

# 情报资料的信息化建设实践

王俊\*, 韩涛

(中国科学院国家科学图书馆总馆, 北京 100190)

[摘要] 本文总结了战略情报研究对数据资源信息化建设的实际需求以及信息化建设为战略情报研究带来的效用, 同时分析目前情报调研数据资源的现状, 针对其中资源不易共享的问题, 提出一套通过信息化建设来改进、改善该问题的设想。通过构建数据资源的原型系统, 对“调研得到的国外创新型国家各项创新指标数据”进行加工、存储、检索与利用, 以这个范例来探索设想的可行性, 尝试通过该系统实现情报研究团体内部战略情报研究数据资源统一的存储、管理与共享。在该系统的基础上根据实际应用的需要, 可以扩充统计分析、授权管理等功能, 还可以集成第三方数据分析软件, 并转换成 B/S 体系结构, 最终形成情报研究成果呈缴与共享、数据管理与分析的平台。

[关键词] 情报调研、数据资源、信息化、共享

[中图分类号] G352

## Data Resource of Information Study for Informatization Construction

Wang Jun\*, Han Tao

(National Science Library, Chinese Academy of Sciences Beijing, 100190)

[Abstract] This article is based on data resources of the information investigation and study unification processing, the memory, construct the storehouse, the retrieval and the use, take 'each innovation target data of overseas innovation country' as the model, through the construction of data resources prototype system, attempting to establish the memory, the management of resources and the sharing mechanism, to further construct the top level application mechanism to lay the foundation.

[Keywords] information investigation and study, data resource, Informatization, Sharing

### 1. 前言

我国的信息化建设已有20余年历史, 得到了政府、科研机构 and 产业部门的广泛响应。从全国到地方, 从高校到图书馆, 从税务征缴到国土资源管理, 各地区、各部门、各行业都纷纷提出信息化、数字化计划, 取得了巨大的进步<sup>[1]</sup>。在信息化建设渗透在各行各业的大环境下, 本文站在战略情报研究的角度, 总结战略情报研究对信息化建设的实际需求以及信息化建设为战略情报研究带来的效用, 在此基础上以“国外创新型国家各项创新指标数据”的调研项目为范例, 构建情报

---

\* 作者简介: 王俊 (1977-), 女, 博士, 助理研究员; 研究方向: 科技情报共享研究。

Email: wangj@mail.las.ac.cn

收稿日期: 2009-12-28, 发表期刊: 《现代情报》2010 年第 3-4 期。

研究调研数据的管理信息系统, 尝试通过该系统实现战略情报研究数据资源的存储、管理与共享。

## 2. 情报研究中问题以及信息化的作用

情报研究部门经过多年的建设与实践, 形成了大量的研究成果, 其中尤其是各种调研数据数量巨大, 种类繁多, 而且随着部门业务量的增加, 数据资源也以惊人的速度与日俱增。然而, 这些通过调研、整理、汇编而得到数据资源通常是以表格的形式分散在各种各样的文件(如 Word 文档、PDF 文档、Excel 文档等)中, 并由课题组中各研究人员独自保管。这种状况使得课题组中的数据管理和共享面临了一定的困难, 也在一定程度上影响着情报研究的交流范围, 影响了情报研究的工作效率。由于情报研究人员各自具有一套数据格式, 情报研究人员通过调研各自掌握的数据资源异构, 导致这些数据无法积累形成一个不断增加、格式统一的数据资源, 因此这些资源也无法在今后的工作中被重用, 当然就更谈不上对这些数据资源进行深度的统计分析。

随着战略情报研究中情报调研广度和深度的提升以及情报研究队伍的日益壮大, 将情报研究过程中沉淀积累下的研究成果进行整合, 实现有效的管理与共享, 这样的需求越来越强烈, 而这正可以借助信息化技术得以解决。结合目前的现状来说, 情报研究中数据资源信息化有以下几点作用:

(1) 研究报告数据(指标数据)集中存储在数据库中实现统一管理。

规划好数据的存储形式后, 把各类研究数据分门别类地集中入库, 一个集中管理的数据库既是资产也是信息化的基础, 同时通过备份来保护这些数据是容易完成的。而且随着数据的逐年增长, 数据库的优异管理能力也会更加凸现。

