

ZHISHI LIANJIE JIQI FUWU YANJIU

知识链接 及其服务研究

■ 曾建勋 著



科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

国家社会科学基金项目(08BTQ022)

知识链接及其服务研究

曾建勋 著



科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

图书在版编目(CIP)数据

知识链接及其服务研究/曾建勋著. —北京:科学技术文献出版社, 2012. 8
ISBN 978-7-5023-7385-6

I. ①知… II. ①曾… III. ①图书情报学 IV. ①G250

中国版本图书馆 CIP 数据核字(2012)第 144919 号

知识链接及其服务研究

策划编辑:周国臻 责任编辑:周国臻 责任校对:赵文珍 责任出版:张志平

出版者 科学技术文献出版社
地址 北京市复兴路 15 号 邮编 100038
编务部 (010)58882938, 58882087(传真)
发行部 (010)58882868, 58882866(传真)
邮购部 (010)58882873
官方网址 <http://www.stdp.com.cn>
淘宝旗舰店 <http://stbook.taobao.com>
发行者 科学技术文献出版社发行 全国各地新华书店经销
印刷者 北京时尚印佳彩色印刷有限公司
版次 2012 年 8 月第 1 版 2012 年 8 月第 1 次印刷
开本 850×1168 1/32 开
字数 187 千
印张 7.75
书号 ISBN 978-7-5023-7385-6
定价 28.00 元



版权所有 违法必究

购买本社图书,凡字迹不清、缺页、倒页、脱页者,本社发行部负责调换

前 言

随着信息技术和网络技术的迅速发展,用户的知识信息需求、科研创新环境、学习行为习惯,以至于生活处世方式都发生了深刻变革。从认知的角度来看,通过“链接”的方式获取所需的相关信息,更符合用户通过“联想”获取信息的本性,用户梦想的是一个覆盖全部学术信息资源的信息系统,所有信息互通互联,信息获取简捷准确。一种基于多语种、多媒体、移动的、跨系统、语义的知识共享空间正在形成。以知识链接为核心的知识服务逐渐成为用户获取知识信息的重要手段。通过知识链接,可以建立一个各种信息资源互相链接的图书馆信息环境,包括图书、期刊、会议论文、专利、科学数据、网页和各种产品信息等。

目前,数字信息资源向两极发展,一方面,各种类型的数字信息资源向聚合方向发展,汇聚成更大的多类型的数字资源聚合体,并能够根据内容、根据需求更加自由地组合;另一方面,数字资源打破了传统物理介质资源的“容器”,可以以更加微小的内容单元呈现。近年来,国内外建设了大量的科学引文数据库,但是引文链接应用于知识服务领域的研究尚不系统深入;虽然进行了大量的语义网、本体方面的研究,但语义关系运用于知识资源组织的实践尚不成熟实用。新型的泛在知识环境迫切需要加强知识链接及其服务研究,实现知识组织、知识构建、知识融合、知识导航、知识评价等知识服务活动,推进用户体验的知识服务。

知识链接及其服务研究,既可以丰富现有知识组织的研究理

论与成果,又提供认识网络信息知识资源内在特征的新视角。在实践层面上,为知识链接系统的构建和知识链接服务提供设计支撑和实践指导。

知识链接是发现、利用知识间的各种关联关系,将具有同一、隶属、相关关系的知识单元按照一定的需要有序地联系在一起,形成序列化或结构化的知识集合,继而构成知识网络的一种行为,是以知识联系为基础的信息组织^①。知识链接基于综合性的数字资源,除了期刊论文,还包括图书、会议文献、学位论文、专利、标准等多类型文献;除了文献之间的引证链接,还包括文献中所包涵的内容抽象,如知识元链接、主题参数链接以及科研实体等的链接和评价。

