基于语义网的企业知识集成方法

谷文栋 夏国平 尤薇佳

(北京航空航天大学 经济管理学院, 北京 100083)

摘 要: 目前,企业知识往往分散于各个业务系统中,不能很好地实现知识管理的目标. 基于现存系统,应用语义网技术提出一个企业知识集成的系统框架模型,详细介绍该模型的数据层、业务系统层、知识集成层和知识应用层 4层架构构成,并对其中的数据知识化、知识的 RDF(Resource Description Framework)表示、知识链构建和知识查询 4项关键技术进行了研究. 最后,给出系统开发方法,并以某大型水利水电施工企业信息系统的建设为例,介绍了该体系框架的应用.

关 键 词: 知识工程; 语义网; RDF; 知识表示; 知识链; 知识查询

中图分类号: TP 182; TP 39

文献标识码: A **文 章 编 号**: 1001-5965(2007)02-0243-05

Method of enterprise knowledge integration based on semantic web

Gu Wendong Xia Guoping You Weijia

(School of Economics and Management, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: In most of today's enterprises, organization knowledge is located in separate operation systems, therefore the knowledge management is not well realized. Based on enterprise's legacy systems, an application framework which supports the EKI(enterprise knowledge integration) by applying semantic web technology was proposed. The architecture of the framework was divided into four layers from bottom to top: data layer, operation application layer, knowledge integration layer and knowledge service layer. Each layer was introduced in detail. Four key techniques for the architecture, i.e. data semantic transformation, unified knowledge representation by RDF (resource description framework)/RDFS (RDF schema), constructing knowledge chain and knowledge query were further studied. An approach with eight definite steps to develop the framework was also proposed for referenced. Finally, application of the proposed framework in a famous Chinese large scale hydropower corporation was further provided to illustrate the effectiveness of the framework.

Key words: knowledge engineering; semantics; RDF(resource description framework); knowledge representation; knowledge chain; knowledge query

知识集成 (KI, Knowledge Integration)最早是由 Demsetz[1]在 1991年提出的. 而正式提出知识集成概念的则是 Grant[2],他指出,“企业的第一角色,以及企业能力的本质,就是知识的集成”. 但迄今为止,知识集成还没有统一的概念. 国内学者顾新建和祁国宁两人给出了一种可供参考的知识集成定义[3],“知识集成一方面是对分散在企业中,主要是存在于企业员工头脑中的意会性知识进行集成,另一方面是对分散在不同协作单位间的知识进行集成”. 有关知识集成的理论和技术的研究,国内外已经取得了一定的成果. 文献[3]概述了国内外关于知识集成方面的主要研究,并对知识集成的概念、集成方式、集成模型和工具等进行了讨论;文献[4]探讨了基于认识论

收稿日期: 2006-03-14

基金项目: 国家自然科学基金资助项目(70371004)

作者简介: 谷文栋(1979-),男,河北邯郸人,博士生, wdgu@tom.com.

的知识集成的 SECI模式和基于本体论的 IIOI模式,并在此基础上建立了知识集成的三维模式;文献[5]重点探讨了在知识集成过程中如何更好地使用信息技术;本体论是指导知识集成的有效方法;文献[6]介绍了一种基于本体进行知识的组织、共享和查询系统的结构与实现,文中重点介绍了使用本地本体、全局本体以及概念词典完成知识查询的方法.这些研究对知识管理系统的应用做了有益的探讨和实践,但大多数仅关注于企业知识集成的单个环节.进一步的工作,需要从系统化的角度,研究如何把知识集成的研究成果应用到企业中,真正为企业创造效益.

本文从企业知识集成的实现方法出发,在企业现有业务系统的基础之上,结合使用最新的互联网研究成果——语义网技术,提出了一套支持企业知识集成的体系框架,包括框架的整体结构及其层次关系,数据知识化、统一的知识表示、构建知识链以及知识查询4项关键技术,给出了可供参考的系统开发方法,介绍了具体的应用实例.

1 企业知识集成框架

在过去的10多年里,随着计算机科学与互联网技术的高速发展,越来越多的学者致力于将知识管理理论与信息技术相结合展开研究[7].企业知识集成(EKI, Enterprise Knowledge Integration),正是这种结合的产物,它从企业整体角度出发,综合运用知识管理理论和方法,借助于先进的计算机和网络技术,实现知识的获取、存储、共享和重用,提高企业的技术创新能力,将企业的各种知识资源有机集成在一起,形成有机的整体.

本文提出了一套基于语义网技术的企业知识集成框架,如图1所示.

