



第6章 空间数据库设计与建设

中南大学地信系 李光强
QQ : 41733233



教学目标

- 掌握空间数据库设计的流程及操作方法
- 了解空间数据标准在数据建库中的作用
- 熟悉空间数据库建设过程
- 掌握空间数据质量的概念及其控制技术

教学重点

- 空间数据库设计流程
- 空间数据库建设内容



教学内容

- 6.1 空间数据库设计概述
- 6.2 空间数据标准
- 6.3 空间数据库设计
- 6.4 空间数据库建设
- 6.5 空间数据库建设实例

□空间数据库设计则是空间数据管理的关键步骤，设计成果直接影响上层应用软件的体系结构。空间数据库的设计既具有传统数据库设计的特点，又具有空间数据独有的特征。

- 6.1.1 数据库设计
- 6.1.2 空间数据库设计概念
- 6.1.3 空间数据库设计原则
- 6.1.4空间数据库设计过程

- 广义上讲，数据库设计是对整个数据库系统的规划，包括数据库相关的软件、硬件、应用环境等的设计。
- 狭义地讲，数据库设计是对数据库自身结构的规划，包括数据库模式目录、数据表命名与结构、功能行为等的设计。
- **数据库设计**是面向业务主题并在给定应用环境约束下，为业务系统规划合理的数据模型与处理功能，确定数据库**存储结构与存储方法**的物理设计，使之能够满足业务管理需要的信息管理和数据操作等功能

□数据库设计的**内容**包括：

- (1) **结构设计**。结构设计是对数据库模式、信息实体结构及其联系的规划，包括概念结构设计和逻辑结构设计。
- (2) **功能设计**。功能设计是在业务系统体系结构约束下，对数据库行为的设计，包括数据库查询、约束条件、事务处理、函数和存储过程等的设计。
- (3) **物理设计**。物理设计是在特定的数据库管理系统环境中，实现数据库结构和功能行为的过程。

□数据库设计的**步骤**包括

- 需求分析
- 结构和功能设计
- 物理设计
- 设计验证
- 建设和运行维护

- **空间数据库的设计**是地理信息**项目总需求目标**的指引下，依据项目设计的要求，分析业务系统用户的数据现状，归纳项目需要管理的数据内容和类型，遵循国家相关标准和规范，并在一定的软件、硬件环境约束下，构造合理的且能满足地理信息业务应用与性能要求的**数据库模型**，并能在特定的数据库管理系统中实现**物理模型**。
- 空间数据库设计目标：
 - （1）依据地理信息系统管理的详细目标和任务，归纳、概括数据建设内容和类型，并设计数据库模型。
 - （2）空间数据库设计人员要细化和分解地理信息系统需求目标，设计出满足当前需求目标规定的数据库内容，也要设计需求目标没有规定且必需的数据内容。
 - （3）设计的空间数据库模型要具有良好的数据完整性、数据安全性和最佳的访问性能。

□空间数据库设计的**内容**：

- (1) **数据库结构设计**首先要描述和定义地理信息系统涉及的数据**实体及其关系**，然后定义数据表的字段结构。
- (2) **数据库行为设计**是指施加在数据库上的动态操作的设计，也是地理信息系统功能在数据管理层的映射。空间数据库的功能设计主要包括表触发器、自定义函数和存储过程等内容。
- (3) **数据库安全设计**也是空间数据库设计的重要内容，包括数据并发访问控制和数据操作日志、数据恢复等的规划。并发控制是多用户同时访问数据时，在保证数据安全的前提下能够独立工作。
- (4) 确定每类空间要素的类型、空间维度和表现形式，依据这些要素的特征选择合适的**表达方式和表达模型**。
- (5) 确定**专题图层对应的空间数据表**，并为每个专题图层设计适当的数据表。

□ (一) 数据库设计规范化要求

- **规范化**实质上是概念的单一化，即一个关系只表示一个实体或联系。
- **规范化**理论是用来改造关系模式，通过分解关系模式消除其中不合适的数据依赖，以解决插入异常、删除异常、更新异常和数据冗余等问题

