

装备领域中语义桌面上的个人主观本体研究

向凯全,王盼卿,陈军广,张剑炜

(军械工程学院 计算机工程系,河北 石家庄 050003)

摘 要:随着计算机资源的爆炸性增长,传统的基于目录的资源管理方式已经不能很好满足用户的实际需求。语义桌面基于语义检索,提高了计算机资源查找的效率,并能挖掘出存在语义联系的不同资源。基于装备领域中各业务参谋分管业务的不同,提出建立满足业务需求和适合个人习惯的个人主观本体,实现基于语义桌面的语义检索功能。应用语义网和本体的相关知识,提出了装备领域的语义桌面体系结构,采用“骨架法”建立了适合语义桌面的个人主观本体,在此基础上实现语义桌面检索功能,大大提高了资源的检索效率。

关键词:语义桌面;本体;个人主观本体

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2011)08-0125-04

Research of Personal Mental Ontology Based on Semantic Desktop in Equipment Domain

XIANG Kai-quan, WANG Pan-qing, CHEN Jun-guang, ZHANG Jian-wei

(Department of Computer Engineering, Ordnance Engineering College, Shijiazhuang 050003, China)

Abstract: Semantic desktop changes the resource management oriented by catalogue to semantic, improves the efficiency of retrieval and location of resources, and reflects the semantic relation of different resources. Based on the difference of managed services in equipment domain, in this paper, the idea to construct personal mental ontology fitted for affair need and personal habit and realize the semantic search function oriented by semantic desktop is proposed. Related knowledge of semantic web and ontology, in this article, semantic desktop architecture in equipment domain is proposed. Personal mental ontology is constructed, and semantic retrieval is accomplished.

Key words: semantic desktop; ontology; personal mental ontology

0 引言

随着军队信息化和装备信息化建设的发展,军队业务主管参谋的计算机资源呈爆炸性增长,基于目录的计算机资源管理方式已经不能很好满足用户的实际需求,主要表现在以下几个方面:

(1)业务计算机存储的信息资源越来越丰富,用户即使知道文件所在目录,也需要花费相当多的时间来定位所需文件;

(2)目录并不能完整反映资源或文件的分类结构;

(3)无法反映不同资源之间的语义联系。

2003年,德国的Leo Sauermann在其学位论文“Gnowsiss Semantic Desktop”中提出了语义桌面(Semantic Desktop)^[1]的概念,将语义网的相关技术应用扩展到个人计算机资源管理领域。从此,基于语义检

索的语义桌面技术迅速引起了人们的关注。最初,语义桌面的研究者们简单地将语义网的标准和技术移植到语义桌面系统中。随着语义桌面研究的深入和实际应用,人们认识到语义网和语义桌面的应用环境存在很大的差异,将语义网上的标准和技术直接应用到语义桌面中效果并不理想。Leo Sauermann提出应当在语义桌面中建立适用个人信息管理需要的本体及其本体语言^[2]。而对于装备领域而言,各业务主管参谋分管业务不同,所关注的信息内容也就不同。因此,需要针对业务不同,建立个人主观本体去构建语义桌面系统,管理计算机资源。

1 语义桌面

语义桌面的定义如下:“语义桌面是一个能够独立存储各种数字信息(如:文档、多媒体、消息等)的设备。所存储的信息都由语义Web资源来解释,被URI标注,所有这些信息都可以通过RDF图来检索或访问;该设备可用于存储或创建Web中的各种资源,并与其他用户的资源实现共享”^[3]。

收稿日期:2010-12-28;修回日期:2011-03-09

基金项目:总装综合计划重点项目

作者简介:向凯全(1982-),男,湖南溆浦人,讲师,研究方向为数据库系统开发和计算机系统安全。

尽管语义桌面的提出已经有很长一段时间,但目前而言,语义桌面系统的开发仍处于初级阶段。Gnows^[4]是最早且最具有代表性的语义桌面系统。它将结构化数据源当作虚拟的 RDF 图来进行处理,并提出了个人主观模型,使得桌面搜索结果更加精确和个性化。但该系统的实际应用研究不够深入。Haystack^[5]系统采用一种集成的方法,使用户用最有效的方式管理个人信息,可以作为字处理、邮件客户端、图片管理等应用程序的替代品。但该系统存在性能问题。

