

中国科技成果数据库的功能与增值服务

朱艳华

(中国科学院计算机网络信息中心, 北京 100190)

摘要: 中国科技项目创新成果鉴定意见数据库(知网版)收录了1978年以来所有正式登记的中国科技成果数据, 共计56万余项, 集成了科技成果著述信息、知网学术文献总库中的期刊文章、硕博论文、会议论文、标准规范和专利文献等在内的各类资源, 最终形成科技成果知识网络, 并提供了导航功能、检索功能、知网节功能、评价功能和用户管理等各项服务。其中, 知网节功能是对检索结果进行挖掘分析的资源整合模式, 将所有具有密切联系的知识集成在一个汇编页面上。此外, 文章基于语义分类导航理念, 探索了数据库增值服务的新模式。

关键词: 科技成果数据库; 增值服务; 知网节; 知识关联; 集成服务

中图分类号: TP392

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2014.02.011

Construction and Value-added Service of National Database of Scientific and Technological Achievements

Zhu Yanhua

(Computer Network Information Center of CAS, Beijing 100190)

Abstract: The database of scientific and technological achievements has great significance to promote the transformation of scientific and technological achievements and knowledge innovation. Nation Achievement of Science and Technology Database (CNKI Version) collected more than 560,000 achievements from 1978 till now. It mainly provided five kinds of function: Navigation function, search function, knowledge network node function, evaluating function and user management function. Knowledge Network Node is a special search-result display page created by CNKI. It is similar to citation links but much more powerful. The purpose is to improve research efficiency and make all closely related articles in contents available on a single page. Compared with usual achievement databases, every node of achievement in Nation Achievement of Science and Technology Database (CNKI Version) collects its bibliographic information and related journal articles, doctoral dissertations, masters' theses, proceedings papers, newspaper articles, standards, patents and so on, which finally formed the knowledge hierarchy of the achievement. In addition, this paper explored the new mode of database value-added services based on semantic navigation.

Keywords: Nation Science and Technology Achievement Database, value-added service, Knowledge Network Node, knowledge connection, integrated service

作者简介: 朱艳华(1982-), 女, 中国科学院计算机网络信息中心工程师, 硕士, 主要研究方向: 数据库技术与标准规范、数据应用服务。

收稿日期: 2013年11月12日。

1 引言

中国科技项目创新成果鉴定意见数据库（知网版），以下简称科技成果数据库（知网版），是同方知网（CNKI）与国家科技成果管理项目组（中国化工信息中心）联合建设的国家级大型权威性科技数据库，内容涉及医药卫生、农业、化工、工业、环境、自动化、电工、机械、建筑、交通、无线电、冶金和社会经济等方面。科技成果数据库收录了1978年以来所有正式登记的中国科技成果，部分成果回溯至1920年，并按行业、成果级别和学科领域进行分类。每条成果信息包含成果概括、立项情况、评价情况、知识产权状况及成果应用情况、成果完成单位信息和成果完成人信息等基本描述数据。截止到2013年，科技成果数据库（知网版）共收录科技成果56万余项，科技成果每月更新，相关科技文献每周更新。产品服务形式包括WEB版（网上包库）、镜像站版和流量计费^[1]。

科技成果数据库收录正式登记的国家科技成果与相关科技文献，收录范围共涉及31个省、自治区、直辖市。相比于同类数据库的简单描述，该数据库的内容信息达50余项，具体字段详见表1。其中，成果编号作为成果数据库的唯一标识，保证库中数据没有重复和冗余。众所周知，数据质量问题及其研究由来已久，特别是伴

随着计算机为主的信息技术发展而逐渐成为被广泛关注的研究热点，在大数据时代来临的当下，其重要性已经比肩数据资源本身。对科学数据资源而言，数据质量对科研活动的重要作用比之产品有过之而无不及，科学数据质量将直接影响到我国科技发展整体水平的提高与独创性成果的产出^[2]。科技成果数据库作为科学数据的重要类型之一，其数据质量的重要性不言而喻。成果数据库的关键字段都由人工进行审查和校对，其中，成果名称、成果编号、成果类别、成果入库时间、关键词、学科分类、中图分类、成果摘要、成果主要应用行业代码、第一完成单位、第一完成单位代码和所在省市等字段完整率达100%，准确率达到95%以上。

