库系统、中间件等通用工具软件。

数据层：通过资源整合、业务沉淀、部门共享、终端采集、互联网抓取等多种手段，构建社会治理主题库，包括综合治理、交通服务、城市管理、市场监管、生态环保、食品安全、文旅服务、民生诉求、城市应急等。实现信息留痕、集中、对称、共享、公开以及应用呈现，以此逐步建立大数据应用机制，支持大数据分析研判，助推创新发展。

支撑层：支撑层针对各类业务应用从基础环节进行统一规划，将应用系统与信息资源、基础设施无缝衔接，构建应用系统提供标准、高效、易于集成的通用公共组件，通过建立运行服务统筹、GIS资源统筹、视频资源统筹、可视化展示统筹、城市物联网接入统筹，为系统提供服务总线、融合服务、消息服务、工作流引擎、智能报表、地图服务、视频接入、任务调度、业务联动、统一视图等支撑服务。

应用层：应用层依托应用支撑平台，支持并运行城市综合管理应用和城市应急指挥应用。

接入层：接入层基于应用层之上，依托电子政务外网和应用支撑平台，提供包括实体大厅、统一服务热线、移动终端在内的多渠道接入服务。用户可在共享的身份管理和授权控制下通过统一门户访问平台各类应用功能。此外，可通过协同、交换、开放接口，实现智慧治理体系内部横向、纵向信息共享与业务协同，实现与其他政府部门、社会服务机构的数据交换。

配套支撑体系：支撑体系包括标准规范体系、信息安全体系和运维保障体系。标准规范体系是以《标准化工作导则》和《电子政务标准化指南》为基础，制定相关信息化行业标准和技术规范，保障规范建设和使用。信息安全体系是根据国家安全等级保护要求和信息安全需求，构建信息安全与业务紧耦合、核心关键技术自动可控、安全工程集约化与过程规范化的信息安全体系，保障系统和数据的安全。运维保障体系系统运维包括数据运维和应用系统运维，确保数据更新及时、系统长效稳定运行。

6.2.3 案例成果/成效

（一）建设成果

平台建设以“业务驱动+场景主导”，采用“1+3+N”的智慧治理模式，汇集了市公安局“天网”、市委政法委“雪亮工程”、市交通运输局、市水务局、市城管局等部门以及峨眉山、乐山大佛两大景区12879路

视频资源，归集了58个部门681类2亿多条业务数据，建成1套基础信息互联互通平台和“心连心服务”、“城市综合管理”和“应急指挥调度”三大功能板块，接入了“N”个部门主题业务系统，实现了跨区域、跨部门数据资源高度整合，以及全市城市治理态势的综合展示和城市应急处置能力的提升。在城市综合治理方面，平台将监管与服务并重，全面提升社会治理能力和服务水平；在心连心服务应用场景中，多位一体受理群众诉求，“闭合循环”及时回应关切；在应急指挥板块，实时感知城市运行情况，防范化解风险，提升应急处置能力。

（二）应用成效

一是实现“智慧精准高效”服务新突破。基于接入系统数据汇聚分析展示，全面掌握社情民意动态，从事后监管向事前防控转变，从分段管理向全链条无缝监管转变，切实为企业和群众排忧解难，打造城市服务民生中心。

二是建成“纵横联动”城市治理新体系。开启跨部门高效协同新局面，实现重点业务领域全流程监测和全过程管理，实现更深层次的信息共享和业务互认，对城市状态进行多维度分析，预判防范城市风险。

三是开创城市应急指挥调度新模式。通过对接入的城市信息资源进行整合与展示，为政府领导平时常规监管，战时应急指挥提供平台级服务。实现城市运行的集中监控，城市应急的统一指挥和协同联动。

6.3 重庆市政府大数据资源中心实践案例

6.3.1 案例背景现状

重庆市政府大数据资源中心项目，基于数据治理体系，构建完善的政务数据资源共享系统和政务数据资源开放系统，升级自然人、法人、自然资源和空间地理、电子证照等基础数据库，围绕应用建设主题数据库，根

据部门责任清单建设部门政务数据资源池，形成“两个系统+四大基础库+N个主题库+N个部门政务数据资源池”的数据资源汇聚体系，构建全方位、多层次、一致性的安全防护体系，建立数据接入、数据处理和数据使用多层次的数据全生命周期管理体系。

6.3.2 案例内容

按照重庆市大数据资源中心建设的标准体系的定位，充分考虑标准体系的纵横关系，设置市大数据资源中心项目建设之标准体系结构：业务规范标准、数据规范标准、技术规范标准、管理规范标准、质量规范标准、安全规范标准。

数据治理矩阵具体包括：数据质量、数据标准、元数据管理、数据安全、数据模型、数据资源目录、数据生命周期管理、数据产品工厂8大模块，同时制定了数据管理制度框架，将最高管理层对数据管理的期待变成“法典”。

6.3.3 案例成果/成效

截至目前，市共享系统已有70个市级部门（含二级部门）挂接数据资源，累计挂接数据资源2922类，资源挂接率为75.52%。其中，实时共享部门48个，实时共享数据2037类。