本书首先阐释了知识链接的概念以及内涵,在知识单元演变分析基础上,阐述了知识链接的产生背景与发展历程,论述了知识链接的表现形式及其功能,继而从知识属性出发,分析知识单元间的关联关系,论述知识链接构建模式;依据现有的知识管理理论、信息计量学理论和社会网络理论,提出知识链接理论基础。并着重探索了知识链接构建方法和基础,论述了知识关系的匹配与关联揭示方法。在此基础上,设计知识链接系统框架,探讨系统技术实现方案。并以中国科学技术信息研究所知识链接门户为对象,从系统构建方案与服务实现两个方面进行知识链接应用的实证研究。最后,进一步分析开放式知识链接服务体系框架及其协同服务策略。

全书由7部分组成:

(1)知识链接研究概述:内容包括知识链接的概念和内涵、知识单元的演变与知识链接的产生背景、研究现状、发展历程、表现形式、作用和功能。

^① 朱震远. 网络信息检索环境中知识链接的设计——基于语用和用户行为研究的视角[J]. 图书情报工作, 2010(16): 130-133

(2)知识链接原理研究:内容包括知识链接理论基础、知识关联关系分析、知识链接构建模式等。

(3)知识链接构建方法及其基础:内容包括面向知识评价的规范文档建设、科学文献中机构要素词典的构建、面向知识链接的概念知识体系构建、DOI与参考链接机制构建、知识关系的匹配与揭示方法。

(4)知识链接系统构建:内容包括知识链接系统的建设目标、建设原则、设计思路、系统架构、系统功能、技术实现。

(5)知识链接系统实证:内容包括中信所知识链接门户系统开发背景与建设目标、系统实现方案、系统的服务业务组织。

(6)开放式知识链接服务体系:内容包括开放式知识链接服务体系的构建、开放式知识链接服务方式、知识链接协同服务策略。

(7)结论与展望:在研究总结基础上,提出下一步研究工作展望。

本书研究重点主要体现在以下方面:

(1)从知识单元演变过程中,提出知识链接的主要链接对象,即基于科学文献、基于知识元、基于科研实体和基于知识要素的知识链接。并阐明了知识链接正不断朝着开放式、动态化、个性化、网络化方向发展。将会更多地实现面向科研环境的开放式跨界链接、基于用户行为的个性化推荐链接、基于本体的语义关系链接、面向科研评价的知识要素链接和基于知识元的知识组织链接等新型的知识链接方式。

(2)通过对知识链接产生背景和需求的调研,提出知识链接的基本概念、内涵、特征、功能和理论基础。从文献、作者、机构、基金、文献来源、学科主题等科研实体和知识要素着眼,实现科研实体和知识单元的关联揭示。

(3)依据知识链接的类型不同,基于不同的知识单元和对象,

面向不同的用户任务,提出构建科研关系网络、引文网络、文献网络、知识元网络、行为关系网络和知识语义网络等不同知识链接构建方式。

(4)提出在知识链接中要进行基础构建,如参考链接机制、知识元库与主题图建设、知识要素的抽取标引、规范文档建设、科研实体词典编制等,以规范各种科研实体,挖掘相关知识要素,实现知识链接。

(5)提出知识链接系统的层次结构,论述知识链接服务方式、服务体系框架和开放式协同服务策略,并结合中信所的具体实践,进行实证研究,设计基于引文的知识链接门户系统和服务应用方案。

总之,知识链接原理依据知识间的特有属性和关联关系,知识链接方法则得益于信息技术和文献标准的发展。知识链接以知识组织为基础,以知识发展脉络为主线,反映了知识内容和概念关系,显示了知识内在联系、知识发展的继承关系和创新主体间的学术关系等。从有关知识链接发展历史的追溯到对目前发展趋势的判断,可以看出,无论知识链接理论研究,还是知识链接实践进展都在不断地进化。知识链接及其服务研究下一步将重点探索基于本体和知识元的概念-实例关系、基于人与人的学术关系和行为关系,推进知识链接系统构建的实证研究,发展科研关系网络、信息行为关系和概念关系网络方面的知识链接研究。