与通常的科技成果库相比，科技成果数据库（知网版）每条成果检索结果界面集成了与该成果相关的最新文献、专利和标准等信息，可以完整地展现该成果产生的背景、最新发展动态、相关领域的发展趋势，还可以浏览成果完成人与完成单位在各种出版物上发表的论文等更多信息。可以说，科技成果数据库通过增值服务功能充分展示了基础研究—成果产出—专利产出—标准应用—产业化的过程及内在的必然联系，每条成果都囊括了该成果及其相关领域的所有科技发展信息。

本文将对科技成果数据库（知网版）的内容、功能及服务作进一步论述。

表1 科技成果数据库内容

数据内容	说明
成果概要信息	成果名称、成果编号、成果类别、成果入库时间、关键词、学科分类号、中图分类号、成果摘要、研究起止时间、成果水平、成果体现形式、成果属性、成果应用阶段、研究形式等
立项情况	课题来源、课题立项名称、课题立项编号、经费实际投入额等
完成人情况	成果完成人姓名、性别、出生年月、工作单位等
完成单位情况	项目联系人名称、项目联系人电话、项目联系人电子信箱等
第一完成单位情况	第一完成单位、第一完成单位代码、所在省市、单位地址、单位属性、邮政编码、联系电话、联系传真、单位网址、联系部门、联系部门电话、联系部门E-mail、单位简介、研究领域、研究学科、研究方向简要说明、优势领域、优秀学科、研发优势简要说明等
完成单位学术带头人	姓名、性别、出生年月、研究领域、技术职称、文化程度、简要说明等
评价情况	评价方式、评价单位、评价日期、评价报告编号等
知识产权状况及成果应用情况	转让范围、知识产权形式、应用状态、推广形式等

2 功能设计

知网版科技成果数据库首先对成果数据库原有字段内容进行梳理,检查字段名称是否规范,数据内容是否完整;确定丰富的检索词;将整理后的成果数据库相关字段和CNKI数据库中元数据字段进行映射,以实现跨库检索和增值服务等功能。表2列出了详细的字段对应关系。

表2 科技成果数据库与CNKI元数据检索字段对应关系

成果库检索字段	CNKI元数据字段
成果名称	中文题名
关键词	中文关键词
成果摘要	中文摘要
中图分类	168专题分类
课题来源	基金来源
成果入库时间	发表时间
成果完成人姓名	作者中文名
成果完成人单位	作者机构
成果合作/协作完成单位	作者机构
成果所在省市	单位地址

基于原有成果数据库的功能特点和检索模式,科技成果数据库(知网版)重新设计了5种主要的服务功能,即导航功能、检索功能、知网节功能、评价功能和用户管理等(图1)。



图1 科技成果数据库功能设计图

2.1 导航与检索功能

原有成果数据库按照《中国图书资料分类法》(第四版)对成果数据进行中图分类;按照GB/T13745《学科分类与代码》进行学科分类。所有CNKI数据库均采用统一的168专题分类体系。该分类体系按中国图书分类法对每个出版物及其每篇文献分类,汇编为8个专业总库、168个专题数据库产品。按规定的学科范围,每个专题数据库均可完整收录在同方知网学术文献总库中,收录所有出版物上发表的该学科专业文献,并在“知网节”整合模式和标准化检索平台支撑

下,构成学术文献的权威性检索工具,形成全国学术文献的增值性集成整合传播媒体^[3]。

科技成果数据库(知网版)对成果库原有的分类进行清理,保留原有中图分类体系和学科分类系统,并由同方知网总编室将中图分类号与CNKI的168专题号进行逐一对应,形成映射关系表,确定每条记录的168分类号。因此,整合后的成果数据库共采用3种导航方式,即中图分类导航、学科分类导航和168专题导航。

科技成果数据库(知网版)实现了初级检索、高级检索、专业检索、在结果中检索和智能扩展功能等多种检索功能,力求做到让用户的操作轻松容易。科技成果数据库检索功能详见表3。

