



中华人民共和国国家标准

GB/T 36625.1—2018

智慧城市 数据融合 第1部分：概念模型

Smart city—Data fusion—Part 1: Conceptual model

2018-10-10 发布

2019-05-01 实施

国家市场监督管理总局 发布
中国国家标准化管理委员会

中 华 人 民 共 和 国
国 家 标 准
智慧城市 数据融合
第 1 部分：概念模型
GB/T 36625.1—2018

*

中国标准出版社出版发行
北京市朝阳区和平里西街甲 2 号(100029)
北京市西城区三里河北街 16 号(100045)

网址：www.spc.org.cn

服务热线：400-168-0010

2018 年 9 月第一版

*

书号：155066·1-61194

版权专有 侵权必究

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 概述	1
5 数据资产	3
5.1 概述	3
5.2 编码规则	3
5.3 数据元素	3
5.4 生命周期管理	3
6 数据采集	3
6.1 概述	3
6.2 数据清洗	3
6.3 数据转换	3
6.4 质量保证	3
7 数据描述	4
7.1 概述	4
7.2 数据源	4
7.3 数据模式	4
7.4 非结构化数据源表示规范	4
8 数据组织	4
8.1 概述	4
8.2 分类系统	4
8.3 标签库	5
8.4 词汇表	5
9 数据交换与共享	5
9.1 概述	5
9.2 交换共享平台	5
9.3 分析模型	5
10 数据服务	5
10.1 概述	5
10.2 数据检索	5
10.3 数据展示	5
11 开放共享要求	5

11.1	概述	5
11.2	业务需求	6
11.3	技术要求	6

前 言

GB/T 36625《智慧城市 数据融合》分为 5 个部分：

- 第 1 部分：概念模型；
- 第 2 部分：数据编码规范；
- 第 3 部分：数据采集规范；
- 第 4 部分：开放共享要求；
- 第 5 部分：市政设施基础数据元素。

本部分为 GB/T 36625 的第 1 部分。

本部分按照 GB/T 1.1—2009 给出的规则起草。

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这些专利的责任。

本部分由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本部分起草单位：北京航空航天大学、深圳市华傲数据技术有限公司、中国电子技术标准化研究院、中国电子科技集团公司信息科学研究所、华为技术有限公司、深圳赛西信息技术有限公司、智慧神州(北京)科技有限公司、中兴通讯股份有限公司、讯飞智元信息科技有限公司、大唐软件技术股份有限公司、中国信息通信研究院、中城智慧(北京)城市规划设计研究院有限公司、深圳市千方航实科技有限公司、北京计算机技术及应用研究所、上海计算机软件技术开发中心、上海交通大学。

本部分主要起草人：吕卫锋、蒲菊华、杜博文、刘棠丽、张红卫、贾西贝、于文渊、黎俊茂、赵菁华、荣文戈、康子路、崔昊、李赞、秦永辉、王琦、苏平、张巧英、袁媛、施媛、梁勇、臧磊、王新颖、万碧玉、吴丽丽、刘天平、刘伟麟、王元宏、刘振宇、姚斌、杨磊、陈伟权。

智慧城市 数据融合

第 1 部分：概念模型

1 范围

GB/T 36625 的本部分规定了智慧城市数据融合的概念模型、总体要求、基本过程及数据采集、数据描述、数据组织、数据交换与共享的基本要求。

本部分适用于智慧城市的数据融合实践。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是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 32909—2016 非结构化数据表示规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

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。

3.2

元数据 meta data

描述数据的数据。

4 概述

在智慧城市的场景下,通过采集与汇聚不同种类、不同来源数据,依次通过数据描述、数据组织和数据交换共享三个过程实现数据融合的功能,最终通过数据服务对外提供数据检索和展示等功能。

智慧城市数据融合概念模型如图 1 所示。图 1 中与数据资产相连的箭头表示数据流向,其余的箭头表示数据融合的流程。

根据图 1,数据融合概念模型包含数据采集、数据描述、数据组织、数据交换与共享和数据服务五个部分及数据资产、开放共享要求两个支撑要素。各部分的功能定位描述如下:

- a) 数据采集:采集、清洗不同种类、不同来源的数据;
- b) 数据描述:对数据源中的实体和关系进行抽象和表述;
- c) 数据组织:依赖分类系统对数据进行分类;
- d) 数据交换与共享:通过数据的交换共享,提升数据价值;
- e) 数据服务:对外提供数据检索和展示功能;
- f) 数据资产:数据资产包含数据以及数据融合中数据产生的其他信息,是上述流程中数据与信息获取、储存和管理对象,参与整个数据融合过程;

g) 开放共享要求:开放共享要求是数据融合相关活动对数据进行开放共享的技术规范与约束。

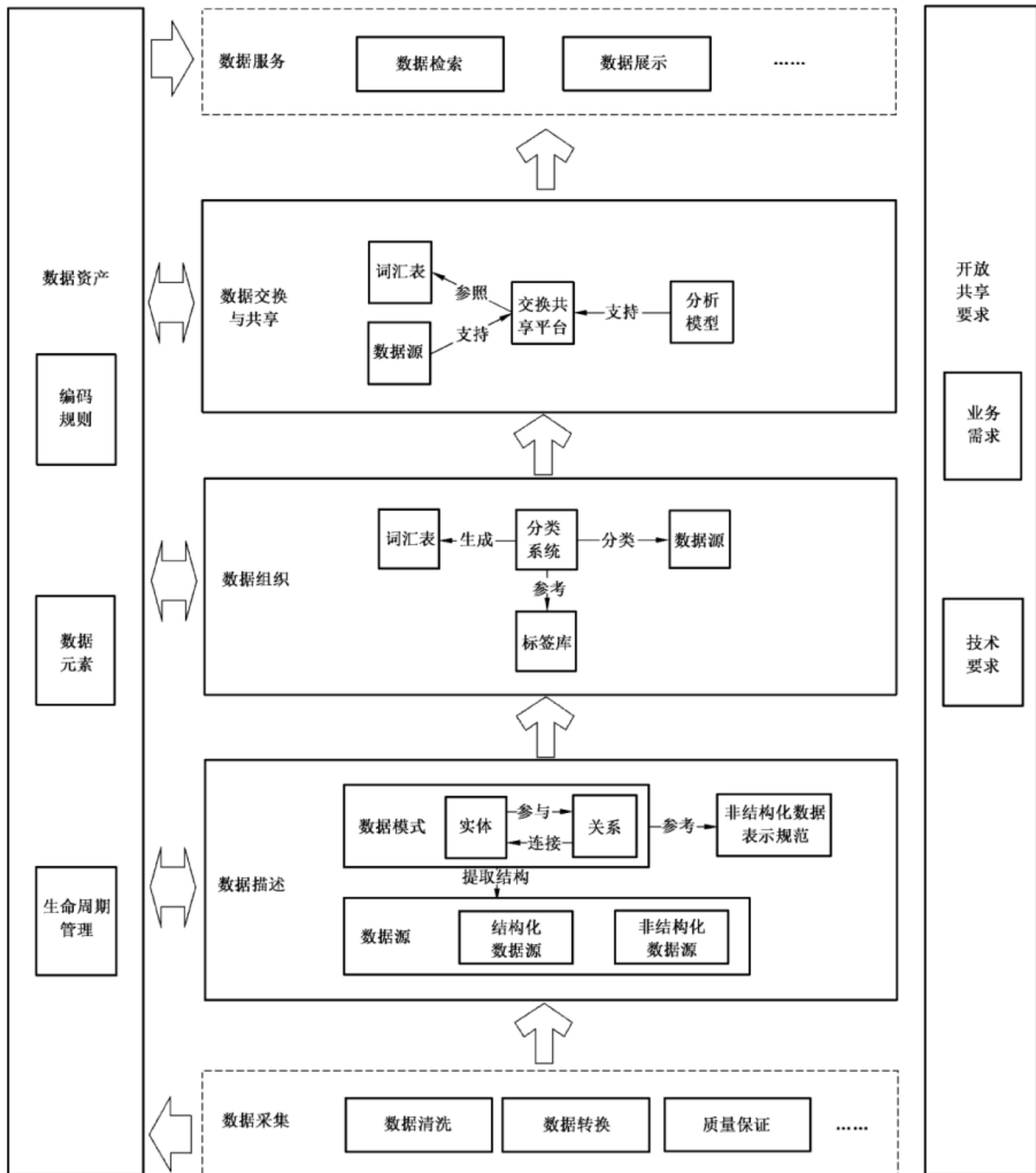


图 1 数据融合概念模型

智慧城市数据融合流程图如图 2 所示。图 1 中数据采集获取原始数据,数据描述产生抽象化数据,数据组织对数据进行分类整理产生信息,数据交换与共享通过分析模型产生知识。

