

DB52

贵州省地方标准

DB52/T 1541.10—2025

政务数据平台

第10部分：主题数据库建设基本要求

Government data platform—

Part10:general requirements for subject database construction

2025 - 03 - 06 发布

2025 - 06 - 01 实施

贵州省市场监督管理局 发布

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 原则	2
5 框架	3
6 要求	4
附录 A（规范性） 主题数据库建设实施流程	9
附录 B（规范性） 主题数据库命名设计	10
附录 C（资料性） 系统表创建相关示例	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。

本文件是DB52/T 1541的第10部分。DB52/T 1541已经发布了以下部分：

- 第 1 部分：建设指南；
- 第 2 部分：数据归集规范；
- 第 3 部分：数据存储规范；
- 第 4 部分：接口规范；
- 第 5 部分：安全技术规范；
- 第 6 部分：面向全网搜索应用的数据处理规范；
- 第 7 部分：运维管理规范；
- 第 8 部分：运营机构服务规范；
- 第 9 部分：数据分发规范。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由贵州省大数据发展管理局提出并归口。

本文件起草单位：贵州中软云上数据技术服务有限公司。

本文件主要起草人：朱永龙、关艳梅、韦超、杨建雄、罗森、钟恺馨、骆进、任凯、顾怀要、秦晓东、杨建国、郑如顺、常丹丹、王班凤、郑洁。

引 言

政务数据平台是承载、治理和利用贵州政务数据的核心载体，是推动贵州一体化政务大数据体系建设的重要基础设施，其建设发展对提升公共服务和社会治理水平、助推政府数字化转型、激活数据要素价值具有重要意义，是贵州数字政府建设的重要组成部分。

为深化政务数据平台建设，贵州省出台了系列标准，以加强大数据创新成果转化，助推政务数据资源体系规范构建。DB52/T 1541系列作为贵州省政务云标准体系的重要组成部分，旨在进一步规范政务数据的采集、存储、整合、共享开放和开发利用等环节有关技术和管理服务要求，拟由10个部分组成，以后根据工作需要，再视情增补。

- 第1部分：建设指南。目的在于为政务数据平台建设提供原则性指导。
- 第2部分：数据归集规范。目的在于为政务数据归集确立相关技术要求。
- 第3部分：数据存储规范。目的在于为政务数据存储提供技术性指导。
- 第4部分：接口规范。目的在于为政务数据平台数据接口确立相关技术要求。
- 第5部分：安全技术规范。目的在于为政务数据平台安全的规划、建设和管理提供原则性指导。
- 第6部分：面向全网搜索应用的数据处理规范。目的在于为政务数据平台全网搜索应用的数据处理活动确立相关技术要求。
- 第7部分：运维管理规范。目的在于指导政务数据平台运维管理工作。
- 第8部分：运营机构服务规范。目的在于为运营服务提供机构的选择评估、服务管理水平的评价和运营机构的自我改进提供原则性指导。
- 第9部分：数据分发规范。目的在于为政务数据平台的数据分发提供技术性指导。
- 第10部分：主题数据库建设基本要求。目的在于为政务数据平台的主题数据库建设提出相关指导和要求。

通过明确政务数据平台主题数据库建设的有关原则和要求，有助于推动各类主题领域数据资源在业务、数据、技术层面实现规范化的汇聚、治理、共享与应用，加快具有贵州特色的政务数据资源体系构建，提升政务大数据应用能力和公共服务水平。

政务数据平台

第 10 部分：主题数据库建设基本要求

1 范围

本文件规定了政务数据平台主题数据库建设的术语和定义、原则、框架和要求。

本文件适用于政务数据平台主题数据库建设的规划、实施、管理与应用。

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 22239—2019 信息安全技术 网络安全等级保护基本要求

DB52/T 1540.2 政务数据 第2部分：元数据管理规范

DB52/T 1540.3 政务数据 第3部分：数据清洗加工规范

DB52/T 1540.4 政务数据 第4部分：数据质量评估规范

DB52/T 1541.2 政务数据平台 第2部分：数据归集规范

DB52/T 1541.4 政务数据平台 第4部分：接口规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

政务数据平台 government data platform

依托政务云计算平台搭建，承载政务数据并支撑数据统筹存储、统筹治理、统筹应用和统筹安全的基础信息技术平台。

[来源：DB52/T 1541.3—2020, 3.1]

3.2

主题数据库 subject database

针对特定应用主题将存在内在逻辑关联关系的各类数据有序汇聚形成的指定数据库。

[来源：DB52/T 1541.6—2021, 3.8, 有修改]

