

# 基于自述研究专长的研究领域 识别与特征差异分析

——以国内图书情报与档案管理专业教师为样本

胡志伟 裴雷

(南京大学信息管理学院, 南京 210023)

**摘要:** [目的/意义] 研究国内图书情报与档案管理学科教师群体的专长分布与自述用词特征, 测算其特征差异与跨学科程度。[方法/过程] 以中国 52 所图书情报与档案管理教育机构的 1016 名教师的个人简历为数据来源, 从其自述研究专长中提取描述词, 采用 LDA 主题聚类、词频分析、词类隶属度等方法识别教师专长的领域分布特征与学科交叉特征。[结果/结论] 国内图书情报与档案管理学科教师的研究专长以信息资源管理生命周期为核心; 信息技术在当前研究中的渗透力度较高, 具有广泛的应用场景; 近五年内新入职教师在研究专长用词上呈现出跨学科与行业导向的显著特性。学科的未来发展, 需要进一步巩固核心知识体系、推动技术赋能研究、提高跨学科能力。

**关键词:** 专长自述 特征差异 LDA 模型 词频分析

**分类号:** G250.2

**DOI:** 10.31193/SSAP.J.ISSN.2096-6695.2020.04.03

## 1 研究背景

专长作为专家在某一领域所拥有的特殊技能, 通常是经过长期的教育培训和实践而获得的丰富经验的体现<sup>[1]</sup>。专长识别不仅能够快速完成学者画像、同行识别、专家遴选和知识推送<sup>[2-5]</sup>, 而且能够有效刻画学者成长、评估领域和学科知识的融合、增长与流变, 从而预测领域前沿、研判学科发展规律, 是学术服务和学科建设中的重要工作。目前, 专长识别领域主要有基于本体或

[作者简介] 胡志伟 (ORCID: 0000-0001-6244-3569), 男, 硕士研究生, 研究方向为图书情报教育、科学计量, Email: mg1914013@smail.nju.edu.cn; 裴雷 (ORCID: 0000-0003-4754-4112), 男, 教授, 研究方向为信息政策、图书情报教育、信息行为, Email: plei@nju.edu.cn。

词表的内容分析<sup>[6-9]</sup>、基于文献特征的聚类算法或主题抽取<sup>[10-12]</sup>、基于共词或社会网络的社区发现<sup>[13-15]</sup>等方法; 而研究资料包括专家成果信息(论文、著作、项目或专利等)<sup>[7-11]</sup>、专家主页或简历等自述资料<sup>[6]</sup>、专家社交网络信息(邮件、社交媒体等)<sup>[14-15]</sup>等。

在图书情报与档案管理学科领域, 在“学者-文献”研究中主要有张晓娟等利用 PLSA 识别图书情报领域的专家专长, 通过“文档-主题”和“主题-关键词”两个矩阵, 确定专家的研究主题<sup>[9]</sup>; 陆伟等利用专家词表进行图书情报与档案管理专家学者检索系统的设计<sup>[7]</sup>。在“教师-简历”分析方法中, 肖希明和杨蕾对国内 37 所图书情报教育机构师资队伍专长调研与归纳分析<sup>[6]</sup>、史雅莉等对国外 41 所 iSchools 院校的师资队伍专长分析<sup>[16]</sup>、赵跃和梅梦娜对 100 名中国高校档案学者的归纳分类<sup>[17]</sup>等。此外, Holmberg 等利用 1168 名教职工在个人简历中的自述研究专长文本, 对 30 所 iSchools 院校教师的研究主题进行了共词网络分析<sup>[18]</sup>。而且, 随着越来越多的高校公开专家个人主页, 展示高校师资、教学及科研水平<sup>[19]</sup>, 在教师个人主页中教师主动公开研究兴趣与研究成果信息越来越多, 成为专长识别的一种重要的途径或资料来源。

从已有简历分析来看, 主要停留在专长标注和词频统计等分析方法, 对词间关系以及词聚类特征的分析还不多见。鉴于此, 本文基于国内图书情报与档案管理教师的个人简历信息, 综合主题聚类与词语定量分析方法, 研究国内图书情报与档案管理学科教师群体研究专长的领域分布特点, 以及在专长自述用词中的特征差异。