(2) 便捷、安全地调阅各类报告数据, 方便共享。

在数据库上构筑基于 DBMS 的上层应用, 借助分布式计算技术并发完成对各类数据进行加工、索引与授权控制, 从而实现各类数据的快速检索和权限范围内的共享。

(3) 随时统计分析, 生成相关图表。

查询数据库中历年各类数据, 方便地对某些指标进行统计分析, 生成报告。也可在最终报告中整合进各类图表从而更直观地表达这些结果, 以辅助进行决策。

(4) 与第三方数据分析软件的适配、集成。

第三方数据分析软件(如德温特分析家(TDA), 我馆与北理工共建的战略情报分析工具)均有自己要求的特定数据格式。为系统开发适配接口, 直接将库中需要分析的数据导出成这些第三方分析软件所需的格式, 并使用这些软件进行更深入的分析。

信息化的内涵围绕着数据统一存储、共享而展开, 这也是一个制定数据结构、数据格式的过程, 起到了对调研专题数据修订、完善、标准化的作用。

### 3. 信息化实施步骤与系统构建<sup>[2-3]</sup>

#### 3.1 实施步骤

##### (1) 数据库规划

选择一个功能强大,易于使用,对应用开发有较好支持且成本低廉的数据库管理系统(DBMS)是实现信息化的第一步。接着规划数据存储空间,以容纳未来若干年的数据。并由于数据库及数据的重要性、安全性,那么就需要专人(DBA)对其管理与维护、定期备份。

##### (2) 应用开发

只有开发出基于数据库的相关应用功能,才能方便操作者查询自己关心的数据、报告并将日常调研数据入库。此时需要开展应用系统需求调研,使系统的开发与未来的使用者之间进行互动,既完善应用系统功能,又利于日后的系统推广。

##### (3) 系统运行维护

有了基于数据库的应用系统,需要广泛听取使用者意见,以便改进系统使之更贴近于用户操作习惯,更适合于当前工作模式,从而切实发挥信息化的作用,满足对数据、经验进行管理、积累、沉淀、共享的初衷。

#### 3.2 系统原型定义与构建

##### (1) 系统原型定义

系统原型是向正式系统过渡的中间产物,定义一个系统原型有助于大家理解未来系统,并可据此展开讨论,为以后系统的迭代开发提供参考。

信息管理系统应具备的基本功能包含数据类别,指标定义,数据录入,检索,浏览打印报表数据,汇总数据等等:

① 调研前期,调研人员应可使用系统自由地制定信息类别与衡量指标,形成调研数据规范。

② 调研期间,调研人员应根据已经制定的数据规范,使用系统进行已录数据查询和数据样本录入。这些数据均应集中存放在一个统一数据库中。

③ 调研期间,系统应允许调研人员修改已有指标和创建新的类别与指标,以更好地完善数据规范。

④ 系统应可以对已经入库的数据进行综合条件的检索与统计,并出具可能的图表,提供给调研人员作为参考。

数据集中存放的意义就在于能够管理、积累、分析、与共享,这也是系统应该着重实现,充分挖掘的功能。

##### (2) 系统原型的初步构建<sup>[4-6]</sup>

基于以上系统原型定义,尝试以“调研国外创新型国家各项创新指标数据”为范例,进行系统的初步构建,开发出一套数据管理原型系统——“科技信息数据管理系统”。以此尝试尽可能地佐证数据资源信息化建设的可行性。该系统目前采用 C/S 模式,基于 Java 开发;允许多用户同时访问并将数据统一存储在指定的

