

项目 1 数据库系统与数据库设计认知

项目导读

数据库技术是现代信息科学与技术的重要组成部分,是计算机数据处理与信息管理系统核心。为了更好地系统掌握数据库的知识,需要先对数据库系统和数据库设计有基本的认知,主要包括数据描述与数据处理、数据管理技术的发展历程、数据库系统、数据库管理系统、数据模型和应用系统数据库设计步骤等内容。

项目素质目标

树立以人为本、以用户需求为核心的理念,理解物质世界的普遍联系性,培养善于发现问题和解决问题的能力,自觉养成认真细致的学习和工作态度。

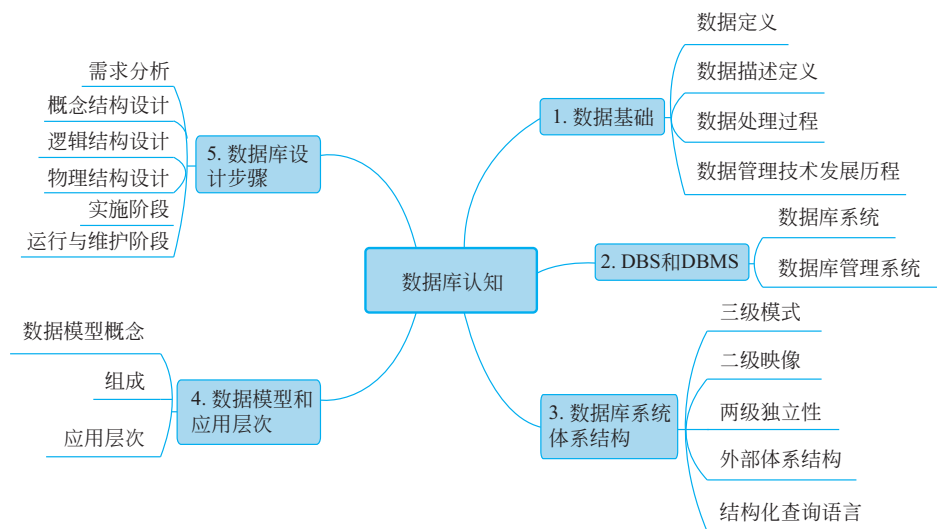
项目知识目标

了解数据管理技术的发展,掌握数据库系统的概念,理解结构化查询语言,理解数据模型的概念,了解逻辑模型的结构分类,理论关系模型和应用系统数据库设计步骤。

项目能力目标

掌握数据模型的理论知识,掌握应用系统数据库设计步骤,了解项目总体实施流程。

项目导图



任务 1-1 数据管理技术与数据库系统

任务描述

本任务主要讲解数据库系统涉及的基础知识,通过本任务的学习,可以对数据库有最基本的认识。依据对数据库管理技术、数据库系统知识点的理解,为数据库应用系统的设计、开发、使用和管理奠定基础,本任务进度如表 1-1 所示。

表 1-1 数据管理技术与数据库系统任务进度表

任务描述	任务下发时间	预期完成时间	任务负责人	版本号
掌握数据描述的概念	8 月 23 日	8 月 24 日	张良	V1.0
掌握数据处理技术发展历程	8 月 25 日	8 月 26 日	张良	V1.0
了解数据库系统体系结构	8 月 27 日	8 月 28 日	李旭阳	V1.0
了解结构化查询语言	8 月 29 日	8 月 30 日	李旭阳	V1.0

任务分析

对于初学者来说,数据库技术涉及的基本概念较多,有些概念较抽象,不易理解。本任务主要介绍数据描述的两种不同形式及其定义,数据处理的概念、数据处理的八个主要方面和数据处理的作用,数据管理技术经历的人工管理、文件系统和数据库系统三个阶段等知识点。