以服务接口挂接的有815类，占比27.89%，（服务已接入数量除以总的接入的数量）库表资源挂接的有1222类，占比41.82%，（数据库已接入数量除以总的接入的数量）挂接文件的有885类，占比30.28%。（文件已接入数量除以总的接入的数量）。以有条件共享方式汇聚的有1293类，占比44.25%，以无条件共享方式汇聚的有1629类，占比55.74%。

数据调用量创新高，2020年内市共享系统共支撑49个市级部门和30个区县，市共享系统日均调用330万次，累计数据调用超10.3亿次，累计访问流量3.08TB。

本项目作为数据治理体系的具体体现，以成熟的实施方法论为基点，以应用创新为目标，取得了不错的成效。

（一）创新性

数据资源共享系统建设，打破了传统的数据目录建设思路，围绕政务部门权力事项，以“互联网+政务服务”、“互联网+监管”工作为重点，全面提高了政务数据质量和数据服务能力，实现了创新行政管理和服务方式。

（二）共享性

依托“数字重庆”云平台集中建设城市大数据资源中心，统一向社会各界提供数据的编目、汇聚、共享、开放、应用等服务，实现数据的“一云承载”和服务的“一窗提供”。

（三）一致性

充分利用现有信息系统成果，编制全市统一的政务数据资源开放清单，实现了数据开放目录和共享目录的一致性。

6.4 山西综改示范区基于数据统筹项目管理体系案例

6.4.1 案例背景现状

山西转型综合改革示范区是国家级经济转型综合改革示范区，定位于“体制机制创新的先行区、营商环境的引领区、创新生态的先导区、战略性新兴产业聚集区”，肩负为国家“在转型发展上率先蹚出一条新路来”的使命。2020年初，示范区发展规划中心启动了“数据融合共享及开发利用为核心的智慧园区信息化项目建设运营统筹管理体系”建设。

山西转型综合改革示范区政务数据融合共享及开发利用，重在解决政务数据赋能招商引资、赋能企业生产经营、赋能区材区用、赋能产业创新转型。

6.4.2 案例内容

该示范区智慧园区建设运营统筹管理体系框架如下图所示。

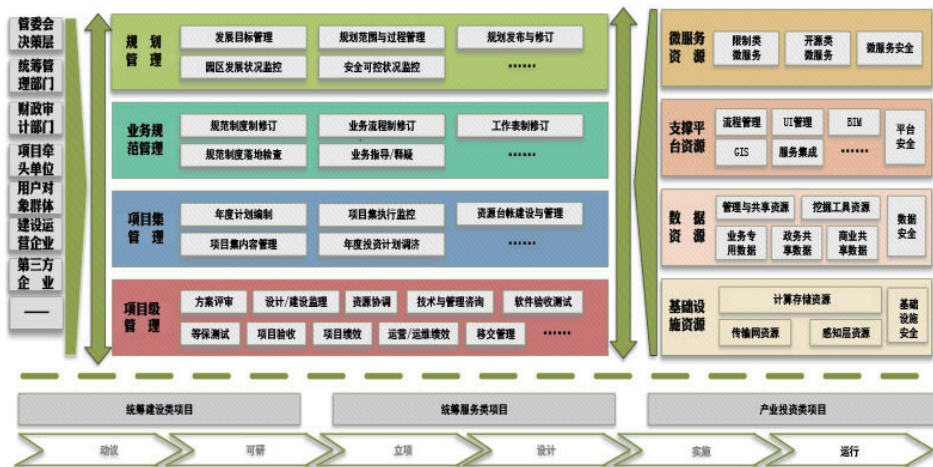


图6-2 智慧园区建设运营统筹管理框架

园区数据统筹管理体系建设的主要内容包括以下几个方面。

（一）标准规范体系建设

园区数据管理相关标准规范建设内容主要包括，数据资源分类方法及分类代码、原数据与二次数据定义方法、数据资源管理属性定义、数据资源台帐技术规范、数据资源归集管理办法、数据治理业务分类方法、数据治理分类操作业务规范、数据资源发布业务规范、数据资源共享审批管理业务规范、数据共享交换节点功能定义、多数据域关联分析业务规范、数据流通审计管理规程等。

其中，与数据开发利用相关的数据资源管理属性包括数据名称、数据格式、数据生产单位、生产时间、数据作用业务域及信用等级、数据源位置、数据访问方式、数据作用域与及流量阈值等核心内容；与数据开发利用相关的数据流通审计管理包括授权系统标识、授权操作人、使用数据的业务系统标识、使用数据的用户标识、访问方式、访问时间、数据流量等。

（二）项目集管理体系建设

项目集管理规范覆盖了专项规划制修订与监控、项目立项审批、项目采购、项目合同签订、项目实施、项目监理与验收、等保测评、运维运营、区域信息安全与应急等项目全生命周期，并明确了项目不同阶段，数据生产计划管理、数据生产过程管理、数据资源归集与共享治理管理、数据流通管理、数据开发利用管理工作中，所有干系方的权责与义务，给出了各干系方的工作内容、工作方法、工作用表。

