



ISSN 1671-2064
CN11-4650/N

09_下

2021

总第366期

中国科技纵横

中华人民共和国科学技术部主管

ISSN 1671-2064



9 771671 206213



187

万方数据数字化期刊群收录期刊

龙源国际期刊网全文收录期刊

中国核心期刊(遴选)数据库收录期刊

中文科技期刊数据库全文收录

面向流程行业工业知识分类方法与建模技术研究构建*

王佳¹ 李媛智¹ 续飞飞² 高慧芳¹

(1.中国软件评测中心, 北京 100048; 2.北京北科亿力科技有限公司, 北京 100102)

摘要: 为流程行业制造企业复杂异构的知识资源提供统一的知识构建方式, 建立一套知识管理范式。基于对流程工业在知识发现、自动推理以及自主决策等方面研究分析, 将知识分为数据类知识、过程类知识、文档类知识和经验类知识 4 种知识表现形式, 建立针对企业不同表现形式的知识资源建模流程。形成基于不同知识类型的制造企业复杂知识资源建模技术, 为知识的固化、凝练和体现提供了一种方法, 满足不同企业用户的多维要素的工业知识软件化产生不同类型的工业软件知识需求。

关键词: 流程工业; 知识分类; 知识管理; 知识模型; 模型构建

中图分类号: TP277

文献标志码: A

文章编号: 1671-2064(2021)18-0146-05

0. 引言

流程工业在国民经济中占有基础性的战略地位, 石油、化工、钢铁、建材等都是典型的流程工业。经过几十年的发展, 我国已成为世界上门类最齐全、规模最庞大的流程工业大国, 流程工业的生产工艺、生产装备和生产自动化水平都达到世界领先水平。目前, 流程工业生产中普遍应用了分布式控制系统 (Distributed Control System, DCS), 该系统可以实时采样流程工业生产过程中各个环节的生产状态, 特别是传感器技术以及数据库技术的快速发展, DCS 系统对流程工业生产状态的监控越来越多。流程工业控制系统采样得到的数据经过长时间的积累形成规模庞大的历史数据库, 在这些数据中包含了可用于生产和管理的大量有价值的信息和知识。除了隐含在直接来自生产现场具有多维度、多尺度、异构的海量数据中的知识之外, 在流程工业中的知识还包括: (1) 具有隐蔽性、非量化的经验知识; (2) 能够反映生产过程的气液固多相变化机理的机理知识^[1-3]。知识资源已经成为支撑流程行业企业生存与发展的重要力量, 流程行业对知识资源的高效应用具有非常迫切的需求。

以面向业务的工业应用软件的开发与应用为实现路径, 工业技术、工艺经验、制造知识和方法等通过软件实现显性化、数字化和系统化。目前工业知识的研究在流程行业中应用已经研究很久, 主要是通过贝叶斯网络^[4]、模糊系统^[5]、以及专家系统^[6]等完成在流程工业对知识的自动推理以及实现自主决策。但是随着实际系统的复杂性不断增加, 上述方法的复杂性会大大增加, 使得大多数理论应用都是停留在研究层面, 实际应用场景非常少。知识发现主要是从专家人员、文件或其他系统等提取知识并向知识型系统转移的过程。常用的方法包括事件关联规则挖掘^[7]、人工神经网络^[8]、数据统计分析方法^[9]、遗传算法^[10]、决策树^[11]等。由于工业数据数量大并且质量难以保证, 工业过程中知识的提取仍然是主要难点。

知识分类和管理是企业发展的的重要组成部分, 但如何做仍然是一道难题, 目前大部分企业都进行过知识相关研究工作。但是多数企业仍然存在无知识、弱知识、死知识的问题, 并且未对相应知识进行分类, 也没有正在将知识融入产品研制和生产过程。因此, 本文对流程工业企业知识自动化的相关的研究现状进行了分析和介绍, 详细阐述流程行业企业知识建模框架和不同类型的知识分类, 并以知识发现为流程, 构建了根据不同知识分类的模型, 扩展了知识发现和构建研究的理论体系。

1. 流程行业知识发现的研究现状与问题

伴随工业 4.0、工业互联网、智能制造等政策的发布, 知识将成为推动企业发展的重要资源。企业制造过程包括 3 个重要环节: 一是研制过程; 二是生产过程; 三是运营过程。知识作为重要的生产要素, 不仅要考虑产品生产周期全过程如何智慧化, 同时还要考虑产品运维过程对研制和生产的反向影响。

