



# 基于 CSSCI 本体的知识检索服务平台构建及应用<sup>\*</sup>

王 昊 苏新宁

(南京大学信息管理系 南京 210093)

**【摘要】** 现有的 CSSCI 信息服务系统只能提供单一、线性的检索功能,检索效果不佳,主要原因在于后台数据组织的不合理。为此,提出使用基于本体的知识服务平台代替现有 CSSCI 检索服务系统的解决方案,提供知识层次的服务,以改善用户的检索环境。在提出基于本体信息检索系统一般模式的基础上,设计一个实用的基于 CSSCI\_Onto 的知识服务平台 KRSP\_CSSCI\_Onto,对该平台的基本功能和系统框架进行探讨,并据此开发一个试验性的知识检索系统,详细阐述本体在知识检索服务中的具体应用,包括用户检索式的语义扩展,查询结果的关联知识推荐,基于学术资源网络模型的知识导航以及基于规则库的知识关系检索等。

**【关键词】** CSSCI 本体 学术资源网络模型 知识检索 语义关联推荐 蕴含推理

**【分类号】** G354

## Services Platform for Knowledge Retrieval Based on CSSCI\_Onto

Wang Hao Su Xinning

(Department of Information Management, Nanjing University, Nanjing 210093, China)

**【Abstract】** The CSSCI information services system can only provide single and linear retrieval, and the retrieval effect is not very good. The main reason lies in the unreasonable data organization background. Therefore, the authors put forward the solution which uses knowledge services platform based on Ontology instead of the existing CSSCI information services system, and provide the services in the level of knowledge, so that to improve the retrieval environment. The paper proposes a general model of information retrieval system based on Ontology, designs a practical knowledge services platform named KRSP\_CSSCI\_Onto, and gives the basic functions and system framework of the platform. An experimental knowledge retrieval system is developed hereby, and the applications of Ontology in knowledge retrieval services is expounded in detail, including semantic extension of search expression, related knowledge recommendation of query results, knowledge navigation based academic resource networks model and knowledge relationships retrieval based rule library.

**【Keywords】** CSSCI Ontology Academic resources networks model Knowledge retrieval Semantic association recommendation Implication reasoning

### 1 引言

中国社会科学引文索引(简称 CSSCI),每年收录 400-500 种人文社科类精品期刊的论文及其相关学术资源,

收稿日期: 2011-01-30

收修改稿日期: 2011-02-26

<sup>\*</sup> 本文系国家自然科学基金项目“面向语义网本体的知识管理研究”(项目编号:09CTQ010)的研究成果之一。

以提供信息检索服务。现有的 CSSCI 检索平台仅能通过元数据检索的方式实现对来源文献和被引文献的查询<sup>[1]</sup>,检索方式和对象都较为单一,既无法利用学术资源间存在的显性或隐性语义关系实现关联推荐,也不能对其他重要的学术资源知识进行相关查询。究其原因,是该平台采用了面向结构的后台数据组织方式,各种资源只是被简单地堆放在一起,相互之间的语义关系没有得到充分表现。随着语义网(Semantic Web)<sup>[2]</sup>概念的提出,作为知识组织与描述基础的本体机制得到了充分的发展。以本体为核心的语义检索开始取代传统的关键词检索,成为信息检索领域研究的主要方向<sup>[3]</sup>,产生了一些著名的检索平台,如 KIM<sup>[4]</sup>、FindUR<sup>[5]</sup>、基于历史领域本体的历史资源检索系统等<sup>[6]</sup>。

随着 CSSCI 收录资源的不断扩充,其所蕴含的知识也越来越丰富。现有的以线性信息组织方式为基础的单一资源检索服务模式无法满足用户日益增长的知识服务需求。为此,本文试图将面向对象的本体机制引入到 CSSCI 知识组织中,充分挖掘各种学术资源之间的语义关联,形成学术资源网络模型;在此基础上,构建 CSSCI 知识检索服务平台(Knowledge Retrieval Services Platform based CSSCI Ontology, KRSP\_CSSCI\_Onto),利用本体中蕴含的知识间语义关系提供更好的检索服务,包括关联推荐、知识导航、关系推理等。

## 2 面向知识服务的 CSSCI 本体及其应用

笔者收集了 CSSCI(2000-2006)共 7 年的数据及其元数据结构,作为 CSSCI Ontology(简称 CSSCI\_Onto)的来源数据。本体构建的基本思路如下:

(1)从来源数据中提取出学术资源概念集合(如主题、文献、期刊、学科等)和概念间关系(包括分类关系和非分类语义关系),建立本体概念模型,实现对 CSSCI 学术资源的抽象描述。CSSCI\_Onto 概念模型包括三层概念层次结构,共 39 个本体类,336 个属性<sup>[7]</sup>。