## 2 研究设计与数据获取

### 2.1 研究方法与数据来源

本文主要采用词频统计和基于 LDA (Latent Dirichlet Allocation) 的主题聚类方法。通过系统采集本学科教职工的简历中的专长信息, 生成标准的词向量模型; 构建合适的 LDA 主题模型并获得专长聚类, 识别其领域特点; 在词频统计基础上计算词类覆盖程度。整体研究思路如图 1 所示。

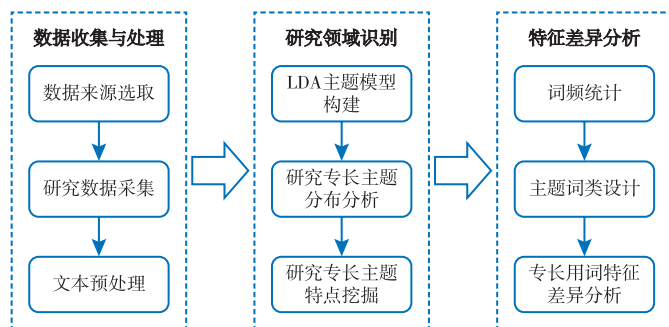


图 1 本文研究思路

在数据获取环节, 基于世界图书情报与档案管理教育机构名录<sup>[20]</sup>与中国教育在线网站中的图书情报与档案管理专业招生院校列表<sup>[21]</sup>, 选取 52 所图书情报与档案管理教育机构作为研究

数据来源,如表1所示。在此基础上,分别通过这些机构的官方网站采集教师的个人资料,涉及教师姓名、职称、所属机构、教育与工作经历、自述研究兴趣等方面。针对教师个人主页信息缺失或更新不及时的情况,以中国知网为入口检索该教师最近发表的2篇期刊论文,采集其于论文“作者简介”部分中描述的个人研究方向,作为对原始数据的补充与完善。数据采集时间为2019年11月25日至12月4日,最终获得1016名教师信息,共提及研究专长4399次。

表1 样本数据来源列表

安徽财经大学	安徽大学	北京大学	北京航空航天大学
北京理工大学	北京师范大学	东北师范大学	东南大学
福建师范大学	福州大学	广西民族大学	河北大学
河海大学	华东师范大学	华南师范大学	华中科技大学同济医学院
华中师范大学	吉林大学	济南大学	江苏大学
兰州大学	辽宁师范大学	南京大学	南京农业大学
南开大学	青岛科技大学	曲阜师范大学	山东大学
山东科技大学	山东理工大学	山西大学	上海大学
四川大学	苏州大学	天津大学	天津工业大学
天津师范大学	武汉大学	西安电子科技大学	西北大学
西南大学	西南科技大学	湘潭大学	云南大学
浙江大学	郑州大学	中国科学技术信息研究所	中国科学院文献情报中心
中国农业大学	中国人民大学	中山大学	重庆大学

## 2.2 数据预处理

分词处理与矩阵构建主要借助 Rwordseg 包和 tm 包完成。因样本中存在中英文多语种来源,首先对英文专长描述进行了英译中处理,但不对中文表述中常用英文词汇(如“Web”、“IT”等)进行翻译;其次,采用哈工大停用词表和四川大学机器智能实验室停用词库构建分词停用词表,删除无意义的词语。最终,得到 1016\*10516 的 DTM (Document Term Matrix) 矩阵。

# 3 研究发现

## 3.1 研究领域识别

借助 R 语言中的 lda 包、topicmodels 包和 LDAvis 包,采用 Gibbs 采样算法对国内图书情报与档案管理专业教师研究专长的主题分布进行求解,并基于困惑度(perplexity)指标随主题数变化的情况将最优主题数  $k$  设定为 15、文档迭代次数 num.iterations 设定为 2000,而对其余参数取默认值。针对输出结果中每个主题下所包含的显著度(saliency)<sup>[22]</sup>最高的 30 个词语,对主题进行标识分析。根据“文档-主题”概率分布中教师聚焦主题的概率,计算主题综合强度,在一定程度上代表当前教师研究专长中的重要程度。由表 2 可知,当前时间窗口上的教师重点研究主题包

