

Research on the Application and Development of Next-Generation Semantic Web in Cloud Environment

Jiaolian Fan

National Experimental Teaching Demonstration Center
Beijing University of Posts and Telecommunications
Beijing, China
E-mail: fanjl@bupt.edu.cn

Abstract—In the current cloud environment, the development of next-generation Semantic Web application indicates that whether it is personal or group oriented, the development of next-generation Semantic Web application in cloud environment will bring about great convenience. In China, there are 1,090 colleges and 796 postgraduate training units, which provide a mammoth foreign language learning community. Foreign language is an important basic course for each college student. No matter the undergraduate, graduate or doctoral students, they are all required to learn English or other languages. Due to the foreign language teaching has the characteristics of wide involvement, great influence and heavy tasks, how to use Internet to improve its personalization and efficiency has always been the focus of attention. Research on the application and development of next-generation Semantic Web in cloud environment is the key to solving this problem.

Keywords—cloud computing; Semantic Web; foreign language teaching

I. INTRODUCTION

Apply the construction, evolution and feedback mechanism of dynamic learner community to Semantic Web in cloud environment, research results can quickly adapt to the needs of social and technological development. It not only uses Semantic Web-based knowledge, resources organization and integration strategies to improve the way of education resources integration in a timely manner so that the resources restricted by physical campus, geographic region or traditional Internet logic area can better adapt to ubiquitous network or ubiquitous learning needs, but uses the Semantic Web-based dynamic learner community model to optimize teaching behavior and develop efficient teaching modes through purposeful and conscious community construction, evolution and feedback guide, ultimately realizing that ubiquitous information resources and knowledge provide service to ubiquitous learning needs.

II. SIGNIFICANCE OF RESEARCH ON THE APPLICATION AND DEVELOPMENT OF NEXT-GENERATION SEMANTIC WEB IN CLOUD ENVIRONMENT

Although there is no denial that next-generation Semantic Web in cloud environment is the development trend of future Internet, research progress in last decade is

not as rapid as the web development in 1991. The reasons are chiefly as follows.

A. *Excessively theoretical and technical specification biased research*

In earlier studies of Semantic Web, W3C and universities focused on the development of technical specifications. They raised seven programs of the Semantic Web and were committed to developing corresponding standard in each layer. Semantic Web technical specification set is very complex, making it impossible for some laboratory methods to be effectively promoted to the actual Web application. For example, the standard emphasizes logical layer must have accurate and complete reasoning capability, but the fact is that many Semantic Web applications only use metadata and it should be multi-technology integrated rather than logical reasoning based.

B. *Too much demand on user ability, resulting in a small audience and lack of replicability*

Ontology modeling and annotation are core technologies of Semantic Web. Existing Semantic Web technical specifications are usually assumed that users can perform ontology modeling. For instance, making a distinction between concepts, relationships and examples, or utilization of ontology editor or some annotation tools to provide metadata, which causes the increasing number of companies engaged in research and development of Semantic Web related products. However, these companies are overly focused on some specific industries, such as bio-pharmacy, medicine, publishing, and some large enterprises and research institutions. The experience acquired from particular target users and special cases is difficult to spread to the majority of ordinary Web users.

C. *Ignore the heterogeneity, low-quality and multi-source features of real data*

Originated from the specification in laboratory, Semantic Web is a Web science rather than a Web project. It puts too much emphasis on experimental data sources, quality, etc., which is compliant. However, under the real Web application environment, heterogeneous and mixed noise data is very common. For example, failure to meet the labeling specifications or inaccurate labeling problems often occurs. The existing Semantic Web technologies don't take

heterogeneity, low-quality and multi-source features of real data into consideration neither from the angle of technical specification nor the new technology convergence.

D. Insufficient use of integrated technologies as for the realization of Semantic Web

Semantic Web realization should integrate multiple technologies, such as information retrieval, machine learning, natural language processing, database, etc. However, as for the technology realization of current Semantic Web, natural language processing and the database technology are only applied in ontology model. Besides, Semantic Web is overdependent on rules and logical reasoning, and it emphasizes the utilization of ontology mapping approach to deal with heterogeneous sources of data sets. However, the integrated utilization of statistical machine learning and information retrieval is rare. So no matter semantic annotation or application stage, there exist great restraint and it requires a lot of manual intervention.

