TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.2, February 2014, pp. 1462 ~ 1470
DOI: http://dx.doi.org/10.11591/telkomnika.v12i2.3860

DSRM: An Ontology Driven Domain Scientific Data Retrieval Model

Jianghua Li*[1,3] Qing Cao[2]
[1]School of Information and Engineering, Jiangxi University of Science and Technology, Ganzhou, China
[2]College of Sciences, Henan Agricultural University, Zhengzhou, China
[3]National Center for Materials Service Safety, University of Science and Technology Beijing, China
*Corresponding author, e-mail: jxsimil@hotmail.com

Abstract

With the development of information technology, a large number of domain scientific data have been accumulated with the characteristics of distribution and heterogeneity. It has important significance to acquire exact scientific data from multiple data sources for cooperative research. The existing data integration and information retrieval techniques cannot solve the problems of data semantic heterogeneity and retrieval inaccuracy very well. In this paper, an ontology driven domain scientific data retrieval model is proposed, which uses domain ontology to describe user query and queried data. User query is posed on domain ontology schema. Data retrieval for distributed and heterogeneous data sources is realized through constructing mapping relations between them and domain ontology schema. We developed a prototype for material scientific data, and the experimental results show that the proposed model is effective. Our model can also provide some means of use for reference to other domain scientific data retrieval.

Keywords: *data retrieval, domain scientific data, mapping relation, ontology, retrieval model*

1. Introduction

In the past two decades, with the development of science and information technology, in some domains, a large number of domain scientific data have been accumulated, such as materials, geography, biology and other fields. Scientific data is thought to be the basic data and information produced during scientific practice and the data processing for different demands [1]. It plays an important role for domain scientific research and industrial activities. For different purposes, some different data sources have been established by different research groups or in different periods, they may be stored in different environments and used in different applications. At present, the main characteristics of domain scientific data can be generalized as follows. 1) Distribution: domain scientific data is always autonomous and distributed in geographical position. 2) Heterogeneity: the data sources are heterogeneous in the aspects of system, syntax, structure and semantics. 3) Diversity: Various types of data sources have the different description standards, such as RDBMS, XML, RDF, XLS, and so on. 4) Variability: the data is increasing and updating continuously. In addition, the relations among domain scientific data are always complex so that it is difficult to associate them together.

Scientific data is the source of scientific and technological innovation.Scientific data integration and sharing is the foundation of the data-intensive applications. With the suggestion and development of e-Science, sharing the global, interdisciplinary large-scale scientific research resources and collaborative work are becoming possible. This would greatly promote the exchanges, cooperation and the development of scientific research. Scientists always cannot acquire enough information from single data source. On the other hand, different with Web document retrieval, the scientific data required in scientific research activities should be accurate. So it is a research hotspot for domain scientists to retrieve desired and exact data from the distributed and heterogeneous data sources.

In this paper, we aimed to provide scientists a domain scientific data retrieval method to acquire desired and exact data from distributed and heterogeneous data sources. We proposed an ontology driven domain scientific data retrieval model (DSRM). The rest of the paper is organized as follows. Section 2 is dedicated to the related work. In Section 3 elaborates the

Received July 8, 2013; Revised August 26, 2013; Accepted September 27, 2013

proposed domain scientific data retrieval model and its working principle. Section 4 introduces the experimental prototype to demonstrate how users can retrieve what they want from our system. Finally, Section 5 gives some conclusions and future work.

2. Related work

2.1. Data Integration Technology

Data integration technology is the key to solve the problem of data access for distributed and heterogeneous data sources, which can be traditionally divided into three types: federated database technology [2], mediator/wrapper-based data integration technology [3] and data warehouse technology [4].

Under the premise of maintaining the autonomy of local member data source, federated database technology only integrates part of the distributed and heterogeneous data sources and provides sharing and transparent access to them. It requires the global view of all accessible data source stored in the federated database server, including the configuration information, the structure of the integrated data table, indexes and so on. In federated database technology data dictionary is constructed as the unified global view, which is mapped to local data source and used to rewrite query according to the type of local data source. In a word, federated database technology adopts data dictionary to resolve the syntactic heterogeneity among member data sources. Federal database server communicates with the local data sources through a wrapper, and the wrapper corresponds to a type of data source. After accepting data access instructions from an information integration server, the wrapper converts them into those supported by the local data sources, and submits them to the corresponding data sources server to execute. Finally, the results are returned to the information integration server for further processing.