(1)所属年份:年份以下拉表的形式列出,范围是1978-2013,可选择。(2)逻辑运算:可增加或减少逻辑行,逻辑关系为“并且/或者/不包含”。(3)检索词扩展:以检索词为中心的相关词,可显示与输入的检索词在意义上相关的词。相关词以3种方式自动添加到检索框中:单词自动增加、多词自动增加或相关词取代原输入词。在相关词前面的方框中以“√”的形式选择相关词,则该词自动以“逻辑与”的关系增加到检索框中;或点击所需要相关词,则该词自动进入检索框并取代原先所输入的检索词。(4)匹配方式:模糊或精确,以下拉表的形式列出,可选择。(5)排序:时间、排序码、无、相关度,以下拉表的形式列出,可选择。(6)每页:以下拉表的形式列出每页保存的记录数:10/20/30/40/50,可选择。(7)中英文扩展:当匹配方式为精确时,可以“√”形式在方框中选择,当匹配方式为模糊时,则方框变灰为不可选择。(8)检索导航:单库检索导航为学科分类和CNKI-168专辑导航两种可选导航,导航以树型方式列出,可以层层展开,可在导航的各栏目名称前以“√”形式选择,以控制检索的范围。(9)二次检索:在第一次检索结果的基础上再次检索,以缩小检索范围。(10)智能扩展功能:检索智能扩展基于概念关系词典相关语义运算技术,实现不同字段间的智能扩展,以提高查全

率，发现新知识，实现知识漫游服务。

2.2 知网节功能

以每一篇文章为知识网络中的一个传播节点，简称“知网节”。知网节将一篇文献的参考文献、引证文献、相似文献、读者推荐文献、同类文献、相关文献以及文献作者与机构有关信息、知识概念等内容链接整合到一起，通过知识网络有序性地集成文献的个性内容与整体性知识。它以知识整合传播为设计理念，基于知识网络、知识挖掘、知识发现和互动传播等网络传播模式与技术，使文献检索、文献传播的基本概念与价值观产生质的飞跃^[4]。

基于知网节理念，科技成果数据库（知网版）以每条成果为基本单元，在整合原有成果库描述信息基础上，集成了包括CNKI期刊文章、硕博论文、会议论文、标准规范和专利文献等在内的各类资源，最终形成科技成果的知识网络，见表4。

