

基于知识图谱的个人信息管理研究热点与前沿分析

谢 笑¹ 李 晶¹ 谢阳群²

(1. 武汉大学信息管理学院 武汉 430072; 2. 淮北师范大学 淮北 235000)

摘 要 利用 CiteSpaceII 软件对 Web of Science 数据库中 1999–2011 年的个人信息管理领域相关文献进行定量分析,研究了国外个人信息管理的重要学科领域、知识基础、研究热点和前沿,为我国的个人信息管理研究提供了有力的参考依据。

关键词 个人信息管理 知识图谱 研究热点 研究前沿

中图分类号 G203

文献标识码 A

文章编号 1002-1965(2012)01-0102-09

Research Frontier and Hotspots Analysis of Personal Information Management Based on Mapping Knowledge Domain

XIE Xiao¹ LI Jing¹ XIE Yangqun²

(1. School of Information Management, Wuhan University, Wuhan 430072;

2. Huaibei Normal University, Huaibei 235000)

Abstract By using quantitative analysis with the information visualization software CiteSpaceII, the paper analyzes a series of overseas academic results and works in the field of personal information management during 1999 to 2011. Relying on analyzing related data which was from the Web of Science, the paper explores foreign important disciplines, knowledge bases, research hotspots and frontier concerning personal information management, which presents some valuable references to the study on personal information management in China

Key words personal information management mapping knowledge domain research hotspots research frontier

0 引 言

个人信息管理(PIM, Personal Information Management)是人们为日常使用和完成任务(与工作有关或无关的)以及履行其他个人职责(作为父母、雇员、朋友、团体成员等)而开展的获取、组织、保存、检索、使用和控制分散的文件(纸质的和电子的)、网页、电子邮件等信息项的活动。20世纪80年代以来,随着信息技术的发展,个人所掌握的信息海量增长,种类和形式更为丰富多样,使个人信息管理变得越来越困难。2005年以来,国际上成功召开了4次个人信息管理专题研讨会(PIM Workshop),来自心理学、计算机科学、经济管理学、图书情报学等领域的专家、学者共同探讨了PIM中的理论和实践问题,并取得了不错的成果。为了更为清晰地展现当今国际PIM研究的热点和前沿,本文利用知识图谱对国外PIM领域的文献进行了

以定量分析为主、定性分析为辅的研究,希望得出的结论能够对该领域的研究提供借鉴和启示。

1 数据来源、理论基础和研究思路

本文使用 Web of Science 数据库和 CiteSpaceII 软件对文献进行分析,在此基础上,得出一些有意义的结论。

1.1 数据来源 本文所使用的数据来源于 Web of Science 数据库,检索时间为 2011 年 4 月 15 日,以“personal information management”为主题词进行检索,共检索到文献 3 028 篇。将时间跨度设定为 1999–2011 年,文献类型设定为期刊论文(article)、会议论文(proceeding paper)和综述(review),并对学科进行筛选,选择与信息相关的学科(包括社会学、心理学等交叉学科),精炼后得到的文献数据为 1 425 篇。

1.2 理论基础与工具 知识图谱(Mapping Knowl-

收稿日期:2011-08-19

修回日期:2011-10-12

作者简介:谢 笑(1987–),女,博士研究生,研究方向:信息资源管理;李 晶(1985–),女,博士研究生,研究方向:信息分析与预测;谢阳群(1962–),男,博士,教授,研究方向:信息资源管理。

edge Domain) 也称为知识域可视化或知识领域映射地图,是近年来新兴的一种科学计量学的研究方法。它是把应用数学、图形学、信息可视化技术、信息科学等学科的理论和方法与计量学引文分析、共现分析方法结合,用可视化的图谱形象地展示学科的核心结构、发展历史、前沿领域以及整体知识架构的多学科融合的一种研究方法。它把复杂的知识领域通过数据挖掘、信息处理、知识计量和图形绘制而显示出来,揭示知识领域的动态发展规律,为学科研究提供切实的、有价值的参考^[1]。美国德雷赛尔大学(Drexel University)的陈超美博士利用 JAVA 语言开发的 CiteSpaceII 可视化软件为知识图谱的研究提供了极大的便利^[2]。CiteSpaceII 使用的光谱聚类算法能够实现对共现和共被引网络的自动聚类 and 标引^[3],采用的 Kleinberg^[4] 突变检测算法可以用于检测某个学科中研究兴趣的突然变化。本文采用了“关键词聚类”和“突变检测算法”来研究个人信息管理这一领域中的研究热点和前沿。

1.3 研究思路 本文采用文献计量学方法对个人信息管理领域中的文献的数量、作者和主要的研究机构进行了统计分析,主要使用 Excel、CiteSpaceII 等软件来绘制这些文献的关键词的聚类图谱,高被引文献的共被引网络知识图谱,从而得出该领域的研究热点。

2 研究现状

本文对 PIM 领域中的文献量、作者情况以及机构的情况进行了统计分析。

2.1 文献信息量统计分析 发表的文献数量在一定程度上代表了某个领域的研究水平与发展趋势。表 1 和图 1 显示的是 1999–2011 年 PIM 领域发表的文献数量及其增长趋势,其中 2011 年的数据只统计到 4 月 15 日,暂不分析。从图 1 可以看出,PIM 领域中的文献量除 2002 和 2010 年略有下降之外,其他年份均呈上升趋势。

表 1 文献信息量情况统计

年份	2011	2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999
数量	18	148	245	202	191	127	111	101	72	52	63	60	35

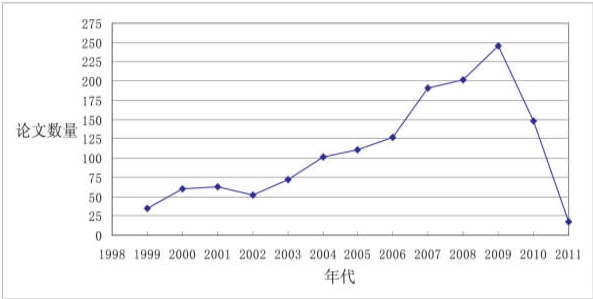


图 1 国际 PIM 领域文献增长趋势

2.2 论文作者统计分析

2.2.1 作者总体情况。经过统计,这 1 425 篇文章来自于 821 位作者(其中发表了 6 篇、8 篇、17 篇和 109 篇文章的均为 IEEE 协会),具体数据如表 2 所示。

