

●毕强 韩毅 刘昆

知识服务——现状、进展及挑战*

摘要 国外对基于语义 Web 技术知识服务研究集中在数字图书馆领域。研究主要围绕用户需求的语义表达,落实于知识服务的“以人为本”上。国内学者对此研究主要集中在语义门户、网络化知识组织系统和基于语义 Web 检索理论方面。基于语义 Web 技术开发知识服务所面临的挑战来自三个方面:知识服务统一标准规范;知识服务主体的语义互联策略;知识服务用户需求的语义描述策略。表 3。图 2。参考文献 28。

关键词 数字图书馆 语义 Web Web 环境 Web 技术 知识服务

分类号 G251.6

ABSTRACT In foreign countries, the researches on knowledge services based on semantic web technologies are limited in the field of digital library. Researches are mainly focused on semantic representation of user demand and the “user-centeredness” of knowledge services. In China, researches are mainly focused on semantic portals, networked knowledge organization system and the retrieval theories based on semantic web. The authors point out three challenges: common standards for knowledge services, semantic interconnection of knowledge services and semantic representation strategies for knowledge service user demand. 3 tabs. 2 figs. 28 refs.

KEY WORDS Digital library. Semantic Web. Web environment. Web technologies. Knowledge service.

CLASS NUMBER G251.6

现代信息环境变化的基础特征是数字化、网络化。信息资源空间数字化和信息交流空间数字化的发展,使信息资源及其存在形态发生了变化,信息获取主体的心理期望与行为发生了变化,情报服务工作的内容发生了变化。这些变化要求信息服务机构具备新的服务模式和新的核心能力,从而使情报服务工作从满足文献需求向满足知识需求发展;从提供信息向提供知识化的资源系统发展;从提供信息的物理入口向构建到知识的智能入口发展;从关注物的因素向关注人的因素发展等^[1]。但是应该看到,走向知识服务,实现知识“服务化”和服务“知识化”的现状不容乐观。究其原因,一是知识服务理论研究滞后于实践,正在成为知识服务的制约因素;二是知识“服务化”的实践,要求知识服务主体能够识别用户需求的语义;三是服务“知识化”的实践,要求知识具有语义特性。为了推动知识服务理论研究更好地指导服务

实践,本文在对目前知识服务研究现状与进展归纳分析的基础上,探讨了基于语义 Web 开发知识服务所面临的挑战。

1 知识服务研究现状及实践进展

1.1 Web 环境下知识服务研究现状及实践进展

20 世纪 90 年代后期,任俊为将知识服务引入国内的图书情报界^[2],拉开了中国图书情报界研究知识服务的序幕。张晓林博士进一步揭示了知识服务概念的本质,即知识服务是以信息知识的搜寻、组织、分析、重组的知识和能力为基础,根据用户的问题和环境,融入用户解决问题的过程中,能够有效支持知识应用和知识创新的服务^[3],这一本质揭示获得国内学者的共识。随后,国内学者对知识服务这一主题展开了深入研究,如表 1 所示。

* 本文是国家自然科学基金资助项目(批准号 70673029)研究成果。

表1 国内学者对知识服务研究一览

研究者	研究成果
任俊为	将知识服务引入国内的图书情报界。
张晓林	揭示知识服务概念的本质。
姜永常	对知识服务与信息服务进行了对比研究,指出知识服务与信息服务在内涵及服务方式等方面的本质差异 ^[4] 。
黎艳	探讨了现代信息环境和知识经济对图书情报工作的深层次影响,认为信息服务向知识服务转变有必然性,并提出“层次化”的参考咨询服务模式、知识管理服务模式、专业化信息服务模式 ^[5] 。
李家清	对知识服务模式进一步研究,认为还应提个性化定制服务模式和垂直服务模式 ^[6] 。
尤如春	探讨了基于网络环境的高校图书馆知识服务的策略,从知识服务策略的内涵入手,探讨知识服务的软件策略和硬件策略,对知识服务的实践运作有一定的指导意义 ^[7] 。
曾民族	认为应面向知识服务重构信息服务模式,并指出构建科技知识服务基础平台的主要方向是构建语义网和发展知识服务的关键技术,但没有具体阐述如何实现知识服务 ^[8] 。