概况地说，知网节功能内容主要包括成果3个方面，知网节功能模块和详细说明见表4。

2.2.1 成果完成人节点功能

(1) 第一成果完成人其他成果：展示第一成果完成人的其他科技成果，并对完成周期进行分析。

(2) 完成单位其他完成人成果：同一机构/合作机构的其他成果完成人的成果，利用第一完成单位和合作完成单位等字段加以分析。

(3) 成果完成人申请的专利：成果完成人申请到的专利信息，主要统计成果完成人在专利库中发表的专利信息^[5]。

(4) 成果完成人完成的标准：成果完成人研发的标准规范；统计成果完成人在标准库中的标准信息^[6]。

(5) 成果完成人发表的文献：成果完成人在期刊数据库、学位论文数据库、会议论文数据库、报纸数据库中发表过的各类文献信息，评价成果完成人的科研能力。

2.2.2 成果完成单位节点功能

(1) 完成单位其他成果：利用成果完成单位字段统计其在成果库中登记的其他成果，了解成果完成单位的科研能力。对企业来说，可以了解竞争对手的情况。

(2) 各省市在学科领域的成果数量：统计成果完成单位所在省市某一学科领域成果完成情况。

(3) 本领域成果地域分布：统计出各省市在某一学科领域完成的成果数量，内各省市成果数量对比，可推知某领域的高科技研发力量集中的区域。

表3 科技成果数据库检索功能

功能		说明
检索方式	初级检索	通过系统提供的初级检索和各种逻辑检索行，选择检索项和输入检索词，能够找到符合条件的成果数据
	高级检索	通过系统提供的逻辑检索行和逻辑关系（并且、或者、不包含），选择检索项和输入检索词，能够找到符合条件的成果数据
	专业检索	通过系统提供的专业检索功能，输入检索条件，能够找到相应的数据
	在结果中检索	二次检索，在第一次检索结果的基础上再次检索，缩小检索结果范围
	概念关系词典	基于概念关系词典设置检索智能扩展功能和概念导航功能
检索项		通过成果名称、关键词、成果摘要、中图分类号、学科分类、课题来源、成果入库时间、成果类别、成果水平、成果主要应用行业、鉴定单位、鉴定时间、成果完成人姓名、成果完成人工作单位、第一完成单位、所在省市、成果合作/协作完成单位名称等检索字段找到相应的成果数据
检索控制		通过对数据库种类、中图分类号、匹配方式、检索结果排序、中英文扩展等控制检索条件，缩小检索范围，更快地定位符合条件的成果数据
检索结果	题录方式浏览	通过题录方式显示检索结果
	保存功能	对检索到的结果进行保存操作
	排序功能	按年份、排序码、相关度等正序和逆序对检索结果进行排序
	分组功能	针对检索结果按年份、名称、关键词等进行内容聚类

(4)完成单位申请的相关专利:成果完成单位申请的专利,统计成果完成单位在专利数据库中的专利。

(5)完成单位研发的相关标准:成果完成单位研发的标准;统计成果完成单位在标准数据库中的标准。

2.2.3 成果相关内容节点功能

(1)相似成果:本领域中与该成果内容相似的成果;利用关键词和中图分类号等字段在成果库中查找相似的数据记录。

(2)相关标准:与本成果内容相关的标准规范;利用关键词等字段在标准数据库查找相关标准规范。

(3)相关专利:与本成果内容相关的专利信息;利用关键词和专题子栏目代码等字段在专利数据库查找相关专利信息。

(4)成果研制动态:本成果研究过程中产生的各类研究文献,如期刊论文、学位论文和会议论文等;利用关键词、课题来源等字段在期刊数据库、学位论文数据库、会议论文数据库中查找

相关文献。这些信息显示本成果所涉及或参考的研究文献。

(5)成果应用动态:本成果有关应用转化动态新闻,利用关键词等字段在报纸数据库中检索相关信息。

2.3 其他功能

评价功能主要对成果完成人或完成单位所产生的成果、专利、文献信息进行统计分析。用户管理包括用户日志统计和用户日志查询,其中,用户日志统计按照IP地址、专辑、专题分库统计日志;超级用户可以按操作类型、时间范围等进行日志查询。

3 增值服务模式

科技成果数据库利用知网节技术将成果数据与标准、专利、期刊、报纸、博士论文、硕士论文、会议论文、图书、工具书、知识元等资源进行整合,深入挖掘科技研究信息与科技产出之间的内在关联关系,形成完善的科技研究与科技成果产出综合信息服务平台。充分反映基础研究-

表4 科技成果数据库知网节

模块	项目	功能说明
成果产出状态分析	成果完成人其他成果	列表显示本成果完成人的其他成果,可直接点击浏览
	完成单位其他成果	列表显示本成果完成单位的其他所有成果;可直接点击浏览;通过年度分布图分析成果完成单位成果产出情况
	相似成果	列表显示与本成果主题高度相似的成果;统计分析各年度相似成果分布;反映成果水平不断提高的过程
	成果核心技术地域分布	图表统计成果核心技术领域所有成果的完成单位地域分布;分析相关核心技术研发力量分布
本领域专利与标准	相关专利信息	列表显示与本成果主题相关的专利发明;可直接点击浏览;反映本领域专利产出情况
	相关标准信息	列表显示与本成果主题相关的中外标准;可直接点击浏览;反映本领域中外标准产出情况
本成果研制背景	期刊论文、硕博论文、会议论文、报纸文章	分别列表显示成果完成前发表的核心技术相关文献;显示成果研制期间的核心态势和成果产生的背景
本成果应用动态	期刊论文、硕博论文、会议论文、报纸文章	分别列表显示引用本成果的各类文献;反映本成果的研究与技术参考价值
完成人发表文献	期刊论文、硕博论文、会议论文、报纸文章	分别列表显示成果完成人各年度发表的文献;显示成果完成人的研究能力
完成单位发表文献	期刊论文、硕博论文、会议论文、报纸文章	分别列表显示成果完成单位各年度发表的文献;显示成果完成单位研究能力
知识链接	相关机构	对相关成果完成单位做进一步的了解
	相关专家	对相关成果完成人做进一步的了解

