

王子舟 王碧滢

# 知识的基本组分——文献单元和知识单元<sup>\*</sup>

**摘 要** 知识组织是当代图书馆学的重要命题,知识组织是以知识组分为基础的。文献单元和知识单元是客观知识的基本组分,二者既有区别,又有联系,不能有所偏废。基于知识组分的知识组织有其涵义、原理和方法。参考文献 19。

**关键词** 知识组分 知识组织 文献单元 知识单元

**分类号** G250

**ABSTRACT** As an important issue in modern library science, knowledge organization is based on knowledge component parts. Document elements and knowledge elements are basic component parts of objective knowledge. They are different and correlated. The knowledge organization based on knowledge component parts has its own implications, principles and methods. 19 refs.

**KEY WORDS** Component part of knowledge. Knowledge organization. Document element. Knowledge element.

**CLASS NUMBER** G250

世界上万事万物的组成离不开构成其成分的基本元素和元素间的组织结构。元素与结构是事物形成的两个要素。事物的元素也可表述为事物的组分。客观知识世界是由众多人类显性知识系统构成的,这些知识系统同时也是由无数知识的基本组分组成的。知识的基本组分可分为两大类,一是具有“硬性”特点的、具有独立载体形态的知识组分(即文献单元),一是具有“软性”特点的、不具有独立载体形态的知识组分(即知识单元)。

## 1 文献单元与知识单元的涵义

### 1.1 文献单元

为了认识文献,人们对其下过种种定义。近几十年中,文献的定义已不下数百种。其中较有权威和具有广泛影响的是国家标准对文献的定义。如《文献类型与文献载体代码》(GB3469-83)与《文献著录总则》(GB3792.1-83)均将文献定义为“记录有知识的一切载体”。这个定义简洁明了,对统一人们关于文献的认识起到了良好的规范作用。但它也有某些疏漏。首先,记录有知识的一切载体并非均是文献,如人的大脑也记录有知识,但大脑并不是文献。文献只能是记录有知识的“人工载体”。尽管人类早期的图书曾使用过泥板、纸莎草、兽骨等自然物

为其载体,可是这些自然物也经过了一些人的加工,而且文献载体在后来的发展中有着不断人工化、非自然化的演变趋势。正如加拿大学者麦克卢汉所说:“拼音字母表用在粘土和石头上是一回事,用在轻盈的莎草纸上就是迥然不同的另一回事了。”<sup>[1]</sup>其次,记录有知识的人工载体很多,除了图书以外,还有钱币、邮票、乃至有着某种艺术风格的建筑物,它们都记录了一定的知识但却不是文献,因为文献必须是专用于传播知识、交流知识的,它在用途、功能上具有知识传播的惟一性。以往我们讲,文献有其组成的三要素:知识内容、记录符号、载体形态。但我们往往忽略了一个问题,就是文献三要素的结合必须满足文献传播、交流的要求与目的,否则就不能成其为文献。如过去我们将载有大段铭文的商周时期的青铜器称作“青铜的书”,这是欠妥的。载有铭文的青铜器皿虽然具备构成文献的三要素要求,也有传播知识信息的某种倾向性,但毕竟与刻有文字的石碑不同,它们有着某种容器方面的特定用途,传播知识并非其惟一的主要功能。所以,这些载有铭文的青铜器皿只能算作“准文献”。当人们将其铭文用纸张拓印下来时,拓片才成了完全意义上的文献。

根据上述理由,我们可将文献重新定义为:“专门记录和传递有知识的人工载体”,而文献单元就是

\*本文为武汉大学“新世纪教学改革研究计划”项目的研究论文之一。

“专门记录和传递有知识的人工载体单元”。应该说,文献虽然是个抽象的概念,但我们所见到的文献却都是具有单元形态的,即文献的客观存在、现实呈现,是以各种各样的文献单元形式表现的,如一部著作、一篇论文、一盒磁带、一张光盘等等,图书馆的文献组织通常就是把一个个独立的知识出版单元视为文献单元的。