以数据开发利用为核心的建设运营管理内容主要包括，项目可研编制要求明确业务、数据、来源关系表；立项评审要求对数据重复采集、数据流通域界定进行审核；实施管理明确规定了数据共享干系方的职责；验收管理明确了项目验收前置条件包括数据备案证明；运维经费申请/支付前置条件包括数据备案证明/数据更新证明，对需要共享其他业务系统数据的运维项目，同时需要提供数据共享质量评价表等。其中，数据备案证明是该项目生产数据明细及管理属性是否在数据统筹管理部门登记的证明材料，数据共享质量评价表是使用共享数据的业务部门对共享数据完整性、正确性、可用性等的的评价，该评价一定程度上影响共享数据输出系统的运维费申报。

（三）数据供需管理

“示范区信息资源管理信息系统”基础数据仓库包括，数据资源库、应用软件资源库、中台能力资源库、传输网资源库等等，其中，数据资源库用于存储包括数据流通与开发利用相关的数据资源属性。

该管理信息系统支撑的数据备案业务用于园区数据资源的登记；数据资源分类检索业务为不同用户群体（例如，业务部门、可研方、设计方/开发方、运维/运营方）提供不同的数据资源检索权限、范围、结果；数据申请审批业务用于数据共享申请的全流程管理；数据对接管理业务用于不同项目之间数据共享的协调管理；数据质量评价业务用于使用共享数据

的业务部门对共享数据质量的评价等。

（四）体制机制保障

山西综改示范区为加速政务数据的共享流通和开发利用，一是将数据备案与项目建设验收挂钩，确保新建或迭代的业务系统数据及时登记；二是将数据备案与运维运营类项目验收挂钩，确保在产业务系统数据及时登记；三是建立数据共享审批机制，促进不同项目间的数据共享；四是建立项目级数据共享质量评价机制，促进被共享数据的质量提升；五是建立常态化的数据需求挖掘机制，持续增强各业务部门的数据共享需求；六是建立定期培训制度，不断提升各业务部门数据请求能力；七是第三方专家团队驻场服务，协助示范区发展规划中心对体制机制运行进行支撑、监控，发现瑕疵及时解决，不断完善相关规章制度、标准规范的细节内容。

6.4.3 案例成果/成效

该项目已输出成果包括《智慧园区 统筹管理体系框架》《智慧园区 数据资源管理属性定义方法》《智慧园区 信息资源分类及共享管理办法》《智慧园区 数据资源备案业务规程》《智慧园区 数据共享审批业务规范》《智慧园区 共享数据接口基本功能要求》《智慧园区 数据共享质量评价方法》《智慧园区 项目分类及编码》等，以及围绕数据管理为核心要素的项目立项、采购、实施、验收、运维、监理、测试、评价等的项目管理规章制度与业务规范30余个。

截止2021年1月，核减新业务系统拟采集数据150余种，以共享方式向新建业务系统提供已有数据；16个业务系统数据共享已打通；13个竣工项目业务系统数据清单均已备案，且具有数据共享输出要求的业务系统数据均已与区级数据共享交换平台对接；面向园区企业开放的各类政务数据500余种。

从“重技术平台类建设”转向“重数据管理与开发利用”，是投资意识的重大转变。近一年，山西综改示范区依托政务数据反哺企业生产经

营，企业转型与产品创新速度明显加快，政府在招商引资与经济运行调度方面施策更加精准，园区招商服务效率及园区经济产值均提升了20%以上。

6.5 唐山市政务大数据开发利用实践案例

6.5.1 案例背景现状

互联网大潮推动数字化背景下小微企业金融服务升级。从2010年开始，移动互联网的大潮势不可挡的奔向社会的各个领域，随着电商和三方支付的成熟，小微商户成为了切入金融服务市场最自然的赛道，这是一件足以“改变银行”的事件，这也预示着互联网金融这一崭新事务的到来。数字化填补了信息的沟壑，小微企业的“硬性”抵押产品不再是唯一信用评价指标，“软性”的数据信息被纳入银行的视野中，数字化加持的银行开始精准触达小微企业“毛细血管”。

巨量小微企业信贷风控缺口仍旧长期存在，客观上仍有数量巨大的小微信贷缺口的现实。小微企业融资难本质上是风控难。风控难体现在两个方面，一是风控效果难提高，小微企业因管理不完善、报表不规范，信用水平难以依靠其提供的信息进行很好地评估；二是风控效果和成本难两全，即如果按照传统评估方式通过纯人工的方式进行信息搜集、风险评估，其投入将大幅提高风控成本。

6.5.2 案例内容

为破解中小微企业信贷融资难题，唐山“春雨金服”平台脱颖而出。在国务院第七次大督查发现的43项统筹疫情防控和经济社会发展典型经验中，唐山市的“春雨金服”平台榜上有名。



图6-3 唐山市企业综合金融服务平台

作为春雨金服的载体，“企业综合金融服务平台”的建设充分发挥了政府组织优势和技术创新优势。运用自主研发的智能分析决策技术对行方、政府机构内外部数据进行融合、加工、挖掘、建模和关联性分析；进行风控建模刻画企业画像，生成企业评价报告，有效解决过去融资过程中信息不实、信用不足、信任不够等问题。平台具有平台化、数字化、智能化和集市化的特征，能有效地拓展供给侧、扩大需求侧、完善管理侧。

