

·专题·

知识组织的科学理论阐释*

苏新宁

(南京大学信息管理学院 江苏南京 210093)

摘要:科学理论是人类智慧的结晶,是对相应科学领域高屋建瓴的总结。鉴于此,文章对系统科学理论中的老三论(系统论、控制论、信息论)、新三论(协同论、耗散结构理论、突变论)对知识组织中的作用进行了阐释,对利用系统科学理论知识组织中的一些重要问题进行了解释。此外,还将图书情报学中的基本定律(布拉德福定律、齐普夫定律、洛特卡定律、小世界理论)对知识组织的适用性和存在问题进行了说明。以期从理论角度为提升知识组织的水平提供参考。

关键词:系统科学理论 图书情报学定律 知识组织 系统论 控制论 信息论 协同论 耗散结构理论 突变论

中图分类号:G254.29 **文献标识码:**A **文章编号:**1003-6938(2013)06-0001-07

Scientific Theory of Knowledge Organization

Abstracts Scientific theory is the crystallization of human wisdom, is also a high-level summary in the field of science. In view of this, author illustrates the role in knowledge organization about the system science theory (system theory, cybernetics, information theory, synergy theory, dissipative structure theory and mutation theory), and explains some important problems in knowledge organization by system science theory. In addition, author also states the applicability and problems of fundamental law in informatics about (Bradford's law, Zipf's law, Lotka's Law, the small world theory) to knowledge organization. The paper aims is to promote the level of knowledge organization from theory angle.

Keywords system science theory; informatics law; knowledge organization; system theory; cybernetics; information theory; synergy theory; dissipative structure theory and mutation theory

1 引言

科学研究需要有科学的理论和方法来支撑,对于每一项研究也需要科学的理论来升华和诠释、需要借助科学的方法来实现。就知识组织研究而言,欲获得高效知识服务需要来自于科学理论指导知识组织,需要科学的方法实现知识组织。长期以来支撑着知识组织研究的除了图书情报学中一些定律之外,还涉及与数学密切关联的理论和方法,尤其是系统科学中一些重要理论对知识组织研究与实践都有着很大的指导意义和解感作用。应该说,这些理论和方法促进了知识组织研究与实践的快速发展和完善。

长期以来,在图书情报领域,学者非常注重应用性研究、方法性研究、技术性研究以及解决图书情报工作中实际问题的研究,在理论研究方面或者说研究理论对研究的指导作用、理论对研究对象的科学阐释相对较少。例如,笔者对 CNKI 检索,发现学界在科学理论对知识组织

的理论价值、指导作用等方面的研究尚缺乏系统和深入。因此,迫切需要这样的探讨来改进和提升知识组织效率和水平,从而提升知识服务的效率。

早在 20 世纪 80 年代,刘植惠教授就对系统科学理论在情报学研究中的应用进行了系列讲座^[1-2],探讨了这些理论对图书情报学研究中的指导作用和有关应用问题;王知津教授等于 2009 年出版的《知识组织理论与方法》^[3],主要探讨了网络环境下知识组织的学科基础、基本原理、方法与技术,以及新概念、新理论和新工具等问题,但尚缺乏深入的知识组织的理论探讨;柯平教授等^[4]从知识资源论的角度研究了图书馆学基础理论体系的三层结构;周文芳等^[5]从系统论的角度论述知识组织的构建原理;李红春^[6]根据系统论的原理探讨了知识组织中的分类系统和主题系统;雷会珠等^[7]利用耗散结构理论等信息资源系统中的信息自组织与熵值变化进行了探索。