SQL Server 服务器上；只需要在本地机器上安装 JRE 作为运行环境即可。该原型系统主要具有以下功能。

① 类别指标管理, 即调研数据规范制定, 辅助操作者完成数据规范的制定, 并约束此后的调研数据录入。如图 1 所示。

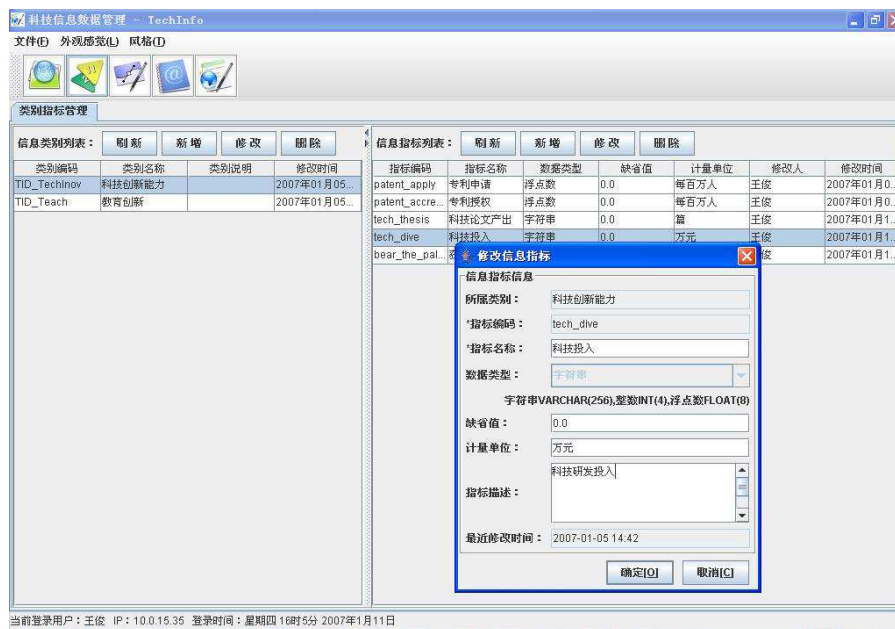


图 1 类别指标管理

② 数据综合查询, 根据时段、国家/地区、数据类别与指标进行数据查询、数据导出、图表分析。如图 2 所示。

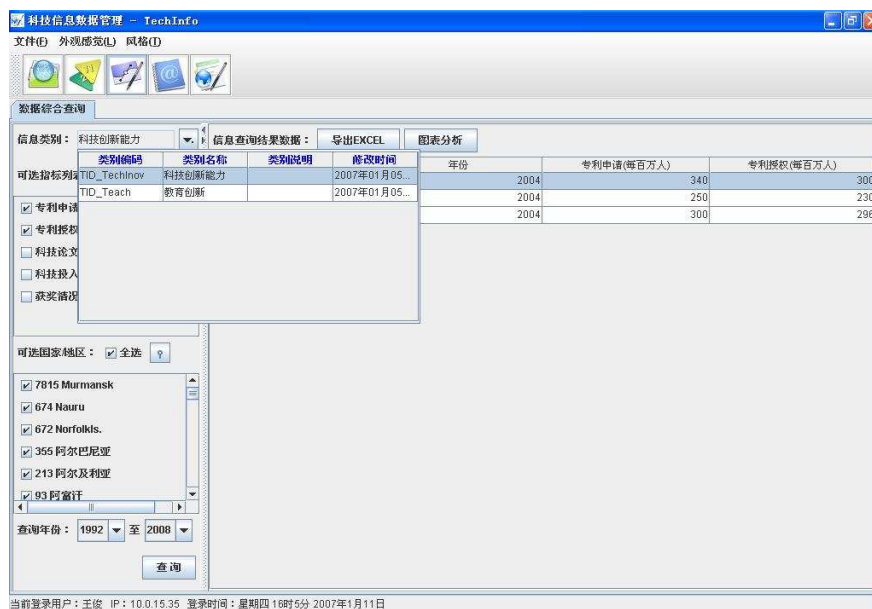


图 2 数据综合查询

③ 科技数据处理, 对调研信息的采编, 进行数据的查询与录入。如图 3 所示。

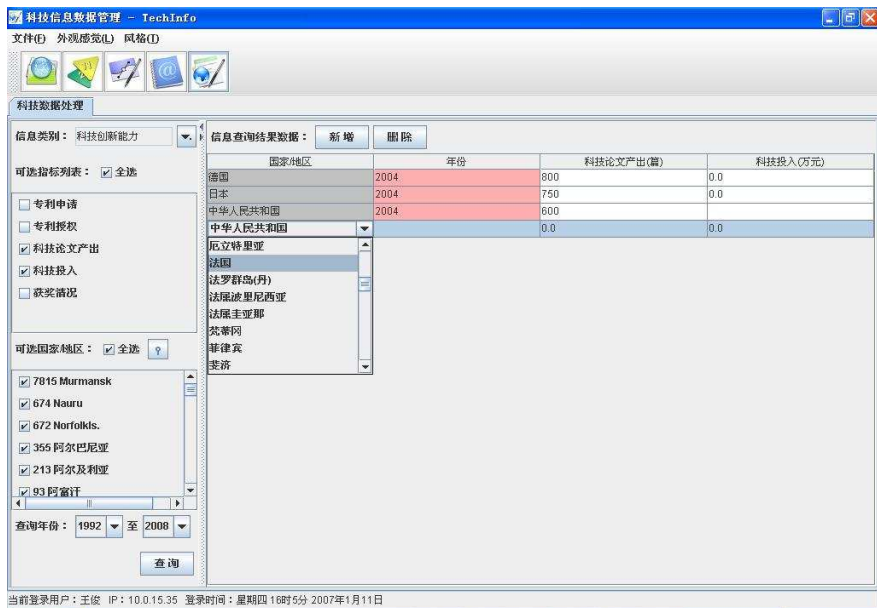


图 3 科技数据处理

④ 信息引文管理, 对调用信息数据的来源进行分类别、分指标地管理。如图 4 所示。

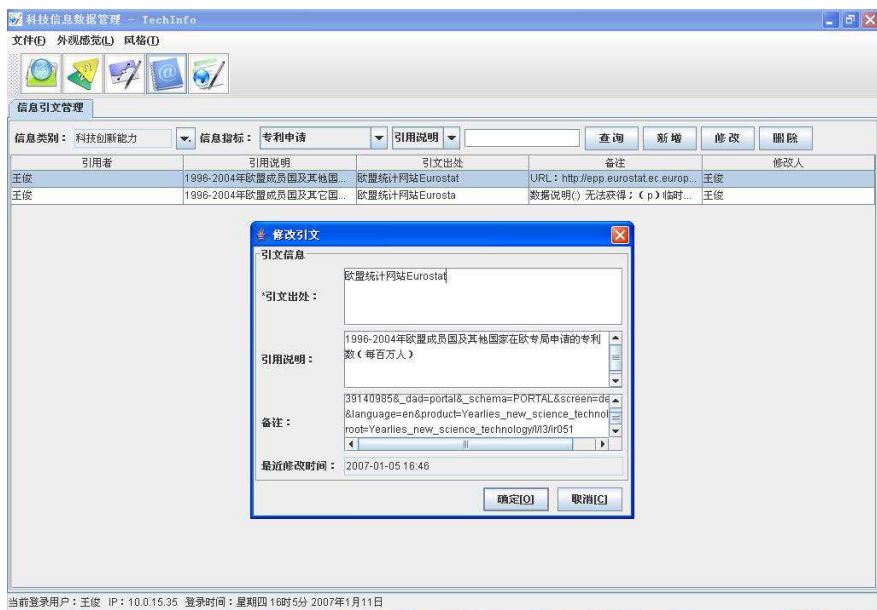


图 4 信息引文管理

通过上述介绍, 可以看到情报研究人员使用类别指标管理功能辅助制定课题调研指标, 调研过程中, 使用科技数据处理功能录入数据样本, 使用引文管理功能保存引文信息。当然, 也可能随着调研需求的明确, 会对类别和指标进行反复调整, 最终会得出一套课题的规范数据标准(元数据)。

