

面向数字环境的《中图法》通用复分表修订思考^{*}

薛春香 侯汉清

摘 要 为推动传统分类法在数字环境中的应用,文献分类法的分面组配化趋势日渐显著,而编制通用复分表是体系分类法分面化改造的主要措施之一。通用复分表在数字信息组织中的应用主要体现在分类浏览信息、提供多个信息检索点、精化检索结果等方面,而《中图法》通用复分表目前存在很大不足,其修订策略有:扩大类目复分范围,增加类目组配自由度;增加通用复分表种类,增加揭示维度;改变标记方法,增加操作灵活性;多重列类,细化复分表;注重与其他分类法的兼容性,等。但在总论复分表与主表类目重合、同一复分表不支持多次复分功能、附表不能单独使用等问题上仍需进一步考虑。参考文献 9。

关键词 中国图书馆分类法 通用复分表 体系分类法 分类法修订
分类号 G254.122

ABSTRACT Faceted transformation is the main method for the adaptation of traditional classification to the digital environment. The common auxiliary tables for the subdivision of CLC are its main facets. The revision of these common auxiliary tables can promote the applicability of CLC to the digital information organization and retrieval. This paper analyzes the changes of applied environments of traditional classification, and the applications of the common auxiliary tables in digital information organization, and makes certain considerations about how to revise the common auxiliary tables of CLC. 9 refs.

KEY WORDS *Chinese Library Classification.* Common auxiliary tables. Enumerative classification. Revision of traditional classification.

CLASS NUMBER G254.122

为推动传统分类法在数字环境中的应用,文献分类法的分面组配化趋势日渐显著,而编制通用复分表是体系分类法分面化改造的主要措施之一。《中图法》对通用复分表的修订历来都比较慎重,每次都限于局部修订,且未能从适应数字环境的角度修订,导致其对数字信息组织的适用性不佳。本文拟从适应数字环境中信息组织的角度来思考《中图法》通用复分表的修订。

1 通用复分表在数字环境中的应用

1.1 文献分类法应用环境的变化

随着信息环境的数字化和网络化,将传统

分类法应用到数字环境中已是不争的事实。但数字信息组织不同于传统图书馆纸质文献资源的组织,数字信息组织具有以下鲜明的特点:

1.1.1 信息组织对象从实体层到内容层,信息组织粒度细化

传统环境中信息组织的对象是以负载信息的物理载体为主体,关注物理实体特征的揭示、组织、存取,当然也包括对信息内容的粗略揭示;但是在数字环境中,信息载体的虚拟化,使人们不必关注其载体形态、表现形式、存储位置,把组织的重心放在信息内容上,更加注重对信息内容的多维度、细粒度揭示,从信息实体管理转向信息内容管理,从注重实体的文献组织转向注重内容的信息组织和知识组织。这是一

^{*} 本文系南京理工大学科技发展专项基金项目“基于集成的领域知识组织系统构建研究”(XKF09026)研究成果之一。

个根本的变化,要求一切用于信息组织的语义工具和信息检索工具,都必须发生相应的变化。

1.1.2 排架不再是分类的主要目的,文献分类从一维走向多维

以往组织分类排架和编制分类目录是文献分类的两大作用。为了保证排架功能,只能建立封闭的线性分类体系。这种分类方法不能全方位地揭示文献主题在人类知识结构中的位置,从而妨碍了检索功能的发挥。如前所述,由于数字信息资源的虚拟性,信息组织不受物理排架的限制,信息分类可以更加关注信息内容的多维度、多视角揭示。数字世界让人们能够超越现实世界最基本的排序规则:同样的东西可以同时被排放在不同的地方,而这种状况好过每一样东西各就各位的情况^[1]。因此,面对纵横交错的人类知识结构,多入口的网状分类体系才能更好地与其相适应。

1.1.3 从单一检索入口向多入口方向发展

数字信息分类与一般的文献分类不同,除了学科主题外,它还以对象、区域、资源类型、媒体类型、语种等作为信息分类的划分标准。允许不同用户从不同的分类途径获取到相同的信息。因此,数字信息分类中,大多采用主从分类体系、多重分类体系并列的模式。如把按照学科主题划分的分类体系作为主分类体系,把面向对象、区域、资源类型、语种等的分类体系作为辅助,主从结合,交替使用。譬如要在搜狐网页目录中找到有关“美国大学”的相关信息,可以采用以下两种途径:

途径一 教育培训 > 高等教育 > 高等院校/大学 > 国外大学 > 美国

途径二 国家与地区 > 世界各国/地区 > 美国 > 教育/科研 > 大学

两种途径分别从主题和地区两个入口找到相应的结果。在类目展开过程中,学科主题分类和地区分类分别作为主从分类体系。总的说来,网络环境中数字信息组织的特点是以一个分类体系为主,多个分类体系辅助;知识分类与主题分类相结合,信息内容列类与信息形式列类相结合^[2]。因此,应当充分考虑传统分类法应用于数字信息环境中对复分表的新要求和新特点。

1.2 通用复分表的编制意义

传统分类法基于简化类表、增强类表细分程度的目的,将分类表中具有共性区分性质的类目抽取出来编成复分表。编制通用复分表的意义体现在两个方面:对分类表本身而言,可以节省主表篇幅、增加类表的灵活性、增强分类法的内在规律性^{[3]55};对分类法应用而言,编制通用复分表能够提供信息的分面、多维度的揭示和组织功能。在秩序化为重的传统环境下,通用复分表的分面揭示功能得不到充分的发挥。鉴于上述数字信息组织的特点,应当重视信息内容分面组织和揭示的意义,摒弃以往把复分表列入辅助表(附表)地位的观念,从而关注复分表的修订,把复分表的修订列为传统分类法分面改造的一个重要内容。

有的学者认为在数字环境中电子分类法可以不实行复分、仿分,其理由有二:一是电子分类法不受载体的限制,可以无限制地增设各种类目;二是随着全文索引和主题标引技术的发展,数字环境中信息只需要实现粗分类,无需采用复分、仿分等方式进行细分。笔者认为,数字环境中可以适当控制分类的级别,应当采用粗略分类与多重分类相结合的方式,即分类深度达到三至五级即可(粗略分类),但在每一层级可以考虑采用多种划分标准,从多个不同维度进行分面复分(多重分类);换言之,即应当控制每一种划分标准的深度,而通过多个不同的划分标准来扩展分类的维度。通用复分表和专类仿分可作为各类进行多重分类的依据,也可增加分类表的规律性和易用性。

在目前国内外的一些网站自编分类法中,除了学科主题外,对象、关联区域、资源类型、媒体类型、语种等都是信息分类的标准和维度。对 AOL、MSN、Yahoo!、Open Directory 等 9 个门户网站分类目录的调查分析发现,这些网站的信息分类至少采用了两种以上的分类维度,并把地域、资源类型等作为与主题同等的大类入口^[4]。由此可见,数字环境中复分表不是要不要的问题,而是如何加强分面组织和揭示的问题。这说明,分面分类在数字信息组织中具有良好的应用前景。通用复分表实质为文献分类

法提供了除学科主题外的信息内容和表现形式的若干组面。

1.3 数字信息组织中通用复分表的应用

1.3.1 基于复分表分类浏览信息

一般的分类浏览是基于分类体系主表建立的学科主题导航,但对于关注某一地区新闻或者某一时期或历史朝代的用户来说,可能更倾向于从地域或时代途径去浏览信息,那么借助于地点或时间复分标识可把同一地区的经济、社会、教育、文化等主题的信息聚合,或将同一时期发生在世界各地各领域的事件进行集中,具有一定的对比、参照意义。例如在 Yahoo 分类目录中可以从地区入口聚合相关的信息,在地区下再按照学科领域主题来细分。

1.3.2 基于复分表提供多个信息检索点

若在信息标引过程中从不同的维度揭示信息的外部特征和内容特征,则这些信息点都可以作为检索入口提供多途径检索。这在一些数字图书馆中应用较多,如英国文化遗产数字图书馆中的图像资源高级检索,除了提供建筑类型、建筑风格等主题检索途径外,还提供了建筑年代、国家地区的检索,这些信息的揭示和检索途径的提供,都可以借助各种复分表来完成。这种分面多维的检索界面给予用户的不再是只有一个框架、一种思路的固定模式,而是允许用户按照自己习惯和喜欢的方式来展开个性化的检索发现之旅,是把分类法、主题词表等传统知识组织系统嵌入到网络信息检索界面中的有效应用^[5]。

1.3.3 基于复分表复分标识精化检索结果

网络信息检索返回结果往往较多,信息筛选的工作量极大,通过聚类形成小的信息集合是精化检索结果的一个主要途径,而时间、地域、信息类型、语种等是最易实现的聚类标准。因此,基于复分表可以实现分面检索和检索结果的分面化。把文献分类法作为主从多维分类体系,将学科主题和各种复分特征分开,实现基于分类的分面浏览,这是一种动态性的浏览和检索结果的高效精化。我们可以在元数据标准中提供这样一些检索点的标注,如 MARC 中《中

