Research on Ontology Modeling of Steel Manufacturing Process Based on Big Data Analysis

Qing Bao, Jian Wang and Jin Cheng

CIMS research center, Tongji University, 201804 Shanghai, China

Abstract. As an important method that steel industries ride the Indutrie 4.0 wave, knowledge management is expected to be versatile, effective and intelligent. Mechanism modeling difficulties, numerous influencing factors and complex industrial chains hinder the development of knowledge and information integration. Using data potentials, big data analysis can be an effective way to deal with knowledge acquisition as it solves the inaccuracy and imperfection mechanism modeling may lead to. This paper proposes a big data knowledge management system(BDAKMS) adhering to data driven, intelligent analysis, service publication, dynamic update principle which can effectively extracts knowledge from mass data. Then, ontology modeling gives the knowledge unified descriptions as well as inference details combined with semantic web techniques.

1 Introduction

Indutrie 4.0 was put forward in Hannover Messe and brings prospects to hasten the strategic transformation of steel industries as well. One of the aspects that Industrie 4.0 propels manufacturing revolution is comprehensive, unified and domain-oriented knowledge collection to instruct business operation. Ref.[1] proposes fast matrix-based approaches to knowledge acquisition in covering information systems and reduce the computational time compared with classical rough set, especially large-scale data sets. However, this method is limited to static information and discretization is not universal. Ref.[2] builds a generic EUEDE contextual architecture for knowledge acquisition, compatible with current DSS. Ontology is used for knowledge formalization, so that the relationship between the knowledge items contribute to the rules for decision making, but EUEDE excludes external factors which are important to decision knowledge extraction. Ref.[3] combines knowledge graph with case analysis to minimize the gap among distinct levels of expertise in access to knowledge. More than the difference or diversity as it has addressed, the inconsistency should not be neglected in multi-source tacit knowledge.

Ontology makes it possible to knowledge interchange and reusability which is essential to knowledge modeling. Manufacturing taxonomy helps ontology modeling and results in ontology based knowledge hierarchy[4]. Top-down development approach can satisfy diversified levels of ontology granularity. Lifecycle[5] management is proposed and brings an impetus to ontology based knowledge update. Ontology has not only been applied in the manufacturing process[6] but also product design and cost budgeting[7]. Ontology is the core of semantic web. While ontology enriches semantic description, knowledge is expressed more comprehensively and integrated from distributed systems. Evolved from RDF Schema, ontology web language(OWL) has a richer vocabulary to define relations for inference[8] and is developed to describe ontologies.

Big data are produced in every segment in manufacturing: For enterprise performance, big data help identify the product demands, productivity and performance via diverse business objectives. For production, big data make it possible to detect the right facility causing the product flaw. Currently, steel making industries have applied big data analysis to quality management, ore blending solution optimization, supply chain management, equipment fault diagnosis and financial, etc. For example, BAO Steel finds better ore blending schemes with analysis of historical data. Anyang steel industry employ big data analysis to find out the key factors that cause the steel cracks[9].

