城市地下空间大数据智能分析与公共服务平台建设及示范应用

数据库设计说明书

西安电子科技大学

目 录

[数据库设计说明书 1](#_Toc18053276)

[第一章 前言 1](#_Toc18053277)

[1.1 编写目的 1](#_Toc18053278)

[1.2 范围 1](#_Toc18053279)

[1.3 参考资料 1](#_Toc18053280)

[第二章 外部设计 2](#_Toc18053281)

[2.1 标识符和状态 2](#_Toc18053282)

[2.2 命名约定 2](#_Toc18053283)

[2.3 设计约定 2](#_Toc18053284)

[第三章 结构设计 3](#_Toc18053285)

[3.1.1 业务建模 3](#_Toc18053286)

[3.1 实体和属性的定义 3](#_Toc18053287)

[3.2.1 ER关系图 6](#_Toc18053288)

[3.3.1 物理设计 7](#_Toc18053289)

[3.1.2 数据建模 9](#_Toc18053290)

[3.1 实体和属性的定义 9](#_Toc18053291)

[3.2.2 ER关系图 13](#_Toc18053292)

[3.3.2 物理设计 14](#_Toc18053293)

[3.1.3 分析建模 18](#_Toc18053294)

[3.1 实体和属性的定义 18](#_Toc18053295)

[3.2.3 ER关系图 20](#_Toc18053296)

[3.3.3 物理设计 21](#_Toc18053297)

# 数据库设计说明书

## 第一章 前言

### 1.1 编写目的

1. 本数据库设计说明书是关于大数据智能开发工具的数据库设计，主要包括数据逻辑结构设计、数据字典及运行环境、安全设计等。
2. 本数据设计说明书读者：系统设计人员、系统测试人员、系统维护人员‘
3. 数据库设计说明书是根据系统需求分析设计所编写的。
4. 数据库设计说明书为开发软件提供了一定的基础。

### 1.2 范围

本数据库设计说明书主要设计的范围包括：

* 1. 数据库逻辑设计描述，主要根据实体关系图（Entity-Relationship）
  2. 数据库物理设计：数据库表的结构定义、存储过程定义。触发器定义。

### 1.3 参考资料

《软件工程导论》（第五版） 张海藩，清华大学出版社

《数据库系统概论》（第四版）王珊，萨师煊，高等教育出版社

《计算机软件产品开发文件编制指南， GB 8567-88》

## 第二章 外部设计

### 2.1 标识符和状态

数据库软件的名称：MySql-5.5.27

数据库的名称为：model

### 2.2 命名约定

所有数据库命名都是以模块的缩写加上具体表的英文词汇组成，这样能够统一数据库表的命名，也能够更好的规范数据库表命名。

### 2.3 设计约定

本数据库的设计采用面向对象的设计方法，首先进行对象实体的设计，然后将对象持久化到数据库中，最后使用Visio设计工具实现表与表之间的关联（ER图），从而实现数据库的设计。

## 第三章 结构设计

### 3.1 业务建模

业务建模包含业务资料、业务主线、业务场景、业务问题个部分，以支持构建新业务时的业务分析。本模块提供业务的资料的上传和下载，支持构建业务主线、场景以及问题，最终形成业务总览图，为分析建模提供业务问题文字描述，公式化表示以及可视化图形，为数据建模提供业务数据类型。

#### 3.1.1 实体和属性的定义

1.【业务模型】用于保存新建的业务名，新业务将包含业务资料、业务主线、业务场景、业务问题，业务标识用于识别业务模型，如图3.1- 1。



图3.1- 1业务模型实体属性定义

2.【业务资料】可以在建模的整个流程中使用，辅助建模，存储内容包括文件名、上传时间、存储路径、以及文件大小，文件标识用于识别业务资料，如图3.1- 2。



图3.1- 2业务资料实体属性定义

3.【业务主线】通过业务名进行业务流程的构建，主线标识用于识别业务主线，流程标识用于识别业务流程，如图3.1- 3。



图3.1- 3业务主线实体属性定义

4.【业务场景】是针对“业务主线”上的某一个具体环节进行刻画，提供“新建业务场景”及“已创建的业务场景”的管理功能，场景标识用于识别业务场景，节点标识用于识别业务主线具体环节，如图3.1- 4。