图法》类号 690 字段,可以提供 \$a 主类号、\$w 总论复分、\$x 地区复分、\$y 时代复分、\$z 民族复分这样一些子字段^{[3]190}。

2 《中图法》通用复分表现状及问题

《中图法》自 1975 年首版,就编制了相应的通用复分表,每次《中图法》修订对复分表也会进行或大或小的增补,其中 1999 年的第四版修订较大,除对原有六个复分表进行完善外,还增补了“世界种族与民族表”、“通用时间、地点表”两个通用复分表,但相较于数字信息分类组织的需求,《中图法》的通用复分表仍显不足。从《中图法》数字信息组织适用性角度分析,其通用复分表目前的不足主要体现在:

(1) 类型偏少。只有总论、时间、地点、种族民族四种类型八个复分表,不仅未达到 DDC、UDC、BC2、《中国图书资料分类法》(以下简称《资料法》)等传统文献分类法通用复分表的种类,更遑论满足网络环境和数字信息多维分面揭示的需求。

(2) 通用复分表不能独立使用。而数字环境中信息资源的地域、时间、资料类型、语种等特征是与学科主题特征并列作为检索入口的,具有独立的地位。如地区综合性信息的分类导航,地区比学科领域主题更适合统领信息,因此,在数字信息分类标引中应当允许单独使用复分表。

(3) 类目细分程度不够。《中图法》通用复分表一直作为附表使用,类目详细程度不够,尤其是地区和时代复分表太过简略,“世界地区表”只复分到国家一级,“中国地区表”只复分到省一级;时代分期过于宽泛,且偏重历史时期。而网络环境中的数字信息更多偏重时效,用户需求往往也更关注年月。

(4) 各表类目设置大多仍采用单一的分类标准,分面揭示程度差,很少使用多个分类标准进行类目的划分和展开。

(5) 缺乏适用于网络环境或数字环境的类目。《中图法》通用复分表类目的设置仍偏重于传统文献,偏重于文献收藏,因此缺乏具有时代

特征和网络特征的类目,比如“网页”、“维基”、“博客”、“论坛”等一些网络中主要的数字资源类型在“总论复分表”中均找不到专指的类目来表示。

《中图法》通用复分表的上述不足导致其在数字信息组织中应用的局限性,甚至影响了《中图法》在网络环境中数字信息组织功能的发挥。

3 《中图法》通用复分表修订的策略

数字时代,实时、交互、数字化分布的信息资源的标引与检索需求有别于传统重藏轻用的文献观念,因此,用于信息整序的分类法也应转向以深度标引与快捷检索为中心、面向终端用户使用、与知识发展同步的现代化管理模式与理念^[6]。DDC21 版把“分面化作为修订的方向和分类法增强主题检索能力的手段^[7]”。同样,在文献工作中发展起来、具有广泛用户基础、体系结构相对完善的《中图法》必须肩负起构建我国数字信息分类体系的重任,其修订必须考虑数字环境下的需求和应用,将复分表的修订作为《中图法》分面改造以适用于数字环境的一项重要措施。

3.1 扩大类目复分范围,增加类目组配自由度

《资料法》第四版规定,10 个通用复分表适用于主表任何一级类目复分组配,还规定无论主表注明或未注明“依……分”,用户需要通用复分表复分时,复分号均需要连同其专用标识符号一起加在主类号之后^[8]。这种方式既有利于自动分类的发展,也有利于基于分类的分面浏览和检索结果精化,主表展现学科专业属性,附表则从其他特征多维揭示。因此,《中图法》可借鉴《资料法》的做法,最大限度地扩大类目复分范围,增加类目复分组配的自由度。

3.2 增加通用复分表种类,增加揭示维度

大量增设复分表和复分类目,以容纳不断出现的新事物、新主题,提高类分信息的专指度,有利于多维度信息检索。对于《中图法》而言,虽然目前已经拥有 8 个通用复分表,但与国

外的许多分类法,如 DDC、UDC 等相比,仍然缺少诸如语种表、人员复分表、材料复分表、环境复分表等处理信息的其它多种有效途径。尤其是语种表,网络信息的多语种性是其显著特征,DDC、UDC、CC 等国外主流文献分类法中都拥有语种表。因此《中图法》应当增加语种表。此外,考虑到网络信息中经济、生活类信息偏多,可以考虑增加行业表、产品表等。

3.3 改变标记方法,增加操作灵活性

《中图法》通用复分表复分规则较乱,例如进行复分时加不加“0”,成为《中图法》长期未能解决的问题。通用复分表使用时,复分标识符使用往往不统一,例如,地区复分除了加“0”的问题外,有时必须使用专用复分标识,有时则不需要,允许直接加复分号;有时直接使用中国地区复分号,有时还需在中国地区复分号前加上代表中国的“2”。分类规定之繁琐复杂,专业分类人员也难以掌握,常常出错。因此建议在不引起与主类号重号的前提下,凡是通用复分,允许自由组配成后组式类号,并统一使用专用复分标识符;同时采用分段标记制度。这样既便于标引操作,尤其是自动标引,便于各个组面的析出和对应类号的转换;同时还便于将组配号码拆解成各个检索点,通过类号轮排,提供多种检索入口,大大提高基于分类号的多途径检索、多因素检索功能;并可以进行检索精化,满足互联网环境下的分类标引和信息检索需要。

