



项目编号: 2016030

CALIS 全国农学文献信息中心研究项目 结题报告

项目名称: 高校图书馆开展深层次信息服务之探索与实践

项目关键词: 信息服务 学科服务 文献计量 知识图谱

项目单位(盖章): _____

通信地址: 南京市龙蟠路 159 号南京林业大学图书馆 210037

项目主持人: 龚 超

联系电话: 025-85427568

电子邮件: gongch@njfu.edu.cn

提交日期: 2017.4.24

项目结题验收单

1 专家验收表（主持人所在单位组织 3-5 名专家对项目进行验收、自评。）

项目名称	高校图书馆开展深层次信息服务之探索与实践				
主持人	龚超	职务/职称	部门主任/馆员		
所在单位	(加盖公章)				
专家意见	本项目以学校和图书馆的需求为指引，把学校重点学科、优势学科和特色学科作为切入点，主动开展服务，强化和拓展情报研究的服务职能，具有重要的实践意义。 通过对 WOS、EI、CSSCI 等我校已采购数据库进行数据采集、筛选、清洗、挖掘和统计分析，逐步构建并完善我校重点学科、优势学科和特色学科以及相关院系的教师科研论文收录数据库。以科研论文收录数据库的建设和利用为突破口，利用文献计量分析方法，借助论文统计和引文分析等系统、软件、工具为学校部分学科和院系提供科技竞争力、影响力、学科现状等学科情报研究服务，提取筛选相关数据，完成了有关四个学科和 7 个院系的学术成果分析研究报告，促进了图书馆服务水平的提高，是图书馆开展深层次信息服务的一次有益探索与实践。 希望课题成员应更多的利用数据和多种工具进行数据深入挖掘分析，得到更多的有用成果为学校科研、学科建设和管理决策提供特色文献信息服务。 专家组经讨论，一致同意该项目结题。 (如需要可增加页数)				
专家签字	刘应华	张	丁	刘	陈
职务/职称	馆/教授	副馆/副教授	副馆/副研	副研	副研

题目：高校图书馆开展深层次信息服务之探索与实践

关键词：信息服务 学科服务 文献计量 知识图谱 CiteSpace

1 研究背景、目的及意义

随着网络、数字资源和信息技术迅猛发展，我们正步入知识经济时代。高校图书馆的服务正从传统的参考咨询逐渐上升到知识服务的层次，向着专业化、知识化、个性化的方向发展。图书馆的核心价值从资源主导型向信息与知识服务主导型转变。

目前我国高校图书馆所提供的知识服务仍存在很多不足，特别是深层情报与信息服务能力还比较薄弱，与图书馆知识服务的发展趋势不相适应。

当前我校正以“十三五”规划为契机，加快学科建设，提升整体实力与核心竞争力，全面推进高水平特色大学建设。以学校需求为指引，把学校重点学科、优势学科和特色学科作为切入点，主动开展服务；根据学校教改、科研、学科建设和管理决策工作的需要，提供特色文献信息服务。

利用现有文献资源和人力资源开展有价值数据的采集、梳理和统计，建立科研论文收录数据库，通过客观数据统计和资源信息挖掘，从文献计量分析角度，开展学科竞争力、科研绩效评估和决策支持服务等深层次信息服务。利用科研评价与分析工具，统计分析我校师生的科研进展态势与发展潜势，科研重点领域，科研机构与研究人员的国际排行等。通过文献计量分析的方法开展科技竞争力、影响力、学科现状、研究进展、科技前沿、发展潜势等学科情报研究服务，撰写学校和相关学科的学术成果分析报告。强化和拓展情报研究的服务职能，努力探索图书馆服务转型的工作模式。

2 研究内容及方法（思路、方法、具体内容）

主要研究思路及方法：

（1）在 WoS 数据库中以我校地址作为机构扩展地址，数据库选择 Web of Science 核心合集：引文索引，进行检索，全字段导出检索记录并把结果数据输出到 excel；利用 Excel 函数组合和 VBA 编程对输出结果数据进行提取、筛选和挖掘，得到我校 WoS 论文信息数据库。