括“信息资源管理”、“信息服务与知识服务”、“电子商务与战略管理”、“信息技术与数字图书馆”和“数字档案管理”(T<sub>1</sub>~T<sub>5</sub>)等领域。

表2 国内图书情报与档案管理专业教师研究专长主题分布

编号	主题标识	高显著度词语	强度
T <sub>1</sub>	信息资源管理	信息 管理 资源 图书馆 服务 检索 组织	0.085
T <sub>2</sub>	信息服务与知识服务	信息 服务 管理 资源 知识 咨询 用户	0.083
T <sub>3</sub>	电子商务与战略管理	信息 电子 商务 管理 技术 企业 战略	0.080
T <sub>4</sub>	信息技术与数字图书馆	数字 信息 知识 图书馆 技术 组织 社会	0.079
T <sub>5</sub>	数字档案管理	档案 管理 理论 文件 信息 电子 数字	0.074
T <sub>6</sub>	科技情报与信息政策	科技 政策 情报 管理 战略 分析 发展	0.071
T <sub>7</sub>	企业竞争情报	信息 情报 管理 竞争 理论 分析 企业	0.067
T <sub>8</sub>	信息分析与信息行为	信息 分析 网络 可视化 文本 用户 交互	0.065
T <sub>9</sub>	电子信息资源与电子政务	信息 资源 电子 政府 数据 政务 文件	0.064
T <sub>10</sub>	数据挖掘与商务智能	数据 分析 智能 挖掘 商务 金融 机器	0.064
T <sub>11</sub>	信息系统与供应链管理	管理 信息 系统 供应链 决策 经济 物流	0.062
T <sub>12</sub>	图书档案史与文献保护	图书馆 历史 文献 档案 保护 中国 阅读	0.056
T <sub>13</sub>	信息计量与科学评价	信息 科学 计量 评价 分析 网络 学术	0.054
T <sub>14</sub>	项目管理与企业信息化	管理 系统 资源 项目 信息化 企业 数据库	0.054
T <sub>15</sub>	数字出版与文化产业	出版 服务 文化 数字 产业 营销 图书	0.042

具体而言, 研究发现:

(1) “信息资源管理”(T<sub>1</sub>)研究主题的关注度最高。信息资源管理涉及信息资源和信息活动进行组织、规划、协调和控制, 以实现组织信息资源的合理开发和有效利用的过程<sup>[23]</sup>, 是图书情报与档案管理学科核心知识的共同点<sup>[24]</sup>。总体而言, 中国图书情报与档案管理学科在信息资源管理理论和实践领域占据了主导地位<sup>[25]</sup>, 而此前学界关于“信息资源管理”学科更名的讨论<sup>[26]</sup>也具有合理性。

(2) “信息服务与知识服务”(T<sub>2</sub>)主题强度0.083, 位列第二。专家学者们发现和解决问题、服务社会的能力不断提升, 在学术研究与社会价值创造方面取得显著成效的基础上, 也促进了学科服务职能的进一步深化<sup>[27]</sup>。随着知识学习、利用、创造与共享的社会取向和运行机制的逐步形成, 信息服务也出现了向知识服务转型的趋势<sup>[28]</sup>。在此过程中, 除传统图书馆服务、情报服务和档案信息服务之外, 当前国内教师也较多关注公共文化服务、开放学术交流服务以及健康信息服务等研究领域, 呈现了更深层次、更广口径的特点。

(3) 技术创新是图书情报与档案管理研究的重要领域。聚类结果显示, “信息技术与数字图书馆”(T<sub>4</sub>)、“数字档案管理”(T<sub>5</sub>)等技术融合的研究主题具有很高的关注度。上述趋势与世界范围内的数据与技术驱动的学科发展相符, 体现了技术工具和技术方法在图书情报与档案管理领域的重要作用。美国图书馆与情报学教育协会 (Association for Library and Information Science

