

# 面向知识服务的知识组织工具： 现状与未来<sup>\*</sup>

谢 靖<sup>1</sup> 钱爱兵<sup>1</sup> 韩 普<sup>2</sup> 苏新宁<sup>3</sup>

<sup>1</sup>(南京中医药大学经贸管理学院 南京 210046)

<sup>2</sup>(南京邮电大学经济与管理学院 南京 210023)

<sup>3</sup>(南京大学信息管理学院 南京 210093)

**【摘要】**从知识服务角度出发,对知识组织工具进行探讨,将其划分为基础知识与知识架构类工具、知识关系建立类工具、知识处理及展现类工具。从这三个分类的角度,对现有的知识组织工具分别进行探讨。基础知识与知识架构类工具主要通过系统化知识组织提供知识点推送服务,充当语义级的知识库角色;知识关系建立类工具则通过对知识点的关联建立提供深层次的知识关联服务,在此基础上通过与基础知识及知识架构类工具相结合,提供相应的知识推理服务;知识处理及展现类工具是知识抽取、关系建立、呈现等工作中所涉及的实用工具,这些工具以用户需求为中心完成对知识的重组。最后,对知识组织工具的未来进行展望,提出多维融合、网络化、分布式、个性化、动态化是其发展方向。

**【关键词】**知识组织 知识服务 知识组织工具

**【分类号】**G25

## Knowledge Organization Tool Catering to Service: Today and Future

Xie Jing<sup>1</sup> Qian Aibing<sup>1</sup> Han Pu<sup>2</sup> Su Xinning<sup>3</sup>

<sup>1</sup>(School of Economics and Management, Nanjing University of Chinese Medicine, Nanjing 210046, China)

<sup>2</sup>(College of Economics & Management, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

<sup>3</sup>(School of Information Management, Nanjing University, Nanjing 210093, China)

**【Abstract】**From the perspective of knowledge service, this paper divides knowledge organization tools into three groups: tools for basic knowledge acquisition and systematization, tools for knowledge relationship establishing, and tools for knowledge processing and visualization. Tools for basic knowledge acquisition and systematization render push services for knowledge elements. Tools for knowledge relationship establishing mainly work on the identification of knowledge relationship and support inference services together with tools for basic knowledge acquisition and systematization. Tools for knowledge processing and visualization are used in the procedure of knowledge extraction, identification and visualization. After the procedure, these tools render user-oriented services by knowledge reorganization. Finally, the paper discusses future trends of knowledge organization tools and points out the characteristics of future tools.

**【Keywords】**Knowledge organization Knowledge service Knowledge organization tools

收稿日期: 2013-06-18

收修改稿日期: 2013-07-30

<sup>\*</sup> 本文系国家自然科学基金项目“面向知识服务的知识组织模式与应用研究”(项目编号:71273126)、高技术研究发展计划(863计划)“以科技文献服务为主的搜索引擎研制”(项目编号:2011AA01A206)和江苏省社会科学基金项目“网络舆情监控预警模式研究”(项目编号:10TQC008)的研究成果之一。

## 1 引言

知识服务,指以信息和知识的搜寻、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之中,提供能够有效支持知识应用和知识创新的服务<sup>[1]</sup>。在此视角下,可以将其理解为以用户需求为中心,对知识内容进行提炼并提供解决方案的过程即为知识服务。知识组织是指对知识客体所进行的诸如整理、加工、揭示、控制等一系列组织化过程及其方法<sup>[2]</sup>。由此可见,在知识组织与知识服务之间存在密切关联,知识组织是知识服务的前提和基础,而知识服务是知识组织的目标与延伸。

知识组织的概念最早由 Bliss<sup>[3]</sup> 提出,国内学者刘洪波<sup>[4]</sup> 于 1991 年将其引入国内,并就知识组织与图书馆工作进行了阐释。马费成<sup>[5]</sup>、王知津<sup>[6]</sup>、蒋永福<sup>[2]</sup> 等学者也在 20 世纪 90 年代对知识组织进行了理论与方法的阐述。进入 21 世纪以来,图书情报领域对于知识组织的讨论,从理论研究逐步拓展至实践,知识组织的相关研究主要集中于知识组织系统(NKOS),其主要标志为 1998 年美国 ACM 数字图书馆大会上“网络知识组织系统工作组”成立。在 2000 年,欧洲数字图书馆会议也将知识组织系统列为独立的工作组。

