

国家工程技术图书馆知识服务模式探究*

□ 王立学 / 中国科学技术信息研究所 北京 100038

摘要: 基于自建和订购的丰富的馆藏知识资源, 面向科研人员、科研机构 and 科研管理机构等用户的需求, 国家工程技术图书馆从知识资源规范着手, 在对文献知识资源要素开展多角度知识分析和知识关联的基础上, 抽取各类文献知识要素构建知识链接系统, 构建了国家工程技术图书馆知识服务体系。文章主要从机构规范文档建设、知识分析角度和要素以及知识服务体系的框架和功能三个方面加以介绍。

关键词: 国家工程技术图书馆, 知识服务, 知识分析, 规范文档

DOI: 10.3772/j.issn.1673-2286.2013.10.006

引言

在信息化、数字化的时代背景下, 文献知识资源的访问与获取越来越便捷, 这既加剧了图书馆面临的到馆用户日渐减少、网络用户也逐渐流失的窘迫处境, 资源优势越来越不明显; 又让许多用户对海量的信息感到无所适从, 在泛滥的信息中去粗取精耗费了大量时间和精力, 难以获得有用的文献和知识。面对信息技术迅猛发展的冲击、用户更加专业化和个性化的需求, 新时代图书馆的使命已经远不只是收集资源并服务, 而是应该加强对文献知识资源的整理、分析和关联, 甚至利用图书情报学科专业的方法和技术手段创建知识资源, 从而为用户提供知识密集型的知识服务。

知识服务是知识与服务的融合。对于知识, 我们可以理解为结构化的、相互链接的、不断增长的信息及其间存在的复杂的相互关系^[1]; 而对于知识服务则有多种理解: 联合国开发计划署 (UNPD) 将其定义为“建立在全球知识技术状态上的建议、专家意见、经验和试验办法, 它被提供来帮助请求者获得对问题的最好解答”^[2]; 学者戚建林认为, 知识服务广义上指一切为用户提供所需知识的服务, 狭义上则特指针对用户专业需求对相关知识进行处理的较深层次的智力服务^[3]; 学者张晓林将其定义为“以信息知识的搜寻、组织、分析、重组的知识和能力为基础, 根据用户的问题和环境, 融入用户解决问题的过程之中, 提供能够有效支持知识应

用和知识创新的服务”^[4], 这一定义获得了许多学者的认同。实际上, 我们无须困惑于准确界定知识服务, 知识服务的关键在于拥有加工分析知识的能力, 能够通过知识层面知识组织的能力, 根据用户的需求解决问题^[5]。

从学术论文来看, 国内学者对于知识服务的研究较为丰富, 据笔者2013年2月从万方期刊的统计, 2003-2012年共有1405篇主题包含或涉及“知识服务”的学术论文, 总共被引4282次。不过, 2012知识服务专家论坛会议纪要中总结到, 目前国内关于知识服务的文章较多, 而实际应用与具体案例则较为匮乏^[5]。比较有代表性的研究有: 学者任树怀基于图书馆2.0构建了包含有学科馆员、学科团队、交互空间、信息及知识库等几个核心元素的学科知识服务平台^[6]; 学者王华伟等以材料学科为试点, 构建了包含学科分类、知识主题、规则、信息点、学科词表和学科专家六种要素在内的高校学科知识服务体系^[7]; 秦晓珠和李晨晖等学者分别探讨了知识服务与当前研究热度很高的大数据的结合^[8,9]; 在集成化方面, 万方数据建设了具有文献检索、知识脉络、论文相似性检测、查新咨询等多项功能的知识服务平台^[10], 中国知网也推出了行业知识服务平台。一般认为, 知识资源、服务能力和用户需求是知识服务的三大要素^[11]。当前, 图书情报服务的核心能力不在于所拥有的资源, 而在于所具备的利用广泛信息资源为用户创造价值的知识和能力^[4], 图书情报机构需要利用分布式的动态资源, 为用户提供集成化、学科化和个性化的知识服务。