（2）从学校人事部门获取学校各部门在职人员名单，把中文姓名转换成多种英文格式，建立我校我校在职人员部门信息表。

（3）把在职人员部门信息和论文信息进行数据关联匹配，经过数据梳理和清洗，录入科研论文收录数据库；

（4）把科研论文收录数据库同 ESI 期刊列表和 InCites 数据库数据进行关联，挖掘出论文及作者的中文相关信息，补充到科研论文收录数据库，为后续学科竞争力、科研绩效评估、ESI 潜力学科分析、决策支持服务等提供基础数据。

（5）在完善我校 WoS 论文收录数据库的基础上，参照其建设思路和方法，总结经验逐步对 EI 和 CSSCI 数据库进行数据检索、导出、筛选和数据清洗，挖掘和完善数据信息，不断更新科研论文收录数据库的基础数据。

(6) 利用科研论文收录数据库,用文献计量分析方法,借助论文统计、引文分析等系统、软件、工具为学校部分学科和院系提供科技竞争力、影响力、学科现状、作者发文、ESI 潜力学科等学科情报研究服务。

研究具体内容:

(1) 在 WoS 数据库中,选择 Web of Science 核心合集:引文索引子库下的 SCIE、SSCI、CPCI-S、CPCI-SSH 四个数据库,在机构扩展字段中以检索词:nanjing forestry univ*进行检索,检索结果全部导出,导出内容为全记录与引用的参考文献。因 WoS 数据库一次最多只能导出 500 条记录,通过多文本合并把导出的所有记录结果合并为一个 txt 文件,再输出到 excel,共得到约 3900 条记录数据;每条记录数据有 62 个数据字段,对每个字段的含义进行中文标注以方便后续分析。通过 Excel 函数组合对 AF(作者全名)、C1(作者地址)、RP(通讯作者地址)、FU(基金资助机构)和 CR(引用的参考文献)等字段数据进行汇总整理和挖掘分析。通过 Excel VBA 编程从 C1 字段查找提取每篇文献中我校作者及作者发文机构地址信息,并统计了每篇文献中我校作者个数、国内合作机构数量、国外合作机构数量等信息。利用 Excel 公式组合提取了每篇文献的参与作者数量、我校作者的数量、第一作者地址、第一作者是否为我校作者、通讯作者地址、通讯作者是否为我校作者、文献是否存在自引,基金资助机构数量等信息;同时还根据 ISSN 刊号通过 ESI 期刊列表查询出每篇文献是否被 ESI 期刊收录以及该收录期刊所属的 ESI 学科;把上面的挖掘信息合并入源表数据,得到我校 WoS 论文信息数据库,为后续进行深入文献计量分析提供多种数据。

(2) 从学校人事部门获取学校各部门在职人员名单约 2000 多人,通过 Excel 函数组合和 VBA 编程,把在职人员中文姓名转换成英文,对 Excel VBA 无法批量转换的姓和名中的生僻字约 146 个也手工转换为拼音并用公式自动构建了 10 种最常见的英文姓名格式,建立了我校在职人员部门信息表,以后新增人员信息也可以很快的更新到该信息表中,为文献的作者和院系定位提供详实的比对数据。

(3) 提取每篇文献(英文)中我校作者及作者发文机构地址数据,对作者发文机构地址数据进行除重,得到 553 条不同的英文机构地址数据,根据英文机构地址和作者信息,通过比对核查得到我校作者发文机构地址的中文信息即作者所在院系,形成我校作者机构地址中文信息表。把我校在职人员部门信息表、我校作者机构地址中文信息表同我校 WoS 论文信息数据库中的我校作者及作者发文机构地址信息进行关联匹配,经过数据梳理后,可以得到每篇文献中我校作者和机构地址的中文信息,方便了文献的定位,为以后按院系进行文献计量分析、学科现状及学术成果分析提供数据参考。

(4) 把我校 WoS 论文信息数据库分别通过国际标准期刊号(ISSN)和 WoS 入藏号同 ESI 期刊列表和 InCites 数据库数据中对应字段进行关联匹配,获取到每篇文献是否被 ESI 期刊收录以及该收录期刊所属的 ESI 学科,还获取到该文献的被引信息和期刊的影响因子等期刊信息。通过以上几步,逐步建立了我校教师科研论文收录数据库,为以后对我校 ESI 潜力学科进行数据分析和预测提供了基础数据。