Boisot<sup>[7]</sup> 认为知识服务是集成了知识管理、知识组织、知识经济、知识市场这些概念而正在形成的新概念。张晓林<sup>[1]</sup> 明确提出图书情报工作的核心应从提供文献服务转向知识服务。毕强<sup>[8]</sup> 指出知识服务主体需满足用户语义需求,同时知识也应具有语义性。

参考张晓林<sup>[1]</sup> 对于知识服务的阐述,知识服务不应局限于提供文献服务,而是应该覆盖从知识查询、组织、分析到为用户提供所需知识的应用,这与知识组织揭示事物本质并实现序化的目标是相辅相成的。在传统纸质图书馆向数字化转化的过程中,知识的载体由纸质文献逐步转向数字化、网络化的电子文献,知识服务的内容也从提供文献服务转为提供用户所需的语义知识。在这种背景下,知识组织不仅立足于对数字化知识的表示,更强调对用户知识需求的知识表示,即为用户提供知识服务成为知识组织的推动力。因此,在数字化、网络化背景下,知识组织与知识服务紧密相关。知识组织是知识服务的框架和基础,知识服务是知识组织的导向与推动力。同时知识组织工具的形式

也更为多样、作用也更为广泛。以提供服务为标准,本文将知识组织工具划分为基础知识与知识架构类工具、知识关系建立类工具、知识处理与展现类工具,如图 1 所示:

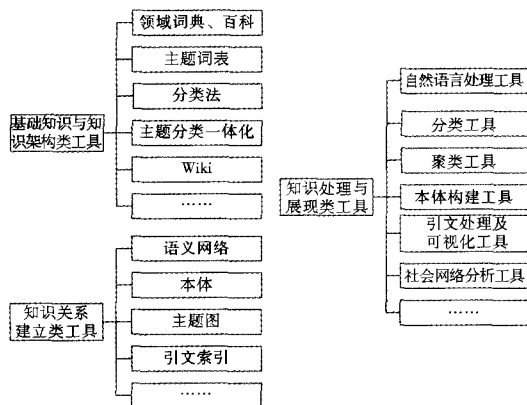


图 1 面向知识服务的知识组织工具分类及架构

其中,基础知识和知识架构类知识组织工具主要作用为架构知识,并为用户提供基础知识释义及查询,即知识点推送服务。在基础架构类知识基础上,针对数据之间的关联,知识关系建立类工具提供了语义关系标注的解决方案,并支持机器理解与推理,即提供知识推理服务。在基础知识构建、知识关系建立的过程中,需要对数据、文本进行处理,如汉语分词、句法标注、本体构建、分类、聚类、网络分析等。一些实用工具为这些处理提供了支持,这些以软件为主的工具是知识处理类工具。

## 2 基础知识与知识架构类工具

基础知识与知识架构类工具在知识组织中实现知识系统化架构及主题标引,进而能够在知识服务中提供知识释义及相关检索。这些工具在概念组织及系统化方面发挥着重要的服务作用,面向不同的知识内容以及用户的知识需求起着规范、释义、传承、指南的作用,其呈现形式也不相同。如知识点的推送服务,通过领域词典、百科全书等基础知识工具来提供相应知识内容检索,这些工具是学科知识释义的重要获取途径。领域词典、百科全书凝聚了编撰专家对收录知识的汇总与规范,具有较高的权威性及影响力,是知识传承的重要组织形式。传统知识组织工具中的主题词表与分类法可以被看作是架构类知识组织工具,体现了领域

专家对学科知识的有序梳理,通过主题词表与分类法对文献资源进行重组排序。基础知识与知识架构类工具是图书情报工作人员从事相关信息服务的主要工具书,在文献编目、标引等工作中发挥重要作用,也是用户解决知识需求的常用工具。基础知识与知识架构类工具的知识组织方式与知识服务内容如表1所示:

表1 基础知识与知识架构类工具对比

工具	提供者	类型	载体形式	服务内容
领域词典、百科全书	领域专家	基础知识类	纸版	知识点释义、检索
主题词表、分类法	领域专家	知识架构类	纸版、电子版	标引、导航
Wiki	用户	基础知识类	电子版	知识点释义

