

Semantics, the Missing Link in Web 2.0

(Position Paper)

Jinlei Jiang

Tsinghua National Laboratory for Information Science and Technology

Tsinghua University, Beijing 100084, P. R. China

jjlei@tsinghua.edu.cn

Abstract. Though 20 years have passed since the birth of CSCW, the original goal of it is not reached as well as people expected. To deliver better collaboration services, new technologies are badly needed. Web 2.0 provides a good basis for wide area collaboration. However, due to the lack of semantics, current Web 2.0 technologies still have difficulties in addressing the need of CSCW. With semantics supplied, various Web 2.0 tools can be integrated effectively to deliver powerful collaboration services. This paper reports our experience in developing a collaborative learning environment based on semantic-enhanced Web 2.0 technology.

Keywords: CSCW, Web 2.0, Semantic Web

1. Introduction

As a milestone in information technology, Computer Supported Cooperative Work (CSCW) or groupware reflects a change in emphasis from using computer to solve problems to using computer to facilitate human interaction [1]. Though 20 years have passed since the birth of CSCW, the original goal of it is not reached as well as people expected [2]. For example, although the first group editor program was developed more than 15 years ago, no commercial running one is seen even till now.

The reasons for such a situation are as follows: 1) cooperation is so complex a social phenomenon that people have not gained enough knowledge and experience with it, and 2) the related technologies are not matured as stated by Mills, K. L. [4]—"CSCW technology generally relies on a big stack of computer and network technology, operating systems and protocols, data formats and user-interface devices", however, "the dissemination of such capabilities, while growing at a rapid pace, is far from ubiquitous, and even where these technologies have penetrated, the systems, protocols, formats, and software are far from homogeneous", and 3) the deployment of CSCW is affected by social factors as also stated by Mills, K. L. [4], "even assuming that the necessary networking and computing technologies achieve complete penetration throughout society, the deployment of CSCW may still be retarded by various administrative and policy decisions".

Today, great changes have taken place both in computer world and in society. On the one hand, people have gained much knowledge and experience on cooperative system design in the past twenty years. What's more important is the increasing requirements on CSCW in real world provide continuous driving force for CSCW research and development. On the other hand, the emerging new technologies such as grid [5] and the so-called Web 2.0 supply a good basis for addressing the problems faced by existing cooperative systems. This paper reports our experience in developing a collaborative learning environment based on Web 2.0 technology.

2. Web 2.0 and CSCW

Web is an ideal infrastructure for collaboration not only because it has penetrated into our society but because it provides ubiquitous yet homogeneous protocols (HTTP), formats (HTML), and software (Browser). However, in the early stage, the Web is only used as a publishing medium and its power in support collaboration is not fully exhibited. As the technology progresses, especially with the emerging of Web 2.0, the power for the Web to support collaboration is opened gradually.

According to Wikipedia [6], Web 2.0 "refers to a perceived second generation of web-based communities and hosted services — such as social-networking sites, wikis and folksonomies — which facilitate collaboration and sharing between users". Though termed with version 2.0, Web 2.0 does not refer to an update to Web technical specifications. Indeed, it is "mostly a social revolution in the use of Web technologies, a paradigm shift from the Web as a publishing medium to a medium of interaction and participation" [3]. In our opinion, Web 2.0 is just the current CSCW paradigm on the Web covering the social aspect of collaboration systems. CSCW and Web 2.0 can benefit from each other.

From CSCW perspective, the Web 2.0 practice not only ships some useful tools such as weblogs, social bookmarking, wikis, RSS, and so on, but provides a new way to study the nature of human behavior and information dissemination, which is the basis for designing more usable and more effective collaborative systems. For example, the back track link provided by weblogs can be used to analyze how certain contents are distributed among different sites as well as interests of people and the version track provided by wikis can be used to study the collaboration process.

From Web 2.0 perspective, on the one hand, CSCW provides the theory foundation (the nature of collaboration, modes of collaboration [1], experience for collaborative system design, etc) for Web 2.0, and on the other hand, it also raises new challenges, thus driving the development of Web 2.0. Indeed, weblogs and wikis only provides support for asynchronous interaction/collaboration whereas RSS only provides a way for Web contents representation that enables information sharing and aggregation and no human collaboration is

taken into account. The ultimate goal of CSCW, in our opinion, is to support pervasive or everyday collaboration. To achieve this goal, different tools and/or protocols should be effectively synthesized to support situated actions — the very nature of human collaboration. This is a non-trivial task and Web 2.0 is obviously unable to fulfill such a great goal.

The main deficiency with We 2.0 is the lack of semantics¹. Thus, difficulties often arise when integrating different (heterogeneous) tools. With the progress in semantic Web, especially with RDF and OWL becoming standards, semantic empowered tools will soon emerge, advancing the realization of CSCW goal stated above. This paper presents our preliminary result on enhancing the semantic aspect of current Web 2.0 tools.