* 本研究得到了国家社科基金青年项目 (编号: 11CTQ019) 和中国科学技术信息研究所预研基金 (编号: YY-201207) 的资助。

1 国家工程技术图书馆知识服务的背景

国家工程技术图书馆作为一个以网络服务为主的图书馆,成立伊始恰逢我国知识服务研究从概念、模型讨论到技术、服务实践大范围兴起^[12]的契机,故一直就将加强知识服务能力建设作为服务工作的重点。为此,国家工程技术图书馆一方面注重基础知识资源建设,为开展知识服务打造稳固的基础;另一方面下大力气开展知识资源的分析和关联,持续打造新型知识服务能力。2001年,中信所研制开发了《中国科技期刊引证报告》^[13],为探索知识服务开了个好头;2007年,基于学术论文间的引用关联,研发了中文知识链接服务系统^[14],能够为用户提供论文、机构、基金、地区、期刊等多个知识检索入口,迈出了探索国家工程技术图书馆知识服务的重要一步;2010年,基于中、外文高被引论文数据,先后研究、开发了机构成果分析、学科成果分析^[15]等多个集成分析系统,基本确立了国家工程技术图书馆的知识服务体系的核心^[16]。从知识到服务,需要在语义、逻辑和功能等方面经过很大的跨越;而建设一个知识服务体系,更是需要在知识资源建设和规范上下足工夫,在知识资源分析上追求全面精准,还要合理设计知识服务体系的功能和用户接口,本文将着重介绍国家工程技术图书馆构建知识服务体系过程中在上述几方面的研究和实践。

2 知识资源组织与规范

国家工程技术图书馆重点采集国内外工程技术领域的期刊、会议文献、科技报告、学位论文、专利等知识资源;在此基础上,建立中文论文引文数据库,投入大量人力、物力每年加工引文数据,目前引文数据量已达到数千万条。借着承担国家“十一五”科技支撑计划项目的契机,重点针对现有文献数据规范工作缺少统一标准、规范流程和处理细则等问题展开研究,研究制定了基础文献库建设的一系列标准和规范,设计了中文期刊、论文及引文数据、作者、作者机构、基金等项目的名称规范规则和元数据集,产出并应用了一整套“现有文献数据的规范、整理和加工”的流程和检验工作流程,每年新增论文数据300多万条。

为了保证知识服务的质量和效果,论文数据中的题名、作者、机构、引文等各要素都必须准确、规范地著录。举例来说,一条论文数据就有可能涉及以下20余

种著录加工项:

作者、机构、摘要、关键词、分类号、所属学科或领域、主题词、论文类型、栏目、参考文献、资助基金、致谢、来源、语种、年、卷、期、页码、收稿时间、录用时间、刊出时间……

由于机构撤销并改名、期刊更名等多种现象的存在,必须建立机构、作者、期刊、基金等专门的规范文档来维护和更新。国家工程技术图书馆建立的机构规范文档涵盖以下5类数据:

- 机构规范名称和非规范名称对照
- 机构基本属性:所在地域、机构类型、所属学科等
- 机构上下级隶属关系
- 机构中英文名称对照
- 机构名称变迁

机构规范文档中的主要元数据项见表1。此外,国家工程技术图书馆还同时建设、维护着作者规范文档、基金规范文档、期刊规范文档,并且应用汉语主题词表对主题词进行规范。这些规范措施有力地确保了知识服务的质量和效果。

3 知识分析角度及要素

知识服务的特点在于面向知识内容和解决方案,这就意味着图书馆的核心能力不再来自所拥有的资源,而来自于全面利用多种信息资源、深入分析知识关联并为用户创造价值的知识和能力。以围绕学术论文展开的知识分析为例,目前国内此类实践比较多,从围绕关键词开展的主题关联分析到对论文作者、机构等文献外部特征的计量和关联分析都有涉及,常见的关联有:期刊、作者或机构的发文量和被引量,作者或机构合著,关键词词频和共现等。

