

国外信息资源管理研究进展及热点分析

——基于 IRMJ 和 JASIS 的分析

毕强 滕广青

摘要 为反映国际信息资源管理领域研究进展与热点内容,本研究对 2003 - 2008 年国际信息资源管理领域 IRMJ 和 JASIS 两大权威期刊文献进行了梳理和归纳。在理论探索与技术应用层面,本体理论、形式概念分析(FCA)、语义学、数据挖掘和以计算机为中介的交流(CMC)这五个方面存在着不同程度的相互交叉和渗透,是目前信息资源管理研究的热点;在人文价值与社会经济层面,知识共享将代表未来知识管理的发展趋势;在用户信息行为和满意度研究中,以用户为核心的理念将成为研究和应用的指导思想。而跨语言与文化差异方面的研究重点则是在保留现有的语言和文化多样性的基础上实现人类异类语言和文化间的无障碍信息沟通和知识交流。关于信息道德、隐私保护与青少年信息安全问题还需要继续研究和探索。参考文献 42。

关键词 信息资源管理 研究进展 研究热点 述评

分类号 G350

ABSTRACT To reflect the progress and hot spots in the international information resource management research, this study conducts an analysis and summarization of IRMJ and JASIS, the two authoritative journals in the international information resource management, published in 2003 - 2008. In the aspects of theoretical exploration and technology application, ontology theory, formal concept analysis (FCA), semantics, and data mining and computer-mediated communication (CMC), there are varying degrees of cross-cutting and penetration, which are the focal points in the study of information resource management. In the aspect of the humanity value and socio-economics, knowledge-sharing will represent the future development trend of knowledge management; in the user information behavior and satisfaction study, the idea of regarding the user as the core will become the guiding ideology of research and application. The research focus of cross-language and cultural differences is to realize the barrier-free information communication and exchange of knowledge between heterogeneous human languages and cultures, based on retaining the existing language and cultural diversity. The issues of information ethics, privacy protection and teenage information security require further research and exploration. 42 refs.

KEY WORDS Information resource management. Research progress. Hot spots of study. Discussion.

CLASS NUMBER G350

通过对 2003 - 2008 年国外信息资源管理领域两大主要文献《信息资源管理杂志》(The Information Resources Management Journal, IRMJ) 和《美国信息科学与技术杂志》(Journal of the American Society for Information Science and Technology, JASIST) 所发表论文的统计和梳理, 分别从理论探索与技术应用、人文价值与社会经济两个层面挖掘和分析国外信息资源管理领域的前沿课题以及研究的热点内容。

1 信息资源管理中理论探索与技术应用层面的研究热点

1.1 基于本体理论的研究

本体(Ontology)在信息科学领域是指共享概念模型的明确的形式化规范说明。随着信息技术的应用逐渐向自动化和智能化方向发展, 本体作为概念模型(Conceptual Models)和概念间视图关系(Conceptual Schemas)的一种规范性

描述,被广泛应用于知识的共享与重用。

在本体理论研究上,通过对 Barry Smith 的“本体论”的讨论,研究者构建了信息系统本体,并提出构建信息系统本体应考虑到来自哲学和科学史的支持,以及加尔达摩(Gadamer)诠释学在理解信息本体和数据库用户任务方面起到重要作用。通过对 Wand、Weber 以及 Guarino 对本体驱动信息系统研究的比较,分析了信息系统的本体(Ontologies of Information Systems)与对信息系统的本体性描述的区别,使本体研究的相关基本概念得到了更清晰的理解^[1]。算法方面,Ahmad Kayed、Nael Hirzallah、Luai A Al Shalabi 和 Mohammad Najjar 提出了不同于先前大多数算法提供所有概念间关系的新算法,能够对概念间关系进行提炼,并已开始应用于实践^[2]。

随着本体理论的不断成熟和完善,国际信息资源管理研究中关于本体的相关研究也逐渐从理论探索走向实际应用研究。在语义 Web 检索方面,比较试验表明,针对不同检索类型、不同内容的检索关键字,结构化的基于任务的索引本体不仅可以改善精度,而且能够减少搜寻时间,提高了检索效率、检索效益以及检索精确性^[3]。在其他领域,研究者通过利用岛屿驱动算法(Island-driven Algorithm)和潜在语义分析(latent Semantic Analysis, LSA)提出基于本体的言语行为识别是有效的对话管理。通过分析语义取向的文本信息产生出一种优于基线法的本体支持的极性挖掘方法。此外,本体理论在知识管理领域和组织记忆评估中也得到应用。目前,基于本体理论的应用除包括信息检索、语义 Web 领域外,还包括信息抽取、异构信息系统的互操作和集成、知识管理以及数据挖掘等领域。

1.2 形式概念分析的应用研究

形式概念分析(Formal Concept Analysis, FCA)最早由 Rudolf Wille 在 1982 年提出。它是对概念的哲学理解的数学化,是以人为中心构造和分析数据的方法,是数据的内在结构、关联和相互依存关系的可视化方法。近几年来,形式概念理论被广泛应用于本体的构建,二者形成了伴生关系,促使形式概念理论的应用领域