Focusing on the difficulties in knowledge acquisition of steel industries, this paper presents a big data analysis based knowledge management structure (BDAKMS). Ontology is used to model the domain knowledge and finally enhance the usability and interoperability. This paper also gives an ontology modeling case based on big data analytical results and achieves accurate inference through SWRL rules and semantic web technologies. This paper focuses on discussions about ontology modeling methodology and semantic deduction. A converter steel making case is given to explain and verify the techniques' validity.
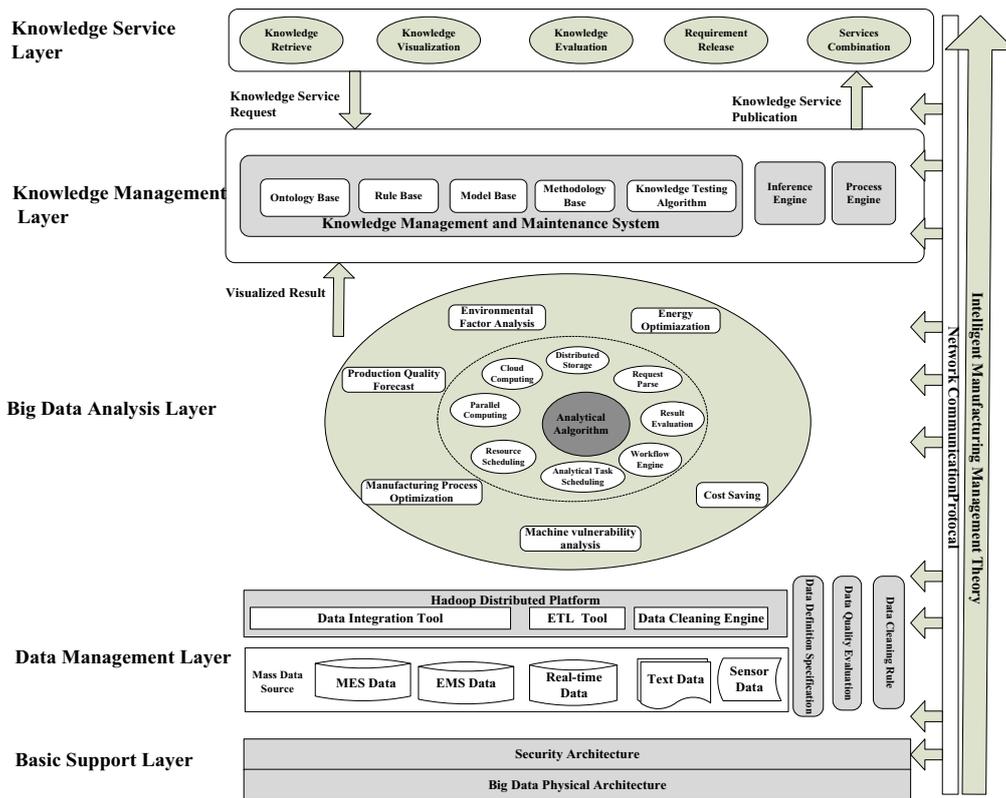
2 BDAKMS

Figure 1. Construction of BDAKMS

2.1 Construction of BDAKMS

BDAKMS adopts data driven, intelligent analysis, service publication, dynamic update pattern to transform data to knowledge: Starting with primary data, knowledge is obtained by intelligent analysis methods. Knowledge is accessed with calling service encapsulation by users, accomplishing the interchange between the knowledge level and data level. The construction of BDAKMS is shown in Fig 1. This structure consists of five layers:

Basic Support Layer. The physical platform adopts Hadoop distributed environment with master slaver structure. NameNode is set to task scheduling while DataNode task execution. Traceability is introduced to security architecture to record the data transmission histories and guarantee the data credibility and privacy.

Data Management Layer. Service data object (SDO) produces the data graph according to the characteristics of source data, which is used in MapReduce for datatransformation and analysis. The ETL tool splits the data and balances the load to integrate preprocessed data into distributed data storage environment. Data access service (DAS) APIs provide data access. DAS is made up of meta data ETL, data registry and data reading. In fact, the data that DAS reads should match the mapping mode which is a set of mapping rules between meta data and data to be loaded. In fact, data information structure, data mapping relations and customized filter are designed in HBase for data unification. For oxygen gas consumption in EMS, the information storage structure is shown in

Table1. Data source IP: port is saved in Row key while column description in Column family.

Table 1. HBase storage structure example

Table Name	EMS_Converter_Oxgen	
Row Key	EMS-TP562039:1521	
ColumnFamily	Column Description	
Column Description	Table Name	EMS_
	Data Property	Consumption
	Object Property	_isLinkedFacility
	Mapping Rule	Rule2344
	Unit	IU