在基础知识与知识架构类工具中,作为领域专家参与的领域词典、百科全书、主题词表、分类法等传统知识组织工具,其特点在于数据质量高。领域词典、百科全书是由专家编撰的知识释义工具,用户只需通过查询词条即可获得较为权威的知识点。主题词表、分类法是领域专家对学科、领域知识的重新梳理,图书情报工作人员可以通过主题词表、分类法对文献进行标引,用户则根据这些标引信息进行查询。在网络环境下,传统知识组织工具在知识组织方面遇到了极大挑战,构建这些工具一般周期较长,虽然信息质量高,但碍于人力、物力、时间限制,在对知识点进行筛选时一般都要控制词条、术语的数量。而在网络环境下发展的开放式百科 Wiki,更加强调用户的参与及分享。在网络环境下,用户参与构建的 Wiki 具有成本低廉、知识组织高效、知识挖掘能力强大等特点,更适应网络海量信息下的知识查询。但其缺点也同样明显,即信息质量参差不齐。Wiki 词条通过链接进行一定组织,借助对词条内容中知识点的机器识别与标识,将词条形成松散的网络,在这点上具有知识关系建立的服务功能。

图书情报领域学者在坚持传统知识组织工具优势的同时,也在探讨网络环境下这些工具的转型,并取得了一定理论与实践成果。Shiri 等<sup>[9]</sup>对网络环境下主题词表的应用进行了总结,并讨论了主题词表的集成与发展;曾建勋等<sup>[10]</sup>提出在网络环境下《汉语主题词表》应基于海量网络数据选择概念数据,更加注重词表间的关系,促进多表融合以及用户的网络参与。在这些研究中,传统知识组织工具在网络环境下的变革中仍能保留其核心价值,即信息质量高、知识组织体系性强。同时,面向用户更为灵活多样的知识需求,其知识

点间的关联加强,用户的协同参与功能也逐渐扩展。但其角色仍主要充当基础知识的提供来源,并提供体系性的知识架构。Wiki 在与用户的协同中,其灵活性是传统知识组织工具所不具备的,作为覆盖面较广、用户接入便捷的基础知识提供工具,在大众知识服务中的作用非常明显,是大众基础知识获取的重要途径。Nakayama 等<sup>[11]</sup>通过对 Wiki 连接的深入挖掘,实现了基于 Wiki 的网络主题词表的构建。

### 3 知识关系建立类工具

知识关系建立类工具通过对数据间关系的建立,将原本无关联的数据连接起来,进而提供基于这些关系的知识呈现,这些关系体现了数据间潜在的、深层的语义关系。通过知识关系的建立,提供更为多元的检索入口,也为机器理解与推理提供了支持,如语义网络、本体、主题图、引文索引、Folksonomy 等。在网络海量数据背景下,通过关系来标识知识一方面可以充分体现多元关系,一方面也更适应于机器理解与推理。与知识架构类不同的是,知识关系建立类工具更多强调知识之间的关联呈现,而不是系统化地构建知识体系。需要说明的是,知识架构类工具可以通过对其上下位类关系、附属索引的转化完成知识关系建立,即知识架构类工具通过转化可以成为知识关系建立类工具,其关系为分类中的上下位关系以及相关附属索引。知识关系建立类侧重于揭示数据、信息、知识间的关系,提供这些关系的呈现,并以此进行知识组织。这些关系的识别和标识,则需通过知识处理类工具完成。知识关系建立类工具的知识组织方式与知识服务内容如表2所示:

表2 知识关系建立类工具对比

工具	知识组织方式	特点	服务内容
语义网络	概念节点及其语义关系构成网络	语义关系更为灵活	支持概念查询,并提供相关概念的链接网络
本体	概念及概念关系规范化表示	语义关系更为规范	智能化检索、推理、知识导航
主题图	以主题为中心建立语义参照	以元数据等作为技术标准	以语法关系呈现主题关联
引文索引	以引用关系为中心	通过引用关系呈现文献间知识源流	共引分析、共现分析
Folksonomy	以用户标签组织文献资源	用户协同性强	通过标签分享知识