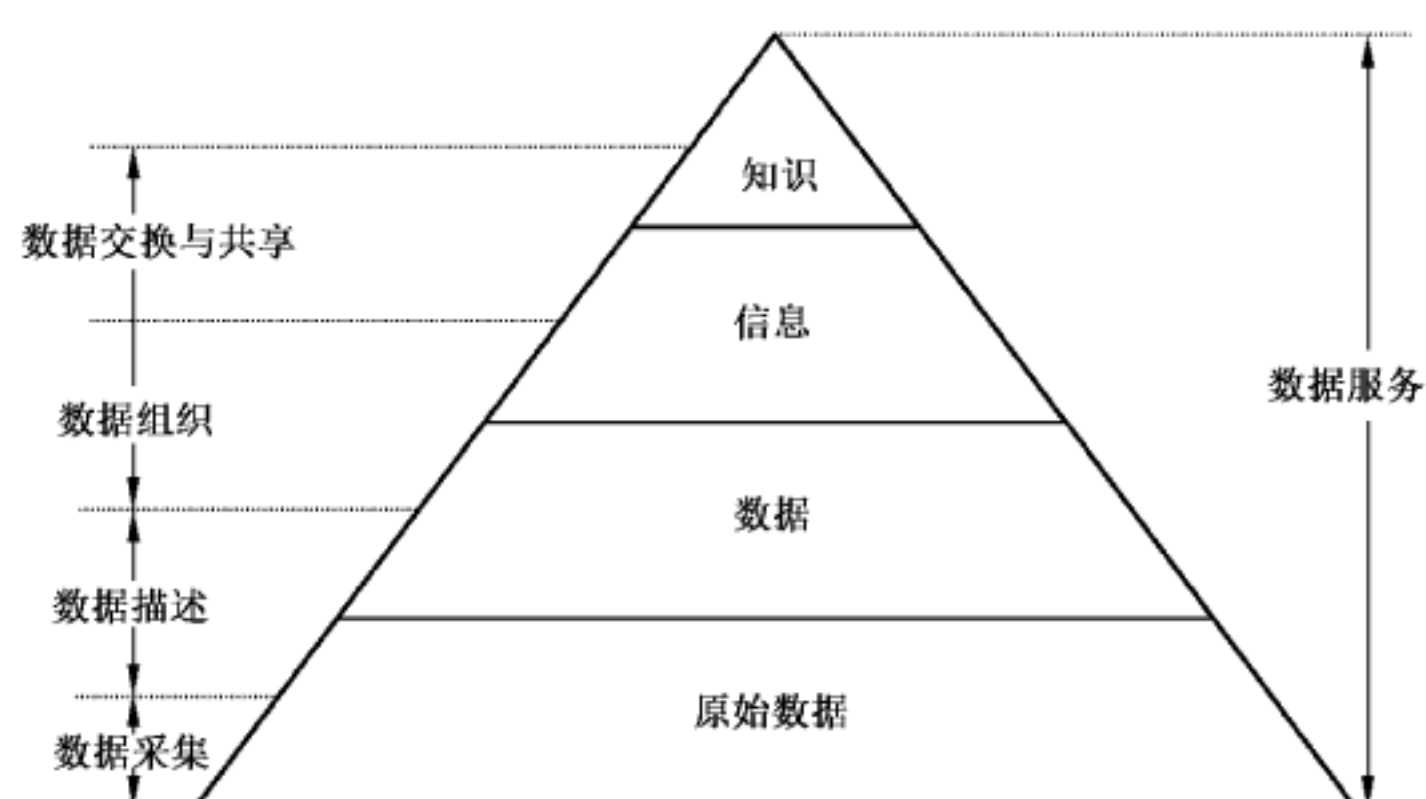


图2 数据融合流程图

5 数据资产

5.1 概述

数据资产包含数据以及数据融合中数据产生的其他信息,是数据采集、数据描述、数据组织、数据交换与共享和数据服务中数据与信息获取、储存和管理对象,参与整个数据融合过程。