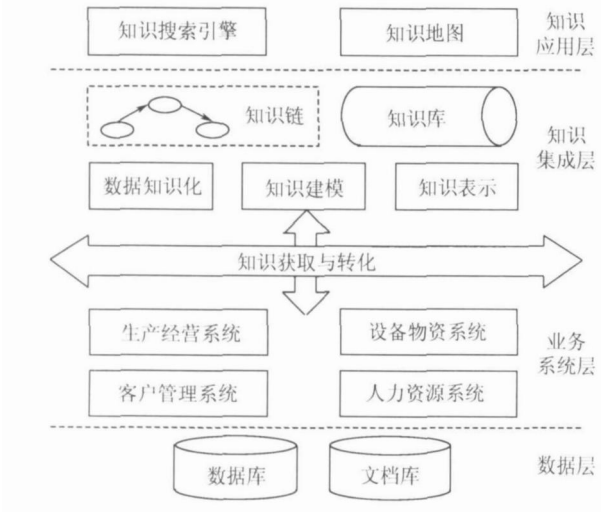


图1 企业知识集成体系框架

1.1 数据层

数据层是整个框架的基础,存储在企业业务系统运行过程中产生的各种数据和文档.数据层需要关注的是可以提升为知识的数据,它们以XML(可扩展标记语言)的形式进行存储.

1.2 业务系统层

业务系统层是企业各个业务系统的集合,如生产经营系统、设备物资系统等.知识的语义需要从各个业务系统中获取,并进行适当地转化,包括:本地知识向全局知识的转化、全局知识向本地知识的转化,以及本地知识之间的转化.

1.3 知识集成层

知识集成层是整个框架的核心,使用语义网技术实现.企业知识集成,是在企业应用集成基础上的进一步扩展.企业知识集成的最终目标是实现企业知识的共享、重用和创新,其中的一个关键问题就是使得知识可以被计算机自动分析、理解和处理,而这恰恰正是语义网所要实现的目标.

1.4 知识应用层

知识应用层是提供给知识工作者的具体应用,如知识搜索引擎和知识地图.这些智能化的服务,能够帮助知识工作者更有效地完成工作.

2 进行企业知识集成的关键技术

语义网(semantic web),被互联网的创始人Tim Berners-Lee称为互联网的第3阶段[8],其目标是通过扩展当前的互联网应用,将网络中的数据加入语义信息,使得计算机能够自动理解和处理,从而增强人与计算机之间以及计算机与计算机之间的交互与合作.

语义网的研究工作,已经取得了一定的成果,并产生了一些国际标准,如RDF(Resource Description Framework)和OWL(Web Ontology Language),下一步的研究热点必然是将其推向实际应用.在知识管理方面,目前研究多的是使用本体对知识进行概念化建模.本体是语义网的核心层次.使用语义网进行企业知识集成,必须解决以下4项关键技术.

2.1 数据的语义提升——知识化

目前,关系数据库是企业中进行数据存储和管理的最主要手段.但是,关系数据只能表达"属性/值"这样的二元关系,本身并不具备语义描述的能力.进行数据的语义提升,目标就是在关系数据中直接抽取数据的语义,弥补关系数据的缺陷.

众所周知,XML已经成为企业进行数据交换和转化的事实标准.但XML并不是严格的语义表

达语言,它缺乏严格的模型理论语义和逻辑基础,对于一份 XML文档所包含语义的解释依赖于解析 XML文档的程序. 使用语义网进行系统集成,对语义的理解是必须的要求. 在语义网层次关系中,进行知识表示的语言,如 RDF和 OWL等,都采用 XML作为其实现语法. 所以,从 XML数据中抽取语义信息,构建本体,是一个可行的方法. 文献[9]在这方面进行了有益的探索. 本文给出了数据的语义提升过程,如图 2所示. 该过程涉及到数据层、业务系统层和知识集成层 3个层次,是企业知识集成的一个关键过程.

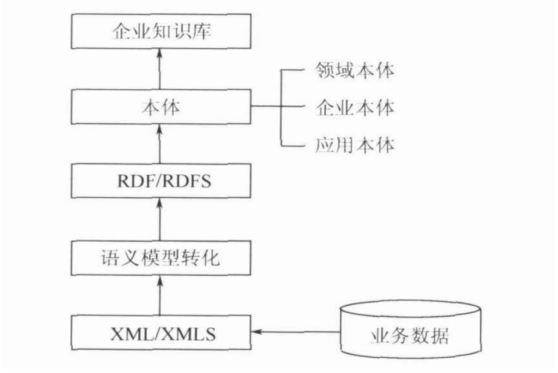


图 2　数据的语义提升过程

2.2　知识的 RDF表示

企业要做到在知识层次上的集成,必须使不同的应用系统对同一领域内知识的理解不存在二义性. 这需要借助于本体来实现. 被广泛接受的本体定义是由 Gruber提出的:本体是一个共享概念模型的明确形式化说明规范[10]. 每个领域必须建立领域本体,领域知识的表示在领域本体之上进行. 基于 Pérez的研究[11],定义领域本体的形式化表示如下:

$$D_O = \{V, I, R, F^D, A^D\}$$

其中,D_O 为领域本体集合;V 为领域中的概念集,通常,概念集以分类层次的形式组织;I 为领域中的实例集,实例,是某个概念类的具体实现;R 为概念之间的关系集,它应该满足条件:$V \cap R = \varnothing$, $R \rightarrow V = V \times V$;$F^D$ 为领域 D 内可以使用的函数集;A^D 为领域 D 内的公理集.

图 3给出了使用 RDF进行知识表示的实现方法. RDF是 W3C组织推荐使用的本体描述语言,它定义了用来描述资源及其之间关系的语义规范,以 XML为实现语言. RDF提供了一种简单但功能较为强大的数据模型,主要包括 3个对象类型:主题(subject),可以是除了基本数据类型之外的任何信息;属性(property),用来描述资源的特定的方面、特征或关系;声明(statement),一个特定的资源和属性名称加上该属性的值就构成了一个 RDF声明.

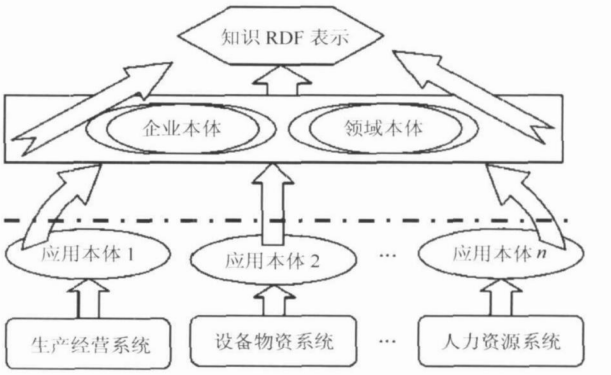


图 3　使用 RDF进行知识的统一表示

虽然对于企业知识表示,RDF仅提供了一些常见的模型限制. 但对于大多数企业知识集成过程,领域知识不会像互联网上的信息那样杂乱繁多,它相对更加集中,结构性更强. 因此使用 RDF来进行企业知识的表示是合适的,够用且足够简单. 而针对复杂度更高的知识建模问题,需要引入 OWL或者其他知识表示的方法,如文献[12]给出了一种面向 Web的知识描述语言.

2.3　构建知识链

所谓知识链,包括知识节点和关系 2个重要元素,多个知识节点以各种类型的关系建立起联系,构成知识的有序组合. 业务系统中的工作流是知识链的来源. 构建知识链,将工作流中的数据进行知识化(包括隐性知识向显性知识的转化、个人知识向组织知识的转化等),是实现企业知识集成的一个关键环节.

在本文的体系框架中,通过使用 RDF对企业知识进行描述,实现了知识在语义级别上的沟通. 基于 RDF建立企业知识链,使得领域内的各种相关知识在语义级别连接成有向网络形式,从而在广度和深度 2个方面增强知识共享和重用的能力,提高知识创新的效率.

2.4　知识查询

知识查询是实现知识共享和重用的关键技术. 根据本文提出的知识集成体系框架,知识查询实质上是对使用 RDF表示的本体库的语义查询. 本体库通常分为企业本体库和应用本体库:企业本体库提供了企业内高层次的公共概念集合,是各业务系统的共享词汇表;应用本体库提供了每个业务系统自己的本体,它们不一定使用相同的概念,但可以通过本体映射和代理的方法,实现互操作. 知识查询过程,可以分解为 3步:企业本体

的查询、企业本体到应用本体的查询转化,以及应用本体的 XML查询.如图 4所示,具体实现为:针对问题建立 RDF知识查询 (RDQL),执行查询获取结果;根据结果再将查询转化为对 XML数据的查询;执行 XML查询 (XQuery),获取数据集合.