In fact, the applied research on the Semantic Web, until now, still uses a top-down planning technology development model, deviated from the normal mode of demand-driven technological development. No matter Yahoo and Google in Web 1.0 era, or Facebook in Web 2.0 era, both followed the normal mode, which made history on the Internet. So Semantic Web in Web 3.0 era will also be no exception. Hence research on the application and development mode of next-generation Semantic Web in cloud environment is conducive to the promotion of technological innovation and expansion of applications scope.

Meanwhile, research found that in China there are three main issues in foreign language teaching system. (1) Neglect of efficient organization and management of teaching resources (heterogeneous media resources, such as tutorial text, image and audio). Although foreign language learning resources are very rich, resources are isolated and closed with each other, which makes it hard for the resources to be shared. Learning resources, education information and advanced personalization features tend to become isolated information islands, resulting in a state of iterative development over a long period of time. (2) At each stage of self-learning, students remain highly dependent on teachers. The widespread lack of self-planning, self-management and a clear learning objective gives rise to the inefficiency of network learning and poor results. (3) Formative assessment lacks reasonable, effective and detailed data, which leads to the difficulty in discovering students' individual difference and achieving the targeted personalized instruction.

Therefore, this article is intended to take personalized integration application with heterogeneous and non-uniform resources in Foreign Language Teaching system as the background. Through building the testbed in Semantic Web environment, it realizes the standardized labeling of resources including foreign language tutorial, courseware, vocabulary, sentence, text, image, listening audio, etc. It also introduces new teaching models and emerging applications, and constructs the effective multi-technology integration framework. At the time of bottom-up researching on Semantic Web technologies in college foreign language

teaching resources, teaching mode and collaborative optimization of self-organized learning behavior, improvement, application, development and effective promotion mode of Semantic Web related technologies should also be explored.

III. APPLICATION VALUE OF NEXT-GENERATION SEMANTIC WEB IN CLOUD ENVIRONMENT

The application value of next-generation Semantic Web in cloud environment is mainly reflected as follows.

A. Large serving group

It is well known that college foreign language teaching is an important component of foreign language education in China. According to the statistics of Ministry of Education, at the end of 2010, in China, there are 1,090 colleges and 796 postgraduate training units, which provide a mammoth foreign language learning community. Foreign language is an important basic course for each college student. No matter the undergraduate, graduate or doctoral students, they are all required to learn English or other languages. Due to the foreign language teaching has the characteristics of wide involvement, great influence and heavy tasks, how to use Internet to improve the personalization and efficiency of foreign language education has always been the focus of attention. Research on the application and development of next-generation Semantic Web in cloud environment is the key to solving this problem.

B. Meet the needs of integration, personalization, collaborative learning of foreign language teaching resources in college

With efforts to expand the scale of education, improve the quality of education, enhance operating efficiency, establish lifelong education system, and implement major education strategic measures, the state has begun to vigorously develop modern distance education. The principle is known as "network construction is the foundation; development is the core; teaching application is the goal; management service is the insurance". In order to provide urgent, high-quality online teaching resources in a series, the state has implemented a series of related projects, such as "New Century Network Course Construction Project" of the Department of Higher Education, "Teachers' Continuing Education Network Curriculum Development of Major Projects" of the Ministry of Normal Education, "Modern Distance Moderate Vocational Education and Adult Education Resources Construction Projects" of the Ministry of Education Vocational Education and Adult Education Department, "Construction Projects of Online Examination System of the National Online Education Exam" of the Department of Higher Education, construction of courses and college English teaching system in "Education Quality and Reform Project" and so on.

However, the target of effective integration of teaching resources and network construction cannot be considered as confined to resources publication and information dissemination on the Internet. This traditional resources integration model is a one-way broadcast mode from

"Teaching Resources" to "Student". In this mode, student must rely on traditional searching engines inefficiently (because traditional ones use a simple operating mode of "PUSH" which does not distinguish between real needs of target users) and manually extract and organize useful information. It does not reflect the need of effective integration of resources extraction that is student-centered, student-initiated and student-personalized. Therefore, foreign language teaching resources are constructed redundantly with low-level quality. It lacks reusability, interoperability and organization. Although the proposal of learning technical standard makes the reuse and management of learning resources possible, in most cases these activities are confined to a particular class of learning management system itself, and there is still a lack of interoperability mechanisms between different learning management systems. Meanwhile, user information, learning strategies and learning activities, which are more educational, are still difficult to be migrated and reused. The concept conversion from "resources provision" to "service provision" is urgent, while the development and application of Semantic Web can effectively provide support for the personalized foreign language teaching, autonomous and collaborative learning.

