

●张 红

面向知识创新的信息组织^{*}

摘 要 在知识创新活动中,信息组织的作用极为重要。面向知识创新的信息组织应从信息的遴选、信息或知识单元的组织、信息检索系统的组织等方面入手。参考文献9。

关键词 知识创新 信息组织 因特网 信息检索系统

分类号 G254

ABSTRACT Information organization is playing a very important role in the activities of knowledge innovations. We should begin the information organization oriented to knowledge innovations with the selection of information, the organization of information or knowledge units and the organization of information retrieval systems. 9 refs.

KEY WORDS Knowledge innovation. Information organization. Internet. Information retrieval system.

CLASS NUMBER G254

1 知识创新与信息组织

知识创新可以分为广义和狭义两种。狭义的知识创新是指通过科学研究获得新的自然科学、社会科学和技术科学知识的过程和行为。广义的知识创新是指为了社会和经济利益获取或创造新的知识的过程和行为。知识创新可以发生在知识的生产、传播和应用过程中^[1]。

实际上,无论在哪一个环节发生的知识创新活动,都和信息密切相关。信息是知识创新必不可少的原料。然而,原始信息本身并不能产生价值,无序的信息也不构成资源,只有将信息有序化组织,把具有必然内在联系的信息和没有必然内在联系的信息形成一个有关联的体系,才能产生效益。信息组织就是这样采用一定的方式,将某一方面大量的、分散的、杂乱的信息经过整理、优化,形成一个便于有效利用系统的过程^[2]。在知识创新活动中,信息组织的作用极为重要。

1.1 信息组织帮助科研人员发现与创造新的知识

知识创新必须充分借鉴和继承前人的科技成果。17世纪丹麦天文学家第谷每夜都辛勤地用仪器精细观察行星的位置,他的目标是绘制1000颗星的星表,但是用毕生的精力只完成了750颗星,不得不在弥留之际将自己未完成的研究托付给助手开普勒。开普勒对第谷堆满几房间的观察资料进行整理、分析、优化和研究,最后归纳出行星运动三大规律^[3]。实际上,开普勒正是通过信息组织去发现与创造新的知识。

1.2 信息组织避免科研人员的重复劳动

知识创新不但要善于借鉴和继承,还要避免进行重复的“知识创新”。如何避免无效劳动,保证成果的创新性,有

一个非常重要的环节,就是科技查新。查新是在信息组织的基础上,帮助科研人员了解某一领域已有的研究成果、研究进展,提供“正在进行中”的科研信息,保证研究人员的创新能力,保证研究课题的新颖性,最终保证创新成果的社会认可价值。

1.3 信息组织全程性保障知识创新的顺利进行

知识创新的全过程是以信息组织为手段,在广泛搜集信息的基础上,对相关信息进行分析研究与考证,从而产生创新思想,制定创新决策,规划创新目标的过程。例如:在知识创新的准备阶段,全面收集反映本学科或专业领域有关方面的研究成果、状况和水平的信息,发现能填补空白、具有重大意义及应用前景的课题,为知识创新提供思路;课题立项阶段,围绕项目论证需要,提供技术、经济、社会等方面的可行性信息支持知识创新的立项决策;研究进展阶段,吸收前人的研究成果,针对研究中要解决的问题提供信息技术;评审鉴定阶段,为研究人员查询与成果有关的鉴定、评审标准等资料;成果传播和转移阶段,要提供成果发表或市场需要等信息^[4]。

2 知识创新中的信息需求

信息组织是以知识创新活动的信息需求为依据的。知识创新是在世界上首次发现、发明、创造或应用某种新知识的过程,在这一过程中产生的信息需求有其自身的特点。

2.1 系统性信息需求

由于科学研究活动的累积性和连续性,决定了研究者要从继承的角度去掌握有关研究课题的系统性信息,以现有的现象、事物、知识、观点、思想为基础,可以推导出一系列新现象、新事物、新知识、新观点和新思想等,获得新的启示。

^{*} 本文系广东省教育厅人文社会科学研究项目“知识创新的信息保障体系建设”的研究成果之一。

2.2 集成性信息需求

当代科学技术的显著特点是学科统一化进程,学科整合和整体性认识的趋势越来越明显。在单学科继续深入发展的同时,向多学科研究、跨学科研究和整体性研究发展。因而,学科交叉极富创造性,学科交叉领域最富生命力^[5]。科学研究不仅需要日益深、与本专业本课题密切相关的信息,也需要日益广泛的相关学科和相关课题的信息。打破学科界限,提供融合多学科的集成性信息是知识创新的需要。

2.3 学术性信息需求

知识创新是根据国内外科学研究动态,掌握最新学科进展,进行前瞻性分析和科学预测,展望未来科技发展趋势,剖析新学科、新技术的增长点,及时调整研究方向,开拓新领域的活动。所需信息类型主要是学科性强学术价值大的原始文献,如图书、期刊、会议文献和研究报告等。研究人员也希望获得二次文献(题录、文摘、索引)和三次文献(综述、述评)。

