

基于 Ontology 的工作流知识管理系统的应用研究

王浩, 武凌, 张海, 徐勇

(安徽财经大学 信息工程学院, 安徽 蚌埠 233041)

摘要: 知识管理已成为企业在信息化条件下提高竞争力的必要手段, 工作流系统是支持工作流定义、执行和监控的计算机软件系统。针对工作流系统应用在企业知识管理时面临的流程数据整合不易、文件内容词汇关系的不明确与存取控制不便等问题, 提出了一种基于本体的工作流知识管理系统架构, 并介绍了其设计及实现机制。该系统以 XML 作为流程文件交换与存储格式, 流程中的知识可以结构化地保存下来, 便于在各部门的异质系统间传递; 并利用了本体论的方法将企业内的组织、计划、人员等不同知识领域所包含的概念及关联特性整合进工作流程中, 改善了部门间的语义差异, 更好地实现了工作流中知识的应用效率与一致性。

关键词: 知识管理; 工作流; 本体; 扩展标记语言

中图分类号: TP393.07; TP311.52

文献标识码: A

文章编号: 1673-629X(2008)06-0023-05

Research and Application of Workflow Knowledge Management System Based on Ontology

WANG Hao, WU Ling, ZHANG Hai, XU Yong

(College of Information Engineering, Anhui University of Finance & Economics, Bengbu 233041, China)

Abstract: Knowledge management is necessary for enterprise's competition. Workflow system is computer software, which supports workflow definition, performance and surveillance. When workflow system is applying in knowledge management, it faces the drawbacks including uneasy data integration, implicit definition of vocabulary, and inflexible control of authentication access. Puts forward a workflow knowledge management system architecture based on ontology, and introduces the designs and implementations. The system uses XML as the format of file exchange and storage, so it can store knowledge structurally and transfers it in heterogeneous system. The system also integrates different domain knowledge's concept and relation into the workflow with the help of ontology. It enhances the semantics and realizes the knowledge applying efficiency and consistency.

Key words: knowledge management; workflow; ontology; XML

0 引言

随着网络及电子数据交换技术的发展, 为了提升企业电子化营运的效率及竞争优势, 企业必须有系统地整合个人与组织的经验, 以知识管理作为企业电子化的核心, 让企业在电子化运作的同时能积累知识、运用知识、并更进一步地创造新知识。根据工作流管理联盟^[1](WfMC, Workflow Management Coalition)的定义, 工作流管理系统是将企业流程的一部分或全部予

以自动化, 其文件、信息或任务会按照预设的程序规则, 传递给参与者(人、组织), 以协同完成工作。

事实上, 工作流管理系统在企业的运作中提供了很多与知识相关的活动, 如知识的交流、知识的储存与重复利用和知识的传承等, 越来越多的研究者提出了将知识管理与流程管理相结合的思想^[2~6], 因此工作流对企业来说是相当重要的知识活动, 更是企业竞争优势的来源, 它增加了知识应用的效率与提供创新的一致性。如果把工作流作为企业知识管理的着手点, 将信息与信息、信息与活动及信息与人连结起来, 能够更好地实现企业的知识管理。

1 基于本体的工作流知识管理系统架构

系统架构如图1所示, 主要分为三层, 分别为用户接口(User Interface)、工作流引擎(Workflow Engine)与数据库(Database)。以XML文件作为系统数据传

收稿日期: 2007-09-19

基金项目: 中华全国供销合作总社重点项目(GXZSKY 06013zd); 安徽省自然科学基金项目(KJ2007B246); 安徽省高等学校青年教师科研资助计划项目(2007jq1083); 安徽财经大学信息工程学院青年项目(xgky2008001)