表 2 作者发文情况统计

篇数	1	2	3	4	5	6	7	8	10	17	22	109
人数	499	238	57	8	8	1	5	1	1	1	1	1
占总人数(%)	60.78	28.99	6.94	0.97	0.97	0.12	0.61	0.12	0.12	0.12	0.12	0.12

这一结果并不是特别符合洛特卡定律,虽然按照平方反比分布,所有生产 1 篇论文的作者的的比例为 60.78% 这一条符合,但是就生产 n 篇论文的作者数量大约是生产 1 篇论文作者数量的 $1/n^2$ 这一点来看,除了发表 3 篇论文的作者符合以外,其他的都不相符。

该结果符合普莱斯定律,即在个人信息管理领域中,全部论文的半数系由该领域中的全部作者的平方根的那些人撰写的,“那些人”就是该领域中的核心作者或高产作者。统计的文章数量共有 1 425 篇,半数为 713 篇,该领域中的全部作者为 821 人,开平方根的结果为 28.65,根据表 2,生产 4 篇及以上论文的作者共有 27 人,符合普莱斯定律。

2.2.2 核心作者情况。核心作者(或高产作者)是指在某一领域中发表了多篇文章的作者。通过对作者的发表文章数量和被引频次进行统计分析,可以得出该领域中的核心作者及其学术影响力。本文选取发表文章数量为 5 篇及以上的作者为高产作者(包括 IEEE 协会、ACM 协会和 ETRI),如表 3 所示。

表 3 高产作者情况统计

排名	作者	发文量
1	IEEE	109
2	ACM	22
3	IEEE COMPUTER SOCIETY	17
4	Tjoa, A. M.	10
5	IEEE COMPUTER SOC	8
6	Ahn, G. J.	7
7	Bergman, O	7
8	ETRI	7
9	Nejdl, W	7
10	Whittaker, S	7
11	IEEE COMP SOC	6
12	Beyth-Marom, R.	5
13	Blobel, B.	5
14	Katifori, A.	5
15	Keil, M.	5
16	Kim, Y.	5
17	Lee, H.	5
18	Mont, M. C.	5
19	Nachmias, R.	5

本文将协会机构与个人作者分开分析。首先,美

国电气和电子工程师协会(IEEE),收录了2000-2010年期间召开的24次会议的论文,共140篇,包括本体论、个人隐私及安全、个人健康信息管理、个人数据空间构造、个人知识管理等。美国计算机协会(ACM),它在2006年1月“Communication of the ACM”上发表了一系列研究个人信息管理的文章,如个人信息管理、个人健康信息管理、个人信息管理与群体信息管理、个人信息管理中的数据统一等。韩国电子通信研究院(ETRI)于2008年10月,在韩国召开了第10届高级通信技术国际会议,主要研究了身份管理系统、数字身份安全、个人健康信息管理、分布式个人信息的管理方法。

就作者来看,发表文章最多的作者是Tjoa,他来自维也纳科技大学,主要是从本体论和语义网络角度研究个人信息管理。Ahn, Bergman, Nejd 和 Whittaker 都发表了7篇文章,Ahn来自美国的北卡罗来纳大学,主要从事PIM中个人隐私和信息安全的研究;Bergman和Whittaker都来自英国的谢菲尔德大学,Bergman主要研究个人信息管理系统和个人信息检索、再现等问题;Whittaker主要研究个人信息的检索与再现、用户检索偏好、个人电子邮件的信息管理问题;Nejd来自德国的汉诺威莱布尼兹大学,主要研究个人信息管理的用户行为和语义桌面搜索。同样来自英国谢菲尔德大学的Beyth-Marom和来自以色列特拉维夫大学的Nachmias与Bergman多次合作撰写论文,他们的研究领域相似。来自德国马格德堡大学的Blobel和来自美国华盛顿大学的Kim,他们主要都是研究个人健康信息管理。Katifori来自雅典大学,主要从本体论的角度研究个人信息管理,也研究用户行为和任务管理。Keil重点研究了个人信息管理的技术问题。Lee来自爱尔兰的都柏林城市大学,主要研究了个人信息管理中的隐私问题,个人健康信息管理以及个人信息管理相关的中介软件。Mont来自惠普公司,主要研究个人信息管理涉及的隐私问题。

2.3 主要研究机构统计分析 统计了发表论文在9篇以上的机构,统计数据如表4所示:

从论文发表的数量上看,华盛顿大学以19篇的数量名列第一。排在第二的是多伦多大学。其它发文数量较多的还有北卡罗来纳州立大学、谢菲尔德大学、IBM公司和普杜大学等。

从地区情况来看,美国(8所)占了发文机构总量的一半以上,这在一定程度上说明美国在个人信息管理领域中具有较强的学术影响力。其它的一些高产机构分布在加拿大、英国、台湾、香港和奥地利,说明该领域的研究机构分布广泛,且主要在发达国家和地区。这些高产机构多数都是高校,只有IBM这一商业公司

排名比较靠前,这表明高校机构是PIM研究的主要阵地。

表4 高产机构情况统计

排名	机构	篇数	地区
1	华盛顿大学	19	美国
2	多伦多大学	16	加拿大
3	北卡罗来纳大学	14	美国
4	谢菲尔德大学	14	英国
5	IBM公司	12	美国
6	普杜大学	11	美国
7	乔治亚州立大学	10	美国
8	国立交通大学	10	台湾
9	国立台湾大学	10	台湾
10	威斯康星大学	10	美国
11	香港理工大学	9	香港
12	科罗拉多大学	9	美国
13	密苏里州大学	9	美国
14	维也纳科技大学	9	奥地利