Moreover, besides the foreign language teaching resources, the same problem also exists in other types of integration of information resources on the Internet. So Semantic Web technology can be used in the annotation and integration of foreign language teaching resources. And for teachers and students, it helps to provide more convenient access to information optimization, promote the formation of learning communities of interest, and can be extended to other equally multi-sourced, heterogeneously integrated resources to provide constructive solutions and recommendations.

C. To achieve integration of new technologies

There is a variety of college foreign language education resources, such as text, image, audio, video, etc. So there is a need for an effective way to label and integrate the resources which is decentralized, loose, highly diverse and dynamic. First, labeling of diverse media resources can enhance the accuracy of information retrieval. More importantly, through introduction of image content retrieval and cross-media retrieval technologies, and combination with ontology construction technology, it can facilitate the release of information resources and self-learning behavior. Second, self-organization and collaborative learning is the trend of teaching in the future. When college teachers and students are acquiring resources, Semantic Web agent can find the resources preferred group on the basis of social attributes quickly by use of statistical machine learning techniques, which can help teachers and students to create multiple communities of interest, and also make it easier to implement communication, sharing and collaboration. Furthermore, in addition to providing efficient resources sharing environment under the virtual community model, the colleges, as gathering areas of mobile phone users, also provide good experimental environment to carry out the research on

foreign language learning upon location-based services (LBS) and knowledge sharing.

D. Research on the application of Semantic Web to the management of college foreign language teaching resources and teaching activities for the first time

As mentioned before, current researches on the Semantic Web application mainly focus on bioinformatics, pharmaceutical, publishing and semantic retrieval areas for news mining. In the field of education, only a few foreign universities are engaged in applied researches with a narrow audience. The University of Wollongong in Australia makes use of Semantic Web technology to optimize the website resources of special class subjects. The University of Murcia, Spain applies the ontology design to analysis students' evaluation of classroom teaching. Open University of the Netherlands applies semantic technology to online learning with certain of guidance. Until now, no research has been found in application of Semantic Web to the promotion of college foreign language teaching.

Meanwhile, the project team has a stronger research foundation in foreign language resources annotation and Semantic Web technology related fields. Now more than 200,000 English entries (about 25 million words) in term of bilingual labeling and more than 650,000 English sentence pairs (about 50 million words) of the parallel corpus have been completed, which has laid a solid foundation for foreign language teaching resources annotation. Furthermore, polytechnic universities gather a large crowd of high-tech followers. They can provide new ideas, technologies and the most direct test feedback to Semantic Web research, which ensures the smooth progress in applied research of Semantic Web.

E. To further promote multi-technology integrated application, promotion and development of Semantic Web under the background of foreign language teaching demand in colleges

Through integration of Semantic Web technology and college foreign language teaching, Semantic Web can promote the optimization of resources organization, task-based teaching activities, personalization and self-organization of learning activities, and establish a more effective bridge of communication between teachers and students. On the other hand, the typical diverse needs of college foreign language teaching environment can directly promote the research on integrated framework of multi-technology in Semantic Web. College foreign language teaching has a wide audience, which provides strong support for bottom-up application, promotion and development research of Semantic Web.

IV. THE TARGET OF NEXT-GENERATION SEMANTIC WEB IN CLOUD ENVIRONMENT

The target of next-generation Semantic Web in cloud environment is mainly focused on the application of Semantic Web and supplemented by innovation of related technologies, which not only fills the research gap, but also prevents duplicated building. Semantic Web has realized

application in some areas, such as database literature retrieval, library management system, etc. But it is the first attempt to apply the Semantic Web technology to foreign language education in China and even international scope. Therefore, based on the foreign language teachers' teaching practice and research findings from Beijing University of Posts and Telecommunications, University of Shanghai International Studies and other first-tier universities, author has got a comprehensive view of weakness and improvement requirements of current college English teaching and learning system. Take advantage of Semantic Web technology, establishment of a personalized college foreign language learning service system is the final objective.