Mediator/wrapper-based data integration technology needs to construct a global data view, on which users pose queries without necessity to know the location of each data source and access methods. But it dosen't store real data before integrating data. Mediator processes a user query, and sends it to each wrapper and composes the returned results from each wrapper together. A wrapper corresponds to a data source to eliminate heterogeneity. It provides the mediator a uniform query access interface and other functions as what is done in the federated database technology.

Data warehouse technology needs to extract, transform and load each data source to a global database, which is inopportune or inappropriate. It is not in the scope of the paper.

The above data integration technologies can retrieve data from distributed and heterogeneous data sources, and solve the problem of data heterogeneity on the system, syntax and structure level. But they cannot resolve the problem of data heterogeneity on the semantic level, which still exists in a variety of data integration systems.

2.2. Information Retrieval Technology

The keyword-based traditional search engines such as Google, Baidu cannot meet user requirements for information retrieval because of the following disadvantages. 1) The user interface does not support semantic extension, so that it can not accurately describe user requirements. 2) Do not support semantic query to RDF and OWL files. 3) In these search engines, vector space model (VSM) is used to represent document by extracting its characteristic words, but this can not be an accurate description of a document, and what is more important is that the knowledge in the document is neglected. 4) The query results are inaccurate. There are always a large number of returns unrelated to user requirements. 5) The ranking method cannot satisfy users. 6) Do not support logical reasoning query.

The accuracy of information retrieval depends primarily on the accurate description and logical representation of queries and query objects, followed by the matching relationships between them. In order to solve the existing problems in traditional information retrieval technology and improve its performance, researchers have proposed a variety of new methods and techniques from the angles of syntax and semantics. Generally these methods are divided into four types: vocabulary-based query extension, semantics-based query extension, ontology mapping and Semantic Web search.

The vocabulary is a set of phrases or words, including various relationships between them, such as hypernyms, hyponyms, alternative words, associative-words and so on. There are two kinds of commonly used vocabularies as follows. 1) Pervasive vocabulary based on the words, such as WordNet [5], HowNet [6] and Cilin [7], the vocabulary contains synonyms or antonyms with semantic relations. 2) Information retrieval-oriented vocabulary based on phrase, such as the Library of Congress Subject Headings [8], MeSH Medical Subject Headings [9], etc. The method uses hypernyms, hyponyms, alternatives and associative-words to extend the keywords input by user and the characteristic words extracted from documents, then computes the included angle cosine between them to query and rank.

With the appearance of Semantic Web, it provides an opportunity to improve the accuracy for information retrieval. The researchers put their hopes on ontology. Ontology is a formal explicit specification of shared conceptualization [10, 11]. Its aim is to provide a common understanding of domain knowledge. Using the common and shared vocabulary and taxonomy defined in ontology to describe the domain resources can exactly express their semantics. Some researchers introduced ontology mapping into the field of information retrieval. They mapped respectively the keywords input by user and the characteristic words extracted from document to the corresponding ontology and replaced them with the concepts defined in the ontology [12-16]. Then the traditional information retrieval technology is used to query and rank.

In order to better represent the semantics of user query and document, some scholars combine the idea of query extension using ontology to research semantics-based query extension [17-22]. Firstly, they mapped user query and document into concepts in ontology. Then they used the hypernym concepts, hyponym concepts, context concepts and their attributes, and synonyms of the corresponding concepts to extend the query and document. Finally the correlation between them is calculated, which is the basis of querying and ranking.