国内在最近几年开始有研究小组对语义桌面进行系统的研究,但研究成果不多。瞿裕忠领导的 XObjects 研究小组曾对语义桌面进行过相应研究,发表多篇研究论文。顾进广提出了语义视图的概念,利用数据挖掘技术、语义网技术等智能信息处理技术,为个人计算机提供一个基于本体的动态语义视图,建立了资源之间的联系,消除了不同应用程序之间的复杂度,提高了资源管理的效率及资源共享的能力。

笔者根据装备领域中业务主管参谋关注装备信息的不同,提出建立装备领域个人主观本体,开发语义搜索引擎,用语义桌面管理个人业务信息,提高资源管理的效率。

2 装备领域个人主观本体建立

2.1 装备领域本体特点

在语义桌面中,对本体一般划分为三种类型:资源本体、领域本体、个人主观本体^[6]。资源本体主要用来对文件资源进行标注和描述,说明该资源是什么,包含了对资源位置的引用^[7]。领域本体描述某一个领域的公共知识,它有着比较明确的定义和该领域特定内涵^[8-10]。个人主观本体描述单个用户主观概念中存在的东西。因此,对于同一个概念,不同的用户因其工作背景和个人习惯等不同,其个人主观模型常常存在着差别。

装备领域信息涉及多个装备业务和装备管理单位,各装备业务部门都按各自的工作需要及习惯确定数据需求。对某一类型的装备,由于其特点类似,可以采用一定的规范,建立通用的资源本体和领域本体信息(此部分工作,在前期的项目中已经完成)。对于同一型号装备,由于业务部门的不同,业务参谋关注的信息也不同。建立语义桌面时,其个人主观模型也就不

同,需要建立不同的主观本体。

2.2 装备领域语义桌面体系结构

根据装备领域中业务参谋管理的信息资源类型和所关注的信息特点,笔者提出了装备领域语义桌面体系结构,如图 1 所示。

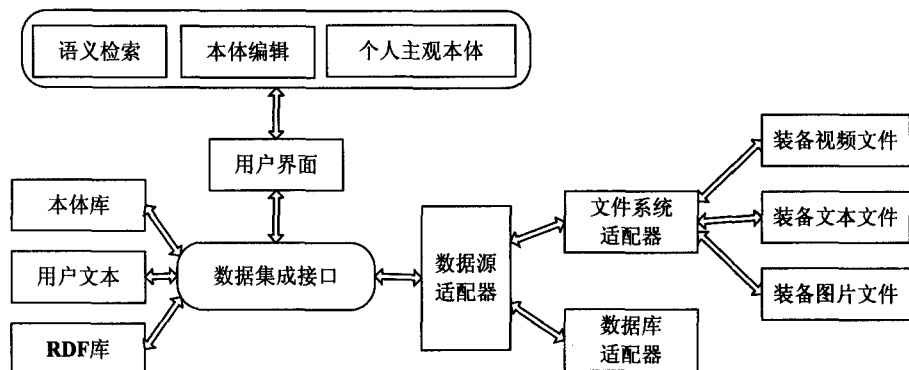


图 1 装备领域语义桌面体系结构

该结构包括以下几个部分:一是数据源适配器,二是数据集成接口,三是用户功能界面。数据源适配器实现从现有的装备数据源(文件系统、数据库系统)中提取数据、转换数据并存储到本体库中。数据集成接口作为通信接口,用来交换数据源适配器、本体库和用户界面的数据。用户界面主要用来编辑本体、建立个人主观模型和实现语义搜索功能。

文中的工作主要针对本体编辑和个人主观模型建立进行。首先建立资源本体和领域本体,然后根据主管参谋的业务特点,建立满足业务需求和适合个人习惯的个人主观本体,构建装备领域个人业务本体库,从而开发基于语义桌面的检索功能,提高计算机资源管理效率。

2.3 个人主观本体建立

在图 2 中,建立了空军装备大类下的歼-6A 和歼-6B 两具体型号装备的相关本体信息。在此处,基本信息属于领域本体的范畴。

在装备领域中,各业务主管参谋对同一个装备,所关注的信息可能是不同的,其主观模型也就不同,因此,需要建立个人主观本体。个人主观本体构建的主要任务是将个人关注的信息用本体描述语言表达出来,以方便实现语义桌面的信息检索,增强用户的满意度和使用效率。