3.4 多重列类,细化复分表

除了主表应当采用多重列类外,复分表也应当实现多重列类。《中图法》第四版在三版“世界地区表”只设立国家的基础上,增加了地貌特征、宗教区域、语种特征等地区列类标准,提高了类目的完备性和专指性。但这在分类实践中仍显不够,还应当借鉴 UDC、BC2、CC 等半分面或分面分类法,结合网络用户检索途径的多样化特征,设置更多分面列类,比如地区表还可以考虑增加历史特征、经济发展程度、方位、气候等特征;此外,在组织过程中还应当增加字顺或笔画、关注度等排序方法。另外,鉴于通用

复分表的类目层次相对较浅,列类不够详细,在适用数字信息组织时可以考虑增加细分类目,比如“中国地区表”可以细分到主要城市,甚至是热点县市。

3.5 注重与其他分类法的兼容性

数字时代网络信息全球化发展,《中图法》通用复分表的编制与修订也应当考虑全球化共享,注重与其他分类法的兼容性。目前,《中图法》复分表在一些类目的划分上仍以中国特色为主,忽视了与其他分类法的兼容。比如“国际时代表”中时代的划分,以我国大部分学者所认定的历史分期来制定时间表,其最大缺陷在于对那些不认可或不知道这种分期概念的人来说,造成了不便。如对“世界史”的划分,我国一般以“原始社会、古代、近代、现代”概念划分,但西方学者却分法不一,英国韦尔斯所著《世界史纲》将“世界史”的时期划分为“人类以前、人类形成时期、最初的文明时期、陆路的时期、文明复兴时期、列强的时代”等^[9]。为避免历史时期划分观点的干扰,《中图法》的时代表一方面可以考虑采用多重标准列类,吸收其他分类法的不同划分方法;另一方面可以采用国际标准和国家标准。目前新闻出版总署发布的《中文新闻信息分类与代码》就直接引用 GB/T2659-2000《世界各国和地区名称代码》作为世界地区复分表,引用 GB/T2260-2002《中华人民共和国行政区划代码》作为中国地区复分表,引用 GB/T3304-1991《中国各民族名称的罗马字母拼写法和代码》作为中国民族复分表。这可作为《中图法》复分表加强与其他分类法兼容性的一种方法来借鉴。

此外,《中图法》复分表修订中还存在一些由来已久的问题,比如总论复分表与主表类目的重合、同一复分表不支持多次复分功能、附表不能单独使用等问题,这些修订意见仍值得进

一步考虑。

参考文献:

- [1] [美]戴维·温伯格. 新数字秩序的革命[M]. 张岩,译. 北京:中信出版社,2008:5.
- [2] 陈树年,李青华,朱连花. 网络信息组织与《中国图书馆分类法》搜索引擎版研制(上)[J]. 中国索引,2005(3):18-24.
- [3] 俞君立,陈树年. 文献分类学[M]. 武汉:武汉大学出版社,2001.
- [4] Chaim Zins. Models for Classifying Internet Resources[J]. Knowledge Organization, 2002, 29(1):20-28.
- [5] 曾蕾. 在浏览和检索界面设计中利用知识组织系统(KOS)[EB/OL]. 2006年数字图书馆前沿问题高级研讨班. 北京,2006. [2008-09-10]. <http://elib.lib.tsinghua.edu.cn:8080/meeting/ppt/zenglei2.pdf>.
- [6] 田书格. 新版UDC与我国三大分类法的通用复分体系比较研究[J]. 山东图书馆季刊,2007(4):57-61.
- [7] 转引自马张华. 《杜威十进分类法》21版修订:特点与启示[J]. 图书情报工作,1997(6):24-27.
- [8] 侯汉清,周冰. 从列举式分类法向组配分类法的转变:试论新版《资料法》的修订特色[J]. 图书馆杂志,2002,21(2):35-39.
- [9] 杨凌云. 我国《中图法》的修订可从国外分类法中得到哪些借鉴[J]. 晋图学刊,2000(3):47-49.

薛春香 中国科学技术信息研究所在站博士后,南京理工大学信息管理系讲师。通讯地址:南京。邮编 210094。

侯汉清 南京农业大学信息科技学院教授,博士生导师。通讯地址:南京。邮编 210095。

(收稿日期:2009-03-22)