针对供给侧，拓展了资金供给侧，将银行、股权投资机构、保险、担保、金融租赁等可为企业提供金融服务的持牌机构纳入，鼓励其通过创新金融产品、开展技术研究等方式满足企业全生命周期的各种融资需求。针对需求侧，助力平台尽可能将各行各业的中小微企业吸纳到平台上，通过大数据、人工智能的手段降低企业进入的门槛，惠及传统金融机构所忽略的企业。针对管理侧，以“店小二”服务精神，主动协助平台帮助政府机关打造主动型服务模式，实现对中小企业应帮尽帮、因企施策、按需施策。不仅是企业单向去找银行、也推动银行找企业、政府帮企业，供需服三方无障碍流动，交叉对接，实现风险共管。

利用大数据技术和风控能力高效处理线上业务并能有效控制风险，保证疫情期间全流程数字化和线上化办公，即使企业业务陷入停滞，企业

与金融机构非接触的情况下，也能实现正常放款。金融机构借助金融平台强大的数字分析能力，实现对受疫情较大影响的客户减免滞纳金，延期还款，因地因时制宜地服务好企业。

6.5.3 案例成果/成效

“春雨金服”平台开展以来，已累计帮助企业解决融资问题2450家次、金额112.98亿元。上线以来，唐山企业综合金融服务平台入驻企业达到5197家、融资金额78.26亿元。“春雨金服”平台切实疏通了资金流向实体经济的渠道，提升了企业获得感和满意度，成为唐山落实国家一系列金融助企惠企政策的有效载体。

首先，有效打通了银企资金融通堵点痛点断点，发挥了金融纾困解难作用。综合金融服务平台“六个一批”诉求窗口，实现了企业需求诉求随时反映、随时办理。截止2020年11月23日，超过2100家企业在金服平台申请融资需求，实际融资金额超过311亿元，有效减轻了企业资金周转压力。

其次，切实降低了企业融资成本，极大地提升了金融服务质效。通过量化考核，激发银行发挥各自优势，开辟绿色通道、降低融资成本、缩短审批链条，贷款业务平均审批时间缩短1/3以上，业务线上办理率增长25%。

同时，切实优化了企业贷款供需结构，探索出了破解“短贷长投”问题的有效路径。通过人工智能、大数据等技术的引入有效缓解了企业信息不对称、信用不充分问题，打通供需错配的堵点痛点，提高贷款投放的有效性和供需匹配度，实现稳增长、调结构、防风险的有效平衡。

此外，切实优化了发展环境，打响了唐山市一流营商环境品牌。海洋牧场、华通线缆、唐鸿重工等一大批企业在助企惠企的春雨金服行动中得到了资金支持，叫响了当地金融品牌，提升了地方知名度。

6.6 漳州市数据汇聚共享服务平台建设项目案例

6.6.1 案例背景现状

漳州市政务数据资源共享服务平台项目以“扩大采集、整合汇聚、创新应用、开放开发”为主线，按照《福建省“十三五”数字福建专项规划》的总体思路要求，从顶层设计出发，充分运用大数据先进理念、技术和资源，推动数据资源从“小”向“大”汇聚，从为主支撑信息化应用向作为重要生产要素转变，从自建自用向共享开放转变，激发数据活力潜力。到2020年，政务信息资源汇聚共享和开放开发格局全面形成，重点领域大数据应用全面推广，基本形成大数据应用体系和能力，有效推动产业转型升级、完善社会治理、提升政府服务和监管能力。

漳州市的政务机房群实现了市级重点政务数据和业务系统的统一承载，但各种业务数据依然“割据”在每个政务业务系统之中，数据可用性和数据质量无法带来良好的用户体验，数据除了满足基本的业务用途之外，无法实现基于数据的创新和数据增值，更难以适应省委省政府开放开发政务数据带动产业发展的目标。根据国家及省市信息化部署建设的要求，组织全面梳理政务数据资源，编制政务数据资源清单和开放内容，集中汇聚数据，制定开放计划，建设市级大数据服务公共平台，向社会开放非密、脱敏政务数据，以特许经营、购买服务等政府和社会资本合作模式推进政务数据增值开发相关任务，提出建设“漳州市数据汇聚共享服务平台”项目。

漳州市政府充分认识到政务数据对于深化“数字漳州”建设、发展智慧城市建设的重大意义。因此，建设漳州市数据汇聚共享服务平台，实现漳州市政务部门数据汇聚和开放管理具有迫切的需求。

6.6.2 案例内容

漳州市数据汇聚共享服务平台项目分为一期和二期，一期和二期项目

建设内容整合如下：

（一）平台功能建设

平台功能建设主要包括平台软件开发、数据库建设两大部分。建设完成数据共享门户、共享服务网关、资源管理系统、资源图谱系统、供需管理系统、运维管理系统、运营管理系统、系统配置中心、通用服务系统、资源目录管理系统、数据交换系统、汇聚各业务部门的共享数据资源，基于API数据服务模式实现数据服务能力共享、建设完成服务平台、资源活化系统、数据治理平台、数据服务系统（ESB）、报表及可视化服务系统、资源监控系统、数据交换节点，实现对数据汇聚共享过程统一监控、汇聚共享数据统一管理、汇聚共享成果统一展现，为全市各部门、各类综合业务系统提供数据支撑。