The above methods can add semantic information to user queries and document, and improve the precision to a certain degree. But there are still the following problems. 1) A concept is usually described with different attributes and taxonomic hierarchy in different domain. Supposed that there are several domain ontologies, how to decide which ontology the concept should be mapped to? Even though a correct mapping is established and a concept is selected to represent a user query keyword, it is still hard to express the real semantics of a user query if there is no other relevant metadata, for example attributes. In addition, only mapping the characteristic words in a document to concepts will lost some knowledge hidden in the document, for example instance and attribute value, which users are usually interested in. 2) In the way of an extension query, the hypernyms, hyponyms, synonyms and their attributes of a concept are used to extend it, which will add some semantic information to the user query and document in a certain degree. However, this extension works at the cost of broadening the retrieval scope. Meantime, it is hard to determine extended hierarchies because of the less of suitable extension principles. This will lead to many keywords unrelated to user requirement in the queries and document, so what can be foreseen is that there are a lot of unrelated documents in retrieval results and the precision ratio is low. 3) The keywords input by user are not usually domain concepts, but the instances and attribute value, which will cause other problems: For example, How to map the keywords to concepts defined in domain ontology and ensure that the mapping is correct? On the other hand, the characteristic words extracted from document are also mapped to concepts or attributes, which results in few matches or wrong matches between original user query and documents. So, the query results could not be exact.

In view of the above problems, we propose an ontology driven domain scientific data retrieval model (DSRM).

3. The Proposed Domain Scientific Data Retrieval Model

The goals of DSRM are as follows. 1) Establish the application on the Internet and integrate the accessible data sources on the Internet. 2) Application is open. New domains and new data sources are permitted to join the application. 3) The flexible user interface can provide domain retrieval metadata model to exactly express user semantic retrieval requirements, and realize accurate retrieval for multiple data sources. 4) Replace VSM with the semantic model for document data, which can be better to represent the knowledge in documents.

To realize the above goals, we adopt the following design principles. 1) Information retrieval is established on the basis of domain. At first, the application guides user to select the retrieval domain from user interface. Then user interface generates domain metadata retrieval model to format user query according to the selected domain. The metadata will add semantics to user query and make it exactly express what the user wants. 2) In view of the mature technology of relational database and query efficiency, in order to improve the response speed to the user query requests, only structured data is the data source type that can be directly accessed by the query engine. That means unstructured data and semi-structured data will be preprocessed and transformed into structured data, such as Web documents, XML files and RDF database. 3) The selected domain determines which data sources are accessible. To improve the query accuracy, the same metadata model that guides user to complete query will be used to describe and store the accessible data resources. Meantime, the mapping relations between them will be stored.

3.1. The Model Definition

Definition 1: Domain Ontology Schema (DOS) is a five-tuple, DOS=<C, H, P, A, R>. C is the set of domain concepts. H is the set of taxonomic relations among concepts, and P is the set of non-taxonomic relations. A is the set of domain axioms. R is the set of Horn rules described by Semantic Web Rule Language (SWRL), which is used by the inference engine supporting rule reasoning to find the hidden knowledge in semantic query.

Definition 2: Domain Metadata Model (DMD) is a quaternion, DMD=<C, H, P, R>. And $C_{DMD} \subsetneq C_{DOS}$, $H_{DMD} \subsetneq H_{DOS}$, $P_{DMD} \subsetneq P_{DOS}$, $R_{DMD} \subsetneq R_{DOS}$. Namely the elements of DMD are the proper subset of the corresponding elements of DOS, because not all the concepts and relations (attributes) are suitable to be metadata for query. Domain metadata and user input together constitute the semantic query.

Definition 3: Relational Data Source (RDS) is the set of directly accessible relational databases on the Internet. Transformed Relational Data Source (TRDS) is the set of the relational databases which are established after non-structured data and semi-structured data are preprocessed according to the corresponding DMD. TRDS is used to store the extracted instances, attributes and their values from unstructured and semi-structured data sources. Each transformed relational database corresponds to a kind of unstructured data or semi-structured data source.

