

我国数字图书馆研究论文 (2005-2009) 的统计分析 ——技术体系与新技术研究*

□ 王建冬 张鹏 王继民 / 北京大学信息管理系 北京 100871

摘要: 基于2005-2009年国内数字图书馆技术体系与新技术研究领域相关论文, 对数字图书馆服务领域研究近5年来的总体情况、核心期刊发文量、热门关键词分布情况以及核心概念网络的分布情况进行统计分析。同时, 采用社会网络分析和计量分析的方法, 定义了领域关键词的专属度指标, 并计算了国内数字图书馆技术体系与新技术研究领域的向心度指标。

关键词: 数字图书馆, 技术体系, 新技术, 统计分析

DOI: 10.3772/j.issn.1673-2286.2010.03.011

1 前言

数字图书馆技术研究是数字图书馆研究的核心领域。美国国家科学基金会(NSF)在开始进行数字图书馆研究计划时, 就确立了三个广泛的关键技术领域: 其一是有关获取、分类、组织信息和知识的表达; 其二是查询、浏览、过滤、总结和可视化; 其三是网络协议和标准、网络信息的发布和知识产权。叶峰等^[1]认为, 对于图书馆和信息中心等部门来说, 在实施数字图书馆项目时在技术方面所要考虑的主要问题是采用合理的信息系统体系结构、选取有效的软件开发方法、进行信息的搜集、加工、组织来向用户提供服务。孙一钢^[2]则指出, 数字图书馆的技术体系结构是建设数字图书馆系统的基础, 是数字图书馆在网络和计算机技术上的具体实现, 是未来信息社会处理、存储和应用数字化信息的基本构架。它的发展目标是符合下一代互联网的发展趋势, 成为一个高度开放、方便可用的体系结构。吴新年^[3]认为, 数字图书馆技术应当包括数字图书馆管理技术、数字化信息生成技术、数字化信息检索技术、数字化信息传播技术、数字化信息安全防护技术及数字图书馆系统开发

技术。

经过十几年发展, 目前数字图书馆技术体系与新技术研究领域不但是数字图书馆研究的核心领域之一, 而且其概念最多, 理论体系最丰富。本文拟对近年来国内数字图书馆技术体系与新技术研究领域的总体发展情况、核心期刊分布情况、领域分布情况等进行统计分析。

2 数据检索方法

本文收集了近五年国内与数字图书馆研究相关的五类专业(包括编辑出版、档案学、图书馆学情报学、新闻传播学、计算机技术、综合性科学技术)的核心期刊近20年的所有文章信息(包括题目、著者、关键词、单位、摘要、基金信息、页码等)共440, 594条。从其中检索标题中包含“数字图书馆”+“技术”, 并且发表年限为2005-2009年的论文, 得到113条数据。从热门关键词列表中手工筛选出以下与管理体系统相关的检索词: 网络技术、数字图书馆技术、信息技术、图书馆自动化、元数据、网络、图书馆软件、技术、存储技术、Web技术、智能信息推拉技术、知识

* 本文受北京大学前瞻性研究课题“基于网络使用的中文科技期刊及论文评价机制和方法研究”资助。

聚类、信息检索、网络附加存储、推荐系统、推荐技术、检索技术、技术支撑体系、技术体系、技术化、互操作、反馈技术、存储系统、存储区域网络、本体技术、P2P技术，并进行扩展检索，得到414条数据。本文将将其作为国内数字图书馆技术体系与新技术研究领域的主要论文。

3 总体情况

2005-2008年4年间，该领域平均发文量为131.2篇^①，占同期数字图书馆领域整体论文的29%。我们进一步考察了2005-2009年5年间该领域研究论文占整个数字图书馆领域研究文章的比例的变化情况，如图1所示（图中纵坐标表示百分比数量）：

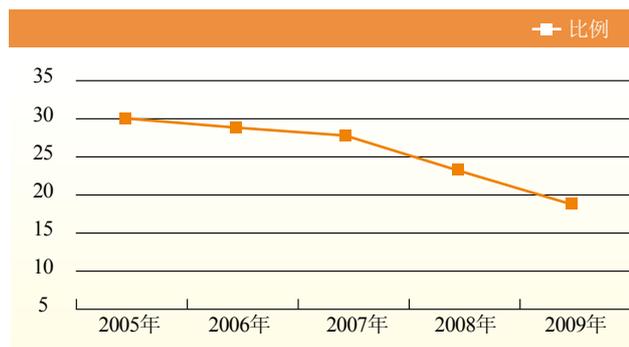


图1 2005—2009年国内该领域研究论文所占比例

从图1可以看出，数字图书馆技术体系与新技术研究占整个数字图书馆领域研究的比重呈现较明显的下降趋势，说明总体来看，国内数字图书馆研究者的研究兴趣开始从集中关注数字图书馆技术研究逐渐扩散到其他领域。

4 核心期刊发文量

该领域发文量最多的20种核心期刊与前文中曾提到的整个数字图书馆领域的20种核心期刊一致。20种核心期刊总发文量为505篇，占总数的92%。为进一步考察20种核心期刊在数字图书馆技术体系与新技术研究领域的偏好程度，我们计算每种期刊在数字图书馆技术体系与新技术研究领域的发文数量占其在整个数字图书馆领域的发文量的比重，将其作为考察期刊在