在知识关系建立类工具中,语义网络、本体、主题

图等工具均围绕语义这一中心,通过知识关系进行组织,并提供以语义关系查询为呈现的知识服务。与基础知识与知识架构类工具的不同在于,知识关系建立类工具对知识点之间的关系进行更深层次的挖掘,同时这些工具的基础知识和知识架构仍可以采用相关知识组织工具。在传统知识组织工具基础上,国内外学者将知识关系建立类工具与主题词表、分类法、领域词典相结合,取得了一定理论和实践成果。Garcia - Torres 等<sup>[12]</sup>提出了基于主题词表及本体开发方法的主题词表 - 本体构建模式;Hahn 等<sup>[13]</sup>在医学主题词表基础上构建了一体化医学本体;Hu 等<sup>[14]</sup>在 WordNet 及中国分类主题词表基础上构建了双语本体;Hepp 等<sup>[15]</sup>以叙词表、分类法为基础提出了基于 OWL 和 RDF - S 本体标准的通用本体构建方法。国内学者曾新红等<sup>[16]</sup>描述了网络环境下叙词表向 OWL 本体转化的思路和实践;王军等<sup>[17]</sup>基于《中国分类主题词表》和书目数据进行了本体自动构建研究。语义网络中影响力较大的有普林斯顿大学开发的 WordNet,基于英语单词同义的认知归类,将词典内词汇链接为词汇语义网络<sup>[18]</sup>。

引文索引呈现了文献之间的引用关系,体现了文献内知识的源流。通过引文索引的追溯,可以把握科研发展的脉络,同时也可以展开共词、共引、共现、科研合作等深层分析。以引文索引为组织方式,提供相关知识服务的有 SCI、SSCI、CSSCI 等,这些引文索引为引文分析提供了数据支持。

Folksonomy 是 Web2.0 环境下基于个人化自由标签的知识组织方式。通过用户自定义标签可以对 URL 资源进行个性化标识,网络信息资源则通过众多用户的个性化标签进行组织。Folksonomy 以用户标签为资源关系的组织方式,其成本低廉、适用范围广,成为网站 Web2.0 应用的主要模式。Folksonomy 由于其出色的应用而受到学者关注,Angus 等<sup>[19]</sup>考察大学内图像群以及 Flickr 图像群并验证了用户标签在图片资源组织中的有效性。

#### 4 知识处理及展现类工具

知识处理及展现类工具主要面向知识的识别、提取、构建等任务,根据任务不同其分类也不同。对于文本中知识提取,首先要利用自然语言处理工具对文本中的知识点词汇和短语进行识别和提取,在中文文本

中需要使用中文分词工具。在词汇基础上,可以对句子中的特定成分进行识别与提取,获得其语法及语义关系,句法分析工具可以提供有利支持。在提取文献特征词汇的基础上,可以通过分类及聚类工具进行标注。在获取文献相关概念知识后,通过本体构建工具、语义网络构建工具,可以实现本体的构建及语义网络的绘制。社会网络分析为语义网络更深层次知识的挖掘提供了可能,而这些知识可以通过可视化工具进行展现。知识处理及展现类工具是计算机技术与相关知识需求的产物,这些工具以相关数据为基础,提供面向知识点的服务。目前广泛应用的知识处理及展现类工具如表 3 所示:

表 3 知识处理及展现类工具对比

工具类别	工具	服务内容
自然语言处理类工具	ICTCLAS、Stanford Parser、LTP	提供词汇、短语、句法知识
分类工具	SVMTools、Maximum Entropy Modeling Toolkit for Python and C++、CRF++、Mulan	提供分类服务
聚类工具	Cluto、Weka	提供聚类服务
本体构建工具	Protégé、Jena	提供本体构建、操作及呈现
引文处理及可视化工具	EndNote、NoteExpress、CiteSpace	提供引文管理服务、可视化服务
社会网络分析工具	Ucinet、Pajek	提供社会网络分析及可视化服务

图书情报知识服务从物理层次的文献服务转向语义层次的知识点服务是知识服务发展的趋势,在这一过程中如何从文本中提取知识点就成为待解决的首要问题,与之相关的知识抽取研究也成为学界关心的热点。这些研究处理对象均从传统的叙词表、分类法、百科全书转向为自然文本,这需要涉及到对词汇、短语、句法等自然语言知识的处理。这种需求促进了自然语言处理工具的发展,并不断拓宽了其应用。这类工具面向自然语言文本主要提供词汇、短语、句法一级的知识服务,是进行分类、聚类、本体构建等处理的基础。通过分词、词性标注、句法分析等服务,对文本中的人名、地名、机构、数词、时间等知识点进行识别,进而实现命名实体识别、信息抽取、本体构建、知识网络构建等深层分析。目前,这类工具中应用较广的有提供中文分词服务的 ICTCLAS<sup>[20]</sup>,提供英文句法分析的 Stanford Parser<sup>[21]</sup>,提供英文短语标注的 OpenNLP<sup>[22]</sup>,以及提供分词、词性标注、命名实体识别、依存句法分析、词义消歧、语义角色标注服务的哈工大 LTP 平台<sup>[23,24]</sup>。