□ (二) 关系范式

- 关系数据库的设计需要遵循一定的规范化原则，即关系范式。**关系范式**是关系模式满足不同程度规范化要求的标准，对关系模式属性间的函数依赖加以不同的限制，就形成了不同级别的范式，包括：
 - (1) 1NF关系是对关系的最基本定义，只要关系中各分量是独立的，就是满足1NF要求的关系。
 - (2) 对1NF关系进行投影，消除原关系中非主属性对键的部分函数依赖，将1NF关系转换成若干个2NF关系。
 - (3) 对2NF关系进行投影，消除原关系中非主属性对键的传递函数依赖，将2NF关系转换成若干个3NF关系。
 - (4) 对3NF关系进行投影，消除原关系中主属性对键的部分函数依赖和传递函数依赖，也就是说，使决定因素都包含一个候选键，得到一组BCNF关系。
 - (5) 对BCNF关系进行投影，消除原关系中的非平凡且非函数依赖的多值依赖，得到一组4NF的关系。
 - (6) 对于4NF 不包含任何连接依赖关系并且连接应该是无损的，则得到5NF。即当所有表都被分成尽可能多的表以避免冗余时，满足5NF。
- 在实际应用中，范式越高意味着数据表的划分越细，原本相关联的数据就需要分配到多个表中存储，从而造成数据表之间的关联与就越复杂。

- (一) 需求分析
 - 主要包括前期设计文档、数据存储格式、数据之间关系、数据使用方式及频度等。
- (二) 概念设计
 - 分类和概括前期分析获取的数据类型和实体，进而建立抽象的概念数据模型的过程。
- (三) 逻辑设计
 - 将概念数据模型映射为特定数据库管理系统支持的数据模型的过程，也就是使用特定数据库管理系统的数据库模型、数据类型，详细定义概念模型中的实体及其属性特征。
- (四) 物理设计
 - 空间数据库的物理结构是指数据库在物理设备上的存储结构、文件类型、索引结构、存取方法等的物理特性。
- (五) 验证设计
 - 验证设计是组织专家组评审、检验空间数据库设计方案的过程，以确定设计方案能否满足地理信息系统目标需求。
 - 验证方法包括包括文档评审和测试验证。

□ (1) 空间数据标准化历程

- **ISO/TC211**是目前最具国际权威性的地理信息标准化组织，已制定了空间数据管理、采集、处理、分析、查询、表示等方面的标准体系
- **OGC**制定的标准包括空间数据模型、空间数据处理、空间数据共享和空间服务模型等
- **CSBTS/TC230**负责我国地理信息国家标准的立项建议、协调组织、编写标准、上报审查

□ 空间数据库设计遵循标准化原则具有以下**优点**：

- (1)标准规范凝聚了大量空间数据管理相关的知识
- (2)借鉴和遵循现有的标准，能节省调研时间，有利于快速形成数据库设计文档
- (3)严格遵循标准规范，有利于数据库设计人员之间的协调和合作
- (4)遵循相关标准规范，有利于后期数据的发布、共享和使用，有助于后期数据维护工作的开展