此外,在图书情报学理论与实践研究方面也有一些

* 本文系国家自然科学基金面上项目“面向知识服务的知识组织模式与应用研究”(项目编号:71273126)研究成果之一。

收稿日期:2013-11-29;责任编辑:魏志鹏

代表性的相关成果。如,武汉大学马费成教授在情报理论方面从事过大量的研究,其中最突出的贡献凝聚他2008年出版的专著《IRM-KM 范式与情报学发展研究》^[8]中,他从知识网络演化及知识网络中的群体行为等方面,提炼出情报学的六个基本原理(离散分布原理、有序性原理、相关性原理、易用性原理、小世界原理、对数透视原理),这些理论对知识组织研究有很大的指导作用;邱均平教授早在1988年就从文献计量学角度深入探讨了布拉德福定律、齐普夫定律、洛特卡定律等,并对这些理论在文献计量学中的应用做了大量阐述^[9],其后20多年他借助这些理论从事了大量的应用研究;中国科技信息所的化柏林博士提出了适合于以信息、知识、情报为对象的三动论(序化论、转化论、融合论)思想^[10],主要探讨了资源的分布、资源转化和资源互操作的理论思想;武汉大学董慧教授等从本体演化角度,结合逻辑学、图论等理论,构建了一个以本体、本体分子为核心的知识组织动态模型^[11];温有奎教授、焦玉英教授从范畴论角度对知识对象的进行了描述,并构建了范畴知识单元信息模型^[12]。

国外虽然也有大量关于知识组织理论与方法上的研究,但真正涉及科学理论与知识组织紧密联系的诠释也不多见。如,Judit Bar-Ilan等^[13]从认知理论出发,根据儿童特点研究了网络知识资源如何适合儿童,并构建了适合于儿童使用Web的知识构建模型;Alon Friedman^[14]结合知识表示、语言学模型和概念理论的知识组织方法,提出了一个融合两者优点的综合知识组织模型;Richard P^[15]基于Idea Collider信息解组合重构理论,提出了在知识检索和分类中,需区分相对于“常识知识”的“多元知识”的理念。

无论是国内还是国外,学者很少涉及科学理论的机理(尤其是系统科学中一些横断理论与方法)研究知识组织的理论价值,确切的说是技术与方法的研究,欲使知识组织更加科学合理,必须将科学理论融入知识组织研究中,用科学的理念指导知识组织研究与实践,从而促进知识组织理论框架的完善。

2 系统科学理论与知识组织

知识组织不同于信息组织,它更强调数据间的有机关联,通过数据间的联系从而形成相关知识。所以,知识组织更关注知识组织的结构和系统性,强调数据质量控制和数据流程管理,注重数据关联的平衡性和稳定性,重视知识间关系的协同作用,发现知识组织中的特殊规律

等,都需要相关的理论来支撑和解释。作为系统科学理论中的“老三论”、“新三论”对一切科学都有着重要的指导作用,同样对知识组织研究过程中一些问题和现象也能够进行科学的阐释和有效地引导。

2.1 系统论与知识组织

中国大百科全书认为:系统论(System Theory)是研究系统思想和系统方法的哲学理论。它包括三部分内容:系统的科学(或称数学系统论)、系统技术论(包括系统工程)和系统哲学^[16]。从图书情报学或信息学角度阐释系统论,认为它是研究系统一般模式、结构和规律的理论,是研究各类系统的共同特征,并用数学方法定量描述其功能,寻求并确立适用于一切系统的原理、原则和数学模型,是具有逻辑和数学性质的一门新兴科学^[17]。

从系统论角度出发,任何事物都具有整体性、关联性、动态平衡性,以及具有等级结构、时序性和可延续性。我们不仅要利用系统论的思想和方法去认识事物的规律和特征,更重要的是借助所认识的规律和特征去创造新事物、管理新事物,使事物达到最优化。

借助系统论思想,我们可以把知识服务系统中的知识资源看成一个整体,利用系统论的事物(知识资源)都具有关联性的思想,认识到孤立的数据不能成为知识。这一思想指导我们必须寻找知识或数据间的联系,并把这些联系上升为语义关系,保证经过关联的知识和数据的知识总量高于知识资源内个体知识之和。

系统论的系统内部事物的动态平衡性思想可以指导我们在进行语义网络构建时,注重语义网络的平衡性,在知识聚类过程中,要保持各类元素的平衡,切忌出现大量的孤立点(或元素很少的类)和稠密网络(元素很多的类),这样既可以避免在进行知识查询时命中很少的结果而不能满足需求,也可以避免命中大量的结果而加重用户筛选结果的负担。

在知识组织的分类系统和词表系统中,其体系的层次结构和词汇间的属分关系,充分印证了系统论中所认为的事物的等级结构之特征。而系统论的这一思想对知识分类体系的建立和词表构建都具有指导意义。同样,系统论中事物的时序性思想,对进行知识组织中的文献链结构和用户行为知识的组织都具有非常重要的参考价值。