作者简介: 王浩(1982-), 男, 安徽蚌埠人, 硕士研究生, 研究方向为计算机网络与软件、实验室管理; 张海, 高工, 研究方向为计算机网络与信息系统。

递的格式,工作流程引擎会将 XML 流程文件按照程序路径传递给每一位参与者,并处理文件的内容,XML 格式不受机器平台的限制且具有结构化的特性。

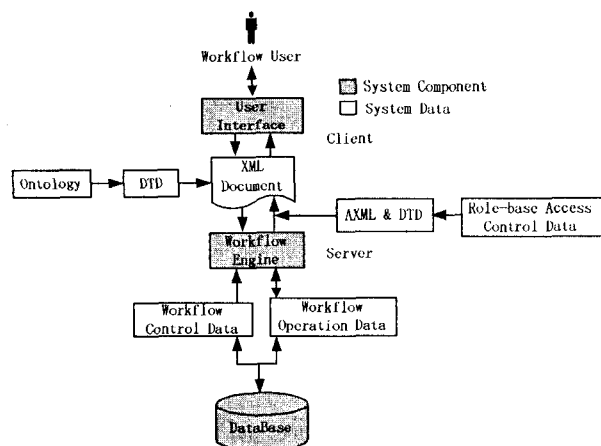


图 1 系统架构

(1) 用户接口为系统的用户端,提供用户输入数据的接口,将数据传递给 workflow 引擎,接收 workflow 引擎发布给用户的信息,并提供用户 workflow 记录搜索的功能,让用户适时地获得工作所需的信息。用户接口会将 workflow 参与者输入的信息或做出的选择通过 DTD (Document Type Definition) 的验证后建立 XML 文件,并传递给 workflow 引擎。此 DTD 是以本体论的概念建立的,不仅能验证用户输入信息的正确性,也能改善词汇差异性,让数据在存储时更为正确,同时也能让用户在搜索 workflow 中的信息时更为快速与正确。

(2) workflow 引擎为系统的服务器端,控制各流程步骤的执行,包括工作的状态(开始、结束、放弃、通知或回复)与路径的规则(启动流程的下一个步骤之前的判断),接收与传递用户工作上的要求,包括传给用户工作清单、工作详细数据或接收处理用户回传的信息。workflow 引擎将根据授权文件(AXML, Authorization XML)中的 DTD,产生一份修改后的 XML 文件给参与者,这样不仅能有效控制文件中细节信息的存取,也能方便管理因 workflow 不断运作产生的大量文件。

(3) 数据库中存放两种类型的数据,为 workflow 控制数据(Workflow Control Data)与 workflow 作业数据(Workflow Operation Data)。workflow 控制数据是指工作传递的路径、规则与工作的状态;workflow 作业数据指的是参与者在进行工作时用到的信息,包括工作上的相关技术数据与目前工作进行的运作资料。这些数据可以在 workflow 自动化下不断地被取用、修改与累积,进而成为企业永久的经验与知识基础。

2 系统实现机制与设计

系统采用 B/S 模式开发,用户端只要有支持 XML

的浏览器就可以使用系统中的各项功能。系统采用 JSP 与 JavaBeans 设计 workflow 引擎,将引擎的各部分重要功能以 JavaBeans 组件的方式开发,使得 workflow 系统的执行组件不仅能有效地重复使用,也能方便地扩充功能或与其它系统服务整合。系统以 XML 作为基础,不仅便于引擎组件之间的内部沟通,同时配合 XSL 与 JSP 将结果通过浏览器提供给用户。

2.1 workflow 引擎的设计

workflow 引擎作为系统的核心运行在服务器端,用于响应处理用户送出的 workflow 相关信息,并控制整个流程系统的运作。引擎主要由三个处理器组成,分别为 workflow 清单处理器(Workflow List Handler)、workflow 程序处理器(Workflow Procedure Handler)与授权处理器(Authorization Handler)。

workflow 清单处理器根据用户的帐户与密码,去搜索各流程中有哪些程序项目正等待该用户去处理,并以清单的格式将程序项目送回用户端。用户能从工作清单中选取其中一个工作项目进行处理,并在处理完之后将处理结果送至 workflow 程序处理器。工作项目信息中包含了 XML 流程文件的链接,当使用此文件时,系统会驱动授权处理器,以用户的授权条件对该文件的内容做适当的修正。