概括说来,学术论文中包含有许多内外部特征项,每个特征项都有可能单独或与其他项联合来开展知识分析。按照知识分析的要素来分,可以大致分为知识内容及其关联分析、知识产出主体分析和综合知识分析;而从分析的角度来看,又可以有数量、时间、内容、领域、结构、交叉性、引用、地域等多种划分,详情见表2。

为支持从内容、领域、结构和交叉性等角度对学术论文进行知识分析,相关的主题标引、文本主题分析、结构和关联分析等工作必须有力地展开;同时,对作

表1 机构规范文档主要元数据项

字段代码	字段名称	字段代码	字段名称
INST_ID	原始机构ID	Pa_INST_ID_UNI	一级机构规范ID
INST_Name	原始机构名称	Pa_INST_Name_UNI	一级机构规范名称
INST_ID_UNI	机构规范ID	Pa_INST_Name_En_UNI	一级机构规范英文名
INST_Name_UNI	机构规范名称	INST_type	机构类型
Superior_ID_UNI	上级机构规范ID	INST_discipline	机构学科划分
INST_grade	机构级别	INST_Name_En	原始机构英文名称
INST_code	机构代码	INST_Name_En_UNI	规范机构英文名称
INST_Province	机构所在省	Register_date	机构成立时间
INST_City	机构所在市	Revocation_date	机构撤销时间
Postal_code	邮政编码	Change_date	变更时间
INST_Address	机构地址	Change_reason	变更原因
Current_Pa_INST_ID_UNI	变更后一级机构规范ID	Original_Pa_INST_ID_UNI	原始一级机构规范ID

表2 知识分析角度和要素概览

分析角度 分析要素		数量	时间	内容	领域	结构	交叉性	引用	地域	
知识 内容	概念	研究热点	沿革及趋势	内涵与外延	背景及关联	属分关系	知识体系	参照关系	主题地域	
	关联	关联强度	知识演变	语义、相关等 关系	知识结构	核心概念、知 识结构	领域融合	知识链接、 领域交叉	分布	
知识 产出 主体	论文	产出能力	趋势、热点、 发文时滞	主题或内容分析	领域分析	文本结构解析	相似性度量	核心主体发现或 识别、影响力评 估、知识链接、 相关主体发现或 识别等	地域科研 力量分布	
	作者	研究活跃度	个人信息变化	研究兴趣点分布	相关(合作/竞争 等) 人员/机构、 领域关注度	科研队伍分析	研究者单位变 迁			
	机构	研究群体	机构沿革	科研布局、 优势方向		属分关系				
	基金	关注度	科研布局、研究重点、学科 交叉、研究前沿等及其演变		战略布局、领域 分布、多学科性	类别设置	相关基金识别			科研投入
	期刊	交流活跃度	期刊沿革	主题分布、 相关期刊	领域交叉	栏目划分	同类期刊识别			科研交流 环境
会议	活跃度、关注 度	会议沿革	热点主题、研究前沿		研究方向分布	相关会议识别				

注：表2引自文献[17]，有修改。

者、机构、基金等项目的著录也必须准确、规范、全面。而这些，正是国家工程技术图书馆前期投入巨大精力所开展的基础性工作。因此，国家工程技术图书馆在利用关键词共现、论文共被引、参考文献耦合和论文合著等文献计量学领域的关联分析方法建立文献知识资源间关联时，才能够较为精准地开展分析。

4 知识服务框架与功能

由于用户群体的行业和地域分布比较广泛且分散，同时也受人力资源的限制，以网络服务为主的国家工程技术图书馆充分发挥自身在知识资源建设和规范方面的优势，规范、严谨地加工和组织文献知识数据，深入、全面地研究知识分析的相关内容，引导用户通过网络进行自助知识服务^[18]。