在古代图书整理活动中,文献单元的辨认与辨别主要是依据载体单元;而在现代社会中,由于文献数量的激增、文献形态的多样,人们对文献单元有了更加深入的了解。在现代图书馆学中,文献单元依据其自身的三要素往往还可分解成若干具体的单元形式,如文献的知识内容单元、知识形式单元与载体形态单元等等。所谓知识内容单元,是针对文献单元中所包含的知识内容体系而言的。任何一个文献单元中的知识内容都具有某些主题,文献单元中的知识内容涉及了多少主题,就可视为具有多少知识内容单元。例如王益、汪轶千主编的《图书商品学》,就具有图书学与商品学两个知识内容单元。所谓知识形式单元,是针对文献单元中的知识符号体系而言的。符号体系往往有题名(包括章节名)、责任者等形式特征,它们可使文献单元的表述系统进一步区分成若干组分。如一部有几个题名与几个责任者的学术论文集,就有几个知识形式单元。所谓载体形态单元,是针对文献单元在“物理”方面表现出来的独立状态而言的。如文献单元的封面、版权页、正文、附录、索引等,都可划分为不同的载体形态单元。正因为现代图书馆学对文献单元的认识越来越深化,图书馆的文献组织方法越来越周密复杂,图书馆学的应用知识体系也越来越细化、丰富。

## 1.2 知识单元

什么是知识单元?不同学科对知识单元的阐释不同。早在 1980 年,我国科学学专家赵红洲先生就曾经提出:知识单元是已经获得科学共同体认同的、堪称“粒子”形态的科学概念,它是科学的细胞,科学大厦的砖块。大量的知识单元经过重组,按一定思路可凝聚成知识纤维,知识纤维又可在更高层次上组成复杂的知识系统<sup>[2]</sup>。此外,人才学研究者王通讯在 80 年代中期提出:“所谓知识单元,系指知识系列的一些基本‘概念’。不同的知识系列之间,可能有相同的知识单元。但是,一个知识系列的形成与规范化,主要体现在它具有一批与其他系列不同的

知识单元。”知识单元组成一个个的知识系列,一个个的知识系列又组成整个知识体系。<sup>[3]</sup>除了科学学、人才学,教育学界也经常提到“知识单元”,不过这一概念一般出现在教学研究领域,如一些研究者将知识单元表述为教材中某章节的基本组分、组元,或某堂课程教学内容的组成单元,它们可以用图表形式表达出来。

相对而言,我国图书情报学界对知识单元的探讨则略显滞后。如 1988 年,张德芳先生撰文指出:知识单元是寓藏于出版单元(某一文献)之中的,一个出版单元可能贮含一个以上的知识单元。“科学知识单元是科学劳动的产品,是在对事物、实物、现象、过程等零散信息基础上,经过系统归纳、分析、整理,能够反映其本质规律的概念、定理或定律。”<sup>[4]</sup>1989 年,马费成先生发表文章称,数据单元也可称知识单元,它是由各种事实、概念和数值等组成的<sup>[5]</sup>。90 年代末期,“知识单元”在图书情报学界的使用频率逐渐多了起来,因为许多图书馆学研究者已经意识到,图书馆如何从传统的文献组织转向文献信息(即客观知识)的组织,如何从文献单元服务深入到知识单元服务,已经是图书馆实践面临着的新的重大课题。

总结上述人们对“知识单元”概念的理解与认识,根据图书情报工作的实践传统,可以为知识单元下这样一个定义:知识单元是客观知识系统中有实际意义的基本单位。这个定义包含以下几层内容:

其一,所谓“有实际意义”,指知识单元或是一个明确的语词概念、一个具体观点,或是一个科学定理、一个数学公式,或是一首歌中的某段旋律、一幅画上的局部构图等。它们是人们对社会存在的感受、认识的反映。如“力”这个概念,当我们停留在“事功曰劳,治功曰力”(《周礼·夏官·司勋》)的认识阶段,只能说它还是生活常识中的一个知识单元;当我们把凡能使物体获得加速度或发生形变的作用称作力,并用某种势场梯度( $F = \partial u / \partial r$ )来表示时,它就变成了科学知识中的知识单元。从认识论的角度看,从“感性的具体”到“抽象的规定”,再到“思维中的具体”,这正是知识单元结晶形成的具体过程<sup>[6]</sup>。

其二,我们所描述、表达出的任何一个知识单元,都是可以归入某个知识系统,成为该系统内容的一个组成部分的。换言之,知识单元正是以知识系

统为参照系提取出的一个概念;离开了知识系统,也就无所谓知识单元了。如我们说“力”的概念是一个科学的知识单元,就是以物理学或力学的知识系统作为参照体系而言。如果没有这样一个参照体系,“力”只能是一个普通的概念,而谈不到“科学的知识单元”了。

其三,知识单元一旦成为知识系统的一个组分,它必须是该知识系统的基本单位。这些基本单位有可分解与不可分解两大类。如诊断学包括物理诊断、电诊断、影像诊断、实验室诊断等基本单位,它们都是诊断学的知识单元。其中物理诊断还可进一步细分成视诊、触诊、听诊、体温及血压测量、呼吸检查等若干知识单元,因此,“物理诊断”就是一个可分解的知识单元。从理论上说,视诊、触诊等知识单元还可再行切割与细分,一直分到不能分解的“知识细胞”为止,这些不能分解的“知识细胞”就构成了诊断学知识系统的不可分解的知识单元。

## 2 文献单元与知识单元的关系

### 2.1 文献单元与知识单元的区别

(1)从知识存在上看,文献单元具有“硬性”特点,知识单元具有“软性”特点。一部科学著作可以在作者的大脑及其著作中间同时贮存、呈现,但科学著作具有物质载体形态,它已经脱离了大脑思维,属于客观知识世界,是一种知识的“凝固”,是一种具有独立存在意义的“硬性知识”。而知识单元既大量地存在于客观知识世界,同时也大量地存在于人的主观知识世界。人的大脑“不失真”地储存文献单元很困难,但“不失真”地储存知识单元却很容易。所谓知识创新,就是人们将储存在大脑中、呈现“游离状态”的知识单元有机重组的一种活动。存在于大脑中的知识单元是隐性的,而存在于客观知识系统中的知识单元也不具有独立的载体形态,而是隐藏在浩如烟海的文献中,是一种“软知识”。

(2)从知识形态上看,文献单元所包含、表现的,大多是完整的知识体系,而知识单元所包含、表现的,更多地是“知识点”。马克思积40年之力撰成《资本论》,从形态上看它只是一个单元文献,但它在内容上却建构了关于资本主义社会发展规律的伟大思想体系。《资本论》中的“商品”、“货币”、“资本”、“剩

余价值”等知识点,形成了这个伟大思想体系中关键性的知识单元。因此,如果说文献单元是知识生产者精心搭建起来的“知识建筑”,那么,知识单元就是那些具体的木石、砖瓦等“知识建材”。从二者的生产数量上来看,“知识建材”的数量显然要大大多于“知识建筑”。

(3)从知识性质上看,文献单元是静止的,而知识单元是衍变的。文献单元的生成是一次成型。新的知识内容取代旧的知识内容,如果通过文献单元形式来表现,必须经过生产新的文献单元来实现更替。知识单元则不同,它具有生命力。在人类知识的积累与传播历程中,知识单元的“内禀智荷”是可以演变的。赵红洲先生曾提出,知识单元的智力(创造力)荷值是由静荷值和动荷值之和构成的。静荷是知识单元内部因子个数及相互逻辑关系构成的,表现知识单元包含智慧多少的度量;动荷则是知识单元在历史进化中积累的智荷,表现知识运动中形成的智慧的积累<sup>[7]</sup>。例如“工业”(industry)这个概念,在工业革命前,其“静荷”意义(即本义)原指一种特殊的人类属性,意译过来是“技术、刻苦、坚毅、勤奋”等,18世纪后期才有了“制造与生产机构”的含义,变成了一个表示“集体”的语词<sup>[8]</sup>。这个附着于“工业”静荷意义上的新的含义,就是历史进化中积累的动荷意义。