在知识框架体系中,分类是重要的组成部分,通过知识分类可以获取其上下位、同位关系。以知识点为分类对象的工作更加强调计算机的参与,这与传统环境下通过人工或人工辅助标引有很大不同。目前,已有一些分类工具可以提供相关应用服务,如基于支持向量机(Support Vector Machine, SVM)模型的 SVM-Tools<sup>[25]</sup>、基于最大熵模型的 Maximum Entropy Modeling Toolkit for Python and C++<sup>[26]</sup>、基于条件随机场模型的 CRF++<sup>[27]</sup>、基于多标签(Multi-Label)的 Mulan<sup>[28]</sup>。通过这些工具,用户只需对输入数据格式进行控制,并选择合适的特征即可进行分类工作,在相关研究中已较多采用这些工具。

在博客、微博等用户更多参与知识共建的环境下,对于用户类型、兴趣、关系等研究对象的分类是不确定的,需要通过对关键词、用户标签等知识进行聚类分析。在聚类相关研究中,Cluto<sup>[29]</sup>、Weka<sup>[30]</sup>等聚类工具为研究人员提供了聚类服务,如 Patterson 等<sup>[31]</sup>在 Cluto 聚类工具基础上对临床记录中的描述文本进行了类别挖掘研究。

在知识关系建立类工具应用中,本体在基于机器理解的概念及关系组织中受到较多关注。在通过本体进行知识关系组织前,需要对其进行构建,本体构建工具就是为此服务的知识处理类工具。在本体构建中,这些工具可以支持层次性概念类本体构建,即由知识架构类工具转化本体,也支持面向语义网的网络本体构建(如 Protégé<sup>[32]</sup>),还有一些工具主要面向本体操作(如 Jena<sup>[33]</sup>)。这些本体构建工具为本体研究提供了技术性解决方案,在相关研究中也已被较多采用<sup>[34]</sup>。

引文知识管理工具服务内容主要有提供在线检索数据库的统一接口、定制个性化检索并保留参考文献数据、管理并辅助用户正确引用等。在引文数据的基础上,一些引文工具还提供可视化分析服务,为用户了解某一主题、学科发展提供参考。由于信息服务商不同,各期刊数据库的检索存在差异,通过引文知识管理工具可以实现多数据检索与下载,并根据用户需求实现定制服务,这些工具中影响力较大的有 EndNote<sup>[35]</sup>、NoteExpress<sup>[36]</sup>。在引文知识管理工具基础上,通过下载引文文本,可以支持对文献源流的分析及可视化, CiteSpace<sup>[37]</sup>是引文分析工具中影响力较大的开源工具。CiteSpace 可以提供基于引文的作者合作、关键词

共现、引用关系等可视化分析,目前支持 Web of Science、CNKI、CSSCI 等引文索引分析。

在知识关系构建类工具中,语义网络使得知识通过有向及无向边连接为相互关联的网络。社会网络分析工具支持对网络中的隐性知识进行研究,如网络中心度、子团体划分、角色分析等,从而更深层次地挖掘语义网络中的知识。在知识组织中,通过关系矩阵构建,可以将语义网络知识转化为相应的矩阵知识。在此基础上,可以对网络整体的各种关系进行研究,如关系的传递性等,也可以对网络个体分析其密度、同质性等指标。在图书情报的应用研究中,对于研究对象间的社会网络关系,如论文的引用网络、作者的引用网络、作者合作网络、期刊引用网络等,可以通过 Ucinet<sup>[38]</sup>、Pajek<sup>[39]</sup>提供相关应用服务。

