

数据银行技术在物探数据管理中的应用现状

张绍红 王永江 刘司红

摘要 随着石油勘探开发的发展,要求人们从大量的新、旧资料中快速获取有用的信息。数据银行技术给用户提供了一个高性能的存贮管理系统,提供更安全的快速联机访问,可以大幅度降低存贮、维护资料成本,节省寻找资料时间,帮助地学家更好地作出决策。它提供了一个经济而有竞争力的服务。本文对数据银行技术的有关方面进行了简要的综述。

关键词 数据银行; 存贮; 管理; 数据发掘

A REVIEW ON THE APPLICATION OF THE DATA BANK TECHNIQUE TO THE MANAGEMENT OF GEOPHYSICAL DATA

Zhang Shaohong

(China University of Mining and Technology, Beijing 100083)

Wang Yongjiang

(Aerogeophysical Survey and Remote-Sensing Center of MLR, Beijing 100083)

Liu Sihong

(Geophysical Institute of Nanyang Oil Field of Henan Province, Nanyang 473132)

Abstract With the development of petroleum exploration and exploitation, people are asked to obtain useful information from large quantities of new and old data as quickly as possible. The data bank technique provides the users with a high-performance storage and management system and safer rapid on-line visit so as to reduce the cost of storage and preservation of data, save the time of searching for data and help geologists to make decision in a better way. It supplies an economical and competitive service. This paper makes a comprehensive review on problems related to data bank technique.

Key words data bank; storage; management; data mining

当今的石油勘探复杂程度越来越大,要想取得成功,不只是需要利用新资料,还要从已经拥有的旧资料中获取大量信息。但是,目前三维地震越来越多,高分辨率地震、四分量地震以及新的测井方法等技术,使勘探数据量呈指数增加^[1]。而且,历史资料数据量也是一个天文数字,同时还存在管理上的问题。如何从这些新、旧资料

中进行分析研究,获取有价值的信息,是勘探界面临的一个重要问题。

1 数据管理存在的问题及新思维

因为缺少一致的格式和标准以及缺少大量的经费,油气公司很难从拥有的资料中取得最大效益。具体来说,传统资料管理存在的问题可以概括为以下几个方面。

1. 数据保存的时间过长,累积的档案错误普遍存在,因而保存下来的资料可能是有错误的井、地震或其它资料,这就降低了资料价值。

2. 数据有很大一部分还是不可机读的(或很难索引),格式化很差,数据缺少标准,没有一致的质量控制标准,通常难以访问。

3. 很多数据的管理责任机制不合理,绝大多数人员只管理与他们有关的资料,很多资料被人为地分割成多门学科分散管理(如电缆压力数据),而且经常是使用资料的人远离资料生产的人。这就很容易造成数据不完整和遗漏,数据不能及时更新。

因此,要从大量的已有资料中寻找有价值的信息,需花费大量的时间和精力。Shell等公司作过统计指出,勘探人员寻找资料的时间一般占整个研究工作时间的50%~60%,而真正进行有意义的资料分析研究的时间只有15%~20%,由此造成智力资源的重大损失〔1,2〕。

更为重要的是,数据采集、存贮管理成本不断上涨。有关的研究指出,存贮、拷贝与补救等管理一盘带的费用是一盘空白磁带价格的30倍〔2〕。现在,光是地震的数据量每年就增加25%,整个勘探数据量呈指数增长,沿用过去旧的管理方式显然是不够的。

目前,数据管理思维同以前有着很大差别。10 a前,许多公司拥有自己的地震队、钻井队,自己开发软件与硬件,各自利用专有的数据库,视占有技术和数据为竞争优势。而现在,勘探开发者认为竞争不应放在如何获得资料上,而应放在对资料的分析 and 解释研究中。很多公司把地震、测井、钻井工作承包出去,只进行少量的开发,利用商业性数据库直接购买数字化数据,不再是同时购买软件及数据,提倡“买而非建”(buy vs build)的经营哲学〔3〕。

随着Internet和万维网(www)技术的发展,人们进行通讯与组织活动的方式也发生了很大变化。光学技术(而不只是电子学)应用于存贮、处理和传输数据信息,面向对象的技术与人工智能(AI)技术也应用于存贮管理中。总之,信息技术革命重塑了石油天然气商业行为,促进了存贮管理技术的发展,使建立的数据库提供了最佳存取方式和最佳分析手段。数据管理系统发展的结晶就是正在兴起的勘探开发数据银行。

