

数据库系统概论

An Introduction to Database System

中国人民大学信息学院

**School of Information,
Renmin University of China
2016.9**

第七章 数据库设计

7.1 数据库设计概述

7.2 需求分析

7.3 概念结构设计

7.4 逻辑结构设计

7.5 物理结构设计

7.6 数据库的实施和维护

7.7 小结



7.2 需求分析

7.2.1 需求分析的任务

7.2.2 需求分析的方法

7.2.3 数据字典



7.2 需求分析

❖ 什么是需求分析——分析用户的要求

- 是设计数据库的**起点**

❖ 需求分析的重要性

- 结果是否准确地反映了用户的实际要求，将直接影响到后面各个阶段的设计，并影响到设计结果是否合理和实用

❖ 需求分析常常被忽视

- 设计人员认为这是软任务，急于进行具体设计
- 用户嫌麻烦
- 领导不重视



7.2.1 需求分析的任务

- ❖ 充分了解原系统（手工系统或计算机系统）工作概况
- ❖ 详细调查要开发应用系统的组织/部门/企业等
- ❖ 明确用户的各种需求

确定新系统的功能

注意：新系统今后可能的扩充和改变

调查的重点是“数据”和“处理”，获得用户对数据库的要求

1. 信息要求

- 用户需要从数据库中**获得信息**的内容与性质
- 由信息要求可以**导出数据要求**，即在数据库中需要存储哪些数据

2. 处理要求

- 用户要什么**处理功能**、对**处理性能**、**处理方式**、**处理周期**等的要求
(批处理 / 联机处理 / 发布处理 / 每月一次 /)