图3.1- 4 业务场景实体属性定义

5.【业务问题】是对业务场景进行定位，并进行问题描述，具体包括文字描述、公式描述及可视化描述，问题标识用于识别业务问题，如图3.1- 5。



图3.1- 5业务问题实体属性定义

6.【业务对象】是对业务场景中的业务节点做进一步的描述。包括对象名、对象属性、对象类型和对象描述，对象标识用于识别对象名，如图3.1- 6。



图3.1- 6业务对象实体属性定义

7.【业务数据】是对业务场景中的业务节点中的数据类型做进一步的描述。数据标识用于识别业务数据名，如图3.1- 7。



图3.1- 7业务数据实体属性定义

8.【业务知识】是对业务场景中的业务节点中的内容做进一步描述。其中包括业务知识名和业务知识描述，业务知识标识用于识别业务名，如图3.1- 8。



图3.1- 8业务知识实体属性定义

#### 3.1.2 ER关系图

业务建模信息模型设计主要包括业务场景总表、业务对象总表、业务数据类型总表、业务节点信息总表、业务建模总表、业务知识总表和业务问题总表。根据业务问题构建出相应的业务场景表，抽取一个或多个业务节点形成各自的业务节点流程。根据节点流程选择业务对象、业务数据类型和业务知识保存场景信息，如图3.1- 8和图3.1- 9。



图3.1- 9业务建模E-R物理模型

图3.1- 10业务建模E-R图

#### 3.1.3 物理设计

1.业务模型（业务标识，业务名，业务资料，业务主线，业务场景，业务问题）

表3.1- 1业务模型表business\_model

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| businessid | int | 5 | 是 | 是 | 业务标识 |
| businessname | int | 10 | 否 | 否 | 业务名 |
| businessmaterial | varchar | 10 | 否 | 否 | 业务资料 |
| businessline | varchar | 10 | 否 | 否 | 业务主线 |
| businessscene | varchar | 20 | 否 | 否 | 业务场景 |
| businessobject | varchar | 20 | 否 | 否 | 业务问题 |

2.业务资料（文件标识，文件名，上传时间，存储路径，文件大小、业务标识）

表3.1- 2业务资料表business\_file

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| fileid | int | 5 | 是 | 是 | 文件标识 |
| filename | varchar | 10 | 否 | 否 | 文件名 |
| filepath | varchar | 10 | 否 | 否 | 上传时间 |
| filetime | datetime | 0 | 否 | 否 | 存储路径 |
| filesize | double | 20 | 否 | 否 | 文件大小 |
| businessid | int | 5 | 否 | 否 | 业务标识 |

3.业务主线（主线标识，主线名，流程标识，业务标识）

表3.1- 3 业务主线表business\_ line

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| lineid | int | 5 | 是 | 是 | 主线标识 |
| linename | varchar | 5 | 否 | 否 | 主线 |
| flowid | int | 10 | 否 | 否 | 流程标识 |
| businessid | int | 5 | 否 | 否 | 业务标识 |

4.业务对象（对象标识，业务对象名，对象类型，对象属性，对象描述）

表3.1- 4业务对象表business\_ object

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| objectid | int | 5 | 是 | 是 | 对象标识 |
| objectname | varchar | 10 | 否 | 否 | 业务对象名 |
| objecttype | varchar | 10 | 否 | 否 | 对象类型 |
| objectattribute | varchar | 10 | 否 | 否 | 对象属性 |
| objectdescribe | varchar | 10 | 否 | 否 | 对象描述 |

5.业务数据（数据标识，数据类型名）

表3.1- 5业务数据表business\_ type

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| dataid | int | 5 | 是 | 是 | 数据标识 |
| datatypename | varchar | 10 | 否 | 否 | 数据类型名 |

