Beijing Ruankexue
Yu Shehui Kexue Yanjiu Chengguo
Xinxi Ziyuan De
Zhenghe Yu Liyong Yanjiu

北京软科学与社会科学研究成果
信息资源的整合与利用研究

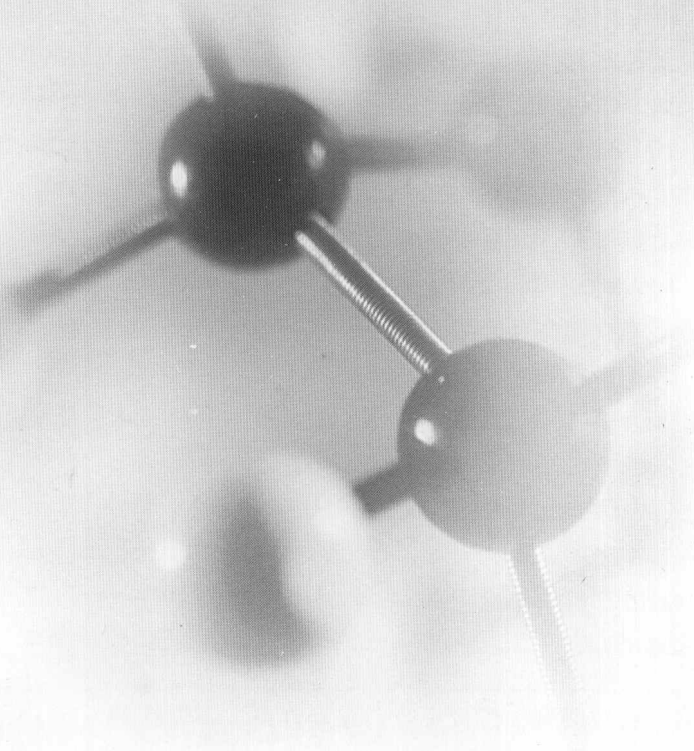
北京市哲学社会科学规划办公室
北京网讯博通信息技术有限责任公司

同心出版社

Beijing Ruankexue
Yu Shehui Kexue Yanjiu Chengguo
Xinxi Ziyuan De
Zhenghe Yu Liyong Yanjiu

北京软科学与社会科学研究成果
信息资源的整合与利用研究

北京市哲学社会科学规划办公室
北京网讯博通信息技术有限责任公司



同心出版社

图书在版编目（CIP）数据

北京软科学与社会科学研究成果信息资源的整合与利用研究/北京市哲学社会科学规划办公室编.
—北京：同心出版社，2007
ISBN 978-7-80716-582-8

Ⅰ.北… Ⅱ.北… Ⅲ.①软科学-研究-北京市②社会科学-研究-北京市
Ⅳ.G322.71 C121

中国版本图书馆 CIP 数据核字（2007）第 165310 号

北京软科学与社会科学研究成果信息资源的整合与利用研究

出版发行：同心出版社
出 版 人：刘霆昭
地　　址：北京市建国门内大街 20 号
邮　　编：100734
电　　话：发行部：（010）85204603（外埠）、85204612（本市）
　　总编室：（010）85204653
E-mail：txcbszbs@bjd.com.cn
印　　刷：北京雅艺彩印有限公司
经　　销：各地新华书店
版　　次：2008 年 1 月第 1 版
　　2008 年 1 月第 1 次印刷
开　　本：787×1092　1/16
印　　张：13 印张
字　　数：220 千字
定　　价：28.00 元

编委会

前 言

随着科学技术的不断发展进步，信息已成为各个行业的重要资源，对于软科学和社会科学的研究成果，在客观上作为本行业的信息资源是相对独立、未能共享的，存在信息分散、资源浪费等情况。诸多问题的存在，在一定程度上影响着社会科学研究事业的发展，是当前亟待解决的重要课题。对软科学与社会科学研究成果资源的整合与利用进行深入研究，找到一种切实可行的资源整合方案并加以具体实施，是解决这一难题的有效途径。

北京作为我国的政治、文化中心，是人文荟萃的地方，有雄厚的科学研究资源。北京地区科研单位承担的各级各类研究项目种类繁多，数量庞大，项目研究所创造出的大量科研成果是一笔巨大的知识财富。为充分开发和利用这些极其丰富的研究成果资源，北京市科委软科学处和北京市哲学社会科学规划办公室等单位都相继建立了在网上运行的科研成果数据库系统，但目前这些科研成果数据库系统大多数都是独立运行的，研究成果资源未能实现共享，造成了科研成果信息资源的极大浪费。