(2)在概念模型的指导下,对 CSSCI(2000-2006)数据进行语义标注,提取知识元及其关系,将各类学术资源实例关联在一起,形成学术资源知识地图,其中包括 552 566 个主题,504 021 个来源文献,558 种期刊,25 个学科以及 215 942 位学者,89 884 个部门机构实例等<sup>[8]</sup>。

(3)基于逻辑推理理论建立蕴含推理规则,用于扩充本体知识库中非显式存在的学术资源间语义关联,例如学者之间的同被引关系可以根据论文同被引关系推理获得,而不需要固化为本体知识库中。

概念库、实例库和规则库构成了完整的 CSSCI\_Onto 知识库,从而实现了 CSSCI 学术资源及其间语义关联的统一、明确和规范的描述。

CSSCI\_Onto 可用于提供知识服务。依据 CSSCI\_Onto 固有的显性知识以及通过对显性知识的挖掘而发现的隐含知识,实现科学评价和学术分析,包括挖掘学术资源间二元和多元关联<sup>[9]</sup>,探测各学科研究热点以及多学科合作热点<sup>[10]</sup>,分析学科热点的发展趋势等;以 CSSCI\_Onto 为数据支持开发知识检索服务平台,以满足用户查询本体中各类知识的需求,具体包括对用户检索条件和检索结果的关联推荐,基于学术资源关联的导航式检索以及对学术资源之间关系的蕴含推理检索等。本文重点探讨基于本体的知识检索服务平台的构建及应用。

## 3 基于 CSSCI\_Onto 的知识检索服务平台构建

在总结基于本体信息检索系统的一般模型的基础上,结合 CSSCI 知识服务的特点和功能要求,构建了一个基于 CSSCI\_Onto 的知识检索服务平台 KRSP\_CSSCI\_Onto。

### 3.1 基于本体的信息检索系统的一般模型

基于本体的信息检索的基本思想是:在领域专家的帮助下,建立相关领域的本体概念模型;以此为指导,对收集的数据进行语义标注,并按规定格式存储在元数据库(如 OWL 文档等)中作为检索源;查询转换器依据本体知识库把用户的查询请求转换成指定的格式并实现扩展,进而在元数据库中匹配出符合条件的数据集合;最后将检索的结果经过处理后返回给用户。整个过程的关键是使用语义标注取代了传统信息检索中的关键字标引或主题标引(文本标引),包括检索资源和检索表达式的语义标注。

基于本体的信息检索系统的一般模型如图 1 所示。该模型包括 5 个模块:

(1)用户界面模块,实现与用户的交互,包括用户输入接口的设计和检索结果的返回。一般情况下,用

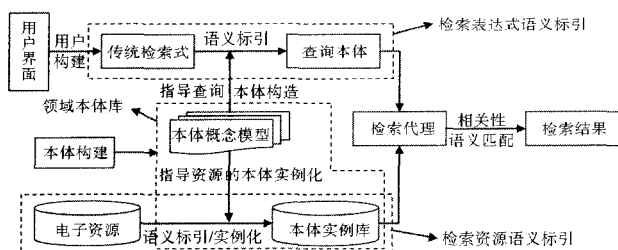


图1 基于本体的信息检索系统的一般模型

户输入的约束条件越多、越精确,则表述的语义就越清楚,用户操作也相对复杂;反之,用户输入的约束条件越少,用户操作越简单,但系统需要花更多的精力来理解用户输入所表达的语义。根据用户输入约束条件的繁简,可将用户界面形式划分为基于复杂表单式的检索、基于查询词的检索和基于自然语言的检索<sup>[11]</sup>。

(2) 查询式语义标注模块,是指依据领域本体,将用户的检索表达式转化为系统能够理解的标准化的本体检索式(查询本体)。上文描述三类用户输入都需要进行语义标引,区别在于:基于复杂表单式的检索中检索项和检索式之间是语义对应关系,语义标注在用户根据检索项输入检索式时即完成,也可以基于知识库对用户输入的检索式进行扩展和修正<sup>[12]</sup>;而基于查询词的检索和基于自然语言的检索则需要由系统根据领域本体自动完成语义标引,其难度随检索式的自然化程度而变化,越接近自然语言,难度越大。由于自然语言处理技术的瓶颈,这两种方式目前多停留在实验阶段,但由于其极大地降低了用户构造查询式的能力要求,已经成为基于本体信息检索系统发展的趋势<sup>[13]</sup>。