2 新的数据管理系统结构与优势

理想的数据管理系统,引用素有“资料存贮管理之父”之称的William H Inmon的定义,就是“面向对象的、综合的、稳定的和随时间变化而变化的数据集系统,能满足管理上各种需要”〔4〕,也就是使研究人员能在任何时候、任何地方均可访问有关信息。油气勘探新的数据管理趋势是:文件由数据库管理,公司各自数据模型被POSC及PPDM公共数据模型取代,集中式数据库向分布式Oracle数据库发展,适应多连接用户体系向服务器体系结构发展。数据管理的工业方法包括构造数据库、应用新的存贮技术、数据库管理软件及新的工业标准。新的数据管理系统能对数据进行检索

和存档，负责数据传输及输入/输出，保证数据质量和安全性。

2.1 基础结构

数据管理系统基本上包括存贮设备(如大容量存贮介质、自动化机器人图书馆系统)、工作站、X终端、PC机、局域网通讯设备、输入/输出设备(如扫描仪、数字化桌、磁带/磁盘驱动设备)等。每一种具体数据库管理系统的结构复杂程度视具体情况而定。例如，图1是CGG公司下属PetroSystems子公司与PECC公司在阿尔及利亚建立的结构〔5〕。

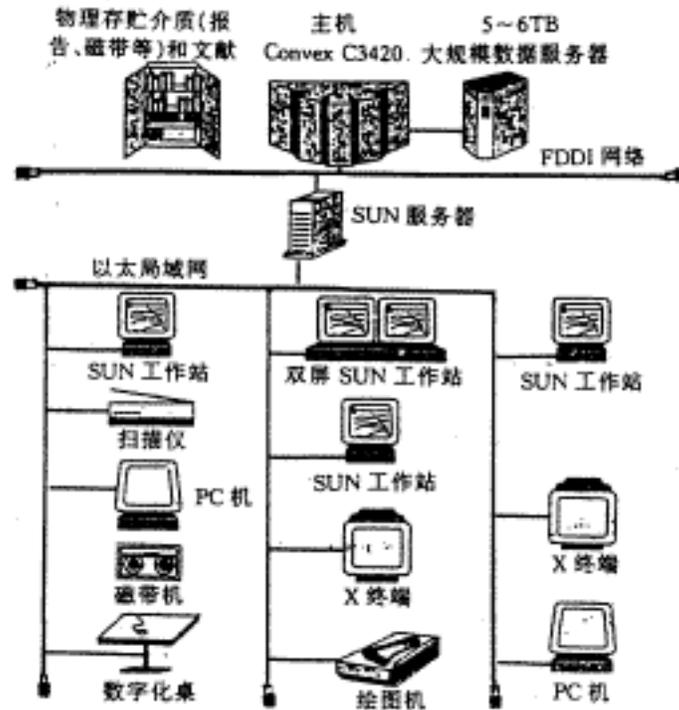


图1 CGG子公司PetroSystems与PECC公司建立的数据管理系统结构〔5〕

目前，数据库存贮空间为TB级(10^{12} 字节级)。系统可在Unix，PC机Windows及Windows NT等不同环境下进行，可把数据直接送到工作站上，供SeisWorks,IESX, Charisma,SeisX等软件应用。

数据管理系统中数据类型更加丰富，有利于研究人员在综合各方面信息的基础上，更快更好地作出评价和决策。数据银行提供多种数据，其中包括大型属性文件(如钻井数据、生产数据)、大型矢量文件(如地震数据和测井数据)、大型图像文件(如卫星图像资料)以及大型文件记录(如文献、报告、地图和图件)等。

数据银行系统安全机制健全。目前，几乎所有数据都是保密的，安全非常重要。数据银行系统采用防火墙、加密技术和其它存取控制措施，成员用户还可以配合使用其它保护性手段。例如，PetroBank安全机制为八层三级安全体系〔1〕。八层安全机制分别为物理运行、网络运行、系统访问Unix注册、应用访问(访问控制模块)、应用功能(Uniface和ArcDis应用接口)、数据库表查询(Oracle)、数据库表(Oracle)和大块数据文件(ADSM和Recall)。三级安全体系为系统运行、PetroBank应用软件及第三方软件。各种勘探开发应用软件都享有相同的项目数据库，从而简化了应用软件与数据库的传

输。

2.2 服务

新的数据银行系统提供多方面的服务。可把存贮在老介质上的大量数据转存到更可靠的大容量介质上。不管使用什么操作平台，通过提供通道服务，可为各种专用数据库提供发送的存取方法(如Oracle, Model 204, DB2, Sybase, Microsoft, Ingres和Informix等，参见图3)。

