

Chap_01 数据库基础概述

- 人工管理 ⇒ 数据库、数据仓库、数据挖掘
 - 数据：描述事物的符号记录
 - 数据处理：将数据加工处理成为信息
 - 信息：数据根据需要进行加工处理后的结果
 - 人工管理：外存储器只有纸带、卡片和磁带，而没有磁盘
 - 文件系统阶段：文件只存储数据，不存储文件记录的结构描述信息。文件的建立、存取、查询、插入、删除、修改等所有操作，都要用程序来实现
 - 数据库系统
 - 分布式数据库系统
 - 面向对象数据库系统
 - 数据仓库
 - 数据挖掘
 - 大数据
-
- 数据库(Data Base, 简称为DB): 数据
 - 数据库管理系统(DataBase Management System, 简称为DBMS): 工具软件
- 数据定义功能
 - 数据操纵功能
 - 数据库运行控制功能
 - 数据库维护功能
 - 数据库通信功能
-
- 数据库系统(DataBase Sysytem, 简称为DBS): 整体系统, 包含 DB 和 DBMS
- 数据库、数据库管理系统、应用程序、管理人员
 - 特点：数据低冗余、共享性高；有统一的数据控制功能；数据独立性高(逻辑独立性、物理独立性)
-
- 数据库系统内部体系结构：
 - 三级模式：
 - 外模式(用户级数据库): 数据库用户的数据视图
 - 概念模式(概念级、应用数据库): 所有用户的公共数据视图
 - 内模式(物理级数据库): 存储模式，一个数据库只有一个内模式
 - 二级映射：
 - 概念模式到内模式的映射 ⇒ 物理独立性
 - 外模式到概念模式的映射 ⇒ 逻辑独立性
-
- 数据模型：数据特征的抽象；数据结构、数据操作、数据约束条件

- 概念数据模型(概念模型): 面向数据库用户的现实世界的数据库模型
 - 数据模型的基础
 - 主要用来描述现实世界的概念化结构
 - 与具体的计算机平台无关、与具体的DBMS无关
- 逻辑数据模型(数据模型):
 - 主要用来描述数据库中数据的表示方法和数据库结构的实现方法
 - 与具体的DBMS有关
 - 层次、网状、关系、面向对象
- 物理数据模型(物理模型): 面向计算机物理表示的模型
 - 描述数据在存储介质上的组织结构的数据模型
- E-R模型
 - 基本概念
 - 实体: 客观存在并可以区别的事物
 - 属性: 描述实体的特性
 - 关键字: 能唯一标识实体的属性或属性的组合
 - 域: 属性的取值范围
 - 实体型: 属性的集合表示一个实体的类型
 - 实体集: 同类型的实体的集合
 - 实体之间的联系
 - 一对一联系(1 : 1)
 - 一对多联系(1 : n)
 - 多对多联系(m : n)
- 关系数据库: 基于关系模型的数据库
 - 关系
 - 一个关系就是一个二维表, 每一个关系都有一个关系名
 - 二维表含有几列就称为几元关系
 - 属性: 二维表的一列称为一个属性
 - 元组: 每一行称为一个元组或记录
 - 分量: 元组中的一个属性值
 - 域: 属性的取值范围
 - 候选关键字: 唯一标识(一个属性或几个属性的集合)
 - 主关键字(主键)
 - 外部关键字(外键)
 - 主属性: 包含在任一候选关键字中的属性
 - “关系”是“元组”的集合, “元组”是属性值的集合
- 完整性约束
 - 实体完整性约束(主键不能为空值)、参照完整性约束(关系中的外键取值只能是关联关系中的某个主键或空值)、用户完整性约束(具体应用中数据的语义要求)

- 关系完整性约束(实体完整性约束、参照完整性约束): 关系模型必须满足的完整性的约束条件
- 用户完整性约束是应用领域需要遵循的约束条件
- 关系规范化
 - 第一范式: 关系模式R的所有属性都是最基本的、不可再分的数据项
 - 第二范式: 非主属性都完全依赖于主键
 - 第三范式: 非主属性对任何逐渐都不存在传递依赖
- 关系运算
 - 并、交、差、广义笛卡尔积等运算
 - 相同的结构
 - 选择 σ 、投影 Π 、联接、除运算
- 数据库设计
 - 需求分析、概念设计、逻辑设计、物理设计