

政务数据治理 流程规范

Process specification for government data governance

(报批稿)

XXXX - XX - XX 发布

XXXX - XX - XX 实施

目 次

| | |
|-----------------------|----|
| 前言 | II |
| 1 范围 | 3 |
| 2 规范性引用文件 | 3 |
| 3 术语和定义 | 3 |
| 4 缩略语 | 3 |
| 5 流程框架 | 3 |
| 6 需求导入 | 4 |
| 6.1 需求调研 | 4 |
| 6.2 需求分析 | 5 |
| 6.3 需求设计 | 5 |
| 6.4 需求变更 | 6 |
| 7 服务实施 | 6 |
| 7.1 数据处理 | 6 |
| 7.2 数据资产管理 | 7 |
| 7.3 项目管理 | 7 |
| 8 考核评价 | 8 |
| 8.1 需求导入考核 | 8 |
| 8.2 服务实施考核 | 8 |
| 附录 A（规范性） 业务调研表 | 9 |
| 附录 B（规范性） 数据调研表 | 10 |
| 参考文献 | 12 |

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由智慧城市产业生态圈提出并归口。

本文件起草单位：深圳市华傲数据技术有限公司、深圳市标准技术研究院、山东省计算中心（国家超级计算济南中心）、中国联通智能城市研究院、深圳市敢为软件技术有限公司、西安电子科技大学。

本文件主要起草人：黎俊茂、胡露、史丛丛、杨叶、梁芳、杨奎、张涛。

版权声明

本标准版权属于智慧城市产业生态圈（SCIE）所有，并受法律保护。转载、摘编或以其它任何方式使用本标准的文字或者观点的，应注明来源。违反上述声明者，著作权方将追究其相关法律责任。

政务数据治理 流程规范

1 范围

本文件提出了政务数据治理的流程框架，规定了政务数据治理中的需求导入、服务实施、考核评价等过程的要求。

本文件适用于政务数据治理的实施与评价，其他行业的数据治理可参照使用。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 34960.5—2018 信息技术服务 治理 第5部分：数据治理规范

3 术语和定义

GB/T 34960.5—2018界定的以及下列术语和定义适用于本文件。

3.1

政务数据 government data

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。

[来源：GB/T 38664.1—2020, 3.1]

3.2

数据治理 data governance

数据资源及其应用过程中相关管控活动、绩效和风险管理的集合。

[来源：GB/T 34960.5—2018, 3.1]

3.3

数据融合 data fusion

集成多个数据源以产生比任何单独的数据源更有价值信息的过程。

[来源：GB/T 36625.1—2018, 3.1]

3.4

数据原子化 data atomization

将数据拆分至最小原子的过程，每个原子只包含实体的某个或某类属性，以方便对数据进行灵活组织。

4 缩略语

下列缩略语适用于本文件。

E-R: 实体-关系图(Entity Relationship Diagram)

SQL: 结构化查询语言(Structured Query Language)

5 流程框架

政务数据治理流程包含需求导入、服务实施和考核评价三部分，具体框架见图1。各部分之间的关系如下：

- a) 需求导入是政务数据治理的源头输入，通过需求调研、需求分析和需求设计，确定数据来源和业务需求，为政务数据治理的服务实施提供数据准备；

- b) 服务实施是政务数据治理的实施阶段，通过数据标准化、数据原子化和数据融合等，对归集的政务数据进行加工处理，形成数据资产并加以管理；
- c) 考核评价对需求导入和服务实施两大过程中的关键节点进行检查与考核，通过制度化的考核评价体系，保障政务数据治理的效果。