数据管理系统提供复制服务。管理系统每隔一定时间将数据从一个数据库复制到另一个数据库，以保证为用户提供最新的数据。

新的数据银行系统利用数据发掘技术(data mining technique)〔6, 7〕，根据多条标准查询和存取数据，把原始资料转变成知识提供给勘探开发各方面技术人员(如图2)。

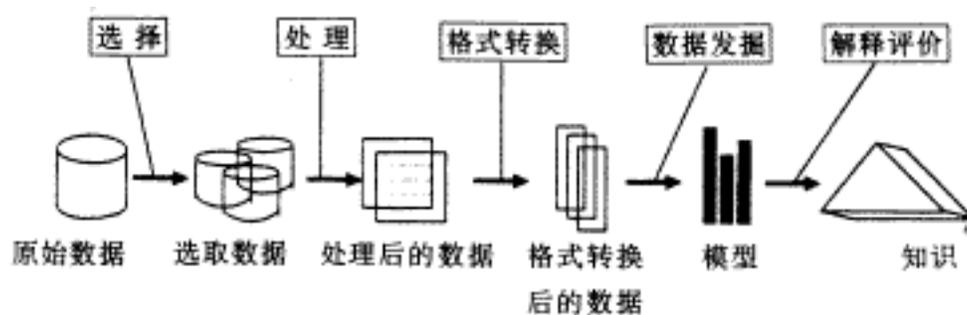


图2 利用发掘技术从资料中提取有价值的知识〔7〕

2.3 数据银行系统工作流程

一旦数据完成再处理加载进入共享数据库后，用户就可以通过高速通讯线路进入数据银行系统，利用先进的图形界面可以选取某地理范围内的所有石油技术数据(如测线、井位等)，然后自动化机器人图书馆管理系统就会根据用户的使用权限提供有关资料。