对于不同的生产过程和知识有不同的结合点, 知识跟企业产品研制过程关系最紧密, 研制是知识的归宿。普通制造过程的知识密度比较低, 但是工艺设计指南、产品使用手册、作业指导书等知识如果能够得到很好的整理, 将对制造过程产生很大的价值。生产过程的知识大多来源于大数据分析结果, 通过制造过程的异常数据可以判断设备、生产线以及产品运营过程的健康程度, 利用大数据技术科研对制造过程、产品运营进行预测、改进甚至自动决策。产品运维过程也会产生大量的数据, 这些数据都可以被知识化, 能够指导企业优化研制和生产过程, 及时修改产品的设计和生产。

流程行业是知识技术非常集中的产业, 对知识资源的应用有非常迫切的需求。当前, 知识分类和知识获取是知识资源应用的基础和关键, 对知识资源的获取、组织和利用知识创造价值的速度是支撑制造企业生存与发展的核心

收稿日期: 2021-08-04

* 基金项目: “2018 年智能制造综合标准化与新模式应用项目——面向流程行业工业技术软件化关键技术标准研制及验证平台建设”

作者简介: 王佳 (1986—), 男, 河北保定人, 博士, 高级工程师, 研究方向: 智能制造评估、安全管理。

146 2021年9月下 第18期 总第366期

横

学技术部主管

从知识发现的本质来看,它集成规律、策略和技术过程。对于国内知识发现研究而言,在大数据驱动的知识发现研究已经从决策树、遗传算法、机器学习挖掘等具体的理论方法转变为如何应用知识发现方法解决各行各业遇到的问题,并且逐渐面向行业服务的综合应用研究转变,存在的问题主要有以下。

实际工业场景问题抽象成数据挖掘问题

工业场景的数据挖掘具有重要意义,在当前流程工业具备高水平业务知识和数据挖掘能力的人员非常少,将实际工业场景抽象成对应数据挖掘问题成为获取和增值过程的一个重大门槛。

数据质量差

生产过程中,采集的数据质量非常差强人意,采集的数据并不会进行数据处理,导致数据的完整性、唯一性和有效性无法保证,严重影响数据挖掘准确性等方面。

领域知识沉淀不足

流程工业知识的特点,工业知识、经验、机理沉淀不足,我国制造业工艺软件化基础薄弱,导致我国机理模型缺失较为严重,无任何机理模型,目前只有少数模型工作者才能完成工业现场复杂分析、精确判工作。

与建模

建模框架结构

企业中,数据、信息和知识是一个彼此关联和提炼的过程。数据在反复关联与使用中提升为信息,信息在反复关联和使用中转化为知识。数据类资源

通过标准化技术,可以使其获得有序化特征,但仍然属于低层次的知识形态。信息类资源对数据知识进行分析整理,通过结构化加工强化共享特征,经过分类、聚类、语义分析等手段结构化后,信息类资源形成各种分类知识库供参考查阅。

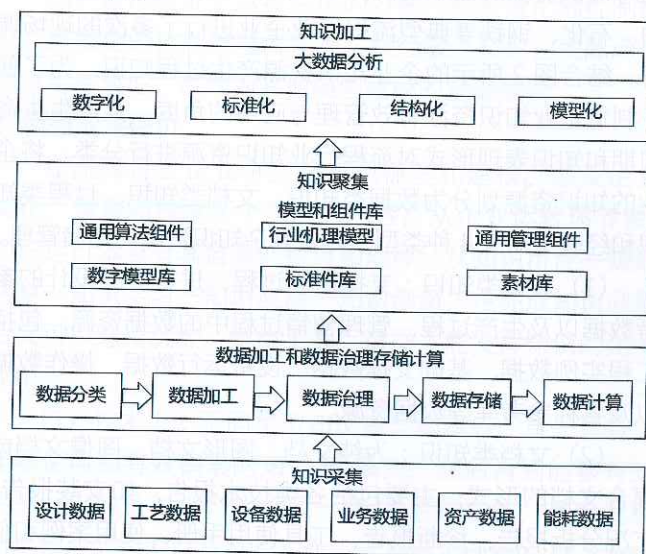


图1 知识建模框架

知识建模框架如图1所示,知识采集针对设计数据、工艺数据、业务数据等类型,是资源数据的基础层;数据加工和数据治理存储计算对采集到的数据进行分类、加工、治理、存储和计算;知识聚集是针对不同的知识形态,设计不同的知识模型和组件库,将知识模块和知识载体管理入库,最终形成各种类型的知识库,包括通用算法组件、行业机理模型、通用管理组件等;最后通过知识加工对知识库中不同类型的知识进行数字化、标准化、结构化、模型化等技术进行加工处理,提升知识层级,实现知