任务目标

- 素质目标:培养对抽象概念的理解能力和善于发现问题、解决问题能力。
- 知识目标:理解数据描述的定义、数据处理的定义和数据处理过程,了解数据管理技术发展历程,掌握数据库系统的构成,了解数据库管理系统,了解数据库系统的体系结构和结构化查询语言。
- 能力目标:能够绘制出数据库基本概念之间的关系图,绘制出本任务的思维导图,发现各概念之间的关系。

任务实施

步骤 1 理解数据描述定义

常用的数据描述有物理数据描述和逻辑数据描述两种不同的形式。物理数据描述是对数据在存储设备中的存储方式的描述,物理数据是实际存放在存储设备中的数据,物理数据通常是指基础数据或元数据,供业务查询和分析使用。逻辑数据描述指对程序员或用户可以操作的数据形式的描述,是抽象的概念化数据。数据描述在数据处理中涉及不同的范畴。从事物的特性到计算机中的具体表示,共经历了概念设计中的数据描述、逻辑设计中的数据

描述和物理存储介质中的数据描述三个阶段。

步骤 2 理解数据处理的定义和数据处理过程

数据处理(data processing)主要是指对数据的采集、存储、检索、加工、变换和传输。数据处理的基本目的是从大量的、可能是杂乱无章的、难以理解的数据中抽取并推导出对于某些特定场合和特定人群来说是有价值、有意义的数据。

数据处理是一项经计算机收集、记录、加工数据,产生新的信息形式的技术。计算机数据处理主要包括以下几个方面。

- (1) 数据采集:采集所需的信息。
- (2) 数据转换:把信息转换成机器能够接收的形式。
- (3) 数据分组:指定编码形式,按有关信息进行有效的分组。
- (4) 数据组织:整理数据或用某些方法安排数据,以便进行处理。
- (5) 数据计算:进行各种算术和逻辑运算,以便得到进一步的信息。
- (6) 数据存储:将原始数据或计算结果保存起来,供以后使用。
- (7) 数据检索:按用户的要求找出有用的信息。
- (8) 数据排序:把数据按一定要求排成次序。

数据处理的过程大致分为数据的准备、处理和输出三个阶段。在数据准备阶段,可以将数据输入各种不同的存储介质中。早期的存储介质主要包括穿孔卡片、纸带、软盘和磁带等,现阶段的存储介质主要是指光盘、硬盘、闪存卡等。数据准备阶段也称数据的录入阶段。数据录入以后,由计算机对数据进行处理,为此预先要由用户编制程序并把程序输入计算机中,计算机是按程序的指示和要求对数据进行处理。所谓处理,就是指上述八个方面工作中的一个或若干个的组合。最后输出的是各种文字和数字的表格。

数据处理系统已广泛用于各种企事业单位,内容涉及薪金支付、票据收发、信贷、库存管理、生产调度、计划管理和销售分析等多个方面。它能产生操做报告、金融分析报告和统计报告等。数据处理技术涉及文件系统、数据库管理系统、分布式数据处理系统等方面的技术。

如今,数据和信息已经成为人类社会中极其宝贵的资源。随着云计算技术、大数据技术和人工智能技术的不断发展,数据处理技术将会进一步发展,它也将进一步推动整个信息化社会的发展。

步骤 3 了解数据管理技术发展历程

数据管理技术先后经历了人工管理、文件系统和数据库系统三个阶段。

1) 人工管理阶段(初等数据文件阶段)

- 经历时期:20 世纪 50 年代中期以前,计算机主要用于科学计算。
- 硬件状况:外存只有纸带、卡片、磁带,没有磁盘等直接存取设备。
- 软件状况:没有操作系统,没有管理数据的软件。
- 数据处理方式:批处理。

2) 文件系统阶段(独立文件管理系统)

- 经历时期:20 世纪 50 年代后期到 60 年代中期。
- 硬件方面:拥有磁盘、磁鼓等直接存取设备。
- 软件方面:操作系统中已经有专门的数据管理软件,一般称为文件系统。



数据管理技术
发展历程

课堂笔记

- 数据处理方式:批处理,联机实时处理。
- 3) 数据库系统阶段
- 经历时期:20 世纪 60 年代后期至今。
 - 硬件方面:拥有大容量磁盘,硬件价格下降。
 - 软件方面:软件价格上升,为编制和维护系统软件及应用程序的成本相对增加。
 - 数据处理方式:使用统一管理数据的专门软件系统,即数据库管理系统。

答疑解惑

问:目前数据管理技术处于哪个阶段?