通过数据汇聚共享服务平台，搭建漳州市人口、法人基础数据平台，为基础数据的共享与应用提供信息基础，并在此基础上，进行专题应用的建设。

（二）标准规范体系建设

标准规范体系建设，以国家相关标准和福建省已有的相关标准为基础，建立规范、完整的标准规范和管理制度体系，满足项目建设对标准规范和管理制度的需求。包括数据汇聚共享服务平台数据管理机制、数据调用监管办法、数据汇聚共享服务平台运行管理制度、市级目录编制标准、数据管理和共享规范、数据接入规范、平台运维规范等。

（三）平台部署实施工作

1.信息资源目录梳理

梳理各部门信息资源目录，确定各部分要汇聚的数据目录，根据确定的汇聚目录，进行各部门的信息资源目录加载和日常维护。

2.数据资源接入汇聚

根据共享资源对接计划，将市级生产单位共享资源接入市级汇聚共享

平台。对于国家级、省级系统的共享资源，并通过国家、省级汇聚共享平台逐层向下接入汇聚。

3.应用需求调研

基于委办局（含生产单位及需求单位）应用数据资源开展梳理实施工作，通过跨部门信息共享工作，打通数据壁垒，解决跨部门系统间数据不通等问题，完成不少于35个市直部门的应用需求调研整理工作。

4.数据资源加工清洗

在确定各数据资源接入的提供方后，对数据进行加工清洗，经过清洗后的数据需要该类信息的提供方进行审查，审核通过后关联入库，作为进一步信息共享的数据基础，要求完成不少于35个市直部门应用数据资源的加工清洗工作。

5.数据资源共享服务

以委办局应用系统或数据用途为维度，生成相应的共享数据接口并完成与各部门系统的对接指导和联调工作，提供数据资源的交换与共享服务。

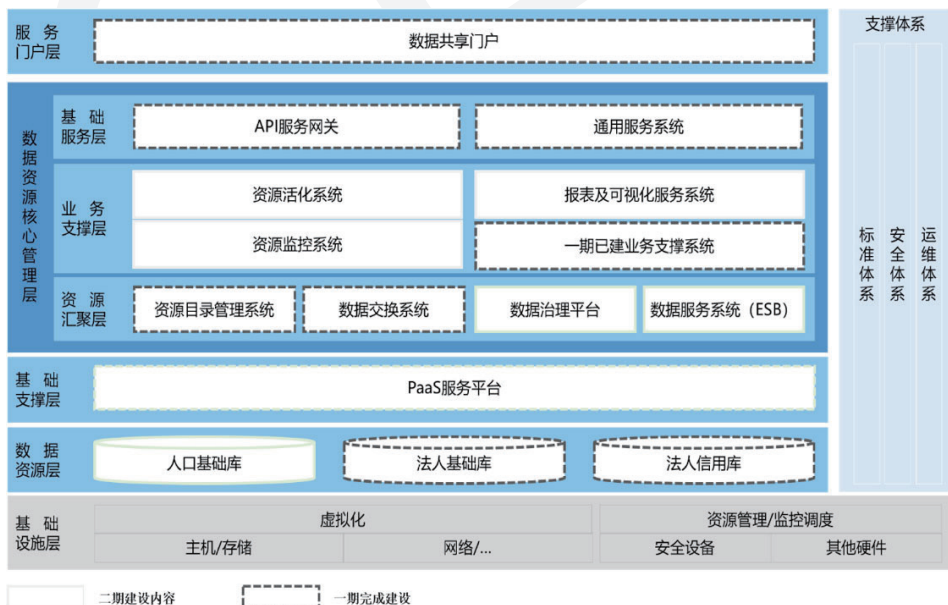


图6-4 漳州市政务数据资源共享服务平台总体架构

6.6.3 案例成果/成效

漳州市数据汇聚共享服务平台建设和应用服务项目自2017项目启动以来，将分散在各个数据中心的“数据碎片”向政务云平台集中汇聚，打破因地域、业务和网络等因素形成的“信息孤岛”现象，基本实现省、市数据互联互通。截止目前，平台已经汇聚了漳州市41个市直部门共计4.3亿条数据，其中涉及公安局、人社局、住建局、民政局等部门，汇聚居民身份证、社会保险缴交信息、婚姻信息、出生医学证明等核心业务数据。

通过对汇聚的人口、法人数据进行融合、重构、云化，建设完成人口数据云和法人数据云，为开展政府个性化主动服务提供技术支撑；通过对汇聚的数据进行归类、分析形成十多个主题数据库，包含交通运输、信用服务、城建住房、生态环境、市场监管等主题，且在部分主题下展开了电子证照信息查询、自然人查询、数字证照画像等“微应用”建设，例如：在“市场监管”主题下，建设了“电子证照信息查询”微应用，企业可根据企业名称查询企业的所有证照信息，包含税务登记证、营业执照、组织机构代码证、不动产权证、医疗器械经营许可证等；个人可根据姓名和身份证查询个人的所有证照信息，包含个人身份证照、学历证书、驾驶证等。

