

知识组织体系的演进路径及相关研究的发展趋势探析^{*}

滕广青 毕 强

摘 要 在对传统的“线形”、“树形”、“盒状”、“链式”知识组织体系以及数字时代的“网状”知识组织体系进行总结和梳理的基础上,认为知识组织结构体系是按照“线形→树形→盒状→链式→网状”的演变进化路径不断“柔化”、“复杂化”的过程。知识组织体系相关研究的发展方向主要包括知识链接与知识关联的研究、数据挖掘理论和应用的相关研究以及用户研究等。图1。参考文献11。

关键词 知识组织体系 演进路径 发展趋势

分类号 G250.76

ABSTRACT This paper summarizes the continually “soften” evolution path of knowledge organization system based on the analysis of traditional “linear”, “tree”, “box”, “chain” knowledge organization system and digital-age “reticulation” knowledge organization system. It also gives an outlook of the trends of knowledge organization system, including research on knowledge linking, data mining theory and application, user study and so on. 1 fig. 11 refs.

KEY WORDS Knowledge organization system. Evolutionary path. Trends.

CLASS NUMBER G250.76

人类对知识组织的探索和追求一直伴随着人类文明史发展的整个进程。从古希腊柏拉图、亚里士多德的精英知识时代,到当今的信息知识多元化、平民化的大众知识社会;从18世纪因为按字母排序而被认为是亵渎神灵的大百科全书,到今天互联网上没有编辑和页码限制也没有顺序的百科全书,人类对知识组织的认识历经变革、不断演进。互联网和计算机的出现与普及颠覆了原有的、人们已经认可和接受了的的知识组织体系。本文在对人类知识组织体系进行梳理的基础上,力求归纳出其演进路径,探寻知识组织体系相关研究发展的未来趋势。

1 传统的知识组织体系

1.1 线形结构的知识组织体系

字母排序法是线形结构的知识组织体系的

典型代表,早在古希腊和古罗马时期,人类就开始按照字母顺序对知识进行组织^[1]。由启蒙作家狄德罗和数学家达朗贝尔主持编撰并在1751—1778年分批出版的法国《百科全书》就是按照字母顺序的线形结构对知识进行组织的。时至今日,这种线形结构的字母排序法仍然在知识组织领域占有一席之地。无论是最新版的牛津词典还是新华字典(按拼音字母排序)都仍然沿袭着这种线形的知识组织体系。线形结构在中文语境中的典型例子就是笔画排序法(尽管中国古代产生的“百家姓”并没有遵循这一规律)。在基于类纸质(包括竹简、皮革、绢丝等)知识载体的时代,在组织复杂的知识体系时,按照字母(笔画)排序的线形结构仍然是一个较为科学的选择。然而按照字母(笔画)顺序组织人类知识的做法同其他线形结构(如时间顺序等)一样,

^{*} 本文系国家自然科学基金项目“基于概念格的数字图书馆知识构建研究”(项目编号:70973044)的研究成果之一。

的确也存在着许多困扰。尽管法国《百科全书》的编撰者未必承认没有将神学排在所有知识领域的最前面是对上帝的亵渎,但是单纯以字母(笔画)顺序为基础的组织体系的随意性的确割裂了许多知识间的关联。特别是当人们试图组织那些尚不完全的、正在发展中的知识体系时,没有人能够知道应该在某个字母或者某个笔画下预留多少空间。

