

海洋科技档案管理中的数据库设计

崔爱菊¹, 周玉斌¹, 苏天赞^{1,2}

(1. 国家海洋局第一海洋研究所 青岛 266061; 2. 海洋沉积与环境地质国家海洋局重点实验室 青岛 266061)

摘 要: 文章对目前海洋科技档案管理的现状和存在的问题进行了综合的论述和分析, 针对海洋科技档案管理的特点和具体的应用需求, 从数据库管理系统、数据分类、概念结构、逻辑结构和物理结构等方面对海洋科技档案数据库 (marine science & technology archives management system, MSTADB) 进行了设计和建设, 旨在为用户提供便捷、高效的海洋科技档案管理、查询和获取服务。

关键词: 海洋; 档案管理; ORACLE; 数据库

海洋是一个沿海国家社会经济发展的重要空间与资源基地, 合理开发、切实保护海洋已经成为关系沿海各国生存、发展和强盛的重大战略问题。我国拥有广阔的管辖海域, 环境条件优越, 海洋资源丰富, 开发利用海洋来缓解我国经济发展中能源、资源的短缺与发展空间的不足具备现实需求的必要性和可能性。

新中国成立以来, 我国政府在海洋调查和海洋科学考察方面做了大量的工作。其调查范围从近海逐步扩展到大洋, 调查方式从海面观测逐步发展到航空航天遥感、海面观测和水下探查。如, 1958年开始的全国海洋综合普查; 20世纪80—90年代的全国海岸带和海涂资源综合调查和全国海岛资源综合调查以及近年来的“908”全国综合海洋调查等。在这些科研工作中归档了大量的海洋科技档案 (如, 调查报告、采集数据和成果图件等), 仅国家海洋局第一海洋研究所目前馆藏了从新中国成立以来各类科技档案4 000余卷。但是, 绝大部分档案以纸介质存储为主, 档案的收集、整理、鉴定、使用和编研等都以手工为主, 管理不科学, 工作效率低, 难以充分利用已有的海洋科技档案资源向科研人员、管理人员和公众提供良好的服务^[1]。因此, 应用先进的网络技术、计算机技术进行濒危档案的数字化拯救, 提高海洋科技档案的科学管理水平, 实现海洋科技档案资料系统、全面、高效的信息化服务是我们工作的着重点^[2-3]。

笔者针对海洋科技档案管理的特点和具体的应用需求, 结合先进的数据库技术, 提出了基于ORACLE的海洋科技档案数据库设计方案, 以期对海洋科技档案资料进行标准化、统一化存储和管理, 提高档案资料的查询效率和质量, 提供高效的数据支撑, 进而促进海洋档案资料的数字化、信息化进程, 实现我国海洋档案资源的共建、共知和共享的目的, 为建设“富饶海洋、生态海洋、安全海洋、数字海洋、和谐海洋”提供基础支撑。

1 需求分析

随着科技的发展和档案管理人员素质的提高, 对档案管理提出了新的要求, 目前我国在海洋科技档案管理遇到如下问题。

1.1 历史档案的拯救

在20世纪90年代中期以前积累下来的海洋科技档案资料主要以纸介质存储为主, 保存不易。同时, 纸介质档案资料的准确快速查找也非常困难, 尤其是历史档案, 索引信息不完整, 如果出现文件丢失、损坏, 将会造成不可弥补的损失。因此, 必须进行档案的数字化拯救, 这涉及大量的数据存储, 以及档案元数据信息的采集。

1.2 新时期快速增长的数据量

近几年来, 我国海洋事业进入了历史上最好的发展时期, 高新技术和先进仪器设备不断

地被应用到海洋调查、研究和开发等各方面,海洋科技档案的收集也改变为纸质资料与电子资料的双归档,因此资料数据量也呈几何级数增长。

1.3 档案存储分散

我国目前有几百家涉海研究单位,大量的海洋科技档案资料都分散在各个项目承担单位内,存放地不统一、管理不规范,并且受各个单位的管理限制,难以实现海洋科技档案系统、统一的管理与提供开放的服务。

1.4 归档标准繁多

海洋开发所涉及的专业种类繁多,目前各个专业之间并没有一个统一的分类编码标准和规范,同时档案资料分散在各个不同的部门,在文档的管理方法和规则上各不相同,因此资料在整理、分卷、编目和归档的标准不统一,造成了对海洋科技档案的整合和查询非常困难,利用率较低。

目前,我国数字海洋信息框架构建工作已全面展开,因此,应用先进的信息化技术、计算机技术实现海洋科技档案的数字化管理,抢救濒临破损和丢失的海洋历史档案资料,提高档案资料的管理水平和工作效率,为“数字海洋”的建设提供基础支撑,是目前我国海洋科技档案工作迫切的任务。