答:数据库系统阶段,当前对数据的管理更具多样性,有更多不同类型的数据库系统。

步骤 4 掌握数据库系统构成

数据库系统(database system,DBS)是采用数据库技术的计算机系统,是由数据库、数据库管理系统、数据库管理人员、支持数据库系统的硬件和软件(应用开发工具、应用系统等)以及用户构成的运行实体。下面从硬件、软件和人员三个角度来说明。

(1) 硬件:构成计算机系统的各种物理设备,包括存储所需的外部设备。硬件的配置应满足整个数据库系统的需要,要求有足够大的空间存放操作系统、数据库管理系统的核心模块、数据缓冲区和应用程序,而且需要较高的通道能力。

(2) 软件:主要包括操作系统、数据库管理系统及应用程序以及核心开发工具。数据库管理系统是数据库系统的核心软件,具有数据库接口的高级语言及其编译系统,便于开发应用程序,科学地组织和存储数据,高效地获取和维护数据。

(3) 人员主要有以下四个类别。

① 系统分析员和数据库设计人员:系统分析员负责应用系统的需求分析和规范说明,他们和用户及数据库管理员一起确定系统的硬件配置,并参与数据库系统的概要设计;数据库设计人员负责数据库中数据的确定、数据库各级模式的设计。

② 应用程序员:负责编写使用数据库的应用程序。这些应用程序可对数据进行检索、添加、删除或修改。

③ 最终用户:最终用户会利用系统的接口或查询语言访问数据库。

④ 数据库管理员(data base administrator,DBA):主要负责数据库的总体管理和控制。通常 DBA 的具体职责主要有:管理数据库中的信息内容和结构;定义数据库的存储结构和存取策略;定义数据库的安全性要求和完整性约束条件;监控数据库的使用和运行,负责数据库的性能改进;对数据库进行重组和重构,以提高系统的性能等。

步骤 5 了解数据库管理系统

数据库管理系统(database management system,DBMS)是一种操纵和管理数据库的大型软件,用于建立、使用和维护数据库。它对数据库进行统一的管理和控制,以保证数据库的安全性和完整性。用户通过 DBMS 访问数据库中的数据,数据库管理员也通过 DBMS 进行数据库的维护工作。它可以支持多个应用程序和用户用不同的方法在相同或不同的时刻去建立、修改和查询数据库。大部分 DBMS 提供数据定义语言(data definition language,DDL)和数据操纵语言(data manipulation language,DML),供用户定义数据库的模式结构与权限约束,实现对数据的插入、删除、更新和查询等操作,它的主要功能包括以下几个

方面。

(1) 数据定义:DBMS 提供数据定义语言,供用户定义数据库的三级模式结构、两级映像以及完整性约束和保密限制等约束。DDL 主要用于建立、修改数据库的库结构。DDL 所描述的数据库结构仅仅给出了数据库的框架,数据库的框架信息存放在数据字典(data dictionary)中。

(2) 数据操作:DBMS 提供数据操纵语言,供用户实现对数据的插入、删除、修改、查询等操作。

(3) 数据库的运行管理:包括多用户环境下的并发控制、安全性检查和存取权限控制、完整性检查和执行、运行日志的组织管理、事务的管理和自动恢复。这些功能保证了数据库系统的正常运行。