3. A Collaborative Learning Environment based on Web 2.0

3.1 Background

China Central Radio & TV University (CCRTVU) is the largest dedicated distance education institution in the world, which delivers as many as 13,000 multimedia courses to nearly 2,000,000 students through radio, TV, print, audio-visual materials and computer software. At present this distance education system is made up of 1 central TV University (CCRTVU), 44 Provincial TV Universities (PTVUs), more than 690 branch schools at prefecture and city level and 1,600 study centers at county level. The main task of CCRTVU is to determine which course should be delivered, set the corresponding teaching plan and requirements, and administer the other subordinate universities, schools or study centers. Under the direction of CCRTVU, the other units may autonomously recruit students, write teaching materials or courseware, and organize teaching. Figure 1 shows the whole organization structure of CCRTVU system.

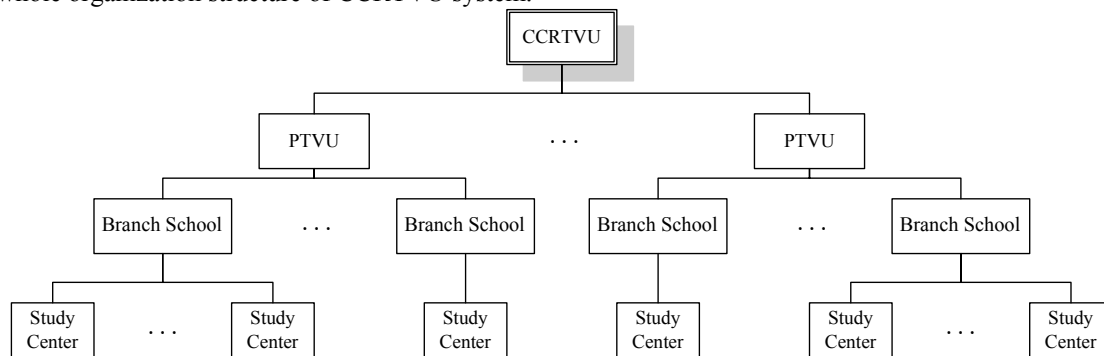


Figure 1. Organization Structure of CCRTVU

In recent years, CCRTVU has begun to adopt Internet and information technology as media for teaching. As the pioneer institution adopting e-learning technologies in China for distance education, CCRTVU has established several e-learning systems, such as CCRTVU online (<http://www.openedu.com.cn>) and CRTVU e-library (<http://library.crtvu.edu.cn/>), to name a few. Today, the requirement for the e-learning systems is increased to support all the education activities online, not only for implementing functions to deliver learning materials or learning administration/management in the early stage, but also for other functions to be used in teaching. For example, there are requirements for providing informal self-assessment to give students ongoing feedback about their progress, and for supporting summative assessment to improve teaching quality. The progress also promotes the requirement to support cooperative learning in these systems, through which student and teacher/tutor community can cooperate in both teaching activities and learning activities.

3.2 Web 2.0 based Solution

With Web 2.0 and semantic technologies, a simple yet powerful solution can be easily obtained. Roughly speaking, RSS can be used as a basis for contents (various courseware, documents, etc) aggregation, and weblogs can be used for students to ask questions and to get answers (from teacher or from peers), and wikis can be used as a tool for students to conduct team work. By providing the missing semantic elements, we can integrate tools of different sites to deliver end users better collaboration experience.

The architecture for the final system is shown in Figure 2. Currently we only connect computer systems of CCRTVU and PTVUs. Each PTVU site provides weblog, wiki, RSS and other necessary services (not presented in the figure) whereas CCRTVU provides services like directory service, metadata service and semantic service. Details of these services are as follows:

- Directory service maintains the information about each PTVU including URL address, names and access points of available services, and so on. The information functions as the basis for different sites to interact.

¹ Some social-networking sites like FOAF have provided support for ontology, however, generally speaking, semantic aspects are lacked in current Web 2.0 tools.

- Metadata service provides functions such as metadata naming, metadata publishing, metadata access, metadata analysis and redundancy cleaning. It forms a solid foundation for transparent data access. Metadata in our system mainly provides description about courseware and documents.
- Semantic service provides a uniform logic view for metadata so as to ensure that users can quickly and accurately find what they need. A semantic service is necessary because different metadata may be used to depict the same thing. For example, student name may be termed “id” or “name” in different systems.