2.1.1 Big Data Analysis Layer.

This layer contributes to relationship discovery among features and prediction observation through data mining. The core of big data analysis is to select the appropriate analytical algorithm through programmed Map and Reduce functions. With support tools and middleware like distributed storage, cloud computing, Linux cluster, analytical task scheduling, resource dispatch and data fusion, the goal of this layer is to achieve environmental factor analysis, product quality forecast, manufacturing

process optimization, machine vulnerability, cost saving and energy optimization, etc.
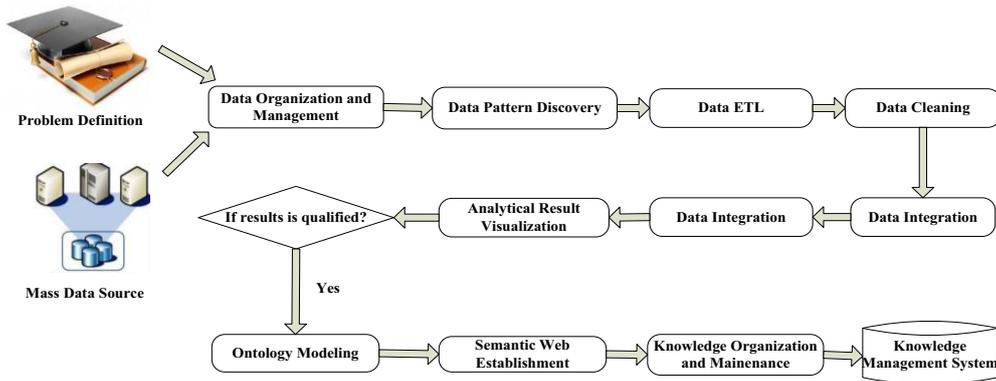
2.1.2 Knowledge Management Layer.

Figure 2. Knowledge acquisition flow based on big data analysis

incorporates the valued knowledge with ontology description language. The new knowledge entities combined with the existing have to be inspected by knowledge testing. Knowledge evaluation module assesses the knowledge items periodically to eliminate the outdated or inaccurate knowledge and make a request to big data analysis for update if possible. Moreover, a completed knowledge management system also needs inference engine and workflow engine to circulate the knowledge resources.

2.1.3 Knowledge Service Layer.

In addition that this layer has knowledge retrieval and inference modules like traditional domain knowledge base, it enables new knowledge publication interface, service composition according to certain business needs and user analysis requests that act as big data analysis objectives.

2.2 Knowledge Acquisition Based on Big Data Analysis

Based on big data analysis, knowledge acquisition is the prerequisite for BDAKMS. As knowledge is organized in lifecycle management which indicates that the process from data to knowledge is iterative and repetitive. Dynamic external and internal environment of enterprises also contribute to the iteration. In general, big data analysis layer keeps on extracting data from MES, ERP, SCADA, sensors and real-time database to do massive analytical works. The knowledge evaluation module assesses the knowledge quality of both newfound and existing knowledge items, and replaces the outdated information with the most appropriate. The knowledge acquisition flow is shown in Fig 2 and it is summarized as three phases.

2.2.1 Phase1 Data preparation.

Ontology base, rule base, methodology base and model base express formalized knowledge resulted from big data analysis. Knowledge testing algorithm will verify the consistency and integrity of the knowledge system. The knowledge is managed in
lifecycle method: The knowledge system only

Mass data source can be divided into SQL data and NoSQL data(log, web, video and sensor data, etc.) that have been stored in distributed memory system or cloud storage infrastructure. The storage system is equipped with fault tolerance, load balancing and cluster mechanism, and it is designed to provide the multi-source and heterogeneous data with uniform environment. Data ETL (Extraction, Transform, Load) extracts data across diverse platforms in preparation for data integration. For SQL data, ETL tool accesses data easily in SQL query. For NoSQL, ETL is implemented in Hadoop for data search. Furthermore, data cleaning improves the data consistency, integrity, usability and accuracy. For instance, tuple neglect or default fill deals with data missing. Clustering or regression method handles noisy data[10]. To provide the heterogeneous data access interfaces, data integration unifies the specification of data definition and constraint.

2.2.2 Phase2 Data analysis and result evaluation.

Data analysis methods are not confined to traditional strategies like reports, statistics, and combination models. Instead, advanced analysis methods like prediction models, data mining, text analysis, page rank, collaborative filtering have been a research hotspot and widely used. Data mining is used to find out predicative, practical and novel results[11], while relation schema and cloud tags visualize the results. However, the analytical results should also match the business process and get merged in knowledge base if they pass the assessment.