(4) 数据组织、存储与管理:包括数据字典、用户数据、存取路径等,需确定以何种文件结构和存取方式在存储级上组织这些数据,如何实现数据之间的联系。数据组织和存储的基本目标是提高存储空间的利用率,选择合适的存取方法以提高存取效率。

(5) 数据库的保护:数据库中的数据是信息社会的战略资源,所以数据的保护至关重要。DBMS 对数据库的保护通过四个方面来实现:数据库的恢复、数据库的并发控制、数据库的完整性控制和数据库安全性控制。DBMS 的其他保护功能还有系统缓冲区的管理、数据存储的某些自适应调节机制等。

(6) 数据库的维护:包括数据库的数据载入、转换、转储,数据库的重组合、重构以及性能监控等功能,这些功能分别由各个应用程序来完成。

(7) 通信:DBMS 具有与操作系统联机处理、分时系统及远程作业输入等操作的相关接口,负责处理数据的传送。网络环境下的数据库系统,其 DBMS 还具有与网络中其他软件系统的通信功能以及数据库之间的互操作功能。

常见的数据库管理系统主要有四种类型:文件管理系统,层次模型数据库,网状模型数据库和关系模型数据库,其中关系模型数据库的应用最为广泛。

步骤 6 了解数据库系统的体系结构

1) 三级模式

1975 年,美国国家标准学会下属的标准计划和需求委员会为数据库管理系统建立了三级模式结构,即外模式、概念模式和内模式。

(1) 外模式:又称关系子模式或用户模式,是数据库用户看见的局部数据的逻辑结构和特征的描述,即应用程序所需要的那部分数据库结构。外模式是应用程序与数据库系统之间的接口,是保证数据库安全性的一个有效措施。用户可使用数据定义语言和数据操纵语言来定义数据库的结构和对数据库进行操纵。对于用户而言,只需要按照所定义的外模式进行操作,而无须了解概念模式和内模式等的内部细节。一个数据库可以有多个外模式。

(2) 概念模式:又称模式、关系模式、逻辑模式,是数据库整体逻辑结构的完整描述,包括概念记录模型、记录长度之间的联系、所允许的操作,以及数据的完整性、安全性约束等数据控制方面的规定。概念模式位于数据库系统模式结构的中间层,不涉及数据的物理存储细节和硬件环境,与应用程序、开发工具及程序设计语言无关。一个数据库只能有一个概念模式。

(3) 内模式:又称存储模式,是数据库内部数据存储结构的描述。它定义了数据库内部

记录类型、索引和文件的组织方式以及数据控制方面的细节。一个数据库只能有一个内模式。

2) 二级映像

外模式/模式映像:模式描述的是数据的全局逻辑结构,外模式描述的是数据的局部逻辑结构,同一个模式可以有任意多个外模式。对于每个外模式,数据库系统都有一个外模式/模式映像,它定义了该外模式与模式之间的对应关系。这些映像定义通常包含在各自外模式的描述中。

模式/内模式映像:它定义了数据库全局逻辑结构与存储结构之间的对应关系,是唯一的。该映像定义通常包含在模式描述中。

3) 两级数据独立性

数据独立性是指应用程序和数据库的数据结构之间相互独立、不受彼此的影响,包括逻辑数据独立性和物理数据独立性。

(1) 逻辑数据独立性:当模式改变时,由数据库管理员对各个外模式/模式映像做相应改变,可以使外模式保持不变。应用程序是依据数据的外模式编写的,因而应用程序不必修改,从而保证了数据与应用程序的逻辑独立性。

(2) 物理数据独立性:当数据库的存储结构改变时,由数据库管理员对模式/内模式映像做相应改变,可以使模式保持不变,因而应用程序也不必修改,保证了数据与应用程序的物理独立性。

特定的应用程序是在外模式描述的数据结构上编写的,它依赖于特定的外模式,与数据库的模式和存储结构相独立。不同的应用程序可以共用同一外模式。数据库的两级映像保证了数据库外模式的稳定性,从而从底层保证了应用程序的稳定性,除非应用需求本身发生变化,否则应用程序一般不需要修改。