Definition 4: Ontology Driven Domain Scientific Data Retrieval Model (DSRM) is a six-tuple, DSRM=<UI, QA, SRKR, DS, DCAP, RET>. The architecture of DSRM is shown in Figure 1. UI is the component of user interface. QA is the component of query analyzer. DS is the set of the accessible data sources. SRKR is the semantic retrieval knowledge repository, which is the core component of DSRM. DCAP is the component of data collecting and preprocessing. RET is the component of knowledge registration and other management tools. The function and principle of every component will be introduced as follows respectively.

The main objectives of the component UI are as folloes: 1) Provide user a semantic query interface constructed dynamically, which guides user to input query keywords and generates a global query G_q. Fist of all, a user needs to select a query domain, which is used to get the domain metadata model DMD from SRKR. According to the DMD, a semantic user query interface is constructed. 2) Receive the query results from QA and present them in the UI according to user ranking requirements.

The component QA is composed of a inference engine, a query decomposer and a result composer. The inference engine provides user a reasoning rules-based semantic query, which is used to find the hidden knowledge in data source. The query decomposer receives the global query G_q and the query domain from UI, and gets accessible data sources and their access methods from SRKR, for example, the configuration information of a data source, web service and so on. If there are multiple accessible data sources, the query decomposer firstly acquires the mapping rules between ontology and every relational schema from O-RS repository in SRKR. And then it decomposes the global user query G_q into different sub-queries against the type of each data source and invokes the corresponding web service. The results composer composes the results of all the sub-queries retrieved from the different data sources together and sends them to UI.

The component SRKR is composed of DOS repository, DMD repository, O-RS repository and UDDI (Universal Description, Discovery and Integration) server, which is the core of DSRM and serves other components. The O-RS repository stores the mapping relations between domain ontology schema and accessible relational data sources in DS so that the decomposer can decompose the global query into different sub-queries against the different data sources. The UDDI server provides web service register and stores the mapping relations on domain—web service—data source, where the mapping relations between domain and web service may be one-to-many.

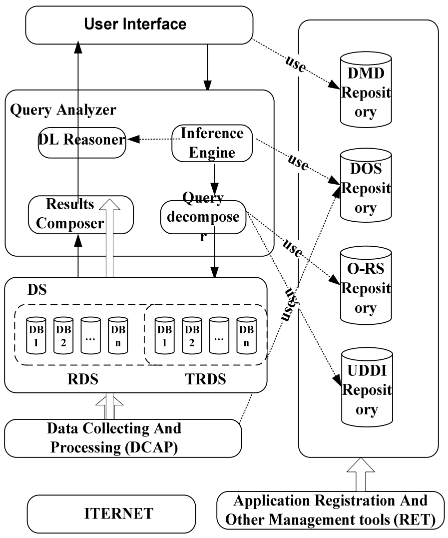


Figure 1. DSRM Architecture

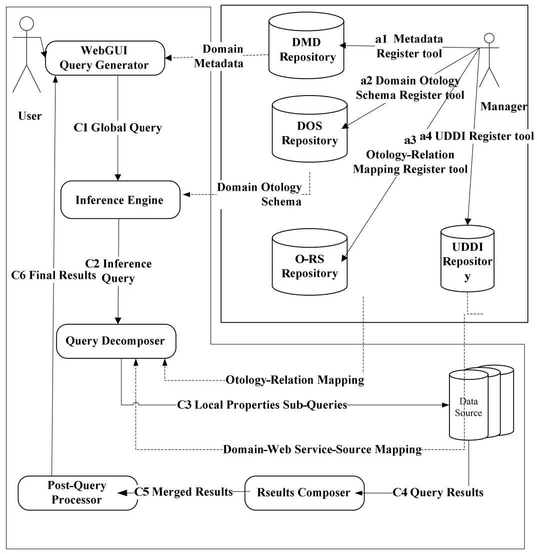


Figure 2. Application Registration and Query Process

The component DS provides application accessible data resources, DS= {RDS, TRDS}.

The component DCAP is composed of a data collecting module and a data preprocessing module. According to the given domain keywords, the data collection module collects periodically the related documents by using a web crawler from the Internet and classifies them in accordance with file type. The data preprocessing module takes charge of processing the collected documents and stores their semantic information into TRDS. The following is the preprocessing algorithm.