2.2.3 Phase3 Knowledge Management.

Qualified knowledge is expressed by concepts, relations, properties, axioms, rules and instance[12] of ontology. On the basis of ontology, semantic web not only makes knowledge service-oriented but also offers an excellent management system which improves knowledge inference efficiency and eliminates Once updated. Above

all, the goal of knowledge management is to form the KMS and publish knowledge service.

3 Ontology modeling and application

3.1 Ontology Modeling Example

Results from big data analysis in BDA KMS should be formalized by ontology and enriched with semantic web. A knowledge management system is finally developed to render knowledge services.

A case of converter steel making process is presented in this paper. Ontology classes are classified into seven groups: environmental factors, product, power, facility, ingredient, energy and process schedule as Fig 3 shows. isA represents a relationship between class and instance. isPartOf represents a parent-child relationship among classes. For example, limestone is one instance of slag making material, while slag-making material is a subclass of non-metal materials. Lime quality is related to humidity, which is an instance of environmental factor. The ingredients of limestone are CaO, MgO and SiO2, etc. K-means cluster analysis was applied in 500 crude ore blending samples to distinguish the better plans form the invalid. One of the better plans is expressed in
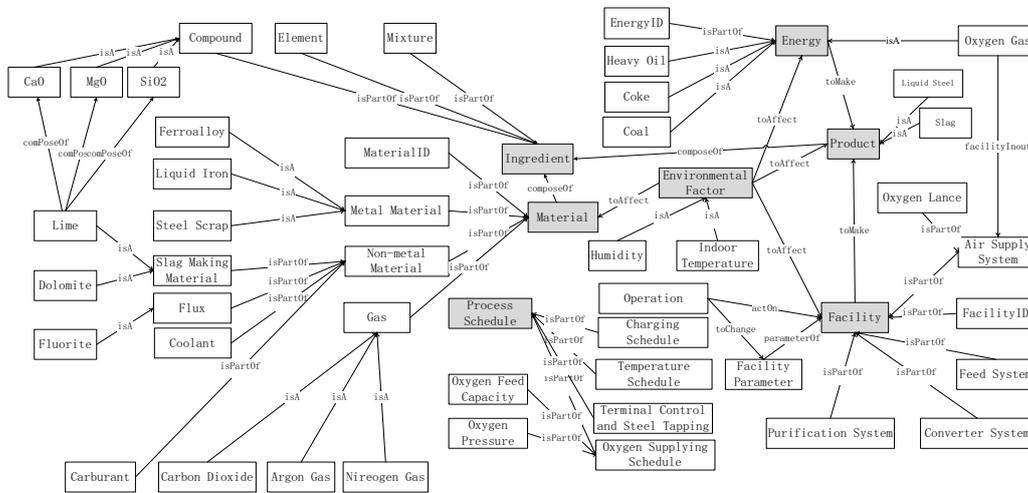

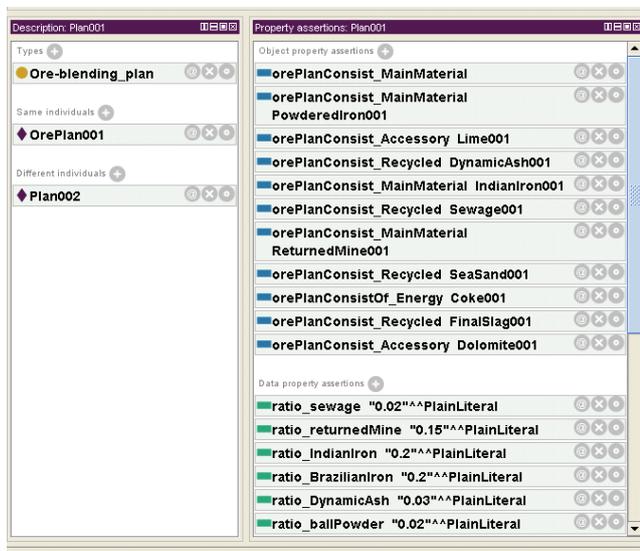**Figure 3.** Ontology structure of converter steel making process

Figure 4. Analysis result formalization in Protégé

Protégé as Fig 4 shows.