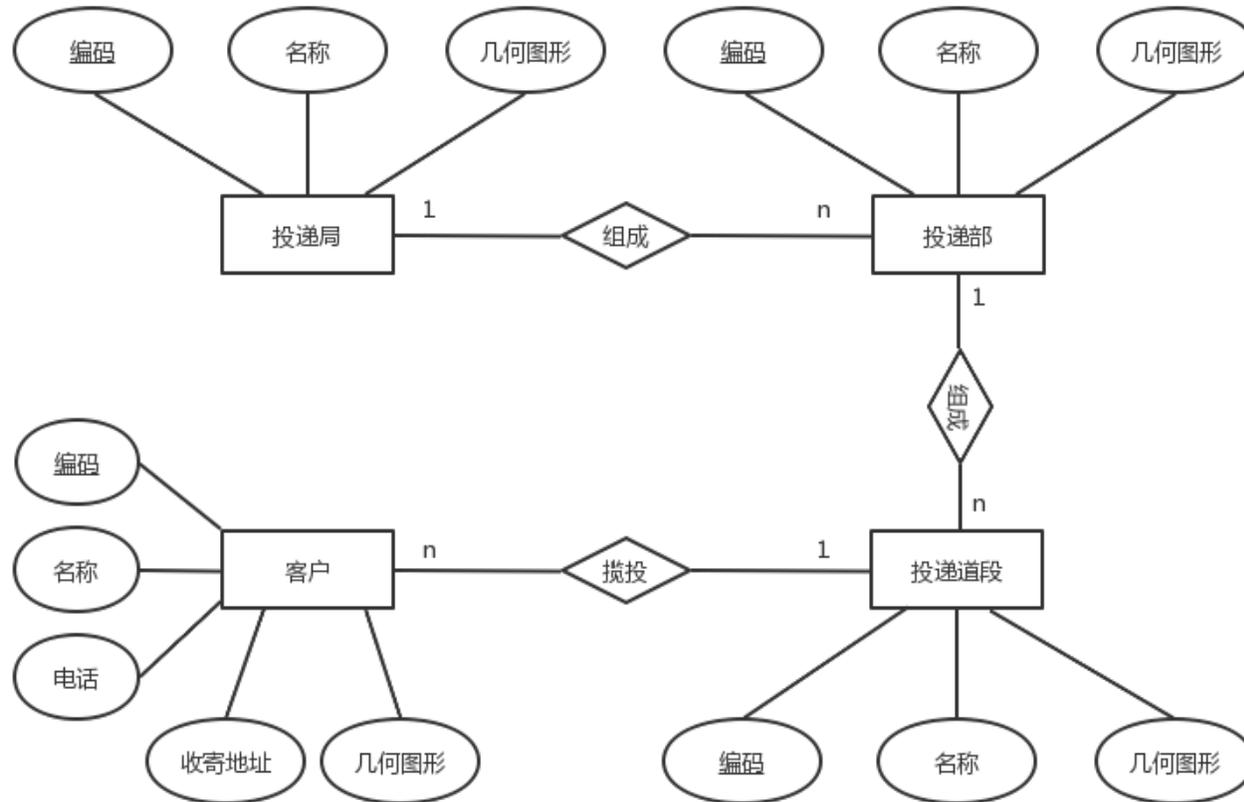
□空间数据库设计包括**结构设计**、**功能设计**和**物理设计**，结构设计又包括概念设计、逻辑设计。

- 6.3.1 概念设计
- 6.3.2 逻辑设计
- 6.3.3 物理模型设计
- 6.3.4 空间数据库设计报告

- 空间数据库的**概念设计**是将地理信息系统的空间实体抽象为信息世界结构模型的过程，常使用**实体-关系模型**表达数据对象之间的关系。
- (一) 实体-关系模型
 - 实体-关系模型 (entity-relationship model , E-R模型) 是一种不受数据库管理系统约束的、面向用户业务管理的表达模型，广泛应用在数据库概念设计中。
 - E-R模型由**实体、属性和联系**三种元素构成，表示方法如下：
 - (1) 实体用矩形框表示，矩形框内写上实体名。
 - (2) 实体的属性用椭圆框表示，框内写上属性名，并用无向边与实体相连。
 - (3) 实体间的联系用菱形框表示，联系以适当的含义命名，联系名称写在菱形框中，用无向连线将参加联系的实体矩形框分别与菱形框相连，并在连线上标明联系的类型，即1—1、1—N或M—N。
 - 在设计**空间数据库概念模型**时，为了体现“空间图形”属性以及要素之间的空间操作关系，一些学者提出了面向空间数据库设计的“**扩展实体-关系模型**”。

