

面向过程的勘探开发知识管理系统

赵艳红 李洪奇 朱丽萍

(中国石油大学(北京)石油数据挖掘北京市重点实验室 北京 102249)

摘要 在石油勘探开发领域,井位部署是其中一个重要的子流程,直接决定着勘探的成败和效益。因此,提出基于本体面向过程的知识模型,基于该模型,设计和开发了一套基于 Web 的知识管理系统,提供了可视化的过程建模工具,实现了业务流程的控制以及经验知识的管理。该平台在胜利油田的应用,充分发挥了已有成果和经验知识的作用,实现了知识的叠加效应,对于新区块的勘探和老区块的二次调整开发具有重要的指导作用。

关键词 本体 过程建模 知识管理

中图分类号 TP3 文献标识码 A DOI:10.3969/j.issn.1000-386x.2016.05.020

A PROCESS-ORIENTED EXPLORATION AND DEVELOPMENT KNOWLEDGE MANAGEMENT SYSTEM

Zhao Yanhong Li Hongqi Zhu Liping

(Petroleum Engineering and Data Mining Lab, China University of Petroleum-Beijing, Beijing 102249, China)

Abstract In the field of oil exploration and development, well positions deployment is an important sub-process in geophysical exploration, which decides the final results and benefit of the exploration. Therefore, we proposed an ontology-based and process-oriented knowledge model, and designed and developed a set of web-based knowledge management systems based on the model. It provides a visualised process modelling tool and thus implements the business process control and the experience and knowledge management. The platform has been applied in Shengli oilfield in China, it brings the roles of existing outcomes and experience and knowledge into fullplay, and realises the superimposition effect of the knowledge. These play an important guidance role to the explorations in new regional blocks and the secondary adjusted developments in old regional blocks.

Keywords Ontology Process modelling Knowledge management

0 引言

知识是人们通过实践对客观事物及其运动过程和规律的认识,是被人们理解和认识并经过大脑思维重新组织和序列化的部分信息,是经验、技能的总结^[1]。在知识经济时代,知识是重要的社会经济资源,企业必须建立有效的知识管理模式,充分利用已有的知识资源,不断增强知识创新能力,才能在竞争中立于不败之地。然而,知识不是孤立存在的,只有与特定的业务流程紧密相连,才能有效地发挥作用。业务流程是企业中一系列创造价值的活动的组合,通过业务流程的管理可以优化企业经营过程,促进各部门协作,提高工作效率,降低业务成本,提高服务质量和用户满意度。

在石油勘探开发领域,包括勘探、测井、钻井、录井、试油、井下作业等二十多个专业,这些专业相互协同,构成石油勘探与开发的全部业务流程。井位部署是油气勘探的重要业务流程,直接决定着勘探的成败和效益。每个环节需要不同专业部门按照井位部署的业务流程相互协调完成,前面环节得出的结论和数据对后面环节的操作都有支撑作用。然而,面对严峻的勘探形式和日益加快的勘探节奏,传统的井位部署工作存在两个主要

问题。一是井位部署业务流程的管理问题。在执行井位部署业务流程的时候,不同的环节需要不同的专业部门来完成,但是各个环节之间没有协同管理平台,决策人员无法及时掌握各个环节的进展情况,各个环节的产生的成果资料无法在各部门之间传递和共享。二是经验知识的存储和共享问题。业务专家通过几十年的勘探开发实践过程,积累了大量的经验知识,这些隐性知识存储在专家意识中,缺乏一个有效的知识表征工具,实现对隐性知识存储、表示和管理。随着时间推移,人员变动,经验知识逐渐遗失,对于企业来说,无疑是巨大的损失。因此,建立集数据、知识、业务过程为一体的知识管理平台,实现业务流程的控制、经验知识的存储和管理,促进企业各个部门之间的沟通协作和知识共享,具有重要的研究意义。

为了实现业务过程与知识管理的结合,首先要对业务过程进行建模。目前,有很多的技术和软件产品支持业务过程建模^[2,3],例如 UML、BPML、Petri-nets、Rational Rose、JBPM 等。虽然这些工具让用户更好地了解业务过程,但是大部分工具仅支

收稿日期:2015-01-07。十二五国家重大专项子课题(2011ZX05020-007-007);测井储层数据挖掘技术研究项目(20112x05020-009)。赵艳红,博士,主研领域:软件工程,过程建模,知识管理,本体。李洪奇,教授。朱丽萍,副教授。