(5) 在完善我校 WoS 论文收录数据库的基础上,参照其建设思路和方法,总结经验逐步对 EI 和 CSSCI 数据库进行数据检索、导出、筛选和数据清洗,挖掘有用数据信息,不断更新科研论文收录数据库的基础数据。

(6) 以科研论文收录数据库为基础,用文献计量分析方法,借助论文统计、引文分析等系统、软件、工具为学校部分学科和院系提供科技竞争力、影响力、学科现状、作者发文、ESI 潜力学科等学科情报研究服务。提取筛选相关数据,完成有关学科竞争力报告,并利用 Bibexcel、CiteSpace、HiteCite、Pajek 和 VOSviewer 等知识图谱绘制工具进行文献分析和可视化展示。

3 结论与建议

面对变革的时代，图书馆的核心价值正从资源主导型向信息与知识服务主导型转变。以学校需求为指引，把学校重点学科、优势学科和特色学科作为切入点，主动开展服务。通过对 WOS、EI、CSSCI 等我校已采购数据库进行数据采集、筛选、挖掘和统计分析，逐步构建并完善我校重点学科、优势学科和特色学科以及重点院系的教师科研论文收录数据库。利用科研论文收录数据库进行客观数据统计和资源信息挖掘，以文献计量分析角度，开展学科竞争力、科研绩效评估和决策支持服务等深层次信息服务。利用文献计量分析方法，借助论文统计、引文分析等系统、软件、工具为学校部分学科和院系提供科技竞争力、影响力、学科现状等学科情报研究服务，撰写学校和相关学科的学术成果分析研究报告。以科研论文收录数据库的建设和利用为突破口，强化和拓展情报研究的服务职能，是图书馆开展深层次信息服务和业务转型的一次有益探索与实践。

本课题现在主要采集挖掘 WOS 和 CSSCI 数据库信息，并且只是针对我校作者，以后应总结经验和方法，构建多种模板，对其他林业院校进行数据采集挖掘，也可针对某一专题进行数据采集挖掘分析，使科研论文收录数据库不断完善，使学科学术成果分析和 ESI 学科分析等学科情报分析服务的数据信息越来越丰富。此外应综合掌握和熟练使用多种知识图谱绘制工具，以更好的进行文献分析和可视化展示。

4 项目成果（发表的文章、开发的软件、取得的实践效果等）

（1）2016 年以我校教师科研论文收录数据库为基础，提取筛选相关数据，完成了我校四个重点学科和 7 个院系的学术成果分析研究报告，受到学校校领导和科技处、研究生院等科研部门的好评，促进了图书馆服务水平的提高。

（2）为研究生院提供我校工程和动植物两个 ESI 潜力学科的发文动态和作者的发文贡献度详细数据。因数据详实全面，有参考价值，得到研究生院的认可并为图书馆的学科服务小组提供专项经费支持。

（3）在 WoS 数据库中，检索并导出我校作者发文数据并输出到 Excel。通过 Excel 函数组合对 AF(作者全名)、C1(作者地址)、RP(通讯作者地址)、FU(基金资助机构)和 CR(引用的参考文献)等字段数据进行汇总整理和挖掘分析。通过 Excel VBA 编程从 C1 字段查找提取每篇文献中我校作者及作者发文机构地址信息，并统计了每篇文献中我校作者个数、国内合作机构数量、国外合作机构数量等信息。利用 Excel 公式组合提取了每篇文献的参与作者数量、我校作者的数量、第一作者地址、第一作者是否为我校作者、通讯作者地址、通讯作者是否为我校作者、文献是否存在自引，基金资助机构数量等信息；同时还根据 ISSN 刊号通过 ESI 期刊列表查询出每篇文献是否被 ESI 期刊收录以及该收录期刊所属的 ESI 学科。把上面的挖掘信息合并入源表数据，得到我校 WoS 论文信息数据库，数据记录约 3900 条。

（4）在 CSSCI 数据库中检索导出 1998 年至 2016 年以我校为来源机构的发文数据，通过 Excel VBA 编程从 txt 文本文件中批量导入 Excel，共得到约 620 条记录，每条记录有 15 个字段数据。