□ (一) 实体-关系模型

- 示例：城市的快递物流管理系统里，整个城市有一个投递局，一个投递局包括多个投递部，一个投递部包括多个投递道段，一个道段负责揽投多个客户快递业务。ER图：



□ (二) 概念设计步骤

- (1) **确定所有的实体集合**。设计人员在收集和确定实体时，需要从系统需求文档和设计文档中分析需要管理的数据实体，得到数据库的实体集。
 - (2) **选择实体集应包含的属性，以及空间图形类型**。实体属性可以直接继承设计文档数据字典的属性描述，或者是类的属性，还可以参考国家或行业的相关标准，直接引用规范中的图层名称及其字段定义。
 - (3) **确定实体集之间的联系**。实体之间关系可以将设计文档描述的数据流连接映射为实体关系，或者根据实际应用场景客观地物之间的关系确定实体关系。
 - (4) 确定实体集的**键 (Key)**，用下划线在属性上表明键的属性组合。
 - (5) 确定**联系的类型**，用线将表示联系的菱形框连接到实体上，在线上注明联系的类型。
- 概念模型的设计可以采用**自底向上、分层制作**的策略，也就是先分层设计各个子系统的E-R图，然后再合并成整个系统的E-R图。在合并的过程中，设计人员需要消除重复的实体或属性、解决“冲突”问题。“冲突”包括：
- 属性冲突
 - 命名冲突
 - 结构冲突

- **逻辑设计**的主要任务是将概念模型转换为特定数据库管理系统支持的数据模型，也就是把概念设计阶段的E-R图转换为与选定DBMS支持的逻辑结构。
- 转换内容主要是将E-R图中的**实体、实体属性和实体间联系转换为关系模式**，即数据表。
- (一) E-R图向逻辑模型转换的原则
 - (1) 一个实体型转换为一个数据表。
 - (2) 一个1:1联系可以转换为一个独立的数据表，也可以作为任意一端对应关系模式的属性。
 - (3) 一个1:n联系可以转换为一个独立的数据表，也可以作为n端对应的数据表属性。
 - (4) 一个m:n联系转换为一个数据表，表的属性是与该联系相连的各实体的码以及联系本身的属性
 - (5) 三个或三个以上实体间的多元联系转换为一个数据表，表的属性是与该多元联系相连的各实体的码以及联系本身的属性
 - (6) 同一实体集的实体间的联系，可以转换成一个关系模式，也可以在实体中引入联系属性。
 - (7) 具有相同码的数据表可以合并。

□ (一) E-R图向逻辑模型转换的原则

- 例如，投递部和投递道段两个实体要转换成逻辑模型的投递部和投递道段两个数据表，实体之间的“组成”关系可以单独设计成数据表，也可以归入投递道段表。逻辑结构有两种方案。

- 第一种方案包括以下数据表：

- ① 投递部数据表（编号，名称，几何图形），
- ② 投递道段数据表（编号，名称，几何图形），
- ③ 投递部-道段组成表（部编号，道段编号）。

- 第二种方案包括以下数据表：

- ① 投递部数据表（编号，名称，几何图形），
- ② 投递道段数据表（编号，名称，所属投递部编号，几何图形）。

□ (二) 模型的优化

- 为了提高数据库管理性能，还应适当地修改、调整数据模型的结构，这就是数据模型的优化。关系数据模型的优化通常以**规范化理论**为指导，优化方法包括：
 - (1) **确定数据依赖**。根据需求分析阶段所得到的语义，分别写出每个数据表的各属性之间的数据依赖，以及不同数据表属性之间数据依赖
 - (2) **消除冗余的联系**。对于各个数据表之间的数据依赖进行极小化处理，消除冗余的联系。
 - (3) **确定所属范式**。按照数据依赖的理论对数据表逐一进行分析，考查是否存在部分函数依赖、传递函数依赖、多值依赖等，确定各数据表分别属于第几范式。
 - (4) 按照GIS业务对数据操作的要求，确定是否要**合并或分解**数据表。

- 空间数据库的**物理模型设计**是选择合适的存储结构、存取路径，设计合理的索引结构，完成数据库内模式的转换。
- 物理模型设计的**目标**是合理规划数据存储模式、有效利用存储空间，保证地理信息系统能够高效地访问空间数据。
- 空间数据库的物理设计包括以下步骤：
 - ① 分析空间数据库服务器的安装环境,针对空间数据量大、更新快、性能要求高的特点，合理设计数据库文件结构和存储策略
 - ② 分析数据表的存储结构，综合考虑空间数据的存取时间、存储空间利用率和维护代价等因素
 - ③ 合理规划数据索引
 - ④ 编写数据库脚本