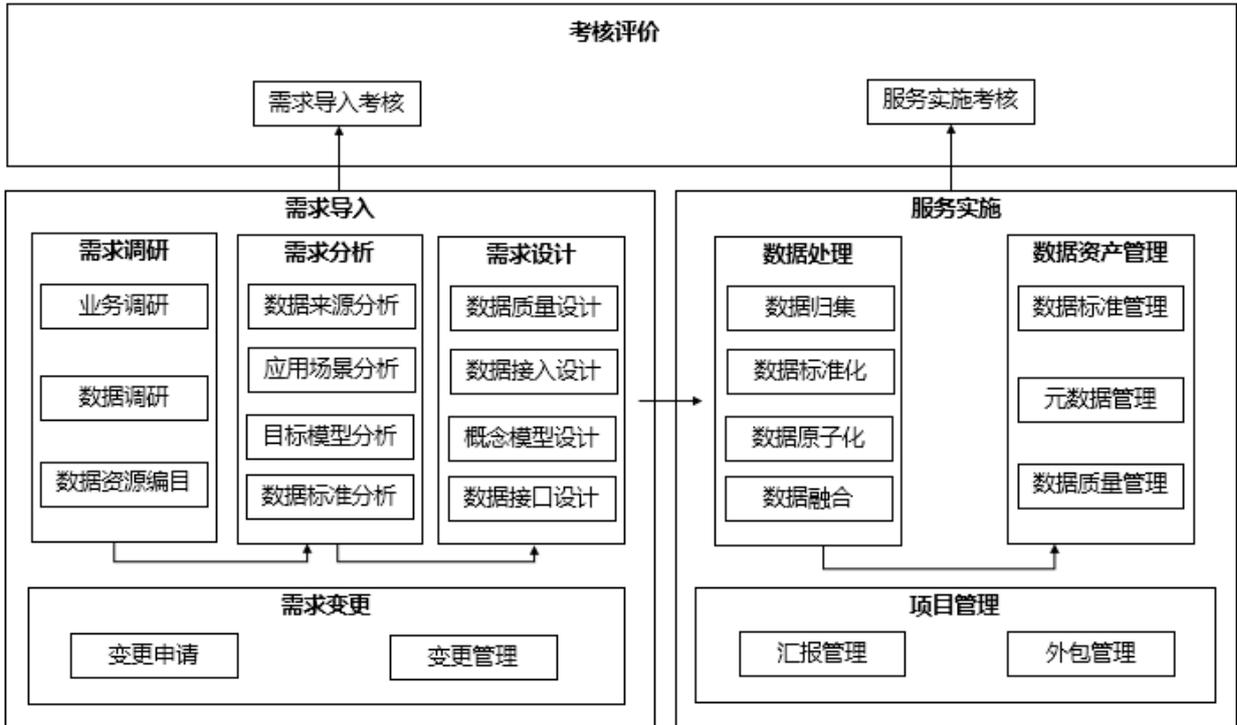


图1 政务数据治理流程框架图

6 需求导入

6.1 需求调研

6.1.1 概述

需求调研包括需求调研步骤、需求调研内容两部分。

6.1.2 需求调研步骤

需求调研一般遵循下列步骤：

- 立项阶段：梳理调研目标，组建调研团队；
- 准备阶段：获取业务部门的“三定”方案、权力清单、责任清单等政策文件，并进行预分析，协调调研发文，制定并确定调研实施方案；
- 实施阶段：按照调研实施方案，通过问卷、电话、面谈等形式开展调研，循环实施调研并执行检查单，定期通报进度及成果；
- 整理阶段：整理输出调研成果，包括数据资源报告、数据需求汇总及安全需求汇总，并组织经验总结和知识归档。

注：“三定”是指“定机构、定职能、定编制”。

6.1.3 需求调研内容

6.1.3.1 业务调研

业务调研应系统梳理业务部门的业务现状，包括但不限于：

- 部门情况：梳理业务部门的组织架构、职责分工、业务开展和各业务处室关联情况等，整理出“三定”方案、权力清单、责任清单和相关规章文件；

- b) 系统情况：梳理业务部门的信息系统及应用情况，包括用户、建设厂商、建设年份、运维厂商、项目文档、使用频率等；
- c) 业务流程情况：梳理业务部门的业务流程，包括各项流程的业务分工、业务环节和产生的业务信息等。

6.1.3.2 数据调研

数据调研应系统梳理业务部门的数据现状，包括但不限于：

- a) 数据情况：梳理业务部门的现有数据情况，包括数据项名称、数据格式、数据类型、值域等；
- b) 数据标准情况：梳理业务部门现有数据对各类标准（国家标准、行业标准和地方标准等）的采用和需求情况；
- c) 数据管理情况：梳理业务部门的数据规范化、数据清洗比对、数据校验等数据管理情况；
- d) 数据应用情况：梳理业务部门的数据统计分析、数据服务调用、数据可视化等数据应用情况。