3.3

数据共享交换平台 government data exchange platform

汇聚发布政务数据资源目录，整合共享数据资源，统一为政务部门提供查询、校验、批量交换等数据共享服务，支持跨地区、跨层级、跨部门的数据资源共享交换与业务协同的基础信息技术平台。

[来源：DB52/T 1540.1—2021, 3.3.5, 有修改]

3.4

贴源层 operational data store

存储由源系统采集而来、未经处理的原始数据。

[来源：DB52/T 1541.2—2020]

3.5

标准层 standard data model

存储经过清洗加工后符合质量要求的可复用的标准化数据。

3.6

主题层 data mart

存储整合后面向业务、应用的数据集合。

[来源：DB52/T 1541.2—2020]

3.7

结构化数据 structured data

一种数据表示形式，按此种形式，由数据元素汇集而成的每个记录的结构都是一致的并且可以使用关系模型予以有效描述。

[来源：GB/T 35295—2017，2.2.13]

3.8

非结构化数据 unstructured data

不具有预定模型或未以预定方式组织的数据。

[来源：GB/T 35295—2017，2.1.25]

3.9

元数据 metadata

关于数据或数据元素的数据（可能包括其数据描述），以及关于数据拥有权、存取路径、访问权和数据易变性的数据。

[来源：GB/T 35295—2017，2.2.7]

4 原则

主题数据库建设遵循以下原则：

- a) 依托统一载体开展建设：宜部署在政务外网并依托政务数据平台进行集约化建设，按不同主题域进行分类整合与归集，基于治理后的数据通过数据共享交换平台提供数据服务与应用；

- b) 应用场景设计明晰：宜围绕经济社会发展的某一主题领域，由来源于同一组织或不同组织的关联业务数据围绕共享应用的重构逻辑进行数据的归集、治理和整合应用，业务需求驱动特点明显，数据应用场景设计宜具有高效、实用、灵活、可持续扩展等特点；
- c) 结合业务和技术需求选型：数据库产品服务选型宜考虑技术适配性、技术栈兼容性和性能效率等特性要求，支持云原生和分布式架构，结构设计宜易于维护和扩展，所选数据库能够支持预期工作负载和数据类型，并兼顾考量未来发展创新的路线及愿景；
- d) 强化数据治理和质量管理：宜利用数据治理技术进行数据采集汇聚和标准化处理，避免数据缺失和遗漏。不能虚构或篡改数据，治理过程中注重数据质量管理，确保数据完整、真实、一致；
- e) 保障安全性与合规性：实施适当的数据访问控制和身份验证措施，避免数据被未授权访问。

5 框架

5.1 主题数据库建设包括需求分析、数据目录编制、数据库设计与创建、数据治理、数据服务与应用5大核心环节和运维管理、安全保障2大体系。主题数据库建设总体框架图见图1。



图1 主题数据库建设总体框架

5.2 主题数据库建设包括以下内容：

- 需求分析：包括应用、数据、管理、安全和功能性能等需求维度分析，为后续设计实施提供基础，确保主题数据库建设设计贴合业务需求；
- 数据目录编制：包括数据目录编制、目录审查、目录发布、共享申请等，为主题数据归集、治理和建设管理提供依据，确保后续各参建单位对主题业务应用理解和协作一致；
- 数据库设计与创建：包括数据库设计、数据库选型和数据库创建等流程，为主题数据库成功创建提供支撑，确定主题数据库建设产出的载体形态；
- 数据治理：包括数据采集、数据入库、数据清洗加工、数据整合、数据接口、数据质量等环节，为提升数据质量和数据管理水平提供保障；
- 数据服务与应用：包括运行监控、查询校验、标签服务、指标统计和个性化场景服务等方式，确保主题数据库建设目标和成效的实现，助推数据价值发挥；
- 运维管理体系：为主题数据库稳定高效运行并保持业务连续性提供保障和支撑；
- 安全保障体系：为主题数据库稳定运行提供必要的安全屏障和防线。

5.3 主题数据库建设的实施流程见附录A。

6 要求

6.1 需求分析

6.1.1 应用需求分析

- 6.1.1.1 明确主题领域的行业管理单位和关联单位，宜由行业管理单位作为主题数据库的建设单位。
- 6.1.1.2 分析应用目标时应对各单位的主题应用需求进行归类整合，明确围绕建设目标涉及的业务范围边界。
- 6.1.1.3 结合应用目标对主题数据库建设涉及的业务对象、对象分类、对象间关系以及与其他关联的数据资源库关系进行初步分析。

6.1.2 数据需求分析

- 6.1.2.1 在规定业务范围内对主题业务关联的数据需求进行梳理，形成初步数据需求清单。
- 6.1.2.2 围绕初步数据需求清单，分析数据的来源单位。来源单位数据应是各业务部门在依法履职过程中生成、采集的原始或经初始化处理的数据资源。
- 6.1.2.3 大数据主管单位宜基于数据共享交换平台共享数据资源目录提供必要的共享协调。