3. 安全性与完整性要求

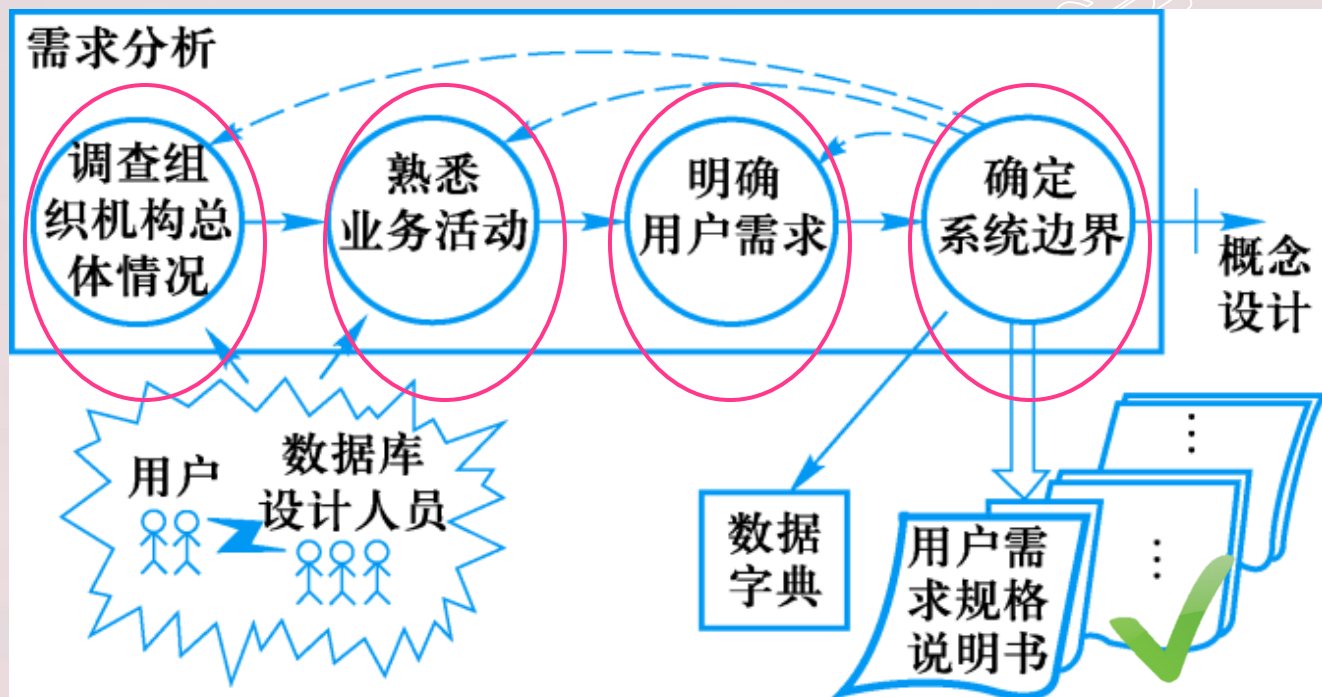
❖ 确定用户需求的难点

- 用户缺少计算机知识，不能准确地表达自己的需求，提出的需求往往不断地变化。
- 设计人员缺少用户的专业知识，不易理解用户的真正需求，甚至误解用户的需求。

❖ 解决方法

- 设计人员必须**不断深入地**与用户进行交流，才能逐步确定用户的实际需求

7.2.2 需求分析的方法



7.2 需求分析

7.2.1 需求分析的任务

7.2.2 需求分析的方法

7.2.3 数据字典



7.2.3 数据字典

❖ 什么是数据字典？

数据字典是关于数据库中数据的描述，称为元数据。

它不是数据本身，而是数据的数据。

❖ 数据字典在需求分析阶段建立，在数据库设计过程中不断修改、充实、完善。

❖ 数据字典是进行详细的数据收集和分析所获得的主要结果。

注意：与 DBMS 中的数据字典的区别和联系



数据字典（续）

❖ 数据字典的内容

- 数据项
- 数据结构
- 数据流
- 数据存储
- 处理过程

❖ 数据项是数据的最小组成单位

❖ 若干个数据项可以组成一个数据结构

❖ 通过对数据项和数据结构的定义来描述数据流、数据存储的逻辑内容



1. 数据项

❖ 数据项是不可再分的数据单位

❖ 数据项描述=

{数据项名,数据项含义说明,别名,数据类型,长度,取值范围,取值含义,
与其他数据项的逻辑关系,数据项之间的联系}

❖ 关系规范化理论为指导,用数据依赖的概念分析和抽象数据项之间的联系——函数依赖

❖ “取值范围”、“与其他数据项的逻辑关系”

定义了数据的完整性约束条件,是模式设计、完整性检查条件、
触发器、存储过程的依据



实例

例：学生学籍管理子系统的数据字典。

数据项，以“学号”为例：

数据项： 学号

含义说明：唯一标识每个学生

别名： 学生编号

类型： 字符型

长度： 9

取值范围：0000 00 000至9999 99 999

取值含义：前4位标别该学生入学年份，第5第6位所在专业系编号，
后3位按顺序编号，例如201615008

与其他数据项的逻辑关系：学号的值确定了其他数据项的值

数据项描述=

{数据项名,数据项含义说明,别名,数据类型,长度,取值范围,取值含义,与其他数据项的逻辑关系,数据项之间的联系}



2. 数据结构

- ❖ 数据结构反映了数据之间的组合关系。
- ❖ 一个数据结构可以由若干个数据项组成，也可以由若干个数据结构组成，或由若干个数据项和数据结构混合组成。
- ❖ 数据结构描述= {数据结构名, 含义说明, 组成: {数据项或数据结构} }

以“学生”为例，“学生”是该系统中的一个核心数据结构：

数据结构：学生

含义说明：学籍管理子系统的主体数据结构，
定义了一个学生的有关信息

组成：学号，姓名，性别，年龄，所在系，年级



3. 数据流

❖ 数据流是数据结构在系统内部传输的路径。

❖ 对数据流的描述

数据流描述={ 数据流名,说明,数据流来源,数据流去向,
组成: {数据结构}, 平均流量,高峰期流量 }

- 数据流来源: 说明该数据流来自哪个处理过程/数据存储
- 数据流去向: 说明该数据流将到哪个处理过程/数据存储去
- 平均流量: 在单位时间(每天、每周、每月等)里的传输次数
- 高峰期流量: 在高峰时期的数据流量