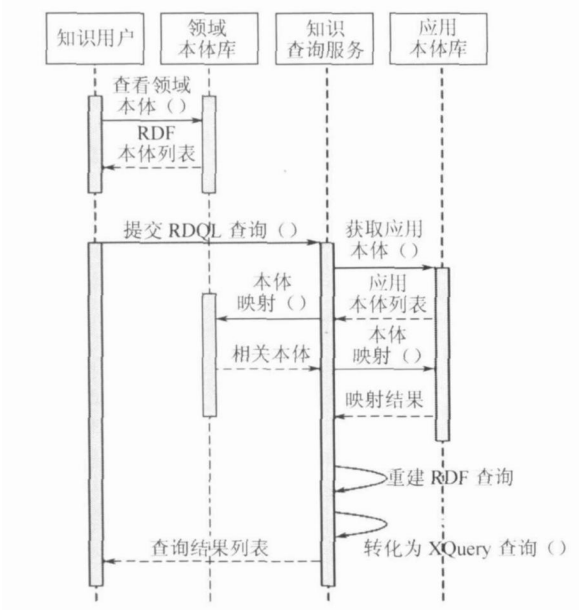


图 4 知识查询执行过程

RDQL语言是 W3C组织推荐的 RDF本体查询语言,具有类似 SQL的语法.其中 SELECT子句用于识别查询中所需返回的变量; FROM子句用于指定 RDF模型所使用的 URI; WHERE子句用于指定 RDF模型中的三元组; AND子句用于指定布尔表达式; USING子句提供了 URI的缩写方式. XQuery是 W3C组织推荐的 XML查询语言,它基于 XPath,同样具有类似 SQL的语法. XQuery提供了一种统一的方法来搜索、提取或者以别的方式来处理 XML数据. RDQL和 XML均具有类似 SQL的语法,在它们之间可以较方便地进行查询转换.

3 企业知识集成系统开发方法

可以参照下述的 8个通用步骤,进行本文提出的企业知识集成系统的开发工作.在实际工作中,可以根据自身的需要对此开发方法进行扩展.例如,对于支持产品设计和制造的企业知识集成系统,建立基于实例推理 (CBR, Case Based Reasoning)的应用就是重要的一项内容.

1) 构造本体.本体构造分为 3个主要阶段:①领域本体构建,由领域专家主导,全面描述领域中的概念及概念相互间的关系;②企业本体构建,由企业知识工作者主导,着重对企业中高层次的知识概念及概念间的关系进行描述;③应用本体构建,由系统开发人员主导,描述具体业务中的概念和它们之间的关系.

2) 定义本体语义映射关系.在构造本体的过程中,使用映射表记录应用知识到企业知识,再到领域知识之间的关系.

3) 使用 XML作为数据源的存储方式.定义业务系统中涉及到的各种数据的 XML Schema,将关系数据库中的数据按照 XML Schema的定义转换为 XML形式.

4) 定义 XML与 RDF之间的转换规则.此步工作完成后,建立了从底层的数据源到应用本体到企业本体最终到领域本体之间的联系,这些联系构成了查询转换和执行的基础.

5) 知识获取和表示.将 XML数据进行语义提升,使用 RDF进行统一表示.

6) 构建知识链,形成知识库.通过使用 RDF三元组,以有向图的形式,将各个层次的知识组成知识链,形成知识库.

7) 知识查询.这里的知识查询包括主动查询和自动查询.主动查询按照使用者提交的查询语句返回结果;自动查询又可以称作知识推荐,使用智能推理技术,自动将有用的知识推荐给使用者.

8) 在知识查询的集成上,建立高级的知识服务和应用.包括知识搜索引擎、知识地图等.

4 结束语

企业的知识往往是离散的、无序的,分散在各个业务系统之中的,通过知识集成,将原先分散的知识有机结合起来,使得知识使用者能够更好地实现知识共享与重用.然而,企业知识集成是一项复杂的系统工程,还有许多问题有待进一步研究,如使用功能更加强大的 OWL进行知识表示,语义查询语言,知识的相似性与相关性定量计算方法、知识推理技术等.

参考文献 (References)

[1] Demsetz H. The theory of the firm revisited [C]//Williamson O, Winter S. The Nature of the Firm. Oxford: Oxford University Press, 1991

[2] Grant R. Prospering in dynamically-competitive environment: organizational capability as knowledge integration [J]. Organization Science, 1996, 7(4): 375 - 387

[3] 顾新建, 祁国宁. 知识集成初探 [J]. 计算机集成制造系统, 2000, 6(1): 8 - 13
Gu Xinjian, Qi Guoning. Research on knowledge integration [J]. Computer Integrated Manufacturing Systems, 2000, 6(1): 8 - 13 (in Chinese)

[4] 王娟茹,赵嵩正,杨瑾. 知识集成模式研究[J]. 工业工程,2004,7(6):26 - 29
Wang Juanru, Zhao Songzheng, Yang Jin. Models of knowledge integration[J]. Industrial Engineering Journal, 2004,7(6): 26 - 29(in Chinese)