为此，我们申请了对“北京软科学与社会科学研究成果资源的整合与利用研究课题”进行研究，并于2005年9月得到立项。课题立项后，我们对北京市科委软科学处的成果库系统、北京市哲学社会科学规划办公室信息管理系统等现有比较成熟的科研成果类数据库进行了调研，并在调研的基础上，经过多次研讨，课题组提出了利用数据仓库技术整合软科学与社会科学成果资源、运用知识管理手段促进成果转化和共享的实现方法。在强化资源共享理念的基础上，研究设计了北京科学研究成果资源共享平台，以科委、规划办等单位成果数据资料为基础数据源，通过构建规范的软科学及社会科学科研成果数据仓库来实现各系统间的“无缝”连接，对科研成果进行全面的资源整合和共享。在技术系统的设计上，我们借鉴国际的信息系统建设方面的先进思想，在保证系统效率的前提下，突出系统的开放性、标准化、模块化、

易用性、高效性、可扩展性、安全性、可管理性以及可靠性等特点。同时，北京科学研究成果资源共享平台采用数据仓库、商业智能等技术，对软科学和社会科学成果资源进行深入挖掘，提高信息资源的可用性。通过充分应用诸多先进技术构建的北京科学研究成果资源共享平台，将为软科学和哲学社会科学成果资源的充分共享、有效利用提供强有力的技术支撑，为科研机构和科研工作者以及党和政府决策机关提供广泛的资源共享及高质量的信息服务，对于推动学术繁荣和理论创新，推动科研管理机制创新，强化科学技术为改革开放和现代化建设服务功能，促进哲学社会科学发展等都具有十分重要的意义。

编著者中一部分是哲学社会科学规划、研究方面的专家，他们具有丰富的哲学社会科学研究经验；一部分是网络信息系统研究开发人员，在信息资源整合利用系统的研制方面掌握着关键技术并有着多年的实践开发经验。这为“北京软科学与社会科学研究成果资源的整合与利用研究课题”的顺利进展提供了强有力的支持与保障。