## 5 知识组织工具展望

在知识组织工具的未来发展中,以用户服务为核心的智能化知识组织工具是发展的主要方向。在基础知识与知识架构类工具的基础上,知识词表建设将更加智能化,即词表更加动态化、个性化。通过文本处理工具,与基础知识与知识架构类工具实现交互,从而完成用户知识词表构建。在大规模用户知识词表的基础上,通过云计算等分布式技术,利用知识关系建立类工具,自动获取用户分类与聚类知识。在此基础上完成本体、语义网络构建,并通过互操作实现知识共享。在相关服务中,更加强调用户对文献的标引、用户知识词典与叙词表、分类法相互结合,通过知识关系建立完成用户知识与专家知识的融合。因此,未来知识组织工具的发展具有多维融合、网络化、分布式、专业化与大众化、个性化及动态性的特征。

(1) 多维融合。对比国内外知识组织相关研究,可以看出国外研究多侧重于实践,其中心为知识组织系统(KOS)。在国外知识组织系统中,基础知识与知识架构类工具中的叙词表、分类法、领域词典与语义网、本体、主题图等知识关系建立类工具相结合,最终目的是提供相关知识领域的智能检索、语义查询等服务,如美国国立医学图书馆的 MeSH<sup>[40]</sup>、UMLS<sup>[41]</sup>。MeSH、UMLS 在原有叙词表基础上与语义网结合,提供语义概念关系表达及查询。国内研究虽已将基础知识与知识架构类工具、知识关系建立类工具相结合,但多

数研究仍是理论与实践探索,能够提供知识服务的系统化工具较少。

多维融合的另一层次是指与自然语言处理(NLP)、社会网络分析(SNA)等研究相结合。从文本中进行知识发现首先需要 NLP 技术工具的支持,在国外相关知识处理及展现类工具中已集成了相关的 NLP 服务,如 CiteSpace 中术语的自动提取、词性标注等功能。由于中文分词工具对于专业知识文本的识别不理想,从而导致目前对于文本中术语识别、提取等工作仍主要停留在研究阶段,构建融合相关 NLP 功能从而提供服务的知识组织工具已成为当务之急。随着基于引文共现、共引、科研合作等知识图谱研究在国内的开展,CiteSpace、VOSViewer 等可视化工具在国内学术界已取得了一定成果。但笔者注意到,这些工具往往都是基于特定数据格式的可视化服务,对国内主流数据库支持不足。以 CiteSpace 为例,其主要支持格式为 Web of Science,虽然已提供对 CNKI 及 CSSCI 的支持,但数据来源仍需经过格式转换。对于此类工具,能够更好地集成国内数据源的获取功能,是其提升知识服务的新挑战。

(2) 网络化与分布式。对于网络环境下的新型叙词表、分类法、领域词典等工具,国外已有许多提供在线服务的实例。国内对于网络化环境下的知识组织工具,理论与探索研究居多,能够提供在线服务的较少。同时基于不同知识关系架构的工具间互映射也是未来研究的重点。网络环境下 Folksonomy 为知识服务带来了新的挑战,基于 Folksonomy 的用户标签具有数量庞大、质量参差不齐、口语化严重等缺点。未来知识组织工具应当融合分类、聚类服务,并通过分布式来解决海量用户标签的组织,提供较为规范、基于大众观点的标记服务。

(3) 专业化与大众化。目前在知识组织服务中具有较多应用资源的主要是医学、生物学等学科。这个现象表明,知识组织工具应当具有一定的专业性。对于专业性较强的领域,知识组织工具的组织理念应以规范为中心,尽量避免歧义。而对于大众化服务,则应在注重规范性的同时保持多样性和协同性。Folksonomy 在图像、视频、社交等网站上的流行表明了大众对于知识表达多样性的诉求,面向大众的知识组织工具应当保持这种多样性,并提供为大众所接受的语义关联服务。

(4) 个性化及动态性。个性化服务是近年来知识服务研究热点,基于个人检索历史的分析为个性化的

知识推送服务提供了有力支持。由于大数据环境下信息来源的海量性与多样性,未来知识组织工具需要根据用户知识需求动态组织信息资源,通过结合分类、聚类技术建立用户知识需求与资源间的相互关联。通过大规模的用户数据聚类分析,获取不同用户知识需求团体,结合个人历史记录进行智能化的知识推送服务。