(3)领域本体构建模块,这里仅指概念模型的构建,即识别领域中的所有概念,并定义概念属性以建立概念之间的关系,如有必要还需建立逻辑推理规则(公理)以约束概念行为。

(4) 检索资源的语义标引模块,是指在领域本体概念模型的指导下,将检索资源按照指定的格式进行组织和存储,建立语义索引。该过程实际上是本体概念模型在现实世界中的具体反映,也称为本体实例化。

(5) 检索处理模块, 查询本体通过检索代理与本体实例库中实例及其属性值进行相关性语义匹配, 并返回相关的检索结果。由于本体中资源之间存在相关性, 可以由系统自动或用户选择等方式返回与检索结果相关的其他信息。此外, 为了给用户提供一个更加

直观的展示效果,在该模块中可以增加信息可视化展示子模块,以图形化的方式丰富本体实例信息的展示。

综上所述,在基于本体的信息检索系统中,本体的作用主要表现在:借助本体概念模型,对检索资源进行语义标注,使资源充分表现出其内涵语义;利用本体描述的实例间关联,对用户检索式进行标准化,使其满足本体实例的格式,同时将查询本体的相关信息推荐给用户,帮助用户把未意识到的或未清晰表达的信息需求进一步显性化;依据本体中描述的实例间关联,对用户检索结果进行扩展或推荐等。

### 3.2 基于 CSSCI\_Onto 的知识检索服务平台的系统框架

知识检索(Knowledge Retrieval),是指在知识组织的基础上,利用知识关联和概念语义检索,智能化地从知识库中检索出知识的过程<sup>[14]</sup>。其区别于信息检索的两个显著的特征在于:基于某种具有语义模型的知识组织体系;对检索资源进行了基于元数据或本体概念模型的语义标注,将其转化为知识。笔者建立了CSSCI本体概念模型,形成了具有语义模型的知识组织体系;并以此对CSSCI(2000-2006)的学术资源进行了语义标注,以面向对象形式组织CSSCI的学术知识。这两个条件的满足为构建知识检索服务平台和向用户提供学术资源知识服务奠定了基础。

KRSP\_CSSCI\_Onto 是一个基于本体的支持多对象和对象关系检索的知识服务平台,可以实现用户检索式的引导式构建,具有检索结果关联推荐功能,能够可视化展示对象实例及其关系。图2描述了该平台的系统架构,整个知识服务平台分为三个模块。

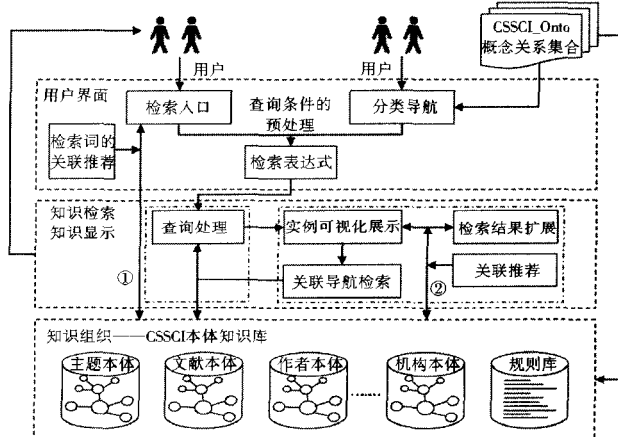


图2 KRSP\_CSSCI\_Onto 平台的系统架构

### (1) 用户检索界面

系统提供了两种用户检索方式,即分类导航和关键字查询方式。

①分类导航是由系统开发者根据用户可能存在的检索偏好或潜在的检索需求而预先设定好的检索表达式,用户可以在系统的指引下逐步查找所需的信息。本体中最常见的关系是 kind-of(继承关系),而描述这种关系的最直观的方法就是采用树型分类。本系统根据 CSSCI\_Onto 的概念模型对学术资源进行分类导航。分类导航由系统开发人员设置,无法兼顾所有用户的需求,只能作为用户在检索需求不太明确时的一种辅助检索策略。

②关键字查询是用户根据自身的检索需要,在系统的帮助下自主地构建检索表达式,进行信息查询,这是目前使用最广泛的具有较好效果的检索方式。在系统中将概念的属性作为检索项供用户选择,实际上是人为地将用户的检索需求进行了语义标注;而且能够根据用户的检索选项和检索关键字,基于实例库推荐相关的资源对象实现查询扩展,以便用户更好地选用检索关键字构建更准确的检索表达式。

