

# Web 环境中的本体构建系统研究综析<sup>\*</sup>

杨 锐 汤怡洁 刘 毅 李 崑

(中国科学院国家科学图书馆武汉分馆 武汉 430071)

**【摘要】**对基于 Web 方式的本体构建系统进行特性分析,通过调研从语义协作、构建知识库、体系结构和可视化特征等几个方面展开深入研究,最后针对有代表性的基于 Web 方式的本体构建系统进行比较和分析。

**【关键词】**本体构建 语义协同 知识库 可视化 系统架构

**【分类号】**G250

## Comprehensive Evaluation of the Ontology Building System in the Web Environment

Yang Rui Tang Yijie Liu Yi Li Wei

(Wuhan Branch of National Science Library, Chinese Academy of Sciences, Wuhan 430071, China)

**【Abstract】**This paper researches the Ontology building system in network environment. Firstly analyzes the characteristic of Web - based Ontology building system, then makes an in - depth study in several aspects, such as the semantic synergy, knowledge base building, architecture and visualization features. Finally, some representative Web - based Ontology building systems are compared and analyzed.

**【Keywords】**Ontology building Semantic synergy Knowledge base Visualization System architecture

随着 Web2.0 技术的应用和发展,互联网已经成为用户信息交流的主要环境,本体在语义网中扮演的关键角色使其覆盖范围越来越广泛,已经在一系列领域应用中体现出重要性。与此同时,本体已经不能由单个人或单个团队进行有效开发,它需要一种基于网络协作的本体环境。在这种情况下,互联网平台成了本体构建系统的首选。基于 Web 方式的本体构建系统是本体使用过程中的一个重要环境,只需要网络环境的辅助就能让领域专家和普通用户不受时空和系统环境的限制在线随时编辑本体和进行本体管理,利用这种方式能够较好地构建和维护相对完善的本体。

### 1 基于 Web 方式的本体构建系统特性分析

很多机构都开发了本体构建系统,通过它们能够较好地降低本体开发难度和加速开发过程。但是目前基于 Web 方式的本体构建系统还很少,使得网络群体在参与本体构建方面的表现很沉默,不利于本体的动态演化<sup>[1]</sup>。基于 Web 方式的本体构建系统主要呈现以下特点:

(1) 使得在网络环境下基于语义网的发展和本体构建人员的需求从静态封闭的构建方式向动态协作的构建方式进行转变。在这一转变过程中,语义协作开发机制使不同的本体构建人员能够共同合作进行广泛交流,降低本体构建难度。

收稿日期: 2011 - 11 - 28

收修改稿日期: 2011 - 12 - 12

<sup>\*</sup> 本文系国家科技支撑计划课题科技知识组织体系共享服务平台建设子课题“科技知识组织体系(STKOS)的开放查询和推理接口建设”(项目编号:2011BAH10B03 - 5)的研究成果之一。

(2) 需要良好的知识库机制和系统体系结构,主要用于辅助本体的创建、编辑、保存或处理以及本体元数据管理等功能。通过知识库机制和良好的系统体系结构保证在一定的生命周期过程中本体自身能够不断创新和演化。

(3) 通过良好的可扩展体系结构能够提供友好的本体构建用户界面,利用可视化技术带来的直观视觉效果可以帮助用户更好地导航和查询本体,从而满足构建本体的各种需求,提高用户满意度。

## 2 基于 Web 方式的本体构建系统理论研究

通过对 Science Direct、Springer、IEEE、AAAI(国际人工智能协会)、中国知网等数据库及相关国际大型项目的研究分析发现基于 Web 方式的本体构建系统研究的期刊论文和会议论文较少,还处于发展阶段。国外综合研究比较少,仅仅针对单个系统的研究比较深入,而国内的研究主要针对传统的非 Web 方式的本体构建系统。研究内容主要包含某个本体构建系统的介绍说明,以及针对某个学科方向的研究应用,比如 WebProtégé 的介绍说明、本体构建系统在知识管理中的应用等。目前基于 Web 方式的本体构建系统都具有自身的特点,研究人员都在不断使之完善,以求为用户提供更加易用的界面和更加完备的功能。