不断扩展和延伸,其中具有代表性的应用包括信息检索和信息工程两个方面。

信息检索方面,有的研究从概念出发将信息检索的焦点看作是一个特定类型的信息检索任务,构建信息研究的概念框架,并从心理学角度通过对概念空间信息流的计算实现从文本语料库自动构建高纬度语义空间,大大提高了查询模型的有效性和精准度。同时,利用赫布型学习算法(Hebbian type of learning algorithm),使用共同出现的数据作为输入,通过对邻近概念的探索找到新概念之间的关联,对探索未知的概念间关系提供了支持^[4]。研究者还提出了多层次动态性质的概念,认为客户相关性是最现实的情景相关性,要求通过交互式信息检索参考其潜在的动态性质。此外,Pablo J Garces、Jose A Olivas 和 Francisco P Romero 通过语义区域从模糊角度看待概念,改善了大多数搜索引擎的词语匹配程度^[5]。

信息工程方面,通过将数据库、实体和文件等多种对象在概念网络中呈现,在知识域中借助语义协同在各个实体之间获得独特的路径。Manuel Mora 等以系统最初的概念形式开发了信息系统文献使用模式和框架,减少当前研究中的焦点冲突并努力寻找学科领域的基本范式^[6]。另外,在类层次设计、数据挖掘和知识发现等方面,形式概念理论都有广泛的应用。

1.3 语义学的相关研究

1.3.1 语义关联(Semantic Association)

有研究者认为,如果实体彼此间的联系能够被发现,数据库中的关系序列能够被明确表示,则就可以构造两个实体之间的语义关联,引入一种集成了语义关联操作与文档操作的查询语言来发现语义关联。在此基础上,根据查询语言中指定的实体是否预先为用户所知以及文本信息是否在文档中被利用,对查询类型进行分类。如果该实体的前提是事先预知,则在查询语言中属于非常低效的模仿行为^[7]。

1.3.2 潜在语义分析(Latent Semantic Analysis, LSA)

LSA 产生的最初目的,就是将 LSA 思想运

用到信息检索中,解决基于关键字检索中遇到的同义词和多义词的问题。目前,新的基于相似概念的双概率模型为潜在语义索引(Latent Semantic Indexing, LSI)提供了更深层次的理解。该模型高峰期的可能性曲线表明了内在语义层面的存在,而遵循齐夫分布的潜在语义索引的重要性表明了 LSI 层面代表着潜在的概念,可以将语义层面含有的噪音信息分离出来,通过统计显著性和可能性来定量揭示语义关联^[8]。而 Miles Efron 则在潜在语义索引中以修正平行分析(Amended Parallel Analysis, APA)作为一种模型选择的新方法用于信息检索等无监督学习的问题。经修正后的平行分析是合恩平行分析的升华,运行于空白特征的信任区间,主张保留比预期的独立性更大的非假设检验参数的相关矩阵特征值^[9]。用词性标注加强潜在语义分析也是文本分类的一个有效算法,在潜在语义分析中使用词性标注器过滤奇异值,可以降低词频矩阵的纬度,将文档在更少、更能表示其特征的语义空间表示出来。潜在语义分析可结合上下文的无关文法从自然语言中推断语义。

1.3.3 多媒体描述(Multimedia Description)

(1) 视频领域。MPEG-7 是国际标准化组织新推出的支持结构化和语义的多媒体描述的标准。在针对 MPEG-7 标准的研究中,研究者采用智能多媒体知识应用(Intelligent Multimedia Knowledge Application, IMKA)构建媒体网络框架,把多媒体作为语义和知觉信息的代表进行表述和处理。研究成果表明,多媒体数据的附加注释可以自动构建媒体网络知识基础,并且用于提高多媒体的检索^[10]。元数据被用于描述多媒体是用来标识一个项目的单纯的物理特性,甚至标识被检查项目本身不能直接确定的属性。这些具有意义的抽象的、解释性的属性,往往是人们搜索的对象。

(2) 音频领域。在音频领域,研究者建立了评价区分情感、结构和动觉等高级别音乐品质描述的模型。提出基于主体特征的因素分析法,并认为主体的性别、年龄、音乐专业知识、音乐修养、音乐品味和音乐熟悉程度对音乐的语义描述具有一定影响力^[11]。

(3) 图形/图像领域。在图形/图像领域,首先利用预先测试的语义属性提取序号。在保证最终提取序号的独特性和作为输入的图像序号的语义依赖的前提下,用于验证发行图片或检测部分修改的可靠程度,并使用更好的已知散列函数数字签名方案,以及产生的字符串来检查不明图片和原始图片的真实性和相似性^[12]。

1.3.4 数据挖掘(Data Mining)

在数据挖掘方面,针对词汇和语义 Web 内容,一个网页往往和类似的网页相链接,并且链接集群将推测到的有关同一主题的网页都聚集在一起,并通过定量研究确认最新 Web 挖掘技术和网络小世界拓扑结构之间联系的成功。同时通过探索非监督学习技术提取语义信息,在段落检索模式中利用语篇中的语义改善文献检索。新的段落检索模型是基于无指导的图形模式,可实现其空间检索模型在主题、概念和术语的分布上方便高效的聚集,在语义表象的多个层面改善了检索结果^[13]。

1.4 数据挖掘应用的研究

数据挖掘从其诞生之日起就开始应用于信息资源管理领域,随着其算法的不断改进,数据挖掘在信息资源管理领域中的应用也越来越广泛,无论是文本挖掘、数据库挖掘还是基于互联网的 Web 挖掘,数据挖掘应用已成为研究的热点之一。