6.1.3.3 数据资源编目

应根据业务调研和数据调研的结果进行数据资源编目，数据资源编目一般遵循下列步骤：

- a) 编目准备：准备业务部门清单、业务部门系统清单、库表资源清单等；
- b) 数据资源编目：通过结构化查询语言（SQL）脚本、手工、库表资源等方式进行数据资源编目；
- c) 数据资源目录管理：通过数据资源目录管理工具对目录实现监控、变更、上报、导出等管理。

6.2 需求分析

6.2.1 概述

需求分析包括数据来源分析、应用场景分析、目标模型分析和数据标准分析四部分。

6.2.2 数据来源分析

应分析业务部门的数据来源情况，梳理数据来源对应的业务部门和业务系统。

6.2.3 应用场景分析

应分析业务部门的应用场景，根据不同的应用场景梳理对应的事项。政务数据的应用场景和梳理事项包括：

- a) 分析型应用场景：梳理分析主题、分析维度、指标、字段需求等；
- b) 操作型应用场景：梳理应用系统、权责事项、填报字段、证明字段、审批字段等；
- c) 信用型应用场景：梳理核对业务部门、核对事项、核对字段等。

6.2.4 目标模型分析

根据业务部门的应用场景，应进行业务抽象和数据抽象，建立业务模型和数据模型。

6.2.5 数据标准分析

应分析业务部门的核心数据对数据标准的采用和需求等情况。

6.3 需求设计

6.3.1 概述

需求设计包括数据质量设计、数据接入设计、概念模型设计和数据接口设计四部分。

6.3.2 数据质量设计

应根据业务部门的数据质量情况，设计质量等级。

6.3.3 数据接入设计

应根据应用场景梳理数据接入场景，确定结构化数据、半结构化数据、非结构化数据等各类型数据的接入方式和接入流程。

6.3.4 概念模型设计

应在需求分析结果的基础上，对数据进行抽象处理，并采用实体-联系图（E-R）等方法进行概念模型设计。

6.3.5 数据接口设计

应对输入接口类型、数据库类型、功能描述、加工字段、输入参数、输出参数、数据样例、测试用例等进行数据接口设计。

6.4 需求变更

6.4.1 概述

需求变更包括变更申请和变更审核两部分。

6.4.2 变更申请

应制定需求变更的申请流程，申请流程至少应包含提交申请单、内部审核和外部审核。

6.4.3 变更管理

应制定需求变更的管理办法，管理办法至少应明确变更申请、变更审核和变更执行过程中的输入、输出和执行部门。

7 服务实施

7.1 数据处理

7.1.1 概述

数据处理包括数据归集、数据标准化、数据原子化和数据融合四部分。

7.1.2 数据归集

数据归集可分为数据采集和数据存储两部分。数据归集应包括下列操作：

- a) 建立数据归集模型；
- b) 执行数据归集统计分析；
- c) 输出数据归集报告；
- d) 执行数据质量稽核；
- e) 执行数据安全审计。

7.1.3 数据标准化

应对已归集的政务数据进行标准符合性测试，输出标准符合性测试报告，依据报告并参考相关数据标准对数据进行标准转换。需进行标准转换的数据问题主要有：

- a) 字段属性不一致；
- b) 值域字典不一致；
- c) 业务含义不一致；
- d) 字段规则不一致。

7.1.4 数据原子化

应对经过标准转换的数据进行原子化处理，实现多源异构数据向同一个模型进行转换，数据原子化应遵循以下步骤：

- a) 建立领域目标模型；
- b) 配置源模型与领域目标模型的字段映射转换关系；
- c) 源模型与领域目标模型转换。

7.1.5 数据融合

对经过标准转换和原子化处理的数据进行整合，包括多源实体对象的整合、多源行为对象的整合、多源属性状态的整合等。可针对具体字段配置不同的数据融合策略，如“一数一源”策略，大多数原则策略、新鲜度原则策略等。