通过对汇聚的数据进行校验比对，验证数据的完整性、重复性，从而提升数据入库质量，提升数据共享服务质量。截止至今，平台已经提供了210项数据共享服务，共生成共享接口243个，接口累计调用次数达764万余次，满足了市直部门对身份证信息、社保信息、婚姻信息、企业信息、电子证照信息、审批办件信息等信息资源的需求。例如：市汇聚平台提供给市卫健委的公安个人信息和户籍信息、民政婚姻信息，实现申请对象的身份信息、婚姻信息、生育信息的自动比对，使得市民可以直接在市卫健委的“漳州市卫生健康委”官方微信平台提交“生育登记办理”，免去窗口材料提交、审核环节；提供给市卫健委的人社局社保卡信息，实现认证用

户与个人社保卡自动关联，市民可以直接用社保卡就医预约挂号。同时，市汇聚平台为“漳州通APP”、“漳州市行政审批统一平台”的部分功能提供了数据支撑，加强政务信息资源跨部门、跨层级互通和协同共享，发挥信息共享支撑多部门协同服务的作用，切实打通政府部门之间的信息壁垒和信息孤岛，确保实现“一趟不用跑”、“最多跑一趟”的政务服务模式。

6.7 贵阳市政务数据共享开放考核实践探索

6.7.1 案例背景现状

政务数据共享开放已成为促进数字经济健康发展，提高政府治理能力和服务水平，激发市场活力和社会创造力的重要手段。然而从国家到地方，对政务数据共享开放缺乏一套成熟可行的考核手段，因此，结合贵阳实际需求，提出“考核体系+考核工具+规章制度”相结合的政务数据共享开放考核新模式，通过构建政务数据共享开放考核指标体系，研发政务数据共享开放考核平台，研究配套工作方案，实现对政务数据共享开放的自动化考核和智能化分析，以评促建，以考核促进政务数据共享开放规范化、制度化、科学化。

6.7.2 案例内容

（一）考核指标体系标准化

科学构建指标体系是开展政务数据共享开放考核工作的前提，本案例根据政务数据共享开放相关政策文件要求，并充分征求政府部门、业内专家、相关企业意见，从基础保障、数据共享、数据开放、数据质量、数据应用、保障机制六方面提出政务数据共享开放考核指标体系（如下表所示）。

表6-2 政务数据共享开放考核指标体系

目标层	一级指标	二级指标
政务数据 共享开放成效	基础保障	信息系统
		数据汇聚
		数据资源目录
	数据共享	有条件共享
		共享规范性
		数据共享度
		共享程度
		重点项目支撑
	数据开放	开放计划
		依申请开放
		数据脱敏
		开放程度
	数据质量	数据完整性
		数据准确性
		数据时效性
		数据可访问性
	数据应用	数据开发利用
		应用成效
保障机制	组织保障	
	制度保障	

（二）考核平台建设

基于涵盖指标管理、算法管理、考核评估、结果分析四大模块的技术架构，研发政务数据共享开放考核平台，提供指标动态管理、实时监测分析、数据流向分析、数据应用分析、考核结果分析等服务，实时呈现政务数据共享交换平台、数据开放平台及考核测评最新资讯，为用户进一步发现共享开放工作短板、制定工作方案提供决策支持。



图6-5 政务数据共享开放考核平台示意

1.指标动态管理是对政务数据共享开放考核指标和标准进行综合管理。鉴于不同地区政府部门在数据共享开放的推进模式、工作重心、发展水平上存在较大差异，政务数据共享开放考核没有完全一致的标准，因此需要根据考核区域、考核对象实际情况建立符合需求的指标体系，并实现对考核指标和标准的动态编辑和综合管理。

2.实时监测分析呈现数据共享开放最新动态资讯。基于政务数据共享交换平台、数据开放平台提供的数据，实现对平台交互情况、使用日志、共享开放数据量等信息的实时监测，展示内容如数据共享开放趋势、数据申请阻塞TOP5、数据申请趋势、热门数据TOP10、共享开放考核指标完成度等。

3.数据流向分析基于部门间数据流通信息，实现对数据流向的挖掘分析和可视化展示。通过对数据流向的跟踪，厘清不同系统间的数据交互关系，梳理各类数据的应用场景和使用情况，利用数据挖掘、网络分析、知识图谱等技术，对数据和应用进行关联分析，挖掘数据应用规律，进而为考核对象智能推荐合适的应用场景，促进数据应用。

4.数据应用分析可实现数据应用成果展示与推介。通过对政府、个

人、企业和社会团体提供的应用开发、应用案例、应用场景数据进行梳理、比较，筛选出在提升政府治理能力、促进创新创业等方面具有良好示范效应的典型案例，并在考核平台进行展示。

5.考核结果分析涵盖自动化考核、考核结果分析、考核报告智能生成与推送等功能。自动化考核，是基于考核指标体系和评分标准，利用API接口自动采集考核数据，实现数据采集、数据处理、指标评分、分值汇总等考核全流程自动化处理；考核结果分析，是以综合评分结果数据为基础，采用比较分析法、分组分析法、回归分析法、因素分析法等数据分析方法，借助图表等表达方式，对评估结果做描述性和推断性统计分析；考核报告智能生成与推送，是基于考核数据及其分析数据，实现考核结果可视化呈现并自动生成考核报告，考核报告包括考核统计结果、智能诊断结果及对策建议等内容，考核报告将通过邮件、用户提醒等方式及时向考核对象推送。