4) 外部体系结构

从数据库最终用户的角度看,数据库系统的结构分为集中式结构(单用户结构、主机/终端结构)、分布式结构(客户机/服务器结构)和多层应用结构,这是数据库系统外部的体系结构。

(1) 单用户应用结构:运行在个人计算机上的结构模式。属于单用户 DBMS 的产品有 Microsoft Access、Paradox、Fox Pro 系列等,它们基本实现了 DBMS 应该具有的功能。单用户 DBMS 的功能在数据的一致性维护、完整性检查及安全性管理上是不完善的。

(2) 主机/终端结构:以大型主机为中心的结构模式,也称为分时共享模式,是面向终端的多用户计算机系统。该结构以一台主机为核心,将操作系统、应用程序、DBMS、数据库等数据和资源均放在该主机上,所有的应用处理均由主机承担,每个与主机相连接的终端都视为主机的一种 I/O 设备。由于是集中式管理,主机的任何错误都有可能导致整个系统瘫痪。因此,这种结构对系统的主机性能要求比较高,维护费用也较高。

(3) 客户机/服务器(Client/Server, C/S)结构:随着计算机网络的广泛使用而出现的结构模式。它将一个数据库系统分解为客户机(即前端)、应用程序和服务器(即后端)三部分,通过网络连接应用程序和服务器。由于 C/S 结构的本质是通过服务功能的分布实现分工服务,因而又称分布式服务模式。人们将 C/S 称为二层结构的数据库应用模式。

(4) 多层数据库应用结构:将应用程序放在服务器端执行,客户机端安装统一的前端运行

环境,通常是浏览器(browser),在客户机和服务器之间增加一层用于转换的服务器,形成三层结构的数据库应用模式,这就是互联网环境下数据库的应用模式。三层结构是由二层(C/S)结构扩展而来的,这种三层结构也称为浏览器/Web 服务器/数据库服务器(B/W/S)结构。

步骤 7 了解结构化查询语言

为了更好地提供从数据库中简单高效读取数据的方法,1974 年博伊斯(Boyce)和钱伯林(Chamberlin)提出了一种称为 SEQUEL 的结构化查询语言。1976 年,这种语言在 IBM 公司研发的关系数据库系统 System R 上实现,将其修改为 SEQUEL 2,即目前的结构化查询语言(structured query language,SQL)。由于它具有功能丰富、使用方便灵活、语言简洁易学等突出的优点,深受计算机工业界和计算机用户的欢迎。1980 年 10 月,经美国国家标准学会(American National Standards Institute,ANSI)的数据库委员会批准,将 SQL 作为关系数据库语言的美国标准,同年公布了标准 SQL。

SQL 集数据查询(data query)、数据定义(data definition)、数据操纵(data manipulation)和数据控制(data control)功能于一体,充分体现了关系数据库语言的特点。

SQL 的核心部分相当于关系代数,但又具有关系代数所没有的许多特点,如聚集、数据库更新等。它是一个综合的、通用的、功能极强的关系数据库语言。其主要有以下四个特点。

(1) 综合统一。SQL 不是某个特定数据库供应商专有的语言,所有关系数据库都支持它。SQL 的风格和语法都是统一的,可以独立完成数据库生命周期中的全部活动,包括定义关系模式、录入数据,建立、查询、更新、维护数据库,以及数据库重构、安全性控制等一系列操作,这就为数据库应用系统的开发提供了良好的环境。

(2) 以同一种语法结构提供两种使用方式。SQL 有两种使用方式:一种是联机交互使用,这种方式下的 SQL 实际上是作为自含式语言使用的;另一种是嵌入某种高级程序设计语言(如 C 语言等)中去使用。前一种方式适合于非计算机专业人员使用,后一种方式适合于专业计算机人员使用。这种以统一的语法结构提供两种不同使用方法的特点,为用户带来了极大的灵活性与方便性。