[5] Maryam A, Dorothy E L. Review: knowledge management and knowledge management systems: conceptual foundations and research issues[J]. Management Information Systems Quarterly, 2001, 25(1): 107 - 136

[6] Fernandes A, de C Moura A M, Porto F. Database and expert systems applications[C]//Proceedings. 14th International Workshop on 1 - 5 Digital Object Identifier 10. 1109/DEXA, 2003: 604 - 609

[7] Liao Shuhsien. Knowledge management technologies and applications-iterature review from 1995 to 2002[J]. Expert Systems with Applications, 2003(5): 155 - 164

[8] Lee T B, Hendler J, Lassila O. The semantic web[J]. Scientific American, 2001, 284(5): 34 - 43

[9] 周武,金远平. XML数据语义映射方法研究[J]. 计算机工程与应用,2003(28): 88 - 91
Zhou Wu, Jin Yuanping. A study of approaches to XML data semantic mapping[J]. Computer Engineering and Application, 2003(28): 88 - 91(in Chinese)

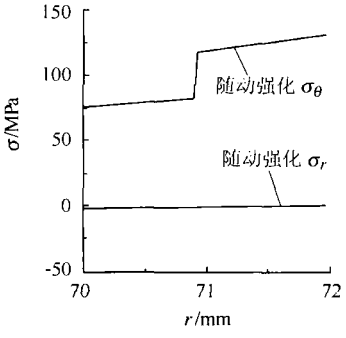
[10] Gruber T. Towards principles for the design of ontologies used for knowledge sharing[J]. International Journal of Human-Computer Studies, 1995, 43(5/6): 907 - 928

[11] Pérez A G, Benjamins V R. Overview of knowledge sharing and reuse components: ontologies and problem-solving methods[C]//IJCAI-99 Workshop on Ontologies and Problem-Solving Methods(KRR5). Stockholm, Sweden: [s. n.], 1999: 1 - 15

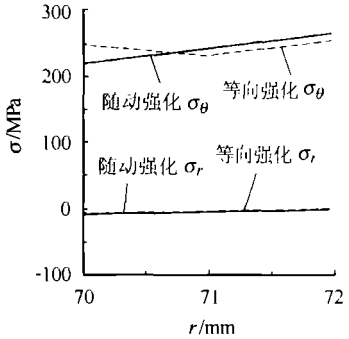
[12] 王向华,何坚,覃征,等. 面向Web的知识描述语言[J]. 北京航空航天大学学报,2004,30(7): 674 - 678
Wang Xianghua, He Jian, Qin Zheng, et al. Web-oriented knowledge description language[J]. Journal of Beijing University of Aeronautics and Astronautics, 2004, 30(7): 674 - 678(in Chinese)

(责任编辑:赵海容)

(上接第 209页)



材料 LY12M, P/P_s = 1
随动强化 r_i' = 72.75mm, α' = 57.66°
a 反向加载区处于弹性变形状态



材料 08, P/P_s = 1.219
随动强化 r_i' = 70.55mm, α' = 59.46°
等向强化 r_i' = 70.1mm, α' = 59.85°
b 反向加载区处于塑性变形状态

图 3 先弯后拉的应力分布曲线

4 结 论

1) 对于先拉后弯,包氏效应是降低回弹的有利因素,但由于预拉伸量较小,或反向加载区的面积很小,包氏效应的影响不大;

2) 对于先弯后拉,包氏效应对减小回弹不利,但由于应变中性层曲率半径与内弯矩的变化相互制约,包氏效应的实际作用十分有限;

3) 按线性随动强化和等向强化条件计算的拉弯回弹,差别较小. 在工程应用上,依据等向强化本构方程计算的回弹,对于先拉后弯,数值偏大;对于先弯后拉,数值偏小.

参考文献 (References)

[1] Gau J T, Kinzel G L. An experimental investigation of the influence of the Bauschinger effect on springback predictions[J]. Journal of Materials Processing Technology, 2001, 108: 369 - 375

[2] 王仁,熊祝华,黄文彬. 塑性力学基础[M]. 北京:科学出版社,1982
Wang Ren, Xiong Zhuhua, Huang Wenbin. Foundations of plasticity[M]. Beijing: Science Press, 1982(in Chinese)

[3] 陈笃. 塑性力学概要[M]. 北京:高等教育出版社,2004
Chen Du. Essentials of plasticity mechanics[M]. Beijing: Higher Education Press, 2004(in Chinese)

[4] Gau J T, Kinzel G L. A new model for springback prediction in which the Bauschinger effect is considered[J]. International Journal of Mechanical Siences, 2001, 43: 1813 - 1832

(责任编辑:文丽芳)