3 研究热点及前沿

利用CiteSpaceII对个人信息管理的学科领域和知识基础、研究热点及前沿和发展趋势进行了分析。

3.1 研究学科领域分析 从数据库的检索结果来看,关于个人信息管理的研究学科共有24类,进一步分析可以看出,主要涉及的学科有计算机科学(包括信息系统、理论和方法研究、人工智能、跨学科应用、软件工程、硬件和结构、控制论)、管理学(公共管理、社会保障)、图书情报学、商学(商业、财政)、医学信息学、运筹学和管理科学、经济学、心理学(交叉学科、应用、实验)和社会学(社会问题、交叉学科)。

本文将相关的数据导入CiteSpaceII中,将网络节点的类型(Node Types)选择为学科主题(category),将时区中的时段选择为1999-2011,时区分割(Time Slicing)选择为1,即将这13年的数据以年为单位切分为13片,将阈值选择为出现频次最高的30个字(top 30),运行软件得到图2所示的PIM学科领域分布图。

从图2可以看出10个明显的聚类,这些聚类代表着个人信息管理研究领域的主要学科,包括计算机科学、工程学、管理学、图书情报学、通信技术、医学信息学、健康护理科学和服务、商学、运筹学与管理科学和心理学。从图2可知,上述学科聚类之间联系紧密,相互交融和重叠,呈现出融合之势。按照图中节点的集中度大小,结合各个学科中的文献数量,可以客观地反映出个人信息管理的各个学科领域重要性。可以得出目前国际上有关PIM的研究主要集中在计算机、图书情报学、管理学、商学、医学信息学以及工程学等学科,且已经形成了一些重要的研究成果。这一分析结果与

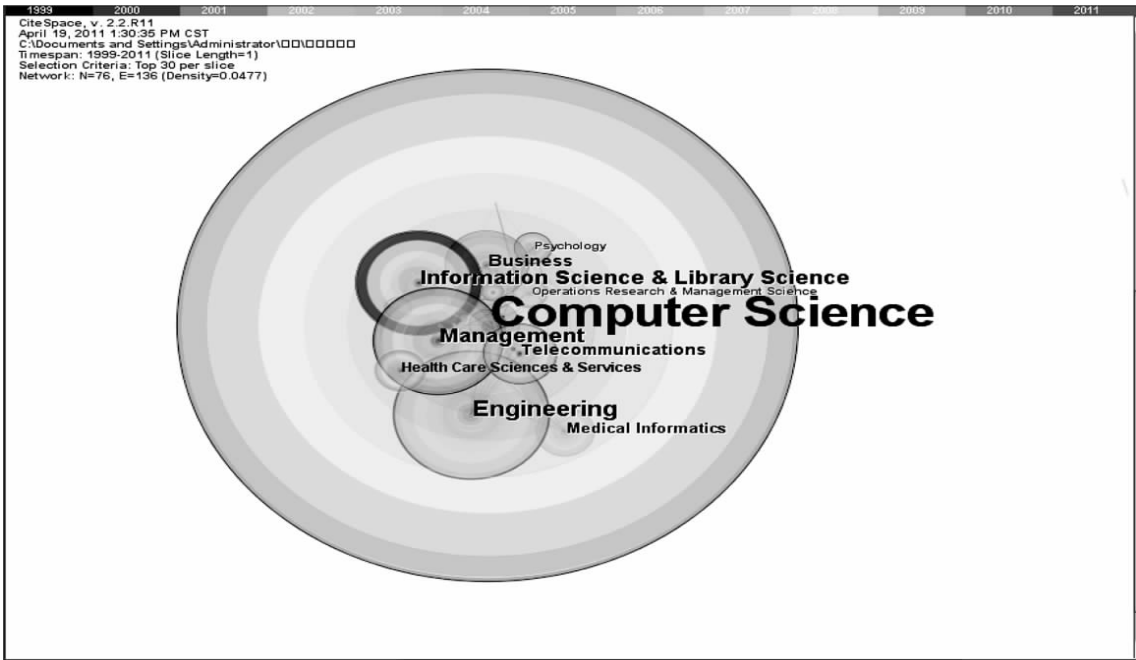


图 2 个人信息管理研究的学科领域

数据库的检索结果基本一致。计算机科学包含各种与计算机和信息处理相关的系统学科,如编程语言、程序设计、软件和硬件、人工智能等。图书情报学在我国一般被归为信息管理专业中,主要包括图书馆学,情报学,文献信息的组织、检索、存储、咨询、分析,信息系统的设计与分析等。管理学是一门综合性的交叉学科,它是系统地研究管理活动的基本规律和一般方法的科学,其目的是在现有的条件下,如何通过合理的组织和配置人、财、物等因素来提高效力和效率。商学包含了范围非常广泛,如会计、财务管理、银行学、市场营销、物流、金融、人力资源管理、经济等学科。医学信息学是医学和计算机等科学的交叉学科,是医学发展的必经阶段,是一门新兴的交叉学科。工程学是一门应用学科,是研究数学和其他自然科学原理在各行业中应用的方式、方法的一门学科,同时也研究工程进行的一般规律,并进行改良研究。从以上的分析结果可以看出,随着科技、信息和网络的发展,个人信息管理涉及的各学科领域都在向交叉学科的方向发展,各学科之间相互交融,有共同的研究方向,也有不同的研究方向,而计算机科学、图书情报学、管理学、商学、医学信息学和工程学等学科中都有某些部分构成了个人信息管理的研究领域。此外,运筹与管理学(operations research & management science)、信息技术这两个学科领域有着较高的突现度(分别为 4.64 和 4.34),这说明,随着科技的发展,网络的普及和个人电脑的广泛应用,运筹学在管理学领域的渗透,运筹学和信息技术在个人信息管理活动中的地位 and 作用日益凸显,在个人信息管理的研究中也得到了越来越多的关注。

3.2 知识基础 知识基础能够反映出某一学科领

域前沿的本质,具有相对的稳定性,有助于明确和预测研究前沿和动态趋势^[8]。本文在 CiteSpaceII 中将节点类型选择为“参考文献(cited reference)”进行分析运行,生成图 3 所示的知识图谱,共有 291 个不同的文献节点和 966 条连线,从中能够得出一些个人信息管理领域中的经典论文或是奠基性文献,如表 5 所示。