(4)从组织方法上看,文献单元的组织方法较为单一,而知识单元的组织方法则复杂多样。当人们将客观知识组织成一个有序的集合时,如果以文献单元为集合的基本组分,那么组织方法则主要是依文献单元的内容属性(类别、主题)或形式属性(题名、责任者)来进行的;如果以知识单元为基本组分,情况就有所不同。知识单元的属性特征多种多样,远较文献单元丰富,故其组织方法也更为复杂多变。例如将化学元素“铁”作为一个知识单元,其属性特征至少有以下这11种<sup>[9]</sup>:

事物名:铁  
类别:化学元素  
颜色:银白  
比重:7.86  
熔点:1535  
沸点:2750  
能否导电:能  
化学性质:活泼

周期表位置:第四周期第 8 族

地壳含量:5 %

原子量:55.847

以上左边所列为铁的属性名,右边所列为铁的相应属性值。可以想象,知识单元的属性特征有多少种,其组织方法(从属性特征角度)也会有多少种。

(5)从获取方式上看,文献单元的获取相对容易,而知识单元的获取相对困难。从知识受众获取知识的经验来说,情形大多如此;就是从图书馆提供知识服务方面而言,提供知识单元的服务较提供文献单元的服务难度也更大。因为,知识单元与文献单元之间往往表现出“知识不对称性”。当我们要找到一个意义准确、完整的知识单元时,可能要阅读众多文献单元,尽管如此仍不能保证这大量的阅读是有效的。通常,在图书馆实践中,提供文献单元服务与提供知识单元服务有着很大差别:前者提供的是有文献载体形态,后者提供的是知识信息(可无文献载体形态);前者属简单劳动(可以依借出文献量计算劳动量),后者属复杂劳动(不能依借出文献量计算出劳动量);前者的专指度有时较低(如读者欲借某类文献),后者的专指度一般均很高(如读者查一个名词、一个数据、一段引文);前者与读者不会“相互进入”(知识信息的切入、交流),后者则常会与读者发生“相互进入”;前者的服务方式是被动式的,后者的服务方式是主动式、互动式的<sup>[10]</sup>。

文献单元与知识单元还有许多区别,以上仅是从主要的几个方面分析论之。

## 2.2 文献单元与知识单元的联系

(1)从知识存在角度看,在客观知识世界里,文献单元与知识单元均是客观知识的基本组分。虽然前者具有“硬性”特征,后者具有“软性”特征,但二者的产生都经过了主观知识的客观化这一过程,而且它们的存在也是互相依赖、互为前提的。知识单元大多存在于文献单元之中,文献单元则是众多知识单元有机组合的成果。

(2)从知识形态来说,文献单元与知识单元既大量地表现为原生态知识,同时也表现为某些再生生态知识。原生态知识是主观知识的原创成果社会化、公开化了的最初知识形态。图书馆学中所谓的“一次文献”就是原生态知识。再生生态知识是为解决人们有效获取知识而对客观知识进行加工、提炼所形

成的知识形态。如《百科全书》中的词条、目录、索引中的款目等,它们也可以通过文献单元、知识单元表现出来。

(3)从知识性质上来看,客观知识所具备的基本性质,如创新性、中介性、可编码性、可传递性、无损耗性、可共享性、可加工性、不确定性等,文献单元、知识单元也都具有。客观知识世界是一个开放的复杂巨系统,从理论上来说,构成这个系统的基本组分,也会不同程度地、或多或少地包含着系统所共有的属性与特征。