1.2 树形结构的知识组织体系

尽管哲学界认为树形的思维方式源自于亚里士多德,但是利用树形结构组织知识的最典型的代表应该是瑞典植物学家卡罗鲁斯·林奈(Carolus Linnaeus)在其1735年出版的《自然系统》(Systema Naturae)一书中提出的植物分类法。林奈依据植物雄蕊和雌蕊的类型、大小、数量及相互排列等特征,将植物分为24纲、116目、1000多个属和10000多个种,通过“纲—目—属—种”的类系搭建起来的植物学参天大树枝繁叶茂,并且这一结构也延伸到动物学领域。在此基础上林奈用拉丁文采用双名制命名法给植物命名。与线形结构相比,林奈的树形结构在一定程度上体现了知识自然的属性,统一了术语,促进了交流。林奈在他的树形结构中用难以界定的“复杂性”为生物划分出分明的等级,他不认为物种会进化。在1735年版的《自然系统》一书中,林奈曾经对“物种不变”观念进行了专门论述^[2]。在今天,很难想象抛开进化论来为生物分类,如果生物之间不存在进化关系,林奈的生物之树就变成了一棵仅仅说明生物间相似性的死树。特别是这种不变的等级关系限制了生物之树的进一步生长。树形结构中的知识叶片具有严谨的排他性,一片知识叶子只能生长在一个枝杈上,非此即彼,这使得树的枝杈只能按一个标准口径生长,气候、环境等其他特征均被忽略。

1.3 盒状结构的知识组织体系

盒状结构知识组织体系的典型代表是美国图书馆学专家梅尔维尔·杜威(Melville Dewey,后改名为Melvil Dewey)于1876年发表的图书馆编目体系“杜威十进制分类法(Dewey Decimal

Classification, DDC)”。美国75%的公立学校图书馆和世界上135个国家的20万家图书馆都在使用杜威十进制分类法,该分类法被誉为是世界上最上流传最广、影响最大的一部分分类法^[3]。“杜威十进制分类法”按学科主题来分类,总共以10个主要的学科(main classes)来涵盖所有的知识体系,每个大类下细分10个小类(divisions),每个小类又再分成10个更小的类别(sections)。这就像把人类的知识分装在10个大盒子中,每一个大盒子中的知识又分装在10个较小的盒子中,每一个较小的盒子中的知识再分装在10个更小的盒子中……

杜威分类体系带有明显的片面性与主观性,其在“杜威十进制分类法”中把200—299都划分给了宗教,下面的10个子类分别为:自然神学、基督教圣经、基督教神学、基督教道德和信仰、基督教仪式和地方教堂、基督教社会神学、基督教会历史和地理、基督教派别、其他宗教与比较。犹太教和伊斯兰教只是在“其他宗教与比较290”这个较小的盒子中分别得到一个更小的盒子296和297^[4],而世界三大宗教之一的佛教则被杜威装在还要小的盒子中,其地位还不如拜火教。但是杜威的体系中更大的问题还不限于此。尽管可以通过将代表主题的数字放在小数点右边来无限地增加和扩展主题数量,但是 $10 \times 10 \times 10$ 的大主题划分方式并不是知识的自然形态,知识组织体系显然不是整齐划一的“十进制”。而且知识本身是不断变化的,杜威体系中的层级结构限制了新的重要主题的产生。在平均每5—6年就要对“杜威十进制分类法”进行的维护、主题的取舍和升降级一直是困扰着管理者的重要问题。

1.4 链式结构的知识组织体系

很多传统知识组织方法体系总是尽可能地罗列现有的主题,而对于未来可能出现的主题则难以预见,印度图书馆学家阮冈纳赞在其1933年出版的著作中提出了“冒号分类法(Colon Classification)”的链式结构很好地解决了这一问题。阮冈纳赞采用冒号作为分段符号,他认为不必详尽列举所有主题,只需列出文献的主题要素,分类时再综合成所需类目^[5]。这种

链式结构与人类知识的不断创新与发展相适应,能够实现将新的主题随时插入到适当位置的需要。

阮冈纳赞的“冒号分类法”提出了五个基本的分类方面:本体、物质、动力、空间、时间。通过五个方面的组合对书籍灵活分类,而不需要事先依次填满每一个可能的位置。其链式结构中并不存在所谓的“树根”,这使得其在基于计算机的系统中空前强大^[6]。尽管链式结构较之此前的知识组织体系已经有了很大进步,并对西方编目理论产生了很大影响,但是在语义描述和体现知识间关联关系方面仍然存在薄弱环节。互联网和计算机的普及颠覆了原有的,人们已经认可和接受了的的知识组织体系,知识不再受物理载体的制约,数字时代必将有新的知识组织体系与之匹配。