(3) 高度非过程化。SQL 是一种第四代语言(fourth-generation language,4GL),用户只需要提出“干什么”,无须具体指明“怎么干”,像存取路径选择和具体处理操作等均由系统自动完成。这不但大大减轻了用户负担,而且有利于提高数据独立性。

(4) 语言简洁,易学易用。SQL 不仅功能很强,而且语言十分简洁,要完成核心功能,只需使用 SELECT、CREATE、INSERT、UPDATE、DELETE、GRANT 等几个命令。SQL 的语法接近英语口语,所以用户很容易掌握。SQL 目前已成为应用最广泛的关系数据库语言。

学习小贴士

SQL 语句主要是对结构化数据进行查询和处理的。当对半结构化和非结构化数据处理时,可以使用 NoSQL,它是一种更加灵活的数据模型。

知识点解析

1. 数据的基本概念

数据(data)是对现实世界的描述。由于计算机不能直接处理现实世界中的具体事物,



微课:结构化
查询语言

课堂笔记

因此必须先把具体事物转换成计算机能够处理的数据。在计算机系统中,各种字母、数字符号的组合、语音、图形、图像等统称为数据,数据经过加工后就成为信息。在计算机科学中,数据是指所有能输入计算机并被计算机程序处理的符号介质的总称,是用于输入计算机进行处理,具有一定意义的数字、符号、字母和各种文字集合的通称。

2. 数据管理技术各个发展阶段的特点

1) 人工管理数据的特点

数据不能长期保存;没有专用的软件管理数据;数据冗余度大,数据无法共享;数据不具有独立性。

2) 文件系统管理数据的特点

数据能够长期保存;由专门的软件即文件系统进行数据管理,文件系统把数据组织成相互独立的数据文件,利用“按文件名访问,按记录存取”的管理技术,可以对文件进行修改、插入、删除等操作;文件系统实现了记录内的结构性,但是整体无结构;数据共享性差,冗余度大。在文件系统中,一个文件基本上对应一个应用程序,即文件仍然是面向应用的,数据独立性差。一旦数据的逻辑结构改变,就必须修改应用程序和文件结构的定义。如果应用程序改用不同的高级语言等,将引起文件的数据结构改变,因此数据与程序之间仍缺乏独立性。

3) 数据库管理数据的特点

数据通过数据库管理系统来管理,数据可以面向整个应用系统,数据的共享性高,冗余度小,数据具有高度的物理独立性和逻辑独立性。

实战演练

1. 任务工单:了解数据描述与数据处理技术发展历程

组员 ID		组员姓名		所属项目组	
硬件配置	CPU:2.3GHz 及以上双核或四核;硬盘:150GB 及以上;内存:8GB;网卡:千兆网卡				
操作系统	Windows 7/Windows 10 或更高版本		软件系统	WPS Office	
任务执行前准备工作	检测计算机软硬件环境是否可用		☐ 可用 ☐ 不可用	不可用注明理由:	
	检测操作系统环境是否可用		☐ 可用 ☐ 不可用	不可用注明理由:	
	检测 WPS Office 软件的启动是否正常		☐ 正常 ☐ 不正常	不正常注明理由:	
掌握数据库系统中的基本概念	了解数据描述定义		完成度: ☐ 未完成 ☐ 部分完成 ☐ 全部完成		
	掌握数据处理概念		完成度: ☐ 未完成 ☐ 部分完成 ☐ 全部完成		
	掌握数据处理过程		完成度: ☐ 未完成 ☐ 部分完成 ☐ 全部完成		
	熟悉数据处理的各个阶段		完成度: ☐ 未完成 ☐ 部分完成 ☐ 全部完成		