(4)在知识组织中,文献单元与知识单元切割、划分均有其相对性。一篇科学论文的抽印本就是一个文献单元,但它发表在某本杂志上时,这本杂志通常会被视为一个文献单元。知识单元也如此,一句唐诗可以是一个知识单元(因此有各种各样的唐诗诗句索引供人们检索),但有时一首完整的唐诗也可作为一个知识单元。因此,文献单元、知识单元在不同的知识组织体系中,根据不同目的及需要,它们的基本单元层次是可分合的。

(5)再从知识获取来看,文献单元和知识单元都能向知识受众有效提供知识服务。20 世纪 90 年代以来,图书情报学界流传着一种观点,认为未来图书馆工作要由文献单元的组织转向知识单元的组织,图书馆服务也会由文献单元服务转向知识单元服务。这个观点有一定道理,揭示了图书馆是一个发展着的有机体。但它也有偏激之处,因为文献的本质就是知识,传统的文献单元服务本身就是提供知识的服务(而我们做得并不好,经常受到各界读者的批评)。通过文献单元汲取知识、提高认识,这是古往今来人们获取知识的最基本方式之一。图书馆应该开发知识单元的服务,但却不能忽略文献单元的服务。况且新事物的出现有时并不排除旧事物的存在,如电子媒介提高了人们知识交流的质与量,但它并不能取代或排除纸质介质的存在,甚至也不能取代和排除“口耳相传”这种古老的知识交流方式的存在和价值。

## 3 基于知识组分的知识组织

### 3.1 知识组织的涵义

将无序的文献组织成一个有序的体系,这是图书馆最基本的活动。从 20 世纪 80 年代起,我国图书馆学研究者已经敏锐地意识到:文献组织的实质

就是一种知识组织。1985年,青年学者刘迅倡导图书馆学应向“知识工程”转变,提出应把知识组织作为图书馆学研究的一个重要领域<sup>[11]</sup>。90年代以后,“知识组织”这一命题成了学术探讨的热点,人们对知识组织的概念也提出不同的见解。如王知津先生指出,将知识组织等同于文献的分类、标引、编目、文摘、索引等一系列整序活动,这是狭义的知识组织;而将知识因子(知识结点)有序化和知识关联(结点间的联系)网络化,这是广义的知识组织。广义的知识组织包含了狭义的知识组织,因为个体文献也可视为是知识因子<sup>[12]</sup>。蒋永福先生则从哲学认识论的角度认为:“知识组织是指为促进或实现主观知识客观化和客观知识主观化而对知识客体所进行的诸如整理、加工、引导、揭示、控制等一系列组织化过程及其方法。”<sup>[13]</sup>这一定义把知识组织划分成了两种类型:主观知识客观化的知识组织、客观知识主观化知识组织。这些观点尽管表述不同,但其中存在着某种共识,它们推进了我国知识组织研究的深入发展。

随着知识社会的来临,“知识组织”渐成多种学科所使用的科学概念,如:知识社会学、知识管理、人工智能、教育学等。甚至在日常语言中,“知识组织”一词出现的频率也越来越多。在广泛应用的层面上,知识组织是指人们对知识(无论隐知识、显知识)的整序、加工、控制、揭示等一系列的组织活动与过程。当然,知识组织这一概念最早产生于图书馆学界。如1929年英国图书馆学家布利斯(H. E. Bliss)《知识的组织和科学的系统》一书就已经使用了这个概念。而就图书馆学来说,知识组织应该是基于显知识中的文献单元、知识单元而言的,其概念可表述为:知识组织是对客观知识组分进行整序的活动与过程。因这个概念具有其专指范畴,故其涵义是狭义的。知识组织的目标是将知识组分整序成能够提供优质服务的知识集合(如图书馆、数据库)。由于知识的有序性、知识的关联性、知识的可检索性只有通过知识集合才能体现出来,因此整序、控制、提供是知识组织的核心点。