这个阶段使用 OWL 本体描述语言创建个人主观本体,对概念和关系进行形式化描述,定义类、子类、属性和它们各自具有的特性^[11]。

在此,以作战参谋为例,其关注的是装备的作战性能,对飞机来说,比如作战半径、爬升高度等。在其个人主观模型中,装备也应该以作战半径等参数描述。在图 2 中,可以在基础信息下添加作战性能属性,表示

作战性能是基础信息属性的子属性,这也是作战参谋重点关注的信息,作为以后信息检索的关键词。

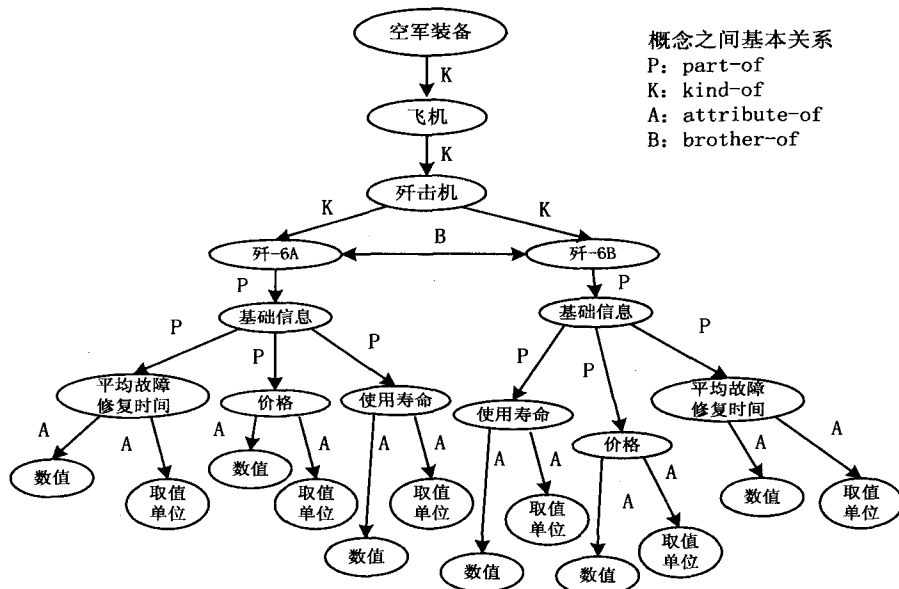


图2 装备基础信息本体图

```

<owl:ObjectProperty rdf:ID="作战性能">//对象属性
  <rdfs:subPropertyOf rdf:resource="#基础信息"/>//“作战性能”是“基础信息”的子属性
</owl:ObjectProperty>

<owl:DatatypeProperty rdf:ID="作战半径">//数据类型属性
  <rdfs:subPropertyOf rdf:resource="#作战性能"/>//“作战半径”是“作战性能”的子属性
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
</owl:DatatypeProperty>

<owl:DatatypeProperty rdf:ID="爬升高度">//数据类型属性
  <rdfs:subPropertyOf rdf:resource="#作战性能"/>//“爬升高度”是“作战性能”的子属性
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
</owl:DatatypeProperty>

```

这样,就建立了飞机这类装备的部分个人主观本体信息。在实际应用中,可以根据个人主观模型的改变,去修改个人主观本体的相关信息,进行个人主观本体的完善和进化,逐步建立满足业务需求和适合个人习惯的个人主观本体,提高信息检索的效率。

3 基于个人主观本体的语义搜索

本节介绍笔者设计的一个管理装备领域资源的语义桌面原型系统,其目的是改变传统的基于目录的计算机资源管理方式,使用户能够根据个人工作背景和习惯,从语义上管理个人计算机信息资源,并能发现不同的资源文件之间的语义联系,更有效地处理个人业务。