6.业务知识（业务知识标识，业务知识名，业务知识描述）

表3.1- 6业务知识表business\_ knowledge

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| knowledgeid | int | 5 | 是 | 是 | 业务知识标识 |
| knowledgename | varchar | 20 | 否 | 否 | 业务知识名 |
| knowledgedescribe | varchar | 20 | 否 | 否 | 业务知识描述 |

7.业务场景（场景标识，场景名，节点标识）

表3.1- 7业务场景表business\_ scene

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| sceneid | int | 5 | 是 | 是 | 场景标识 |
| scenename | varchar | 20 | 否 | 否 | 场景名 |
| nodeid | Int | 11 | 否 | 否 | 节点标识 |

### 3.2 数据建模

基于业务建模，数据建模模块包括感知、筛选、关联、索引、调用等阶段，具体划分为数据链、数据区、数据块、数据包、数据集等节点，涵盖了从数据产生、处理、应用的整个过程，支持全过程可视化展示，为分析建模奠定数据基础。

#### 3.2.1 实体和属性的定义。

1.【数据模型】用于保存新建的数据模型和模型的信息，包括数据链、数据区、数据包和数据集。其中模型标识用于识别新建的模型名，如图3.2- 1。



图3.2- 1数据模型实体属性定义

2.【数据链】即数据的感知阶段，主要是基于业务问题，描述数据以及数据的形成过程。包括数据链名，和数据链流程。其中数据链标识用于识别数据链，如图3.2- 2。



图3.2- 2数据链实体属性定义

3.【数据源】用于定义数据的来源，名称，大小等数据的基本信息。包括数据源名，数据上传时间，存储路径，数据文件大小。其中数据源标识用于识别具体的数据源文件，如图3.2- 3。



图3.2- 3数据源实体属性定义

4.【数据区】即筛选阶段，用于解决数据链中数据的质量问题，包括合并和处理数据链中的数据。其中数据区标识用于识别数据区，如图3.2- 4。



图3.2- 4数据区实体属性定义

5.【算法库】用于存储数据处理过程中用到的算法。包括算法名，算法分类和算法简介。其中算法标识用于识别不同的算法，如图3.2- 5。



图3.2- 5算法库实体属性定义

6.【数据块】即关联阶段。主要是用于对数据进行特征工程、数据聚类、数据关联等深度处理，形成多个数据块。包括数据块名，数据块形成时间等基本属性，其中数据块标识用于识别数据块，如图3.2- 6。



图3.2- 6数据块实体属性定义

7.【数据包】即融合阶段，主要是对数据块进行数据特征融合，形成多个数据包。包括数据包名，数据包形成时间。其中数据包标识用于识别不同的数据包，如图3.2- 7。



图3.2- 7数据包实体属性定义

8.【数据集】即索引阶段，主要是对数据包建立索引，形成贴合业务需求的数据集。包括数据集名，数据集的划分类型，数据集划分时间。其中数据集标识用于识别不同的数据集，如图3.2- 8。



图3.2- 8 数据集实体属性定义

#### 3.2.2 ER关系图

数据建模信息模型设计包括数据模型总表、数据逻辑关系图库、数据链图库、数据区总表、数据块总表、数据包总表、数据集总表、数据库总表。基于数据链图库抽取数据逻辑关系图库，并形成数据区，通过数据处理依次形成数据块、数据包、数据集，最后形成数据库总表，如图3.2- 9和图3.2- 10。



图3.2- 9数据建模E-R物理模型  


图3.2- 10数据建模E-R图

#### 3.2.3 物理设计

1.【数据模型】用于保存新建的数据模型和模型的信息，包括数据链、数据区、数据包和数据集。

表3.2- 1数据模型表datamodel\_info

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| informationid | int | 5 | 是 | 是 | 模型信息标识 |
| datalink | varchar | 5 | 否 | 否 | 数据链 |
| dataarea | varchar | 5 | 否 | 否 | 数据区 |
| datablock | varchar | 5 | 否 | 否 | 数据块 |
| databao | varchar | 5 | 否 | 否 | 数据包 |
| dataji | varchar | 5 | 否 | 否 | 数据集 |
| modelid | int | 11 | 否 | 否 | 模型标识 |