近年来,许多信息服务机构为应对数字环境都在进行着知识服务的探索和实验,努力超越传统模式,拓展新的服务内容和形式。如知识导航服务、数字参考咨询服务、咨询公司模式的团队知识服务、各种个性化的定制服务、网络学习中心以及社区信息枢纽等。其中比较典型是上海图书馆推出的一个旨在向各专业技术和科研人员提供高质量专业知识导航的新型知识服务项目^[9]。另一个是 CNKI 知识服务平台,它使用知网节点主动推送知识,根据一个知识点进行知识扩展,所扩展的信息通过概念相关等方法提示知识之间的关联关系,达到知识扩展的目的,有助于新知识的学习和发现,帮助实现知识获取,知识发现^[10]。这些研究,为开发知识服务提供了视角和基础。

1.2 语义 Web 环境下知识服务研究现状及实践进展

基于现有的 Web 技术由于受限于技术能力,只有信息资源定位描述而无信息资源内容含义描述,因而在服务的广度、深度、自适应程度和灵活性等方面都不能满足用户对知识发现、知识内容挖掘、以及根据其内在特征和价值进行鉴别、关联、重组、识别和创造新知识的需求,不能提供真正意义上的满足用户知识需求的知识服务。语义 Web 技术的诞生,为挖掘、链接和分析信息资源中的知识内容,直接支持用户利用信息解决问题和创造知识以及提供用户工作任务空间中的知识捕获、发现、组织和交换等服务方面提

供了契机^[11],为知识服务的发展提供了新的平台。

以语义 Web 为代表的语义技术,以其严格的逻辑理论基础和标准化的技术路线,正逐渐成为未来 Web 信息系统的一项支撑技术。语义 Web 采用的基本路线是明确表示数据的语义,定义领域相关的本体,使得 Web 实体(例如用户, Agent 等)之间能交互、动态地发现资源、提取知识和问题求解^[12]。语义 Web 的知识服务研究是在开放的、大规模的 Web 语义互联环境中,探讨如何有效地共享、组织和管理分布的语义资源,构建动态的、跨平台的、面向社区的交互模式,提供协作的知识环境以支持问题求解和决策支持。

语义 Web 环境下实现知识服务的过程分三部分:语义理解、知识管理和知识服务。知识库是实现知识服务的基础和核心。知识库提供的是语义理解中最终将要提供给用户的结果。和互联网的状况相同,人类的知识结构和容量都在飞速膨胀,知识库也需要有良好的适应能力。知识服务可以利用语义分析的结果,对知识库进行概念级的服务,对用户需求提供高相关度的结果。

近年来,国内外许多学者认识到语义 Web 技术对知识服务有巨大影响,都在进行着知识服务的探索和实验。国外对基于语义 Web 技术的知识服务研究集中在数字图书馆领域(如表 2)。这些研究主要围绕用户需求的语义表达上,落实于知识服务的“以人为本”上。

表2 国外学者对语义 Web 环境下知识服务的研究与探索

研究者	机构	研究成果
Brendan Rousseau	荷兰埃丁霍温科技大学	认为应使用语义 Web 的层次结构来构建数字图书馆的语义互联基础和 信息表达框架,从而为用户提供知识服务 ^[13] 。
York Sure, Rudi Studer	德国 Karlsruhe 大学 AIFB 研究中心	认为语义 Web 技术在数字图书馆中的知识库和对象语义描述起到了 核心作用,有助于解决分布式异构知识库的互操作以及用户接口和人 机交互、信息空间中用户的统一模型、个性化服务及用户交流协作 ^[14] 。
Deborah L, McGuinness	美国斯坦福大学	讨论了利用增加语义解释层的办法处理语义 Web 环境下用户提出 的问题 ^[15] 。
F. Giunchiglia, P. Shvaiko	意大利 Trento 大学	就语义匹配进行了深入研究 ^[16] 。
R. Guha	斯坦福大学知识系统实 验室	研究了语义检索问题 ^[17] 。
S. Deil ^[18] , B. Hammond	IBM 阿尔马登研究中心、 美国佐治亚大学 LSDIS 实 验室	提出了一系列语义标注的新方法 ^[19] 。
A. Shet, C. Ramakrishnan	美国佐治亚大学 LSDIS 实 验室	报告了他们关于本体驱动的信息检索、分析与整合应用系统的研究 情况 ^[20] 。