首先,在文本挖掘方面。近年来数据挖掘主要应用于文本自动综述、搜索引擎的发展、文档聚类、文件分类、词库建设以及浏览界面等。其目标是在适当的文本集合中帮助自动发现一组有意义的模式或假设,在发现框架基础上用元数据配置文件的主题构建算法,成功地解决关键术语之间的关系和排名问题。Yi-fang Brook Wu 等通过关键短语识别程序(Keyphrase Identification Program, KIP)利用人们已事先识别确定短语的样本配置候选短语的权重来提取文档的关键短语。KIP 的学习功能可以通过自动添加新的关键短语标识丰富术语表,其个性化功能还可以帮助用户建立满足自身兴趣领域的术语表^[14]。此外, A Durfe 等提出了一种新的文

本挖掘方法——原型匹配文本聚类(Prototype-matching for Text Clustering),该方法应用于基于摘要的文档聚类、财务数据分析、作者身份鉴别、相似译版评价等方面^[15]。

其次,在数据挖掘方面。信息资源管理作为数据挖掘应用最早的领域,数据挖掘在这一领域中的应用也是发展最成熟的。以生物医学数据库为例,Andrés Rodríguez 等在数据库中,通过遍历搜索和自动模式识别策略确定信号,利用关联规则使标签序列排列保持相应信号功能的关键字,从而设计出一套发现关联规则的算法。该规则提取算法应用的有效性已经被其在相关领域中发现未知知识和有用知识的能力所证实^[16]。Jiexun Li 等则建议使用基于内核的学习方法(Kernel-based Learning Method)从专业语料库文献中自动提取所需的关系,通过纳入跟踪内核捕获丰富的背景信息来修改标准树内核功能^[17]。

第三,在 Web 挖掘方面。近年来,Web 挖掘一直引导着数据挖掘应用的潮流,关于 Web 挖掘应用的相关成果也是最多的。

(1)Web 结构挖掘。通过用数据挖掘调查主页关系,系统确定网页导航、相对位置和共同项目的特征,将主页连同其支持的网页视为一个 Web 单元,通过从 Web 单元中提取网页间的特征从而实现更好的 Web 精度挖掘^[18]。

(2)Web 用法挖掘。利用语言分析技术借助齐夫分布等统计方法对网站的查询结构、词汇和词间关联进行分析,通过数据存储和迭代分析建立数据模型粒度,从而确定查询行为的模式和趋势。另一种新的网络日志分析方法是借助识别“查询转化”序列的相似性,将“查询转化”序列描绘成图形化网络,通过对转化代码相关性的观察分析说明行为为习惯,进而对检索行为进行序化描述^[19]。

(3)Web 内容挖掘。基于学习的数据清洗算法主要用于筛选目标网页,通过独立查询功能和清洗算法,可以集中少量高价值的网页来满足大部分用户需求,帮助搜索引擎更好地利用有限的存储和计算资源,改善搜索效果。此外,使用增强模式的相关规则将一组查询日志中的相关查询用于初步查询,用户可以利用相

关查询和基于关联度的排序调整或重新定向搜索过程。在搜索引擎的帮助下,新的挖掘 Web 数据的分割算法在检索中可以有效地处理分割模糊、未知词检测、截止词等问题,在精度上优于传统算法^[20]。

1.5 以计算机为交流中介的研究

随着人类科技的进步与创新,知识的积累与传播,人与人(组织与组织)之间的交流越来越被人们所重视。如何将现有的信息和知识有序化,以及将交流机制和模式智能化,使之有利于交流与创新,成为摆在信息工作者面前的重要问题。目前,计算机作为中介作用于人类交流特别是知识交流在信息资源管理研究中越来越引起人们的重视。以计算机为中介(Computer-Mediated Communication, CMC)使得人类知识的交流具有瞬时性、广泛性、准确性和便利性的特征,促进了科技进步。

算法与理论方面,在互动一致性分析(Interactional coherence analysis, ICA)的基础上产生了混合互动一致性(Hybrid Interactional Coherence, HIC)算法,可以更准确地标识和构建以计算机为中介的交流。该算法采用了一组系统和语言功能,包括邮件标题信息、报价、直接地址、词汇关系,以及基于相似性的词汇匹配算法(Lexical Match Algorithm, LMA)等。混合互动一致性算法比较明显地在精度、回叫、论坛和线程测量等方面表现出优势^[21]。为了解决电子媒介与面对面媒介相比对协作型任务的沟通造成的妨碍,产生了补偿适应理论^[22](Compensatory Adaptation Theory)。这一理论把重点放在认知效果、交流模糊性、信息准备、流畅程度和任务效果质量上。

在实践应用方面,团队结构和通信模式被认为会影响虚拟团队的进程演变,包括团队导向、分担工作量、寻求和交换信息的倾向等。这些团队的进程将决定小组成员的信息交流模式,进而影响团队效率和进程的满意程度。Jung-ran Park 认为,语言礼貌现象是人类沟通的关键组成部分,可以很好地为分析社会交往和人际之间的话语与交流参与者变量提供一个框架,因为它不仅适用于面对面的社会交往,而且

适用于在线交流的互动。通过定性分析得出,积极的礼貌表现和直接的语言行为策略都是立足于密切参与者之间人际关系的本质,这类礼貌策略表明,有效和充分实现人际沟通在在线社会互动方面发挥着重要作用^[23]。