持数学和图形建模^[4],而对业务过程的表示缺乏语义层次的描述。 workflows 技术是业务过程建模、重组和执行的重要技术,但是 workflows 技术的主体是业务过程,它在实现对显式的业务过程知识进行管理(通过 workflow 过程定义)的同时并不支持对普遍意义上的知识进行灵活管理^[5]。在国内外,关于知识管理、本体、业务过程的结合有一些研究^[6-8],为实现基于过程的知识管理提供了一些可借鉴的理论和方法,但是仍存在一些问题:(1) 采用本体描述元信息,但是大部分本体都是手动构建的;(2) 采用 UML、OWL-S 等描述业务过程,但是这些过程是预先定义好的,用户无法个性化地对业务流程进行组合和配置;(3) 部分研究提到了 KM 与 BPM 的集成,但是只是提出了一些理论性的研究方法。

本文提出基于本体面向过程的方法,解决了石油勘探开发领域业务流程的控制和经验知识的管理问题。首先,本文第一次基于 POSC 的 Epicentre 本体模型建立了勘探开发领域井位部署的本体模型。POSC 发布的 Epicentre 本体模型定义了勘探开发领域所有的对象以及对象之间的关系,是一个全球性的石油工业标准。其次本文建立了基于本体的过程模型。在业务建模过程中,采用本体描述业务流程中各个活动所需要的对象,例如活动的输入输出对象。基于 Epicentre 建立的本体为该领域不同的应用程序之间提供了共同的公共数据解释,解决了业务流程执行过程中数据交换和数据共享的问题。同时,业务过程各个环节所需要的知识、产生的知识都采用本体进行表示。通过本体将不同业务过程之间的知识联系起来,形成了一个过程知识网络。在业务过程执行的时候,可以主动推送与该业务相关的知识。最后,本文实现了面向过程的知识管理系统,并将其应用于勘探开发领域的井位部署过程。该平台主要包括本体管理、过程管理和知识检索三大模块。本体管理模块主要用于维护基于 Epicentre 构建的本体。过程管理模块提供了一个可视化的知识表征工具,不仅可以对专家的经验知识进行存储和管理,而且可以对业务过程进行个性化的配置和组合。知识检索模块向用户提供本体知识和过程知识。通过知识管理平台在井位部署业务中的应用,实现了勘探开发已有工作成果的回溯和知识重用,对于指导新一轮的油气勘探开发发挥了重要的作用。

1 研究方法

1.1 本体构建

本体是描述概念以及概念之间关系的概念模型,是知识表示的一种重要方式。构建本体可以形式化地描述领域知识结构,为知识获取和表示奠定了基础,避免重复的领域知识分析;统一的术语和概念使得知识共享成为可能。目前,不同领域信息化程度不同,有些领域的信息化程度较高,不但有大量的科学数据资源还具有领域标准规范。如基因本体 GO (Gene Ontology)、开放生物学本体 OBO (Open Biological Ontologies) 等。有些领域还不具有领域内通用的本体,需要进行人工构建。而通过人工方法构建本体是一项费时费力的过程,不仅需要系统化、工程化的方法论提供支持,而且需要领域专家的参与,开发效率非常低,质量更难保证。因此,国内外许多研究者正在积极探索支持本体(半)自动化构建的学习方法。利用知识获取技术从已有的数据源中获取期望的本体备受关注,数据源可是关系数据库、Web 页面、机器可读的词典、纯文本等。企业或者互联网中的大部分数据存储于关系数据库中,关系数据库到本体的映

射采用的主要技术有:逆向工程、模式映射和数据挖掘。通过关系数据库到本体的映射可以快速地构建领域本体^[9,10],国内外关于此项技术的研究也取得了很大的进展,但是也存在一些问题。本体的构建依赖于数据模型,数据模型的规范性和标准化程度直接影响着本体的质量。大部分关系数据库是基于不同的业务需求而建立的,表的重复、属性的重复、元组的重复、属性值的重复等数据冗余现象是不可避免的,通过映射技术建立的本体是粗糙的,需要领域专家不断地进行调整和维护,增加了本体构建的成本。因此,一个标准化、规范化、能够被领域认可的数据模型是构建领域本体的关键。

POSC 是 BP Exploration 等五大石油公司为了解决不同公司之间信息共享和各专业的协同工作而建立的石油技术开放联盟,其技术目标是为石油勘探开发应用软件提供一套标准规范。为了解决应用程序之间协同工作的问题,给所有的应用程序提供共同的公共数据解释,POSC 发布了 Epicentre 数据模型。Epicentre 采用面向对象的技术定义了油田整个生命周期中各阶段涉及的活动、对象及特性,并采用国际标准信息描述语言 Express 对模型进行定义,是一个结构化的文本文件。Epicentre 将勘探开发数据中的一些共同特性抽象地表示出来,并将他们归为 activity、property、object_of_interest、association、geological_process、ref_data 等。图 1 描述了 Epicentre 顶级本体的继承关系。