国内学者对基于语义 Web 的知识服务的探索和 于语义 Web 的检索理论等方面(如表3所示)。实验主要集中在语义门户、网络化知识组织系统和基

表3 国内学者对语义 Web 环境下知识服务的研究与探索

研究者	研究成果
张振海	提出利用 Ontology 解决知识获取、知识重用和知识共享,但未给出具体实现的方法与模型 ^[21] 。
颜端武	借助于智能检索提高知识服务水平,给出了基于语义 Web 和数据挖掘的智能推荐检索机制,其他与知 识服务相关的研究主要集中在对基于语义 Web 检索的理论探讨 ^[22] 。
余正涛	以数字图书馆领域个性化服务为例,以空间向量模型表示用户兴趣和资源特征。借助于构建的领域本 体和“知网”知识词典对向量进行概念上的扩展,形成用户和资源特征概念空间向量,并通过向量相似度计 算寻找最优的资源,有效地将用户和资源虚拟组合在一起 ^[23] 。
曾铮	认为语义 Web 和知识管理两者相结合可以为知识服务的研究与发展开辟新道路,阐述了语义 Web 技术 在知识服务中的角色,探讨了语义 Web 与知识服务系统相结合的应用前景,论述了基于语义 Web 技术构造 知识服务系统的主要实现模式 ^[24] 。
饶弋宁等	利用自然语言处理和理解技术,提出并实现了一种可以对网页中的中文信息进行处理,获取并存储知 识,有自我扩展特性和支持中文智能搜索功能的知识库系统模型 ^[25] 。
张佩云等	提出了一种基于本体的知识检索框架。在这种框架里,知识检索主要由语义检索和基于规则的推理检 索两部分组成。通过对检索方法的研究分析和算法实现,开发了一个基于本体的文档知识管理系统,有一定的 智能性 ^[26] 。

上述探索与实验,虽然还难以明确描述知识服务 现、组织和交换等问题解决的程度,但为基于语义实现过程及支持用户工作任务空间中的知识捕获、发 Web 开发知识服务提供了广泛的可能性和发挥空间。

2 基于语义 Web 技术开发知识服务所面临的挑战

知识的分享和利用是知识获取和掌握知识的基本目的。基于语义 Web 开发知识服务就是要改进知识的利用过程,提高知识利用和共享效果。这主要包括两个方面:让用户能够高效地使用知识;能够让机器计算、存取和使用知识,进而产生新知识。为高效地使用知识需满足两个条件:让知识服务主体中的知识更加标准化、多样化和增值;在第一个条件的基础上提供一个形式化、标准化、表达能力丰富且可扩展的知识推理模式。基于语义 Web 来开发知识服务面临以下三个方面的挑战。

2.1 知识服务的统一标准规范

基于语义 Web 的知识服务统一标准规范包括:统一的知识库标注、统一的用户需求标注、统一的知识操作以及统一的知识视图,如: Dublin Core/Metadata; XMI, RDF, 专业 XML; DAML + OIL, SOAP, UDDI, WSDL, Z39. 50; METS, OAI, SCORM 等。

语义作为知识服务系统之间共享内容和进行交流以及知识服务系统与用户之间进行交流的基本要求 and 最终目的,其互操作涉及多个层次、多种复杂因素。现有的技术可以对具体的、个别的知识服务系统进行语义互操作开发,但很难建立普遍意义的语义互操作规范。对于实现知识服务的语义互操作来说,语义描述规范、形式编码语言和互操作协议三者缺一不可^[27]。