信息与知识的激增对人类传统的交流模式提出了挑战,互联网迅速普及促使知识交流进入一个高速、高效、高容的全新阶段。尽管专门针对以计算机为中介交流的研究目前不是很多,但由于人类社会信息与知识交流的普遍性和多样性,针对这一领域的研究必将会成为信息资源管理领域中新的热点。

2 信息资源管理中人文价值与社会经济领域的研究热点

2.1 知识产权保护与知识共享

在知识产权保护(Intellectual Property Protection)领域,无论是从技术上还是从法律规范上,早期研究已经取得了一定的成果,数字水印(Digital Watermark)技术和各国知识产权保护法的不断完善就充分证明了这一点。美国专利商标局在2002年发布了《21世纪战略纲要》,其战略目标是建立保持美国发明人在全球的竞争优势所需要的专利商标制度,将专利商标局发展成一个以质量为核心,对市场变化反应灵敏的市场驱动型知识产权机构。同年,日本也发布了《知识产权战略纲要》,提出“知识产权立国”的口号,在日本全社会推行知识产权保护制度。近年来,理论界关于知识产权保护的研究有以下两方面的成果。一方面,研究者通过对欧洲专利局和几大信息提供商的比较和分析,认为对于一些特殊功能的需求,受到知识产权保护的专利信息可以实现排序、制图和信息流的可视化,在书目专利耦合、共同专利引文、专利转让、技术专门化等方面可以进一步改善^[24]。另一方面,也有人注意到数字技术的不断进步对以传统材料为基础的媒体(书籍和杂志等)形成了冲击,因此对知识产权保护提出了新的要求^[25]。

对知识共享(Knowledge Sharing)的研究近年来取得了较丰富的成果。Lawrence M Oliva 认

为,信息共享的问题主要是非技术性的问题,目前信息共享的非技术性壁垒主要包括:①法律限制;②行政权限;③安全分类;④经济问题;⑤信息过时;⑥背景。在此基础上,Oliva提出了通过调和与协商政府、商界和学术界信息共享标准消除信息共享壁垒的建议,具体包括:①信息共享标准;②法律保护;③经济补偿;④背景标记;⑤信息退出机制^[26]。

在知识共享理论研究方面,有学者认为,知识共享的倡导和壁垒问题主要存在于个体之中,倡导者往往以集体主义和互惠为出发点,而壁垒者的动机则不尽相同。对于知识管理的知识共享而言,人类个体的执行方式并不总是与他们此方面的信仰相一致,这种矛盾被解释为自我效能(Self-efficacy)。基于知识管理的知识共享已经超越了知识共享的实际行为和意志控制结构,管理必须着眼于组织设置的社会和文化属性,从而加强人们在知识共享实践中的意志控制^[27]。

知识共享的实证研究集中在各类企业经营活动领域。从企业供应链管理(SCM)角度对信息共享效力进行仿真研究得出,分销商和批发商在最终库存和后备订货量等方面从信息共享中获得巨大收益。在企业资源规划(ERP)方面,对默许知识共享(Tacit Knowledge Sharing)进行的研究表明,通过对ERP执行过程中默许知识共享的要素进行识别和分类,有助于探索默许知识共享发展的方向和途径。目前,在基于资源的理论基础上已经有研究者建立起知识共享模型,随着软信息资源的补充,信息共享和自发学习将成为企业智力资本。企业必须同时关注在作为知识共享基础的硬信息资源和智力资本等软信息资源上的投入^[28]。此外,W A Taylor 和 G H Wright 还率先对公共服务领域的知识共享进行了研究,通过要素分析和回归模型识别出影响有效知识共享的独立因素,即从失败中学习的能力和好的信息质量有效地预测成功的知识共享^[29]。

2.2 用户信息行为分析和用户满意度

在用户信息行为分析方面,国外学者在研究中超越了信息资源管理学科本身的范畴,从社会学、心理学、组织行为学等多个学科角度和层面对用户的信息行为进行研究。研究所采样

的数据包括社会学数据(角色、规范、任务)、心理学数据(知识状态、认知状态、作风状态、情感状态)、信息来源数据(来源的性质、来源的行为)和行为数据。强调信息行为的非线性、动态性和层次性,并把互动与反馈纳入了信息行为的研究中,研究重点在于应用,而不是单纯获取技术。Amanda Spink 和 Charles Cole 依据进化心理学的理论,将人类的寻求信息的行为分为三个层次:①日常生活信息寻求常识的决策方法;②信息搜索方式;③信息寻求观念的解决方案^[30]。正是基于人类信息行为的非线性、动态性和层次性,作为信息活动主体的人的个体差异也凸现出来,许多研究者纷纷致力于探究信息搜索策略与人类个体差异之间的关系,采用回归分析为主要技术手段的因素分析法。而基于人类个体差异强调用户个体偏好的个性化信息搜索模型也随之诞生。Chen Ding 和 Jagdish C Patra 在自组织映射(Self-Organizing Map, SOM)的基础上建立了个性化信息搜索模型,从而使使得个性化信息搜索更加有效^[31]。