### 2.1 基于 Web 方式的语义协同特性

Web2.0 技术的普遍应用,使各个领域本体发展不断壮大,其中个人或小型团体的开发逐渐变得困难,协同支持成为基于 Web 方式的本体构建系统的特点。语义协同支持主要利用网络开发环境,支持物理上分散的用户共同合作构建本体。这种方式的特点主要表现在:

(1) 负载性强:系统和用户之间能够进行友好交互,使大量用户可以同时访问,降低了使用难度,使熟悉一般网络的普通用户也能和领域内的专家一起对本体进行构建;

(2) 开放性强:与以往对用户了解本体操作要求较高的本体构建系统不同,它不需要安装软件就能浏览运行,提供了一个使不同用户都能够轻易定制本体的语义协同平台;

(3) 同步性强:允许多个用户同时在同一个本体上进行构建工作,即系统通过控制协调策略,采用版本

控制机制能够处理不同用户构建同一个本体时的冲突问题。

过去,构建本体需要计算机科学家和软件工程师的帮助,用户也必须具有表达本体的丰富经验,这些构成了领域本体发展的巨大瓶颈。而现在基于 Web 方式的本体构建系统大大降低了用户门槛,面对初学者和有经验的用户,系统通过用户界面参数使不同级别的用户能够设置符合自己水平的个性化平台。同时,有一些本体构建工具还包含了插件功能,可以轻松扩展用户界面的功能。美国南加州大学信息科学研究所(USC/ISI)开发了基于 Web 方式的本体浏览工具 Ontosaurus,它提供了一个与 Loom 知识库链接的图形接口以及相应的知识库编辑功能。在用户进入编辑模式后,将阻止其他希望进入本体编辑模式的用户操作。这样就保证了多个用户能够协同对一个本体进行操作。同时该系统还具有有一致性自动校验、推理和多重继承等功能<sup>[2]</sup>。

### 2.2 基于 Web 方式的本体构建知识库

本体构建工作不同于一般的数据处理,在 Web 协作支持环境下更需要一个足够庞大、内容丰富的知识库来处理本体相关信息。本体构建系统的知识库服务器主要用于处理本体操作,将本体的存储和维护工作从客户端分离出来,减轻客户端的负担。同时在服务器上创建不同的本体构建功能模块能够实现对本体的一系列比较复杂的操作,例如检索知识库、一致性和隐含关系验证<sup>[3]</sup>。知识库不仅为本体构建系统提供理论知识和事实数据,其中包含的定理和运算法则还能为推理引擎提供语义问题的解决途径,从而检测出本体错误和冗余,还能够通过知识库不同版本比较来查看本体构建过程中产生的变化<sup>[4]</sup>。美国斯坦福大学开发的 Ontolingua Server 系统具有一个资源丰富的知识库,用户通过整合知识库中的若干模块来形成一个新的本体。Ontolingua Server 系统可以使用户从知识库中浏览、检索、整合、定制和扩展本体。它提供知识库的同时也可以改进原有的本体,通过包含、多态、完善和限制 4 种方式从知识库中复用其已有本体<sup>[5]</sup>。

### 2.3 基于 Web 方式的本体构建系统体系结构

和其他基于 Web 方式的系统一样,基于 Web 方式的本体构建系统分为客户端和服务端两部分。它们无论采用 B/S 还是 C/S 网络结构,在系统体系结构方

面通常都采用三层(系统表现层、业务逻辑层和数据服务层)架构。三层结构将数据处理过程分为三个部分:

(1) 系统表现层通常由 JSP、JavaScript、JavaApplet 和 .Net 等技术实现<sup>[6]</sup>,通过网络发送构建本体等各种请求,由于功能模块被分布到业务逻辑层使得显示和业务分离,降低了客户端的载荷。

(2) 业务逻辑层中提供各种本体构建服务,如对本体的各种操作包括协作支持、版本控制、推理引擎和一致性验证,同时能够根据用户的需求将本体转换成

所需的格式导出。

(3) 数据服务层主要用来负责知识库数据的存储、检索、更新和优化。

三者相对独立,当某一部分改变时不影响其他部分,提高了系统模块的复用性,同时三层体系架构可以对某些层次进行扩展演变为 N 层结构。美国密歇根大学医学院的学者在对 OntoFox 系统的技术说明中对三层架构的实现方式进行了详细阐述,系统工作流程如图 1 所示:

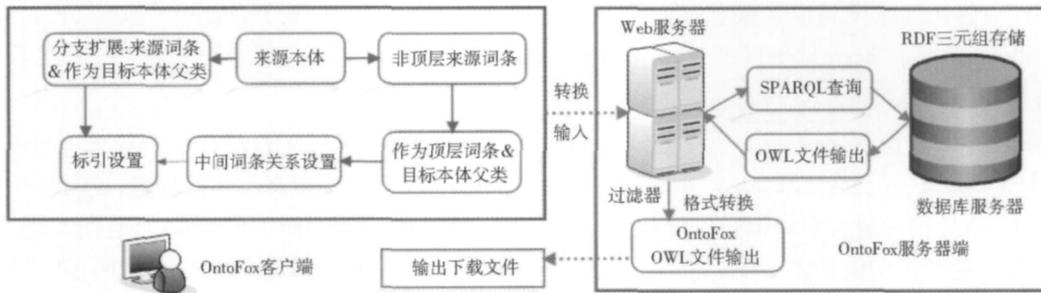


图 1 OntoFox 系统工作流程

OntoFox 服务器对 Web 方式输入的本体和词条按照功能要求进行内部转换,生成 SPARQL 查询语句访问 RDF 格式存储的三元组知识库,访问执行成功后通过 OntoFox 服务器生成 OWL 文件供用户下载<sup>[7-8]</sup>。

### 2.4 基于 Web 方式的本体构建系统可视化特性

基于 Web 方式的本体构建系统中本体表达视图的可视化技术目前处于发展阶段,在 Web 浏览器中可视化功能显得更加重要。可视化能够对本体中的各种概念、类、属性和关系,以及不同问题的状态、构建过程中发生的变化等用多种方式加以区分,从而协助本体概念的分类和概念之间关系的构建。可视化技术主要有 4 种表现形式:展开-折叠布局、源泉布局、曲线图和树状图,如图 2 所示:

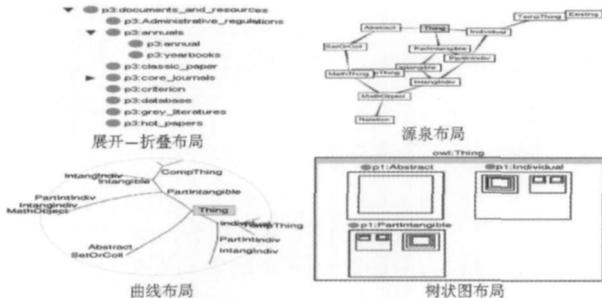


图 2 本体可视化技术<sup>[9]</sup>

每种形式各有其优缺点<sup>[9]</sup>:

(1) 展开-折叠布局是最常见的表现形式,实现简单,用户也比较熟悉。但是由于受屏幕高度限制和上下关系的单一性,无法适用于大规模的本体构建。

(2) 源泉布局中作为根源的节点不断吸引周边节点,最后形成节点群。然而这种布局不能反映节点连接的紧密程度,其在全局位置中的定位也不可预测。

(3) 曲线图布局在上述技术的基础上,遵循节点的连通性关系来布置节点。缺点是这种基于几何学的映射容易造成焦点转移和降低边缘标签的可读性。

(4) 树状图布局的嵌套视图可以让用户使用鼠标点击放大本体的分支,缺点是经过几次焦点的变更,容易使得用户迷失方向,节点的深度也难以直观确定。

### 3 基于 Web 方式的本体构建系统项目分析

基于 Web 方式的本体构建系统目前很少,但呈加速发展的趋势。它为本体创建和管理提供了良好的用户环境,能够支持高品质本体开发。同时这些工具在设计理念、体系结构、标准规范、实现技术和系统规模等很多方面都不尽相同。

#### 3.1 项目介绍

笔者选取了国外的三个具有代表性的基于 Web

方式的 本体构建系统,分别是美国斯坦福大学开发的 WebProtégé 系统、英国开放大学开发的 WebOnto 系统和巴基斯坦国立大学开发的 TODE 系统,从系统整体和特性等方面进行分析和比较。