目 录

1 知识链接研究概述	1
1.1 知识链接的概念与内涵	1
1.2 知识链接的产生背景	5
1.3 知识链接研究现状	11
1.4 知识链接的发展历程	25
1.5 知识链接的表现形式	28
1.6 知识链接的作用与功能	29
1.7 知识链接的发展趋势	31
2 知识链接原理研究	36
2.1 知识关联关系分析	36
2.2 知识链接构建模式	43
2.3 知识链接的理论基础	56
3 知识链接构建方法及基础	67
3.1 面向知识评价的规范文档建设	67
3.2 科学文献中机构词典的构建	78
3.3 面向知识链接的概念知识体系构建	88
3.4 DOI与参考链接机制构建	98
3.5 知识关系的匹配与关联揭示	105
3.6 科研关系网络构建	118

4 知识链接系统构建	123
4.1 知识链接系统的建设目标与原则	123
4.2 知识链接系统的设计思路	126
4.3 知识链接系统的架构	130
4.4 知识链接系统的功能	134
4.5 知识链接系统的技术实现	138
5 知识链接系统实证	160
5.1 中信所知识链接门户	160
5.2 基于知识链接门户的高被引分析应用	179
5.3 基于知识链接的资源聚合应用探索	189
6 开放式知识链接服务体系	198
6.1 开放式知识链接服务要求	198
6.2 开放式知识链接服务体系框架	201
6.3 开放式知识链接服务方式	207
6.4 开放式知识链接服务推进策略	211
7 结论与展望	216
7.1 研究总结	216
7.2 研究展望	219
参考文献	221
后记	237

1 知识链接研究概述

1.1 知识链接的概念与内涵

1.1.1 知识链接概念

什么是“知识链接(knowledge linking)”,目前学术界还没有统一明确的定义。国外学者多以“Citation Linking”^①和“Reference Linking”^②为主题对其进行研究,并没有提出完全与知识链接对应的专有名词。而国内学者在吸收国外研究理念的基础上,从不同角度对知识链接的概念进行了不同界定:

王越(2003)^③认为,节点和有向边构成了知识链接的结构,可以用来表达两个资源间的有序关系和语义关系,从而实现 Web 资源的语义表现,使其可被计算机所理解。而知识链网络是用知识链替代现有 Web 超链接结构的语义互联网模型,其中节点表示资

① S. Hitchcock, L. Carr, S. Harris et al. Citation Linking: Improving Access to Online Journals [EB/OL]. [2011-02-25]. <http://journals.ecs.soton.ac.uk/acm-dl97.htm>

② Van de Sompel H, Hochstenbach P. Reference Linking in a Hybrid Library Environment Part 1: Frameworks for Linking [J/OL]. [2011-3-7]. D-Lib Magazine, 1999, 5(4), http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt1.html

③ 王越. 基于耗散结构理论的知识流分析[J]. 科学管理研究, 2003, 21(3): 86-89

源,有向边表示类型化的知识链,节点和边组成一个有向网络,此有向边表达了 Web 资源及其语义关系。

孙洪波(2004)^①认为,所谓知识链接,就是采用类似超链接的方式,将互相关联的一些知识链接在一起,使原本各自独立的知识关联到一起,形成由点到线,由线到面的知识结构,wiki(维基)是一个典型的例子。

赵蓉英(2007)^②认为,知识链接包括知识元链接、引证链接和作者链接等,它实际上是构成知识网络的一种技术措施、手段和行为,以知识之间的关联关系作为构建依据,实现从一个信息物体链接到另一个信息物体,一般特指期刊文章之间的链接。

贺德方(2008)^③认为,组成多功能的行业性知识服务系统的知识链接方式是按照知识组织的规律,从一个信息点链接到另一个或多个相关联的信息点,其价值不仅体现在通过“关联方式”直接而准确地获取信息,还在于为特定行业或企业建立一个各种信息资源相互链接的知识服务环境,涵盖专业期刊、科学数据、论文作者、专利信息和经营信息等。