2.2 控制论与知识组织

中国大百科全书对控制论(Control Theory, Cybernetics)的定义是:研究生命体、机器和组织的内部或彼此之间的控制和通信的科学^[18]。我们可以理解为,控制论是关

于工程系统、生命系统和社会系统等领域中有关反馈现象和控制机制共同理论的一门学科。控制论的主旨思想是,通过反馈机制,采用系统方法或过程方法对复杂事物或现象进行控制,通过调节使被控对象达到预定目标。因此,作为一门横断理论的控制论,其理论几乎可以应用于任何复杂过程。

在知识组织领域,控制论可以全面运用于数据的质量控制、数据和知识生产流程控制,并借助控制论基本原理解决知识组织过程中大量复杂问题。例如,在进行知识组织的架构时,需要针对系统的服务目标,借助控制论的系统分析方法,辨识数据或知识间的关系,使知识的关联达到最优化;在数据组织与操作过程中,控制论可以将人们对数据组织的计划、组织、激励、领导等职能与纠正工作中偏差紧密结合起来,使数据的管理工作形成一个有机的系统,以确保对数据质量的有效控制。

总之,强调系统方法、过程方法和微观调整的控制论,在知识组织的实施中发挥着重要作用。具体而言,控制论指导人们对数据结构和数据关联所进行的宏观控制思想,为实现知识服务目标提供了保证;过程化的控制思想对数据加工处理环节的数据质量控制、加工处理环节的时间、进度等控制,为知识组织工程的时间、进度和质量的控制提供了有效地理论与方法;为了对事物有效控制,控制论还强调反馈和调节功能,在知识组织的一些细节中,通过系统使用或测试的反馈信号,加强了对数据或知识间关系的调节,确保知识的组织能够提供更为有效的知识服务。

2.3 信息论与知识组织

信息论(Information Theory)是利用数学的方法来讨论信息本质、信息度量、信息识别和信息传递与交换共同规律的科学理论。正如中国大百科全书对信息论定义所述:信息论是研究信息的表示、传输、处理、存储和利用的理论科学。它的经典内容包括信息的数量定义(信息量)、信道容量、信源和信道编码以及信息失真率理论等^[19]。由此可见,知识组织所涉及的知识表示、知识传递、知识的处理与存储、知识的利用等研究与实践,都可以得到信息论思想的指导,并促进知识组织研究的升华。

信息论将信息描绘成在现实世界中客观、广泛地存在着的对象,信息空间充斥着大量杂乱无章的信息子,当信息子有规律的排布时,就会产生能量,这些能量经过传递就会使确定的事件发生。而知识组织正是遵循这一原理,将数据或信息(可比做信息子)有机联系,建立数据、

信息之间的语义关系后,就可能确定知识或产生新的知识,当这些知识传递给人们解决问题时,就激发出知识的能量。

知识组织的一个重要因素就是将数据进行关联后,使之能够进行推理获得知识,以满足知识服务的要求。信息论中所述信息子的摩擦碰撞产生能量,信息子的排布形式决定事件的性质,这一信息论机制非常雷同于知识组织中数据仓库的数据间的多维联系、本体结构数据间的语义关系,这种联系与关系完全可视为信息子的碰撞和排列,其产生的知识恰恰可视为解决问题的能量。

信息论思想告诉我们,知识系统内部存在着大量知识熵,为了控制知识熵的增长,我们必须将知识系统内的知识序化,也就是进行有效地知识组织。信息论有助于我们去探索知识组织的本质,去观察知识组织中的全新世界,让我们跳出传统的数据组织思想去建立一个全新的知识组织领域,促使系统提升了知识服务的能力,为我们构造出更加高效的知识服务系统提供理论基础。

2.4 耗散结构理论与知识组织

耗散结构理论(Dissipative Structure Theory)是指“寻求系统在非平衡状态下形成稳定有序结构的条件和机制,探索自组织过程的一般原理的理论”^[17]。在知识组织过程中,无论是知识聚类、分类,还是语义网络的构建,亦或数据仓库的组织等,均存在大量非平衡状态,而这些知识组织体系本身就潜在着稳定的有序结构,我们需要认识这些有序结构形成的条件和机制,完成这些知识组织结构的自组织过程。