2.数据链（数据链标识，数据链名，数据链流程图，模型标识）

表3.2- 2数据链表datamodel\_ link

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| linkid | int | 5 | 是 | 是 | 数据链标识 |
| linkname | varchar | 255 | 否 | 否 | 数据链名 |
| linkpicture | varchar | 255 | 否 | 否 | 数据链流程图 |
| modelid | int | 11 | 否 | 否 | 模型标识 |

3.数据源表（数据源标识，数据源名，上传时间，存储路径，文件大小）

表3.2- 3数据源表datamodel\_ source

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| sourceid | int | 5 | 是 | 是 | 数据源标识 |
| sourcename | varchar | 10 | 否 | 否 | 数据源名称 |
| sourcetime | datetime | 0 | 否 | 否 | 上传时间 |
| sourcepath | varchar | 50 | 否 | 否 | 存储路径 |
| sourcesize | double | 0 | 否 | 否 | 文件大小 |

4.数据区表（数据区标识，数据区名，数据逻辑关系，数据预处理，流程标识，数据链标识）

表3.2- 4数据区表datamodel\_ area

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| areaid | int | 5 | 是 | 是 | 数据区标识 |
| areaname | varchar | 20 | 否 | 否 | 数据区名称 |
| linkid | int | 5 | 否 | 否 | 数据链标识 |

5.算法库（算法标识，算法名，算法分类，算法简介）

表3.2- 5算法库表datamodel\_ algorithm

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| algorithmid | int | 5 | 是 | 是 | 算法标识 |
| algorithmname | varchar | 10 | 否 | 否 | 算法名称 |
| algorithmtype | varchar | 20 | 否 | 否 | 算法分类 |
| algorithmdescribe | varchar | 20 | 否 | 否 | 算法简介 |

6.算法处理表（表标识，表名，创建时间，处理前后数据类型，算法标识）

表3.2- 6算法处理表datamodel\_ process

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| processed | int | 5 | 是 | 是 | 表标识 |
| processname | varchar | 10 | 否 | 否 | 表名 |
| processtime | datetime | 0 | 否 | 否 | 创建时间 |
| areaid | int | 5 | 否 | 否 | 处理前后数据类型 |
| algorithmid | int | 5 | 否 | 否 | 算法标识 |

7.数据块（数据块标识，数据块名，创建时间，表标识，算法标识）

表3.2- 7数据块表datamodel\_ block

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| blockid | int | 5 | 是 | 是 | 数据块标识 |
| blockname | varchar | 20 | 否 | 否 | 数据块名 |
| blocktime | datetime | 0 | 否 | 否 | 创建时间 |
| processid | int | 5 | 否 | 否 | 表标识 |
| algorithmid | int | 5 | 否 | 否 | 算法标识 |

8.数据包（数据包标识，数据包名，创建时间，数据块标识，算法标识）

表3.2- 8数据包表datamodel\_ bag

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| bagid | int | 5 | 是 | 是 | 数据包标识 |
| bigname | varchar | 5 | 否 | 否 | 数据包名 |
| bigtime | datetime | 0 | 否 | 否 | 创建时间 |
| blockid | int | 5 | 否 | 否 | 数据块标识 |
| algorithmid | Int | 5 | 否 | 否 | 算法标识 |

9.数据集（数据集标识，数据集名，数据集划分，数据划分时间，数据包标识）

表3.2- 9数据集表datamodel\_ collect

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| collectid | int | 5 | 是 | 是 | 数据集标识 |
| collectname | varchar | 20 | 否 | 否 | 数据集名称 |
| collecttype | varchar | 10 | 否 | 否 | 数据集划分 |
| collecttime | datetime | 0 | 否 | 否 | 数据划分时间 |
| collectid | Int | 5 | 否 | 否 | 数据包标识 |

