

工程知识管理系统模型研究

王宏君¹ 蔺娜¹ 鬲玲¹ 陈金梁¹ 敬石开^{1,2}

(1. 北京神舟航天软件技术有限公司, 北京 100094; 2. 北京理工大学机械与车辆学院, 北京 100081)



摘要: 针对工程知识管理系统模型多样化的现状, 从需求、功能和应用三个方面论述了对工程知识管理系统模型的理解和认识以及需求、功能及应用模型的含义和组成。提出工程知识管理系统模型由知识需求、系统功能和知识应用构成。在工程知识管理系统模型的基础上, 成功设计开发出工程知识管理系统实例, 并在航天企业中形成了一定规模的应用。

关键词: 语义引擎; 横向管理; 知识助理; 过程知识; 推荐引擎

Model Research of Engineering-knowledge Management System

Wang Hongjun¹ Lin Na¹ Ge Ling¹ Chen Jinliang¹ Jing Shikai^{1, 2}

(1. Beijing Shenzhou Aerospace Software Technology Co. Ltd, Beijing 100094;

2. School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081)

Abstract: Targeting at the various models of engineering knowledge management system, the paper describes the author's understanding of engineering knowledge management system model from three aspects including requirement, function and application. The definition and components of such three aspects are discussed in detail. It is proposed that an engineering knowledge management system is composed of knowledge requirement, system function and knowledge application. Cases of engineering knowledge management system are successfully designed and developed based on the model of engineering knowledge management system, and from a certain scale.

Key words: semantic engine; horizontal management; procedure knowledge; knowledge assistant; recommending engine

1 引言

工程知识^[1]数量众多、类型多样、存储分散、价值密度高, 内容广泛。知识管理以知识为管理对象, 管理知识的生命周期。工程知识管理是知识管理在复杂产品制造领域知识管理的分支, 以工程实践知识为管理对象, 管理工程实践知识的产生、应用与消亡的全过程。工程知识管理需要管理一切工程知识, 管理对象、范围不确定, 因此工程知识管理是开放式的管理, 必须包罗万象又必须井然有序。

工程知识管理属于横向信息管理。由于工程知识数量、类型、存储和价值方面的特点, 工程知识管理

与产品、工艺等数据管理截然不同。产品、工艺等数据管理属于产品生命周期中的一个管理环节, 管理对象确定, 管理范围明确、管理理论成熟, 属于纵向信息管理; 工程知识分布在产品、工艺等业务系统和专家头脑中, 它的任务是将处于产品生命周期中分散状态的同种知识的不同表达形式统一地组织管理起来, 将单元工程知识组成一条知识链和知识串。工程知识管理范围、对象广泛, 属于横向信息管理。

工程知识管理以概念为核心。零件是产品数据管理的核心, 工序是工艺数据管理的核心, 工程知识管理的核心是概念。概念是工程知识的浓缩, 一个工程

基金项目: 国家高技术研究发展计划(863计划) 2013AA040605 汽车及模具产业链设计制造协同服务平台研发与应用。

作者简介: 王宏君(1972-), 工程师, 机械制造工艺与设备专业; 研究方向: 知识管理、可制造性分析、故障预测。

收稿日期: 2015-05-05

概念代表一条或者多条工程知识；由一组专业工程概念组成的样本空间能够涵盖和描述相应的一系列工程知识。一般地，概念从属于特定的专业分类，与概念关联的工程知识顺理成章地隶属于此分类，概念建立了专业分类与工程知识之间的关系。知识检索的检索式由概念组成，知识订阅的需求最终分解为概念，知识地图以概念为纽带关联工程知识。

工程知识管理以语义^[2]管理为基础。工程知识标注、传输和展示是工程知识管理系统的基本数据处理方式。通常，利用三元组的形式表达工程知识的语义，通过三元组完成工程知识语义信息的标注、传输、展示。缺少语义的工程知识管理是缺乏灵魂的知识管理系统。

2 需求模型

工程知识管理需求模型^[3]是来源于工程实践的业务成果和业务活动的知识管理需求描述。在工程实践中产生结论型知识和过程型知识。结论型知识是工程实践中形成的共识、经验、禁忌、理论等对某种现象、问题的陈述性描述。过程型知识是工程实践活动先后顺序、作业流程的过程性描述。在工程实践中，过程型知识贯穿结论型知识；结论型知识支撑工程实践活动。工程知识需求模型表达如下：