典型的耗散结构需具备四个基本特征:系统的开放性、系统的非平衡性、系统与外界物质和能量的可交换性、系统的自组织性。也就是说,具有耗散结构的系统必定能和外界进行某种交换,孤立或封闭的系统不可能产生耗散结构;其次,耗散结构是远离平衡态的,平衡态或近平衡态都不可能将已有的有序走向更高级的有序,只有远离平衡的开放系统,通过与外界进行物质与能量交换时,所积聚的能量达到某个临界值将会引起突变,形成稳定有序的结构;第三,耗散结构具有自组织能力,这种自组织能力使系统具有柔韧性和鲁棒性,具备抗拒外界的干扰能力,确保系统走向自我完善。

因此,知识组织可视为一种耗散结构,如何借助耗散结构理论促进知识组织的稳定有序、提升知识组织的知识服务能力,需要将耗散结构理论充分运用于知识组织的研究中。知识组织是一个开放并远离平衡的系统,它需

要外部知识的介入,语义关系的链接激活了知识,随着内外部知识交换的频繁,新的知识大量增生,产生了质变,构成了自成一体的知识组织系统,形成了知识组织的耗散结构,为实现理想的知识服务提供了保证。

随着社会、文化和科学技术的发展,原有知识体系将不断被新生知识所取代,如果原有知识组织不更新,很快就会衰退到惰性的平衡,又从有序走向无序。所以,欲保证知识组织对知识服务的持续有效性,需要不断加强内外部的能量(数据、知识、关系)交换,增添新的知识和语义关系,使知识组织处在有序—无序—有序的转化过程中,保证知识服务系统始终处于高效的活動状态。

如此看来,知识组织虽然是一种耗散结构,但仍需要在耗散结构理论指导下不断提升和完善,只有这样才能在知识服务中发挥更大作用。

2.5 协同论与知识组织

协同论也称协同学(Synergetics)或协和学,是研究系统从无序到有序转变规律和特征的一种理论方法。它是研究系统中子系统之间怎样合作产生宏观的空间结构、时间结构或功能结构的^[16]。协同论认为,尽管各子系统之间千差万别,但它们在一个大的系统下,依然存在相互影响、相互配合以及相互作用的关系,也存在着相互干扰、相互制约等因素。因此,协同学研究的任务就是,探索人类社会各类系统共同本质,研究它们发展演变的共同规律,以及寻求它们之间的协作机理。

协同论所研究的对象在知识组织的研究中大量存在,信息聚类组织时依赖的是信息的主题,而同一主题的信息我们可以看作是这些信息所具有的共同本质。在关联信息的组织时,相关联的数据可能会相互影响、相互作用,甚至是相互制约。将信息的这种影响或制约关系加以描述并组织起来,可使信息上升为能够提供知识服务的知识,实际上这就是协同论中所阐述的系统从无序转变为有序。然而,协同学的思想更希望我们去寻找信息关联前后信息转化为知识的转变规律和特征。

协同学应用于知识组织告诉我们这样一个道理,孤立的数据或一个个数据个体可能价值很小甚至没有价值,但是数据“协同”(关联)起来,却能展现出数据中隐藏的大量知识,也就是说“协同”可以帮我们找到了数据中大量的“客观性质”。因此,在知识组织中,我们应当借助协同学思想,针对知识服务的目的和要求,建立探索数据间各类关系的机制,解决知识组织中的复杂问题,更有效地开展知识服务。

2.6 突变论与知识组织

突变是指事物发展过程中的间断或突然转换的现象,从理论角度探讨突变,就是用数学的方法研究自然界中事物发生演变出现的不连续现象。如,中国大百科全书认为:突变论(Catastrophe Theory)是描述自然界中发生的各种突变现象的数学理论^[20]。突变论使我们对人类许多社会现象的发展变化有了更深刻的理解,为我们战胜愚昧和无知提供了新的武器。突变理论使事物发生的质变或飞跃(革命性的变化)有了科学的解释依据。例如,互联网的急剧膨胀,网络地址资源远远不能满足互联网发展的需要,并制约了互联网的发展,使接近崩溃边缘的互联网处于极不稳定状态,需要一种网络技术的革命性变化(突变)使互联网达到一种新的平衡,所以新一代互联网技术完成了互联网发展历程中的突变。

