Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Engineering 15 (2011) 3220 – 3226
Procedia Engineering
www.elsevier.com/locate/procedia

Advanced in Control Engineering and Information Science

The Research of Ontology-based Scientist Resources Service Platform

Gan JianHou[a,b,c], Wen Bin[c], Li ShuRen[b], Xu ShiTing[b], Yan BaoPing[b,a]*

[a]*College of Metallurgy and Energy Engineering, Kunming University of Science and Technology, Yunnan, Kumming, 650093;*
[b]*Computer Network Information Center, Chinese Academy of Sciences, Beijing, 100190;*
[c]*College of Information, Yunnan Normal University, Yunnan, Kumming, 650092*

Abstract

By the combination of scientist resource areas, the Ontology-based service platform for scientists, namely, resource SRSP has been studied in this paper, in which the scientist resources MIS systems has been designed, the Ontology-based scientist resources domain has been constructed, and the automatic generation technology for scientist information resource website and the personalized recommendation technology has been proposed.

Selection and/or peer-review under responsibility of [CEIS 2011]
Keywords: Scientist resources; Ontology; management information system; Website automatic generation

1. Introduction

Through the construction of Ontology-based service platform of scientist resources, it improves, updates and manages the existing data to construct the domain Ontology of scientist resource. By mining the data of scientist resource information, it does the research and development of Ontology-based comprehensive integration platform for the scientific research information customer service to enhance

* Corresponding author. Tel.: +86-871-594-1322; fax: +86-871-594-1311.
E-mail address: kmganjh@yahoo.com.cn

doi:10.1016/j.proeng.2011.08.605

the customer service level of the scientific research information and offer aid in decision making for the enhancement of the scientific research information in Chinese Academy of Sciences.

The Ontology-based service platform of scientist resources mainly includes: ① The service platform of scientists' information resource data with management, improving and updating. ② The establishment of scientist information resources model. ③ The construction and knowledge reasoning of s semantic web-based scientist resource domain Ontology. ④ The research of automatic generation technology for scientist information resource website and the personalized recommendation technology for scientist resource. ⑤ The research of the integrated service platform for Ontology-based scientists' resources.

The development tools of Ontology-based scientist resource service platform (SRSP) mainly consists of Protégé, RacePro, Jena, Lucene, and etc. Protégé is used to create and maintenance of the domain Ontology; RacerPro is used to do one-time test of the domain Ontology, class hierarchy relationship reasoning and equivalence class reasoning. Jena is regarded as the development package for the analysis and operation of Ontology files. Lucene is regarded as the development package for full-text search. Mysql is for the database management system .

2. The Main Functions of Scientist Resource Service Platform

2.1. Acquisition of Scientist Resource Information

Getting the information is the basic work for the purpose. The channels includes: ① Making use of all resource date in relation to scientists in ARP system in Chinese Academy of Sciences and maintaining consistency of information with it. ②Collecting and updating scientist resource information with all kinds of manual methods. ③ Acquiring the relevant scientific research information by searching the researchers on the internet and by relationship mining technology. ④Using the literature search and data mining technology to get the scientists' resources information. ⑤ Establish scientists' resources data updating mechanism with the relevant institutions and research institute of the CAS and keep data updating timely and correctly by system innovation[1].

2.2. The Construction of Semantic Web-based Scientist Resource Domain Ontology

The application of semantic web technology makes the network resources understood by machines, which makes up for the deficiencies of the existing technologies. This paper is to research the mapping problem among Ontology languages in the Semantic Web (such as the mapping among XML, RDFS and OWL), and put forward some solutions for the difficulties in mapping such as multi-valued element mapping and recursive element mapping; In addition, it builds the interfaces with other Ontology library and import the Ontology which has been established into the system[2].

Based on the research of the traditional information search techniques and Ontology technology, the semantic & Lucene-based retrieval system model is proposed. The model adds Ontology in the traditional full-text retrieval system to provide semantic support. Tt also presents the calculation method of the similarity between concepts and the related method of query expansion and applies the algorithm to the retrieval system of scientist resource[3].

It introduces the Ontology technology to the scientist resource service platform and does the Ontology construction for scientist resource areas to achieve the organization, management and service of scientist resource. The concept of field scientists, resource hierarchy tree is illustrated in Fig.1.