□ 数据库的设计报告包括

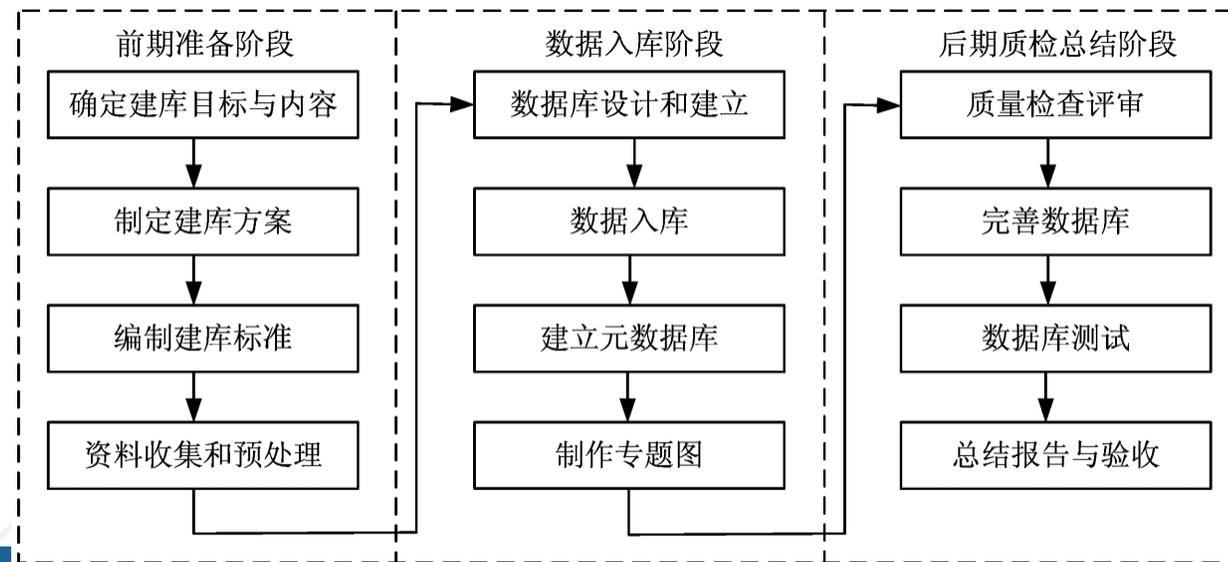
- 前言
- 数据库环境说明
- 数据库命名规则
- 数据库的概念设计
- 数据表清单及表结构字典
- 安全性设计

□ **空间数据库建设**是按照应用语义类型和几何维度类型**划分空间要素类别**、为每个要素类**创建数据表**、向数据表**录入相应空间数据**的过程。

□ **空间数据库建设内容**包括：

- 收集和整理GIS业务管理所需的数据、图件、文档等资料
- 按照相关规范要求，将资料分类处理、加工、转换、存入数据库
- 对存入结果进行数据质量检查和验收

□ **空间数据库建设流程**



- **空间数据库质量**是指空间数据在表达空间位置、专题属性、时间属性等信息的准确性、一致性、完整性等方面的特性。空间数据库的数据质量是一个相对的和动态的概念，空间数据质量标准都是在特定时代针对特定应用领域制定的。
- **空间数据库质量保证措施**
 - (1) 空间数据库需求分析阶段的质量保证措施
 - (2) 数据库设计阶段的质量保证措施
 - (3) 数据预处理过程中的质量保证措施
 - (4) 建库入库过程的质量保证
 - (5) 数据质量评估

6.5 空间数据库建设实例

□ 空间数据入库实例

- https://www.bilibili.com/video/BV1nh411379y/?spm_id_from=333.999.0.0&vd_source=733009f32a4b2ed8e1d58c11fabceb40



中南大學
CENTRAL SOUTH UNIVERSITY



感谢您的观看!