实例

数据流“体检结果”可如下描述：

数据流： 体检结果

说明： 学生参加体格检查的最终报告

数据流来源： 体检（处理过程）

数据流去向： 批准（处理过程）

组成： { 学号, {血常规}, {尿常规}, {血液生化}, {心电图},
{B超}, {其他体检} }

平均流量： 每天200

高峰期流量： 每天400

数据流描述={ 数据流名,说明,数据流来源, 数据流去向,
组成:{数据结构}, 平均流量,高峰期流量 }



4. 数据存储

❖ 数据存储

是数据结构停留或保存的地方，也是数据流的来源和去向之一。

❖ 数据存储描述={数据存储名,说明,编号,输入的数据流,输出的数据流,组成:
{数据结构}, 数据量, 存取频度, 存取方式}

- 存取频度：每小时、每天或每周存取次数，每次存取的数据量等信息
- 存取方法：批处理 / 联机处理；检索 / 更新；顺序检索 / 随机检索
- 输入的数据流：数据来源
- 输出的数据流：数据去向



实例

数据存储“学生登记表”可如下描述：

数据存储： 学生登记表

说明： 记录学生的基本情况

流入数据流： 每学期5000

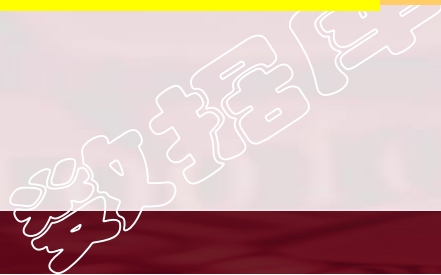
流出数据流： 每学期5000

组成： {学号，姓名，性别，年龄，所在系，年级，{学习成绩}，{体检结果}，
{奖惩记录} }

数据量： 每年10000张

存取方式： 随机存取+按照专业系/班级打印

数据存储描述={数据存储名,说明,编号,输入的数据流 ,输出的数据流, 组成: {数据结构}, 数据量, 存取频度, 存取方式}



5. 处理过程

❖ 处理过程

具体处理逻辑一般用判定表或判定树来描述。

❖ 数据字典中只需要描述处理过程的说明性信息

❖ 处理过程描述={ 处理过程名, 说明, 输入:{数据流}, 输出:{数据流}, 处理:{简要说明} }

❖ 简要说明：说明该处理过程的功能及处理要求

- 功能：该处理过程用来做什么

- 处理要求：

处理频度要求，如单位时间里处理多少事务，多少数据量、响应时间要求等

- 处理要求是物理设计的输入及性能评价的标准



实例

处理过程“分配宿舍”可如下描述：

处理过程：分配宿舍

说明：为所有新生分配学生宿舍

输入：学生，宿舍

输出：宿舍安排

处理：在新生报到后，为所有新生分配学生宿舍。

要求同一间宿舍只能安排同一年级同一性别的学生

一个学生只能安排在一个宿舍中。每个学生的居住面积不小于6平方米。

安排新生宿舍其处理时间应不超过15分钟。

处理过程描述={ 处理过程名, 说明, 输入:{数据流}, 输出:{数据流}, 处理:{简要说明} }

需求分析小结

- ❖ 把需求收集和分析作为数据库设计的第一阶段是十分重要的。
- ❖ 第一阶段收集的基础数据用数据字典来描述是下一步进行概念设计的基础。
- ❖ 强调两点
 - (1) 设计人员应充分考虑到可能的扩充和改变，使设计易于更改，系统易于扩充
 - (2) 必须强调用户的参与，领导的重视