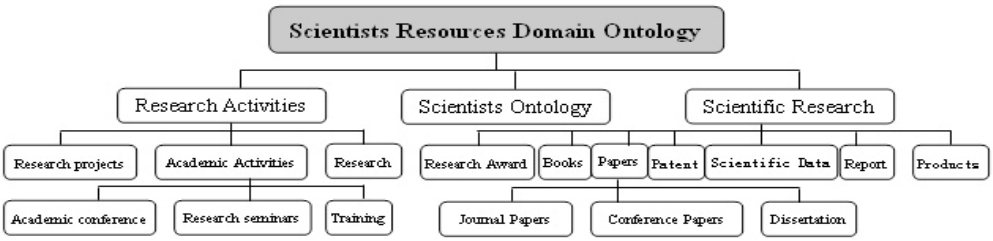

Fig.1. The concept of field scientists resource hierarchy tree

2.3. The Problems in Knowledge Reasoning of Scientist Resource Service

The common problems in knowledge reasoning of scientist resource services include:

2.3.1. Class (Concept)-- Instance Relationship Inference

Give knowledge base X, C is one class (concept) in K, while i is an individual object in K, the following instance relationship can be reasoned:

(1) to judge whether an individual is an instance of C: for example, if academician Sun is a class, judge whether scientist Lee belongs to this team;

(2) to judge all C instance in knowledge base K: for example, if the paper published by academician Sun is a class, judge all the papers of academician Sun;

(3) to judge what class' instance i belongs to in the knowledge base: for example, scientist Lee is what research teams' instance;

(4) to judge the relationship between two instances or judge the specific relationship with an special instance: for example, to judge the relationship of scientist Li and Yang, like student, teacher, co-worker, friend, subordinate, superior and other relations.

2.3.2. Class (Concept) Relationship Reasoning

Given class C and D, to judge the relationship between them (subclass, member, part, etc.), such as the relationship between academician Sun's research team and academician Lu's research teams relationship.

2.3.3. Reasoning in the Systematic Structure of Class

Given class C, return to all or all relevant super classes of C in the knowledge base K; or all or all related subclasses of C in the knowledge base K. For example, give scientist Lee the subject class, judge all or all relevant super classes or subclasses in the knowledge base K.

2.3.4. SAT Reasoning of Class

Given a class C, judge if C is satiable in the knowledge base K. For example, if given scientist Lee the subject class, judge if this subject class is satiable in the knowledge base K.

2.3.5. Attribute-based Reasoning

There are similar reasoning between attribute and class (instance), including: attribute – instance relationship, attribute contains, attribute systematic structure and attribute SAT, and so on.

For in-depth knowledge which can summarize or completely cover the entire field of scientist resource service, it requires to clarify the relationship between data field in scientist resource relational database Ontology reasoning. Through analysis, Table 1 summarizes the application of data field in scientist resource relational database in knowledge reasoning[4].

Table 1. The application of data field in scientist resource relational database in knowledge reasoning

Data Field	Subjection Data Tables	The Application in Knowledge Reasoning
Master, Doctor,t or Postdoctoral	Basic information data tables	Reasoning the university-oriented relationship between scientists
Work Unit	Basic information data tables	Reasoning the colleague relationship between scientists
Duties & Work Unit	Basic information data tables	Reasoning the subjection relationship between scientists
Name & Work Unit	Basic information data tables	Considering the name and work unit of scientists to prevent same-name situation happening in reasoning
Start-stop time & Resume	resume data tables	Reasoning the development history of scientists
Start-stop time & Educated information	school work information tables	Reasoning the academic course of scientists
Project member	Project data tables	Reasoning the team relationship between scientists research projects
Signed order、amount of work & Contribution rate	project member data tables	Reasoning the members relationship between scientists
Project subject code	project data tables	Reasoning all scientists in the field
Project subject code,item category 、Appraisala chievement & project-based Awards level, etc.	project data tables Research Award Tables ……	Reasoning the influence of all scientists in the field
scholarly topic, time, subsidize project	Academic Activities data tables	Reasoning the overall condition of scientists academic activities in project process
Corresponding author and other author or not	Dissertation data tables	Reasoning the members relationship between scientists
Patent owners	Patents data tables	Reasoning the members relationship between scientists
Cover news, reporting agencies level，Publishing location， Report number，frequency, etc.	Cover news、report data tables	Reasoning the influence of scientists
Related data fields	All data tables	Automatic generation scientist personal information documentation，can be divided into Summarized and detailed
Related data fields	All data tables	Automatic generation scientist information resource website SRWeb

2.4. The Automatic Generation Technology for Ontology-based Scientist Information Resource Website

Based on database query technology and Ontology reasoning technology, it builds Ontology-based automatic generation model for scientist information resource website, mainly including: personalized topic map learning navigation and the acquisition and learning of the intelligent knowledge. the acquisition and learning of the intelligent knowledge can be realized by intelligent knowledge acquisition in massive amounts of data and the system can automatically assemble the retrieved contents into a single interface like a portal site. This interface is not a linear list; it can greatly improve the efficiency of users.

Main problems to solve: in the use of Ontology technology's control and guidance on the conversion process, many semantic problems may occur. Through the use of human-computer interaction technology, dialogue consultation window and the corresponding domain expert knowledge base provide the necessary help to solve it.

2.5. The Personalized Recommendation Technology for Scientist Resource

By in-depth study of the being personalized information service technology, the data mining, natural language understanding and expert system technology be used in personalized information service technology. Through research for the structure of personalized information service system, can be recommended to the user information relevant to the content and not just a word that contains the information reflects the resources of individual scientists, intelligence information services.