信息组织同样也经历过面临从量变到质变的突变过程。从二次文献信息到全文信息、从文本信息到多媒体资源都带来信息存储技术、信息获取技术跃变和质变。目前,在大数据和云计算环境下,数据的组织更是发生了巨大的变化,大数据产业的出现,数据的计算已由GB级的计算上升到TB、PB甚至YB级的度量,信息组织架构已不再局限于某个局部,而是拓展到全球,形成大区域甚至全球协作战略。

在知识服务环境下,随着知识服务需求的不断提升,知识组织也将会发生突变,如何应对这种突变?如何控制需求变化带来的知识组织的突变?我们可以借鉴突变论中应对突变现象的数学方法,对知识组织由于知识服务需求带来的突变进行有效的控制和科学应对。首先,我们可以针对知识服务要求分析刻画知识需求参数和满意度控制参数;其次,进行知识标引并建立知识间的关联关系,构造知识间的势函数;第三,分析关联,构建实现平衡态的知识空间;第四,研究知识空间到知识需求参数间的投影,确定知识需求可能发生的变化。总之,在突变论的指导下,知识组织研究与实践可以迅速适应用户知识需求的不断升级和变化。

3 图书情报学相关定律与知识组织

图书情报学是研究事实、数据、信息、知识和情报的产生及其有效收集、组织、存储、传递、转换和利用规律,并运用科学技术手段对其进行有效的管理和利用的一门科学^[17]。因此,图书情报学中的定律注重探索信息、知识的活动特征,强调探究信息、数据中价值情报的规律。

著名的图书情报学基本定律(布拉德福定律、齐普夫定律、洛特卡定律)虽然是针对文献、词汇以及作者和文献关系给出的一般性规律,但同样对知识的组织有很大的指导意义。另外,在图书情报学中广为应用、阐述社会网络的小世界理论也对知识组织研究有着启示作用。

3.1 布拉德福定律与知识组织

布拉德福定律(Bradford's Law or Bradford's Distribution)是英国文献学家、化学家布拉德福先生^[17]创立的,主要描述文献分散规律的经验定律。布拉德福定律认为,如果按科技期刊刊载某专业论文的数量多少,以倒序排列,则可以分出一个核心区 and 相应的几个区,每个区的论文数量相等,则核心区期刊数量和相继区期刊数量成 $1:n:n^2 \dots$ 的关系($n>1$)^[17]。布拉德福定律用专业文章数量的多少确定期刊核心区的概念,虽然这一定律不一定适合当今选择核心期刊,但他的这一思想却对当今选择核心期刊仍有很大的指导作用。

布拉德福定律的主要思想告诉我们,文献有相对集中和普遍分散两个方面,这一思想对开展知识组织研究同样具有帮助作用。我们可以将文献分散原理应用于知识组织的信息聚类研究中,聚类核心区的思想可以指导我们确定阈值,从而帮助我们确定类的规模。另一方面,无论是知识仓库还是知识库,亦或本体数据组织,都存在着知识点的集中和分散规律。布拉德福定律可以指导我们去发现知识点的集中区,了解分散的知识点,这些信息可以帮助我们更科学有效的构造各类知识库的逻辑结构。

布拉德福定律强调,对于某一专业领域的论文通常分布在众多期刊中,虽然有少量的期刊刊载了较多的该专业领域论文,但仍存在大量的论文分散在众多的期刊中,而这一经验总结提示我们,每一学科都有本学科重要知识点的关键词,也有其他学科知识点的关键词散落在自己学科,这就提醒我们在进行相关知识库组织时,对散落在不同学科的知识点(关键词)做好学科关联处理,并确定关键词的主要学科,以保证我们可以通过关键词的学科关联关系发现交叉学科,寻找学科新的增长点,以此为学科规划提供知识服务。