2.4 网络化信息需求

网络技术为科学研究人员提供更加便捷的交流手段,从事知识创新的科学研究人员将更多地网络化环境中相互竞争、交流与合作。网络信息资源作为一种新型的信息资源,在知识创新中发挥着愈来愈重要的作用。

3 面向知识创新的信息组织

3.1 信息的遴选

信息的遴选是信息组织的首要任务,信息组织首先是一个遴选过程,只有有价值的信息才可能被有效组织。知识创新对信息的选择有特定要求,通过一定的组织和技术手段,为知识创新用户建立有效的信息遴选机制非常必要。无论是传统的文献信息还是网络信息都存在着价值效用明显差异的问题:既有学术前沿的会议文献,也有雅俗共赏的大众读物;既有学科进展的研究报告,也有失去时效的科技专著;既有知识重组的三次文献,也有效用不大的二次文献。良莠不齐的信息,不但不能给知识创新提供保障,反而会给知识创新制造障碍。

传统文献信息的遴选,在出版环节就有严格的审查制度;毫无参考价值的文献,一般不会有其产生的渠道。在传播环节,如图书馆的文献采集人员,会选择有价值的文献信息入藏,传统信息的开发人员,也会筛选用户最有用的信息进行加工、重组。而网络信息的遴选要困难得多。由于网络开放性的特点,网络信息的数量与质量呈反变关系,上载的信息不受任何控制,网上的信息海量,人工也无法加以选择。为了有效地过滤无用信息,可用因特网内容选择平台(the platform for internet content selection, PICS)来实现因特网信息内容的选择。这种选择方式是利用 PICS 为标准,在信息的接受者和信息源之间插入阻塞软件的方法来控制因特网的内容获取。PICS 采用类似电影分级审查制度的方法来进行内容筛选,它提供了一种价值中立的因特网标记的基本数据结构^[6]。

3.2 信息或知识单元的组织

知识创新所需信息类型主要是原始文献。目前网络信

息资源中的全文数据库是经过数字化了的信息,也称为网上一次信息。全文数据库的信息组织方式不对文献的特征进行格式化描述,而是利用自然语言揭示文献中的知识单元,按文献全文的自然状况设置检索点。这种数据库在全文检索软件的支持下可以对数据库文本中的各知识单元进行检索。如“中国学术期刊网”全文数据库就可以按篇名、作者、关键词、机构、引文、基金、中文刊名、年、期进行检索,并根据用户需要,将检索结果以不同形式输出。虽然全文数据库一次性地解决了原始文献的直接查阅问题,但严格地讲并没有达到真正的全文检索,即文献中的任意字符或字符串不能成为检索对象。而这些貌似没有检索价值的信息,很可能是隐含在信息单元或知识单元内的信息,一旦被发现和标引出来,其价值往往不可估量。全文数据库信息的组织要注重隐含信息的揭示,把存在于文献中但未被人们认识的信息发掘出来,使其成为新的信息或知识,这是信息组织研究的一个前沿性课题,也是知识创新的一条途径。

3.3 信息检索系统的组织

人类的知识创新活动对知识信息保障体系提出了更高要求,以往存储和利用知识的状态在发生变化,如何更有效地组织、控制、传递知识,提高知识的利用率,已成为当前知识信息组织研究的重点和方向。信息组织的另一个任务是在一次文献信息组织的基础上,将经过优化的二、三次文献信息组织成一个有序的整体,成为检索工具信息系统。这种信息系统的形式可以是手工使用的信息检索咨询工具,也可以是计算机化的信息数据库。信息系统的功能有检索型(二次信息)和参考型(三次信息)之分。前者如书目、索引、网络搜索引擎等,这类信息系统只提供原始信息的线索;后者如辞典、手册、百科全书、综述、述评等,这类信息系统可直接检索到具体信息。

3.3.1 网络搜索引擎

因特网涵盖了人类知识的所有领域,它不仅是一个浩瀚无穷、包罗万象的信息海洋,而且也为新观点、新概念、假设、推理以及因为内容太新或太专未被纳入正式文献交流渠道的灰色文献或边缘文献提供了传播平台^[7]。由于这些新颖的信息反映了最新的科研动态和科技水平,所以备受科研人员的青睐。而搜索引擎作为网络环境下重要的信息组织工具,自它出现之日起就引起了人们的关注,到目前为止搜索引擎是人们检索网上资源最常用的工具。但实践证明,搜索引擎检索的查准率不尽人意,往往是检出的结果太多,以致于淹没了真正有价值的信息。这种状况严重制约了知识创新对网上资源的有效利用。对网上有用的信息资源进行有序化组织,已成为网络时代最具有挑战意义的重要课题之一。

中文汉语分词技术,可以有效解决搜索引擎查准率低的问题。搜索引擎采用的是作为自然语言的关键词标引,关键词检索的前提是将检索条件分解成若干关键词,而中文不同英文,中文词与词之间没有分隔符,需要人工切分,一旦分词不正确,就会带出许多不符合检索条件的结果,那么查准率就可想而知。如果根据语料库进行总结,获得每