注1：“一数一源”策略，同一数据来自多个部门时，以权威部门来源为准，如“姓名”以公安部门为准，“婚姻状态”以民政部门为准。

注2：大多数原则策略，同一数据存在多条不同记录时，以大多数记录为准，如某人的“身高”数据，三条记录为1.7米，一条记录为1米，取1.7米为其身高信息。

注3：新鲜度原则策略，同一数据存在不同的时间属性时，以时间最近的数据为准，如某人的“学历”信息，两年前为“硕士”，两年后为“博士”，取“博士”为其学历信息。

7.2 数据资产管理

7.2.1 概述

数据资产管理包括数据标准管理、元数据管理和数据质量管理三部分。

7.2.2 数据标准管理

数据标准管理服务至少应包括标准文档管理、标准查阅服务、标准地图服务等：

- a) 标准文档管理：提供数据标准文档、标准数据元、标准代码、标准术语等管理；
- b) 标准查询服务：支持以数据元名称为查询条件，展示与此数据元相关的标准文档、规则、表、数据元版本、变更等信息；
- c) 标准地图服务：应能通过数据标准地图了解当前数据标准的覆盖情况，如哪些表/字段是否使用了标准，使用哪些标准，哪些数据未被标准覆盖等。

7.2.3 元数据管理

元数据管理服务应包括元数据采集、元数据分析、元数据统计等服务：

- a) 元数据采集：通过自动采集、手动采集等方式满足元数据采集需求；
- b) 元数据分析：满足元数据映射管理、元数据地图展示、元数据血缘分析、元数据影响分析等服务需求；
- c) 元数据统计：满足元数据查询和不同维度的元数据统计需求。

7.2.4 数据质量管理

数据质量管理服务应实现对不符合规则的数据进行探查、发现、修复、报告的闭环管理，至少应包括数据质量探查、质量规则管理、质量问题管理、质量报告管理等：

- a) 数据质量探查：通过列分析、比对分析等手段，对数据基础质量进行探查，发现数据的一致性、完整性、重复性等质量问题；
- b) 质量规则管理：提供数据一致性、规范性、正确性、关联性、及时性、完整性、重复性等数据质量度量指标的管理；
- c) 质量问题管理：对问题数据进行分析、修正、跟踪、统计等，提供对数据质量的全生命周期管理，支撑数据管理部门与数据责任部门协同对数据质量问题进行修复；
- d) 质量报告管理：提供不同场景下的数据质量报告生成和发布，以帮助用户理解数据质量详细情况。

7.3 项目管理

7.3.1 概述

项目管理包括报告管理和外包管理两部分。

7.3.2 报告管理

应制定项目定期报告管理制度，并根据项目交付的里程碑节点和时间计划开展汇报，项目汇报的信息应至少包括汇报时间、参与方、汇报主题、汇报材料、会议纪要以及汇报后的任务事项等。项目管理单位应将报告成册留档。

7.3.3 外包管理

当涉及第三方外包团队参与服务实施工作，如大规模调研、人工维护数据标签、人工录入文件信息资源等，应制定外包管理制度，外包管理制度应至少明确外包方的参与人员、参与时间、责任事项、外包方培训方案、外包方成果检查方案等。

8 考核评价

8.1 需求导入考核

应结合政务数据治理项目情况与业务要求，从需求调研、需求分析、需求设计和需求变更四个方面，制定需求导入规范化评价方案，对数据治理项目的需求整理工作进行考核。需求导入考核评价的要点如下：

- a) 业务调研情况检查，业务调研表应按照附录 A 的要求填写；
- b) 数据调研情况检查，数据调研表应按照附录 B 的要求填写；
- c) 数据资源编目情况检查，数据资源编目可参照《政务信息资源目录编制指南（试行）》的要求执行。

8.2 服务实施考核

应结合数据治理项目情况与业务要求，从数据归集、数据处理、数据资产管理、项目管理四个方面，制定服务实施规范化评价方案，对数据治理项目的服务实施工作进行考核。服务实施考核评价的要点如下：

- a) 数据归集统计检查；
- b) 数据模型设计检查；
- c) 标准符合性测试检查；
- d) 数据质量稽核检查；
- e) 数据血缘分析检查；
- f) 数据安全审计检查。

附 录 A
(规范性)
业务调研表

表A.1给出了业务调研相关信息。

表A.1 业务调研表

| | | | |
|-----------------------|---|----------|--|
| 单位（必填） | | | |
| 联系人（必填） | | 联系电话（必填） | |
| 业务情况（必填） | 权力事项： | | |
| | 责任事项： | | |
| 系统情况（有系统必填） | 系统用户： | | |
| | 系统承建方： | 建设年份： | |
| | 运维厂商： | 联系方式： | |
| | 项目文档： <input type="checkbox"/> 有 <input type="checkbox"/> 无 | | |
| 业务流程情况（必填） | 填报事项： | | |
| | 证明事项： | | |
| | 核对事项： | | |
| | 审批事项 | | |
| | 其他事项： | | |
| 备注： | | | |
| 填报说明： 1、每个单位填报一张表。 | | | |

