

数字资源知识整合及其发展方向研究

石 建

(第四军医大学图书馆
西安 710032)

石苗茜

(第四军医大学唐都医院
心内科 西安 710032)

赵 晔

(第四军医大学图书馆
西安 710032)

[摘要] 从数字资源知识整合的现实意义入手,对图书馆数字资源知识整合的基本概念、基本模式、基本技术发展方向等问题进行梳理与探讨。认为数据整合、信息整合和知识整合是图书馆数字资源整合过程中逐层递升的 3 个阶段,只有知识整合才能满足用户知识查询需求,是资源整合的最终目标。

[关键词] 数字图书馆;数字资源;知识整合;知识服务

Research on Knowledge Integration of Digital Resources As Well As Its Development Direction SHI Jian, Library of Fourth Military Medical University, Xi'an 710032, China; SHI Miao-qian, Department of Cardiology, Tangdu Hospital, Fourth Military Medical University, Xi'an 710032, China; ZHAO - Ye, Library of Fourth Military Medical University, Xi'an 710032, China

[Abstract] The paper starts with the realistic significance of knowledge integration of digital resources, discusses the basic concept, mode, technology development direction of knowledge integration of digital resources in libraries. It concludes that data integration, information integration and knowledge integration are three phases in integrating process of digital resources in libraries, whose levels are gradually increasing. Only knowledge integration will meet the requirements of knowledge query of users, it is the ultimate target of resource integration.

[Keywords] Digital library; Digital resources; Knowledge integration; Knowledge services

随着计算机技术、网络通信技术、数字资源管理技术的快速发展,在图书馆数字资源建设与发展过程中,遭遇到两大技术瓶颈:一是如何将网上海量的、随机的、缺乏逻辑关联的零碎信息,整合为具有内在联系的统一体,以解决信息有效利用的问题;另外则是如何将不同软硬件环境支持下的异质、异构数字资源进行有机整合,以实现相对独立数字资源的无缝链接。在这一背景下近年同行专家对资源整合进行了深入研究,在取得成效的同时,也遇到了一些需要进一步克服的难题。首先是数字资源的数据整合问题,只解决了异构数据库中信息实体的合并问题,未能有效地反映信息实体间的诸

联系。这样从整合的数字资源系统中获得的还是信息,并不是知识的集合,也不能满足用户深层次的知识需求。如何实现知识的有效获取和利用,成为数字资源知识整合的终极目标,也正是当前数字资源整合优先研究的领域。

1 数字资源知识整合的基本涵义

1.1 知识整合的本质

数字资源知识整合概念体系的形成,是一个动态的、不断完善的过程,基点则是人们对图书馆数字资源服务需求层次的提升。知识整合是图书馆数字资源整合的高级阶段,是以有效利用知识为目的,以知识管理为基础,知识组织方法为依托,在已有数据整合与信息整合的基础上,充分展示知识

[修回日期] 2010-12-01

[作者简介] 石建,副研究馆员,主要研究方向为知识组织与检索,发表论文 30 余篇。

体系、知识内容、知识概念及概念关系的一种整合方式。知识整合是一个动态的过程,最终目标是能够有效利用知识,需要对不同载体上的显性与隐性知识,实现有效挖掘与重新组织。其中包括摒弃无用的知识,将不同系统中的知识融合起来,使融合的知识更具柔性、条理性、系统性,必要时进行知识重构,以此来形成新的知识体系。知识整合的主要对象为知识体系,其中包括体现知识主题概念及概念间的各种关系;也包括知识体系中的知识元、知识点、知识链等。从知识体系的结构看,其中的知识元、知识点、知识链并不处在同一等级。在一个体系中是重要的,在另外一个体系中可能就没那么重要。所以,将知识结构进行等级划分是知识整合的必然要求。知识体系不是各种知识内容的汇集,而是相互间具有某种逻辑联系的层次结构。知识整合的本质其实就是通过一定的技术手段与相应的知识组织系统(例如本体、主题图等),将存在于知识体系中、能反映知识结构的概念体系,组织成一个能够有效获取知识的网络系统^[1-2]。

1.2 知识整合与数据整合、信息整合的差异

数据整合、信息整合、知识整合是图书馆数字资源整合系统发展的3个不同阶段,其区别主要是数字资源整合目标上的差异,其次还有数字资源整合模式以及整合技术方法上的差异。按照用户对数字图书馆服务需求层次及深度的不同划分:数据整合满足的是用户一般性的低层次服务要求;信息整合满足的是用户特殊需要的“中等层次”的服务需求,是对数字整合服务功能进一步的改进与提升;知识整合又是对信息整合服务功能体系进一步完善和优化,能满足不同用户在特定环境、特定时间、特别需求状态下的服务,满足的是高层次的知识服务需求。按照整合的对象划分:数据整合的对象是异构数据源,即分布的、异构的和自治的数据库,由于用户不能实现一站式访问和查询,数据整合主要用来解决不同种类的数据管理系统之间的互操作、相互融合问题;信息整合的对象是信息实体之间的关系,由于数据整合只解决了多源异构数据的合并问题,内在的逻辑关系和其它链接未能有效揭