3.2 Knowledge Inference

SWRL describes the correlations of distributed and heterogeneous information so that the interoperability is facilitated[13]. Principle of SWRL inference is shown in Fig 5 SWRL should meet various levels of

ontology granularity (top-level ontology, domain ontology, task ontology and application ontology), and it describes rules matching these ontologies.
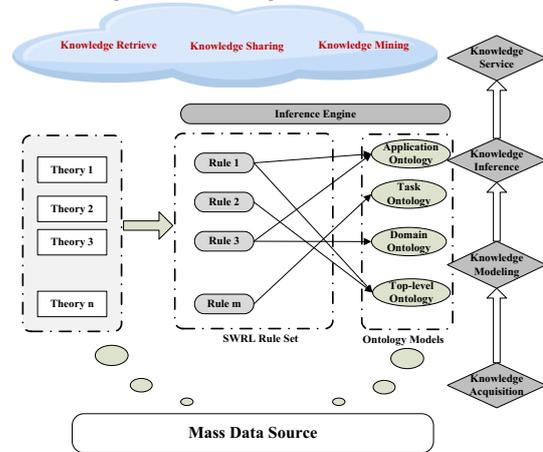

Figure 5. SWRL rule inference frame

Table 2. Parameter and its description

Parameter	Paramete description	Value
IW_{Si}	Si content of iron water	0.85%
IW_p	Pi content of iron water	0.2%
LM_{CaO}	CaO content of lime	89%

LM_{SiO2}	SiO2 content of lime	1.2%
LM_{MgO}	MgO content of lime	3.0%
DM_{CaO}	CaO content of dolomite	32%
DM_{MgO}	MgO content of dolomite	21%
DM_{SiO2}	SiO2 content of dolomite	1.3%
SG_{MgO}	MgO content of final slag	6%
µ	Basicity of final slag	3.5
LMneed	Demand of lime	To solve

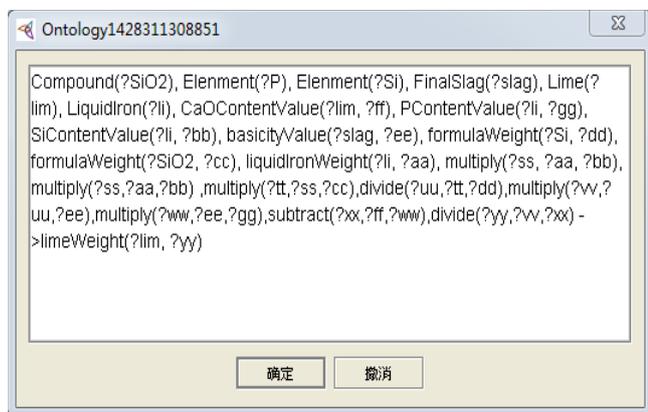|	
Metalnonmetal(?li,?lim)	?lim mixes with ?lim
liquidIronWeight(?li,?aa)	The weight of ?li is ?aa
SiContentValue(?li,?bb)	The Si content of ?li is ?bb
formulaWeight(?Si,?dd)	The formula weight of ?Si? is ?dd
limeWeight(?lim,?yy)	The quantity demand of ?lim is ?yy

Figure 6. SWRL rule editor

Figure 7. Inference results in knowledge management system

Then, the inference engine(Jess, Pellet, Racer or Jena, etc.) combines
ontologies with rules, reasoning out related knowledge. This paper takes a converter steel-making as an example to illustrate the SWRL rules establishment. The formulation is:

$$LMneed = \frac{IWweight \times IW_{Si} \times M_{SiO_2} / M_{Si} \times µ}{LM_{CaO} - µ \times IW_P} \quad (1)$$

Parameters specification in this formula is explained in Table 2. Fig 6 shows the edited rule and the atoms in the edited SWRL atoms are explained as Table 3 shows. Similarly, lime addition, dolomite demand and dolomite addition can be calculated in other formulas. Pellet, as inference, is adopted and initializes the following instance shown in Table 3. This paper also develops a prototype knowledge base system in B/S pattern to the inference interface and the results of this case are shown in Fig 7.