填报人：

年 月 日

附 录 B
(规范性)
数据调研表

表B.1给出了业务调研相关信息。

表B.1 数据调研表

| | | | |
|----------------------|--|----------|--|
| 单位（必填） | | | |
| 联系人（必填） | | 联系电话（必填） | |
| 业务系统（必填） | <input type="checkbox"/> 有（系统名称_____） <input type="checkbox"/> 无 | | |
| 系统归属（有系统必填） | <input type="checkbox"/> 区级 <input type="checkbox"/> 市级 <input type="checkbox"/> 省级 <input type="checkbox"/> 国家级 | | |
| 系统使用情况（有系统必填） | <input type="checkbox"/> 在用 <input type="checkbox"/> 废弃 <input type="checkbox"/> 其它：_____ | | |
| 系统承建方（有系统必填） | | | |
| 系统承建方联系人和联系方式（有系统必填） | | | |
| 数据项名称（表名、必填）及描述 | | | |
| 数据存储方式（必填） | <input type="checkbox"/> Ms SQL Server <input type="checkbox"/> DB2 <input type="checkbox"/> Oracle <input type="checkbox"/> Mysql <input type="checkbox"/> excel 文档 <input type="checkbox"/> 纸质文档 <input type="checkbox"/> 其他_____ | | |
| 当前数据量（条、存储容量、必填） | | | |
| 当前数据是否共享（必填） | <input type="checkbox"/> 已共享 <input type="checkbox"/> 未共享 | | |
| 当前数据提供方式（必填） | <input type="checkbox"/> 导出 Excel 文件 <input type="checkbox"/> 前置机数据库对接 <input type="checkbox"/> 其他（如接口等）_____ | | |
| 当前前置机数据库 | <input type="checkbox"/> Ms SQL Server <input type="checkbox"/> DB2 <input type="checkbox"/> Oracle <input type="checkbox"/> Mysql <input type="checkbox"/> 其他_____（如其它数据库、接口地址、文件服务器地址等）_____ | | |
| 当前数据更新频率 | <input type="checkbox"/> 实时 <input type="checkbox"/> 每天 <input type="checkbox"/> 每周 <input type="checkbox"/> 每月 <input type="checkbox"/> 每季度 <input type="checkbox"/> 每年 <input type="checkbox"/> 其它_____ | | |
| 当前共享使用单位 | | | |
| 未来数据共享类型（必填） | <input type="checkbox"/> 无条件共享 <input type="checkbox"/> 有条件共享 <input type="checkbox"/> 不予共享（原因：_____） | | |

表B.2 (续)

| | | | |
|---|--|----------------|---------------------|
| 未来数据共享方式 (必填) | <input type="checkbox"/> Ms SQL Server 前置机 <input type="checkbox"/> DB2 前置机 <input type="checkbox"/> Oracle 前置机 <input type="checkbox"/> Mysql 前置机 <input type="checkbox"/> 其他 (如其它数据库、接口地址、文件服务器地址等) | | |
| 数据字典 (必填) | <input type="checkbox"/> 有 (表名 _____) <input type="checkbox"/> 无 <input type="checkbox"/> 其它 _____ | | |
| 数据标准 (必填) | 标准采用: | | |
| | 标准需求: | | |
| 数据管理 (必填) | | | |
| 数据应用 (必填) | | | |
| 需要其他单位提供哪些数据 | | | |
| 提供数据的单位 | 数据项名称 | 数据项内容 (字 段) | 该数据在您单位工作中的应用场 景 |
| | | | |
| | | | |
| 备注: | | | |
| 填报说明: 1、每项数据填报一张表。 2、数据为纸质文档应提交相应文档或照片。 | | | |

填报人:

年 月 日

参 考 文 献

- [1] GB/T 34960.5—2018 信息技术服务 治理 第5部分：数据治理规范
 - [2] GB/T 36625.1—2018 智慧城市 数据融合 第1部分：概念模型
 - [3] 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
 - [4] 《政务信息资源目录编制指南（试行）》（发改高技〔2017〕1272号）
-