3.2 齐普夫定律与知识组织

齐普夫定律(Zipf's Law)是美国学者 G.K.齐普夫于 20 世纪 40 年代提出的关于词汇在文献中出现频次的分布规律,也称词频分布定律、最省力法则、齐氏分布定律。齐普夫定律的主要思想:如果把文献中词的频率(f)从高至低递减排序,并用自然数给这些词编上等级序号(r),

则有 $f \cdot r = C$ (C 为常数)。这个定律与 80/20 规律有异曲同工之效。这条定律提示我们做任何事情尽可能以较少的付出获得较大的收益。当然,这条定律并不是暗示我们“偷工减料”,而是指导我们如何利用事物的客观规律,发现可能产生较大影响的并较为“省力”部分作为重点突破,以后再向全局突破。

在知识组织研究中,我们可以依照这条定律的原则,归纳出社会对知识服务的需求,并分析用户对各类知识服务需求的渴望程度,再根据分析结果针对相应的知识组织的难易程度、工作量大小,选择率先建设的知识组织系统。例如,根据用户使用信息系统的行为信息,分析大家关注的领域,并针对这些领域率先构建相关知识库。如,知识点链接知识库、用户行为推荐知识库以及相关主题的知识仓库,以后逐渐扩大知识库的构建领域。再如,该定律提示我们可以针对国家、地区或企业关注的领域,选择重点建设相关知识库,确保知识服务能够首先和国家、地区重点关注的问题紧密结合起来。

在具体的知识组织方面,我们可以根据齐普夫定律将知识组织体系中工作量大小以及各项数据组织工作在知识服务中影响的大小分别进行排列,并权衡工作量大小和知识服务中发挥作用大小选取先进行的工作,使知识组织工作达到效率和效果的平衡。例如,在语义关系的构建工作中,我们可以花费较小的力气建立实体间的简单关系,那些大量的复杂语义关系可以交给计算机实现,计算机根据人们设计的推理算法就可以完成各种复杂关系的组合。这恰恰是最省力法则的精髓。

3.3 洛特卡定律与知识组织

洛特卡定律(Lotka's Law)揭示了科学研究人员数量与所著文献数量之间的关系,认为发表两篇论文的作者数量是发表 1 篇论文作者数量的 1/4,发表 3 篇论文的作者数量是发表 1 篇论文作者数量的 1/9,如此类推,发表 N 篇论文的作者数量是发表 1 篇论文作者数量的 $1/n^2$ ^[17]。洛特卡定律又称倒平方定律,描述了科学生产率的经验规律,是美国学者 A.J.洛特卡于 20 世纪 20 年代提出的。

洛特卡定律并未有涉及知识组织领域,但该定律能否符合知识表达上的一些规律,确切的说,在学术成果的影响上是否遵循这一规律。我们做了实验,论文发表三年以后,没有被引用的论文是被引一次论文的 4 倍,被引 2 次的论文是没有被引过论文的 1/9,如此下去,基本符合洛特卡定律的倒平方定律。这一规律的出现为我们构建引文索引的数据结构,计算文章发表后的知识量提供了

一个非常有效的途径。

同样,洛特卡定律也为我们探索知识词库中使用词的规律,发现常用词汇以及分析各词汇在知识服务中的贡献(热词)提供了新的思路,我们曾经对文章中汉语词汇的分布情况进行了探索,发现汉语词汇在文章中呈现洛特卡现象^[21]。这一结论为我们更加科学有效地构建知识词库提供了新的依据,为知识服务系统在分析用户行为用词上提供更加科学的方法。在引文知识组织中,洛特卡定律为我们构建作者知识库,探求核心作者群提供了新的思路。

3.4 小世界理论与知识组织

小世界理论(Small world theory)也称为六度分割理论(Six Degrees of Separation theory),它来自数学领域的猜想和社会学领域的实验。六度分割概念始于1929年的一则小故事。匈牙利作家 Frigyes Karinthy 在这个小故事里论断,地球上任何两个人都可以通过一条由5位联系人组成的链条联系起来^[17]。也就是说,世界上任何两个人,均可以通过不超过6个人建立联系或搭上关系。即,一个人预想结识世界上任何一个陌生人,可通过自己所熟识的那个可能认识这个陌生人的人去结识他,这个熟识的人再通过他所熟识的、可能更靠近那个陌生人的人,依此联系不断进展,这个人可能最多经过6个人就可能和那个陌生人搭上关系。这个故事后来成为数学家的猜想,也得到社会学家的实验证明^[17]。六度分割理论让世界变“小”,所以也被人们称为“小世界理论”。