表 5 个人信息管理领域高被引核心作者及其经典文献

排名	被引频次	中心度	作者	被引著作/论文题目	出版年
1	43	0.06	Nonaka I. & Takeuchi H.	The Knowledge Creating Company: how Japanese companies create the dynamics of innovation	1995
2	37	0.19	Fornell C. & David F. L.	Structural equation models with unobservable variables and measurement error - algebra and statistics	1981
3	31	0.09	Davis F. D.	Perceived usefulness, perceived ease of use, and user acceptance of information technology	1989
4	26	0.02	Malone T.	How do people organize their desks - implications for the design of office information - systems	1983
5	24	0.00	Lansdale, M. W.	The Psychology of Personal Information Management	1988
6	23	0.02	Davis F. D., et al.	user acceptance of computer - technology - a comparison of 2 theoretical - models	1989
7	22	0.00	Dumais S. T., Cutrell, E., et al.	Stuff I' ve seen: a system for personal information retrieval and re-use	2003
8	20	0.00	Whittaker, S. & Sidner, C.	Email overload: exploring personal information management of email	1996
9	19	0.00	Nunnally J. C.	Psychometric theory	1978
10	18	0.01	Anderson J. C.	Structural equation modeling in practice - a review and recommended 2-step approach	1988
11	18	0.02	Barreau, D. & Nar-di, B. A.	Finding and reminding: File organization from the desktop	1995

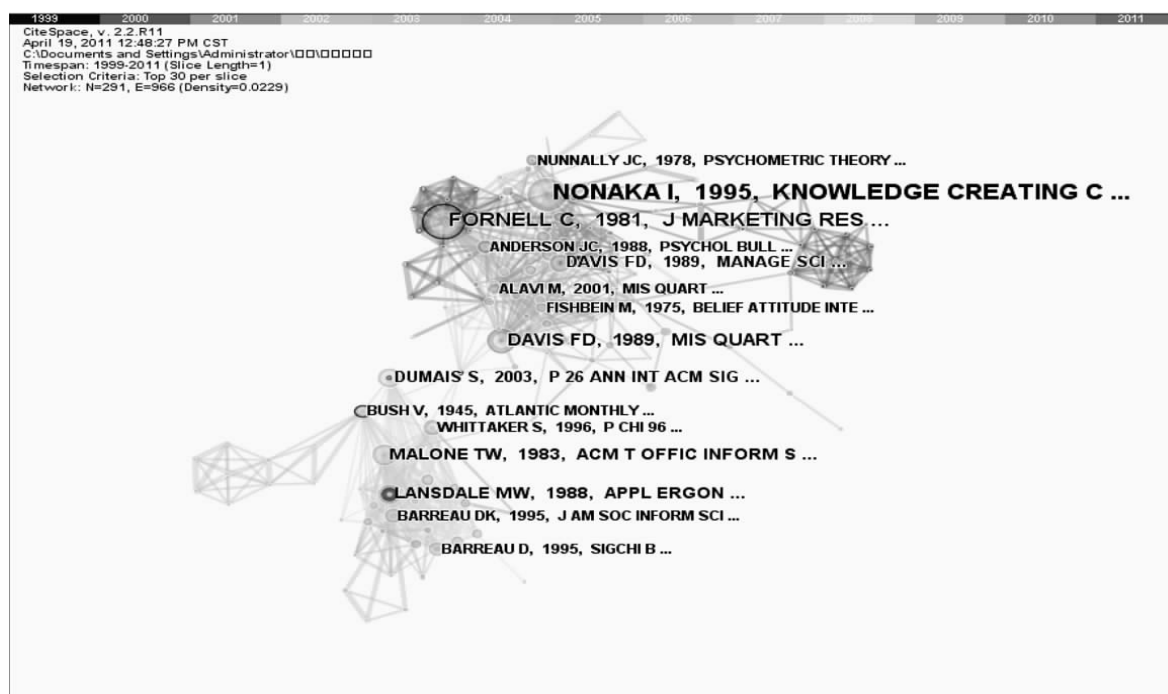


图3 文献共被引网络知识图谱

Nonaka 等所著的《知识创造企业: 日本企业如何创造动态性的创新》可以被视为是知识管理领域的奠基之作, 作者提出知识创造可以是个人层次、组织层次, 知识转换包括 4 个主要过程: 从隐形到显性, 显性到显性, 显性到隐性, 以及隐性到隐性; 并提出组织应该打破传统的自上而下或自下而上的管理方式, 采用更有利于知识创造和交流的管理方式。可以说本书也是 PIM 的奠基之作, 为个人层次的知识管理提供了理论和实践上的指导^[1]。Fornell 等针对 Bagozzi 提出的收敛与区分标准提出了质疑, 以《不可观测变量和测量误差的结构方程模型——代数和统计》进行了回应, 作者认为解释统计结果必须依据研究目的, 此外作者还证明了 Fornell-Larcker 测量体系内部具有一致性, 也符合数据与提取变量的对应规则^[2]。Davis 在《感知有用性、感知易用性和用户对信息技术的接受》一文中, 对 152 位用户使用的 4 种应用程序进行了有效测试, 实证分析的结果表明: 与易用性相比, 有用性对用户的使用行为影响更大^[3]。Malone 在其《人们如何管理自己的桌面——办公室信息系统设计的含义》一文中对人们的办公室行为进行了研究, 指出信息分类会降低人们检索和回忆该信息的速度, 而且对于那些可以归入多个类目的信息, 人们存在认知上的困境^[4]。Lansdale 在其《个人信息管理心理》一文中指出, 随着信息技术的发展, 信息资源的种类、形式越来越丰富, 原先用于处理传统纸质信息的方法不再适用, 需要从收集、整理、组织、检索等方面探究信息管理自动化的方法^[5]。Davis F. D., et al. 在其《用户对计算机技术的接纳——两种理论模型的比较》一文中, 对 107 位用户