周晓英(2010)^④认为,知识链接就是根据知识体之间的关联关系将它们联系起来的过程、方法和技术。广义的知识链接可以指知识体之间点到点,点到面以及面到面的关联;而狭义的知识链接主要是指知识体之间严格的点对点的关联关系,连接知识点和微观的知识集合。知识之所以能够链接和需要链接是因为知识网络的存在,知识链接是形成知识网络的技术和条件,知识网络是知

① 孙洪波. 构建知识库(四)内容管理与维护[J]. 软件工程师,2004(9)

② 赵蓉英. 论知识网络的结构[J]. 图书情报工作,2007,51(9):6-10

③ 贺德方. 知识链接发展的历史、未来和行动[J]. 现代图书情报技术,2005(3):12-15

④ 周晓英. 知识链接的发展阶段、发展动因和类型特征分析[J]. 图书情报工作,2010,54(12):36-37

识链接的目标和结果。

王菊^①认为,知识链接是指将不同的知识节点,根据其上下文背景和知识组织规则,实现语义层次上的知识关联的一种行为。其中,知识关联是知识链接的本质与基础。知识关联就是指大量的知识点之间存在的知识序化的联系,以及所隐藏的、可理解的、最终可用的关联。

作者认为,知识链接是以知识联系为基础的信息组织,是信息组织的深化。知识链接体系能够显示知识元之间或者说知识信息群之间的网络、结构、互动、交叉、演化或衍生等诸多类复杂关系^②。现代信息技术尤其是网络技术的发展是知识链接基础,它以知识组织为基础,通过语义关联,把知识体系连接为清晰的知识网络结构,为知识检索和知识共享提供知识资源,可以方便用户即时获取有效的知识。

1.1.2 知识链接内涵

知识链接作为一种新型的信息组织方式,正在从面向用户转到面向机器应用,从文献计量转到情报计算,从信息描述转到知识展示/表现,从语义隐含转到语义揭示,从以主题为中心转到以概念关系为中心,从信息表示转到智能推理方面发展。主要具有以下特征:

①以知识组织为基础。知识组织是对资源内容概念及概念间的相互关系进行描述与组织的机制(工具),能够支持对资源对象按照知识内容进行组织和描述,并支持基于语义和推理的知识检索。

②以语义标引为中枢。借助语义标引工具,对资源对象进行

① 赵蓉英,王菊. 知识链接及其实现模式探讨[C]. 新信息环境下的知识链接与知识服务研讨会,2010,1

② 曾建勋,赵捷,吴雯娜等. 基于引文的知识链接服务体系研究[J]. 情报理论与实践,2009,32(5):1-5

概念分析、分类、标引、描述和处理,使相对独立、没有语义的信息集合形成具有语义关联的知识组织系统,是实现基于知识链接的基础。

③以知识构建为手段。知识构建是在信息构建的基础上,通过一定的知识表示技术,将基于知识的整合系统按照一定方式,清晰有序地在一个统一的界面上展示出来。

1.1.3 知识链接对象

知识链接对象非常丰富,抽象地讲凡是有实质性涵义的数字对象和知识单元都可以作为知识链接对象。具体地说,既包含文献及其外部特征,如期刊论文、图书、会议文献、学位论文、专利、标准等,还包涵文献资源中的内容抽象,如知识元、主题、参数,以及科研实体、知识要素等。归纳起来主要有以下几种类型:

科研实体:包括科研主体(科研人员、科研机构、国家地区)、科研活动(科学会议、科学实验、科研项目)、科研条件(科研方法、科学仪器、科研经费、科学期刊)、科研产出(科研成果、专利、标准)以及科学文献、科学主题等。