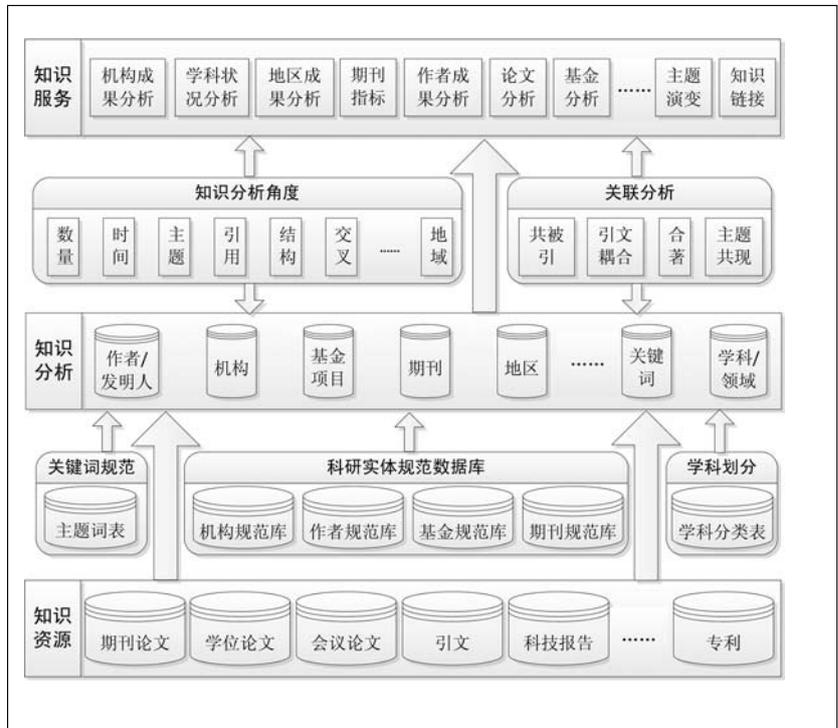


图1 国家工程技术图书馆知识服务架构

4.1 知识服务框架

经过近十年的实践和摸索，国家工程技术图书馆建立了如图1所示的知识服务体系架构，简单说来可以归结为：三层数据、两种分析，一个窗口。

①三层数据。底层数据是国家工程技术图书馆采集和加工的中文期刊论文及其引文、中外文学位论文、中外文会议论文、美国科技报告、全球专利全文、外文期刊论文及其引文（主要是SCI和EI等数据）等知识资源；中间层数据是日常应用且持续更新维护着的机构规范库、作者规范库、期刊规范库和基金规范库，以及用于关键词规范的主题词表，此外还有相对固定的学科分类表；上层数据则是从知识资源中抽取出的、经过各个规范文档规范过的知识分析要素，包括作者、机构、基金、期刊、主题词等。

②两种分析。第一种分析针对底层知识资源开展，经过实体识别、数据挖掘、语义分析等步骤，将有关的知识分析要素识别并抽取出来；第二种分析是关联分析，通过文献计量学特有的共被引、合著、引文耦合等关联方法，对抽取出的知识分析要素进行关联，再基于各种分析角度组织相应的分析数据。