1) Classify the document according to its domain.

2) If the domain that the documents belong to appears for the first time, create relational schema in TRDS according to the corresponding DOS.

3) If the file type of a document is html, xml, PDF, or doc, go to 4), otherwise if the type is RDF or OWL, go to 5).

4) Extract instances, attributes and their values from the document according to DOS by using natural language processing algorithm, and describe them with the format of RDF.

5) Store the instances and attribute values into relational schema created according to DOS in TRDS, as well as the mapping relations between extracted knowledge and the document.

The component RET is provided for system administrators, which has two main functions. One is application registration, namely to add the new accessible retrieval domains, data sources and their corresponding access methods, new DOS, DMD, mapping rules and so on to SRKR. The other is to maintain the knowledge in SRKR, for example, modifying domain ontology schema or mapping rules etc.

3.2. Application Registration and Query Process

The application registration and query process of DSRM are shown as the step labels in Figure 2.

1. Application registration process

Application registration is designed for system administrators, which can be divided into two cases. One is to register the newly increased retrieval domain to SRKR; the other is to register the newly increased accessible data resources to SRKR. The following steps are used to describe the application registration process.

Step a1: If the registered domain is not in the DOS repository, register the domain ontology schema of the newly increased domain to DOS repository, otherwise go to Step a3.

Step a2: Register domain metadata to DMD repository according to the domain ontology schema.

Step a3: If the accessible data source is RDS, construct the mapping relations between the relation schema of the RDS and the corresponding domain ontology schema in O-RS repository, otherwise go to Step a4. This is because the relation schemas of the original relational data sources on the Internet are heterogeneous with the domain ontology schema apparently. But user query is posed on the domain ontology schema. In order to access distributed and heterogeneous data, the mapping relations between them must be constructed and stored in O-RS repository. Compared to TRDS, it is not necessary, because the schema of TRDS itself is transformed from the domain ontology schema, which ensures the consistency between them.

Step a4: Register the web service for newly increased data source to UDDI server, as well as the mapping relations on domain—web service—data source.

2. Query process

Step c1: First of all, a user selects a query domain from the application interface, then the application generates the user query interface according to DMD. Next, the user selects the query attributes and inputs the corresponding values, and then the user query described in SPARQL is submitted to QA.

Step c2: If there are reasoning rules in the domain ontology schema, the inference engine supporting description logic reasoning will be activated to get hidden query semantics in user query.

Step c3: Query Decomposer gets the data sources and their types that user query needs to access. If the accessible data source is RDS, then the mapping relations between DOS and the RDS will be got from the O-RS repository. If there are multiple accessible data sources, Query Decomposer will decompose the global user query G_q and translate them into different sub-queries Lq_1, Lq_2 … Lq_n against the type of corresponding data source. The corresponding web service WS_1, WS_2 … WS_n in the UDDI server will be invoked to execute the queries.

Step c4: Every data source returns the results CS_1, CS_2 … CS_n in accordance with the local grammar to query results composer.

Step c5: Query results composer merges the results CS_1, CS_2 … CS_n with the same grammar together.

Step c6: The results are ranked by UI and presented to user.