workflow 程序处理器提供了五种处理 workflow 的程序(Response、Back Response、FYI、And Split 与 And Join),每种程序以不同的方法处理不同形式的路径。在接收到用户的响应之后会先判断此程序为何种类型,并以该类型的方法去处理程序项目,这些方法在处理程序时会先判断该程序是否已被处理完成,若是未处理完成则不会存入或改变 workflow 数据库中的数据,若处理完成则将该程序项目的处理结果存入数据库中,并结束该程序的执行。在处理完使用者响应的程序项目后,会判断下一程序的执行项目,并将欲启动的下一程序项目初始化,把执行权转至下一程序项目。

授权处理器会根据用户的角色,修正授权文件的内容,并再根据授权文件对 XML 流程文件进行存取权限的控制,最后用户接收到的 XML 文件,就是属于该用户的角色所被赋予的存取内容。

2.2 知识定义与运作

本系统强调两种知识的运作,分别是作业知识与组织知识,这两种知识都以 XML 的格式表示。作业知识是指 workflow 的自动化运作过程,包含工作传递的路径、规则与信息。路径与规则会被嵌入系统中,而信息则会因为工作上的需要被不断地传递给下一位参与者分享运用或修正,直至工作的结束;组织知识是指企业内相关的技术文件、研讨资料、规章等,另外当一流

程实例结束时,其包含的工作处理事项与事件原因等历史信息也将成为组织知识的一种。本系统将作业相关的组织知识整合进来,是因为当用户在进行工作时常常需要取得一些额外的信息来协助问题的解决,因此系统提供了一个完整的知识再利用机制,让需要获得该信息的用户能更方便地取得。信息再利用的过程可分为搜索与获得两个部分,在搜索部分,本系统的词汇由于应用了本体论的概念,在搜索时能不受 XML 文件结构上的限制,且根据词汇间的推论与概念定义能得到更正确的信息。在获得部分,系统将角色授权的概念应用在 DTD 中,能有效地维护数据的读取安全性,通过 DTD 规定的存取限制修改数据的内容,再回传给搜索用户。

2.3 基于本体论的 DTD

在设计 XML 流程文件时,如果其应用领域的领域知识无法有效地被整合进 XML 文件中,数据在传递与整合时会产生不一致的情况。另外,虽然 XML 提供用户根据自己的需求定义标签语义及文件结构,但也会造成在搜索 XML 文件数据时,由于文件结构的复杂性与词汇的不一致而带来的搜索难度,这样的情况也同样发生在工作流程中。

2.3.1 本体论的知识领域表达

为了对 workflows 系统中的作业知识与组织知识进行管理,采用了本体论描述领域知识,并将本体论的描述转换成流程 XML 文件的 DTD,为 workflow 系统的知识传递、整合与搜索提供良好的基础。本体论提供了领域知识的完整描述,利用这些领域知识能得到较为完善的工作流系统的文件,并且文件中的标签词汇也更为适当并保持一致性。经过本体论取得的流程 XML 文件,事实上就是被概念化且分类的文件,在对其内容作搜寻时将更为快速,而且查询的语法将不受限于文件的结构,甚至可以通过标签与标签中词汇关系的定义,推论出更多有用信息。以工程设计蓝图审查与检验流程为例,定义了工程图处理和员工两个知识领域,如图 2 所示。

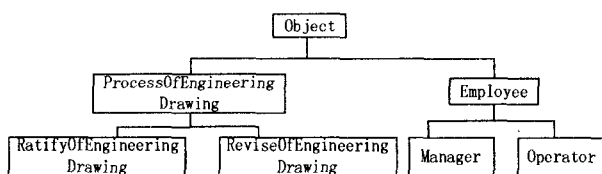


图 2 概念的继承关系

图 3 所示为采用了 Frame-Logic 表示法的本体论,作为 workflow 系统建构知识领域文件的基础。Frame-Logic 是面向对象的逻辑学的表达方式,通过类别与属性集合,将本体论的概念、概念与概念的关系、概念