下面以Sybase公司实际使用管理系统进行具体说明(图3)。该系统基本工作流程可以概括为四个方面：

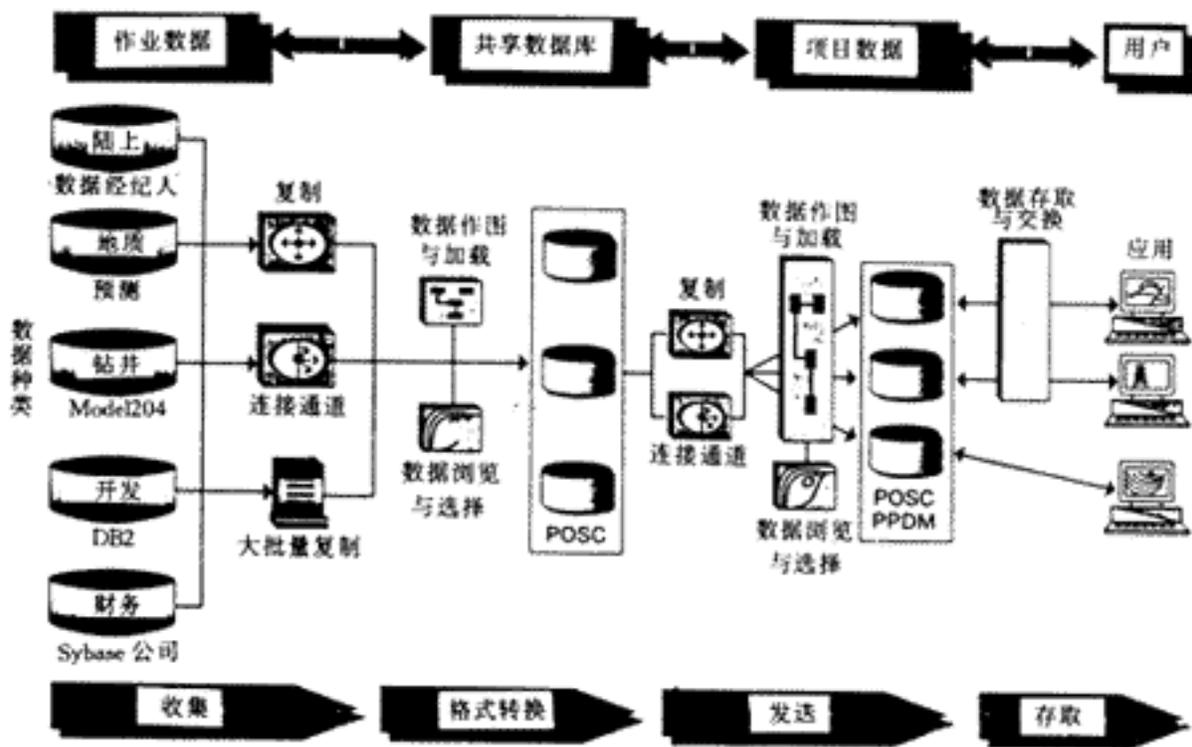


图3 Sybase公司数据管理系统处理程序〔4〕

1. 资料收集。收集勘探开发生命周期内最广泛的资料，包括物探、测井、地质、钻井和生产多方面数据，以及外部数据库等媒体的资料。

2. 对数据进行格式转换后加载进共享数据库，形成格式一致的共享资料。其中对原始资料进行再处理，保证多种格式的多源数据质量一致性至关重要。再处理时，通过扫描发现数据质量问题，按POSC开放标准信息交换数据模型，把数据格式转换成工业标准格式(如SEG Y, UKOOA, ESSO-V2, GEOSHARE), 然后转存到IBM3590或其它大容量介质上。

3. 把资料及时发送给用户。根据各个项目用户的要求，以低廉的成本快速向用户传送有关资料。

4. 用户利用有关设备，根据特定的业务从数据库提供的资料中选取最佳信息，并把最新的研究成果再输入数据库中。

2.4 数据模型工业标准

标准是资源和利益的统一体，制定工业标准是为了协调不同系统之间的联系。目前，在数据管理领域2个通用的数据模型标准是：POSC(petrotechnical open software corp,石油技术开放软件公司标准)和PPDM(public petroleum data model assoc,公共石油数据模型协会标准)〔3〕。POSC是一个面向对象的宽作用域综合逻辑数据模型，PPDM是一个物理关系数据模型。由于POSC较难理解，因此，一些油气公司和计算机厂家把POSC和PPDM相结合，推出工程开发计划(project discovery), 目的是把POSC扩展成PPDM所接受的物理关系数据模型。另外，有些组织或公司根据各自业务需要，还制定专用内部标准。例如，SEG采用RODE数据封装标准，以便实现不同介质间快速有效数据移植〔8〕。

2.5 利益

数据银行技术的经营效益主要表现在以下三方面：

1. 节约大量寻找资料时间。由于资料管理应用高速网络通讯传输、自动化存取等新技术，实行综合管理模式，可提供在线(on-line)、脱机(off-line)和近线(near-line)方式服务，比手工分散管理节省大量寻找资料时间。PGS公司实践表明，如果以前花两周时间寻找资料的话，现在只需要2~3 h。一旦接通世界范围内的勘探数据库，利用PGS Exsci数据管理系统就能获取数据，所选数据的作图和报告就会在几分钟内产生。IBM公司专家称，应用PetroBank可以减少20%~30%寻找资料时间〔9〕。

2. 节约数据管理成本。这里首先以挪威应用为例。在1992年以前，挪威各石油公司每年在数据管理上共花费近1亿美元，每个公司平均花费约700万美元。1992年挪威海上12家石油公司应用IBM的PetroBank技术实施Diskos项目〔9〕，建立全国性的共享数据库，各个公司在数据管理上分摊的费用只有18~25万美元。据统计，现在处于该项目的初期，参与的12家公司数据管理总成本就节约30%。如果该项目完成了，预计将节约50%的管理成本，也就是每年可节约4 000万~8 000万美元，并且到那时可使全挪威的E&P成本减少40%~80%。又如英国开展CDA项目(common data access共享项目)把多家公司的多个拷贝整理，去掉重复，将200万项数据减少到约25万项目，大大降低了管理成本〔10〕。

3. 促进数据贸易。通过网络使数据交易更方便。目前，已完成多起数据购买合同。

总之，数据银行具有一个灵活多变、适应性强的信息技术结构，适应多种操作系统、网络、介质类型和数据格式。勘探学家将能方便地选择世界各地感兴趣的地震等数据。

3 国内外数据银行发展现状

世界上许多石油公司，如IBM，PGS，CGG，Schlumberger等都在建立石油勘探开发银行。这些数据银行有全球性的、各国的或企业级的不同规模。它们以适应市场的不同形式建立起来，其经营哲学是“共同参与”(partnering)，包括技术上和商业上合作。技术上，就是利用先进的网络通讯技术和管理技术，多个公司共同研究、开发，形成公共数据银行和数据中心，多个公司共享更多的数据和计算成果。商业上就是共同分担数据服务费用，加强数据贸易。