Scientists resource personalized recommendations process includes the following three steps: (1) to collect user information and Web resources, information about users, including user browsing behavior, user logs, etc.; (2) information modeling and classification; (3) Information Optimization processing and recommended information. The methods of analysis for gathering information include content-based filtering, collaborative filtering, rule-based filtering and Web usage mining and so on.

3. The Overall Structure of Ontology-based Scientist Resource Service Platform

Scientist resources service platform is based on database technology, data mining technology, machine learning, Ontology technology, knowledge engineering and other advanced technologies for organization, management and service applications of scientist resources.

Scientist resources service platform includes five layers: the infrastructure layer, the data center layer, the service platform layer, the user application layer and an extending service layer.

The infrastructure layer includes network facilities, servers, storage, super computers and Internet, it can provide basic hardware support for the basic data communications, computing, storage and management of scientist information service comprehensive integration platform.

The data center layer includes database of scientific resources, domain Ontology knowledge base, scientist resource document base and other resources database. It's the core of the whole system for covering all the data of scientist resource. The data is mainly accessed by scientist information documentation and literature resources.

The service platform can support database applications, data mining, document mining, Ontology acquisition and mapping, log mining and decision.

The user application layer includes user information retrieval, scientist resource site SRWeb, scientist personal affairs and personalized recommendation; technology support includes personalized recommendation technology, visualization technology, intelligent search, knowledge reasoning and services, the automatic generation technology for Website.

The extending service layer includes e-mail systems, video desktop systems, VOIP systems, online communication systems and collaboration platforms. Based on the scientists resources service, the extending service layer is applied to the existing, mature application systems.

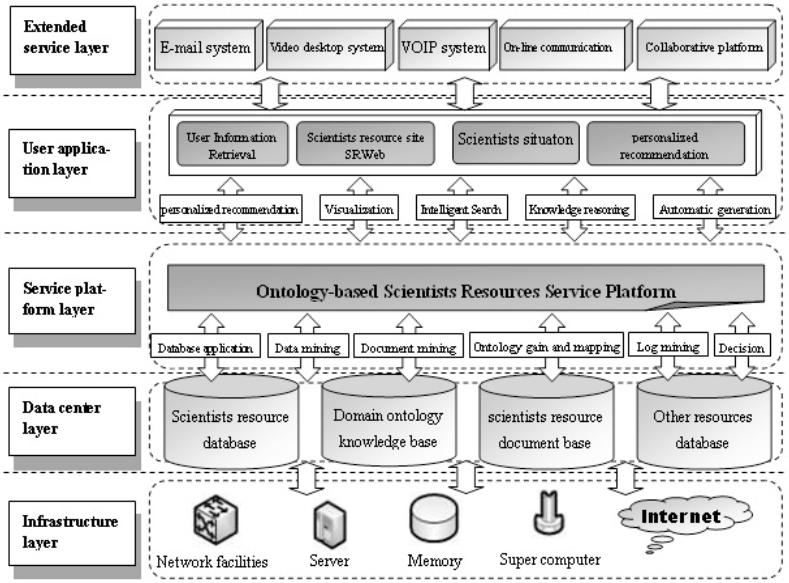

Fig. 2. The overall structure of scientist resources service platform.

4. Conclusions

Scientist resources information is the important foundation for promoting the informatization of science research. Relying on scientists resources of Chinese Academy of Sciences and through the semi-automatic, automatic, continuous data accumulation of scientist resources, an authoritative service platform can be built. This work will have an important theoretical value and a wide range of applications. The future work will be focused on studying and developing the Ontology-based scientist resources platform.

Acknowledgements

The work is supported by the Talent Planning Project in Yunnan Province (2009CI062), CNIC Creative Foundation Project of Chinese Academy of Sciences, the General Project of Social Development in Yunnan Province (2011), the National High Technology Research and Development Program of China ("863"Program) of China (2008AA04A105), the General Projects of Social Development in Yunnan Province (2008CD113), and the Foundation Project of Education Office in Yunnan Province (2010Z089).

References

[1] Gan JianHou & Yuan LinYun. Study on Knowledge Base Construction for the Network Teaching Resource Based on Semantic Web[J]. Applied Mechanics and Materials, Vols. 16-19 (2009) pp 758-763.

[2] Gan JianHou. Semantic Web and its application-based on Ontology, description logic, semantic networks[M]. Yunnan Science and Technology Press, 2007.11.

[3] Duan SouJian & Yang ChaoFeng & Gan JianHou. Domain-based semantic similarity and the relevance of the concept of comprehensive quantitative study[J]. New Technology of Library and Information Service, 2009(11):40~43.

[4] GAN Jian-hou & WEN Bin & Zhang Shu.Research of the knowledge reasoning based on extensional Description Logics ALC-plus[J]. IEEE,CiSE2009.