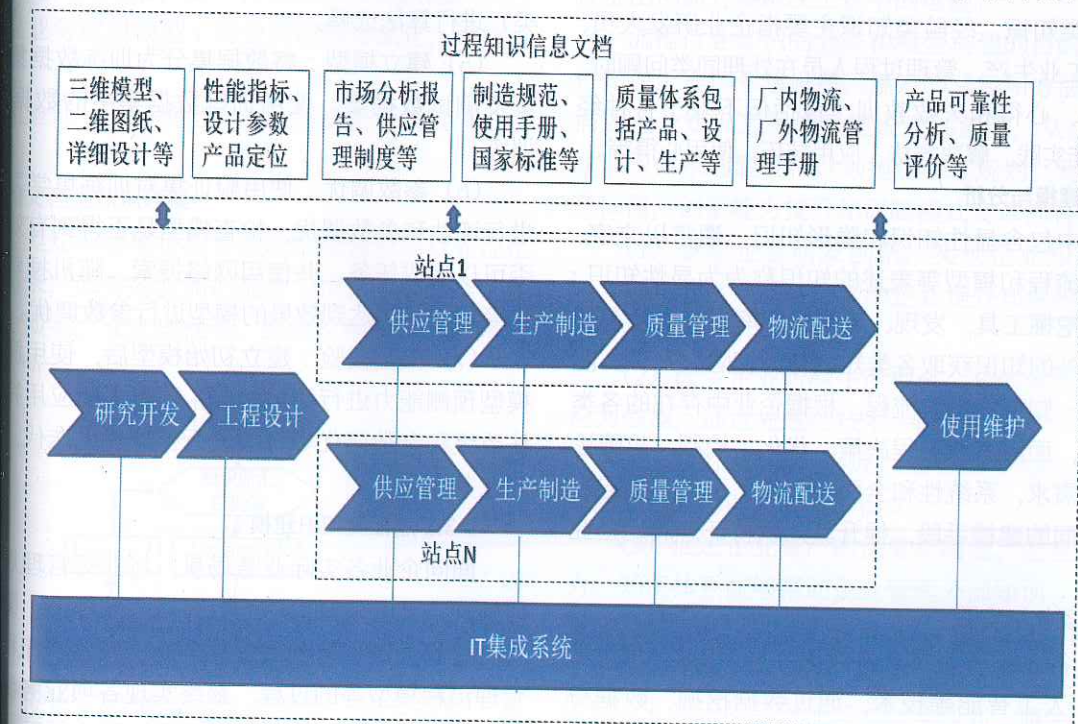


图2 企业供应链流程

识增值。

2.2 知识表现形式与分类

合理的知识分类是有效获取、管理与利用知识的基础。为深入了解和掌握流程行业企业复杂知识资源的特征及对知识资源的需求情况,完成对知识资源的梳理,对石油、石化、钢铁等典型流程行业企业进行了多次的现场调研。结合图2所示的企业知识资源产生过程知识,为了便于制造企业知识资源有效管理与应用的角度,从产生生命周期和知识表现形式对流程行业知识资源进行分类。将企业的知识资源划分为数据类知识、文档类知识、过程类知识和经验类知识4种类型,方便复杂知识资源的存储管理。

(1) 数据类知识:支撑研制过程、进行产品设计的参考数据以及生产过程、管理营销过程中的数据资源,包括工程实例数据、基础支撑数据、设备运行数据、操作数据以及各种参考库等数据资源。

(2) 文档类知识:为纯文档、图形文档、图像文档或混合文档的形式,主要包括各类技术报告,如安装报告、工况分析报告、诊断报告、工具使用手册、使用案例和使用方法等;工业活动过程中需要遵循的标准和约束条件,如相关国家标准、行业标准、企业标准、规范作业等。