3.1 本体存储

对于装备领域的本体存储,我们设计了一个 RDF

存储库,用来管理资源本体、领域本体和个人主观本体信息。根据语义桌面的特点和系统开发需要,将个人计算机上待存储的本体进行分类,并将它们的定义与具体的实例分开进行存储。在所开发的原型系统中,基本存储单位是一个由 Subject, Predicate, Object, Type, Number 组成的五元组 (S,P,O,T,N)。其中,S 表示主语,P 是谓词,O 是宾语。T 描述的是该元组的类型。T 值为 0 表示本体的定义,1 表示文本实例,2 表示图片实例,3 表示音频实例,4 表示视频实例,9 表示数据库实例。N 代表检索次数,用来调整存储单位在本题库中的位置,提高检索效率。

同时,为提高资源检索效率,设置一个 64 位长整数 NodeID 作为索引,用来描述五元组 (S,P,O,T,N) 中的各个元素。参照了 kowari 的实现方法,建立了索引来对元素字符串进行排序和组织。

同时,为提高资源检索效率,设置一个 64 位长整数 NodeID 作为索引,用来描述五元组 (S,P,O,T,N) 中的各个元素。参照了 kowari 的实现方法,建立了索引来对元素字符串进行排序和组织。

3.2 语义检索

个人主观本体建立以后,得到了装备信息的语义数据,利用语义数据和 W3C 的 SPARQL 查询语言,就能为用户提供语义查询,实现对数据的精确查找^[12-17]。

例如,用户想要查询“内容是关于 Su-37 的所有图片”,对应的 SPARQL 语句如下:

```

<! ----名字空间声明,dc 和 foaf 为本体名字空间缩写---->
PREFIX dc: <http://www.w3.org/dc/elements/1.1>
PREFIX foaf: <http://www.w3.org/foaf/0.1/>
<! ----将图片 URI 从数据集中选取出来---->
SELECT? URI
<! ----WHERE 子句指明检索条件---->
WHERE {
  <! ----内容为 Su-37---->
  ? content foaf: source "Su-37".
  <! ----文件格式为图片---->
  ? file format foaf: type "2"
}

```

通过语义查询,用户无需关心文件的存放位置,只需要给出个人主观模型中的数据特征,例如文件类型、

关键字、作者等信息,就可以精确找到想要的信息。

3.3 功能实现

笔者在 C# 工作环境下,基于个人主观本体的相关信息,设计实现了装备信息的语义搜索功能。用户只要在搜索栏输入个人主观模型中的关键字,即业务参谋关注的装备信息,例如射程,点击“搜索”按钮,系统就能根据提供的语义,搜索出与之相关的资源文件,包括文档、图片、视屏等文件,在文件列表中列出。用户点击想关注的文件,在下方的文件信息中就会展示该文件的详细信息,并提供文件位置链接功能。用户点击后直接转到该文件的具体存储位置。

通过语义搜索,用户无需知道文件的存放位置,只需给出个人主观模型中的数据特征,就可精确地找到想要的信息。并且,通过语义搜索,可将存在语义关联的资源显示出来,揭示了这些资源之间的语义联系。

4 结束语

语义桌面系统为了提高个人信息检索的效率,对本体的存储和管理提出了新的需求。为此,根据作战参谋的业务特点,笔者建立了适合作战需求的个人主观本体。在完成个人主观本体的构建之后,笔者设计完成语义桌面的检索功能。随着本研究的深入,语义桌面在装备领域中的应用,必将大大提高业务主管参谋的工作效率,对我军信息化建设也是一个很好的补充。

参考文献:

- [1] Sauermann L, Bernardi A, Dengel A. Overview and outlook on the semantic desktop [EB/OL]. 2008-04-18. http://www.ischool.utexas.edu/~i385t-sw/readings/Sauermann-2005-Semantic_Desktop.pdf.
- [2] Sauerman L, Grimnes G, Kiesel M, et al. Semantic Desktop

2.0: The gnowsiss experience [C]//Pro. of the ISWC Conference. [s. l.]:[s. n.], 2006:887-900.