6.7.3 案例成果/成效

建设政务数据共享开放考核平台，实现考核全流程的自动化处理和智能化分析，对于强化政务数据共享开放督导管理、提升考核工作效率、促进数据共享开放和开发利用具有重要意义。鉴于当前国内外主要城市还未建立从政务数据共享开放考核体系设计到考评平台研发的系列方案，本研究将为政务数据共享开放考核工作提供可借鉴的经验，并具有良好的示范带动效应。

6.8 成都市公民大数据平台案例

6.8.1 案例背景现状

为有效促进公民信息数据资源共享利用和跨部门的业务协同，2010年，成都市委、市政府提出建设以公民身份号码为唯一标识，集居住、

婚育、就业、纳税、信用、社会保险等信息于一体的公民大数据平台的要求。成都市政务服务管理和网络理政办公室牵头建设覆盖全市人口的成都市公民大数据平台（一期），该平台建成后，有效提高了成都市各级政府部门管理和服务的效率与质量，为政府宏观决策提供了有力支持。

2018年11月，为贯彻落实《成都市公共数据管理应用规定》，促进政务信息资源整合共享，助推“放管服”改革，成都市启动公民大数据平台（二期）建设，助力提升公民信息管理应用水平。

6.8.2 案例内容

成都市公民大数据平台主要建设内容是构建“五大体系”，形成“一库、一平台、四应用”。“五大体系”指构建公民信息数据采集治理体系、系统支撑体系、政策与标准规范体系、信息安全保障体系、公民信息应用和服务体系。“一库”指建立全市公民信息数据库，“一平台”指建成公民信息数据目录和共享交换平台，“四应用”指建设决策分析支持、跨部门业务协同、公民服务、个人信用等四大类应用。公民大数据平台总体框架图如下。

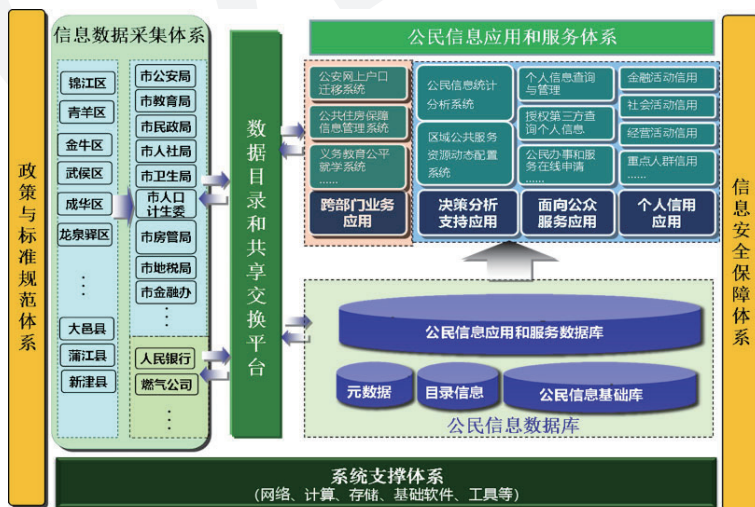


图6-6 公民大数据平台总体框架图

（一）构建标准规范体系

制定公民信息归集、使用、隐私保护等管理办法，建立公民信息数据共享和应用的长效机制建立数据维护更新、错误发现及纠错、目录适时调整机制和安全保密等管理制度，确保数据及时、准确、完整、有效。

标准规范为信息资源一致性和技术平台的互联互通互操作提供了基本的技术保证，成都市公民大数据平台标准规范的研究制定过程经过体系规划、标准研制、试用验证、专家评审、培训、宣传、贯彻、应用示范、完善与实施等阶段，而每个阶段的工作都在统一的标准化工作平台上完成。

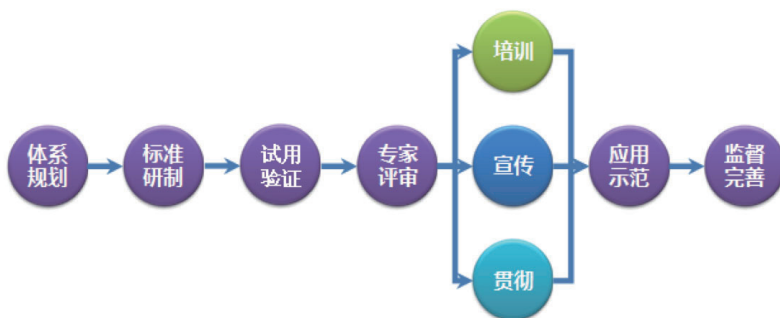


图6-7 标准规范编制过程流程图

（二）构建公民信息数据采集治理体系

1.建立公民信息数据目录和共享交换平台。按照“有标贯标、无标定标”的原则，制定公民信息的数据目录、分类编码、数据归集、数据字典、数据元等规范与标准，建立公民信息数据目录。制定公民信息数据共享交换规范，建立公民信息数据共享交换子平台，提供数据交换、目录管理、认证授权等服务，实现跨部门公民信息数据的共享和交换。