### (1) WebProtégé 系统

WebProtégé 是成熟的基于 Web 方式的本体构建开源系统,能比较好地支持网络环境下联合编辑本体的工作,支持 RDF(S) 和 OWL 等格式语言。它使用 Google Web Toolkit (GWT) 作为系统开发框架,以 Protégé 组件作为后台用来支持本体服务。WebProtégé 通过调用 Collaborative Protégé<sup>[10]</sup> 组件实现本体编辑,使不同的用户通过 WebProtégé 系统进行协同操作,同时能够查看其他用户的修改记录情况。

### (2) WebOnto 系统

WebOnto 是由英国开放大学知识媒体研究所开发的基于 Web 方式的本体编辑系统。它能够提供更比 Ontolingua 系统更为复杂的浏览、可视化和编辑能力。WebOnto 系统基于可操作概念建模语言 (Operational Conceptual Modeling Language, OCML) 的知识模型,提供多重继承和提供锁机制,支持用户合作浏览、创建和编辑本体。WebOnto 系统作为综合的本体编辑工作平台,可以减少本体建模以及与其他本体工具、本体应用程序进行本体交换等工作的工作量,同时为本体创建和编辑提供多用户协同工作的环境<sup>[11]</sup>。

### (3) TODE 系统

TODE (Tool for Ontology Development and Editing) 是由巴基斯坦国立大学研究人员共同开发基于网络的符合 W3C 标准的本体开发和构建系统。其最大特点不同于其他采用 C++ 或者 Java 开发的本体编辑系统,它采用 .Net 技术进行开发,通过 AJAX 技术提供简单操作界面实现知识域建模,包括类、实例、属性和关系<sup>[12]</sup>。同时本体的导出支持 RDF、OWL-Lite、N-Triple、N-3 和 RDBMS 等格式。TODE 系统通过运用明确定义的方法和易于使用的 GUI (Graphical User Interface) 来为本体开发人员提供更为简明的本体创建方式。

## 3.2 项目功能比较分析

针对以上三个系统,从特征方面做了详细的分析比较,如表 1 所示。

可以发现,这些系统都具有比较好的特性,都能够为本体构建创建多用户协同工作的环境。在本体基本

表 1 基于 Web 方式的本体构建系统特征比较

比较项	WebProtégé	WebOnto	TODE
最新版本	0.5Beta	V2.4	2.0rc5
免费软件	是	否	是
开放源码	是	否	否
体系结构	B/S	C/S	B/S
程序语言	Java	Java	.Net
输入语言	RDF, OWL	OCML	RDF, OWL-Lite, N-3, RDBMS, N-Triple
输出语言	RDF, OWL, Logic, CLIPS	Ontolingua, GXL, RDF(S), OIL, OCML	RDF, OWL-Lite, N-3, RDBMS, N-Triple
推理功能	支持	支持	不支持
本体库	不支持	支持	支持
可扩展性	支持	不支持	支持
版本控制	支持	不支持	不支持
多重继承	不支持	支持	不支持

概念的操作方面包括本体的编辑、导入和导出等,能够满足用户本体构建的基本需求。它们各自采用不同的技术现实,WebProtégé 采用 Java 的 GWT 框架构建,WebOnto 采用 Java 的 Swing 框架构建, TODE 采用 .Net 的 C#语言进行构建。

## 3.3 项目语义特性分析

三个系统应用方面也表现出各自的特性:

### (1) WebProtégé 系统

WebProtégé 以开源方式使用 GWT 为用户接口,并且使用 Protégé 来支持本体服务。同时利用 Collaborative Protégé 组件以及注释来扩大本体编辑环境,从而形成 WebProtégé 的整体架构。它的特性主要表现在:

①利用 GWT 框架将系统的每一个功能作为门户组件(如类、实例等)进行呈现。用户通过拖拽以及显示或隐藏组件实现布局和外观的个性化定制。中国科学院国家科学图书馆数字知识资源环境开放组织引擎系统里面集成的 WebProtégé 页面如图 3 所示:



图 3 WebProtégé 系统集成页面

②用户可以共享本体进行协同工作,通过 Collaborative Protégé 组件与本体服务器交互参与本体注释的操作。系统

在 CHAO (Changes and Annotations Ontology) 中定义了注释的结构, CHAO 包含了定义注释类型(如评论、提议)的类,以及用户在编辑本体时能够操作的不同修改类型的描述。