知识要素:包括相关技术、项目、配方、剂量、产品及其参数、指标、相关的图形表格等。

科学文献:包括期刊、学位论文、会议录、科技报告、图书专著、专利、标准等,同时,还包括这些科学载体中的科学论文或论著中的引文。

科学文献中既包含标题、作者、来源出处、作者机构、正文等相关项目,又包含论文或引文中的知识单元,需要用知识单元来表达引文本身的內容特征,以及构成引文的相关来源及外在特征的科研实体等等。

科学主题:包括各种概念、术语、专有名词、实例等表述事物及其属性的知识单元。

1.2 知识链接的产生背景

1.2.1 知识链接的演化

知识单元是在知识管理中可以对知识进行独立、自由、有效地识别、处理与组合的基本知识单位^①。知识单元是构建知识链接的基础,知识链接的目的就是通过建立知识单元之间的有效关联关系实现知识服务、知识获取和知识创新。

知识单元的演变经历了从文献单元到信息单元继而到知识单元三个重要阶段,在这一过程中,图书情报机构也逐渐将知识关联的研究从文献单元和信息单元过渡到知识内容本身上,由此推动了知识链接的产生,如图 1-1 所示。

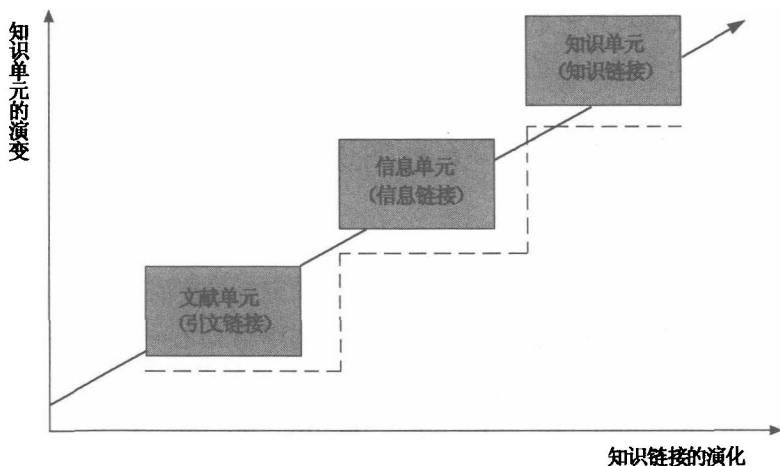


图 1-1 知识单元的演变与知识链接的产生

^① 文孝庭. 知识单元的演变及其评价研究[J]. 图书情报工作, 2007(10):

联机数据库出现之后,图书情报机构主要借助目录、文摘和索引来提供知识信息的检索服务,这时的链接是以文献单元为实体的引文链接;网络信息技术的飞速发展使以信息单元为实体的链接服务成为图书情报服务的主要内容,尤其是在知识管理理论被引入图书情报领域以后,实现信息服务向知识服务的转变,提供以知识单元为核心的知识链接服务成为图书馆情报机构面临和需要解决的主要问题。

(1)以文献单元为实体的引文链接

以文献单元为实体的引文链接是指以知识的载体——文献——作为知识管理的基本处理单元^①。依据是文献单元之间存在着相互联系,通过各种文献单元之间的关联关系可以形成复杂的知识网络。

① 布什(Vannevar Bush)的“memex”机制。以文献单元为实体的引文链接思想最早是由 Vannevar Bush 提出的,Bush^② 指出要确认两个概念相联系的作品之间可操作性的链接可以利用所谓的“联想索引”(associative indexing),并描述了一个谓之“memex”的理论性工具。基于这一构想,出现了许多与知识链接相关的尝试。

② 加菲尔德、普赖斯和科学引文网络。SCI 的引文链接分析在某种意义上可以看作是 Bush 所构想的理论模型的一定程度实现^③。SCI 是加菲尔德科学引文索引^④(利用论文之间的相互引用

① 文孝庭. 知识单元的演变及其评价研究[J]. 图书情报工作, 2007(10): 72-76

② Vannevar Bush. As We May Think [EB/OL]. [2011-2-25]. <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/3881/>