6.1.3 管理需求分析

- 6.1.3.1 对主题数据库建设各阶段涉及的建设单位和关联建设单位管理职责分工进行梳理。
- 6.1.3.2 对主题数据库的数据采集管理、归集管理、共享管理、应用管理需求进行规范，规划配套相应的管理规范。

6.1.4 安全需求分析

- 6.1.4.1 基于政务数据平台的安全保障体系，结合应用、数据、管理需求按照“同步规划、同步建设、同步使用”原则进行强化安全防护的需求分析。
- 6.1.4.2 结合政务数据分类分级有关要求，主题数据的安全级别不宜低于一般数据第二级。
- 6.1.4.3 梳理需配套的数据安全管理制度，包括数据安全保密、数据备份与恢复、数据安全应急管理的要求。

6.1.5 功能性能要求

- 6.1.5.1 结合需求分析，比对数据类型格式支持和分布处理数据、多线查询数据、在线事务处理、数据分析等能力，进行主题数据库选型分析。
- 6.1.5.2 分析主题数据库的事务响应时间、吞吐量、并发用户数、数据访问和处理效率等性能指标。
- 6.1.5.3 规划主题数据库对外提供服务应用的方式，宜包括但不限于应用程序接口（API）、库表、文件等方式。

6.2 数据目录编制

6.2.1 目录编制

6.2.1.1 建设单位应在数据分析的基础上组织编制主题数据库数据目录。比对初步数据需求清单和数据共享交换平台的共享数据目录，对已上架数据共享交换平台的数据进行订阅。对未上架数据共享交换平台的数据，应协同大数据主管单位对其可能来源单位进行目录数据来源确认。

6.2.1.2 主题数据库数据目录的信息项应包括但不限于数据资源中文名称、来源单位、来源系统名称、数据共享属性、数据开放属性、数据级别等。

6.2.2 目录审查

建设单位对经确认的数据目录进行汇总检查后报大数据主管单位审查。

6.2.3 目录发布

6.2.3.1 大数据主管单位将审查通过的数据目录在数据共享交换平台上发布。

6.2.3.2 审查未通过的数据目录由建设单位落实修正后再提交大数据主管单位审查直至通过并发布。

6.2.3.3 主题数据库数据目录由建设单位负责维护。

6.2.4 共享申请

6.2.4.1 数据来源单位依据发布的数据目录将主题数据上架数据共享交换平台。

6.2.4.2 建设单位对已上架的主题数据进行使用申请，并通过数据共享交换平台获取接口文件、库表数据、文件数据等。

6.3 数据库设计与创建

6.3.1 数据库设计

6.3.1.1 在需求分析的基础上，应进行概念设计、逻辑设计和物理设计等系统化设计。

6.3.1.2 概念设计应结合业务需求分析、数据目录及分类、业务关系、语义关系等，提取业务实体和实体属性，使用实体关系模型来表示业务实体、属性，以及实体与属性之间的关系。

6.3.1.3 逻辑设计应在概念设计基础上进行数据规范化与结构优化，指定符合每个属性的数据类型，使用适当的连接表优化实体之间的关系，初步设计索引策略，确定具体的表结构、键和索引。

6.3.1.4 应结合政务数据平台的云化环境对逻辑设计进行物理实现设计，包括但不限于设计数据存储方式、优化性能和索引策略、制定数据备份计划等。

6.3.1.5 为增强主题数据库设计的可读性和可维护性，宜提升命名设计的规范性。主题数据库命名设计见附录 B。

6.3.2 数据库选型

6.3.2.1 主题数据库应结合业务需求和数据库设计要求进行数据库产品选型。

6.3.2.2 对以结构化主题数据为主的数据，应优先选择行列式表的关系型数据库进行存储；对于非结构化数据，如文本、图片、视频等，可以采用非关系型的数据库或云上文件服务器进行存储。

6.3.2.3 主题数据库建设应优先选择具有统筹安全、统管统维、性能良好和维护成本相对低等特点的云化数据库产品。数据量小的主题数据库应优先选择小型数据库产品；针对数据处理量大和并发请求高的大规模主题数据库，应优先选择分布式数据库产品。

6.3.2.4 主题数据库故障时其对外服务应能够自动迁移到备份节点并及时恢复服务，保证服务不受到影响。

6.3.2.5 主题数据库选型应考虑可扩展性，以满足业务增长对数据存储、数据处理和数据应用等方面的扩展需求。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：

<https://d.book118.com/036103023130011103>