### 3.2 知识组织的原理

(1)知识组织的假设前提。当我们将知识组织定义为“是对客观知识组分进行整序的活动与过程”时,就隐含了一个前提、假设:客观知识世界是

无序的。客观知识世界的无序状态是由知识组分的无序呈现的,它是由知识生产的无序、知识分布的无序、知识流通的无序造成的。以知识分布的无序为例,每一专业知识领域中,通常存在着核心知识、外围知识、虚假知识三个区域。核心知识增长缓慢,所占区域最小;外围知识增长相对较快,所占区域稍大;虚假知识增长最快,处于边缘地带,所占区域也最大。虚假知识高度无序,且有泛滥之势,就会造成“知识过剩”和“知识污染”,它是客观知识世界“熵”值增大的活跃因子。知识存在的无序性与人们知识需求的有序性之间形成了尖锐的矛盾。知识组织的理论与方法,就是平衡、克服二者之间矛盾而生成的<sup>[14]</sup>。

(2)知识组织的内在机制。就本质而言,知识组织的内在机制是对人脑记忆机制的一种模拟。人脑的记忆有识记、保持、再现三个环节(即信息的编码、存储、提取),知识组织也如是。记忆的第一环节是“识记”,包括视觉识记、听觉识记、视听识记,知识组织也有这些方式。记忆的中间环节“保持”包括空间保持、系列保持、联想保持、网络保持、更替保持等方式,知识组织也在使用这些方式;记忆的保持会有“遗忘”,知识组织则会有“遗漏”。记忆的第三环节是“再现”,体现为再认与回忆,知识组织的“再现”体现在检索与输出。当然,在人工智能领域里,知识组织已经发展出专家系统这样高级的形式,但这种形式本质上仍是对人脑机制的一种模仿。未来的知识组织形式,在飞速发展的信息技术的支持下,将呈现对人脑记忆机制、思维机制交互模仿的趋向。

(3)知识组织的需求基础。强调知识组织以知识的基本组分为基础,是因为人们获取知识的需求与行为常常诉求于知识组分。学习型的知识需求倾向于得到文献单元,而研究型知识需求倾向于获取知识单元。因为文献单元可以提供学习者所需求的系统性知识,而知识单元则可以提供研究者所需求的概念、公式、定理、数据、事实等。另外在现实生活中,人们对知识的需求,有时从表面看来是系统的、整体的,但实际上是可分解成诸多单元的。就好比买房子,人们在作出是否购买的决定时,往往考虑的是房子的建筑质量、面积、楼层、朝向、位置、环境等各种商品房中隐含的特性,而不会只买一个叫“房子”的东西<sup>[15]</sup>。相反,人们对知识

的需求,有时从表面上看只是针对某一知识单元、知识结点,但实际上只有掌握了一个系统知识后,才能真正了解这一知识单元、知识结点。因此,人们知识需求的不同,以及文献单元、知识单元承载功能的不同,决定了知识组织者对二者不能有所偏废。

(4) 知识组织的基本工具。知识必须通过符号系统将其记录下来,并附着于一定的物质载体上,才能进行传递、交流、共享。从符号学角度看,用于人们之间的相互交流的符号系统很多,如航海信号、军用信号、宗教仪式、礼节仪式、语言等,其中语言是最为重要的符号系统。其他符号系统在一定时间内只与一定数量的人发生关系,而语言却是人们时刻都要使用的、参与其事的。人的感情交际可以用手势、表情、音乐、绘画来表达,而知识的传递,即人们思想观念的表达,一般是通过语言来进行的。故而,文献单元、知识单元大多情况下是使用语言符号系统进行记录,以实现其外在化、客观化、社会化。由于语言是个人以外的、社会的东西,个人不能创造语言,只能使用语言,语言较之其他符号系统又具有最广泛的社会性,因此用语言记载的知识才能在最大范围内实现共享。著名语言学家索绪尔提出,语言是由“所指”(概念)和“能指”(音响-形象)构成的,二者结合才成为符号<sup>[16]</sup>。而符号是有差别、有单位的,所以也是可编码的。至少能以其“联想关系”、“句段关系”编码。由此我们可以得到这样的结论:语言是知识组织的最基本的工具。各种知识体系、知识集合基本上是用语言材料组织起来的。即使众多知识集合使用了大量人工语言,人工语言也只不过是自然语言的代替物。