③将分类导航和关键字检索结合使用,具体的做法是先行进行有限的分类导航,在缩小范围之后再行关键字检索。

### (2) 知识检索和知识显示

无论是分类导航还是关键字检索,最后都是形成检索表达式交由知识检索和知识显示模块进行处理。该模块又可以分为知识检索处理子模块和知识展示子模块,如图 2 中虚线框所示。知识检索处理子模块是接收来自用户界面的检索表达式,将其解释成为含有条件的 SPARQL 语句,在本体库中将用户检索关键字和实例属性值进行相关性匹配,确定检索请求与本体实例及关系的对应,获得符合条件的实例对象或实例关系,并将结果返回给知识展示子模块;知识展示子模块负责知识实例集合和知识实例信息及关系的显示(包括列表和图形等显示方式),并在实例库的作用下,完成对检索结果的知识扩展和知识间的关联导航。

### (3) 知识组织(CSSCI\_Onto 知识库)模块

在本体构建中, CSSCI\_Onto 分为三个部分,即概念模型、实例集合以及推理规则集合。概念模型是对实例的声明和定义,能够提供实例分类的标准,实现系统的分类导航;实例集合则是对概念模型的具体实现,是 CSSCI\_Onto 的主要内容,用户检索的主要对象,包括主要概念实例如主题、文献、作者等和辅助概念实例如基金、时间等;规则集合则包含了由本体工程师定义

的推理规则,能够实现对本体库固有知识的有效扩展。本系统提供对本体中主要概念实例的检索,至于辅助概念实例则可以通过分类导航或资源关联导航进行查找。CSSCI\_Onto 实例不仅是用户检索的对象,而且在检索关键词推荐和检索结果推荐中都具有重要应用,如图 2 中①和②所示。CSSCI\_Onto 是 KRSP\_CSSCI\_Onto 的知识基础,为知识检索的实现提供了知识组织体系的保障以及对学术资源语义标注的支持。

## 4 基于 CSSCI\_Onto 的知识检索服务应用

CSSCI\_Onto 在 KRSP\_CSSCI\_Onto 平台中的应用主要表现在学术资源推荐、知识导航以及学术资源关系检索中。

### 4.1 基于语义关联的学术知识推荐

CSSCI\_Onto 描述了同类实例间的语义关联,借此可以实现学术资源的关联推荐,包括对用户检索词的语义扩展和对检索结果的语义推荐,这是知识检索的基本功能之一<sup>[15]</sup>;在学术资源推荐的同时,能够从一个知识转移到推荐的知识中,以此为导航,实现用户最终检索目标的定位。

#### (1) 检索表达式的语义扩展

在 KRSP\_CSSCI\_Onto 的关键字检索系统中,如果以描述实例间关系的对象属性(Object Property)作为检索项,那么用户输入的检索关键字(即对象名称)在知识库中就可能存在其他同类实例与之相关,而这些相关实例特别是关联度比较高的实例很有可能也是用户感兴趣的对象,这些对象可以作为用户原始查询式的语义扩展。例如,在来源文献检索中,用户使用其属性“关联主题”查询与主题相关的所有来源文献,用户输入“数据挖掘”,此时向用户推荐与该主题相关的其他主题,如“数据仓库”、“知识发现”、“数字图书馆”等,如图 3 所示:

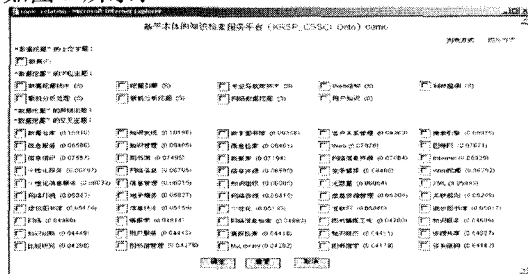


图 3 “数据挖掘”的关联主题推荐

其中不乏用户感兴趣的主体。选择这些相关主体进行文献检索,即可实现查询式的语义扩展。对于不同类型的对象属性,可以提供不同的扩展方向。本文主要以“主题”概念为例阐述检索词语义扩展方式。

在 CSCI\_Onto 中,“主题”概念具有上下位主题、共现主题以及交叉主题等属性,并且基于来源数据为每一个主题实例设置了上述属性的属性值。系统可以基于主题之间的这 4 种关系为用户提供相关主题实例,实现 4 个方向上的语义扩展,见图 3。用户在使用主题名称、关联主题或文献篇名等属性作为检索项检索实例时,输入检索式后即可显示关联主题。用户可以根据检索意图从中选择相关主题以完善查询式。

这种主题语义扩展在用户对自己的检索意图不明确,或试图进行查询扩展时具有重要意义。例如用户原意是检索有关“网络上进行数据挖掘”方面的来源文献,但不知如何表达,只知道与“数据挖掘”相关,在

这种情况下,主题语义扩展可以指导用户找到合适的主题。图3中的“网络数据挖掘”以及“Web挖掘”等都可以成为用户的候选检索关键字,甚至可以进一步查看它们的关联主题,发现“Web挖掘”的交叉主题中“网络信息挖掘”也是符合用户检索需求的主题。基于关联主题的语义扩展为用户做出正确选择提供了有意义的知识导航。