Model_{需求}：[result_EKM, process_EKM, relation_EKM]

其中，result_EKM 为结论型知识，process_EKM 为过程型知识，relation_EKM 为 result_EKM 和 process_EKM 的关系。

结论型知识广泛蕴藏于工程实践产生的成果文档、参与者的头脑中。工程实践成果文档种类多，有设计文件、工艺文件、生产文件等；成果文档的数量呈现急速上升趋势，尤其近几年以几何级数的速度增长；工程实践活动专业性强、要求高，知识密度分布不均衡，其中专家头脑中的经验禁忌知识价值较高。面对种类多、数量大、增长速度快、价值密度^[4]不均衡的知识来源，需要采取相应的自动化手段获取知识。因此，结论型知识需求模型是面向大数据的工程知识管理需求的描述，表达如下：

Model_{result_EKM}： [result_EKM_{volume} , result_EKM_{velocity} result_EKM_{variety}, result_EKM_{veracity}, result_EKM_{auto}]

其中 result_EKM_{volume} 代表结论型知识需求数量大，result_EKM_{velocity} 代表结论型知识增长需求

速度快，result_EKM_{variety} 代表结论型知识需求种类多，result_EKM_{auto} 代表结论型知识需要自动化处理。

过程型知识是工程实践中提炼、固化和验证的作业活动的过程描述。过程型知识大多编入工程实践指导性的规范汇总，例如作业程序文件、质量程序文件等。过程型知识由活动参与者整理、提炼而成，具有很强的指导性。过程型知识还有一项更重要的特性，它是结论型知识的贯穿线索，是结论型知识的应用链路。因此，过程型知识需求模型是工程实践作业过程指导和贯穿结论型知识的需求描述，表达如下：

Model_{process_EKM}：[process, relation_EKM]

其中，process 代表工程实践过程，relation_EKM 代表 result_EKM 和 process_EKM 的关系。

综上所述，工程知识管理需求模型是结论型知识管理需求和过程型知识管理需求模型描述总和。

3 功能模型

工程知识管理功能模型^[5]由四大引擎为核心构建，它们分别是语义引擎、评价推荐引擎、知识片段抽取引擎和业务流程引擎。语义引擎主要承担在结论型知识中完成语义标注的工作；评价推荐引擎主要承担用户行为分析、结论型知识和过程型知识的推荐；知识片段抽取引擎主要承担从半结构化、非结构化的知识源中定向建立结构化结论型知识和过程型专项知识库的工作；业务流程引擎主要承担结论型知识、过程型知识融入工程实践活动的工作。工程知识管理系统功能模型表达如下：

Model_{功能}：[Engine_{ontology}, Engine_{evaluate&recommend}, Engine_{extract}, Engine_{business}]

其中，Engine_{ontology} 为语义引擎，Engine_{evaluate&recommend} 为评价推荐引擎，Engine_{extract} 为知识片段抽取引擎，Engine_{business} 为业务流程引擎。

3.1 语义引擎

语义是工程知识管理系统的基本特征，用三元组的方式描述。语义引擎是工程知识管理系统的基础引擎，承担向普通文本注入语义信息的工作。它由前处理、实体识别、语义标注^[6]和后处理等环节构成。文本归一化是语义引擎的前处理环节，它将结构化、半结构化和非结构化的文档处理成相同格式的文本；实体识别是语义引擎识别归一化文本中的特定文本，并标明词性；语义标注是实体识别发现的文本与关系型实体关联，将这些信息附加归一化文本，共同存储；

后处理是语义引擎消除实体识别中发现的二义实体。归一化文本经过语义引擎处理后,具有语义特征。在知识检索、订阅、地图的数据传输、存储和展现中,利用三元组作为数据的表达形式,完整的表达语义特征见图1。