示,信息整合则是通过相应的机制、标准实现数字资源的全方位整合,满足特定的服务要求;知识整合的对象是内在于知识体系结构中的概念及概念关系,目的是实现知识的有效获取,满足人们对知识的特定需求。另外,整合的方式也不尽相同:数据整合采用的是基于数据仓库集成、中间件集成、XML数据模型集成等方式;信息整合采用基于信息链接、信息门户的整合方式;知识整合的特殊性在于采用知识概念关系的整合方式。

1.3 知识整合的特征

知识整合必须与知识服务相结合,脱离知识服务要求的整合属于无的放矢。因此,在整合的过程中必须体现动态性与系统性要求。其中动态性主要体现在知识的转化和转移,并能完成显性知识的组织和隐性知识的挖掘任务,同时还要以动态的知识资源网络为基础,实现个性化、专业化、增值性的知识服务。这就要求知识整合必须适应外部环境的变化,当外界环境发生变化时,知识整合的内容、范围和方式都将做出适当调整。

2 数字资源知识整合的基本模式

2.1 基于知识本体的知识整合

数字资源知识整合,不能简单地理解为把不同载体、不同类型、不同系统中的知识资源聚集在一起,而是涉及到对知识的描述、组织、处理、整序、检索、服务等方方面面。这些都需要在一定的技术模式支持下得以实现。建立在知识本体基础之上的知识整合模式主要依赖于知识本体拥有的某些独特属性和功能。这些属性和功能使其在知识表示、知识结构与逻辑关系的揭示、知识查询系统的性能等方面更具优势。也正是这些优势促使了它在知识整合过程中得到充分运用。知识本体其实是知识领域中的概念及概念之间关系的规范化描述。这种描述是规范的、明确的、形式化的,同时又是可共享的。“明确”意味着所采用概念的类型和它们应用的约束,实行明确的定义;“形式化”指知识本体是计算机系统可读与可识别的;“共享”即反

映知识本体应捕捉该领域中一致公认的知识。基于本体的知识整合目标,就是要捕获相关领域的知识,提供对该领域知识的共同理解,确定该领域内共同认可的知识概念词汇,并从不同层次的形式化模式上,给出这些知识概念词汇和词汇间相互关系的明确定义,以便于知识共享。目前,在知识本体的构建过程中,已有丰硕的成果,比较有效的方法是以现有本体为基础构造出新本体。分类表、词表等传统的知识体系属于简单的知识本体,元数据方案也是本体的简单表现形式^[3-4]。

2.2 基于知识网格的知识整合

美国信息学家 Fran Berman 较早提出的知识网格概念是这样的:知识网格是一个智能互联环境,能使用户或虚拟角色有效地获取、发布、共享以及管理知识资源,并为用户提供所需要的知识服务,辅助实现知识创新等。我国知识网格研究领域的领军人物诸葛海研究员^[4]对知识网格技术的特征做了进一步的诠释,即人们可以通过单一语义入口,获取和管理广泛分布的知识,而不需要清晰知道知识的所在位置。同时,在知识网格环境中,知识不是以静态的形式存贮,而是能动态演化,并能自动更新。即知识网格中的知识服务在使用过程中可以不断自动演化与改进。知识整合的目标是为用户提供有效的知识服务,这与知识网格的内涵与特征相吻合。知识网格在知识整合中的应用在于为知识整合提供统一的技术平台。在此基础上,知识整合借助于知识网格强有力的知识处理能力及海量的同构知识,实现知识的自动积累,进而达到知识挖掘、知识推送,为用户提供智能导航。

2.3 基于主题图的知识整合

主题图是目前国内知识整合领域的热门话题,是多种技术融合后而展示出的一种新兴数字资源处理技术。通过对国内已有的相关文献研究发现,人们主要是将主题图作为一种复杂的元数据、数字化的知识组织方式、智能化的知识索引技术、模型化的知识表示和导航技术、灵活的分布式资源链接技术,在数字资源知识整合领域加以应用,其范围遍

及知识库的建构、知识分类与检索、知识概念关系的揭示、学科知识导航、语义网等多个方面。按照国内学者马文峰教授^[3]对主题图涵义的界定,主题图的主要功能特征,就在于可充分描述和展示知识概念间的相互联系与层次,可定位某一知识概念所在知识体系中的具体位置。

从整体功能上看,主题图作为近年新生的知识整合技术,除吸收了索引、叙词表、分类法等思想,还吸收了数字整合、信息整合的某些技术方法并有所发展。主题图在知识整合中的应用方法可以总结为:构建领域本体、知识关联量化、主题图合并和主题图可视化等。通过主题图,可将抽象的知识内容组织成一个有坐标概念的、立体化的多维知识地图,并形成结构化的语义网络。借助链接技术,将用户引导到相关的知识资源,从而能从总体上把握知识的层次结构,方便地定位知识资源的出处,并可获得与之有密切关系的一切知识点^[5]。