③ 陈兰杰. 知识链接理论与实践的三次嬗变探究[J]. 图书情报工作, 2010, 54(12): 46-47

④ Garfield E. Citation indexes for science; a new dimension in documentation through association of ideas [J]. Science. 1955(122): 108-111

关系来探讨和分析科学活动)和普赖斯引证网络^①理论在实践中的应用。SCI的发展,使通过引文链接来实现科学出版物之间知识链接的意义在国际科学学与科学计量学界得到了广泛的共识。

③布鲁克斯(Brookes B. C.)与认知地图。布鲁克斯关于知识地图的思想对知识链接的发展及其理论和方法的形成也起到了积极的推动作用^②。认知地图就是分析文献中的逻辑内容,按知识的逻辑结构找出人们在创造和思考中相互影响和联系的结点,找出人们思维中相互影响的连接点,按照知识的逻辑结构,把它们直观地表示出来,就像地图一样展现知识的有机结构。

引文链接遵循科学研究间承前启后的内在逻辑,通过揭示科学技术之间引证与被引证关系,展示科学文献中内容主题之间的相互关联,由此形成了一种新的检索途径,这一检索途径的最大优点在于降低检索结果的不相关性,因此在科学研究中具有重要意义。

(2)以信息单元为实体的信息链接

以信息单元为实体的信息链接是以知识的某些属性特征,包括外形特征和内容特征(比如文献的主题词、关键词、概念等)作为知识管理的基本单元^③。与文献单元相比,信息单元不但反映了文献单元的基本特征,而且是在文献单元基础上进行了提炼、加工、组合和转换,因此成为知识的重要来源。信息单元通过提炼、加工、组合、转换等过程形成知识,信息单元是联结文献单元和知识元的一座桥梁。

① Price D. J. S. Networks of scientific papers [J]. science, 1965, 149 (3683): 510-515

② Brookes B C. Foundations of information science (Part IV) [J]. Journal of Information Science. 1981(3):3-12

③ 文孝庭. 知识单元的演变及其评价研究[J]. 图书情报工作, 2007(10): 72-76

①超文本信息链接。Gardner 认为,超文本信息链接是一种附加增值服务,并把这种思想首先介绍到心理学领域。基于这一思想,King 和 Roderer 于 1978 年提出一种构想,即把科学文章的超文本链接结构作为电子档案的长期目标之一。1990 年, Frank G. Halasz 提出给超文本建立文摘链接的模型^①。一般来说,信息单元之间不会出现类似于参考文献链接的关系,即用一个信息单元去解释或参考另外一个信息单元的关系。但是描述文献的信息单元,如关键词、主题词等,很可能就是一个知识元的名称或者与知识元有着密切关系。建立信息单元之间的超文本信息链接可以引导用户去阅读知识元,有助于用户更好地理解信息单元。

②数据库链接。网络技术的发展使以数据库为依托的服务也发生了变革,主要表现为数据库服务商在自己的数据库内甚至数据库之间建立知识链接。从实现链接的方式看,主要包括 5 种形式:从目录链接到全文、从书目检索结果链接到文章全文、从引文链接到被引文章的文摘或全文、从文章链接到后来的引文(前向链接)、从一个数据库链接到其他属于不同机构的数据库。此外,如果将论文中的基因系列与基因数据库中的记录建立链接,将论文中二维的化学式与三维的化学结构链接,或者将论文中的实验结果描述与科学数据结合等可以提供更智能化的链接方式,这也是目前数据库链接发展的方向。数据库链接原理如图 1-2 所示。

(3)以知识单元为实体的信息链接

知识单元是指文献中相对独立的、表征知识点的元素,是知识管理、知识计量与知识评价的最小单元。知识单元之间的关联关系是人们利用语义网发现新知识的重要途径,是进行知识链接的基础。一般来说,以知识单元为实体的知识链接的实现需要

^① Halasz F, Schwartz M. The Dexter Hypertext [J]. Communications of the ACM, 1994(37):30-39