2 数据库设计

2.1 数据库管理系统

由于在海洋开发活动中形成的科技档案种类多,涉及的专业广,分类编码方法和管理工作各不相同,而且随着海洋开发力度加大和先进技术的应用,海洋科技档案数量的增长速度非常快,档案的信息化管理实现起来存在较多困难。因此,在数据库管理系统的选择上,需要满足对结构复杂、海量的档案信息进行快速高效的存储、维护和检索的需求,具备大容量、高性能、全方位资料保护和系统备份架构设计等优点,能符合网络存储快速增长的需求,使应用系统可以方便、及时、准确地从数据库中获得所需的信息。

ORACLE数据库是一种对象关系型数据库

管理系统(ORDBMS),它提供了强大的数据管理功能、分布式功能和网络功能,支持多个用户并发访问,具有良好的开放性和可移植性^[4]。因此,笔者在海洋科技档案数据库的设计上采用ORACLE数据库来组织和管理种类繁多、格式复杂的海量档案资料。

2.2 数据分类

经过分析,笔者针对数据类型和特点将海洋科技档案资料分为指令性项目、开发服务项目、仪器设备、声像和实物等几大类,各大类数据又细分为若干子类。例如,项目信息按照课题的专业类型细分为地球科学、海洋科学和生物科学等,海洋科学又细分为海洋物理、海洋化学和海洋地球物理等子学科。依据数据分类规则,创建数据字典,编制分类代码,开发数据字典管理模块,便于对数据类别名称和代码的管理和维护。

2.3 概念结构设计

概念结构设计主要对海洋科技档案数据对象之间的相互关系和属性结构进行设计和定义,它是独立于数据库管理系统和应用操作系统的。海洋科技档案资料包括元数据、项目、人员、单位、调查船、仪器设备和档案文件等对象,每一个对象除了自身具有属性结构之外,相互之间还具有一对多或者多对多的关系。如,“元数据”对象除了具有档案名称、档案类别、语言、摘要、数据量、格式等属性信息之外,还与“项目”对象具有多对一($m:1$)的关系(即一个项目可能包含多个档案资料的元数据),同时又与“档案文件”对象具有一对多($1:m$)的关系(即一个元数据对应多个档案文件)。另外,“项目”对象除了具有项目名称、项目代码、起始时间、终止时间、坐标系统和投影方式等属性信息外,还与“仪器设备”对象、“调查船”对象、“单位”对象分别是多对多($m:n$)的关系。这样,通过概念结构可以清楚地对海洋科技档案涉及的数据对象进行划分,定义其属性结构的名称和类型,进一步明确数据对象之间的关系;验证数据结构的一致性,为下一步逻辑结构的设计提供依据。

2.4 逻辑结构设计

逻辑结构设计是依赖特定的数据库管理系统或应用操作系统的。在海洋科技档案数据库概念结构设计的基础上,基于 ORACLE 10g2 数据库管理系统规定的标准和格式,对数据库的逻辑结构进行设计。在逻辑结构设计中,对概念结构中的数据对象向 ORACLE 10g2 所支持的关系模式进行映射。如,表和视图等。对具体属性结构的名称和类型进行定义,转换成能够被 ORACLE 10g2 直接识别的关系模式。如,对于时间类型的属性在 ORACLE 中定义为 DATE,变长度类型的属性在 ORACLE 中定义为 VARCHAR2,长二进制类型的属性在 ORACLE 中定义为 BLOB。

另外,在逻辑结构设计中,需要根据数据对象之间的关系(如一对多,多对多等)转换成不同的数据库关系表。概念结构向数据库关系表转换主要遵循以下原则。

(1) 每一个对象可以映射为一个或多个关系表,多个对象也可以对应同一个关系表,对象的属性即为关系表的属性字段,对象的唯一标志(ID)是关系表的关键字;

(2) 关联关系(一对一、一对多以及多对多)可以映射为一个包含被关联对象主键的关系表,同时也可以不进行映射,而在相关联的对象关系表之间定义相应的外键;

(3) 对映射后的数据库关系表进行冗余控制调整,使其达到合理的关系范式。

例如,“项目”对象与“仪器设备”对象在概念结构中是多对多的关系,转换成逻辑结构之后,除了各自转换成相应的关系表,还定义一个“项目—仪器设备”关联表,分别通过外键关联将两个关系表连接起来。

2.5 物理结构设计

海洋科技档案数据库的物理结构设计主要针对数据库物理存储空间以及外设配置进行设计,使得海洋科技档案资料数据库响应时间少,存储空间利用率高,事务吞吐量,数据安全性高,维护代价低。前已述及,海洋科技档案数据库采用 ORACLE 数据库管理系统对数据进行管理,ORACLE 数据库提供了表空间管理模