Education, ALISE) 的最新研究分类将数据管理、数据科学、人机交互与设计以及信息技术也设定为一级类别<sup>[29]</sup>, 并在二级分类中涵盖了大数据、人工智能、机器学习等热门技术, 体现了当前图书情报与档案管理领域与前沿技术的快速采纳与融合发展。

(4) 跨学科研究主题介入程度高。聚类结果也包括“电子商务与战略管理”(T<sub>3</sub>)、“项目管理与企业信息化”(T<sub>14</sub>)、“数字出版与文化产业”(T<sub>15</sub>)等具有鲜明跨学科特色的研究主题。体现了在学科实践中, 研究学者广泛采纳跨学科方法, 参与跨学科应用和实践。

## 3.2 特征差异分析

### 3.2.1 主题词类设计

研究专长用词方面的特点与趋势能够直观地反映研究内容的集聚与分布特征。对专长词频的统计发现, 669个词语累计出现了10516次。如表3所示, 将词语出现频次划分为5个区间, 依照帕累托法词频 $x>5$ 的141个词语(约占总体的前21%)为国内图书情报与档案管理教师简历中研究专长文本的重点词, 主要围绕信息链和信息资源管理生命全周期展开。

表3 国内图书情报与档案管理教师研究专长词频分布

词语频数(x)	1	2	3	4	5	$x>5$
词语个数(n)	297	96	60	46	29	141

基于HowNet语义词典<sup>[30]</sup>分别标记141个重点词的语义特征(包括义项与义原), 通过对义原逐级回溯(具体过程示例如图2所示), 提取“信息技术”、“经济管理”和“行业领域”共3个主要主题词类。其中, 信息技术词类包括“技术”、“数据”、“数字”等30个指向以计算机技术为代表的现代信息技术的词语; 经济管理词类包括“商务”、“经济”、“企业”等12个具有鲜明经管特色的词语; 行业领域词类是指面向传统图书情报与档案管理应用领域和商业之外的应用场景的词语集合, 具体涉及政府信息资源管理、健康信息学、农业图书情报等研究领域, 包括“政府”、“政务”、“媒体”等9个词语。

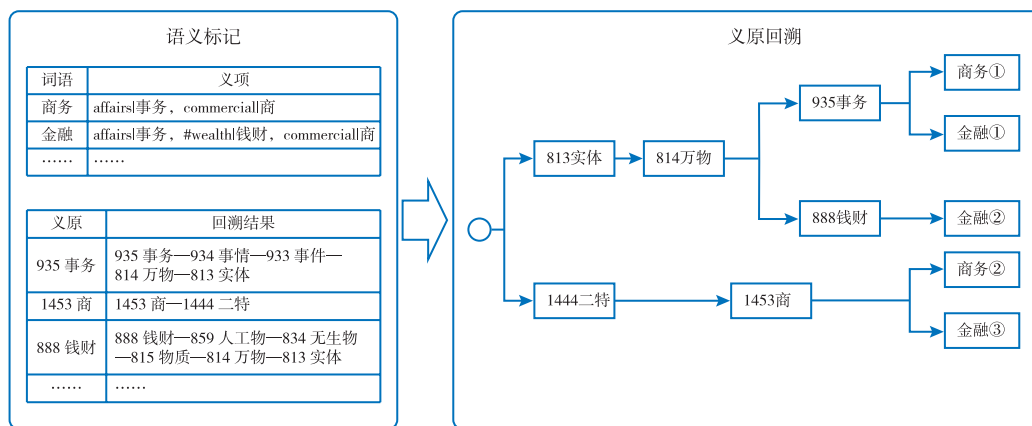


图2 研究专长重点词义原分析过程(部分)



**表 4 国内图书情报与档案管理教师研究专长主题词类设计**

词类名	所含重点词
信息技术	技术、数据、数字、系统、电子、网络、挖掘、智能、工程、开发、语义、互联网、数据库、计算、机器、Web、计算机、智慧、人机、算法、虚拟、图谱、集成、模式、多媒体、软件、本体、个性化、文本、保密
经济管理	商务、经济、企业、供应链、金融、物流、营销、市场、商业、客户、贸易、产品
行业领域	政府、政务、媒体、健康、医学、卫生、医疗、农业、农村