2 数字时代的网状结构知识组织体系

20 世纪末期,随着计算机和互联网的普及,人类社会开始向数字时代迈进,知识的载体也由原本具有鲜明物理特征的一类纸质材料而变为没有固定形式的比特流。与此同时,知识呈现多元化趋势,知识不再只是少数社会精英的专属物品,而是社会大众的共享财富。

在任何一种传统的知识组织体系中,由于只能采取一种或有限几种人们认为最主要、最本质、最鲜明、最易用的口径特征对知识进行排序、分类等工作,导致知识组织活动中被掩盖和遗漏的信息总是要多于被揭示和显现的信息。《埃及王子》(The Prince of Egypt)(梦工厂 1998 年出品的一部动画片,该片以圣经旧约中“出埃及记”为故事蓝本)到底是应该挂在“基督教”这根大枝杈上,还是应该装入“文学作品”或“卡通片”这样的小盒子中?或者由于好莱坞当红影星方·基墨和桑德拉·洛克等为影片配音,著名歌手玛莉亚·凯莉和惠特尼·休斯顿携手为该片演唱了主题曲而归入某一明星的作品集?亦或由于其获得了奥斯卡最佳音乐奖而将其纳入“奥斯卡获奖影片”或者“音乐”的范畴?……知识的诸多缤纷的、有用的特征信息被湮没在权威、固定且强大的传统知识组织体系中。

在数字时代,由于知识的组织不再受其载体物理特性的制约,因此,可以摆脱必须编排页码的纸质图书限制,不必被刻板地排序,不必被固定地挂在严格的枝杈上或装入苛刻的盒子里,也不必被非此即彼地分类。知识排序与知识归属的特征逐渐模糊,知识关联与知识链接的特性日趋鲜明。数字时代的知识组织体系呈现出一种网状结构^[7],并且这个网状的结构不是少数知识精英和权威们缔造的,而是由社会大众共同搭建的。需要明确的是,这里所说的网状结构并非是简单的二维平面结构,而是一种多向度的、多元化的、不断生长的、发展变化的、知识间有着内在与隐含的关联关系的错综复杂的立体结构。

这种网状结构的知识组织体系打破了传统知识组织体系中一片叶子只能被挂在一个枝杈上,或者一种知识只能被装入一个盒子中的教条的组织格局。在网状结构的知识组织体系中,知识可以多向度地延展其触角,构成错综复杂的知识网络。在网络百科全书 Wikipedia 中,截至 2010 年 4 月中旬,仅英文版的词条就达到 325 万余条,编撰次数超过 38053 万次,注册用户超过 1212 万人^[8]。Wikipedia 中的“埃及王子”页面没有固定的编辑人员,也没有页码顺序,更没有所属的类别。布满页面的则是历史、地理、宗教、文艺等的词条,以及各种各样的评论和引文等指向不同向度的链接。在 Wikipedia 的用户中,没有人知道、也不需要知道、甚至不会关心这部网络百科全书究竟是怎样组织“书”中的知识的。某一个具体的页面究竟是存储于哪一个服务器或者哪一块磁盘,已经同用户检索和获取知识的行为和需求没有任何关系。Wikipedia 所做的,就是打破了传统的知识界限的束缚,将某一个知识的叶子通过链接挂在了多个枝杈之上。

Amazon 与 Wikipedia 不同的是其目的是向用户推销图书。用户检索《埃及王子》的时候,Amazon 在其“Frequently Bought Together”链接下推介给用户的《奇迹创造者》(The Miracle Maker)(讲述耶稣的故事)^[9]不是因为 Amazon 认为二者都与基督教有关,也不是因为在 Amazon 的分类表中二者都属于卡通片。Amazon 在用户检索《埃