③WebProtégé 以文件形式保存本体,对于每一个本体 WebProtégé 服务器端将保留当前版本。当本体做出任何修改时,本体的版本号发生变化。同时每一位用户都有自己的版本号,它不同于服务器的本体版本号,用户将自身的本体信息传送给服务器时,服务器将返回给客户两个版本之间的差异数据。当客户端接收到变化时便会更新本体模型和本版本号<sup>[13]</sup>。

(2) WebOnto 系统

WebOnto 是一个与网络服务器相连接的 Java 应用程序,允许用户浏览和编辑本体知识模型,WebOnto 用一种图形化的概念建模语言 OCML 来编辑本体,不支持 W3C 推荐的 本体描述语言标准。WebOnto 支持用户合作浏览、创建和编辑本体。

①WebOnto 系统采用 C/S 的体系结构,用户通过注册能够登录系统对本体进行浏览和编辑,包括本体的类、属性、关系、实例和本体标注,同时能够以可视化图形方式显示本体,在图形上对本体进行操作。WebOnto 在进行协同编辑本体时,采用解锁机制进行知识模型的编辑。在 WebOnto 服务器端对知识模型进行锁定,保证在同一时刻只有一个人进行知识模型的编辑<sup>[14]</sup>。当服务器发生问题时,知识模型仍然保持锁定状态,只有系统管理员能够解除。

②WebOnto 系统是基于 OCML 语言开发的。OCML 是由英国开放大学知识媒体研究所开发的一种本体描述语言,其特点是通过函数、关系、类、实例和规则来规约知识系统的建模,它还包括定义本体及问题解决方法的机制。此外,该语言具有一定的分类推理 (Classification) 功能,能够提供诸如非操作性声明包含 (Inclusion of Non-operational Statements)、公理约定 (Specification of Axioms) 和关系约定 (Specification of Relation) 等功能。一组标准的 OCML 关系操作<sup>[14]</sup>如图 4 所示:

关系定义	关系定义说明
:iff-def	为包含参数的关系定义充要条件
:sufficient	为包含参数的关系定义充分条件
:constraint	为关系的定义和实例产生约束表达式
:def	通过约束表达式定义局部关系
:axiom-def	提供关系之间的关联描述

图 4 OCML 关系定义说明

(3) TODE 系统

TODE 采用 C#语言实现使它成为一个比较特别的

本体编辑系统,它的设计采用了目前较为流行的 MVC 的框架结构,整个系统不需要对业务逻辑做出很大的改变就能很好地适应用户的扩展。

①TODE 系统分为三层,包括模型层、视图层和控制器层。模型层主要包括本体存储模型,当本体模型发生改变时能够通知视图层进行响应。视图层主要负责处理所有本体编辑操作的描述逻辑。控制器层主要负责所有本体创建和维持所要求的业务逻辑的调度和实现。控制器层主要管理模型层和视图层之间的交互,所有本体编辑请求通过控制器接收,然后调用模型层进行业务逻辑处理,同时将处理结果转发给视图层进行业务逻辑处理。TODE 系统架构如图 5 所示<sup>[12]</sup>:

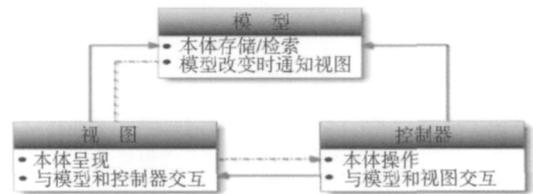


图 5 TODE 系统体系架构

②由于本体底层驱动方面不够成熟等原因,TODE 系统没有选择在 .Net 环境里面使用的 Owl - Dot Net API,而是采用 Java 的 Jena 库进行开发。相对来说,Jena 库更加成熟,它能够提供更完善的本体操作方法集合,可以用来进行本体构建、操作、推理以及支持 RDF、OWL 等本体语言<sup>[15]</sup>。由于 Jena 使用 Java 语言进行开发,它不能直接在 .Net 环境下运用,因此 TODE 系统使用 IKVM Conversion Utility 将 Jena 代码转换为 .Net 代码来解决这个问题<sup>[12]</sup>。