的特性与推论知识做出完整表达。

Concept hierarchy

Object[]

ProcessOfEngineeringDrawing :: Object

ReviseOfEngineeringDrawing :: ProcessOfEngineeringDrawing

RatifyOfEngineeringDrawing :: ProcessOfEngineeringDrawing

Employee :: Object

Operator :: Employee

Manager :: Employee

Inherited Properties

ProcessOfEngineeringDrawing[drawingName = >>STRING;

drawingNo = >>STRING; drawingPath = >>STRING;

date = >>STRING; material = >>STRING; hardness = >>STRING;

heatTreatment = >>STRING; weight = >>STRING; tensileStrength = >>STRING; appraise = >>STRING; dealWithPerson = >>Employ].

ReviseOfEngineeringDrawing [reviseDealWithPerson = >>Operator].

RatifyOfEngineeringDrawing [ratifyDealWithPerson = >>Manager].

Employee[name = >>STRING; title = >>String; section = >>STRING;

job = >>ProcessOfEngineeringDrawing].

Operator [job = >>ReviseOfEngineeringDrawing].

Manager [job = >>RatifyOfEngineeringDrawing].

Axioms

FORALL Pro1, Emp1

Pro1: ProcessOfEngineeringDrawing [dealWithPerson - >>Emp1] <- >

Emp1: Employ[job - >>Pro1].

FORALL Rev1, Ope1

Rev1: ReviseOfEngineeringDrawing [reviseDealWithPerson - >>Ope1] <- >

Ope1: Operator[job - >>Rev1].

FORALL Rev1, Rat1

Rev1: RatifyOfEngineeringDrawing[ratifyDealWithPerson - >>Ope1] <- >

Ope1: Operator[job - >>Rev1].

图 3 采用 Frame-Logic 表示法的工程

图处理和员工本体论

2.3.2 由本体论取得 DTD

文献[7]利用本体论提出了一套建构与使用 XML 文件的方法,主要思想就是将本体论转换成文件的验证机制 DTD,再根据 DTD 验证 XML 文件的结构与词汇。对上述文献提出的方法加以修正,将本体论的概念应用到工作流程的文件验证上。

为了将本体论的阶层概念定义转换成 DTD 的格式,系统采用参数实体(Parameter Entity)的方法表示

概念与子概念间的继承关系。参数实体主要目的是为了简化 DTD 的语法,其语法如下所示:

```
<! ENTITY % entityName entityValue>
<! ELEMENT elementName ( %entityName; )>
```

在定义实体名称时须先加上 % 符号,并定义实体值,实体值的数量为一个至多个,每个实体值为一个元素,因此能将参数实体的定义看成是一个元素集合的定义。当在定义 DTD 的某元素,需要用到大量且相同的元素集合时,就能通过调用参数实体取用这些元素集合,调用的方法是以 % 为开头,以 ; 为结尾,中间则是被调用的实体名称。

通过参数实体的方法能表达本体论中的概念继承关系和概念的属性。对于本体论的属性值将它分为两类:一种为一般的属性值,如 STRING 或 NUM,这类属性值较为简单,可通过 #PCDATA 的方法转换;另一类的属性值较为特别,其属性值为概念,这时可以通过子元素的方法定义。图 4 是图 3 中的本体论转化而成的 DTD。

概念的继承定义

```
<! ELEMENT ProcessOfEngineeringDrawing ( % ProcessOfEngineeringDrawing; )>
<! ELEMENT ReviseOfEngineeringDrawing ( % ReviseOfEngineeringDrawing; , % ProcessOfEngineeringDrawing; )>
<! ELEMENT Ratify Of Engineering Drawing ( % RatifyOfEngineeringDrawing; , % ProcessOfEngineeringDrawing; )>
<! ELEMENT Employee ( %Employee; )>
<! ELEMENT Operator ( %Operator; , %Employee; )>
<! ELEMENT Manager ( %Manager; , %Employee; )>
```