六度分割理论并非是准确的两个人之间需通过6个人才能建立联系,随着现代人类社会交往的密切,网络的普及而增添了更多的交往渠道,人类关系联系更加紧密,可能会使人类的联系链更短。但无论怎样,它却说明了一个问题,世界上任何事物都存在或近或远的联系,这种联系为我们建立信息、数据间的关联提供了理论依据,也为我们建立语义网实现推理性知识服务打下基础。

既然人与人之间通过联系链总能找到某种联系,那么数据、信息、知识也一定存在某种关联。小世界现象在学术文献间的引用网络、知识间的语义网络、词汇(主题词表或叙词表中)间词语关系网络、作者合作研究网络、研究主题的交叉网络等都有很深刻的体现。如果我们把每篇文献、每一个词汇、每一位作者、每一个主题等看作是一个网络的节点,并根据关联关系实现两者之间的联系,构成的网络就是一种小世界现象。

小世界理论的世界事物都存在关联的思想,为我们

在寻找数据间关联提供了依据,使我们清楚地认识到,数据、信息或知识在关联或传递中不断提升其价值。将数据、知识间的关系有效地组织,将会大大提升知识服务系统的服务效率。因此,我们可以借助小世界理论,建立复杂网络,并通过复杂网络中数据间的关联,进行数据挖掘和深度分析,发现新的知识,提供解决问题的素材、知识和方案。

当然,在图书情报学领域对小世界理论运用还处在初步阶段,在知识组织领域中还需要我们进一步研究。例如,小世界理论只给出了人与人之间的一般联系,并没有强调人与人之间的关系程度,实际上人与人之间的关系强度在人的交往中是非常重要的一个方面,可能决定两人合作事件的走向。这也提示我们,知识间的联系也有强弱关系,这种强弱关系可能会影响我们对知识网络中的知识发现,因此我们可以设定一个阈值屏蔽数据或知识间的一些弱关系,将强关系凸现出来,这也就是要求我们在知识组织方面如何区别关系的强弱问题。

再者,在小世界理论中,只强调人与人之间存在联系链,但人与人之间具体有什么关系并没凸显。例如,一位社会学家为证实小世界现象,进行一项连锁信件实验,让发件人把信件传给其他城市的陌生人,每一位发信人将信件发给自己熟识并认为最有可能认识目标人的人,最终信件经过不到6个人到达了陌生的收信人手中^[17]。这一实验说明小世界现象并没有改变人与人之间的原有关系,传递的信息并没有得到升华。因此,在知识组织研究中我们必须认识到,知识间的关系是普遍存在的,如果欲实现知识升华,在知识服务中发挥更大效能,需要建立知识间语义关系。

小世界现象在知识组织中还有一个方面需要注意,在人的传递网中,忽略了关系链中两人互逆问题。如,某人有两个朋友A和B,分别支撑着此人的前后朋友链,但A、B两人矛盾很深,如果此人将这两个朋友同时引入朋友圈,这条朋友链就可能崩溃。同样的道理在知识组织中也会出现,这就需要在知识组织时需要把各知识元间语义关系都要建立起来,尤其是将知识元之间的主要关系建立起来。

4 结语

科学理论是人类智慧的结晶,是高屋建瓴对相应科学领域的总结。严格地说来,科学理论并不像许多人所认为的那样,是完全抽象的东西,实际上它是实实在在的指

导着我们开展科学研究、支持着我们进行科学实验。本文通过对系统科学理论的阐释与理解,对这些理论在知识组织研究中指导作用、理论渗透以及应用价值给出了深刻的解答。如分析了系统论对加强知识组织的整体性、关联性、动态平衡性、等级结构性提出了理论依据;控制论为我们加强数据质量控制和知识的生产流程控制提供了科学有效地方法;信息论为我们探索信息的本质、控制信息熵的增长给出了科学的理论解释;耗散结构理论告诉我们,知识组织就是一种耗散结构,它是在有序—无序—有序的转化过程中使知识得到增长;协同论使我们认识到数据需要协同(关联)才能产生新的知识和发现隐藏规律;突变论解释了在知识存量不断增长、知识爆炸的突变环境下,如何有效地进行知识架构和知识组织。此外,文中还对图书情报学相关理论对知识组织研究的渗透进行了科学阐释,探讨了布拉德福定律、齐普夫定律、洛特卡定律、小世界理论对知识组织提供的理论思想和科学方法。