表1 20种期刊在本领域内的发文量和比重

期刊名	该领域 发文量	总发文量	比例
兰台世界	15	41	36.6%
中国图书馆学报	27	81	33.3%
图书情报知识	9	28	32.1%
现代图书情报技术	100	320	31.3%
情报杂志	46	149	30.9%
图书馆杂志	21	70	30.0%
情报学报	8	27	29.6%
图书与情报	11	38	28.9%
图书馆学研究	40	140	28.6%
情报科学	29	102	28.4%
国家图书馆学刊	12	43	27.9%
情报理论与实践	15	54	27.8%
图书情报工作	53	198	26.8%
大学图书馆学报	14	62	22.6%
图书馆工作与研究	20	89	22.5%
图书馆论坛	30	137	21.9%
图书馆理论与实践	27	124	21.8%
图书馆建设	9	61	14.8%
情报资料工作	12	82	14.6%
图书馆	7	49	14.3%

该领域的偏好程度的指标，计算结果如表1所示：

可以看出，20种核心期刊对数字图书馆技术体系与新技术研究领域的关注程度呈现较明显的不同，对该领域关注最多的几种期刊包括《兰台世界》（36.6%）、《中国图书馆学报》（33.3%）、《图书情报知识》（32.1%）、《现代图书情报技术》（31.3%）、《情报杂志》（30.9%）、《图书馆杂志》（30.0%），关注较少的期刊则包括《图书馆建设》（14.8%）、《情报资料工作》（14.6%）、《图书馆》（14.3%）。

5 热门关键词分布情况

表2显示了数字图书馆技术体系与新技术研究领域

^① 2009年数据尚不完整，因此在计算时并未考虑在内。

设关键词 K_i 在数图领域出现总词频为 f_i ，在数字图书馆技术体系与新技术研究领域出现词频为 f_i 。所有关键词在数图其他领域出现总词频的平均值 $F = (\sum_{i=1}^n (F_i - f_i)) / n$ ，在数字图书馆技术体系与新技术研究领域出现词频的平均值 $f = \sum_{i=1}^n f_i / n$ 。

则关键词 K_i 的专属度 $C_i = (f_i \times F) / ((F_i - f_i) \times f)$ 。

当 $C_i > 1$ 时，则可以认为该关键词“专属”于数字图书馆技术体系与新技术研究领域。更进一步，还可以通过计算专属度 >1 的关键词在分支领域所有关键词

中所占的比例考察该领域的向心性。

7 结语

本文通过对国内数字图书馆技术体系与新技术研究领域的相关论文的计量分析，从整体上考察了国内数字图书馆技术体系与新技术研究领域的发现现状。希望本文的研究能够为领域内研究者的进一步探讨提供有益借鉴。

表3 该领域热门关键词专属度分析

关键词	数字图书馆技术体系与新技术研究领域出现词频	数图领域出现词频	其他领域出现词频	专属度	关键词	数字图书馆技术体系与新技术研究领域出现词频	数图领域出现词频	其他领域出现词频	专属度
数字图书馆	298	1148	850	0.347136	OAI-PMH	13	15	2	6.435987
元数据	48	48	0	NA	知识组织	13	22	9	1.430219
信息检索	23	23	0	NA	XML	13	22	9	1.430219
网络	20	20	0	NA	信息网络传播权	12	12	0	NA
图书馆	20	108	88	0.225035	网络环境	12	12	0	NA
互操作	19	19	0	NA	网络存储	12	12	0	NA
数字图书馆技术	16	16	0	NA	SAN	12	13	1	11.88182
网格技术	15	15	0	NA	个性化服务	12	52	40	0.297046
推荐系统	14	14	0	NA	信息服务	12	54	42	0.282901

参考文献

- [1] 叶峰,刘小兵. 数字图书馆技术:研究现状和建议[J]. 现代图书情报技术,1999(3).
- [2] 孙一钢. 数字图书馆的技术体系结构[J]. 现代图书情报技术,2001(5).
- [3] 吴新年. 数字图书馆技术之进展[J]. 图书与情报,2000(2).

作者简介

王建冬 (1982-), 北京大学信息管理系博士研究生, 研究方向: 信息服务与情报分析。通讯地址: 北京大学信息管理系 100871。E-mail: zs.wagner@yahoo.com.cn

王继民 (1966-), 北京大学信息管理系副教授, 博士, 研究方向: 搜索引擎与Web挖掘。通讯地址: 同上

张鹏 (1984-), 北京大学信息管理系硕士研究生, 研究方向: 信息检索与分析。通讯地址: 北京大学信息管理系 100871。E-mail: 同上

A Statistical Analysis on China Digital Library Research Papers (2005-2009)--- Technology System and New Technologies

Wang Jiandong, Wang Jimin, Zhang Peng / Department of Information Management, Peking University, Beijing, 100871

Abstract: Based on Technology System and New Technologies field literature of digital library research from Wanfang Data between 2005 and 2009, we carry out a statistical analysis of digital library Technology System and New Technologies' 5-year study, which includes the overall situation of the study, the related articles of core journals, the distribution of popular keywords, as well as the core concept network. At the same time, using social network analysis and quantitative analysis methods, we define the exclusive degree of keyword and calculate the centripetal degree of the digital library technology system and new technologies field.

Keywords: Digital library, Technology system, New technology, Bibliometric

(收稿日期: 2010-01-14)