用户满意度研究方面,较早的研究在本质上仍然没以产品(系统)为核心。迪洛涅和麦考林模型(DeLone and McLean's model)研究的焦点集中在以下四层关系:系统质量的感知和用户满意度、信息质量感知和用户满意度、用户满意度和使用意向、用户满意度和个人影响感知。模型中对质量的高度强调,暴露出其根本出发点仍然是产品而不是用户。随着学术界和产业界普遍接受以用户为核心的指导思想,感知易用性(Perceived Ease of Use)对用户满意度影响的权重越来越为人们所认识。感知易用性是强调以用户为核心,提倡用户参与、互动和控制,其决定因素包括个体的差异、系统的特点以及抵制变革(Resistance to Change, RTC)^[32],从易用的角度来考虑用户满意度,真正实现了以用户为核心的指导思想。

2.3 跨语言与文化差异

人类的文化是多元性的,以互联网迅猛发展为特征的信息时代要求信息交流遵从便捷、高效、畅通的原则,而文化的多元性造成的语言、文字交流障碍显然有悖于这一原则。另一

方面,面对互联网的英语化(Internet 上 90% 的内容为英语),许多国家采取一定的措施来保护本国、本民族的文化不被吞噬。这些问题和矛盾促使国际信息资源管理领域中关于跨语言和文化差异的相关研究日渐兴盛。

国外信息资源管理领域中跨语言和文化差异的研究主要集中在跨语言信息检索(Cross-Language Information Retrieval, CLIR)方面。针对跨语言信息检索中超出词典和词汇不匹配的问题,目前主要采取建立检索中介和借助平行语料库两种方式。

(1) 建立检索中介的方法是在跨语言信息检索过程中先建立一个检索中介以弥补功能性和交互性的不足。这个检索中介可以是一个中继性的词汇知识库(Lexical Knowledge Base, LKB),也可以是类似词库的一个初级检索结果集。根据自然语言处理(Natural Language Processing, NLP)的策略建造词汇知识库或传递字典,定制可用的翻译资源用于翻译查询。也可以依据有效的标准向量空间模型算法,通过自动分类程序查询术语和相关专业概念,利用专业概念建立一个语言抽象,使信息的检索以独立的方式进行,以使语言歧义最小化。还可以采用基于 Web 的术语翻译方法,借助一个真正的网络搜索引擎对双语结果页面进行挖掘,从而确定未知查询项的有效翻译^[33]。

(2) 利用平行语料库技术进行跨语言信息检索时,更强调检索内容的相似性。在掌握了海量的所涉及语言已严格标注的语料基础上,根据相似性原则提取恰当的含义。再根据翻译规则和相似矩阵,从平行语料库中自动学习。最后根据不同的相似性提取适当的词汇,提高检索精度。对于存疑之处采用类似字符串匹配技术,参考相似之处的上下文关系,排除不适合含义^[34]。

此外,文化的多样性和文化差异问题在学术界已是一个不可回避的问题。人们一方面强调网络和计算机技术越是高度发达,就越应该更多地考虑系统主体的文化差异问题。另一方面,通过研究证实,团队中文化的多样性是一把双刃剑。这也决定了文化差异不是简单的是否应该被消除或如何消除,而是如何既保留文化

多样性又协调了文化差异的问题。信息社会应立足于并促进对文化特征、文化和语言的多样性、传统以及信仰的尊重,推动不同文化和文明间的对话。

2.4 信息道德与隐私保护

网络社会是由于计算机网络的兴起而出现的包括信息、知识、情感的一种崭新的生存环境。在这种全新的生存环境中,各种异质的价值观念、生活方式通过多种信息传播渠道相互交汇、碰撞,人们不得不面对道德伦理关系的挑战。专业的信息工作人员需要掌握其工作的道德尺度,而即使缺乏信息素养的普通参与者(生产者、提供者、传播者、汇集者)也应该努力减少未验证信息。近年来,对于如何防范和识别违反道德甚至犯罪的信息行为的研究已取得了一些进展:①技术上,采用共生分析(Co-occurrence Analysis)、最短路径算法(Shortest Path Algorithm)、启发式处理方法、可视化技术、灌溉算法(Flooding Algorithm)等建立的一系列信息系统可以有效地克服信息过载和搜索的复杂性,帮助进行犯罪调查^[35];②制度上,研究者发现借助威慑理论(Deterrence Theory)和计划行为理论(Theory of Planned Behavior)建立的惩罚和道德培训机制可以有效地减轻潜在的软件和信息安全的威胁^[36]。

另外,我们也应该看到,在信息道德问题上,人们既认可依据道德准则和伦理原则的标准进行判断,同时也强调信息道德的主观性,认为现代信息社会中道德和不道德行为之间的界线在不断变化,对网站所有者的潜在影响是习惯于支持新的道德规范,人们的判断往往可能随时间而改变。互联网已经改变了人们对隐私的传统认识,但由于许多与隐私有关的法律问题还没有解决,研究人员、决策者、使用者都开始担心网上的隐私,隐私问题在互联网上已经越来越多的被人们所关注。目前,学术界关于隐私问题的研究成果主要集中在隐私关注问题(Privacy Concern)、隐私保护制度和隐私保护技术三个方面。

(1)隐私关注问题。研究者通过多种方式对互联网隐私问题进行了测度,在此基础上

Mike Z Yao等依据二元性、分层回归和结构方程建立模型,得出隐私的心理需要和隐私的信仰与权利是影响对在线隐私关注的主要因素^[37]。

(2)隐私保护制度。有学者从政府和社会的层面提出采用类似数字版权的管理规范建立扩展隐私框架,用于管理个人资料的隐私。从企业组织的层面,通过了解隐私权威威胁的来源、确定隐私防线和识别工具及隐私防范程序、监测这一提案的效力^[38]三个步骤来监督隐私问题和确保符合相关法律。