2.建设公民信息数据库。依据公民信息数据目录，以公民身份号码为唯一标识，通过共享交换平台采集城市各政务部门负责管理的公民信息数据；通过数据接口、离线导入等方式采集公共企事业单位涉及的公民信息数据；通过个人申报方式获取公民个人掌握的公民信息数据。经过对数据整理、关联、清洗，分类入库形成集中统一的公民信息数据基础库、公民

信息应用和服务库。

（三）构建平台支撑体系

建设数据存储系统、前置交换系统、数据库系统、中间件系统、地理信息系统等软硬件基础平台。运用数据检索、报表统计、联机分析处理、商务智能等技术和工具，搭建应用和服务支撑平台，提供数据核查、查询、比对、引用、分析、报表、数据模型、空间分析等基础服务，形成安全可靠、功能完备的平台支撑体系。

（四）构建公民信息应用和服务体系

基于公民信息基础数据，重点开发决策分析支持、跨部门业务协同应用、公民服务、个人信用四大类应用和服务。

6.8.3 案例成果/成效

成都市公民信息大数据平台是以公民身份号码为唯一标识的全市人口大数据资源库，现已成为成都市打造新型智慧城市的信息新基建基础设施，具体成果如下：

（一）建设成效

1.数据归集方面

系统归集成都市46个部门和公共企事业单位的公民信息数据，建立了统一的公民信息数据库，解决各部门数据一致性、准确性和完整性的问题。帮助制政府各部门有效避免和减少公民信息重复采集，节约政府行政成本。截止目前，已累计完成266类4042项共计26.8亿条数据的入库工作，覆盖了全域成都3642万人口[包含户籍人员，居（暂）住证人员，涉蓉人员（非户籍、未有办理居住证、但在成都参保），就医，房产交易，投资经营等]。

2.支撑跨部门协同应用服务方面

建立了全市公民数据目录体系和共享交换平台。支持公民信息资源共享、全域协同，为各级、各类政务部门行政审批、公共服务和社会管理提

供跨部门数据支撑。现已为22家市、区县政府部门提供了跨部门数据应用服务和5家企业及金融机构的普惠金融等公共数据应用业务提供了1033余万次脱敏信息查询服务，累计提供服务2.03亿次。

针对群众反映强烈的“奇葩证明”等问题，公民信息大数据平台大力推进政府内部数据共享，变“群众跑腿”为“数据跑路”，公安户籍信息、人社低保审核、社保管理、公共住房保障、公积金综合业务、人口计生服务等直接关系民生的系统通过公民系统实现了业务协同办理，为服务流程再造、窗口快捷办事提供了数据支撑，让群众切实感受到更多的便利和“获得感”。

3.数据决策支持方面

帮助政府建立了全市人口数据可视化及管理看板，实现公共资源的配置分析，帮助政府进行宏观决策分析。现已开发了“劳有应获”、“老友颐养”、“学有良教”、“住有安居”、“人有普保”五大民生主题分析，形成了《农业转移人口市民化分析专题报告》《成都义务教育学校资源配置分析》《成都市精准扶贫情况分析研究》等20余份公民信息资源应用研究报告。

4.公民信用应用方面

整合公民个人参与社会活动、经营活动、金融活动等的信用信息，目前已探索公民信用信息在政务领域和普惠金融方面的应用，下一步计划深化公民信用信息在政务领域应用，并拓展至社会各个领域，逐步形成公民信用数据清单、行为清单、应用清单，推动形成事前告知承诺、事中评估分类、事后联动奖惩的公民信用应用机制，让公民信用无时不有、无处不用，为社会诚信体系建设提供有力的支撑。

（二）社会价值

成都市公民信息大数据平台可以为城市管理者提供决策依据，为政府部门提供数据共享和协同平台，并为广大公民提供随时、随地、随需的个

人信息服务。既可以提升政府管理水平，又可以提高部门办事效率，推进服务型政府发展进程，社会效益显著。

一是提高行政效能。通过政府部门间数据共享利用和业务协同，提高部门工作效率和质量，减少老百姓办事所需材料、时间和成本，提升政府社会管理和公共服务水平。

二是支持科学决策。综合利用多部门集中的公民信息大数据，为政府及各部门掌握全域公民信息的多维度特征及变化趋势，充分发挥综合信息系统检索历史、感知现在、预测未来的功能，促进教育、医疗、文化等公共服务资源优化配置，及时做出科学决策提供数据支撑。

三是加强社会管理。整合公民社会活动、经营活动、金融活动等信息，建设公民社会信用信息系统，构建公民个人社会信用体系，推动社会诚信体系建设，大力促进社会管理水平提升。

四是全面增加民生福祉。通过面向公民信息大数据平台的建设，有利于促成政府、市场、社会多元供给的主体格局，提高公共服务供给质量和效率。提升教育、就业、住房、公共交通、医疗卫生服务、收入和社会保障、人口管理与服务优化、养老服务建设、残疾人福利、文化、体育水平等领域的公共服务均等化、精准化、优质化水平，更好地满足人民群众对优质公共服务的需求。



全国信标委大数据标准工作组

联系方式: 010-64102858