概念的属性定义

```
<! ENTITY % ProcessOf EngineeringDrawing ( drawingName | drawingNo | drawingPath | date | material | hardness | heatTreatment | weight | tensileStrength | appraise | dealWithPerson ) * >
<! ENTITY % ReviseOfEngineeringDrawing ( reviseDealWithPerson ) * >
<! ENTITY % RatifyOfEngineeringDrawing ( ratifyDealWithPerson ) * >
<! ENTITY % Employee ( name | title | section | workName ) * >
```

```
<! ENTITY % Operator ( job ) * >
```

```
<! ENTITY % Manager ( job ) * >
```

概念的属性

```
<! ELEMENT drawingName ( #PCDATA )>
<! ELEMENT drawingNo ( #PCDATA )>
<! ELEMENT drawingPath ( #PCDATA )>
<! ELEMENT date ( #PCDATA )>
<! ELEMENT material ( #PCDATA )>
<! ELEMENT hardness ( #PCDATA )>
```

```
<! ELEMENT heatTreatment ( #PCDATA )>
<! ELEMENT weight ( Event )>
<! ELEMENT tensileStrength ( #PCDATA )>
<! ELEMENT appraise ( #PCDATA )>
<! ELEMENT dealWithPerson ( Employ )>
<! ELEMENT reviseDealWithPerson ( Operator )>
<! ELEMENT ratifyDealWithPerson ( Manager )>
<! ELEMENT name ( #PCDATA )>
<! ELEMENT title ( #PCDATA )>
<! ELEMENT workName ( #PCDATA )>
<! ELEMENT job ( ReviseOfEngineeringDrawing | RatifyOfEngineeringDrawing )>
```

图 4 工程图处理和员工本体论转换后的 DTD

2.4 流程 XML 文件的设计

图 5 是工程设计蓝图 XML 文件,其标签必须遵循上述 DTD 的定义。工程设计蓝图文件随着流程中的各程序提供给用户读取,同时根据用户角色的权限修正文件的内容,让文件内容的保护更为完善。此文件中包含的信息有工程蓝图设计信息,如图名(drawingName)与图号(drawingNo)等,及处理者的相关信息如姓名(name)、称谓(title)等。这些数据能完整表达工程图在流程处理时被存取审核的信息,对文件搜索时能有效地找出工程图处理的相关使用者或工程图设计等相关信息。

```
<? xml version="1.0" ? >
<ProcessOfEngineeringDrawing>
  <drawingName>motor bracket</drawingName>
  <drawingNo>No. 66</drawingNo>
  <drawingPath>c: \ </drawingPath>
  <date>2007-5-29</date>
  <material>steel</material>
  <hardness>135(BHN)</hardness>
  <heatTreatment>quench</heatTreatment>
  <weight>23(kg)</weight>
  <tensileStrength>99(1000psi)</tensileStrength>
  <appraise>750(NT)</appraise>
  <quantity>2</quantity>
  <dealWithPerson>
    <name>justin</name>
    <title>designer</title>
    <section>designing section</section>
    <workName>drafting</workName>
  </dealWithPerson>
  <dealWithPerson>
    <name>jimmy</name>
    <title>examiner</title>
    <section>designing section</section>
    <workName>examine</workName>
```

```

</dealWithPerson>
.....
</ProcessOfEngineeringDrawing>

```

图5 工程设计蓝图文件

3 结束语

针对工作流程中的知识管理问题,提出了一种基于本体论的工作流知识管理系统架构,将工作流程文件的格式与结构建立在可交换与可重复使用的基础上,同时改善了词汇差异的问题,整合各领域事物间的关系与特性至工作流程中,借助领域性知识的明确定义,使得数据在传递与储存的过程中能正确无误地被处理。文中采用的本体论并非自动化的产生,在以后的研究中,可以考虑设计本体论的建构处理器,利用该处理器对领域的概念作分类,自动产生本体论,并采用文中转换本体论的方法,自动产生合适的 DTD。还可以通过本体论的表示法建立搜索处理器,搜索文件中的信息给用户或应用程序加以处理。

参考文献:

- [1] Hollingsworth D. The Workflow Reference Model[EB/OL]. 2006. <http://www.wfmc.org/standards/docs/tc003v11>.