(1) 语义描述规范:指对知识服务系统中数字对象的语义和结构等的描述规范。对于知识服务系统中的资源以及用户需求的表达,不同的抽象程度、或者关注不同的方面,就可能有不同的描述规范。一般把这些规范分为与数字对象一定数据格式(微观数据结构)、有关的数字对象规范、与资源集合和服务有关的资源集合描述规范(通常包括资源内容描述、服务描述和代理 Agent 描述三个方面)和与知识、资源组织体系有关的规范(如分类法、主题法、各类词表等规范或半规范的主题框架 Scheme)。在知识服务系统中,这些规范兼有语义、语法、语用结构描述的功能。

(2) 形式编码语言:指知识服务系统语义描述规范中使用的形式编码语言。目前各类语言大都基于 XMI,但 XML 只是提供了一种结构化的形式和手段,并不能满足各类领域应用多样化的编码要求,也不能保证所编码的信息内容在结构或语义上的互操作,而

且基于 XML 的不同功能的编码语言还在不断产生,其中大多数和语义的编码直接有关,如 XML/XMLS, RDF/RDFS, OWL 等。在基于语义 Web 的知识服务系统中,形式编码语言关系到知识资源的语义以及用户需求语义能否得到忠实的表达。

(3) 互操作协议:指实现支持知识服务互操作所需要的协议或者协议体系。互操作协议是对于知识服务系统之间以及与用户之间互操作(跨库检索、元搜索等)实现所需应用的协议或者体系框架。目前因特网的大多数应用协议都可以用来传达语义信息,但是高层互操作协议一般对语义信息的编码进行了具体的规定,所以特别适合进行语义互操作。这些协议中的高层协议建立在下一层协议基础之上,相互之间也有着复杂的关系,适合不同的软件和网络系统架构,如 Z39. 50, SRU/SRW, OAI 等。

2.2 知识服务主体的语义互联策略及应用模型

实现服务“知识化”要求,即在统一知识服务规范标准基础上,制定基于语义 Web 的知识服务主体的语义互联策略。因为知识服务主体中的数字对象具有层次结构性,且带有多义、模糊、噪声等各种不确定性,所以用计算机处理这类对象所面临的主要问题之一就是它们进行形式化描述和度量。因此有必要建立一种多层次、多通道、多模式、多可能的数字对象语义模型,应用有效的知识标识语言(RDFS, DAML + OIL, OWL, Loom, OKBC, CycL 等),对知识库进行语义分析、语义描述和语义表达,通过单一语义映像和多语义互联,将知识库从多个不同类型的语义空间变换并整合到一个统一的语义空间,构建基于语义 Web 的知识库的语义互联模型,实现多个知识服务主体间的语义互联与语义互操作(如图 1 所示)。

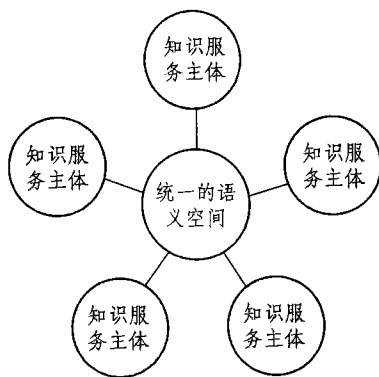


图1 多个知识服务主体间的语义互联与语义互操作

从图1看出:在知识服务主体的数字资源空间中,通过建立本体到数据的映射,本体在局部数字资源中构成一个语义空间。语义空间屏蔽局部数据在结构上的不同,使得用户或程序可以面对数字对象的内容,支持基于语义的数字对象的访问。在语义空间中,通过建立本体之间的映射,多个语义空间形成物理上分离但内容聚合的社区,这就是统一语义空间。在统一语义空间中,社区成员之间能相互理解。统一语义空间能屏蔽本体之间的异质性,协调“多个”资源的共享,是对多个语义空间的有序化^[28]。

实现多个知识服务主体之间的语义互联和语义互操作涉及以下问题:①为语义 Web 提供语义级共享的知识库本体;②应用有效的知识标识语言对知识库中的资源进行语义描述,并兼顾语义描述的兼容性、丰富性和适用性及语义描述的规范性;③数字对象和数字对象之间、数字对象和描述性元数据之间以及不同类型数字对象之间的关联及其组织方式、技术支撑和协议;④知识库之间的语义互联。