由于成书时间仓促，加之作者水平有限，书中疏漏之处在所难免，恳请各位专家和读者批评指正。

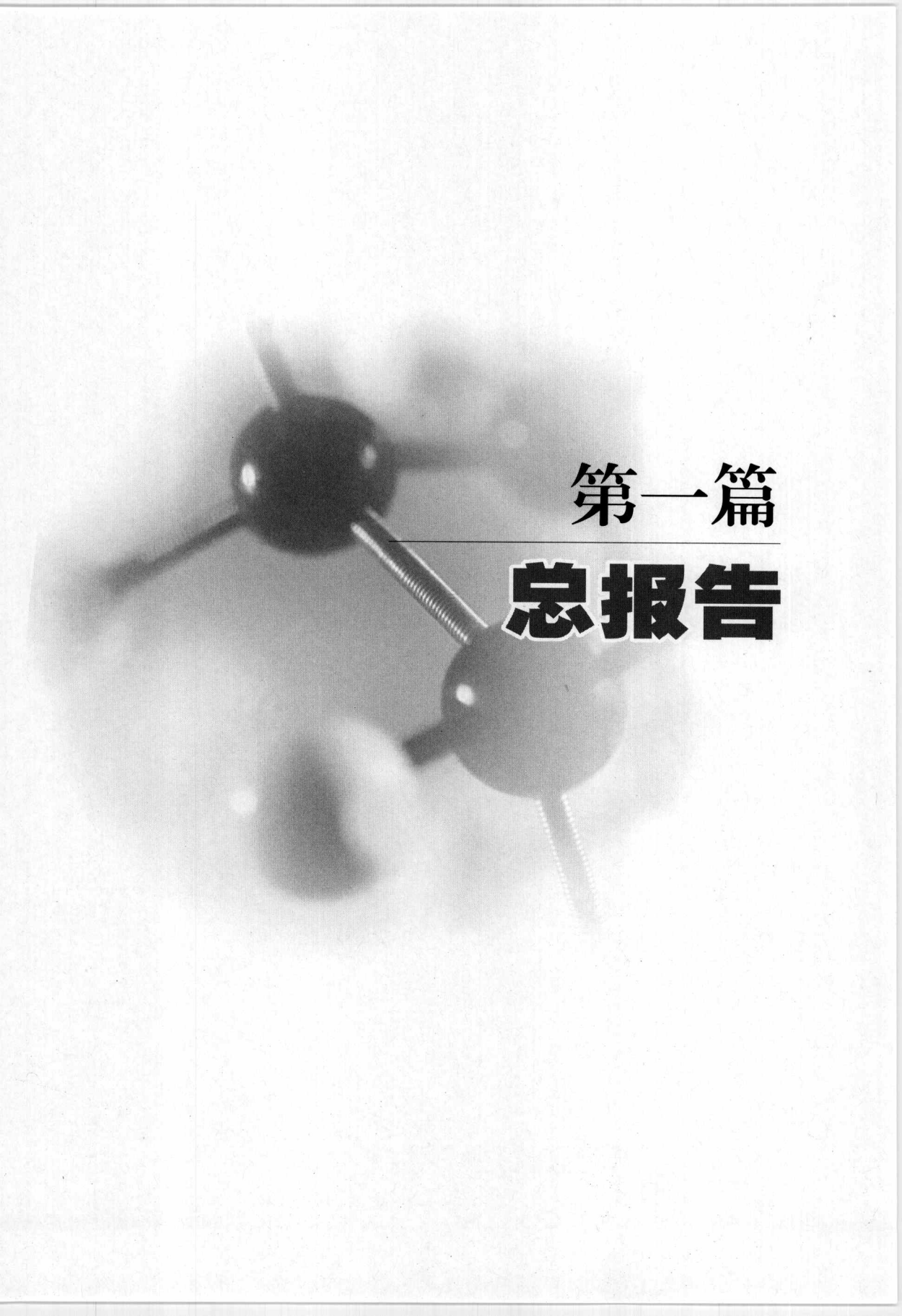
编　者

2007 年 8 月

目 录

第一篇 总报告

第二篇 专题报告



第一篇

总报告

第一章　绪论

一、软科学与社会科学成果整合的必要性

北京作为我国的政治、文化中心，是人文荟萃的地方，具有雄厚的科学研究资源。其中有中央和地方的各类专业研究院所上百个，高等院校近百所，中央和地方的党、政、军政策研究单位和党干校几百个。北京从事科研活动的人员占全国的10％以上，两院院士占全国一半以上，国家的重点科研院所大多在北京。北京地区科研单位承担的各级各类研究项目种类繁多，数量庞大，仅北京市科委软科学处自成立以来设立的项目就达上千项，北京市哲学社会科学规划项目自1983年设立至今，已有1100多项，各委办局、科研院所、高等院校所设立的研究项目更多。广大科研人员通过这些研究课题的研究，创造出了大量的科研成果，在阐释现实、预见未来、传承文明等方面做出了贡献，这是一笔巨大的知识财富。这些研究成果大多数已经在首都的政治、经济、文化、社会建设中发挥了重要的作用，为党和政府的决策提出了许多可行性、可操作性的意见和建议，有的科研成果信息已经辐射全国。

为充分开发和利用这些极其丰富的研究成果资源，北京市科委软科学处和北京市哲学社会科学规划办公室都相继建立了在网上运行的科研成果数据库系统，北京市教委、中共北京市委研究室等单位也建有本系统研究项目的数据库或统计表。但是，目前这些科研成果数据库系统大多数都是独立运行的，研究成果资源未能实现共享，造成了科研成果信息资源的极大浪费。如何避免重复研究，实现成果资源的广泛共享已经成为科研管理工作中的重要工作环节和重要难题。

从未来的发展趋势看，软科学与社会科学成果资源的整合与利用是大势所趋。整合北京地区的丰富的软科学和社会科学研究成果资源，形成成果转化和利用的良好机制，达到研究资源、研究成果的共享，为科研机构和科研

工作者以及党和政府决策机关提供广泛的资源共享及高质量的信息服务，对于推动学术繁荣和理论创新，推动科研管理机制创新，强化科学技术为改革开放和现代化建设服务功能等都具有十分重要的意义。

二、资源整合面临的主要问题

以数据库为基础构建的成果数据库系统，特别是在单位内网或互联网上运行的数据库系统，已经能够在一定范围和一定程度上实现了研究成果信息的共享。但是，由于各单位的成果库系统均独立建设，在系统设计的指导思想、实现目标、技术结构、功能特点、性能指标、维护管理等方面都不尽相同，因而无法实现各系统的成果信息的统一检索和利用。要实现软科学与社会科学研究成果的整合与共享，目前面临的主要问题是：