## 参考文献:

- [1] 张晓林. 走向知识服务:寻找新世纪图书情报工作的生长点[J]. 中国图书馆学报, 2000, 26(5): 32-37. (Zhang Xiaolin. Toward Knowledge Services: Seeking for the Development Opportunities for Library and Information Services in the 21st Century[J]. Journal of Library Science in China, 2000, 26(5): 32-37.)
- [2] 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[J]. 中国图书馆学报, 1999, 25(5): 19-23. (Jiang Yongfu. Library and Knowledge Organization——Understanding Library Science from the Perspective of Knowledge Organization[J]. Journal of Library Science in China, 1999, 25(5): 19-23.)
- [3] Bliss H E. The Organization of Knowledge[M]. H. Holt and Company, 1929.
- [4] 刘洪波. 知识组织论——关于图书馆内部活动的一种说明[J]. 图书馆, 1991(2): 13-18, 48. (Liu Hongbo. Knowledge Organization Theory——An Explanation about the Internal Activities in Library[J]. Library, 1991(2): 13-18, 48.)
- [5] 马费成. 知识组织系统的演进与评价[J]. 知识工程, 1989(2): 39-43. (Ma Feicheng. The Evolution and Evaluation of Knowledge Organization System[J]. Knowledge Engineering, 1989(2): 39-43.)
- [6] 王知津. 从情报组织到知识组织[J]. 情报学报, 1998, 17(3): 230-234. (Wang Zhijin. From Information Organization to Knowledge Organization[J]. Journal of the China Society for Scientific and Technical Information, 1998, 17(3): 230-234.)
- [7] Boisot M H. Knowledge Assets: Securing Competitive Advantage in the Information Economy [M]. Oxford University Press, 1998.
- [8] 毕强, 韩毅, 刘昆. 知识服务——现状、进展及挑战[J]. 中国图书馆学报, 2007, 33(6): 41-45. (Bi Qiang, Han Yi, Liu Kun. Knowledge Services: Present, Improvement and Challenges [J]. Journal of Library Science in China, 2007, 33(6): 41-45.)
- [9] Shiri A A, Revie C. Thesauri on the Web: Current Developments and Trends[J]. Online Information Review, 2000, 24(4): 273-279.
- [10] 曾建勋, 常春. 网络环境下新型《汉语主题词表》的功能定位与发展[J]. 情报学报, 2010, 29(6): 973-977. (Zeng Jianxun, Chang Chun. Function Orientation and Development of New Edition of Chinese Thesaurus in Web Environment[J]. Journal of