成果产出—专利产出—标准应用—产业化的过程及内在的必然联系，并综合分析和评价机构（个人）的科研与科技产出绩效。

除了利用知网节技术提升成果数据库增值服务能力之外，为推进未来科技成果的转化，响应中共中央、国务院提出的《关于深化科技体制改革加快国家创新体系建设的意见》，充分发挥企业在技术创新决策、研发投入、科研组织和成果转化中的主体作用，我们还要进一步探索成果数据服务的新模式，加强数据内容挖掘力度，强化产学研用紧密结合；开发针对技术开发、科技创新活动和企业创新活动的特色服务产品；扩展成果数据库开放共享范围，增强支撑产业创新能力，吸引各类创新主体协同合作，加快研发成果的转化工作。

基于成果数据库已有内容，以增加多种语义分类导航为切入点，探索数据库增值服务的新模式。语义导航作为知识组织的方式之一，比传统的学科分类导航或主题词分类导航具有更强的逻辑性和内容分析挖掘能力，能够针对特定服务对象，有效引导企业用户进行快速检索和需求定位。举例如下。

（1）增加“成果高新技术领域分布”导航，方便用户迅速查找到其感兴趣的领域成果。成果所属高新技术领域可以划分为：电子信息；软件；航空航天；光机电一体化；生物、医药和医疗器械；新材料；新能源与高效节能；环境保护；地球、空间与海洋；核应用技术；农业等。

（2）增加“成果体现形式”导航，准确定位成果类型。成果形式主要包括：新技术；新工艺；新产品；新材料；新装备；农业、生物新品种；矿产新品种；其他应用技术等。

（3）增加“成果推广形式”导航，明确成果知识产权状况和成果应用情况。具体推广形式主要包括：产权转让；资金入股；技术入股；合作开发；技术服务等。增加“成果水平评价”导航，使用户对成果水平有一个整体的认识和评估。成果评价等级具体可分为国际领先、国际先进、国内领先、国内先进、国内一般等。一般来

讲，成果得到的评价越高，未来转化的潜在价值就越大。

4 结语

全面收集和整理科技计划项目产生的科技成果具有重要意义，科技成果数据库不仅用于成果查新和技术转让，还可以为技术咨询和服务提供重要信息来源，是技术改造、新产品开发以及革新工艺的重要依据。我国经济高速发展很大程度上来自科技成果的有效转化，因此，加强科技成果数据库建设，推动成果产业化，特别是将具有自主知识产权的科技成果转化成真正的产能具有现实意义。

科技成果数据库收集整理正式登记和上报的中国科技成果信息，提供查新服务和技术咨询信息，对于推动成果转化具有重要意义。科技成果数据库（知网版）旨在打造产学研综合信息平台，揭示科技发展与创新奥秘。成果数据库提供了导航功能、检索功能、知网节功能、评价功能和用户管理等多种服务模式，同时还将成果数据与标准、专利、CNKI期刊、报纸、博士论文、硕士论文、会议论文和知识元等资源进行整合，深入挖掘科技研究信息与科技产出之间的内在关联，形成完善的科技研究与科技成果产出综合信息服务平台。

参考文献

- [1] 中国科技项目创新成果鉴定意见数据库（知网版）[EB/OL].[2013-07-01]. <http://epub.cnki.net/KNS/brief/result.aspx?dbprefix=SNAD>.
- [2] 胡良霖,黎建辉,刘宁,等.科学数据质量实践与若干思考[J].科研信息化技术与应用,2012,3(2):10-11.
- [3] 万方知网学术文献总库平台[EB/OL].[2013-06-10]. <http://www.cnki.net/>.
- [4] 王明亮.整合传播——网络时代科技文献检索工具的本质特征[C]//第二届国际科学编辑研讨会论文集,2005.
- [5] 中国专利数据库[EB/OL].[2013-06-18]. <http://dbpub.cnki.net/Grid2008/Dbpub/Brief.aspx?ID=SCPD&subBase=all>.
- [6] 中国标准数据库[EB/OL].[2013-06-18]. <http://epub.cnki.net/kns/brief/result.aspx?dbPrefix=CISD>.