2.3 知识服务受体(用户)需求的语义描述策略及应用模型

实现知识“服务化”所面临的关键问题就是用户需求的语义分析与描述。基于语义 Web 的用户需求语义描述策略是指对用户需求进行语义分析的基础上,应用有效的知识标识语言,进行语义描述,构建基于语义 Web 的用户知识需求语义描述模型。如图2所示,通过对用户需求的语义分析与语义描述,然后在统一语义空间进行语义匹配,并将结果返回到用户。

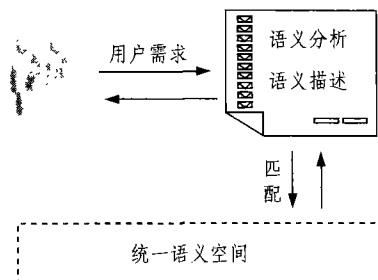


图2 用户需求的语义描述与统一语义空间的匹配

实现对用户需求进行语义描述涉及如下问题:①用户需求的语义分析,即用户需求的学科领域分析与识别,包括学科领域静态范畴识别、情景领域识别、背景领域识别和对语言学识别;②基于语义 Web 的用户需求语义描述策略;③基于语义 Web 的用户需

求语义描述模型的构建。

实现知识服务要解决主体(知识服务提供者)与受体(用户)交互中主体“读懂”用户需求的问题和一个知识服务主体“读懂”另一个知识服务主体的问题。需要技术、知识内容和用户三方面的相互作用,用户创造或增值知识,技术提供知识组织和与用户交互的手段,基于知识内容的互动满足用户需求,而语义是用户与知识服务主体以及知识服务主体之间共享知识内容和进行交流的基本要求和最终目的,知识服务理想状态即通过语义 Web 技术实现主体与受体之间的语义互操作。

语义互操作涉及多个层次、多种复杂因素。现有的网络技术可以对具体的、个别的知识服务主体进行语义互操作开发,但很难建立普遍意义的语义互操作规范。基于语义 Web 技术,可以根据信息对象间的语义结构和语义关系制定知识服务策略,强调信息对象的语义维度。在语义 Web 环境下,知识服务系统将自身所获得的语义信息存储在知识库中,并利用知识库中的知识引擎对语义内容进行重构,形成关于某个特定学科领域全面的知识网络。在用户开始浏览检索活动时,系统将根据用户和系统交互方式与交互内容等一系列变量,综合判断用户的语境,分析和描述用户需求的语义,并将经形式化描述后的需求传递给知识库中的推理引擎,由后者进一步判断用户需求,进而将与用户需求相匹配的知识内容以可视化的方式呈现在用户面前,并在用户反馈的基础上修改检索策略,达到知识和用户互联互通。

参考文献

- 1 周晓英. 数字时代情报学学科发展动向. 数字图书馆论坛, 2006(10)
 - 2 任俊为. 知识经济与图书馆的知识服务. 图书情报知识, 1999(1)
 - 3 张晓林. 走向知识服务:寻找新世纪图书情报工作的生长点. 中国图书馆学报, 2000(5)
 - 4 姜永常. 论知识服务与信息. 情报学报, 2001(5)
 - 5 黎艳. 信息服务向知识服务转变的探析. 图书情报工作, 2003(2)
 - 6 李家清. 知识服务的特征及模式研究. 情报资料工作, 2004(2)
 - 7 尤如春. 论网络环境下的知识服务策略. 图书馆, 2004(6)
 - 8 曾民族. 构建知识服务的技术平台. 情报理论与实践, 2004(2)
- (下转第 50 页)