类似“主题”,当用户使用“关联作者”、“关联期刊”、“关联部门”等对象属性作为检索项时,也可以根据用户提供的检索词显示与其关联的实例供用户选择,实现查询扩展。例如,图4以图形化方式展示了学者“杨建林”的交叉知识推荐,以连线的粗细表达关联学者对于中心学者的关联程度,连线上标以数字表明具体的关联度。在图4列出的知识中,可以发现与中心学者研究方向类似的、且用户可能关心的其他学者,并可进一步了解该学者的关联情况,实现在学者知识地图中的任意导航。

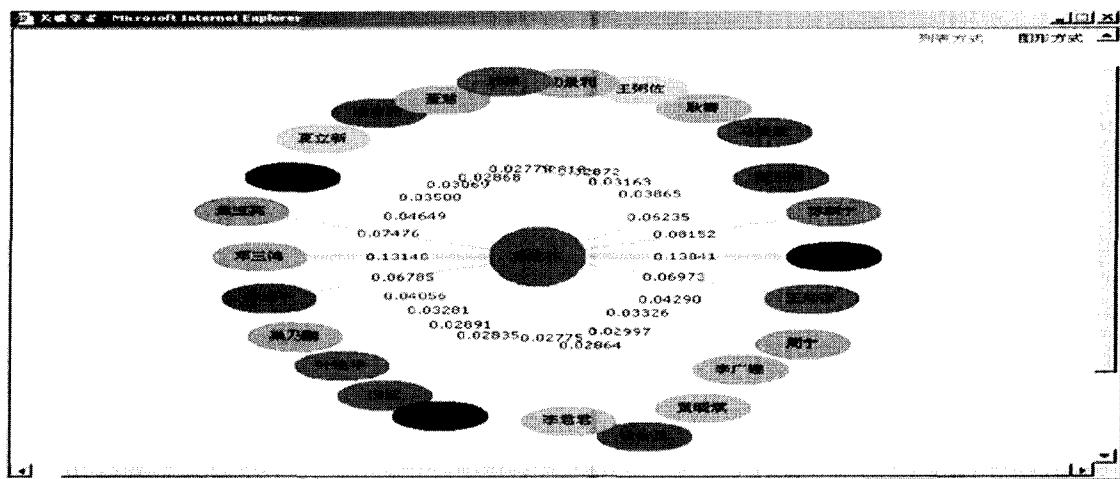


图4 图形方式展示“杨建林”的关联学者

## (2) 检索结果的语义扩展

除了在输入端对用户检索词进行关联推荐外,也可以根据实例间的关系,在结果返回端对用户的检索结果进行语义扩展。这里的检索结果并不是指满足检索表达式的所有返回对象的集合,而是指用户进行重点考察,浏览其属性信息的单个对象。基于本体中定义的对象间关系,可以提供用户可能感兴趣的相关对象。例如,用户通过检索定位到“苏新宁”作者实例,此时系统认为用户可能对与“苏新宁”相关的其他作者也感兴趣,如“邱均平”、“许鑫”等,于是便以关联作

者的形式向用户进行推荐,供用户参考。本文以来源文献检索为例,阐述检索结果语义推荐的具体实现。

来源文献是 CSSCI 中的重要学术资源。通过对来源文献的检索定位,用户可以了解指定文献的作者、出处(包括期刊、发表时间以及页码)等基本信息。然而用户检索来源文献的目的往往是为了得到一组内容相似的文献用于某课题研究。通过构建合理的检索表达式可以达到目的,但这要求用户具有较高的课题理解程度和检索能力。大多数的普通用户希望当他们查询到一篇合适的文献时,即可以同时获得内容类似的其

他文献,这就要求检索系统能够实现合理的语义扩展。在 CSSCI\_Onto 中,“来源文献”概念具有“引证文献”、“被引证文献”、“相似文献”、“同引证文献”以及“同被引文献”等属性<sup>[16]</sup>,即每一个来源文献实例都有 5 组文献集合从不同角度与其保持内容关联,这些文献可以作为查询扩展推荐给检索用户。

序号	标题	作者	第一作者	来源期刊	卷(期)	被引次数	相似度
1	中国社会科学引文索引设计(作者:苏新宁,2000)	苏新宁	苏新宁	《中国社会科学》	2000(1)	1	0.0000
2	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
3	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
4	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
5	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
6	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
7	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
8	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
9	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
10	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
11	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
12	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000
13	中国社会科学引文索引设计(作者:董慧,2003)	董慧	董慧	《中国社会科学》	2003(1)	1	0.0000