4 结 语

近年来,国外的大学、研究机构、公司企业对基于 Web 方式的本体构建系统研究的关注程度越来越高,它与语义技术一起成为各国重点研究的新兴技术之一。目前,基于 Web 方式的本体构建系统的开放性、组织性和协作性研究已经日趋完善,并积极应用于各种实践中,使得泛在知识环境下知识本体的创造更加易于实现。随着相关领域理论和技术的发展,基于 Web 方式的本体构建系统将最终创造和实现本体语义网络服务的无障碍应用。

参考文献:

[1] 张智雄,马建霞,文奕,等. 中国科学院国家科学图书馆战略研究调研成果[R]. 2008(9): 98 - 100.

- [ 2 ] Ontosaurus: Loom Web Browser [EB/OL]. [2011 - 11 - 20]. <http://www.isi.edu/isd/ontosaurus.html>.
- [ 3 ] Bechhofer S ,Horrocks I ,Goble C ,et al. OilEd: A Reasonable Ontology Editor for the Semantic Web [C]. In: *Proceedings of KI2001 , Joint German/Austrain Conference on Artificial Intelligence*. 2001:396 -408.
- [ 4 ] d' Aquin M ,Bouthier C ,Brachais S ,et al. Knowledge Editing and Maintenance Tools for a Semantic Portal in Oncology [J]. *International Journal of Human - Computer Studies* ,2005 ,62( 5 ) : 619 - 638.
- [ 5 ] 田晓迪. Ontolingua Server: 全球第一个本体服务器 [J]. *现代图书情报技术* ,2006( 2 ) : 21 - 25.
- [ 6 ] Thoméré J ,Barker K ,Chaudhri V ,et al. A Web - based Ontology Browsing and Editing System [C]. In: *Proceedings of National Conference on Artificial Intelligence*. 2002:927 - 934.
- [ 7 ] Xiang Z S ,Lin Y ,He Y Q. OntoFox and Its Application in the Development of the Brucellosis Ontology [C]. In: *Proceedings of the International Conference on Biomedical Ontology* , New York ,USA. 2011:304 - 306.
- [ 8 ] Xiang Z S , Courtot M , Brinkman R R ,et al. OntoFox: Web - based Support for Ontology Reuse [J]. *BMC Research Notes* 2010 , 3: 175 - 186.
- [ 9 ] Liebig T ,Noppens O. OntoTrack: A Semantic Approach for Ontology Authoring [J]. *Journal of Web Semantics* ,2005 ,3( 2 ) : 116 - 131.
- [10] Tudorache T ,Noy N F ,Tu S ,et al. Supporting Collaborative Ontology Development in Protégé [C]. In: *Proceedings of the 7th International Semantic Web Conference( ISWC 2008)* ,Karlsruhe ,Germany. 2008.
- [11] WebOnto [EB/OL]. [2011 - 11 - 20]. <http://projects.kmi.open.ac.uk/webonto/>.
- [12] Islam N ,Siddiqui M S ,Shaikh Z A. TODO: A Dot Net Based Tool for Ontology Development and Editing [C]. In: *Proceedings of the 2nd International Conference on IEEE Computer Engineering and Technology ( ICCET)* . 2010:229 - 233.
- [13] Tudorache T ,Vendetti J ,Noy N F. Web - Protégé: A Lightweight OWL Ontology Editor for the Web [OL]. [2011 - 11 - 20]. [http://bmir.stanford.edu/file\\_asset/index.php/1712/BMIR-2008-1467.pdf](http://bmir.stanford.edu/file_asset/index.php/1712/BMIR-2008-1467.pdf).
- [14] Domingue J ,Motta E ,Garcia O C. Knowledge Modelling in WebOnto and OCML a User Guide [OL]. [2011 - 11 - 20]. [http://projects.kmi.open.ac.uk/webonto/user\\_guide.2.4.pdf](http://projects.kmi.open.ac.uk/webonto/user_guide.2.4.pdf).
- [15] McBride B. Jena: A Semantic Web Toolkit [J]. *IEEE Internet Computing* 2002 6( 6 ) : 55 - 59.

( 作者 E - mail: yangruiwuhan@126.com)