- 27]. <http://www.gmw.cn/01gmr/2007-02/27/content-559079.htm>
- 4 邹志仁,黄奇,孙建军. 情报研究定量分析. 南京:南京大学出版社,1992:264~274
- 5 张跃,邹寿平,宿芳. 模糊数学方法及其应用. 北京:煤炭工业出版社,1992:218~227
- 6 刘普寅,吴孟达. 模糊理论及其应用. 武汉:武汉大学出版社,2002:75~83
- 7 彭祖赠,孙温玉. 模糊(Fuzzy)数学及其应用. 武汉:武汉大学出版社,2002:122~131
- 8 张卫国,徐维军. 企业创新系统的模糊评价模型及实证研究. 数学的实践与认识,2006(1)
- 9 丁日佳,信春华,陈佳鹏. 科技成果转化国际标准潜力的模糊综合评价模型. 世界标准化与质量管理,2006(10)

李丕仕 中国矿业大学图书馆副研究馆员。通讯地址:江苏省徐州市。邮编 221008。

(上接第45页)

- 9 杜也力. 知识服务模式与创新. 北京:北京图书馆出版社,2005
- 10 [2006-09-21]. <http://www.cnki.net>.
- 11 张晓林. 数字信息环境下的图书情报服务:挑战、应变与再造. 四川图书馆学报,2002(4)
- 12 Dieter Fensel, Mark A Musen. The semantic web: A brain for humankind. IEEE Intelligent Systems, 2001, 16(2)
- 13 Brendan Rousseau, Ronald Rousseau. Some Ideas Concerning the Semantic Web. Library and Information Service, 2002(8)
- 14 York Sure, Rudi Studer. Semantic Web Technologies for Digital Libraries. [2006-02-24]. http://www.aifb.uni-karlsruhe.de/WBS/ysu/publications/2005_sw_for_dl.pdf
- 15 Deborah L McGuinness, Paulo Pinheiro da Silva. Explaining answers from the Semantic Web: the Inference Web approach. Journal of Web Semantics, Vol. 1 No. 4, October 2004
- 16 F. Giunchiglia, P. Shvaiko. Semantic matching. The Knowledge Engineering Review journal, 2004, 18(3)
- 17 R. V. Guha, R. McCool, E. Miller. Semantic search. Proceedings of the 12th International World Wide Web Conference, Budapest, Hungary, ACM Press, 2003
- 18 S. Deill et. al. Sem Tag and Sem Seeker: Bootstrapping the Semantic Web via automated semantic annotation. Proceedings of the 12th International WWW Conference, Budapest, Hungary, May 2003
- 19 B. Hammond, A. Sheth, K. Kochut. Semantic Enhancement Engine: A Modular Document Enhancement Platform for Semantic Applications over Heterogeneous Content, in Real World Semantic Web Applications, V. Kashyap and L. Shklar, Eds., IOS Press, December 2002
- 20 A. Sheth and C. Ramakrishnan. Semantic (Web) Technology in Action: Ontology Driven Information Systems for Search. Integration and Analysis. IEEE Data Engineering Bulletin, Special issue on Making the Semantic Web Real, December 2003
- 21 张振海. Ontology 与网络计算. [2006-04-24]. <http://tpi.cnki.net/meeting/tpi40/pdf/ontology.pdf>
- 22 颜端武等. 数字图书馆中基于语义 Web 和数据挖掘的智能推荐检索机制研究和探讨. 见 数字图书馆网格应用模型研究. 武汉:湖北人民出版社,2005
- 23 余正涛等. 基于本体的个性化领域信息服务. 计算机工程, 2005(5)
- 24 曾铮. 基于语义网技术构造知识服务系统. 情报学报, 2005, 24(3)
- 25 饶弋宁等. 支持智能搜索的自扩展知识库模型的研究和设计. 计算机应用研究, 2006(6)
- 26 张佩云等. 基于本体的知识检索研究及实现. 情报学报, 2006(5)
- 27 数字图书馆语义互操作空间. [2006-07-20]. <http://my.donews.com/keven/2006/04/11/interoperationphase/>
- 28 黄河. 语义 Web 中知识服务的研究(博士论文). 中国科学院计算技术研究所, 2006

毕强 吉林大学管理学院教授, 博士生导师。通讯地址: 长春市。邮编 130022。

韩毅 空军航空大学馆员, 博士研究生。通讯地址同上。

刘昆 吉林大学图书馆馆员, 硕士研究生。通讯地址同上。

(来稿时间: 2006-12-25)