### 3.3 分析建模

分析建模模块包括分析流程、分析单元、模型构建、模型融合、分析调试和分析评估六个部分，支持了分析模型构建、运行计算、结果分析和可视化展示，完成对业务问题的全方位分析。

#### 3.3.1 实体和属性的定义

1.【分析流程】主要是获取业务问题的文字描述、数学描述和可视化图形；将分析目标做以分解细化，给出数学描述的公式、公式中各项的含义及各项求解步骤和可视化图形，建立分析流程。包括文字描述，数学描述等基本信息。其中流程标识用于识别分析流程，如图3.3- 1。



图3.3- 1分析流程实体属性定义

2.【分析单元】根据分析流程输入需要分析的单元，包括对数据包的要求、确定分析类型和结果定义。包括分析单元名，分析单元类型和结果定义。其中分析单元标识用于识别不同的分析单元，如图3.3- 2。



图3.3- 2分析单元实体属性定义

3.【分析模型】主要用于保存构建的分析模型。包括分析模型名。模型创建时间，模型标识用于识别不同的分析模型，如图3.3- 3。



图3.3- 3分析模型实体属性定义

4.【分析评估】用于对模型进行管理。包括评估标识，评估说明，评估描述等。其中评估标识用于识别不同的分析评估结果，如图3.3- 4。



图3.3- 4分析评估实体属性定义

#### 3.3.2 ER关系图

分析建模信息模型设计包括分析建模总表、分析流程库、分析单元库、分析结果库、可视化组件库、数据库、算法库、模型库。基于分析单元库创建分析流程，通过绑定数据库、算法库以及模型库信息进行计算，获得分析结果，并通过可视化组件库进行可视化展示，如图3.3- 5和图3.3- 6。



图3.3- 5分析建模E-R物理模型



图3.3- 6分析建模E-R图

#### 3.3.3 物理设计

1.分析流程表（流程标识，文字描述，数学描述，业务问题标识）

表3.3- 1分析流程表analyse\_flow

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| flowid | int | 5 | 是 | 是 | 流程标识 |
| worddescribe | varchar | 10 | 否 | 否 | 文字描述 |
| mathdescribe | varchar | 20 | 否 | 否 | 数学描述 |
| businessobjectid | int | 5 | 否 | 否 | 业务问题标识 |

2.分析单元表（单元标识，分析单元名，分析类型，结果定义，数据包标识，分析流程标识）

表3.3- 2分析单元表analyse\_ unit

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| unitid | int | 5 | 是 | 是 | 单元标识 |
| unitname | varchar | 5 | 否 | 否 | 分析单元名 |
| result | varchar | 10 | 否 | 否 | 结果定义 |
| type | varchar | 10 | 否 | 否 | 分析类型 |
| baoid | int | 5 | 否 | 否 | 数据包标识 |
| flowid | Int | 5 | 否 | 否 | 分析流程标识 |

3.分析模型表（模型标识，分析模型名，创建时间，分析流程标识，算法标识）

表3.3- 3分析模型表analyse\_model

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| modelid | int | 5 | 是 | 是 | 模型标识 |
| modelname | varchar | 10 | 否 | 否 | 分析模型名 |
| modeltime | datetime | 0 | 否 | 否 | 创建时间 |
| flowid | int | 5 | 否 | 否 | 分析流程标识 |
| algorithmid | Int | 5 | 否 | 否 | 算法标识 |

4.分析评估（评估标识，模型名称，创建时间，说明，结果）

表3.3- 4分析评估表analyse\_ asses

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 字段名 | 数据类型 | 长度 | 主键 | 非空 | 描述 |
| assessid | int | 5 | 是 | 是 | 评估标识 |
| assesname | varchar | 10 | 否 | 否 | 分析评估名称 |
| assestime | varchar | 10 | 否 | 否 | 创建时间 |
| modelid | int | 5 | 否 | 否 | 分析模型标识 |
| algorithmid | int | 5 | 否 | 否 | 分析单元标识 |