个词的出现概率以及词与词的关联信息,就可能有效地排除各种歧义,大幅度提高分词的准确性,从而准确地表述查询要求和文档信息,提供智能型检索^[8]。

科研工作者从事知识创新活动,利用网络信息资源时尤其注重信息内容与自己所需信息的关联程度,希望所查询出的信息主题内容充分接近所要研究的主题内容。信息检索过程中的相关性选择或者说是查准率,应成为衡量信息组织是否优化的基础标准。

3.3.2 信息综述和信息述评

当今社会信息数量庞大,即使经过优化了的二次信息,数量仍嫌过大,以至于超过了科研工作者的吸收利用能力。因此,对原始信息进行内容浓缩加工,在提供大量知识因子的基础上,对其进行分析与综合,形成新的知识关联,从而产生更高层次上的新生知识源,组成各种集约化的知识产品,成了信息组织活动的重中之重。经过知识关联重组后的三次信息(综述和述评)不仅可以使科研人员在组织化的信息系统中快捷、准确地获得所需信息,而且由于改变了知识因子间的原有联系,往往能够提供新知识,也只有再生的新信息才是知识创新的直接动因。

信息综述应对某一领域或课题一定时期的有关信息进行归纳、整理、分析和加工。综述要具有高度浓缩性、简明性、研究性特征,目的是帮助研究人员用较少的精力和较短

的时间对有关领域或课题的内容、意义、历史、现状及发展趋势有一个完整、系统、明确的了解。

信息述评必须对某一领域、研究课题的现状、动态、发展水平及影响进行全面而系统的分析、评价或提出建议。述评应具有研究性和创造性特征,要能够控制和鉴别一次信息的质量和价值,为研究人员确定研究方向,提供知识创新的合理方案^[9]。

参考文献

- 1 何传启,张凤.知识创新——竞争新焦点.北京:经济管理出版社,2001
- 2 尚克聪.信息组织论要.图书情报工作,1998(11)
- 3 范贤民.知识创新工程与研究所文献结构的调整和建设.见:中国科学院第十一次图书馆情报学科学讨论会文集(1999)
- 4 黄彬.知识创新的信息消费及其保障.图书情报工作,2001(9)
- 5 张焘.促进知识创新的科学系统.中国科学院院刊,1998(4)
- 6 倪莉.网络环境下信息组织面临的新问题.情报理论与实践,2001(4)
- 7,8 燕惠兰,桂筱丹.网络环境下信息资源的组织与检索.情报科学,2001(9)

张 红 佛山科技学院图书馆副馆长,副研究馆员。通讯地址:广东佛山。邮编 528000。(来稿时间:2003-09-18)

第三届“信息化与信息资源管理”学术研讨会 征文启事

6 为了加强信息资源管理及信息化领域的理论与应用研究,促进科技创新,推动我国信息业的快速和健康发展,武汉大学信息资源研究中心将主办第三届“信息化与信息资源管理”学术研讨会。本次研讨会将围绕教育部批准立项的第三批人文社会科学重点研究基地重大项目来进行,其主题为:“文献计量与内容分析的综合研究”、“中国企业信息竞争能力的研究与评估”。欢迎广大理论研究人员和实际工作者撰文与会。现将会议征文内容及有关事项通告如下:

一、征文内容:1. 国内外文献计量与内容分析的新进展和发展趋势;2. 文献计量与内容分析的综合原理和基础理论研究;3. 文献计量与内容分析的综合方法研究;4. 文献计量与内容分析的综合应用研究及案例分析;5. 文献计量法与内容分析法的比较研究;6. 计算机辅助内容分析法的研究与应用;7. 内容分析软件工具的开发与应用;8. 知识网络与内容分析的研究;9. 基于知识元的内容分析法的研究;10. 内容分析与科学评价的关系研究;11. 企业信息化的体系结构研究;12. 企业信息化与企业竞争能力的关系研究;13. 中国企业信息化竞争能力的理论与应用模型;14.

企业信息竞争能力评价体系的研究;15. 企业信息竞争能力的实施方法研究;16. 国内外企业信息竞争能力的实证分析与比较分析;17. 与会议主题有关的其他问题研究。

二、征文要求与截止日期:1. 论文须主题突出并未公开发表过,字数为5000~10000字左右,含300字以内的中、英文摘要。2. 论文按如下顺序排列:论文标题、作者、单位及邮编、摘要、关键词(其后列出以上各项对应的英文)、正文(小标题以数字编号,如1.1,1.1.1……)和参考文献。具体格式可参考《图书情报知识》论文。文尾请附作者详细通信地址、联系电话、传真、E-mail地址。3. 论文要求以中文WORD97或WORD2000格式存入软盘,并提供便于通信、编辑和出版的电子版本。4. 论文请寄:武汉大学信息资源研究中心查荔收(请在文稿左上角标注“研讨会论文”)。邮编:430072;电话:(027)87665541;传真:(027)87665541;电子邮箱:irc@whu.edu.cn。5. 征文截止日期:2004年9月15日;录用通知:2004年10月15日前寄出。本次研讨会录用的论文将在会前正式出版论文集。

三、研讨会时间与地点:2004年11月,武汉大学。