续表

掌握数据库系统中的基本概念	掌握数据管理技术经历的阶段	完成度: ☐ 未完成 ☐ 部分完成 ☐ 全部完成
	掌握人工管理数据的特点	完成度: ☐ 未完成 ☐ 部分完成 ☐ 全部完成
	掌握系统管理数据的特点	完成度: ☐ 未完成 ☐ 部分完成 ☐ 全部完成
	了解数据库系统阶段的特点	完成度: ☐ 未完成 ☐ 部分完成 ☐ 全部完成
任务未成功的处理方案	采取的具体措施:	执行处理方案的结果:
备注说明	填写日期:	其他事项:

2. 任务工单:理解数据库系统体系结构和结构化查询语言

在完成上述工单“了解数据描述与数据处理技术发展历程”后,扫描右侧二维码下载并完成此工单。



任务工单:理解数据库系统体系结构和结构化查询语言

任务评价

组员 ID		组员姓名		所属项目组			
评价栏目	任 务 详 情	评 价 要 素	分 值	评价主体			
				学生 自评	小组 互评	教师 点评	
数据描述与 数据处理技术 发展历程的 了解情况	了解数据描述定义	是否完全了解	5				
	掌握数据处理概念	是否完全了解	3				
	掌握数据处理过程	是否完全了解	5				
	熟悉数据处理的各个阶段	是否完全了解	5				
	掌握数据管理技术经历的阶段	是否完全了解	4				
	掌握人工管理数据的特点	是否完全了解	5				
	掌握系统管理数据的特点	是否完全了解	4				
	了解数据库系统阶段的特点	是否完全了解	4				
数据库系统体 系结构和结构 化查询语言的 理解情况	理解数据库系统构成	是否完全理解	5				
	掌握数据库系统中有哪几类人员	是否完全掌握	5				
	理解什么是数据库管理系统	是否完全理解	5				
	了解数据库管理系统(DBMS) 主要功能	是否完全了解	5				
	理解数据库系统的体系结构的三级模式	是否完全理解	5				

续表

评价栏目	任务详情	评价要素	分值	评价主体		
				学生自评	小组互评	教师点评
数据库系统体系结构和结构化查询语言的理解情况	理解数据库系统的体系结构的两级映像	是否完全理解	5			
	理解数据库系统的体系结构的两级数据独立性	是否完全理解	5			
	理解数据库系统的外部体系结构	是否完全理解	5			
	理解什么是结构化查询语言	是否完全理解	5			
	理解结构化查询语言的特点	是否完全理解	5			
掌握熟练度	知识结构	知识结构体系形成	2			
	准确性	概念和基础掌握的准确度	2			
团队协作能力	积极参与讨论	积极参与和发言	2			
	对项目组的贡献	对团队的贡献值	2			
职业素养	态度	是否认真细致、遵守课堂纪律、学习积极、具有团队协作精神	2			
	操作规范	是否有实训环境保护意识,实训设备使用是否合规,操作前是否对硬件设备和软件环境检查到位,有无损坏机器设备的情况,能否保持实训室卫生	3			
	设计理念	是否突显以人为本的设计理念	2			
总分			100			

拓展训练

1. 数据记录 and 文件

(1) 为了方便用户从大量的数据中查询出指定的内容,这些数据通常以文件的形式存储在磁盘或者其他外部存储设备上。一个文件是一个被命名的、存储在设备上的信息的线性字节流。文件在需要的时候可以读取这些信息或者写入新的信息。计算机中通常存有很多文件,需要对诸多的文件进行归类、存储、查找等操作。早在 1965 年开发的 Multics (UNIX 的前身)就详细地设计了文件系统,这使得文件系统成为多用户单结点操作系统的重要组成部分。存储在文件中的信息必须是永久的,也就是说,它不会因为应用的创建/终止而受到影响。只有当用户显式地删除它时,文件才会消失。

对文件的管理,包括文件的结构以及命名、存取、使用、保护和实现方法,这些称为文件