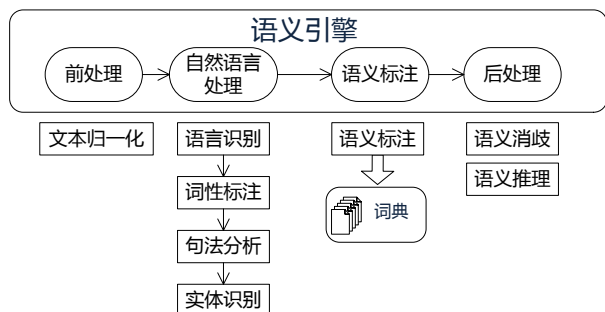


图1 语义引擎原理图

3.2 推荐引擎

用户行为数据是工程知识管理系统中第二类知识数据,它是评价推荐引擎^[7]的输入数据。评价推荐引擎根据用户行为数据中记录的信息,通过评价算法分析数据,生成推荐指标,将推荐数据返回用户。评价推荐引擎属于数据挖掘范畴,比较耗费资源,一般采用离线间歇式评价推荐,以独立服务器的方式提供评价推荐服务。评价推荐引擎应该提供一系列的WebService API供应用系统调用。为了适应评价指标变化,允许不同业务编写不同的算法,实现评价算法可扩展,如图2所示。

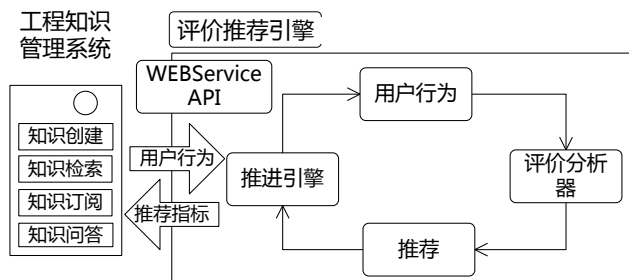


图2 推荐引擎原理图

3.2.1 知识片段抽取引擎

知识片段^[8]抽取引擎属于文本挖掘范畴,通过文本分词、实体识别、词典标注、本体标注、规则标注等过程,从半结构、非结构化文本中抽取结构化知识片段,解决专业知识库建立困难的问题。知识片段抽取引擎是定向处理引擎,在使用过程中需明确抽取知识片段的范围,有针对性地收集词汇建立词库,收集

专业概念建立本体库,收集文本表达规则建立规则库。在此基础上,知识片段抽取引擎自动建立专项的专业知识片段库。知识片段抽取引擎处理过程如图3。

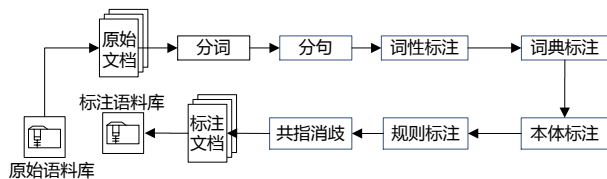


图3 知识片段抽取原理图

3.2.3 业务流程引擎

工程知识管理主要承担工程实践活动中知识管理、知识助理、知识助手的职能。业务流程引擎^[9]是过程型知识和结论型知识注入工程实践活动的桥梁和纽带,充分发挥知识助手的职能,实现用知识支撑业务活动的目标。业务流程引擎是用流程形式描述作业过程流程,作业活动的输入与输出由业务流程引擎描述,作业活动涉及的工具与信息系统入口通过业务流程引擎整合和绑定,同时伴随以结论型知识和过程型知识。在作业活动中,业务流程引擎推荐工具、信息系统和工程知识到参与者桌面。在工程实践中,不存在孤立的信息系统和孤立的工具,全部由业务流程引擎统合在企业业务活动的运行过程中,所有的企业信息化系统和工具利用业务流程引擎整合为一个整体。业务流程引擎将会建起以流程为导向,统合工程知识管理与业务信息系统的业务运行平台。

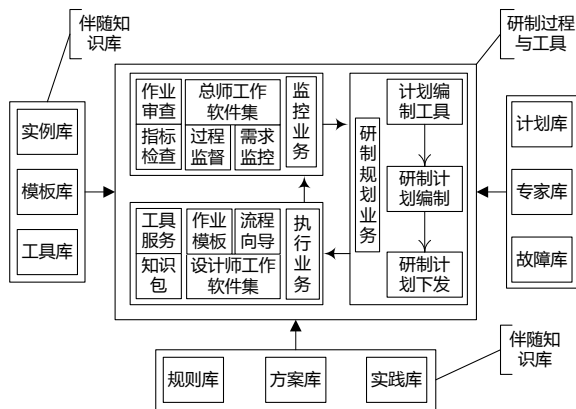


图4 业务流程引擎原理图

4 应用模型

工程知识管理系统应用模型是将工程知识应用到业务活动中,用知识支撑业务活动方式的描述。工

程知识用到业务活动中有多种方式: a. 用户登录工程知识管理系统检索、订阅知识; b. 将工程知识管理系统中知识检索、订阅与业务系统集成, 实现业务活动能得到知识管理系统支持; c. 用工程知识管理系统 APP 将知识检索、订阅驻留在用户桌面上; d. 通过工程知识服务门户, 保障用户在业务活动中得到实时知识服务。因此, 工程知识管理系统应用模型由工程知识管理系统、工程知识管理系统 REST API、工程知识管理 APP 和工程知识服务门户构成, 表达如下:

Model 应用: [System, REST, APP, Portal]

其中, System 代表工程知识管理系统, REST 代表工程知识管理系统 REST 接口, APP 代表工程知识管理 APP, Portal 代表工程知识服务门户。工程知识管理系统以应用为导向, 以丰富的应用方式见长。

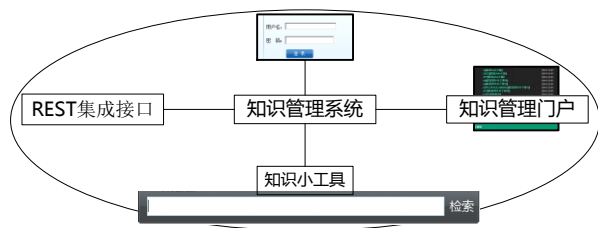


图 5 应用模型原理图

5 结束语

工程知识管理系统模型由需求模型、功能模型和应用模型构成。三种模型互为基础、互相影响。处于反复迭代的动态过程中。通过对三种模型的认识和理解, 实现了第一代工程知识管理系统, 在典型用户的示范应用中, 验证了三种模型的合理性和全面性。第一代工程知识管理系统在用户现场为用户在知识获取、管理和应用中提供了一系列的便利和支撑作用, 得到用户的认可。企业的工程实践数据初具规模, 显现出大数据的一些特性; 而且企业在硬件环境方面采

用了虚拟化的运行方式。这两个因素给工程知识管理系统模型提出了新要求, 要求工程知识管理系统适应大数据^[10]和云计算^[11]的发展需要。工程知识管理系统模型需要适应云端化的要求, 实现管理功能分布化, 执行功能云端化。在硬件资源无限的条件下, 软件执行能力也应无限。工程知识管理系统模型在大数据和云计算的推动下, 反复迭代且持续发展。未来, 工程知识管理系统模型需要面向大数据和云计算开展研究。

参考文献

- 1 胡熠, 周雄辉, 李从心. 基于特征模型和 Web 服务的工程知识库共享[J]. 上海交通大学学报, 2010, 9: 1229~1234
- 2 马斌, 王金虹, 闫娟娟, 等. 基于本体的智能语义检索模型设计与研究[J]. 情报科学, 2015, 2: 46~49
- 3 谢仲文, 李晓燕, 等. 一种将需求模型转换为软件体系结构模型的方法[J]. 计算机科学, 2014, 5: 196~203
- 4 侯健卫, 杨波, 李栓林, 等. 基于 MPP 和 Hadoop 的低价值密度结构数据联合处理方法的设计与实现[J]. 中国新通信, 2015, 1: 123~124
- 5 于菲, 檀润华, 曹国忠, 等. 基于系统功能模型的元件裁剪优先权研究[J]. 计算机集成制造系统, 2013, 2: 338~347
- 6 邱均平, 牟楠, 楼雯, 等. 国内外语义标注研究进展分析[J]. 计算机集成制造系统, 2014, 5: 12~16
- 7 袁义龙. 基于 Mahout 的电影推荐引擎的设计与实现[D]. 北京邮电大学, 2014
- 8 戴璐, 丁立新, 薛兵. 一种摘要中隐含的知识片段的挖掘方案[J]. 计算机科学, 2013, 2: 218~221
- 9 武德港. jBPM5 工作流引擎在业务流程管理系统中的研究与应用[D]. 北京邮电大学, 2014
- 10 张引, 陈敏, 廖小飞. 大数据应用的现状与展望[J]. 计算机研究与发展, 2013, S2: 216~233
- 11 林闯, 苏文博, 孟坤, 等. 云计算安全: 架构、机制与模型评价[J]. 计算机学报, 2013, 9: 1765~1784