### 3.2.2 专长用词特征差异分析

本文引入覆盖率 (coverage rate) 作为测度不同主题词类覆盖程度的统计指标, 并对其作如下定义:  $CR_{ij} = n_{ij}/N_j \times 100\%$ , 其中  $CR_{ij}$  表示词类  $C_i$  在群组  $G_j$  中的覆盖率,  $n_{ij}$  是群组  $G_j$  中提及词类  $C_i$  中所含词语的教师数量,  $N_j$  是群组  $G_j$  的教师总量。为进一步分析新入职教师自述研究专长用词特征, 构建两个教师群组  $G_1$  和  $G_2$ , 其中  $G_1$  表示 2015 年及以后入职的教师 ( $N_1 = 279$ ),  $G_2$  表示 2015 年以前入职的教师 ( $N_2 = 737$ )。

如表 5 所示, 经济管理、行业领域用词在群组  $G_1$  和  $G_2$  中的覆盖率具有显著差异。2015 年以来入职的教师同 2015 年前入职的教师相比, 在研究专长描述中更多地使用了经济管理和行业领域类词语, 且明确提及经济管理类研究方向的新入职教师超过总体半数 ( $P_{21} = 50.2\%$ )。上述两个词类在两组教师群体之间的差异均是显著的 ( $p < 0.05$ ), 体现了新入职教师在研究专长上的跨学科应用与行业导向的趋势更明显。从跨学科的范围和广度看, 52 所图书情报与档案管理教育样本机构中, 有 27 所机构具有经济管理、公共管理或医学等专业背景, 客观上需要跨学科或者复合型学者的加入, 也加剧了图书情报与档案管理机构新入职教师的跨学科背景比例。

**表 5 各主题词类的覆盖率及其组间差异检验**

CR <sub>ij</sub> 词类 \ 组别	G <sub>1</sub>	G <sub>2</sub>	$\chi^2$	p	假设检验结果
信息技术	56.2%	58.6%	0.37	0.55	Rejected
经济管理	50.2%	22.0%	75.69	$2.20 \times 10^{-16}$	Accepted
行业领域	17.6%	12.1%	4.73	0.03	Accepted

同时, 信息技术类词语在国内教师研究专长文本中出现的频率整体处于较高水平 ( $P_{11} = 56.2\%$ ,  $P_{12} = 58.6\%$ ), 并且在两组教师群体之间不存在显著性差异 ( $p = 0.55$ ), 说明了这两组教师群体对待信息技术的态度基本一致。

此外, 在专业表述的术语使用上, 青年学者更愿意选择具有跨学科或没有鲜明图书情报与档案管理学科背景的术语作为研究方向, 具有一定的模糊规避 (ambiguity aversion) 现象。模糊规避是指在相同情况下, 相较于从主观上判断具有相同精确概率的事件, 决策者更加偏好有模糊概率的事件<sup>[31]</sup>。将教师简历描述视为一种自我行为决策, 采用中性词或涵义跨领域的术语表述研究方向的倾向, 在青年教师群体中更加普遍。据样本数据统计, 2015 年以来入职的教师直接

使用本学科或专业名称作为自述研究专长的教师占比仅为 11.8%，而 2015 年以前入职的教师有 25.2%，二者差异显著。

## 4 研究结论与展望

本文选取国内 52 家图书情报与档案管理教育机构中的 1016 名教师作为样本，针对教师简历文本利用 LDA 主题模型和词频统计方法，对其研究专长进行了主题聚类与描述用词分析。研究发现：第一，国内图书情报与档案管理专业教师的研究方向具有一条相对清晰的研究主线，15 个主要研究领域均以信息生命周期及其管理活动为主线，围绕信息利用、分析与服务展开；第二，信息技术仍然是图书情报与档案管理研究的重要挑战和创新领域，尤其是应用于具体应用场景的理论创新，形成了本学科跨学科融合创新的特点；第三，新入职教师在研究专长用词上展现出了更多的跨学科和模糊规避倾向，且更倾向于细粒度、具体精细的专长表达。