1. 资源共享理念问题。

首先，各单位对于成果资源价值的认识存在差异；其次，一些单位或部门存在一定的本位主义，对于资源共享带来的综合效益的提升认识不足。

2. 标准化问题。

现有的成果数据库在建设时考虑的是本系统、本单位使用的方便，没有更多地考虑今后的共享，所以各种数据库资源没有统一标准、格式。

3. 数据维护难以持续进行。

很多数据库系统由于设计思想、技术架构等原因，数据添加、更新不方便，最后数据库变成“死库”，信息资源陈旧，逐步失去利用价值。

4. 信息管理、检索等技术不够完善。

其查准率、查全率等信息提取的指标不理想，影响用户对信息资源的使用。

当今世界信息技术的飞速发展，为软科学与社会科学成果资源的整合、共享、管理、发布等提供了技术条件，中文自然语言处理技术的研究能够为中文成果资源的有效利用提供智能、高效的理论支持和实现手段。

三、本课题研究的目标

本课题研究的总体目标是：在强化资源共享理念的基础上，研究设计“北京科学研究成果资源共享平台”（下面也简称“系统”），以科委、规划办

等单位成果数据资料为基础数据源，通过构建规范的软科学及社会科学科研成果数据仓库来实现各系统间的“无缝”连接，对科研成果进行全面的资源整合和共享。

“北京科学研究成果资源共享平台”不仅应为广大科研人员提供信息查询、资源获取、项目查新的便捷渠道，而且能够为各参与单位（如科研管理部门等）提供成果资源有序管理、资源价值分析、制定项目规划及进行科研决策的有效工具。同时通过平台在互联网上的公开运行，也能够起到对科研成果的宣传作用，最大化地发挥科研成果的社会经济效益。

1. **近期目标。**

充分利用“北京社科规划”这一北京市哲学社会科学门户站点，建立“北京科学研究成果资源共享平台”，实现北京市科委软科学与北京市哲学社会科学研究成果资源的整合。

2. **中长期目标。**

邀请北京地区研究机构和高等院校中有研究成果数据库的单位，参加“北京科学研究成果资源共享平台”，通过平台使现有的北京地区网上运行的社会科学成果资源实现整合。

3. **远期目标。**

经过多年的努力，争取达到与外省社会科学研究成果数据库的接轨。

四、方案设计的指导思想和原则

（一）方案设计的指导思想

课题研究将运用知识管理的理念，对北京地区巨大、丰富的软科学和社会科学研究成果资源进行调研，在此基础上，提出对成果资源的整合与利用方案，使“无序”变成“有序”。力求通过调研和深入研究，提出研究成果信息资源整合、开发、利用的途径和技术手段，最后形成一个完整的、具有可操作性的软科学与社会科学两个科研管理单位对成果信息资源的整合与利用方案。

在技术系统的设计上，我们将借鉴国际的信息系统建设方面的先进思想，在保证系统效率的前提下，突出系统的开放性、标准化、模块化、易用性、高效性、可扩展性、安全性、可管理性以及可靠性等特点。

在相关信息系统建设的发展趋势上，目前比较流行效率较高、安全性较强的多层架构体系。本系统也将采用多层架构设计，即数据层（Database Tier）、业务逻辑应用层（Application Tier）、业务逻辑表达层（Presentation Tier）与客户端展现层（Client Tier）。系统采用封装与构件化的技术架构，包括 J2EE 或 .NET 标准的系统体系架构等。前端访问将主要采用 B/S 架构，而对于后端运行的成果数据同步和数据备份恢复部分将使用 C/S 架构。在本系统的信息资源服务方面，我们引入中文自然语言处理技术，为成果资源的有效利用、管理提供条件。同时，本系统将采用数据仓库、商业智能等技术，对软科学和社会科学成果资源进行深入挖掘，提高信息资源的可用性。

（二）方案设计的原则

软科学与社会科学成果资源的整合与利用是基于一系列的原则与规范而开展的。这些原则如下：

1. 效率原则。

效率原则体现在成果数据及信息的获取、传输、处理、施用等各个环节都要有一定的时间限制。在其他原则可容忍的情况下，应尽量通过软硬件优化等手段提高各环节的运作效率。

2. 高可用性原则。

高可用性通过以下三方面体现：

（1）面向业务功能设计，以需求为主导。技术系统设计基于当前信息整合的业务需求，同时充分考虑目前系统现状与未来发展，分析需求、现状、发展三者之间的关联，定位系统的架构与开发范围等。