式,从逻辑上将数据存放在不同的表空间中,从物理上将数据存放在与表空间相关联的数据文件中(图1)。

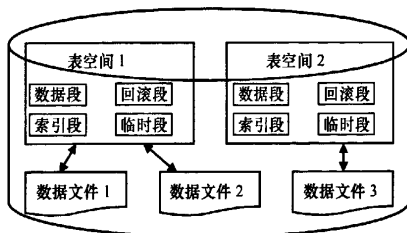


图1 数据库、表空间和数据文件的对应关系

笔者利用这种表空间管理模式,对海洋科技档案资料进行分类,将各个类别的档案资料放在不同的表空间进行管理。同时,根据各类数据的大小和读写频率设计相应的数据段、回滚段和索引段,规定每个段中数据块的参数值,优化海洋科技档案数据库的性能。其中,数据段存放数据表的数据;回滚段存放被每个事务修改的数据的原有值,用于回滚事务和恢复数据库;索引段存放索引信息,用于快速检索数据。

3 结论

本文基于海洋科技档案资料管理的应用需求分析,从数据库管理系统选择、数据分类、概念结构、逻辑结构、物理结构和安全性能等几个方面对海洋科技档案数据库进行了设计,建成了海洋科技档案数据库,对海洋科技档案资料进行标准化和集成化存储和管理。另外,在该数据库的基础上,基于 C/S 与 B/S 混合结构开发了海洋科技档案管理信息系统^[5],实现了档案资料信息的自动化和便捷化的管理、检索、浏览、统计、输入、输出及格式转换等功能,满足我国海洋科学研究和工程建设对档案资料管理的应用需要。

参考文献

- [1] 侯秀生. 试论我国海洋档案的管理机制与模式[J]. 海洋通报, 2003, 22(4): 60-65.
- [2] STIELOW, FREDERICK J. Building digital archives, descriptions, and displays: a how-to-

- do-it manual for archivists and librarians[M]. New York: Neal-Schuman Publishers, 2003.
- [3] CRESPO A, GARCIA-MOLINA H. Modeling Archival Repositories for Digital Archives[R]. Stanford Computer System Laboratory Technical Report, 2004.
- [4] 丁铨. ORACLE8/8i 数据库系统管理[M]. 北京: 人民邮电出版社, 2001: 1-5.
- [5] 崔爱菊, 卢铭, 梁运萍. 基于 C/S 与 B/S 混合结构的海洋科技档案管理系统的设计和实现[J]. 海岸工程, 2005, 24(4): 73-82.

(上接第 55 页)并计算相关的特征数据,提供初步的分析结果,自动生成决策所需的图形和量算数据等;可根据事件特征,查阅与此有关的法律、法规、公约、案例、军事、经济等相关法律条文,辅助实现突发应急事件的处理。对海洋灾害突发事件的处置提供一个快速可靠的应急机制。

海洋调查项目管理:建立申请项目审批信息库,可管理和提取海洋科学调查和国际海洋科学合作项目申请、审批、实施、执行、检查等基础性信息,为国家对相关项目的立项、审批和组织实施提供辅助决策的科学依据,满足我国在海洋科学研究管理的需要。

案例法规管理:实现海洋倾倒区、自然保护区、特别保护区、排放标准、水质标准、排污总量和污染损害赔偿等信息的管理,查询检索与损害我国海洋权益的国际法和我国相关法律法规的条款,以及国际上对此类违权事件处理的案例,从而为我国处置海洋环境污染受损个案提供法律依据和国际惯例的决策。

总之,渤海是我国沿海及诸海域中生态环境最为脆弱的海域。加强这一海域生态环境的综合治理,是保护水体质量和生物多样性,维护渤海健康生命的需要,是提高该区域环境承载力,建设新的区域经济增长点的需要,是人与海洋和谐相处,实现工程治理区经济社会可持续发展的需要。保护好、治理好渤海,不仅对渤海具有重

要意义,而且可以为我国其他海域的综合治理提供有益经验和借鉴。

我们应该同所有环渤海地区的利益相关者一道,携起手来,制定环渤海地区社会各阶层可执行的、实现未来发展美好设想的环境管理一般性战略,共同分担责任,解决依靠任何单一力量,如政府、管理部门、社区团体和个人所不能解决的复杂环境问题。我们相信,在各部门、有关利益相关者的共同努力下,渤海环境一定能得到有效的治理与保护,渤海和环渤海地区的明天一定会更加美好。

参考文献

- [1] 刘元旭,罗博. 渤海治污越治越污,多部门合作竟成“群龙闹海”[N]. 经济参考报, 2006-06-29 (4).
- [2] 许长新,邱珍英. 沿海滩涂开发与环境保护的可持续发展[J]. 海洋开发与管理, 2004, 21(6): 9-10.
- [3] 杜碧兰. 21 世纪中国面临的海洋环境问题[J]. 海洋开发与管理, 1999, 16(4): 67-72.
- [4] 刘文剑. 环境压力下海洋经济可持续发展研究[D]. 青岛:中国海洋大学, 2004.
- [5] 徐秦,方照琪. 船舶对海洋环境的污染及对策[J]. 中国水运, 2003(11).
- [6] 张玉珍,张丽玉,曾悦,等. 我国近岸海域环境现状及保护对策[J]. 山东环境, 2003(2): 32-34.