针对专长自述中的“聚与散”、“认同与规避”，图书情报与档案管理学科在师资队伍建设中，可能需要进一步加强学科认同感和图书情报的学科话语体系张力。相较于学术论文明确的关键词选取流程，学者对自身研究专长的描述过程具有一定的主观性，出现了自述专长与研究专长在方向或粒度上的不一致现象，甚至一定程度的模糊规避现象，尤其在青年学者的专长表达中，跨学科术语采用更为广泛。上述现象反映出，目前青年学者在学科认同上存在一定的模糊空间。当前，在图书情报与档案管理学科共同体的核心知识业已明朗、研究主体围绕信息资源管理展开的背景下，未来仍需进一步凝练核心知识与理论内核，提升学科吸引力和学科术语的传播力。

此外，本文对教师简历的数据采集主要来自机构网站，可能存在教师信息更新不及时、数据老旧的情况，影响分析结果。今后，可进一步通过教师简历、论文注释及其学术成果等多来源资料进行综合分析，以更好地刻画教师专长特征。

### 【参考文献】

- [1] 杨瑜. 基于专长视角的学术信息检索认知能力研究 [D]. 福州: 福建师范大学, 2016.
- [2] 袁莎, 唐杰, 顾晓韬. 开放互联网中的学者画像技术综述 [J]. 计算机研究与发展, 2018, 55(9): 1903-1919.
- [3] 靳健, 杨海慈, 李凝, 等. 基于主题契合度的专家推荐模型研究 [J]. 数字图书馆论坛, 2017(4): 47-55.
- [4] 冷昕阳. 基于专家知识的评审专家推荐算法研究 [J]. 环球人文地理, 2016(22): 314-315.
- [5] 李江, 李东, 冯培桦, 等. 基于专长吻合度、学术影响力与社会关联值的专家推荐模型研究 [J]. 情报学报, 2017(4): 338-345.
- [6] 肖希明, 杨蕾. 我国 LIS 教育机构师资队伍建设的调查分析 [J]. 图书情报知识, 2014(1): 4-10.
- [7] 陆伟, 刘杰, 秦喜艳. 基于专长词表的图情领域专家检索与评价 [J]. 中国图书馆学报, 2010, 36(2): 70-76.
- [8] 胡月红, 刘萍. 基于本体概念的专长表示研究 [J]. 图书情报工作, 2012, 56(4): 17-21, 40.
- [9] 张晓娟, 陆伟, 程齐凯. PLSA 在图情领域专家专长识别中的应用 [J]. 现代图书情报技术, 2012(2): 76-81.
- [10] 刘晓豫, 朱东华, 汪雪锋, 黄颖. 多专长专家识别方法研究——以大数据领域为例 [J]. 图书情报工作, 2018, 62(3): 55-63.
- [11] 陈翀, 李楠, 梁冰, 等. 基于成果特征的学者学术专长识别方法 [J]. 图书情报工作, 2019, 63(20): 96-103.

胡志伟, 裴雷. 基于自述研究专长的研究领域识别与特征差异分析——以国内图书情报与档案管理专业教师为样本 [J]. 文献与数据学报, 2020, 2(4): 040-048.