图 5 来源文献检索结果的语义推荐

如图 5 所示, KRSP\_CSSCI\_Onto 从 CSSCI\_Onto 中获取用户选定文献“中国社会科学引文索引设计(作者:苏新宁,2000)”的相关文献,图中显示的是 13 篇相似文献集合。可以通过选择“引证文献”等标签,获得对应的关联文献集合,并允许用户转移到指定的文献中,实现知识导航。除了向用户推荐关联文献外,还可以向用户提供其他知识服务,包括:

①推荐当前文献第一作者的关联作者、第一作者机构的关联机构、所在期刊的关联期刊以及当前文献所属学科的关联学科等,这些可能都是用户感兴趣的知识,例如用户可能会试图到关联作者发表的文献中寻找与当前文献主题类似的文献;

②根据当前文献的引证和被引证情况,绘制被引证曲线和引证树,图 6 为“基于本体论和数字图书馆的信息检索(作者:董慧,2003)”的被引证曲线,明确显示了该文献引用了 1995、1998、1999 以及 2001 年各一篇文献,一共被引 12 次,其中 2005 年被引 7 次,2006 年被引 5 次;而引证树则以树图方式展示文献的引证情况或被引证情况,如图 7 所示。

类似的,基于本体中描述的来源作者之间、来源期刊之间、主题之间以及部门之间等的关系,在定位到用户需要的实例时,可以向用户推荐其相关实例,因为这些对象都可能是用户的潜在目标。

## 4.2 基于蕴含推理的学术知识关系检索

KRSP\_CSSCI\_Onto 能够提供学术资源实例知识之间“关系”检索的功能。例如,输入两个来源作者(以前者为主体,后者为客体),系统能够自动返回两者的关联信息:若两者存在直接关联,即后者是前者的关联

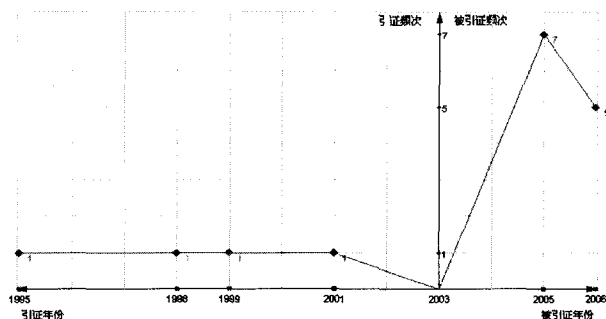


图 6 来源文献检索结果的被引证曲线示例

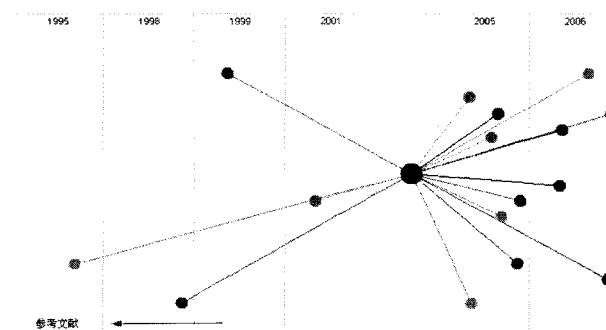


图 7 来源文献检索结果的引证树示例

作者,则返回关联次数、关联度等;若两者存在间接关联,如后者是前者的同被引作者、被引作者等,则返回关联路径,即表明通过哪些实例建立了两者的间接关联。

实例知识关系检索的基本过程如图 8 所示:

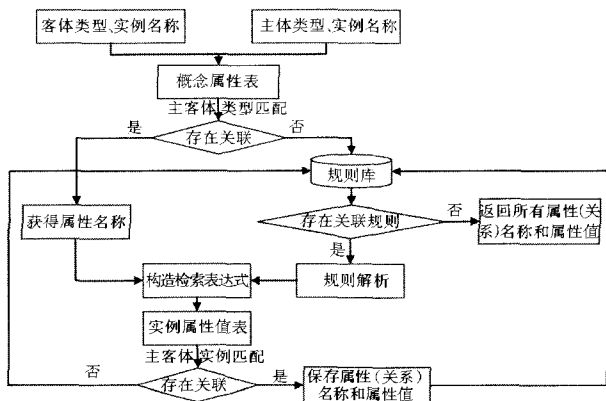


图 8 实例关系检索流程图

(1)根据用户选择的主客体类型,检索概念属性表(包含在本体概念模型中),若客体概念是主体概念的对象属性,则表示客体概念对主体概念存在直接关

联,进而执行(2),否则执行(4);

(2)根据主客体之间的关联属性名称或规则的解析结果构造检索表达式,由于实例之间的关系可能不止一种,因此可能要构造多个检索表达式;