(3) 过程类知识:包括科研、生产管理、营销活动外围的流程知识及内部的过程知识。科研、生产管理及营销活动外围的流程知识为科研、生产管理、营销活动和相关科研活动的逻辑关系,以数据流图的形式表示,还包括对科研活动的输入输出的选取原则、要求等进一步描述,一般以文字描述的形式表示。科研、生产管理及营销活动内部的流程知识,主要描述执行上述相关活动的设计子流程、执行步骤的说明,以 workflows 或文字描述的形式表示。

(4) 经验类知识:经验类知识主要指企业研发人员、研究人员等为工业生产、管理过程人员在处理同类问题时,提供成功经验、心得和失败教训。包括但不限于各种经验、教训、最佳实践、解决方法、应用技巧、使用心得等。

2.3 知识的建模与分析

知识分类中包含显性知识和隐性知识,通常以文字、图表、公式、流程和模型等表述的知识称为显性知识;通过自动知识挖掘工具,发现、收集知识库以外以及游离于管理制度之外的知识获取各类知识称为把隐性知识。区别于以往的单一知识的提取流程,根据企业中存在的各类知识表现形式,面向工业过程决策、优化、控制、故障诊断等实际应用需求,系统性和全面的提出了不同类型的资源信息采用不同的建模手段,提升其知识特征,满足企业的发展需求。

2.3.1 数据类知识建模

采集企业各业务环节中产生数据,基于工业大数据、工业区块链、人工智能等技术,通过数据挖掘、数据分析、数据处理等方法建立数据与业务的关联模型,实现以

数据支撑业务管理的决策优化。主要提取过程包括:数据处理、数据分析、特征提取、算法选择、建立模型、参数调优,模型验证等。数据建模过程可用如图3所示。

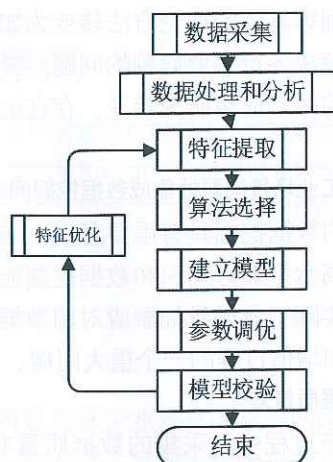


图3 数据知识建模流程

(1) 数据采集:收集生产管理运营过程中与模型相关的业务数据、操作数据、设备数据、监测数据、管理数据等。

(2) 数据处理和分析:在建立模型数据集的基础上,对数据进行预处理,包括:缺省值、异常值处理,以及数据的标准化等。通过初步数据探索,发现数据特征和规律,常用的数据分析方法有数据特征描述、相关性分析、主成分分析等。

(3) 特征提取:通过过滤法、包裹法、嵌入法等方法选择或者构建与模型相关性较大的一部分特征,作为模型的输入,以减少数据噪声对模型结果的影响。

(4) 算法选择:根据业务的任务类型(监督学习、非监督学习、强化学习)以及涉及的问题(回归、分类、聚类)进行算法选择。

(5) 建立模型:将数据集分为训练数据集、验证数据集及测试数据集,使用训练数据集中的数据对模型进行训练。

(6) 参数调优:使用验证集对训练集学习得到的模型进行评估和参数调优,检查模型是否得到有效的训练或是否可以完成任务。并使用网络搜索、随机搜索或贝叶斯优化等方法对为达到效果的模型进行参数调优。

(7) 模型校验:建立初始模型后,使用测试数据集对模型预测能力进行测试验证,同时不断应用特征优化、算法改进和参数调优等方法进行模型优化迭代,直至满足功能业务需求。

2.3.2 过程类知识建模

面向企业各实际业务场景,将业务管理规范、业务流程管控、业务信息流转等业务以工业组态、图形化建模等方式提炼为数据管理流程模型、工艺生产流程模型、业务管理流程模型等的过程,最终实现各项业务活动的信息化管理。过程类知识建模过程可用如图4所示。