Applied goals:

- To establish resources semantic annotation system and metadata management platform with flexible and open features. Both of them should satisfy the application requirements of foreign language teaching
- To establish the ontology model of view discourse genre in accordance with attitudes and rank difference
- To construct network-based self-learning framework of college foreign language which is intelligent, open, personalized, and to achieve dynamic teaching evaluation system with immediate effect feedback
- To build the heterogeneous resources organization and management platform which integrate Semantic Web with traditional Internet, and to establish an effective access control system
- To establish a task-based ontology model with flexible expansion capability in foreign language teaching field
- To form a bottom-up multi-technology integration and promotion mode of Semantic Web application

Academic goals:

- To build foreign language education resources annotation mode with semantic descriptions and expression capability
- To propose technological discourse ontology model and genre analysis model based on machine learning techniques
- To establish foreign language learners' experimental teaching model based on Semantic Web
- To propose automatic semantic annotation and detection methods based on machine learning techniques, aiming to improve the efficiency of resources retrieval
- To explore multi-feature extraction and binding of cross-media resources, and the vectorization model of Semantic Web knowledge
- To establish cross-media information retrieval model based on Semantic Web
- To explore new ideas of trust management and secure data mining in Semantic Web environment

V. CONCLUSION

In summary, although in the last decade, the research on Semantic Web has made considerable achievement and has established a relatively complete theoretical system and technical specifications. Successful application in some areas

also indicates that it is bound to be the supporting technology in Web 3.0 era. At this time, it is of great theoretical and application value to investigate the applications, development and integration framework with new technologies of Semantic Web with bottom-up and application-driven model in the new situation and new environmental, and promote a specific application based on its core technology. However, specific application research was often restricted to the research on data integration, specific areas or events. The various studies are unrelated, and progress on service integration is slow. In fact, Semantic Web application actually requires more integrated use of technologies, but until now the current studies are still a top-down research model and lack new technologies and demands consideration, which makes it difficult to promote the application of research results. Therefore, since Semantic Web has received more attention and has been approved across International scope to be the technical basis of Web 3.0 era, it is necessary and urgent to implement a bottom-up research approach to foreign language education in colleges based on the application of resources integration, and support research on application and development of Semantic Web.

ACKNOWLEDGMENT

"Twelve Five" Quality Engineering, National Experimental Teaching Demonstration Center

REFERENCES

- [1] M. P. Cuellar, M. Delgado and M. C. Pegalajar. Improving learning management through semantic web and social networks in e-learning environments [J]. *Expert Systems with Applications*, 2011, 38(4).
- [2] Q. A. Memon and S. A. Khoja. Semantic Web Approach to Academic Program Assessment [J]. *International Journal of Engineering Education*, 2009, 25(5): 1020-1028.
- [3] E. Liliana Tovar and M. Esther Vidal. A language to express reactivity in the Semantic Web [J]. *Revista Tecnica De La Facultad De Ingenieria Universidad Del Zulia*, 2010, 33(1): 68-76.
- [4] E. A. Gavrilova. Implementation features a university data portal using the Semantic Web technology [J]. *Programming and Computer Software*, 2011, 37(1): 48-55.
- [5] Zheng Jiamao. Experiment on University Teaching PCT Issues Analysis [J]. *Laboratory Research and Exploration*, 2007, (12).
- [6] Department of Higher Education. College English Curriculum Requirements [M]. Beijing: Foreign Language Teaching and Research Press, 2007.
- [7] Wang Shou-ren. Universities College Foreign Language Education Development Report [M]. Shanghai: Shanghai Foreign Language Education Press, 2008.
- [8] Jia Guodong. On college English teaching reform achievements, problems and measures [J]. *Foreign sector*, 2006: 42 a 47.
- [9] Jia Guodong. Classroom and web-based college English teaching mode [J]. *Chinese university education*, 2009, (6).
- [10] Chen JianLin. College English Teaching Network Concept and Application Analysis [J]. *Foreign Language Education*, 2004, (6).
- [11] M. Cai, W. Y Zhang and K. Zhang. ManuHub: A Semantic Web System for Ontology-Based Service Management in Distributed Manufacturing Environments [J]. *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans*, 2011, 41(3): 574-582.
- [12] V. Janev and S. Vranes. Applicability assessment of Semantic Web technologies [J]. *Information Processing & Management*, 2011, 47(4).