Table 3. SWRL atom description

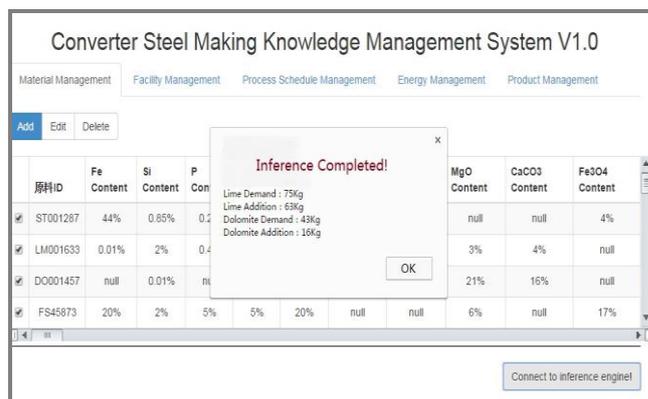| Atom	Description
Compound(?Si)	?Si is one of compound instance
Element(?P)	?P is one of element instance
LiquidIron(?li)	?li is one of liquid iron instance
Lime(?lim)	?lim is one of lime instance
FinalSlag(?slag)	?slag is one of slag instance

4 Summary

This paper proposes a big data analysis adopting the data driven, intelligent analysis, service publication, dynamic update patterns to transform data to knowledge: integrative service. This paper focuses on ontology modeling in BDAKMS and SWRL rule establishment. Finally, this paper gives a converter steel making case for validity verification. However, before this structure is utilized, three aspects are in need of improvements: 1. More accurate and effective analytical algorithms should be developed catering for steel industries' changing needs. 2. Big data analysis and should be integrated in the whole BDAKMS, but how to translate the user analysis to big data analysis objective is an issue to be considered.

References

1. Anhui Tan, Jinjin Li, Guoping Lin, et al. Fast approach to knowledge acquisition in covering information systems using matrix operations, Knowledge-Based Systems. **79**, 90-98 (2015)
2. Shah Jahan Miah, Don Kerr, John Gammack, et al. A generic design environment for the rural industry knowledge acquisition. Knowledge-Based Systems. **21**, 892-899 (2008)
3. Mara Abel, Luís A. Lima Silva, John A. Campbell, et al. Knowledge acquisition and interpretation problem-solving methods for visual expertise: S study of petroleum-reservoir evaluation, Journal of Petroleum Science and Engineering. **47**, 51-69 (2015)
4. A. Giovannini, A. Aubry, H. Panetto, et al. Ontology-based system for supporting manufacturing sustainability. Annual Reviews in Control. **36**, 309-317 (2012)

5. Qiu Ming, Chen Gang, Dong Jinxiang. Ontology based adaptive product lifecycle management system.Computer Intergrated Manufacturing Systems. **11**, 826- (2005)

6. S.Lemaignan, A.Siadat, J. Y Dantan, et al. MASON: a proposal for an ontology of manufacturing domain. IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications. 195-200 (2006)

7. H.Panetto, M.Dassisti, A.Tursi. ONTO-PDM: product-driven ontology for product data management interoperability within manufacturing process environment. Advanced Engineering Informatics. **2**, 334-348 (2012)

8. Grit Denker, Lalana Kagal,Tim Finin. Security in the semantic web using OWL.Information Security Technical Report. **1**, 51-54(2005)

9. Feng Weizhi,Cen Honghong.The realization of grade Ⅱ computer controlling software in making steel, Metallurgical Collections. **4**, 35-36 (2004)

10. J.Fan, H.Liu, Statistical analysis of big data and opportunities with big data. Proceeding of the VLDB Endowment. **12**, 2032-2033 (2012)

11. Muhammad Usman.Rusesel Pears, A.C.M.Fong. A data mining approach to knowledge discovery from multidimensional cube structures. Knowledge Based Systems. **40** ,39-41 (2013)

12. Berners L T,Hendler J,Lassila O. The Semantic Web. Scientific American. **284**,34-43. (2001)

13. Lezcano, Leonardo, Sicilia, Miguel Angel, et al. Integrating reasoning and clinical archetypes using OWL ontologies and SWRL rules. Journal of Biomedical Informatics. **2**, 343-350 (2010)