及王子》的时候把《奇迹创造者》推介给用户,仅仅是建立在 Amazon 对既往用户购买行为的挖掘之上。这说明“Frequently Bought Together”链接下的内容不是一成不变的,而是处于动态之中的。Amazon 还鼓励用户为商品添加各种各样用户认为适合的标签,并以数字反映用户对某一标签的认可程度。《埃及王子》被用户贴上了包括动画、圣经故事、迪士尼、摩西、DVD、中东、王子、基督教、音乐、埃及、家庭影片、犹太教、十诫、出埃及记、幻想、历史、灵感等多达 54 个标签。这意味着,即使用户不知道该商品的名字,也不必担心选择错误而误入歧途。用户可以沿着任何一个可能的标签链接在网状结构中找到其目标商品。传统刻板的、权威精英式的知识组织体系在数字时代的网状结构中被彻底颠覆了。

3 知识组织体系演进的路径

人类在构建组织知识体系的过程中总是试图呈现知识最原始、最本质的特征。然而不可否认的是,任何一种知识组织体系都不同程度地也反映了人类自身的偏好。在 Web 2.0 技术普及的网络环境中,Wiki、Tag、Blog 等新技术的应用在一定程度上进一步丰富了信息资源,同时由于用户的参与也使得草根终于漫过了权威之树。数字时代知识组织体系中的人类偏好终于不再是少数权威的“真知灼见”,取而代之的则是大众纷繁复杂、色彩缤纷的“群体智慧”。

甲骨、石板、竹简、兽皮、纸张、比特,知识的载体在一步步扩充其容量的同时也在逐渐“柔化”,乃至趋近于“无形”。当然,更重要的是柔化的不仅仅是知识的载体,还包括知识组织的体系结构。在各种结构的知识组织体系不断演变进化的过程中,其路径依稀可辨(如图 1 所示)。

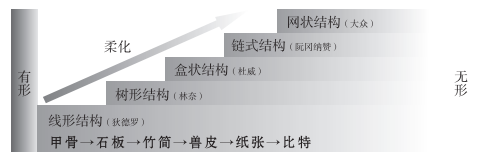


图 1 知识组织结构体系演进的路径

由图 1 看出,伴随着知识载体从“有形”到“无形”的变迁,知识组织结构体系“线形→树形→盒状→链式→网状”的演变进化路径与其说是一个不断复杂化的过程,不如说是一个不断“柔化”的过程。传统的、僵硬的知识组织体系在不断地进化为能够依据人们多样性的需求和偏好而调整的、柔化的知识组织体系。因此,网状结构的知识组织体系是一个能够因人而异的柔性化的知识组织体系。人们不必再为寻找新知识、新主题在体系中唯一确定的位置而烦恼,同时还能够以多种方式和形式呈现知识的多元化特征。

4 知识组织体系研究的发展趋势

二十多个世纪以来,人类的精英学者们前赴后继,曾经努力试图去建立一个权威的、准确的、清晰的、完整的知识组织体系,也取得了一些成就。然而回过头来审视这些既往的成就和走过的道路,人们的视线似乎变得模糊了。从字母顺序的大百科全书到按照主题分类的百科全书,人们似乎在实现着知识序化的理想;但是没有编辑和页码限制也没有顺序和属类制约的网络百科全书的出现和兴盛,似乎又让人们觉得构建一个权威、准确、清晰、完整的知识组织体系的初衷几乎遥不可及。知识精英一夜之间丧失了统治知识的权威,未来的知识组织体系又将如何发展?

目前来看,研究知识的相关问题总是难免要涉及到“内隐知识”和“外显知识”的理论^[10]。多年以来,人们的研究焦点主要集中于“外显知识”,而对“内隐知识”则讳莫如深。尽管有些知识根本就无法彻底显性化,但随着柔性化的知识组织体系的构建和对知识体系研究的不断深入,超越传统的顺序、等级、属类的知识链接和知识关联为探索存在于知识之间的“内隐知识”提供了新的途径。因此,知识链接与知识关联的研究将成为未来知识组织体系研究的一个主要方向。