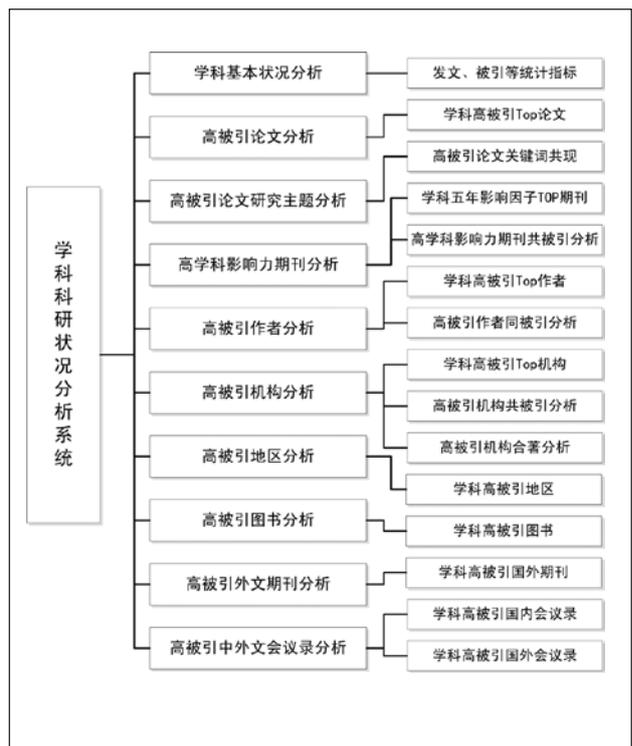


图2 学科科研状况分析系统的分析项

③一个窗口。上述所有的后台操作的结果最终都将体现在国家工程技术图书馆网站的知识服务窗口上。

4.2 知识服务功能

目前,国家工程技术图书馆知识服务体系中的知识链接、期刊指标、论文分析、作者分析和基金分析等功能已经上线服务(<http://www.istic.ac.cn>),学科科研状况分析和机构学术成果分析两个子系统正在内测。其中,学科科研状况分析系统的分析项如图2所示。

系统以中文学术论文及引文作为分析数据源,结合中图分类法和各学科实际情况来划分学科,并分别统计学术论文的发表和被引情况,既从整体上展现学科内论文发表和被引的数量分布和地区分布等概况,又从期刊、作者、机构、图书、会议等方面反映学科内的学术影响力情况,还进一步用共现、同被引以及合著等方法揭示各学术主体之间内在的主题关联。主要分析指标和分析项目有:

①发文/载文指标和被引频次指标。从实际数据统计来看,中文期刊论文的被引主要集中在发表后5年以内。因此,系统按学科划分来统计5年来的发文情况以及在最近1年的被引情况,分别从总量和学科内的角度统计作者、期刊和机构等对象的发/载文量、被引频次和相关的比例数据(被引率、篇均被引、影响因子等)。借助这些基本的统计指标,用户可以快速了解学科5年来的概况。

②高被引分析项目。按学科划分,从最近1年内发表的论文中统计各种分析对象的被引频次,首先按被引频次高低将前1%的论文定义为学科高被引论文,再进一步统计筛选出各学科内位列Top N的高被引论文、高学科影响力期刊、高被引作者、高被引机构、高被引图书、高被引学术会议录、高被引国外期刊以及h指数等项目。学科内的各类高被引对象反映了最近一年中科研人员热点研究主题的分布,以及学科内影响力较大的学者、期刊、机构、图书、学术会议和国外期刊,便于科研人员掌握最近一年学科的关注焦点及重要的科研主体。

③文献计量关联分析。同样从学科视角出发,借助关键词共现分析、论文共被引分析和论文合著分析等三种文献计量方法,分析、挖掘并揭示学科内科研主题间的潜在学术关联。

● 关键词共现分析。除了统计关键词的文档词频,系统还抽取了各学科中高被引论文的施引文献,从论文的作者关键词字段中两两统计关键词共现关系,希望以学科内的基础性论文为线索,分析获得学科内最近一年的研究前沿。

● 论文共被引分析。利用论文共被引分析,获得高学科影响力期刊的载文主题关联、高被引作者发文主题关联和高被引机构研究主题关联,为科研人员揭示出关联期刊、优势学者和优势科研机构,为他们寻求合作对象等需求提供线索和依据。

● 论文合著分析。以高被引作者和高被引机构为线索,揭示他们之间及他们与其他作者和机构之间的科研合作关联,向科研人员揭示出学科内已经或正在形成的科研团队。

5 问题与展望

在知识服务的概念兴起之初,始终追随着知识经济这一概念,但经过十余年的发展,我们没有停留在概念和模型论证上,而是深入研究、积极开拓、勇于实践,我国各级各类图书情报机构的知识服务现已开展得有声有色。通过前面的介绍,国家工程技术图书馆与知识资源相关的文献主题特征和文献外部特征都加以规范,无论是对知识元的统计、分析和关联还是对多个知识要素的计量、关联和评价都提供了重要支撑,使得从不同角度开展的查询、统计、分析、关联挖掘和影响力评价等知识服务活动得以顺利开展。从研究实践中归纳得出的常见知识分析的角度及要素,为知识服务提供了放矢之的。为了跟上应用环境的不断变化和信息技术的进步,国家工程技术图书馆必须继续加强知识分析,并且需要不断研究、探索和形成新的知识服务能力。