笔者搜集到的国外几个较大数据银行有：

1. PetroBank〔1, 9〕。全球最大的数据银行之一，IBM公司注册，1995年正式投入商业应用，功能强大，应用范围遍及全球。1996年收入为3 300万美元。IBM希望占有全球此类市场40%的份额。

2. GeoBank。PGS公司注册，已在伦敦、休斯顿和澳大利亚的珀斯建立了多用户数据银行。有商用软件PGS Tigress和Exsci数据管理软件。其中PGS利用GeoBank技术与英国HIM公司(Hay Information Management)合作，共同建立全球性数据管理系统，已在英国、美国等12个国家的50多个数据中心安装。

3. GeoQuest Finder数据管理系统。由Schlumberger GeoQuest公司开发，是全球最大的数据库之一。包括LogDB, AssetDB和SeisDB软件包，应用遍及30多个国家，已安装500多套。其中在委内瑞拉建立一个全国综合数据库。

4. 英国Lynx信息系统公司开发的ArcView数据管理系统。该系统把包括伊拉克、利

比亚、西伯利亚、委内瑞拉和东南亚等国家多家公司拥有的数据连接在一起，形成了一个全球性的大信息包。

5. Landmark公司OpenWorks项目数据管理系统及OpenExplorer企业级数据管理系统。

6. CGG与Shell公司合伙建立了世界范围E&P数据管理系统。

在国内，石油天然气总公司正在建立一个全国性的大型综合数据管理系统，而且石油等勘探单位目前也在建立各自的企业数据库，加强资料数字化、电子化的工作。

其中物探局于1997年建立了企业级“陆上石油物探管理信息系统”〔11〕。

4 结束语

数据银行提供了一套综合管理各种勘探数据的方法，可以帮助地学家有效地利用大量新旧资料，快速获取有关信息，减少因数据不充分引起的决策失误，更快更好地作出决策。

尽管在浏览、查找和分析可视化数据方面，数据银行是一个既简便又功能强大的工具，但目前仍存在以下挑战：

1. 定义新的类型(如空间域、时域和图形)，并和传统数据库集成，其规模要扩充到 10^{15} 字节(PB)级。

2. 多源数据集成(组合)，使用何种资料来源，哪里可能有多种来源。如何自动发现数据的趋势和异常(数据发掘、数据分析)。

3. 根据资料的经营性质和最终用户性质，对资料的定位，进行人类组织的工作流(即运行的一组作业)和数据流的刻划和管理。

4. 自动数据库设计和管理。由于计算机技术可以廉价地存贮各种数据，大多数数据将成为在线数据，因此，也许最富挑战性的问题是对数据的理解。

石油勘探数据银行思想新、方案周全，放眼全球，今后对石油工业的作用必将越来越大。

第一作者简介 张绍红，男，1965生，1988年毕业于中国地质大学(武汉)应用地球物理系，后在河南省南阳油田计算中心从事石油勘探与研究，任工程师。1997年在中国地质大学(北京)攻读硕士学位，1999年到中国矿业大学(北京)攻读博士学位，公开发表论文数篇。

作者单位：张绍红 中国矿业大学，北京 100083
王永江 国土资源部航空物探遥感中心，北京 100083
刘司红 河南南阳油田物探研究所，南阳 473132)

参 考 文 献

[1] Nick Wreden. Norwegian-PetroBank offers modern data handling solutions. World Oil, 1996, 217(9): 60 ~ 64

[2] Lesslar P C, et al. Managing data assets to improve business performance. JPT, 1998, may: 54 ~ 58

- [3] Stark P H,Hodgson R N.Data management:future and directions.World Oil,1996,217(9):57 ~ 58
- [4] Rich Morris.Data warehousing leads to improved business performance.World Oil,1995,216(9):51 ~ 52
- [5] Jean-Paul Marbeau,Ugar algan.Databank solution for E&P data management.World Oil,1995,216(5):66 ~ 67
- [6] Joyce Bischoff.数据仓库技术.北京：电子工业出版社，1998
- [7] Fan Jianhua ,et al.An overview of data mining and knowledge discovery.Journal of computer science and technology.1998,13(4):349 ~ 358
- [8] Allen Johnson.Effective data management provides a competitive advantage.World Oil,1995,216(10):151
- [9] Chip Nemesi.IBM's off and running with its data management solution.First Break,1997,15(2):49 ~ 50
- [10] News.CDA project enters new phase.First Break,1997,15(6):201 ~ 202
- [11] 岳永霞.勘探数据库的建立与发展.石油地球物理勘探，1998，33(增1)：175 ~ 176
- [12] Dennis Denney.Information-technology integration at Petroleum De Venezuela S.A.JPT,1998,April:98 ~ 106
- [13] Keney J S.Enhancing the effectiveness of management systems through a TQM approach.JPT,1988,June:62 ~ 68

1998年9月9日收稿。