图 1 POSC Epicentre 顶级本体

随着本体技术的发展和应用,POSC 建立了 Express 与 OWL DL 1.0 之间的映射规则,提供了 OWL 格式的数据模型规范。Epicentre 数据模型的 OWL 文件可以导入本体编辑工具 protégé 中,并通过内置组件 OntoGraf 实现本体可视化。Epicentre 是一个全球性的石油工业数据模型,基于 Epicentre 数据模型建立的勘探开发领域本体是标准的、规范的、被领域认可的。本文基于已有的 Epicentre 本体,从中抽取勘探开发领域的核心概念集、属性集以及概念之间的关系,然后交由领域专家进行确认、修改和完善。勘探开发领域本体的构建确认了对该领域知识的共同理解,为实现知识的共享和重用奠定了基础。

1.2 基于本体的过程建模

知识和数据不是孤立存在的,伴随着业务活动的执行而产生或者再造,数据、知识应该和业务过程相联系。在本文中,将本体应用于过程模型,为过程模型的活动结点提供标准的规范化的输入对象和输出对象。基于本体的过程模型框架如图 2 所示,实现了业务流、数据流和知识流的融合。

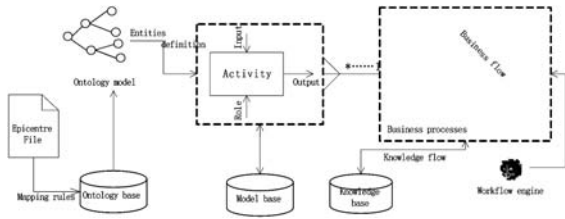


图 2 基于本体的过程模型

业务流程表示为完成特定的目标而按序执行的一系列活动的集合。业务流程可以不断进行分解细化,分解出的流程称之为二级流程、三级流程等。业务流程中不能再分的最小单元,称之为活动节点。活动节点包括输入、输出、角色和操作集合。在活动节点执行过程中,需要从各种专业数据库中获取活动节点所需的业务数据。由于数据源的异构性,数据模型的不一致,导致数据很难统一地映射到活动节点中。为了实现活动节点输入输出的通用性和普适性,必须建立一个标准的规范化的逻辑模型,实现数据源由多样性到统一性的转换。本体作为一个领域认可的规范化的概念模型,独立于任何应用程序和物理存储实现,可以将不同领域不同专业的数据集成到一起,解决了异构数据源之间的互操作性。过程模型表示的是活动与活动之间的关系、活动与对象之间的关系,本体模型表述的是对象之间的关系,将本体应用与过程模型,实现了活动节点输入输出数据的一致性和规范性。本体不依赖于过程模型,过程模型的修改不影响领域本体的稳定。采用 BNF 给出流程和活动节点的形式化概括:

```

business process ::= < activity > { , < activity > } | < business process >
< activity > ::= < input > < output > < role > < processing logic >
< input > ::= { < data entity > } | { < document entity > }
< output > ::= { < data entity > } | { < document entity > }
< data entity > ::= < ontology >
< document entity > ::= < ontology >
< ontology > ::= < concept - name > < attribute - list > < relations > < individuals >
< processing logic > ::= < service > { , < service > }
< service > ::= < service - name > < operation > { , < operation > }
< operation > ::= < operation - name > [ ( < input > , < output > ) ]
  
```

活动节点的输入输出对象主要由两种:数据对象和文档对象。活动的输入输出对象采用本体进行表示,这样保证了同一个活动不同实例的输入输出在逻辑模型上的一致性。定义活动节点输入输出对象的时候,直接从本体库中选择需要的本体。当业务流程执行,活动节点被实例化的时候,通过本体和关系数据库的映射,数据转换为本体实例,为活动节点提供所需的数据。本体是与活动相关的,同一个本体可以被不同的活动使用,为区分同一个本体在不同活动中的实例,在本体定义中添加了一个属性字段活动 id,记录了该本体实例所关联的活动。

活动节点的操作集合是一系列服务的集合。在本文中,采用基于 SOA 架构的将领域内一些固化的算法(例如数据挖掘算法等)、图版或者经验图版封装成服务。通过服务名和操作名用户可以直接调用该服务完成特定的功能。