进行了长期的调查研究后发现, 感知有用性对用户的意图有强烈影响, 感知易用性对用户的意图影响较小, 且会随着时间的流逝而减弱, 但是也很重要, 主观规范对用户目的没有任何影响, 作者提出了简单有效的用户接受度决策模型, 该模型能够引导管理干预, 从而降低计算机使用的不充分^[6]。Dumais 在其《Stuff I've seen: 一个个人信息检索和再利用系统》文章中分析了用户的个人信息管理行为, 并设计出了个人信息检索和再利用的 SIS (Stuff I've Seen) 搜索工具, 他认为如果能够使用 SIS 找到所需要的信息, 那么在保存文件时, 能够节约添加标签和注释的时间^[7]。Nunnally 所著的《心理测量学理论》可以看作是心理学与行为科学的经典著作, 奠定了心理测量理论的基础, 同时对工程学、医学以及其他领域的研究都有所启示^[8]。个人信息管理离不开对用户行为的追踪、分析与研究, 是行为科学的重要分支领域, 因此用户行为和心理学理论也是 PIM 的基础理论。Anderson 的《实践结构化方程模型——综述及两步法推荐》一文为使用结构方程建模的研究者提供了理论测量和模型开发指导, 提出了一种综合性的、两步建模方法, 使用了一系列的嵌套模型, 以及相应的卡方差异检验, 并讨论了该方法与一步法相比的优势, 以及评估方法的开发^[9]。Barreau 在 1995 年发表的《寻找和再现: 桌面文件的组织》是早期研究个人信息检索行为的经典论文, 论文分别对使用个人计算机的早期用户 (使用 DOS, Windows 3.1 和 OS/2) 和使用 Macintosh 的经验用户进行了采访, 发现用户在寻找和再现信息时, 往往更喜欢使用导航浏览文件夹来寻找信息而不是通过搜索来寻找信息^[10], 后

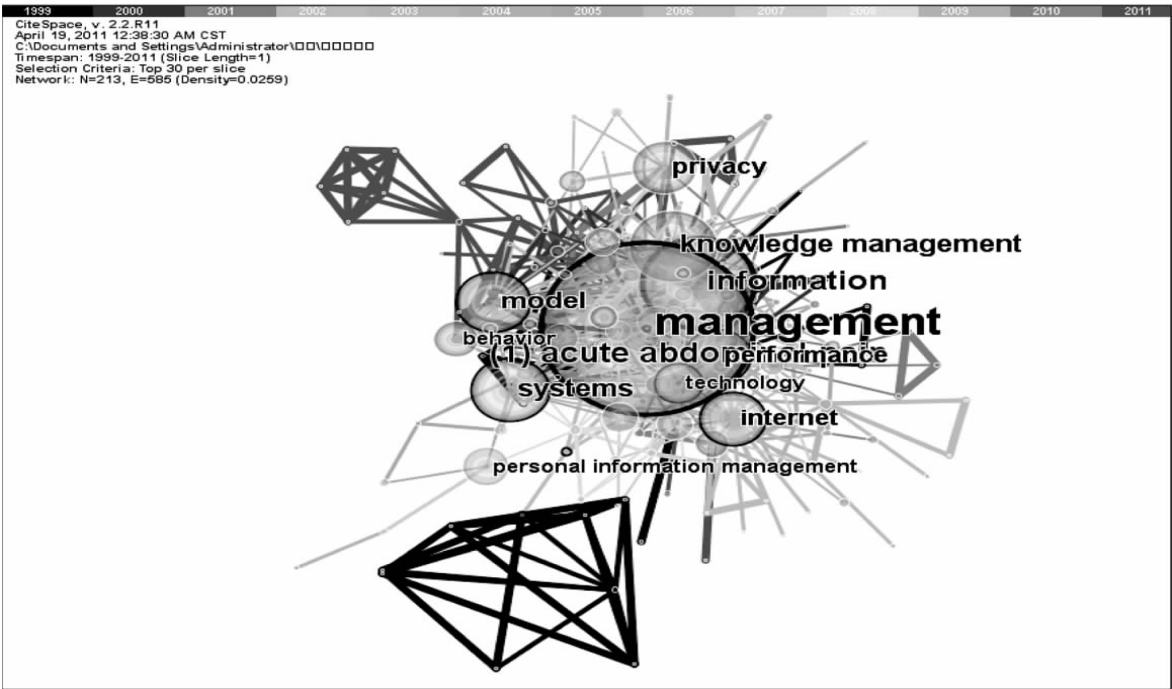


图 4 关键词共现图谱

来的许多学者通过研究调查也得出了类似的结论(如 Boardman&Sasse; Capra&Pérez – Quiñones 2005; Kirk et al.; Teevan et al.)。这 10 位高被引核心作者的论文涉及了知识管理、商学、心理学、行为科学、信息技术、计算机等多个学科领域,证实了上文得出的个人信息管理涉及多学科领域的结论,他们的研究为 PIM 提供了重要的理论支持。

3.3 研究热点 关键词是文章的核心,是作者对文章的主题进行精炼之后得到的结果,有着很强的代表性,因此,在分析时,往往使用关键词来研究某一学术领域中的研究热点。本文将网络节点的类型(node types) 选择为关键词(keyword), 其他参数不变,运行软件得到图 4 所示的关键词共现图谱,共有 213 个不同的关键词节点和 585 条连线:

根据社会网络分析理论,中心度是衡量网络中某节点对经过该节点并彼此相连接的另外两个节点的控制能力,一定程度上表征某节点与其他节点之间的联系以及在整个网络中的重要地位和作用^[4]。在图 4 中,具有较高中心度的词用紫色的圈圈表示出来,这些关键词包括: 管理、信息、系统、技术、网络、绩效、模型等。关键词年环的厚度与关键词出现的频次正相关,表 6 列出了频次最高的 10 个关键词及其中心度(其中排名第 10 和第 11 的关键词出现的频次一样,按照字母顺序排列):