- [12] 巩军, 刘鲁. 基于个人知识地图的专家推荐 [J]. 管理学报, 2011, 8(9): 1365.
- [13] 刘萍, 周梦欢. 基于共词网络的专家专长挖掘 [J]. 情报科学, 2012, 30(12): 1815-1819.
- [14] SONG X, TSENG B L, LIN C Y, et al. ExpertiseNet: relational and evolutionary expert modeling [C]// ARDISSONO L, BRNA P, MITROVIC A. Proceedings of User Modeling 2005. Edinburgh: Springer. 2005: 99-108.
- [15] YANG K W, HUH S Y. Automatic expert identification using a text categorization technique in knowledge management systems [J]. Expert Systems with Applications, 2008, 34(2): 1445-1455.
- [16] 史雅莉, 司莉, 姚瑞妃, 等. 国外 iSchools 图情学科教师队伍结构与研究领域调研及启示 [J]. 图书情报知识, 2018(5): 33-40.
- [17] 赵跃, 梅梦娜. 我国档案学研究的坚守、转向与进路——基于档案学者研究方向变化的审视 [J]. 档案学通讯, 2019(3): 4-11.
- [18] HOLMBERG K, TSOU A, SUGIMOTO C R. The conceptual landscape of iSchools: examining current research interests of faculty members [J/OL]. Information Research, 2013, 18(3): C32 [2020-12-8]. <http://InformationR.net/ir/18-3/colis/paperC32.html>.
- [19] 连镏. 高校个人主页系统助力教师形象传播 [J]. 中国新通信, 2018, 20(8): 210.
- [20] Schniederjürgen A. World Guide to Library, Archive and Information Science Education [M]. Berlin: Walter de Gruyter, 2007: 66-120.
- [21] 中国教育在线. 全国研究生招生学科、专业目录索引 [EB/OL]. [2020-09-17]. [http://www.eol.cn/e\\_ky/zl/common/zym/](http://www.eol.cn/e_ky/zl/common/zym/).
- [22] CHUANG J, MANNING C D, HEER J. Termite: visualization techniques for assessing textual topic models [C]. Proceedings of the International Working Conference on Advanced Visual Interfaces. ACM, 2012: 74-77.
- [23] 王贻志, 丁波涛. 信息资源管理: 学科发展与应用前景 [M]. 上海: 上海交通大学出版社, 2012: 9.
- [24] 闫慧. 青年学者论图情档一级学科核心知识及发展方向——2019年图书情报与档案管理青年学者沙龙会议述评 [J]. 中国图书馆学报, 2019, 45(1): 121-127.
- [25] 马费成, 宋恩梅, 张勤. IRM-KM 范式与情报学发展研究 [M]. 武汉: 武汉大学出版社, 2008: 37-38.
- [26] 赖茂生. 信息管理学科发展如何破解瓶颈悖论? [J]. 大学图书情报学刊, 2019, 37(5): 3-7.
- [27] 冯惠玲. 学科探路时代——从未知中探索未来 [J]. 信息资源管理学报, 2020, 10(3): 4-10.
- [28] 杨海燕. 知识经济环境下大学图书馆的职能转变 [J]. 图书与情报, 2010(1): 118-121.
- [29] ALISE. ALISE Research Taxonomy [EB/OL]. [2020-09-17]. <http://www.alise.org/research-taxonomy->
- [30] HowNet [EB/OL]. [2020-09-17]. [http://www.keenage.com/html/e\\_index.html](http://www.keenage.com/html/e_index.html).
- [31] 张军伟, 徐富明, 刘腾飞, 等. 行为决策中模糊规避研究的回顾与展望 [J]. 应用心理学, 2009, 15(3): 245-250.



# Research-Field Identification and Characteristics-Difference Analysis Based on Self-reported Expertise in CVs: Evidence from 1016 Faculties in 52 Chinese LIS Schools

HU Zhiwei PEI Lei

(School of Information Management, Nanjing University, Nanjing 210023, China)

---

**Abstract:** [ **Purpose/significance** ] This paper is intended to study topic distribution and term characteristics of domestic LIS teachers' expertise, and to measure its characteristics difference and interdisciplinary degree.

[ **Method/process** ] After extracting descriptive words from self-reported expertise in the CVs of 1016 teachers from 52 institutions, we identified their domain distributional and interdisciplinary characteristics using methods such as LDA topic clustering, word frequency analysis and membership degree with word categories. [ **Result/conclusion** ] It is found that domestic LIS teachers' expertise is centered on the lifecycle of information resource management, and information technology has a high penetration in the current research with a wide range of application scenarios. Specifically, new teachers employed in recent 5 years bear significant characteristics of interdisciplinary and industry-oriented, when referring to the words used in their self-reported expertise. We accordingly put forward some suggestions, including consolidating core knowledge system, promoting information technology to empower contemporary research, and improving interdisciplinary research ability.

**Keywords:** Self-reported expertise; Characteristics difference; LDA; Word frequency analysis

---

( 本文责编: 王秀玲 )