角色指明了该活动执行者的类别,例如分析家、评价家或者分析家等。当业务流程实例化的时候,角色被指派给具体的业务人员,同时该活动会出现在指派用户的代办任务列表中。当活动执行完毕的时候,用户将输出结果上传到系统中,并发送消息通知流程引擎,控制业务流程继续运行。角色实现了不同部门之间的相互协作,使得数据流和知识流在不同的活动节点中进行传递和共享。

2 系统实现

基于如上所提的知识模型,开发了 POKMS(Process-oriented Ontology-based knowledge management system)知识管理系统,并将其应用于勘探开发的井位部署领域。系统采用 J2EE 架构,并使用开源框架 spring 和 hibernate 实现事务管理。开发环境细节如下所示:

- Web Server: Apache Tomcat 6.0
- Programming Language: Java, Flex
- Development Platform: Eclipse 3.5, Adobe Flash builder 4
- RPC Service: BlazeDS
- Database: Oracle 10g
- Library: RDF, Jena
- Graph Visualization Software: Prefuse, BirdEye

业务流程是系统的核心,主要包括活动、参数、端口、执行单元等。本文采用面向对象的方法对流程进行建模和表示,如图 3 所示。

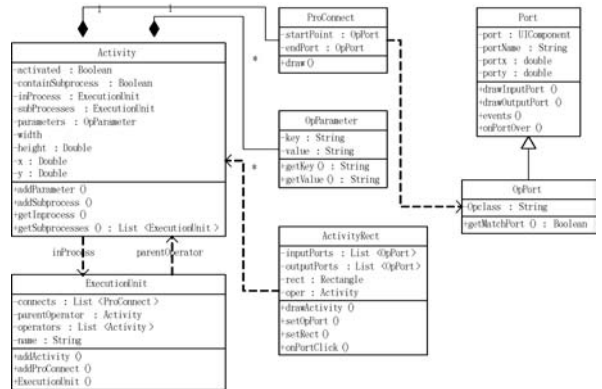


图 3 业务过程核心类图

POKMS 主要包括本体管理、流程管理和知识管理三部分,还有一些通用的模块例如权限控制、服务管理、人员管理等。

本体管理子系统:本体维护、本体可视化和本体查询等模块。系统提供了本体的维护界面,实现本体以及本体实例的查看、修改和删除。本体可视化方式有三种:树形结构、网状结构和 OWL 形式。其中树形结构和网状结构使用了开源的 Adobe Flex 图表制作组件 BirdEye,如图 4 所示。在业务流程的活动节点被实例化的时候,对应的输入端本体、输出端本体亦被实例化。本体查询模块可以查看单个本体的所有实例、与实例关联的活动、与活动关联的流程等,实现了对业务数据生命周期的监控和追溯。



图 4 井位部署顶层本体的可视化(网状结构)

特定的设备,开发平台设计和实现,可能无法满足普遍需求。相信在此基础上不断改进,系统的功能也会越来越完善。

参 考 文 献

[1] 刘记. 无线指纹考勤系统的设计与实现[D]. 河北工程大学,2007.

[2] 林培杰,朱安南,程树英. Android 数据库 SQLite 性能优化[J]. 计算机系统应用,2014,23(4):193-196.

[3] 何锡标,陈淑荣. 一种基于无线定位技术的 LBS 应用[J]. 微型机与应用,2014(9):7-10.

[4] 孔勇平,钟致民,杨广龙. 基于多运营商基站信号和 Wi-Fi 信号的混合定位技术探讨[J]. 移动通信,2013(11):32-37.

[5] 史斌斌,张琴,王宜怀. GPRS 无线指纹身份验证系统的应用研究[J]. 单片机与嵌入式系统应用,2011(7):71-73.

[6] 彭邦伦. 基于 Web 服务的接口平台的设计与实现[C]//第五届中国软件工程大会,2008.

[7] 章通,陈金玉. 网络指纹考勤系统的设计与实现[J]. 计算机应用与软件,2011,28(2):73-75.

[8] 博客园. 数据同步解决方案[R]. 2008. <http://www.csharpwin.com/csharpspace/1940.shtml>.

[9] 彭曙蓉,王耀南. 一种基于指纹识别的网络考勤系统[J]. 长沙电力学院学报:自然科学版,2006,21(2):56-58.

[10] 李悦,刘广荣,林锡龙. 一种便携式自动指纹识别系统的设计与实现[C]//2004 全国光学与光电子学学术研讨会,2005 全国光学与光电子学学术研讨会,广西光学学会成立 20 周年年会论文集,2005.