(3)隐私保护技术。企业的商务活动离不开收集和交换数据的信息行为,因此如何既满足企业商务活动的的数据需求,又为企业间的数据交换提供隐私保护是十分重要的。针对这一问题,最新的隐私保护技术方面的研究成果表明,分布式数据的二元评级置顶(Binary Ratings-Based Top-N)方案可以被用于在保护数据所有者隐私的前提下提供准确的推介,为企业数据隐私保护提供了一种可行方案^[39]。此外,也有学者认为隐私保护的根本在于立法而不在于技术。鉴于目前国际上隐私问题立法的多样性以及人们对隐私问题认识的复杂性,人们在隐私保护问题上的探索和研究还会继续下去。

2.5 青少年信息安全

随着计算机和网络技术的发展,互联网越来越成为青少年学习、娱乐和社会交往的重要平台。网络给孩子们的影响是双向的。网络在为青少年带来自己动手观念、多元化知识、反权威、独立性、探索性、开放性与合作性等诸多利益的同时也带来了某些有碍他们健康成长的危险因素。根据国际儿童网络(Childnet International)的看法,目前儿童在网络中遇到的危险主要有以下三类^[40]:①联系性危险(Contact Danger),指通过网络聊天和电子邮件等交互方式,某些人可以联系上儿童,从而引诱或者伤害儿童。②内容性危险(Content Danger),指儿童接触到不适宜的内容,这些包括了有关色情、种族歧视和种族仇恨、暴力等文字、图片和影像资料。③商业性危险(Commercialism Danger),指儿童隐私受到市场商业行为的侵犯。国际电信

联盟 (ITU) 指出,青少年是互联网最积极同时也是处境最脆弱的用户群之一,因此必须加强对上网青少年的保护。国际电信联盟的秘书长 Hamadoun Touré 博士于 2009 年 2 月 17 日宣布,2009 年世界电信和信息社会日 (5 月 17 日) 的主题为“保障儿童网上安全”。

在理论研究方面,最新的研究成果表明,不同的社会经济、种族、文化和地理界限背景下的青少年具有相似的信息需求,互联网并没有影响青少年是否访问公共图书馆,只是影响他们访问的频率。使用互联网降低了青少年把公共图书馆作为个人信息来源的需求,但并没有影响公共图书馆在他们的学业和娱乐方面的作用,随着年龄的增长使用互联网的需求逐渐上升。互联网的优势在于能够提供更详细的信息,但由于互联网从业人员都是成年人,青少年使用互联网在检索目标定位上却不尽人意。尽管传统的认知理论认为层级目录导航更适合少年的信息检索,但是青少年更喜欢按关键字进行检索,原因是目录中的术语往往由成人设计。如果在逻辑检索和休闲浏览的界面中提供多层次的、连续的层级导航,发展与其年龄相适应的受控词汇表,另外不仅在界面设计中而且在内容和结构设计中考虑青少年因素,就能够更好地理解其检索和浏览的技能和偏好^[41]。在具体措施上,针对青少年网络安全进行专题教育,增加信息安全常识,这种教育可能会成为青少年学习社会信息学和各种社会信息处理技能的主要因素。另外,针对《儿童互联网保护法案》(CIPA) 规定的信息过滤为图书馆带来的一系列实际问题,有学者提出必要信息研究这样一个视角,从而更好地理解 CIPA 的冲击^[42]。

在实际应用方面,美国的公共图书馆都已安装了信息过滤装置,像美国在线 (AOL) 和家庭点击 (FamilyClick) 公司还提高了商业化网络内容过滤服务的标准,在精确性和效率性方面有较大的提高。YahooLigans、AJKIDS 等专为青少年服务的搜索引擎应运而生,Hello Kitty 的美国网站 Sanriotown 以及 Disney.com 等面向少年儿童网站都增加了相应的控制功能,努力避

免和降低自己面临的法律风险。

3 总结与展望

通过以上对国外信息资源管理领域近年来研究成果的归纳和梳理,在理论探索与技术应用层面可以得出以下结论:

本体理论、形式概念分析 (FCA)、语义学、数据挖掘、以计算机为中介的交流 (CMC),这五个方面都存在着不同程度的相互交叉和渗透。本体理论、形式概念分析和语义学结合的产物——潜在语义分析,广泛应用于 Web 语义检索和对话管理中,必将成为未来信息检索领域的研究重点。本体理论和形式概念分析同时也对数据挖掘和知识管理产生了新的影响,而数据挖掘也通过语义提取发展出更新颖的算法并被更多地应用于改善信息检索的结果,未来的数据挖掘显然会与本体理论和形式概念分析更深入地结合。本体理论和语义学与以计算机为中介的的结合使基于本体的潜在语义分析在以计算机为中介的交流方面实现有效的对话管理,这为日后进一步在计算机为中介的交流领域展开研究奠定了语义学层面的基础。可以预见,未来信息资源管理领域在理论探索与技术应用层面的宏观趋势必将是以跨学科、广理论、多技术的交叉融合为特点,各项理论与技术在各自的发展过程中不断地从其他理论与技术中汲取营养,然后在完善和发展自身体系的同时,再回馈到其他相关理论与技术的发展中,交叉融合,相辅相成。