（5）从学校人事部门获取学校各部门在职人员名单，通过 Excel 函数组合和 VBA 编程，把在职人员中文姓名转换成英文，对 Excel VBA 无法批量转换的姓和名中的生僻字约 146 个也手工转换为拼音并用公式自动构建了 10 种最常见的英文姓名格式，建立了我校在职人员部门信息表，表中人员信息约 2000 多条。

(6) 提取每篇文献(英文)中我校作者及作者发文机构地址数据,得到我校作者机构地址中文信息表,共有记录约 550 条。

(7) 把我校 WoS 论文信息数据库分别通过国际标准期刊号(ISSN)和 WoS 入藏号同 ESI 期刊列表和 InCites 数据库数据中对应字段进行关联匹配,获取到每篇文献是否被 ESI 期刊收录以及该收录期刊所属的 ESI 学科,还获取到该文献的被引信息和期刊的影响因子等期刊信息,建立了我校教师科研论文收录数据库,数据库记录数约 3900 条。

5 参考文献

- [1]周健,仇玉芹,宋振世,等.高校图书馆开展决策支持服务之探索与实践_以华东师范大学图书馆为例 [J]. 上海高校图书情报工作研究, 2015, v.25;No.99(3): 10-14.
- [2]陆婷婷. 高校图书馆学科化知识服务模式研究 [J]. 兰台世界, 2014
- [3]钱鹏. 高校科学数据管理研究 [D]. 南京大学, 2012.
- [4]孙翌,郭晶. 基于博客的高校图书馆学科化知识服务平台实证研究 [J]. 图书与情报, 2009, No.129(5): 110-113, 125.
- [5]吴卫华,李亚君,秦荣环,等. 基于文献计量的高校图书馆学科化知识服务模式探析 [J]. 图书馆杂志, 2011, v.30;No.248(12): 53-54, 75.
- [6]张子石,吴涛,金义富. 基于CiteSpace的网络学习知识图谱分析[J]. 中国电化教育, 2015, No.343(8): 83-90.
- [7]马晓佳.网络引文分析与传统引文分析的比较 [D]. 南京大学, 2011.
- [8]顾东蕾,武莹,邱家学,等. 基于WOS的学科发展预测研究 [J]. 现代情报, 2014, v.34;No.276(6): 34-42.
- [9]王攀. 高校教师科研评价研究 [D]. 武汉理工大学, 2006.
- [10]王丽萍. 基于服务科学视角下的学科化知识服务系统机制研究 [J]. 图书馆学研究, 2014, No.347(24): 60-63.
- [11]周群.文献分析软件bibexcel介绍 中国农业大学图书馆[J].
- [12]王岩,刘雅娟. 运用引文分析进行论文评价的方法初探_引文分析中期刊影响因子的运用 [J]. 科研管理, 2001
- [13]冯花朴. InCites数据库科研绩效评估的有效工具[J]. 现代情报, 2014, v.34;No.271(1): 149-156.
- [14]徐璟,郭晶. 高校图书馆学科化知识服务模式探究 [J]. 图书情报工作, 2010
- [15]陈锦波. 高校图书馆学科化知识服务能力再造探究 [J]. 大学图书情报学刊, 2014, v.32;No.143(3): 5-9.
- [16]杨思洛,韩瑞珍. 国外知识图谱绘制的方法与工具分析[J]. 图书情报知识, 2012, No.150(6): 103-111.
- [17]张群,张逸新,吴信岚. 面向研究生信息素质教育的学科化知识服务 [J]. 图书情报工作, 2010, v.54;No.418(21): 74-77, 151.
- [18]徐恺英,刘佳,班孝林. 高校图书馆学科化知识服务模式研究 [J]. 图书情报工作, 2007
- [19]秦萍,李雪琛,梅秀秀. 高校潜力学科发展预测研究 [J]. 情报杂志, 2015
- [20]何丹青. 高校图书馆学科化知识服务发展策略研究 [J]. 图书馆工作与研究, 2010, No.167(1): 15-17.
- [21]王芳,庞德盛,杨错. 高校图书馆开展文献计量服务的探索与思考 [J]. 河南图书馆学刊, 2015, v.35;No.179(3): 115-117.