(上接第 83 页)

流程管理子系统:流程管理子系统包括业务过程建模、存储、管理和可视化、过程知识管理等。本体库是该子系统的基础,活动节点输入输出对象的定义、核心算法的参数对象的定义等都是基于本体进行的,这样保证了活动定义的普适性和通用性,以及不同活动节点之间数据传输的一致性。该模块提供了一个可视化的知识表征工具,专家可以将自己的经验知识以过程模型的形式录入到系统中,实现了对经验知识的存储、管理和回溯,如图 5 所示。

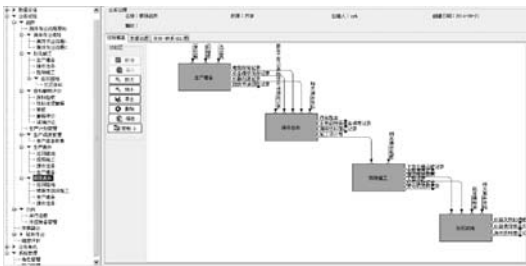


图 5 裸眼测井业务流程

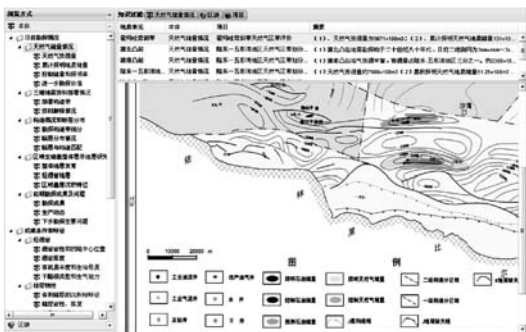


图 6 知识管理子系统

知识管理子系统:主要包括知识检索和知识可视化,如图 6 所示。基于本体面向过程的知识模型,实现了数据、知识和业务的一体化管理,提高了知识检索的查全率和准确率。用户输入检索关键词,系统向用户提供本体知识、数据知识和过程知识。

3 结 语

知识的创造、共享和再利用不是在真空中发生的,知识的收集和再利用只有与特定的业务流程密切联系,才能有效地发挥作用。为了解决企业业务流程控制和经验知识的管理问题,本文提出了基于本体面向过程的知识模型。本体模型为业务模型的输入输出提供了规范化的概念模型,过程模型定义了主题业务的具体流程、资源以及输入输出对象。POKMS 的建立,实现了数据流、知识流和业务流的融合,实现了专家知识和经验知识的存储、管理和回溯。基于流程化的管理,将引导企业的经营方式由以职能为中心向以流程为中心转变,促进企业内部各个部门之间的沟通协作和知识共享,提高企业生产效率。该系统在胜利油田井位部署过程中的应用,充分发挥了已有成果和经验知识的作用,对于新区块的勘探和老区块的二次调整开发具有重要的指导作用。

参 考 文 献

[1] 陈晰明. 信息管理 with 知识管理的比较研究[J]. 情报科学,2003,21(4):381-384.

[2] 张树玲,孙波,田艳琴. 过程建模技术研究综述[C]//Proceedings of 2010 Third International Conference on Education Technology and Training (Volume 7). 2010.

[3] Laguna M, Marklund J. Business process modeling, simulation and design[J]. CRC Press Taylor and Francis Group Boca Raton Florida Usa, 2005.

[4] Vergidis K, Tiwari A, Majeed B. Business process analysis and optimization: beyond reengineering[J]. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 2008, 38(1):69-82.

[5] 张晓刚,李明树. 基于工作流的知识流建模与控制[J]. 软件学报, 2005, 16(2):184-193.

[6] Li D, Kang L, Cheng X, et al. An ontology-based knowledge representation and implement method for crop cultivation standard[J]. Mathematical and Computer Modelling, 2013, 58(3):466-473.

[7] Hai R, Theißen M, Marquardt W. An ontology based approach for operational process modeling[J]. Advanced Engineering Informatics, 2011, 25(4):748-759.

[8] GarciaCrespo A, RuizMezcua B, LopezCuadrado J L, et al. Semantic model for knowledge representation in e-business[J]. Knowledge-Based Systems, 2011, 24(2):282-296.

[9] Santoso H A, Haw S C, Abdul-Mehdi Z T. Ontology extraction from relational database: Concept hierarchy as background knowledge [J]. Knowledge-Based Systems, 2011, 24(3):457-464.

[10] Jia C X, Hu W, Bai W Y. SMap: Semantically mapping relational database schemas to OWL ontologies [J]. Journal of Computer Research and Development, 2012, 49(10):2241-2250.