3 数字资源知识整合技术的发展方向

3.1 更加准确与全面

随着知识整合技术的不断完善与进步,用户对数字资源知识整合的服务要求越来越高,不仅需要清楚不同物理环境和虚拟环境下的异构知识源的分布情况,还要求展示和把握不同知识体系中各知识簇、知识点、知识元之间的相互关系,从而更好地获取知识。知识整合的技术应用,在建立知识关联、整合异构知识源的过程中,进行透明访问、无缝链接以及更加深入、全面的知识要素揭示将成为必然。

3.2 整合工具需要创新与完善

面对具有多样性、复杂性特征的知识源,传统的知识整合工具在知识整合过程中显得力不从心。人们要求对传统的知识组织或整合工具改造与创新。虽然知识组织和整合系统的类型越来越多,所显现的功能也越来越强大,但知识整合体系本身决定了它对技术的依赖,有关知识整合工具及相关技术的研究始终是知识整合体系研究的主要内容。知

识整合工具多以用户为中心进行设计研究,在以用户为中心的基础上,设计和研究新技术、新工具,研究如何使知识概念和知识术语的描述更易理解,更易表达概念与概念间的关系,开发能够完全显示知识组织体系中丰富语义关系的可视化工具^[6-7]。

3.3 研究重心依然是构建领域本体

在网络环境下知识整合实现的重要基础条件之一就是知识本体。知识本体的表现则是知识共享概念模型的形式化及规范化说明。一个知识本体就是关于某一领域知识概念的一套规则及对规则的清晰描述,还包含了知识概念、每个知识概念的属性,同时还有知识概念属性的限制条件。知识本体的构建方法直接决定了本体对知识概念的表示与逻辑推理能力。研究者出于对不同学科领域知识的差异和对实践的不同考虑,导致构建本体的过程也各不相同。而且目前尚没有一套标准的知识本体构建方法,所以构建领域知识本体依然是重心。

3.4 语义标引是研究与应用的重中之重

建立在本体之上的语义标引,是数字资源知识整合语义化处理和语义查询的基础,同时也是改进与修正传统标引方法弊端的出路,在知识组织、知识管理和知识检索等方面,具有比较重要的理论和实践应用价值。目前,以 Web 内容为主的语义标引工具,在国外的一些大学和研究机构尚在研究与开发之中,国内对本体标引技术的研究也在进行。为实现知识整合的精确要求,理想的本体标引系统应该实现标准化与通用性,应该具有知识表达能力,支持多语言标引,自动或半自动的知识获取,以及支持本体的学习与进化^[8-10]。

3.5 趋于多技术融合应用

在当前语义网、知识网格大环境的逐步形成过程中,整合技术越来越成熟,整合模式除了早期基于 OPAC、跨库检索、资源导航、超级链接外,还出现基于知识本体的知识整合模式,所涉及的技术领域涵盖了网格、P2P (Peer to Peer)、SOA (Service Oriented Architecture)、Web 服务技术等。对同

一问题的解决往往采用不同的技术方式,以求最佳效果。按功能放大原则,多技术融合才能取长补短,协同使用。因此,知识整合过程中的多技术融合应用是技术应用研究的主要方向^[10]。

4 结语

综上所述,知识整合是数据、信息整合发展到一定阶段的产物,是用户对数字图书馆需求层次由低级到高级逐步提升的结果。知识整合水平的提高离不开知识整合技术的发展,知识整合技术的发展促进知识整合更加系统、更加深入、更加准确、更加全面。为了满足用户知识服务最大化的要求,就必须对知识整合理论进行系统、全面的梳理,明晰数字资源知识整合的方向,把握影响知识整合未来发展的关键技术,这对于数字资源知识整合的可持续发展具有十分重要的意义。期待设计出更加科学、结构更加完备的知识整合系统,为数字图书馆的知识服务开辟更加广阔的发展空间。

参考文献

- 1 黄传慧,李娟.我国图书馆数字资源整合研究[J].图书与情报,2009,(4):66-69.
- 2 田玉晶.数字资源整合核心领域研究进展[J].农业图书情报学刊,2009,(3):35-39.
- 3 马文峰,杜小勇,卢晓惠.基于知识的资源整合[J].情报资料工作,2007,(1):51-56.
- 4 陆静.基于知识网格的知识整合研究[J].农业图书情报学刊,2009,(10):19-22.
- 5 黄晓斌.网络主题图的原理及应用[J].情报理论与实践,2004,(3):305-307.
- 6 姜永常.知识构建的基本原理研究(下)——知识构建的技术支撑[J].图书情报工作,2009,(6):106-109.
- 7 刘雪竹.数字资源整合研究综述[J].现代情报,2008,(2):4-6.
- 8 焦丽.数字资源整合研究回顾与展望[J].科技信息(科学教研),2008,(15):289-290.
- 9 陈军.数字资源整合研究热点分析[J].科技情报开发与经济,2007,(35):6-8.
- 10 蒲筱哥.数字资源整合技术研究现状述评[J].情报探索,2009,(3):28-31.