- [3] Sauermann L. The gnowsiss—using semantic web technologies to build a semantic desktop [D]. Viennea: Technical University of Vienna, 2003.
- [4] Sauermann L, Grimnes G A. Semantic Desktop 2.0: The Gnowsiss Experience [C]//The 5th International Semantic Web Conference at the ISWC. [s. l.]:[s. n.], 2006.
- [5] Quan D, Huynh D, Karger D R. Haystack: A platform for authoring end user semantic web applications [C]//International Semantic Web Conference. [s. l.]:[s. n.], 2003.
- [6] 余翔宇. 语义桌面上的本体存储研究 [J]. 计算机科学, 2008, 35(8): 158-160.
- [7] 刘琼, 李宝敏. 一种果品领域本体库的构建方法 [J]. 计算机技术与发展, 2009, 19(1): 197-199.
- [8] 邢玉萍, 吴飞, 邢桂芬. DW 中一种基于本体的语义集成方法研究 [J]. 微计算机信息, 2007, 23(4): 187-189.
- [9] 鲁宁, 刘磊. 本体的语义变化表示策略 [J]. 吉林大学学报, 2009, 47(5): 969-976.
- [10] 赵夷平. 传统搜索引擎与语义搜索引擎服务比较研究 [J]. 情报科学, 2010, 28(2): 265-270.
- [11] 王金环, 李宝敏. 基于本体 DL 的语义推理研究 [J]. 计算机技术与发展, 2009, 19(11): 94-100.
- [12] 花开明, 陈家训, 杨洪山. 基于本体与元数据的语义检索 [J]. 计算机工程, 2009, 33(24): 220-224.
- [13] 陈红红, 李辉, 李新春. 基于领域本体的概念格语义匹配 [J]. 郑州大学学报, 2010, 42(2): 70-73.
- [14] 黄果, 周竹荣. 基于领域本体的概念语义相似度计算研究 [J]. 计算机工程与设计, 2010, 28(10): 2460-2463.
- [15] 时念云, 杨晨. 基于领域本体的语义标注方法研究 [J]. 计算机工程与设计, 2010, 28(24): 5985-5987.
- [16] 冉婕, 孙瑜, 昌霞. 基于 OWL 的成语典故本体构建研究 [J]. 计算机技术与发展, 2010, 20(5): 63-66.
- [17] 王继东, 张瑜, 李娜. 基于本体的语义检索技术研究与实现 [J]. 计算机技术与发展, 2009, 19(10): 134-137.

(上接第 124 页)

参考文献:

- [1] 郑少仁, 王海涛, 赵志峰, 等. Ad Hoc 网络技术 [M]. 北京: 人民邮电出版社, 2005.
- [2] 王金龙, 王呈贵, 吴启晖, 等. Ad Hoc 移动无线网络 [M]. 北京: 国防工业出版社, 2004.
- [3] 陈林星, 曾曦, 曹毅. 移动 Ad Hoc 网络: 自组织分组无线网络技术 [M]. 北京: 电子工业出版社, 2006.
- [4] 肖书成, 唐学文, 王康, 等. MANET 路由协议及其性能研究 [J]. 计算机工程与设计, 2004, 25(7): 1133-1136.
- [5] 屠样浩, 吴荣泉, 钱立群. 无线 Ad Hoc 网络 DSR 路由协议的优化设计 [J]. 计算机工程, 2009, 35(4): 97-99.
- [6] 张登银, 姬广芹. Ad hoc 网络中基于 DSR 的节能路由协议研究 [J]. 计算机技术与发展, 2009, 19(8): 15-18.

- [7] Kunz T. Ad-Hoc, Mobile, and Wireless Networks [C]//5th International Conference. Ottawa: Springer, 2007.
- [8] 柯志亨, 程荣祥, 邓德隽. NS2 仿真实验—多媒体和无线网络通信 [M]. 北京: 电子工业出版社, 2009.
- [9] 黄化吉, 冯穗力, 秦丽娇, 等. NS 网络模拟和协议仿真 [M]. 北京: 人民邮电出版社, 2010.
- [10] 常莉, 吴蒙, 王立. Ad Hoc 网络中能量有效路由协议的性能研究 [J]. 计算机技术与发展, 2010, 20(4): 47-50.
- [11] 胡汀, 裴廷睿, 朱晓瑜, 等. 使用新判据的改进型 DSR 协议 [J]. 计算机工程与应用, 2010, 46(8): 92-95.
- [12] Johnson D B, Maltz D A. Dynamic source muting in ad hoc wireless networks [M]//Mobile Computing. [s. l.]: Kluwer Academic Publishers, 1996: 153-181.