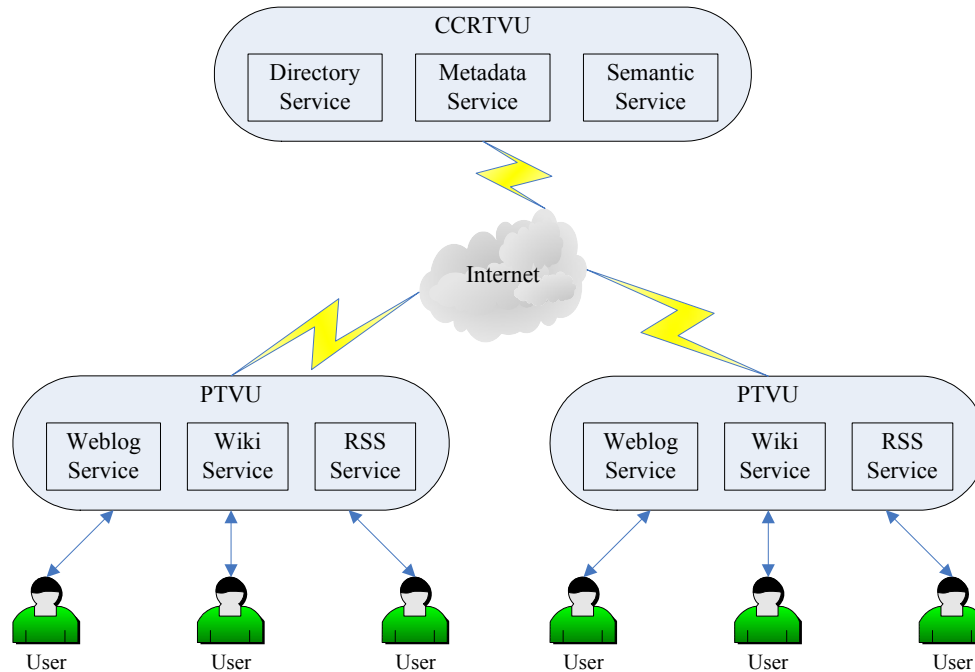


Figure 2. System Architecture for Web 2.0 based Collaborative Learning Solution

Such an architecture provides great freedom to PTVUs — each PTVU is free to choose the implementation of services and maintains its own resources to serve local or global (via other PTVU site) users. Users interact with the system via Web browser. Two types of end users in our system are student and teacher. Interaction supported includes student-to-student, student-to-teacher, and teacher-to-teacher. In the following, we mainly concentrate on asynchronous collaboration.

3.3 An Application Scenario

The advantage of our system can be demonstrated by the following content subscription scenario which is very popular on the Web. Content subscription is also very useful in e-learning environment occurring not only between students and teachers but also between teachers and teachers and students and students. The process that our system deploys to handle users' request on certain contents is as follows:

Firstly, users specify their interests on a certain type of content (e.g., physics courseware) through the interface provided by the local site, say S1.

Secondly, S1 resorts to CCRTVU site to get all the information about the peer (other PTVU) sites and subscribes to the sites related to users' request instead of users. During this step, semantic service is utilized to find out the most related sites.

Thirdly, when the desired content is available, S1 is notified and the content is then aggregated and put into the user's personal space (a RSS reader). To remove duplicated or similar contents, semantic service is also utilized before content aggregation to evaluate the similarity.

Fourthly and the final, users obtain the desired contents from their own space.

From the above statement, we can see that our system brings great convenience to users. Indeed, different sites are separated in current Web 2.0 practice and as a result, users have to login to many sites to specify their interests once and again — a really tedious work. Moreover, duplicated or similar contents are often obtained, wasting the time of users. With semantics supplied, the system can do the most work instead of users, not only sparing users' time, but also enhancing the accuracy of the contents got. Therefore, a better collaboration experience is gained.

4. Conclusion and Future Work

In this paper, we propose a semantic-enhanced Web 2.0 solution for collaborative learning mainly focusing on content sharing. Compared with grid based approaches, our solution has the following advantages:

- It is cost-effective. The heavy work in our system is to define the semantics lacked in existing tools. However, for grid based solution, middlewares are also needed to be developed and to be deployed besides

the semantic definition work. As agreed by most people, it is a challenge to develop reliable grid middlewares and it is not an easy task to maintain grid systems.

- It is user-friendly. This embodies in two aspects. From end users' perspective, their working environment is kept unchanged. From developers' perspective, they don't need to learn more about different sites, which is but necessary to develop and to deploy grid middlewares.

Though our system presents the advantages as stated above, it can still be improved to support collaboration from the following aspects: 1) supply application tools that can transform plain text (in HTML or other formats) into semantic-rich formats (e.g. RDF or OWL) and that can append semantic information to contents in rich media formats, and 2) provide online synchronous collaboration support using existing or emerging standards without changing system architecture.

5. Reference

- [1] Ellis, C., Gibbs, S. and Rein, G.: Groupware: some issues and experiences. *Communication of the ACM*, 1991, 34(2): 38-58
- [2] Jinlei Jiang, Shaohua Zhang, Yushun Li, Meilin Shi: CoFrame: A Framework for CSCW Applications Based on Grid and Web Services. In *Proc the 2005 IEEE International Conference on Web Services*, 570-577
- [3] Lassila, O. and Hendler, J.: Embracing "Web 3.0". *IEEE Internet Computing*, 2007, 11(3): 90-93
- [4] Mills, K. L.: Computer-Supported Cooperative Work Challenges. *Encyclopedia of Library and Information Science*, 2003, 678-684
- [5] Foster, I., Kesselman, C. and Tuecke, S.: The Anatomy of the Grid: Enabling Scalable Virtual Organizations. *International Journal of High Performance Computing Applications*, 2001, 15(3): 200-222
- [6] Wikipedia. Web 2.0. http://en.wikipedia.org/wiki/Web_2