横

学技术部主管

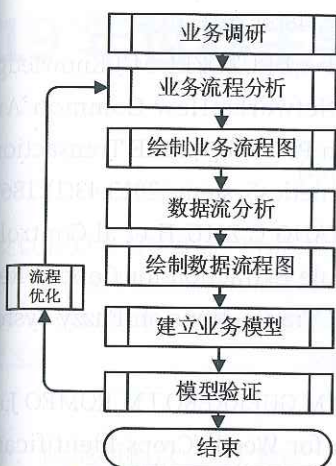


图4 过程知识建模流程

调研：通过业务调研方式完全理解业务流程类别（营销流程、采购供应流程、计划运流程以及企业管理流程），流程涉及的组作规范、管理章程、规章制度、行业标准等

流程分析：针对具体业务流程，开展业务流止节点、过程活动节点及步骤，模型涉及的门分析，以及活动流程中的控制流。

业务流程图：根据业务流程分析，使用跨职ML活动图等方法绘制。

流分析：梳理完成业务流程的数据输入项以而逐渐细化为最细粒度数据，一般包含数据、传递存储等功能。

流程图：依据数据流图的基本规则：外部实、数据流和数据存储，建立描述数据流动、逻辑关系图，即：将业务流程抽象化。

业务逻辑模型：将业务流程抽象化，建立业

验证：验证模型产生的结果是否实现其业务

文档类知识建模

类知识建模将相关领域的经验知识以显性化来，如将经验知识提炼为预警报警标准、操预案、安全规程等。经验文档知识类建模过所示。

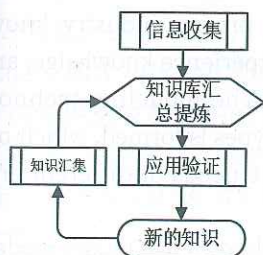


图5 经验知识建模流程

收集：汇集与经验类知识（相关经验、教使用心得等）梳理相关的数据、文件、图

片等数据。

（2）知识库汇总提炼：建立相关问题的知识库，通过简单的数据分析、数据对标、业务对标、流程分析的方式，提炼与目标问题相关的新的结论。

（3）应用验证：以显性化的方式，如量化的控制标准、逻辑化的控制管理流程等。将经验类知识在实际场景的应用并验证。或作为其他模型建模的基础数据。

3. 模型成果评价

在流程行业企业中实施知识梳理和建模，涉及企业文化，管理和技术3个层次的内容，从企业运转环节看，知识运行本身有其自身的特点和规律。从知识沉淀—知识共享—知识学习—知识应用—知识创新，在完成知识创新后又可以继续沉淀，螺旋上升，这样才能发挥知识的无穷作用。

（1）技术上：不同类型知识模型构建后，首先具备了能够使用和有效积累的知识IT系统，然后再到能够有效共享、支持员工学习的更深一层IT系统功能，最后发展到具备获取知识以及协同工作、支持员工创新的IT系统。

（2）管理上：不同类型知识模型构建后，首先具备相应的知识管理机制和模式来保障有效知识的积累，然后具备保证企业员工采用一定方式共享经验、建立导师、传帮带等机制，最后发展到具有相应的流程和制度来保证积累的知识得到充分应用和知识创新。

（3）文化上：不同类型知识模型构建后，首先从员工养成及时总结工作经验和教训的习惯，然后到企业员工愿意将经验给其他同事分析，并付诸行动，最后发展到员工在工作中具备了充分利用过去经验和知识的意识，并能积极提出创新方案解决工作中的难题。

流程行业知识的分类和模型构建的目的在于为企业进行知识工程体系建设提供高效的路径和方法，该研究通过描述企业知识工程体系建设规划蓝图，从企业现状出发，对知识体系的建设和计划提供合理规划，对人才发展、资产保值，创新能力提升和研制模式变革都有明显的价值。

（1）驱动企业研发创新。实现企业知识的产生、应用、创新的良性循环，帮助企业提升知识的共享程度，由此带来产品和机制的创新。

（2）改变企业研制生产模式。知识体系以科学的方法论为依据，通过先进的技术手段在企业中的组织、流程和IT系统上进行建设，形成以业务为中心的，以知识流和信息流为支撑的研制生产模式。