(3)对实例属性值表(包含在本体实例库中)执行检索表达式,以获取实例之间所有存在的关联,并保存实例属性以及相应的属性值;

(4)在推理规则库中查找是否存在适应主客体概念的推理规则,若存在则依次进行规则解析后执行(3),直至符合条件的规则全部使用完毕后,返回所有保存的实例属性(即关系,包括本体中存在的关系和在规则库中定义的关系)以及对应的属性值。

以查询来源作者实例知识间关系为例,输入主体“杨建林”,客体“邓三鸿”。系统在查询了概念属性表后认为来源作者之间存在名为“关联作者”的关系,随即到“关联作者”实例表中查询两个作者实例之间的关联度,可以返回结果“关联作者(0.13140)”,其中数字为客体实例对主体实例的关联度;此后系统会继续搜索规则库,查询在规则库中定义的新关系,例如蕴含推理规则“( ? A 发表文章 ? X ), ( ? B 发表文章 ? Y ), ( ? X 是同被引 ? Y )  $\rightarrow$  ( ? A 是同被引 ? B )”表示作者 A 发表文章 X, B 发表文章 Y, 而且 X 是 Y 的同被引文献,那么可以推出 A 是 B 的同被引作者,这条规则适合两个作者实例,于是系统对推理规则进行解析,构造查询语句对实例库进行查询,查看“邓三鸿”是否是“杨建林”的同被引作者,如果是,则返回关联的整个路径,否则认为两者在同被引关系上无关,按照这种方式遍历规则库中的所有规则(与作者相关的基于规则推理的新关系还包括合作、被引、引用、共引等);最后返回所有的查询结果,如图 9 所示:

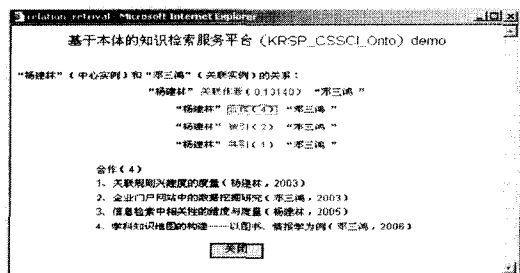


图9 实例知识关系检索结果示例

图9中,下半部分为点击“合作(4)”时的关联路

径,即“杨建林”和“邓三鸿”通过4篇文献建立了合作关系。

在 CSSCI\_Onto 中为来源作者、来源期刊、学科、部门等概念建立了18条蕴含推理规则,用于概念同类实例(知识)之间的关系推理。

## 5 结语

现有的 CSSCI 信息服务系统只能提供单一文献对象的检索,机械地接收用户查询并返回毫无推荐意见的结果,对于一些检索意图不是很明确的用户而言,很难找到其所需的信息。造成检索效果不佳的主要原因在于后台数据组织的不合理。为了改善用户的检索环境,提供一种知识层次的服务,笔者设计了一个实用的基于 CSSCI\_Onto 的知识服务平台 KRSP\_CSSCI\_Onto,详细阐述了该平台的基本功能和系统框架,并据此开发了一个试验性系统。本文探讨了本体在提供和改善信息检索服务,实现知识检索服务中的具体应用,包括基于本体中定义的实例间的语义关系,对检索式进行引导式的语义扩展,帮助用户构造满足检索意图的检索式;根据查询的结果返回本体中的关联知识,推荐给用户作为查询扩展;基于网络化组织的各种学术资源知识,引导用户查看所有关联资源,实现知识导航;查询学术知识之间的关系,包括在本体中直接定义的关系和在规则库中定义的推理关系。

在 KRSP\_CSSCI\_Onto 检索入口的设计中,本文采用了最简单的基于表单的查询式输入方式,即由用户根据系统提示对检索式进行语义标注,完成查询本体设计。事实上,为了减少用户对检索式构造的思考,可以考虑更加智能的检索入口:即由用户输入使用自然语言表达的检索需求,然后使用自然语言处理技术和本体概念模型对检索语言进行机器语义标注,由系统自动完成查询本体构造。笔者认为自然语言式的检索方式是基于本体信息检索系统的发展趋势,随着汉语自然语言处理技术的完善,这方面的探索也有待于进一步深化。

## 参考文献:

- [1] 南京大学中国社会科学评价中心[EB/OL]. [2011-01-02]. <http://cssci.nju.edu.cn>.
- [2] Semantic Web Roadmap[EB/OL]. [2011-01-02]. <http://>