数字化时代的网状知识组织体系中,用户的知识检索和获取是由目的驱动的,用户的需

求和偏好会通过其在网络上的点击流体现出来。更重要的是,通过对用户点击流的数据挖掘,可以获得用户的行为模式,而这种行为模式在一定程度上则反映出存在于用户头脑中以及用户与用户之间的“内隐知识”。因此,数据挖掘理论和应用的相关研究也将成为知识组织体系研究的一个重要内容。

在柔性化的知识组织体系中,尽管知识机构仍然要提供一个事先建立好的体系结构,但是随着知识精英和权威们的地位被逐渐淡化,用户在一定程度上接管和控制了自己所检索的知识组织结构。用户可以根据自身的需求和偏好任意地组织知识。虽然某一个用户自行搭建的知识结构只会赢得少数用户的认同,但是众多用户搭建的众多知识结构的长尾效应^[11]则会产生不可估量的价值。因此,未来知识组织的研究重心将会向用户研究方面偏移,这种偏移也许恰恰让知识组织的相关研究迫近了知识的本质。

在关于知识组织体系的具体研究工作中,以上三个方面并不是完全孤立的。随着其它相关领域理论与技术的不断成熟以及知识组织体系相关研究的不断深化,集成多领域、多方向理论与技术的综合研究必将成为主要的趋势。因此,在知识相关研究中具有独特优势的概念格技术、本体理论与技术、数据挖掘技术、语义标注技术等必将在未来的知识组织体系相关研究中互补借鉴、融合发展。

5 结语

伴随着人类社会的进步与发展,人类知识组织的体系也是在不断地演变与进化。“线形→树形→盒状→链式→网状”的演变进化路径不但是一个逐渐复杂化的过程,也是一个逐渐柔化的过程。柔性化的网状知识体系结构打破了人们试图建立一个权威、准确、清晰、完整的知识组织体系的梦想,淡化了知识精英的权威作用和地位,以多视角、多形式,因人而异地呈现知识的多元化特征。对知识组织体系的演进路径的研究为下一步知识组织体系相关研究的

开展提供了新的思路,为知识组织体系研究的发展趋势作出了探索,为今后的研究工作提供了更多的视角。概念格、本体、数据挖掘、语义标注等理论与技术的互补融合有待于我们在未来的知识组织体系相关研究中进一步努力探索。

参考文献:

- [1] 杨子竞. 国外图书馆史简编[M]. 天津:南开大学出版社,1990.
- [2] 林周.“植物之王”卡尔·林奈[J]. 科学大观园,2009(4):64-66.
- [3] 程鹏.“《中图法》和《杜威法》对照系统”的研制[J]. 图书馆学研究,2001(2):55-57.
- [4] Dewey Decimal Classification Index [EB/OL]. [2010-04-11]. <http://librarian.school.net.hk/ddc000s.htm>.
- [5] 马张华,陈志新. 阮冈纳赞对现代编目理论的贡献[J]. 图书馆论坛,1998(2):24-26.
- [6] 温伯格 D. 新数字秩序的革命[M]. 北京:中信出版社,2008.
- [7] 赵蓉英. 论知识的网络结构[J]. 图书情报工作,2007,51(9):6-10.
- [8] Statistics [EB/OL]. [2010-04-15]. <http://en.wikipedia.org/wiki/Special:Statistics>.
- [9] The Prince of Egypt (1998) [EB/OL]. [2010-04-12]. http://www.amazon.com/Prince-Egypt-Val-Kilmer/dp/B00000JGOQ/ref=sr_1_1?ie=UTF8&s=dvd&qid=1271304583&sr=8-1.
- [10] Nonaka I, Takeuchi H. The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation[M]. Oxford: Oxford University Press,1995.
- [11] Anderson C. The Long Tail[EB/OL]. [2010-01-22]. <http://www.wired.com/wired/archive/12.10/tail.html>.

滕广青 吉林大学管理学院博士研究生。通讯地址:吉林省长春市人民大街 5988 号。邮编:130022。

毕强 吉林大学管理学院教授、博士生导师。通讯地址同上。

(收稿日期:2010-04-17)