参考文献

- [1] 董颖. 知识服务机制研究[D]. 北京: 中国科学院软件研究所, 2003: 3-9.
- [2] 邱红. 基于项目的知识服务体系研究[J]. 图书馆学研究, 2012, (5): 90-93.
- [3] 戚建林. 论图书情报机构的信息服务与知识服务[J]. 河南图书馆学刊, 2003, 23(2): 37-38.
- [4] 张晓林. 走向知识服务: 寻找新世纪图书情报工作的生长点[J]. 中国图书馆学报, 2000, 26(129): 32-37.
- [5] 刊物记者. 知识服务推动图书馆转型: “2012知识服务专家论坛”纪要[J]. 图书情报工作, 2012, 56(3): 5-11.
- [6] 任树怀, 高海峰, 季颖斐. 基于图书馆2.0构建学科知识服务平台[J]. 大学图书馆学报, 2007(3): 58-62.
- [7] 王华伟, 张然, 任红超. 高校学科知识服务体系构建初探: 以武汉理工大学材料学科为例[J]. 情报理论与实践, 2012, 35(9): 1-4.
- [8] 秦晓球, 李晨晖, 麦范金. 大数据知识服务的内涵、典型特征及概念模型[J]. 情报资料工作, 2013, (18-22).
- [9] 李晨晖, 崔建明, 陈超泉. 大数据知识服务平台构建关键技术研究[J]. 情报资料工作, 2013(2): 29-34.
- [10] 万方数据. 万方数据知识服务平台[EB/OL]. [2013-08-12]. <http://www.wanfangdata.com.cn>.
- [11] 万蔚萍. 知识服务内在深层隐性的交互联动关系: 知识服务根本属性的重新界定[J]. 情报科学, 2013, 31(6): 77-82.
- [12] 靳红, 程宏. 图书馆知识服务研究综述[J]. 情报杂志, 2004(8): 8-10.
- [13] 庞景安, 马峥. 《中国科技期刊引证报告》的研制与应用[J]. 情报学报, 2001, 20(4): 495-503.
- [14] 曾建勋, 赵捷, 吴雯娜, 等. 基于引文的知识链接服务体系研究[J]. 情报理论与实践, 2009, 32(5): 1-4, 8.
- [15] 曾建勋. 中国高被引指数分析[M]. 2011年版. 北京: 科学技术文献出版社, 2011: 1-7.
- [16] 曾建勋, 丹英. 国家工程技术图书馆服务策略研究[J]. 情报理论与实践, 2009, 32(9): 56-59.
- [17] 曾建勋, 王立学. 面向知识评价的规范文档建设方法[J]. 图书情报工作, 2012, 56(10).
- [18] 李霞, 樊治平, 冯博. 知识服务的概念、特征与模式[J]. 情报科学, 2007, 25(10): 1584-1587.

作者简介

王立学, 男, 博士, 助理研究员, 研究方向: 引文分析. E-mail: wanglixue@istic.ac.cn

Research on Knowledge Service System of National Engineering and Technology Library

Wang Lixue / Institute of Scientific and Technical Information of China, Beijing, 100038

Abstract: Based on self-constructed and ordered collection of document and knowledge resources, the National Engineering and Technology Library (NETL) constructs a knowledge services system (KOS) to meet the need of researchers and research institutes. The KOS begins with knowledge resources standardization, multi-angle knowledge elements analyzing and linking, and consists of several sub-systems, such as Knowledge Linking, Scientific Research Analysis, and Institute Research Analysis. This paper mainly discusses authority files, knowledge analysis, and function of NETL KOS.

Keywords: NETL, Knowledge services, Knowledge analysis, Authority files

(收稿日期: 2013-09-09)