4. The Experimental Case

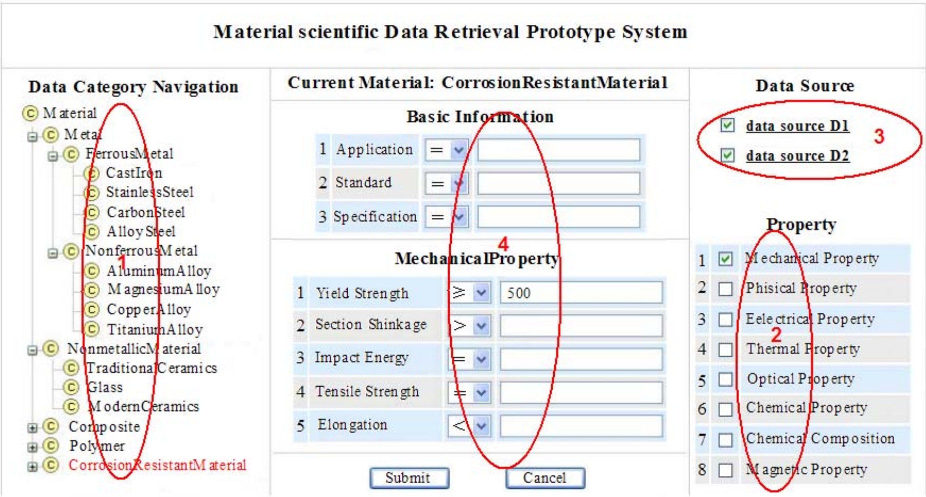
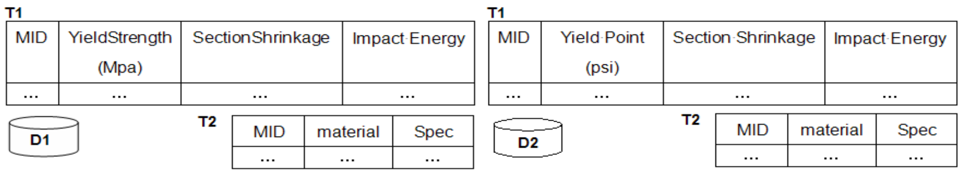


Figure 3. Material Domain Scientific Data Retrieval Model Prototype System

In order to test the proposed domain scientific data retrieval model, we developed a prototype for materials scientific data shown as Figure 3. The prototype is developed using Java and MySql, in which the web-based GUI is implemented using JSF. Jena API is invoked to manipulate ontology and Jess is used as inference engine. The accessible data sources are distributed in University of Science and Technology Beijing, Central South University, Northwestern Polytechnical University, Chinese Academy of Sciences, including the types of relational database, XML files, MatML, RDF, OWL and Web documents.

A user can exactly retrieve material scientific data that he wants. Firstly, he can select the domain concept what he wants from the concept list shown in the area 1 in Figure 3. Then the corresponding attributes (i.e. metadata) and accessible data sources are shown in the area 2 and 3 in Figure 3 respectively. The user can select his concerned attributes and the data sources he needs to access. Finally, the selected attributes are listed in the area 4, and then the user can input the attribute values in the textbox and submits the query to query analyzer. Because the global query G_q is posed on ontology schema, described in SPARQL. Next, G_q will be translated into different sub-queries against the type of the selected data source, and the corresponding web services will be invoked to retrieve data. Finally, the results returned from different data sources will be merged and presented to user.



(a)

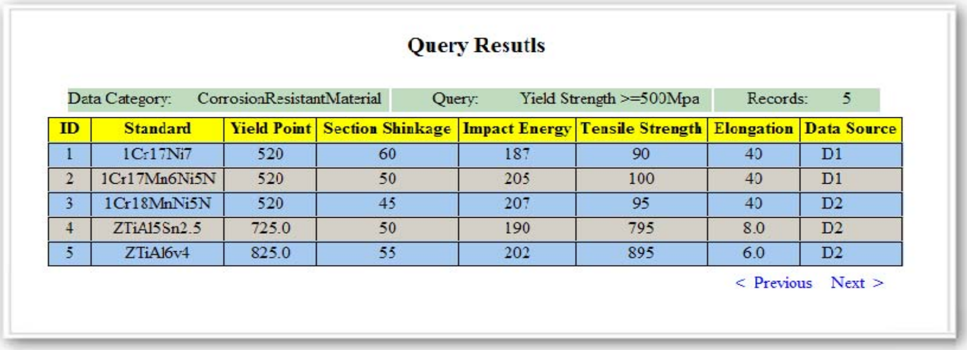
DOS	hasYieldStrength	hasSectionShinkage	hasImpactEnergy	hasTensileStrength	hasElongation
D1	YieldStrength	SectionShinkage	ImpactEnergy	TensileStrength	Elongation
D2	Yield Point	Section Shrinkage	Impact Energy	Tensile Strength	Elongation

(b)