### 3.3 知识组织的方法

(1) 知识组分的标识。在一个产品密集存放的仓库里,如果每个产品没有标号登录,就不可能形成有序的存取与保管。同理,知识组分如不经过标识记录,也不可能形成有序的知识集合(如图书馆、数据库)。标识记录是对知识的某种描述,自身不显示任何智能行为,其惟一作用是携带知识。传统的标识记录,主要是对知识文本人工著录与标引进而形成款目来实现的。但是,由于网络信息资源的出现和激增,图书馆不得不越出自己的“围墙”来链接外部世界,与之进行知识资源的交换。而传统的记录

方法面对具有动态性、开放性、非线性、不规范性和多媒体性很强的网络资源,又显得无法适应,力不从心。因此,新的标识记录方法(如元数据)就应运而生了。

元数据(metadata)是应用于互联网环境的一种关于数据的结构化数据。其任务是对数字化的知识组分进行格式化描述,以便使计算机或人能对其进行识别和处理。目前元数据已有许多种类,其中在图书情报界最有影响的是 DC(都柏林核心元素集)。DC由 15 个元素组成(分属内容描述、知识产权描述、外形描述三个类型),现已被翻译成几十种语言。其之所以有影响,是因为它是一种支持简单而含有足够信息量的元数据,并能提供一个共享的语义学标准,使那些不同规则编目的元数据(如 MARC)可以直接交换<sup>[17]</sup>。

(2) 知识组分的简化。根据知识受众最省力原则的要求,对内容量较大的知识组分(主要是文献单元)进行压缩、提炼,形成文摘、提要等。它们可以附着在知识组分上,或单独为序形成二次文献。知识简化在我国起源很早,至少可以追溯到刘向整理皇家藏书所作“书录”的汉代。当然,知识简化方法并不局限于图书馆领域,如宋代朱熹因司马光《资治通鉴》294 卷浩繁难读,就曾用提要加小注的方法将其压缩改编成 59 卷的《资治通鉴纲目》,实际上这已是一种知识的加工重组了。与知识简化相对应的是知识繁化。如左丘明为《春秋》作传,裴松之为《三国志》作注,这些传注文字要多出原文本的文字数倍。但这些现象出现在“知识稀薄”的年代,随着人类知识的指数增长,知识文本的繁化现象会越来越少,并作为一种特殊需要而局限于很小的领域。

(3) 知识组分的整序。知识组分经过标识记录拥有了自己的“身份”与“地址”之后,便开始进入了整序过程。知识组分的整序方法主要有以下几种:

知识分类。即按科学知识的系统性分门别类地组织与揭示知识组分,使之成为一个树形的有序化结构。常用的基本工具是先组式的各种文献分类法。在知识组织的手工时代,文献分类法居知识组织方法的主导地位。在网络环境下,这种分类法因其符合人类认识事物的思维方式,有较强的系统性、族性检索功能,故而在知识组织中依然发挥着不可

替代的作用。如网络资源的浏览、非文本信息的组织、超文本信息的管理等,都离不开文献分类法。近几年,自动分类技术也日趋成熟。自动分类不仅适用于文献单元,而且还能关注文本中“文字的重要性”<sup>[18]</sup>,进而使知识单元的分类也能得到充分的实现。

知识聚合。即从具体事物、对象和问题的主题来组织和揭示知识组分内容,使之成为一个线形的有序化结构。这是一种后组式的方法,人们习惯称之为主题法。主题法常用的基本工具为标题法、单元词法、叙词法、关键词法等。由于用主题法组织与揭示知识组分较分类法更具有直接和直观的特点,目前网络资源相当一部分浏览器与搜索引擎都使用之,或者分类法、主题法混用。另外,知识聚合的方法能深入到文献的“果肉”中提取单元词、关键词、主题词,并用参照系统来表示知识关联,所以这种方法更适用于知识单元的组织。传统图书馆学领域编制的各种人名索引、地名索引,各类型词典、字典,乃至众多数据库,大多是使用这种方法。