在这个工作流程中, 数据规范(类别指标)被严格定义, 数据按规范被存储, 并可被检索分析, 操作员可以查询到其他使用者建立的规范与数据。这初步起

到了信息化中数据存储、信息共享、有效管理的目的,并为以后可能的数据挖掘奠定了基础。

#### **4. 系统功能迭代与展望**

通过该系统可以基本实现战略情报研究中数据资源的统一管理和有效共享。在该原型系统的基础上,根据实际应用的需要,可扩充统计分析及授权管理(人员管理、角色管理及系统授权)等功能,还可以将第三方数据分析软件整合其中。最终建设成 B/S 结构,形成情报部门成果呈缴与共享、数据管理与分析的平台。

#### **参考文献**

- [1] 李景文. 国土资源信息化体系结构及相关技术[J]. 桂林工学院学报. 2005, 25(3),313-316.
- [2] 王珊. 数据仓库技术与联机分析处理[M].北京:科学出版社, 1998.
- [3] 王珊, 刘怡, 晋良颖, 麻占全. 数据组织与管理[M]. 经济科学出版社, 1996.
- [4] 李晓喆, 张晓辉, 李祥胜. SQL Server 2000 管理及应用系统开发[M]. 人民邮电出版社. 2003.
- [5] 徐国智, 汪孝宜等编著. SQL Server 数据库开发实例精粹[M]. 电子工业出版社. 2006.
- [6] Cay S. Horstmann, Gary Cornell 著, 程峰, 黄若波, 章恒翀译. Java 2 核心技术(第 6 版)[M]. 机械工业出版社.