Figure 4. The Structures of Data Sources and the Their Mappings with DOS

Assumed that a user wants the corrosion resistant materials whose yield strength is greater than 500Mpa and some of their attributes: section shinkage, impact energy and so on. And there are two accssible data sources D1 and D2 shown in Figure 4(a). Simply, the mapping relations shown in Figure 4(b) between the accessible data sources and DMD have been strored in O-RS repository. The rule 1MPA=145psi has been stored in DOS too. Accoding to DMD, the global query G_q can be generated in SPARQL. Next, According to the mappings in O-RS and the rule 1Mpa=145psi, the global G_q is decomposed sub-queries Lq_1 and Lq_2 against D1 and D2 respectively by Query Decomposer. G_q, Lq_1 and Lq_2 are shown as follows. Then the parameters in Lq_1 and Lq_2 are passed the corresponding web service WS_1 and WS_2 in UDDI server respectively. Finally, the query results from D1 and D2 are merged by Results Composer and presented to the user shown in Figure 5.

```
Gq :SELECT ?material ?hasYieldStrength ?hasSectionShinkage ?hasImpactEnergy
  WHERE { ?material hasProperty ?MechProp.
       ? MechProp rdf:type MechanicalProperty.
       ? MechProp hasYieldStrength ?X.
       FILTER (?X >500).}
Lq1: Select material,YieldStrength, SectionShrinkage, ImpactEnergy, TensileStrengh
  From T1, T2
  Where T1.MID=T2.MID and T2.YieldStrength>500;
Lq2:Select material Yield Point, Section Shinkage, Impact Energy, Tensile Strength
  From T1, T2
  Where T1.MID=T2.MID and T2.Yield Point>500*145;
```



Query Resutls

Data Category: CorrosionResistantMaterial | Query: Yield Strength >=500Mpa | Records: 5

ID	Standard	Yield Point	Section Shinkage	Impact Energy	Tensile Strength	Elongation	Data Source
1	1Cr17Ni7	520	60	187	90	40	D1
2	1Cr17Mn6Ni5N	520	50	205	100	40	D1
3	1Cr18MnNi5N	520	45	207	95	40	D2
4	ZTiAl5Sn2.5	725.0	50	190	795	8.0	D2
5	ZTiAl6v4	825.0	55	202	895	6.0	D2

< Previous Next >

Figure 5. The Query Results

5. Conclusions and Future Work

To improve the precision and efficiency for information retrieval, combining the Semantic Web technology, we researched the semantic query in the paper, and proposed a domain scientific data retrieval model, as well as its architecture, basic functions and working principle. The proposed model can retrieve data from multiple distributed and heterogeneous data sources, which also is used to realize data semantic integration. We thought that, under the premise of the same domain ontology schema, using domain metadata to describe user query and queried data sources could add semantic information to them. This will exactly express user query demands and the logical view of queried data, and improve the matching accuracy between them.

It is a complex and polyfunctional system to retrieve data from multiple distributed and heterogeneous data sources, involving several techniques. And these techniques are not yet mature, and need for further research and exploration. For example, increasing a new retrieval domain in the model needs to obtain the domain ontology schema first, which will need a large number of ontologies. How to acquire domain ontologies efficiently by automatic or semi-automatic methods? How to exactly extract the semantics and hidden knowledge from document? And how to transformed ontology schema into relational schema and realize relational storage of large-scale RDF data? In addition, besides document and relational database, there are a mass of pictures, video and other unstructured data. How to annotate them automatically using the domain ontology and put them into the retrieval model? These are issues that require further study, and we will also continue to improve the proposed domain scientific data retrieval model.

Acknowledgements

The work in this paper is supported by the Key Science-Technology Plan of the National ‘Twelfth Five-Year-Plan’ of China under Grant No.2011BAK08B04. And the work is also partly supported by the Scientific Research Found of Jiangxi Provincial Education Department of China under Grant No.GJJ13411, and the Jiangxi Provincial Ministry of Education Research Fund under Grant No.12770.

References

[1] J Gray, DT Liu, MN Santisteban, A Szalay, DJ DeWitt, G Heber. Scientific data management in the coming decade. *SIGMOD Record*. 2005; 34(4): 34-41.