- the China Society for Scientific and Technical Information*, 2010, 29 (6): 973 - 977. )
- [11] Nakayama K, Hara T, Nishio S. Wikipedia Mining for an Association Web Thesaurus Construction[C]. In: *Proceedings of the 8th International Conference on Web Information Systems Engineering (WISE'07)*. Berlin, Heidelberg: Springer - Verlag, 2007: 322 - 334.
- [12] Garcia - Torres A, Pareja - Lora A, Pradana - Lopez D. Reusing Thesauri: Documentalists Face the Semantic Web Challenge[J]. *Profesional De La Informacion*, 2008, 17(1): 8 - 21.
- [13] Hahn U, Schulz S. Building a Very Large Ontology from Medical Thesauri[A]. // *Handbook on Ontologies* [M]. Berlin, Heidelberg: Springer, 2004: 133 - 150.
- [14] Hu H, Du X Y. Building Bilingual Ontology from WordNet and Chinese Classified Thesaurus[C]. In: *Proceedings of the 2nd International Conference on Knowledge Science, Engineering and Management*. Berlin, Heidelberg: Springer, 2007: 649 - 654.
- [15] Hepp M, de Bruijn J. GenTax: A Generic Methodology for Deriving OWL and RDF - S Ontologies from Hierarchical Classifications, Thesauri, and Inconsistent Taxonomies[C]. In: *Proceedings of the 4th European Semantic Web Conference*. Berlin, Heidelberg: Springer, 2007: 129 - 144.
- [16] 曾新红, 明仲, 蒋颖, 等. 中文叙词表本体共建共享系统研究[J]. *情报学报*, 2008, 27(3): 386 - 394. (Zeng Xinhong, Ming Zhong, Jiang Ying, et al. Research on OntoThesaurus Co - Construction and Sharing System[J]. *Journal of the China Society for Scientific and Technical Information*, 2008, 27(3): 386 - 394. )
- [17] 王军, 程煜华. 基于传统知识组织资源的本体自动构建[J]. *情报学报*, 2009, 28(5): 651 - 657. (Wang Jun, Cheng Yuhua. An Automatic Approach to Ontology Building by Integrating Traditional Knowledge Organization Resources[J]. *Journal of the China Society for Scientific and Technical Information*, 2009, 28(5): 651 - 657. )
- [18] WordNet [EB/OL]. [2013 - 02 - 05]. <http://wordnet.princeton.edu/wordnet/>.
- [19] Angus E, Thelwall M, Stuart D. General Patterns of Tag Usage Among University Groups in Flickr[J]. *Online Information Review*, 2008, 32(1): 89 - 101.
- [20] ICTCLAS2013 [EB/OL]. [2013 - 02 - 05]. <http://ictclas.nlpir.org/>.
- [21] The Stanford Parser: A Statistical Parser[EB/OL]. [2013 - 02 - 05]. <http://nlp.stanford.edu/software/lex-parser.shtml>.
- [22] Apache OpenNLP [EB/OL]. [2013 - 02 - 05]. <http://opennlp.apache.org/>.
- [23] 刘挺, 车万翔, 李正华. 语言技术平台[J]. *中文信息学报*, 2011, 25(6): 53 - 62. (Liu Ting, Che Wanxiang, Li Zhenghua. Language Technology Platform[J]. *Journal of Chinese Information Processing*, 2011, 25(6): 53 - 62. )
- [24] Ltp: Language Technology Platform[EB/OL]. [2013 - 02 - 05]. <http://ir.hit.edu.cn/ltp/>.
- [25] SVMTool [EB/OL]. [2013 - 02 - 05]. <http://www.lsi.upc.edu/~nlp/SVMTool/>.
- [26] Maximum Entropy Modeling Toolkit for Python and C++ [EB/OL]. [2013 - 02 - 05]. [http://homepages.inf.ed.ac.uk/lzhang10/maxent\\_toolkit.html](http://homepages.inf.ed.ac.uk/lzhang10/maxent_toolkit.html).
- [27] CRF++: Yet Another CRF Toolkit [EB/OL]. [2013 - 02 - 05]. <http://code.google.com/p/crfpp/>.
- [28] Mulan [EB/OL]. [2013 - 02 - 05]. <http://sourceforge.net/projects/mulan/>.
- [29] Cluto [EB/OL]. [2013 - 02 - 05]. <http://glaros.dtc.umn.edu/gkhome/cluto/cluto/>.
- [30] Weka [EB/OL]. [2013 - 02 - 05]. <http://www.cs.waikato.ac.nz/ml/weka/>.
- [31] Patterson O, Hurdle J F. Document Clustering of Clinical Narratives: A Systematic Study of Clinical Sublanguages[C]. In: *Proceedings of the AMIA Annual Symposium*. American Medical Informatics Association, 2011.
- [32] Protégé [EB/OL]. [2013 - 02 - 05]. <http://protege.stanford.edu/>.
- [33] Apache Jena [EB/OL]. [2013 - 02 - 05]. <http://jena.apache.org/index.html>.
- [34] 王昊, 苏苏宁. 基于本体的 CSSCI 学术资源网络模型构建及其应用研究[J]. *情报学报*, 2010, 29(2): 331 - 341. (Wang Hao, Su Xinning. Research on Construction and Applications of CSSCI Academic Resources Networks Model Based on Ontology[J]. *Journal of the China Society for Scientific and Technical Information*, 2010, 29(2): 331 - 341. )
- [35] EndNote [EB/OL]. [2013 - 02 - 05]. <http://endnote.com/>.
- [36] NoteExpress [EB/OL]. [2013 - 02 - 05]. <http://www.reflib.org/CompanyWeb/>.
- [37] Chen C. CiteSpace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature[J]. *Journal of the American Society for Information Science and Technology*, 2006, 57(3): 359 - 377.
- [38] Ucinet [EB/OL]. [2013 - 02 - 05]. <http://www.analytictech.com/ucinet/>.
- [39] Pajek [EB/OL]. [2013 - 02 - 05]. <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>.
- [40] MeSH [EB/OL]. [2013 - 02 - 05]. <http://www.ncbi.nlm.nih.gov/mesh>.
- [41] UMLS [EB/OL]. [2013 - 02 - 05]. <http://www.nlm.nih.gov/research/umls/>.

(作者 E-mail: bmy\_xj@163.com)