5.2 编码规则

依照统一的编码规则对数据和数据源进行编码。

5.3 数据元素

对智慧城市应用领域的的数据制定统一的、规范的数据元素描述规则。

5.4 生命周期管理

用于数据在整个生命周期内的流动:从创建和初始存储,到它过时被删除。

6 数据采集

6.1 概述

数据采集是利用各种装置,采集数据并进行质量控制等操作后输入到数据资产的过程。

6.2 数据清洗

通过数据清洗,完成查重纠错,剔除数据文件中可识别的重复和错误。错误包括但不限于:检查数据一致性、处理无效值和缺失值等。

6.3 数据转换

通过数据转换,实现对标准代码、格式、类型等信息的标准化转换处理。

6.4 质量保证

通过质量保证,保证数据的真实性、完整性、可靠性、可用性、可再用性、可追溯性。

7 数据描述

7.1 概述

数据描述指对获取到的数据进行统一的描述。

对于结构化数据源,直接使用数据模式对其进行描述;对于非结构化数据源,可先依据非结构化数据源表示规范对其进行描述,再根据其描述信息来抽取数据模式,最后将数据源和其对应的数据模式汇聚到数据资产中。

7.2 数据源

数据源分为结构化数据源和非结构化数据源。结构化数据源是结构化实体的载体,对于同一类型的结构化实体具有统一的访问接口,与所存储的数据无关;非结构化数据源是除结构化实体以外的其他实体的载体。

7.3 数据模式

7.3.1 概述

数据模式是数据的一种表示方式,一般是逻辑数据模型或概念数据模型。

数据模式主要由实体和关系组成。实体具有一定的属性,属性被数据类型约束,不同的实体通过关系连接。

7.3.2 实体

实体通过元数据进行描述。

实体是依据共有属性,对人、地点、物体、事件及概念的抽象表示。

实体之间存在着一定的关系,一般可以用树形结构来表示实体之间的关系。数据类型是属性取值的约束。属性是构成实体的基本元素,是同一种实体共有的特征。

7.3.3 关系

关系是指实体之间的互相联系。

实体中的层次结构也是关系的一种表现(包含关系)。由关系连接实体所得到的数据源可以嵌入智慧城市实体的层次化结构中,并且同一种数据源可以被不同的实体所包含。

7.4 非结构化数据源表示规范

按照 GB/T 32909—2016 中规定的非结构化数据表示规范对实体、数据模式进行描述。

8 数据组织

8.1 概述

数据组织指对数据源进行分类并生成词汇表。

首先从数据资产中获取各数据源的实体关系描述,再利用分类系统,参考已定义好的标签库,对各数据源中的实体进行分类,最后生成数据的词汇表,存储到数据资产中。

8.2 分类系统

分类系统是参考标签库中不同的类别,对包含不同的实体的数据源进行分类的系统。

8.3 标签库

标签库是分类系统类别的参考,包含一系列实体所共有的特征。

8.4 词汇表

词汇表是数据资源目录,记录标签库中各标签对应的数据源的索引。

9 数据交换与共享

9.1 概述

数据交换与共享基于词汇表,对不同种类的数据源进行统一管理和分析。

9.2 交换共享平台

交换共享平台是实现数据交换、共享功能平台,是数据资产的统一管理平台。

9.3 分析模型

分析模型对数据资产进行知识发现。

分析模型主要包括但不限于:

- a) 描述性分析:描述性分析是对一组数据的各种特征进行分析,以便于描述测量样本的各种特征及其所代表的总体特征。
- b) 预测性分析:预测分析为每一个个体以概率的形式,提供一个预测评分,从而决策,反馈或影响针对于大量上述个体的组织性流程。
- c) 诊断性分析:诊断性分析是测试一个因子与相对应的测度项之间的关系是否符合研究者所设计的理论关系的方法。
- d) 因果性分析:因果性分析是使用因果图对数据进行分析的方法。

10 数据服务

10.1 概述

数据服务对外提供数据资产中的数据、信息和知识的检索和展示等功能。

10.2 数据检索

数据检索是将存储的数据、信息和知识根据特定需求提取出来。

10.3 数据展示

数据展示是将特定数据在不同设备、平台上展示。

11 开放共享要求

11.1 概述

开放共享要求规定了数据融合整个过程中数据交换、共享和开放所需遵从的规范。其主要涉及数据融合开放共享业务需求和技术要求。

11.2 业务需求

根据业务场景,数据开放共享应满足政府相关部门、城市基础设施、企业信息系统和个人设备的业务需求。

11.3 技术要求

数据开放共享技术要求主要包括数据质量、数据安全、隐私保护等方面的要求。