从图 4 可以看出,管理是最大的节点,信息次之。个人信息管理,是对个人的信息进行管理,信息和管理自然成为个人信息管理的基础。信息管理这一领域中的许多理论、方法和工具都可以应用到个人信息管理

中,但是由于个人信息管理的核心是个人信息,因此,信息管理在个人信息管理中应用,还需要做出一些调整和改进才能够适用,这也是许多学者研究的重点之一,是该领域的研究热点。

表 6 高频关键词表

排名	频次	中心度	关键词
1	154	0.22	Management
2	82	0.23	Information
3	68	0.06	Knowledge Management
4	63	0.17	Systems
5	58	0.13	Model
6	54	0.14	Performance
7	50	0.16	Internet
8	48	0.05	Privacy
9	39	0.17	Technology
10	38	0.09	Behavior
11	38	0.04	Personal Information Management

排名第三的是知识管理。知识管理是信息管理发展的第四个阶段,是一种重视与人打交道的信息管理活动,其实质是将结构化与非结构化的信息与人们利用这些信息的规则联系起来,它重视团队成员的联合并鼓励经验和知识的共享,其核心是知识的创造、应用、学习、理解和协商^[7]。此外,个人信息管理的一个重要的研究分支就是群体信息管理(Group Information Management, GIM), 它是指个人信息在公共领域中的作用,解决的是个人信息在群体内共享以及共享的基础标准等问题,其实质与知识管理类似,因此,从知识管理的角度研究个人信息管理成为当前国际上的研究热点,这也印证了上文“高被引核心作者及其著作”中包含大量知识管理学领域经典文献。

排在第四位和第五位的是系统和模型。个人信息管理是从搜索(search for)、遇见(encounter)、翻译(interpret)、组织(organization)、保存(keeping)、寻找(finding)/再现(refinding)、再利用(re-use)、再获取(re-access)到信息的最终利用(ultimately use)这样一个过程,在这个过程中,会出现各种各样的问题,包括难以预估信息价值、存储位置分散、难以寻找和再现等,解决这些问题的最佳方法就是通过建立模型、设计出合适的个人信息管理系统来实现个人信息的有效管理。而根据数据库的搜索结果,其中有 477 篇文献是计算机科学和信息系统学科的,这都说明个人信息管理系统和建模的研究是当前研究的热点问题之一。

排名第六位的是绩效。设计出合理的个人信息管理系统(或工具),还要对其进行评价,评价的指标包括可用性(效力、满意度、易用性、易学性)、绩效、采纳和使用等。对个人信息管理的系统(或工具)进行评价,指出问题并加以改进,能够促进这一系统(或工具)的广泛使用,因此,个人信息管理系统(或工具)的评估是目前该研究领域非常活跃的热点之一。

排名七、八、九位的分别是网络、隐私和技术。随着因特网和信息技术的发展,网络上的个人信息不断增加,个人信息在群体和组织中共享,个人信息管理中的个人隐私和信息安全问题变得越来越重要。如何使用先进的信息技术更好的管理个人信息,确保网络共享中的个人信息安全是当前研究的难点,也是热点之一。

排名第十的是行为和个人信息管理。个人信息管理是本文的检索词,个人信息管理活动过程有各种各样的个人信息管理行为,这些行为也是个人信息管理的研究热点。

由此可以看出,管理、信息、知识管理、个人信息管理构成了个人信息管理的研究理论基础,同时也属于信息资源管理领域。信息资源管理的概念出现于 20 世纪 70 年代末、80 年代初;进入 20 世纪 90 年代以后,随着以网络为代表的现代信息技术的飞速发展,形成了“信息高速公路”,促进了信息资源管理的进一步发展。知识管理是 20 世纪 90 年代出现的新名词,它的存在主要依赖于信息管理和对工作实践的管理,现今,知识管理经过 20 多年的发展,已经在很多机构和部门中实施,并逐渐与其他的社会活动建立了紧密的联系,并取得了不错的进展。个人信息管理在最近的十几年里不断发展,成为新的学术领域的研究热点。系统、模型、绩效、网络和技术则属于信息系统和信息技术研究,主要包括信息系统的分析、设计、建模、运行、检测、维护和更新的过程,现有系统的评价与改造,网络环境下系统的运行与发展,网上个人信息资源的共享与开发,信息技术的应用及发展等,其中信息技术是信息系统和网络得以产生和发展的基础。隐私是个人信息中的一个重要方面,会影响到群体信息管理的发展,Ahn 和 Lee 都对个人信息管理中的这一问题进行了研究。行为是个人信息管理的影响因素之一,主要包括对用户行为的研究,许多学者进行过这一方面的研究,如 Malone 研究了人们的办公室行为。

3.4 研究前沿 突变词是指短时间内增长速度突然加快或使用频率突然提高的术语。CiteSpaceII 与 CiteSpace 相比,新增了突变检测的功能,能够利用突变词的动态变化性,研究某一学科领域的研究前沿。本文选用突变检测算法来检测 PIM 领域的突变专业术语,同时选择时间区域(timezone) 聚类视图,可以生成图 5 所示的知识图谱。