（3）促进企业资产的保值和增值。企业知识属于软实力，知识体系能够帮助企业管理外部知识、内部知识、专家经验知识等，实现企业智力资产的持续积累和增值。

（4）加速人才培养、补充企业技术人员断层。实现对流程企业知识的系统梳理和科学规划，避免因人才流动造成的知识损失，为企业的知识传承奠定基础。

4. 结语

不同的知识有不同的知识分类理论,不同的目的有不同的知识分类标准,工业知识体系是一个多维度矩阵结构。针对流程行业企业内容复杂、格式异构、数量庞大的知识资源,以企业调研为基础,从企业知识资源管理的实际需求出发,提出了一般性和通用性的知识分类方法。在此基础上研究了知识建模技术,为制造企业复杂异构的知识资源提供了统一的知识获取方式。文中提出的知识分类方法和复杂知识资源获取模板,通过标准化的知识资源获取界面,结合人工获取和半自动化获取技术,实现了对制造企业复杂知识资源的高效获取,对企业未来较长时间内知识资源的建设给予指导。

加大对知识技术的研究和运用,优化数字技术路线。工业知识是构建工业APP的重要组成部分,工业APP作为工业互联网平台的业务应用是软件技术和工业技术深度融合的产物,将人的经验、知识提取出来,通过软件运行。全面采集工业数据,让生产线与数字孪生系统高度融合,不断提升生产与制造流程;通过知识图谱,将不同类别的制造业基础数据进行知识分类和建模,加强知识提取、关系挖掘、多领域融合,搭建知识服务平台,提高企业流程中各类问题的预见和解决能力。

参考文献

- [1] 桂卫华,王成红,谢永芳,等.流程工业实现跨越式发展的必由之路[J].中国科学基金,2015(5):337-342.
- [2] HEISIG P,ADEKUNLE S O,KIANTO A,et al.Knowledge Management and Business Performance:Global Experts' Views on Future Research Needs[J].Journal of Knowledge Management,2016,20(6):1169-1198.
- [3] 桂卫华,陈晓方,阳春华,等.知识自动化及工业应用[J].中国科学:信息科学,2016,46(8):1016-1034.
- [4] ZAGORECKI A,DRUZDZEL M J.Knowledge Engineering for Bayesian Networks:How Common Are Noisy-MAX Distributions in Practice[J].IEEE Transactions on Systems, Man, and Cybernetics Systems,2013,43(1):186-195.
- [5] ZHANG B,YANG C,ZHU H,et al.Controllable-domain-based Fuzzy Rule Extraction for Copper Removal Process Control[J].IEEE Transactions on Fuzzy Systems,2018,26(3):1744-1756.
- [6] MONTALVO M,GUERRERO J M,ROMEO J,et alAutomatic Expert System for Weeds/Crops Identification in Images from Maize Fields[J].Expert Systems with Applications,2013,40(1):75-82.
- [7] MANDA P,OZKAN S,WANG H,et al.Cross-ontology Multi-level Association Rule Mining in the Gene Ontology[J].Plos One,2012,7(10):e47411.
- [8] MARTINSEN K,DOWNEY J.Human-machine Interface for Artificial Neural Network Based Machine Tool Process Monitoring[J].Procedia CIRP,2016(41):933-938.
- [9] NAYAK A K,SHARMA K C,BHAKAR R,et al.ARIMA Based Statistical Approach to Predict Wind Power Ramps[C]//Power & Energy Society General Meeting.Denver,CO,USA:IEEE,2015:1-5.
- [10] BI W,DANDY G C,Maier H R.Improved Genetic Algorithm Optimization of Water Distribution System Design by Incorporating Domain Knowledge[J].Environmental Modelling & Software,2015(69):370-381.
- [11] WU T C,HSU M F.Credit Risk Assessment and Decision Making by a Fusion Approach[J].Knowledge-based Systems,2012(35):102-110.

Research and Construction on Knowledge Classification and Modeling Technology for Process Industry

WANG Jia¹,LI Yuanzhi¹,XU Feifei²,GAO Huifang¹

(1.China Software Testing Center, Beijing 100048;2.Beijing Beike Yili Technology Co., Ltd., Beijing 100102)

Abstract:It provides a unified way of knowledge construction for the complex and heterogeneous knowledge resources of the process industry, and establishes a set of knowledge management paradigm. Based on the research and analysis of knowledge discovery, automatic reasoning and independent decision-making in process industry, knowledge is divided into four forms: data knowledge, process knowledge, document knowledge and experience knowledge, and the knowledge resource modeling process for different forms of enterprises is established. The modeling technology of complex knowledge resources of manufacturing enterprises based on different knowledge types is formed, which provides a method for the solidification, refinement and embodiment of knowledge, and meets the needs of different enterprise types of industrial knowledge.

Key words:process industry;knowledge classification;knowledge management;knowledge model;model construction