（2）面向组织架构设计，用户、角色、权限分明。明确技术系统的用户范围、各层次用户拥有的角色，以及对各角色用户提供功能的确定。

（3）面向稳定运行设计，保证业务永续。保证系统运行稳定，有较强的防错、抗错能力。有很强的故障恢复和应急处理措施，使系统核心功能模块不间断地正常运转。

3. 安全性原则。

安全性原则至少包含以下两个层面：

（1）数据安全。在成果数据的传输、存储、加工以及应用的各环节，都必须采取必要的安全措施。尤其在数据访问层，要做好用户认证与权限管理工作，保证用户只能进行权限内的业务操作、获取权限内的业务信息。对于

重要信息要做身份认证，通过用户一次登录识别用户权限，系统为不同权限的用户提供不同的服务，防止越权操作发生。系统分级分层授权，数据分级分层管理，以保证信息的安全保密。

（2）技术系统运行安全管理。对于信息系统来讲，技术手段只是保证其安全性的一部分，建立一套严格完善的安全管理规章制度，并严格遵守，才能从根本上确保系统的安全。应该考虑的安全因素有：机房的安全管理、计算机病毒的预防和清除、关键操作的跟踪和审计等。

4. 可靠性原则。

可靠性原则包含以下三个方面：

（1）网络可靠性保证。

（2）主机、存储系统以及系统软件等可靠性保证。

（3）应用系统可靠性保证，如运行监控、异常和错误处理、数据的备份和恢复等。

5. 可扩展性原则。

成果信息整合目前只涉及软科学和哲学社会科学领域，未来必须可扩展到其他领域，在技术系统层面，还有性能、功能等方面的扩展考虑：

（1）通过增加新增硬件对系统进行扩展，满足不断增加的性能要求。

（2）在应用软件结构上采用多层平台设计，分别支持应用系统的横向扩展和业务系统的功能扩展。

（3）系统模块化设计，可根据需要拆分、组合，系统模块及应用模块的设计开发尽可能为后续功能预留接口。

6. 可定制原则。

可定制性包含以下三个方面：

（1）系统设计采用结构合理的多层架构，将系统的应用服务集中管理。

（2）系统设计采用构件技术，在应用构件的基础上，搭建业务系统。

（3）灵活设计业务数据模型，能适应信息系统的扩展。

7. 易用性原则。

易用性主要包括以下三个方面：

（1）网络、主机、系统软件均要求提供相应的维护工具。

（2）应用系统通过统一的管理界面，清晰的管理向导，简化系统的维护难度。

（3）良好的人机操作界面，用户可以方便快捷地查找到自己所需的信息。

8. **实用性原则。**

实用性是在符合成果资源整合的业务需求的情况下，从使用角度进行技术系统的设计与开发，包含以下四个方面：

（1）方案设计要符合系统总体规划的要求和设计原则。

（2）充分利用成熟的技术。

（3）尽量节省经费的投入。

（4）防止因应用系统在设计上的缺陷而造成系统处理能力不足。

9. **先进性原则。**

先进性包含以下四个方面：

（1）方案设计具有前瞻性、先进性，符合当今社会信息化发展的趋势。

（2）在技术上采用目前流行的多层结构设计，以成熟的开发方式，保证技术系统的先进性和成熟性。

（3）设计时充分考虑到信息技术的发展趋势，采用先进的监测、预警理念。

（4）保证系统能够适应技术的未来发展趋势。

五、本课题研究的主要创新点

本课题的主要创新点包括以下几方面：

（1）软科学与社会科学研究成果整合的相关数据标准规范的研究制订，如成果数字化形式、结构化规范、数据库数据字典规范、成果分类标准等。

（2）提出基于角色管理的成果资源分级授权使用、管理的完整方案，资源整合的系统架构。

（3）在中文自然语言基础研究和信息资源管理方面，进行一些创新研究，如自动分词技术、自动分类技术等。

（4）数据仓库技术在成果库数据挖掘方面的创新应用探索，完成软科学与社会科学成果资源整合与利用的方案设计及相关问题的创新研究，为后续开发相关信息系统提供方法论依据。在文本资源的信息处理和利用上，目前的趋势是逐步深入应用数据仓库技术、商业智能技术（BI）构造知识管理系统，对信息资源进行深入挖掘，极大地提高信息资源的可用性。