在人文价值与社会经济层面,虽然仍旧是以那些亘古不变的话题为主,但研究内涵和研究重心都已悄然发生转移。知识产权保护与知识共享的研究此消彼长,研究的重心已经转向了知识共享领域,知识共享也必将代表未来知识管理的发展趋势。在用户信息行为和用户满意度研究中,无论是个性化信息检索还是感知易用性都强调人类个体的思想、行为、需求的个性化发展,以用户为核心的理念必将成为指导未来这一领域研究和应用的核心指导思想。而跨语言与文化差异方面的研究将继续深入下

去,只是未来的研究重点是在保留现有的语言和文化多样性的基础上实现人类异类语言和文化间的无障碍信息沟通和知识交流。信息道德与隐私保护问题在未来仍将是一个争论的焦点,人类科技的发展总是不断地触动人类的道德神经,关于道德层面的研究必将是一个持久的话题。青少年的发展关系到人类的未来,但目前无论是学术界还是产业界尚不能从信息需求或信息组织方面找到一个一劳永逸的妥善解决办法,在未来的一个时期内,关于青少年信息安全问题还需要继续研究和探索。

随着人类科学技术的不断发展和进步,人们观念的不断更新和提高,信息资源管理领域中的前沿和热点不会是一成不变的。而且,本文所列举的10个方面也只是信息资源管理学科诸多分支中的一小部分,并不足以反映国际信息资源管理领域研究成果的全貌。假以时日,文中所列的10个方面或者不断发展进步乃至成为信息资源管理领域中的显学,亦或衰退、没落甚至被替代或消亡。

参考文献:

- [1] Fonseca F. The double role of ontologies in information science research[J]. *Journal of the American Society for Information Science and Technology*, 2007, 58 (6):786-793.
- [2] Kaye A, Hirzallah N, Shalabi L A, et al. Building ontological relationships: A new approach[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (11):1801-1809.
- [3] Kabel S, De Hoog R, Wielinga B J, et al. The added value of task and ontology-based markup for information retrieval[J]. *Journal of the American Society for Information Science and Technology*, 2004, 55 (4):348-362.
- [4] Van der Eijk C C, Van Mulligen E M, Kors J A, et al. Constructing an associative concept space for literature-based discovery[J]. *Journal of the American Society for Information Science and Technology*, 2004, 55 (5):436-444.
- [5] Garces P J, Olivas J A, Romero F P. Concept-matching IR systems versus word-matching information retrieval systems: Considering fuzzy interrelations for indexing Web pages[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57 (4):564-576.
- [6] Mora M, Gelman O, Forgionne G, et al. Integrating the fragmented pieces of IS research paradigms and frameworks: A systems approach[J]. *Information Resources Management Journal*, 2007, 20 (2):1-22.
- [7] Niemi T, Jämsen J. A query language for discovering semantic associations, Part I: Approach and formal definition of query primitives[J]. *Journal of the American Society for Information Science and Technology*, 2007, 58 (11):1559-1568.
- [8] Ding C H Q. A probabilistic model for Latent Semantic Indexing[J]. *Journal of the American Society for Information Science and Technology*, 2005, 56 (6):597-608.
- [9] Efron M. Eigenvalue-based model selection during latent semantic indexing[J]. *Journal of the American Society for Information Science and Technology*, 2005, 56 (9):969-988.
- [10] Benitez A B, Zhong D, Chang S F. Enabling MPEG-7 structural and semantic descriptions in retrieval applications[J]. *Journal of the American Society for Information Science and Technology*, 2007, 58 (9):1377-1380.
- [11] Lesaffre M, De Voogdt L, Leman M, et al. How potential users of music search and retrieval systems describe the semantic quality of music[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (5):695-707.
- [12] Poulos M, Bokus G, Vaioulis F. Towards the semantic extraction of digital signatures for librarian image-identification purposes[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (5):708-718.
- [13] Menczer F. Lexical and semantic clustering by Web links[J]. *Journal of the American Society for Information Science and Technology*, 2004, 55 (14):1261-1269.
- [14] Wu Y F B, Li Q Z, Bot R S, et al. Finding nuggets in documents: A machine learning approach