语义网络。语义网络是模仿人类联想记忆的一种图式组织方法。它由节点(node)和链(link)组成一个网状结构。节点表示知识单元;链则表达知识关联,如单项链( )、双项链(→)等。目前超文本(Hypertext)、超媒体(Hypermedia)信息组织方法就是一种典型的语义网络组织法。超文本中的“链”定义了超文本的非线性网络结构,提供了以联想方式浏览、查询节点的功能。因此,“链”是超文本的灵魂<sup>[19]</sup>。有了语义网络组织法,文献单元、知识单元就可以组织成超文本、超媒体数据库;知识单元就能按布鲁克斯所说的“知识地图”联结在一起,供人们“按图索骥”。

总之,知识组织的方法是不断发展的。一些新的知识组织方法,如自动文摘(automatic abstracting)、自动分类(automatic classification)、机器翻译(machine translation)、专家系统(expert system)、数据挖掘(data mining)、推送技术(push technology)等,都是目前正在研究并有广泛探索空间的领域。知识组织向智能化发展,是可以预见的、不容置疑的发展方向。同时,知识组织研究也超出了图书馆学情报学界,成为人工智能、脑科学、心理学、术语学、教育学、知识社

会等众多学科共同关心的命题。有着悠久知识工作历史积淀的图书馆学,应对未来知识组织的发展作出新的贡献。

#### 参考文献

- 1 [加]麦克卢汉,秦格龙编;何道宽译. 麦克卢汉精粹. 南京:南京大学出版社,2000
- 2 赵红洲. 初论“潜科学”. 潜科学,1980(1)
- 3 王通讯. 论知识结构. 北京:北京出版社,1986
- 4 张德芳. 激发和活化凝固在文献中的知识:论图书馆改革. 四川图书馆学报,1988(6)
- 5 马费成. 知识组织系统的演进与评价. 知识工程,1989(2)
- 6 赵红洲. 科学能力学引论. 北京:科学出版社,1984
- 7 赵红洲等. 论知识单元的智荷及其表示方法. 知识工程,1991(3)
- 8 [英]雷蒙德·威廉斯著;吴松江等译. 文化与社会. 北京:北京大学出版社,1991
- 9 陆汝钊. 人工智能(上册). 北京:科学出版社,1989
- 10 王子舟. 文献信息服务与文献服务区别论. 图书馆理论与实践,1996(3)
- 11 刘迅. 知识工程:未来图书馆学情报学教育内容变化的学科归宿. 情报科学,1985(5)
- 12 王知津,王乐. 文献演化及其级别划分:从知识组织的角度进行探讨. 图书情报工作,1998(1)
- 13 蒋永福. 图书馆与知识组织:从知识组织的角度理解图书馆学. 中国图书馆学报,1999(5)
- 14 蒋永福,付小红. 知识组织论:图书情报学的理论基础. 图书馆建设,2000(4)
- 15 熊秉元. 灯塔的故事. 北京:社会科学文献出版社,2002
- 16 [瑞士]德·索绪尔著;高名凯译. 普通语言学教程. 北京:商务印书馆,1980
- 17 吴建中主编. DC元数据. 上海:上海科学技术文献出版社,2000
- 18 陈树年. 网络信息的自动分类问题. 图书馆杂志,2001(10)
- 19 刘国明,夏祖勋. 超文本和超媒体中的知识表示方法. 华东船舶工业学院学报,1994(3)

王子舟 武汉大学信息管理学院图书馆学系主任,博士,副教授. 通讯地址:武汉市. 邮编 430072.

王碧滢 武汉大学图书馆. 通讯地址同上.

(来稿时间:2002-07-11)