[2] Gonçalves, Eder Mateus Nunes. Specifying knowledge in cognitive multiagent systems using a class of hierarchical Petri nets. *Journal of Software*. 2012; 7(11): 2405-2414.

[3] Hidalgo Justo, Pan Alberto, Álvarez Manuel, Guerrero Jaime. *Efficiently updating cost repository values for query optimization on web data sources in a mediator/wrapper environment*. Proceedings of the 6th International Conference on Next Generation Information Technologies and Systems. 2006; 1-12.

[4] Ni Zhiwei, Guo Junfeng, Wang Li, Gao Yazhuo. An efficient method for improving query efficiency in data warehouse. *Journal of Software*. 2011; 6(5): 857-865.

[5] Huang Guangjun, Wang Shuili, Zhang Xiaoguo. Query expansion based on associated semantic space. *Journal of Computers*. 2011; 6(2): 172-177.
[6] Pei-Ying Zhang. A HowNet-Based Semantic Relatedness Kernel for Text Classification. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(4): 1909-1915.
[7] Liu Yuling, Zhu Yinghong, Xin Guojiang. Chinese text watermarking method based on TongYiCi Cilin. *International Journal of Digital Content Technology and its Applications*. 2012; 6(19): 465-473.
[8] Lu Caimei, Park Jung-Ran, Hu Xiaohua. User tags versus expert-assigned subject terms: A comparison of Library Thing tags and Library of Congress Subject Headings. *Journal of Information Science*. 2010; 36(6): 763-779.
[9] Leydesdorff Loet, Rotolo Daniele, Rafols Ismael. Bibliometric perspectives on medical innovation using the medical subject Headings of PubMed. *Journal of the American Society for Information Science and Technology*. 2012; 63(11): 2239-2253.
[10] Gruber TR. Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*. 1993; 5(2): 199-220.
[11] Lirong Qiu. Website Resource Monitoring Platform Supporting Tibetan and Uyghur Language based on Semantics. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(8): 4766-4773.
[12] Li Wen, Chen Yewang, Peng Xin, Zhao Wenyun. Effective Term-Concept Mapping Method Based on Ontology. *Computer Science*. 2010; 10(10): 138-142.
[13] Chen Yewang, Li Haibo, Yu Jinshan. A Semantic Retrieval Model Based on Agricultural Field Ontology. *Journal of Huaqiao University*. 2012; 33(1): 27- 32.
[14] Mohamed Hazber. A Mapping Rule from Relational Databases to Semantic Web. Hua Zhong Normal University. 2011.
[15] Wang Zhihua, Zhao Wei. Research on semantic web retrieval model based on ontology and key technologies. *Computer Engineering and Design*. 2011; 32(1): 145-148.
[16] Guo Guangjun, Zhou Xianguang, Chen Daiwu. Network Education Resources Management and Retrieval System Based On Ontology and Lucene. *Computer Applications and Software*. 2012; 29(7): 115-119.
[17] Yuan Hui, Li Yanxiang. The Knowledge Base Based on Ontology Semantic WEB Expanded Search Method. *Electronic Test*. 2013; 5: 69-71.
[18] Du Jinyang, Yi He, Yang Chun. Retrieval strategy based on semantic expansion of keywords. *Journal of Computer Applications*. 2009; 29(6): 1575-1577.
[19] Yang Qinglin, Li Taoshen, Nong Jian. Semantic query expansion based on domain ontology knowledge base. *Computer engineering and Design*. 2011; 32(11): 3853-3856.
[20] Hu Chuanlie, Fu Yunqing, Zhong Mingyang. Semantic Query Expansion Based on Domain Ontology. *Computer Systems & Applications*. 2012; 21(7): 83-89.
[21] Li Weijiang, Zhao Tiejun, Wang Xiangang. Context-Sensitive Query Expansion. *Journal of Computer Research and Development*. 2010; 47(2): 300-304.
[22] Tian Xuan, Li Dongmei. Survey on Contextual Information Retrieval. *Computer Science*. 2011; 38(9): 18-24.