- [J]. *Journal of the American Society for Information Science and Technology*, 2006, 57 (6):740-752.
- [15] Durfee A, Visa A, Vanharanta H, et al. Mining text with the prototype-matching method[J]. *Information Resources Management Journal*, 2007, 20 (3):19-31.
- [16] Rodríguez A, Carazo J M, Trelles O. Mining association rules from biological databases[J]. *Journal of the American Society for Information Science and Technology*, 2005, 56 (5):493-504.
- [17] Li J X, Zhang Z, Li X, et al. Kernel-based learning for biomedical relation extraction[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (5):756-769.
- [18] Sun A X, Lim E P. Web Unit-Based Mining of Homepage Relationships[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57 (3):394-407.
- [19] Wang P L, Berry M W, Yang Y H. Mining longitudinal web queries: Trends and patterns[J]. *Journal of the American Society for Information Science and Technology*, 2003, 54 (8):743-758.
- [20] Liu Y Q, Zhang M, Cen R W, et al. Data cleansing for Web information retrieval using query independent features[J]. *Journal of the American Society for Information Science and Technology*, 2007, 58 (12):1884-1898.
- [21] Fu T J, Abbasi A, Chen H C. A hybrid approach to Web forum interactional coherence analysis[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (8):1195-1209.
- [22] Kock N. Compensatory adaptation to media obstacles: An experimental study of process redesign dyads [J]. *Information Resources Management Journal*, 2005, 18 (2):41-67.
- [23] Park J R. Linguistic politeness and face-work in computer-mediated communication, Part 1: A theoretical framework [J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (13):2051-2059.
- [24] Stock M, Stock W G. Intellectual property information: A comparative analysis of main information providers[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57 (13):1794-1803.
- [25] Himma K E. The justification of intellectual property: Contemporary philosophical disputes [J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (7):1143-1161.
- [26] Oliva L M. 21st century challenges to sharing information[J]. *Information Resources Management Journal*, 2005, 18 (4):i-iv.
- [27] Kuo F Y, Young M L. A study of the intention-action gap in knowledge sharing practices[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (8):1224-1237.
- [28] Widén-Wulff G, Suomi R. Utilization of information resources for business success: The knowledge sharing model[J]. *Information Resources Management Journal*, 2007, 20 (1):46-67.
- [29] Taylor W A, Wright G H. Organizational readiness for successful knowledge sharing: Challenges for public sector managers [J]. *Information Resources Management Journal*, 2004, 17 (2):22-37.
- [30] Spink A, Cole C. Human information behavior: Integrating diverse approaches and information use[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57 (1):25-35.
- [31] Ding C, Patra J C. User modeling for personalized Web search with self-organizing map[J]. *Journal of the American Society for Information Science and Technology*, 2007, 58 (4):494-507.
- [32] Nov O, Ye C. Users' personality and perceived ease of use of digital libraries: The case for resistance to change[J]. *Journal of the American Society for Information Science and Technology*, 2008, 59 (5):845-851.
- [33] Wang J H, Teng J W, Lu W H, et al. Exploiting the Web as the multilingual corpus for unknown query translation[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57 (5):660-670.
- [34] Bellaachia A, Amor-Tijani G. Proper nouns in English-Arabic cross language information retrieval [J]. *Journal of the American Society for Information Science and Technology*, 2008, 52 (12):1925-1932.

《中国图书馆学报》英文年刊创刊发行

《中国图书馆学报》英文年刊借着中国国家图书馆百年华诞的喜庆诞生了!

随着中国经济建设和文化事业的迅猛发展,信息技术的不断渗透,中国图书馆事业发生了前所未有的深刻变化,理念的更新、技术的应用、建设的规模、发展的速度等等,逐渐引起世界范围的广泛关注,中国图书馆事业成为全球图书馆事业发展中越来越不可被忽视的一个重要组成部分。为了便于世界及时了解中国图书馆事业的发展状况和中国图书馆学研究的动向,《中国图书馆学报》英文年刊在《中国图书馆学报》中文版双月刊的基础上创刊了。

《中国图书馆学报》英文年刊以发表图书馆学及其密切相关领域的高层次学术成果为己任,以理论性、前沿性及深度的实践研究为特色。她关注国际图书馆学研究和图书馆事业的发展趋势,侧重于对中国图书馆事业的研究,反映中国图书馆学的最高学术研究成果,旨在向同行提供一个国际化的学术交流平台。

《中国图书馆学报年刊(英文版):2009》已于2009年9月由国家图书馆出版社出版发行,欢迎大家订购。

订购地址:北京西城区文津街7号国家图书馆出版社

邮编:100034 E-mail:btsfxb@nlc.gov.cn(邮购) 电话:010-66139745 66151313 66175620 66126153



- [35] Schroeder J, Xu J, Chen H, et al. Automated criminal link analysis based on domain knowledge[J]. Journal of the American Society for Information Science and Technology, 2007, 58(6):842-855.
- [36] Workman M, Gathegi J. Punishment and ethics deterrents: A study of insider security contravention[J]. Journal of the American Society for Information Science and Technology, 2007, 58(2):212-222.
- [37] Yao M Z, Rice R E, Wallis K. Predicting user concerns about online privacy[J]. Journal of the American Society for Information Science and Technology, 2007, 58(5):710-722.
- [38] Sipior J C. Ethically responsible organizational privacy protection[J]. Information Resources Management Journal, 2007, 20(3):i-iii.
- [39] Polat H, Du W L. Privacy-preserving top-N recommendation on distributed data[J]. Journal of the American Society for Information Science and Technology, 2008, 59(7):1093-1108.
- [40] Childnet International. E_safety: An introduction for trainee teachers[OL]. [2009-02-25]. <http://www.childnet-int.org/downloads/ttleaflet.pdf>.
- [41] Bar-Ilan J, Belous Y. Children as architects of Web directories: An exploratory study[J]. Journal of the American Society for Information Science and Technology, 2007, 58(6):895-907.
- [42] Jaeger P T, Bertot J C, McClure C R. The effects of the Children's Internet Protection Act (CIPA) in public libraries and its implications for research: A statistical, policy, and legal analysis[J]. Journal of the American Society for Information Science and Technology, 2004, 55(3):1131-1139.

毕强 吉林大学南岭校区管理学院信息管理学系教授,博士生导师。通讯地址:吉林大学管理学院。邮编130022。

滕广青 吉林大学南岭校区管理学院信息管理学系博士研究生。通讯地址同上。

(收稿日期:2009-03-11;
最后修回日期:2009-03-30)