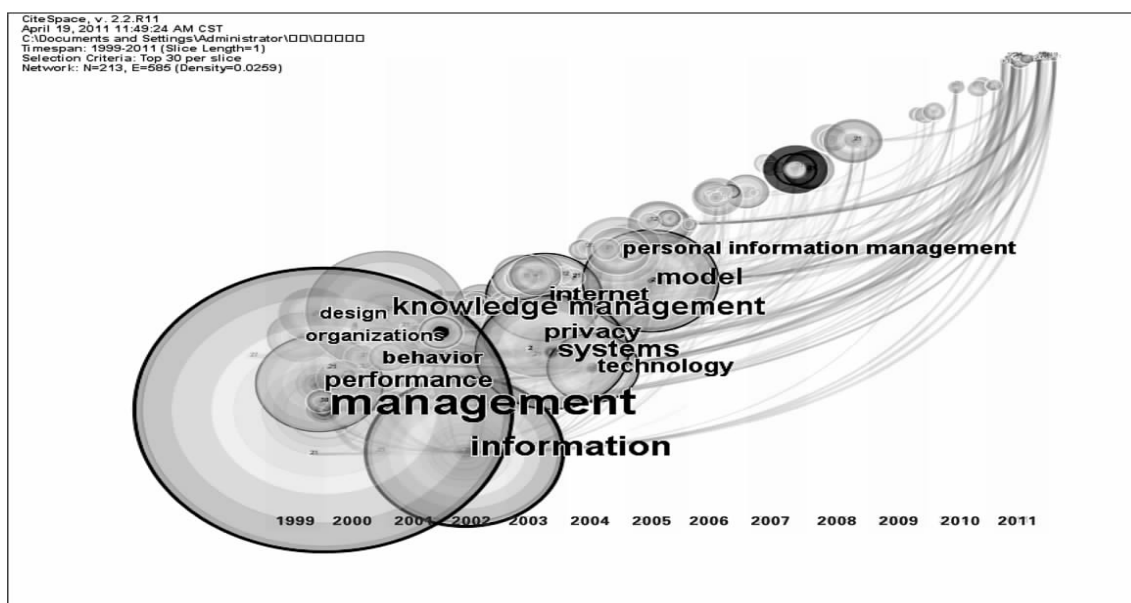


图 5 关键词聚类图谱

按照突现度这一量度指标从大到小排列,一共探测出 5 个突变词,如表 7 所示。

表 7 突变词分布列表

突现度	频次	中心度	突变词	年份
3.56	29	0.02	Information-technology	2007
3.18	7	0.01	Human factors	2008
3.10	15	0.05	Location management	2002
2.95	21	0.06	Strategy	2001
2.81	18	0.01	Work	2000

其中 2007 年出现的“信息技术”一词突显度最高,表明在 2007 年,个人信息管理领域中的研究开始关注于信息技术在这一领域中的应用,如何充分地利用信息技术来实现有效的个人信息管理,包括信息系统、数据库、信息检索、人工智能、软件工程等技术方法已经成为前沿热点。2008 年出现的“人的因素”,表明个人信息管理领域中的信息用户研究,包括信息用户的分类、分析和教育,用户信息需求的决定因素、表现形式和满足方式,用户利用信息资源的行为与心理,用户潜在的信息资源开发等问题,其中用户信息行为的研究较多。用户的信息行为由于受到任务、工具和环境等因素的影响,会产生个体差异,这成为个人信息管理研究的前沿领域。因此,在 PIM 工具的设计中,我们必须要以用户为中心,根据用户的需求和个体差异来设计出个性化的个人信息管理工具,为不同类型的用户设计出不同类型的个人信息管理工具,并增加工具的自适应性。2002 年出现的“存储位置管理”、2001 年出现的“战略管理”、2000 年出现的“工作”都成为新的研究重点。首先,随着信息技术的发展,存储信息的载体变得越来越多,如计算机、移动硬盘、笔记本电脑、平板电脑、PDA、纸张等,而人们在存储信息时,可能会出于害怕遗忘重要信息的缘由,将相同的信息存储在多个不同的位置,这导致了信息存储位置的日趋分散,难以在需要的时候及时地再现所需信息;这种在不同存储位置保存相同信息的做法还会带来信息的更新和同步的不一致问题;因此,信息存储位置的管理成为个人信息管理的研究前沿之一。其次,个人信息管理的战略选择,也就是个人信息管理的方法选择,在信息的存储、组织和再现的过程中,如何使信息序化和优化,选择适当的信息组织、检索的方法与工具,能够在需要的时候及时再现和再利用信息是很重要的,因此这一方法的选择就成为研究的重点与前沿。此外,“工作”一词包括工作有关的信息、记忆的管理,也可以理解成个人信息管理中的任务管理,如何提醒人们目前未完成的任务、跟踪进行中的任务状态、校对和保存这些与任务相关的信息、创建有效的任务管理工具,实现工作任务的管理也是当前研究的前沿之一。目前的任务管理工具

主要是电子邮件,虽然其最初设计时并不是出于任务管理的目的,但是由于它有着工作和信息导管的功能,所以被用于个人信息管理,尤其是任务管理,当然,这一工具是有缺陷的,尤其是在提醒人们回复邮件和校对信息方面。最近,人们也提出了几种改进电子邮件中的任务管理的方法,如信息结构化,即将收件箱与某一任务相关的所有信息聚集在一起的信息组织方法;嵌入任务管理支持,这一技术能够识别和组织任务相关的信息构建,直接支持提醒和校对;工作流程,依据基本任务顺序和人们在执行任务过程中的角色来组织信息。

4 研究的不足与展望

本文用 CiteSpaceII 对 Web of Science 数据库中的个人信息管理相关文章进行了统计分析,得出了目前国际上个人信息管理领域的研究热点和前沿,但是,这一方法也有一定的局限性,它只能对某一特定数据库,如 Web of Science 收录的文献进行分析,不能实施跨库分析,这就可能存在分析不全面的情况。例如,国际上成功召开了 4 次 PIM 专题会议,并取得了许多有意义的研究成果,但是由于这些成果未被 Web of Science 数据库收录而没有纳入本文的研究范围;一些学者出版了专门研究个人信息管理专著,如 2007 年华盛顿大学出版社出版的《个人信息管理: 机遇与挑战(personal information management: challenges and opportunities) ^[3]》、Willian Jones 和 Jaime Teevan 的《个人信息管理(personal information management) ^[3]》、William P. Jones 的《保持发现(Keeping Found Things Found) ^[3]》等。此外,还有不少大学和研究机构推出的 PIM 研究项目(如《Keeping Found Things Found》、《My Life Bits ^[3]》等) 及其成果等也由于 CiteSpaceII 软件的局限性未被纳入本项研究中。因此,针对上述研究的不足,在未来还应综合利用其他书目统计和文献计量方法对 PIM 热点和前沿进行更全面、深入的研究。