- www.w3.org/DesignIssues/Semantic.html.
- [3] 杨建林. 基于本体的文本信息检索研究[J]. 情报理论与实践, 2006, 29(5): 598-601.
- [4] A General KIM Platform Presentation [EB/OL]. [2011-01-02]. <http://www.ontotext.com/kim/KIMPlatform.pdf>.
- [5] Franz B, Calvanese D, Deborah M, et al. The Description Logic Handbook: Theory, Implementation and Applications[M]. Cambridge, UK: Cambridge University Press, 2003: 1-100, 436-459.
- [6] 董慧, 余传明, 杨宁, 等. 基于本体的数字图书馆检索模型研究(Ⅲ)——历史领域资源本体构建[J]. 情报学报, 2006, 25(5): 564-574.
- [7] 王昊, 苏新宁. CSSCI 本体概念模型的构建和描述[J]. 中国图书馆学报, 2009, 35(3): 43-51.
- [8] Wang H, Su X, Liu F. Semantic Annotation for CSSCI Academic Resources Based on Ontology[C]. In: *Proceedings of the 6th International Conference on Fuzzy Systems and Knowledge Discovery*, Tianjin, China. 2009: 374-385.
- [9] 王昊, 苏新宁. 基于 CSSCI 本体的学科关联分析[J]. 现代图书情报技术, 2010(10): 10-16.
- [10] Wang H, Su X, Deng S. Study on Hotspots Detection Based on CSSCI Academic Resource Ontology [J]. *Advanced Materials Research*, 2011, 171-172: 19-26.
- [11] 余传明. 语义检索的原理及其实现[J]. 情报理论与实践, 2007, 30(2): 182-184.
- [12] 朱庆生, 邹景华. 基于本体论的论文检索[J]. 计算机科学, 2005, 32(5): 172-173, 176.
- [13] 陈康, 武港山. 基于 Ontology 的信息检索技术研究[C]. 见: NCIRCS2004 第一届全国信息检索与内容安全学术会议. 2004: 99-105.
- [14] 马文峰, 杜小勇. 知识检索研究[J]. 情报理论与实践, 2006, 29(2): 157-160, 219.
- [15] 张佩云, 孙亚民, 吴江. 基于本体的知识检索研究及实现[J]. 情报学报, 2006, 25(5): 553-558.
- [16] 王昊. 信息资源网络模型及应用[M]. 南京: 南京大学出版社, 2010: 108.
- (作者 E-mail: ywhaowang810710@sina.com)

## ALA 发布“我们的作家, 我们的拥护者”宣传工具

2011 年 1 月 7 日, 美国图书馆协会 (ALA) 发布了由“我们的作家, 我们的拥护者”制作的公益广告视频, “我们的作家, 我们的拥护者”是一场全国图书馆为启发公众意识而发起的宣传活动。图书馆的拥护者将可以从 [ilovelibraries.org](http://ilovelibraries.org)、诸如 Sharon Draper、Brad Meltzer、Sara Paretsky 和 Scott Turow 这些畅销书作家那里下载公益广告音频和视频。该网站还将提供静态的、可定制的英语和西班牙语的公益广告脚本, 用以协助图书馆的拥护者。

随着图书馆资源需求的持续高涨, 大量的预算削减无疑迫使许多图书馆做出艰难的抉择。大多数 (56%) 公共图书馆的运作经费预算报告与 2010 年年度财政持平或下降, 仅有略超过 40% 的图书馆比 2009 年年度财政稍高一些。ALA2010 年关于美国图书馆状况的报告数据也显示, 41% 的州在财政预算中减少了公共图书馆的资金。由于在地方一级的削减资金, 使得图书馆的资金缺乏状况加剧——形成的雪球效应已迫使许多图书馆缩减时间, 减少人员配备, 进行收费和罚款, 或者探索私有化。

“我们的作家, 我们的拥护者”是 ALA 主席 Roberta Stevens 所采取的关键举措之一。通过“我们的作家, 我们的拥护者”, 作家将借助媒体采访、播客、公共服务宣言和其他营销材料支持图书馆。这些资料将免费提供给 ALA 成员在地方一级使用, 用以协助其进行宣传工作。

ALA 主席 Roberta Stevens 说: “作家了解图书馆所发挥的关键作用, 是困难时期图书馆的天然盟友。‘我们的作家, 我们的拥护者’将突出强调那些知名的、充满激情的作家, 他们将发表关于图书馆价值和维持图书馆服务重要性的言论”。

参与活动的作家录音时长不等, 有 15、20 和 30 秒等, 内容涉及图书馆资金、图书馆对教育的支持、图书馆的价值等问题。作家根据自身对图书馆的热衷程度分享其个人故事, 这些内容也将提供下载。

(编译自: <http://ala.org/ala/newspresscenter/news/pr.cfm?id=5932>)

(本刊讯)