(上接第22页)

确性如图5所示。

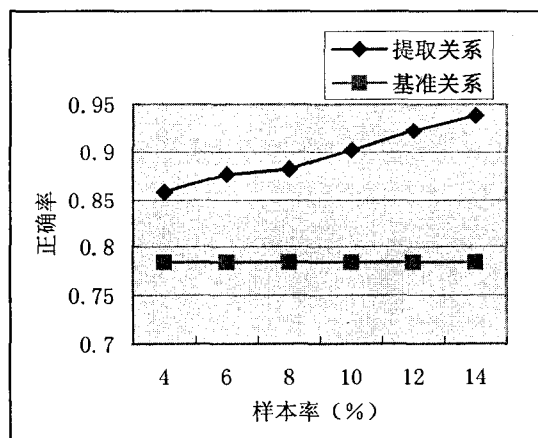


图5 基于提取关系的群体挖掘的正确性

实验证明,改进的算法可提取基于用户查询(标记数据)的最优关系。利用标记信息,可优化关系提取。在多关系网络中,用户信息需求是多样的。这使得关系提取在社会网络分析的预处理过程中越来越重要。

3 结束语

社会网络中存在着多样、异构的关系。对这些关系进行有效的合并可以产生重要的更能适合用户信息

pdf.

- [2] Andersson B, Bider I, Perjons E. Integration of Business Process Support with Knowledge Management - A Practical Perspective[C]// Practical Aspects of Knowledge Management - 5th International Conference. Vienna, Austria: [s. n.], 2004: 227-238.
- [3] Mou Yu-jie, Zhang Shen-sheng, Cao Jian. Providing knowledge support in business process: A context based approach [C]// Proceedings of IEEE International Conference on Systems, Man and Cybernetics. Hague, Netherlands: [s. n.], 2004: 2143-2149.
- [4] Chung P W H, Cheung L. Knowledge-based process management - An approach to handling adaptive workflow[J]. Knowledge-Based Systems, 2003, 16(3): 149-160.
- [5] Moore J, Inder R, Chung P, et al. Combining and Adapting Process Patterns for Flexible Workflow[C]// IEEE Database and Expert Systems Applications, Proceedings 11th International Workshop. London: [s. n.], 2000: 797-801.
- [6] 沈兵虎,王坚,潘瑞芳,等. 基于工作流技术的知识管理系统研究与设计[J]. 制造业自动化, 2007, 29(3): 23-27.
- [7] Erdmann M, Studer R. How to structure and access XML documents with ontologies[J]. Data & Knowledge Engineering, 2001, 36(3): 317-335.

需求的新关系。基于这样的想法,提出一种基于用户查询的新算法应用于关系提取。应用此算法进行关系提取和群体挖掘,可以获得精准的语义,提高群体挖掘效能。可以预见,这种依赖于查询的关系提取和群体挖掘在社会网络分析中将会引起许多潜在的新应用。

参考文献:

- [1] Wasserman S, Faust K. Social Network Analysis: Methods and Applications [M]. Cambridge, UK: Cambridge University Press, 1994.
- [2] Adamic L A, Adar E. Friends and neighbors on the web[J]. Social Networks, 2003, 25(3): 211-230.
- [3] Schwartz M F, Wood D C M. Discovering shared interests using graph analysis[J]. Communications of the ACM, 1993, 36: 78-89.
- [4] Domingos P, Richardson M. Mining the network value of customers[C]// In: Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining. [s. l.]: ACM Press, 2001: 57-66.
- [5] Björck A. Numerical Methods for Least Squares Problems [M]. Philadelphia, PA: SIAM, 1996.
- [6] Washio T, Motoda H. State of the art of graph-based data mining[J]. SIGKDD Explor. Newsl, 2003, 5(1): 59-68.