虽然,知识组织研究所涉及的理论与方法远远不止本文所述,但本文所论及的理论和方法对知识组织研究的指导作用和针对性更加显著,对面向知识服务的知识组织理念形成,提升知识服务系统水平都有很大的帮助。我们也期待着更多的学者加入到知识组织理论的研究中,促进和完善知识组织的理论与实践研究,从而提升我国知识服务的水平。

参考文献:

- [1]刘植惠.系统论、控制论和信息论简述及其在情报学研究中的作用[J].情报理论与实践,1988,(5):41-43;(6):41-44.
- [2]刘植惠.耗散结构理论、协同论、突变论概述及其在情报学研究中的作用[J].情报理论与实践,1989,(1):45-48;(2):45-47.
- [3]王知津.知识组织理论与方法[M].北京:知识产权出版社,2009:1-212.
- [4]柯平,王平.基于知识资源论的图书馆学基础理论体系研究[J].中国图书馆学报,2006,(2):9-14.
- [5]周文芳.论知识组织的系统论原理[J].情报资料工作,2007,(6):50-51,61.
- [6]李红春.系统论与知识组织[J].现代情报,2010,(6):20-23.
- [7]雷会珠,赵慧清.图书馆信息资源系统的自组织与熵值变化[J].图书情报工作,2010,(增刊):251-253,4.
- [8]马费成.IRM-KM 范式与情报学发展研究[M].武汉:武汉大学出版社,2008:341-364.
- [9]邱均平.文献计量学[M].北京:科学技术文献出版社,1988:109-195.
- [10]化柏林.情报学三动论探析:序化论、转化论与融合论[J].情报理论与实践,2009,(11):21-24,41.
- [11]董慧等.基于数字图书馆的本体演化和知识管理研究(II)——动态知识组织[J].情报学报,2009,28(4):483-491.
- [12]温有奎,焦玉英.基于范畴论的知识单元组织与检索研究[J].情报学报,2010,29(8):387-392.
- [13]Judith Bar-Ilan,Yifat Belous. Children as architects of Web directories: An exploratory study [J]. Journal of the American Society for Information Science and Technology,2007,(6): 895-907.
- [14]Alon Friedman. Concept Theory and Semiotics in Knowledge Organization[J]. Journal of Documentation,2011,67(4):644-674.
- [15]Richard P. Idea Collider: From a Theory of Knowledge Organization to a Theory of Knowledge Interaction [J]. Bulletin of the American Society for Information Science and Technology Volume,2011,37(4):43-47.
- [16]中国大百科全书编辑委员会.中国大百科全书(第二版)(第24卷)[M].北京:中国大百科全书出版社,2009:24-556.
- [17]《中国情报学百科全书》编委会.中国情报学百科全书[M].北京:中国大百科全书出版社,2010:303,78,191,13,149,308.
- [18]中国大百科全书编辑委员会.中国大百科全书(第二版)(第13卷)[M].北京:中国大百科全书出版社,2009:13-130.
- [19]中国大百科全书编辑委员会.中国大百科全书(第二版)(第25卷)[M].北京:中国大百科全书出版社,2009:25-161.
- [20]中国大百科全书编辑委员会.中国大百科全书(第二版)(第22卷)[M].北京:中国大百科全书出版社,2009:22-350.
- [21]王东波,朱丹浩,苏新宁.洛特卡现象在汉语词汇句法功能分布复杂度中的呈现[J].情报学报,2012,(3):297-302.

作者简介: 苏新宁,男,教育部长江学者特聘教授,南京大学信息管理学院教授,博士生导师,研究方向:信息智能处理与检索,信息分析与科学评价。