5 总 结

本文利用 CiteSpaceII 软件对 Web of Science 数据库中 1999-2011 年的个人信息管理领域相关文献进行定量分析,研究了国外个人信息管理的重要学科领域、知识基础、研究热点和前沿。结论得出,国际 PIM 的研究领域集中在计算机、图书情报学、管理学、商学、医学信息学和工程学等学科领域,研究的热点主要是信息管理、知识管理、系统及建模、工具与评估和个人隐私问题,未来的发展方向是研究 PIM 的信息技术、个体差异、存储位置管理、战略方法选择等方面。

参考文献

- [1] 知识图谱 [EB/OL]. [2011-04-16]. <http://baike.baidu.com/view/5361633.htm#sub5394976>
- [2] The CiteSpace Homepage [EB/OL]. [2011-04-16]. <http://cluster.cis.drexel.edu/~cchen/citespace/>
- [3] Chen C, Ibekwe-San Juan F. The Structure and Dynamics of co-citation Clusters A Multiple-Perspective co-citation Analysis [EB/OL]. [2011-04-16]. <http://www.pages.drexel.edu/~cc345/papers/papers.html>
- [4] Kleinberg J. Bursty and Hierarchical Structure in Stream [J]. Data Mining and Knowledge Discovery, 2003 (10) 393 - 397
- [5] 任红娟, 张志强. 基于文献记录的科学知识图谱发展研究 [J]. 情报杂志, 2009(12) : 86-90
- [6] Ikujiro Nonaka, Takeuchi Hirotaka. The Knowledge-creating Company: How Japanese Companies Create the Dynamics of Innovation [M]. New York: Oxford University Press, 1995
- [7] Fornell C, David F L. Structural Equation Models with Unobservable Variables and Measurement Error—algebra and Statistics [J]. Journal of Marketing Research, 1981(8) : 382
- [8] Davis F D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology [J]. MIS Quarterly, 1989(9) : 319
- [9] Malone T. How do People Organize Their Desks—implications for The Design of Office Information—systems [J]. ACM Transaction on Office Information Systems, 1983, 1(1) : 99
- [10] Lansdale M W. The Psychology of Personal Information Management [J]. 1988, 19(1) : 55
- [11] Davis F D, et al. User Acceptance of Computer—technology—a Comparison of 2 theoretical—models [J]. Management Science, 1989(8) : 982
- [12] Dumais S T, Cutrell E, et al. Stuff I've Seen: a System for Personal Information Retrieval and Re-use [C]. In Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM Press: 72-79
- [13] Nunnally J C. Psychometric Theory [M]. New York: McGraw Hill, 1978
- [14] Anderson J C. Structural Equation Modeling in Practice—a Review and Recommended 2—step Approach [J]. Psychological Bulletin, 1988(3) : 411
- [15] Barreau D, Nardi B A. Finding and Reminding: File Organization from the Desktop [J]. SIGCHI Bulletin, 1995, 27(3) : 39
- [16] 刘军. 社会网络分析导论 [M]. 北京: 社会科学文献出版社, 2004: 122
- [17] 马费成. 信息管理学基础 [M]. 武汉: 武汉大学出版社, 2005: 38
- [18] William Jones, Jaime Teevan. Personal Information Management: Challenges and Opportunities [M]. The University of Washington Press, 2007
- [19] William Jones, Jaime Teevan. Personal Information Management [M]. The University of Washington Press, 2007
- [20] Keeping Found Things Found [EB/OL]. <http://kftf.ischool.washington.edu/index.htm>
- [21] My Life Bits [EB/OL]. <http://research.microsoft.com/en-us/projects/mylifebits/> (责编:刘武英)

(上接第49页)

由表3可知,1987-2007年间我国电子信息产业规模居于前三位的地区分别是广东、江苏和上海,且这三个省的电子信息产业占全国比重大体上呈上升的态势,到2007年占全国比重高达67.72%,说明我国电子信息产业的集中度越来越高,也从侧面反映出电子信息产业在我国其他地区的分布是相当零散的,两极分化问题越来越严重。主要原因在于这三个省均位于东部沿海地带,改革开放较早,凭借优越的区位条件、宽松的政策环境和健全的市场经济体制吸引大量的外商投资进入电子信息产业领域,逐渐形成了具有大规模电子信息产业相对聚集的市场。

3 结论

本文从偏度和峰度角度构建了产业聚集指数,在此基础上对1987-2007年我国电子信息产业的聚集水平进行了评价与分析,得到以下结论:第一,从整体来看,我国电子信息产业的聚集指数值在整个研究区间上均大于18.42,说明我国电子信息产业的聚集水平较高,产业聚集现象十分明显。第二,从聚集指数的变动趋势来看,1987-2007年间我国电子信息产业的聚集指数一直保持在较高的水平,且从各个年份的变化

趋势上看,呈现出先快速上升,后缓慢下降的倒U型曲线变化趋势。第三,从聚集地的分布来看,广东、江苏和上海这三个东部沿海省市是我国电子信息产业的主要聚集地,且1987-2007年间这三个地区的集中度越来越高,反映出电子信息产业的两极分化问题越来越严重。

因此,在电子信息产业聚集水平不断提高的同时,更要注重电子信息产业的创新扩散和梯度转移,逐步缩小地区之间的差距。

参考文献

- [1] Krugman P. Increasing Returns and Economic Geography [J]. Journal of Political Economy, 1991, 99: 483-499
- [2] Yang X. Economics: New Classical Versus Neoclassical Frameworks [M]. Blackwell Publishers, 2001
- [3] 王子龙, 谭清美, 许箫迪. 高技术产业聚集水平测度方法及实证研究 [J]. 科学学研究, 2006(5) : 708-710
- [4] 赵玉林, 魏芳. 基于熵指数和行业集中度的我国高技术产业集聚度研究 [J]. 科学与科学技术管理, 2008(11) : 123-125
- [5] 杨洪焦, 孙林岩, 梁冬寒. 我国高技术产业聚集度的变动趋势及区位因素分析—以电子及通信设备制造业为例 [J]. 科学学研究, 2009(9) : 1335-1337
- [6] 林晶. 高技术产业集聚变动趋势的实证研究 [J]. 中国高新技术产业, 2009(8) : 1-2 (责编:刘武英)