# Consistency Enforcement Using Ontology on Web

Neelam Goyal, Mukesh Kumar, Renu Vig
University Institute of Engineering and Technology, Panjab University, Chandigarh, INDIA

***Abstract*-As the web size grows handling the problem of consistency enforcement becomes very difficult. These days Consistency enforcement is a major challenge in front of World Wide Web. In recent years this field attracts the attention of many researchers. Users are continuously uploading, updating and removing the pages from the Web which makes consistency enforcement a necessity to retrieve relevant and up-to-date information from the Web. In this paper a new approach to handle the problem of consistency enforcement is proposed along with web personalization. In the proposed system ontology is used for managing the data on the web, navigating the web, providing personalized contents to users and for handling the problem of consistency enforcement.**

**Keywords**: Semantic Web, Ontology, Content Management, Consistency Enforcement, Web Personalization

## 1. INTRODUCION

During the past few years the World Wide Web has emerged to become the biggest and most popular way of communication and information dissemination. Every day, the Web grows by roughly a million electronic pages, adding to the hundreds of millions pages already on-line. Because of its rapid growth, it is difficult to access consistent and relevant information from the web. Users often feel disoriented and get lost in that information overload that continues to expand. An ultimate need nowadays is that of predicting the user needs in order to improve the usability and user retention of a Web site.

The Semantic Web is a vision of a new architecture for the World Wide Web. We can think of the Semantic Web as a mesh of information linked up in such a way as to be easily process able by machines. The core idea is to create the Meta data describing the data, which will enable computers to process the meaning of things. The Semantic Web simply means the web of meaning. We can simply say that the Semantic Web is a web of data rather than a web of documents. Semantic Web is an extension of the “existing” Web in which information is given a well-defined meaning. The ultimate goal of the Semantic Web is to transform the web into a medium through which data can be shared, understood and processed by automated tools. The Semantic Web promises a solution in which the web becomes one big knowledge base and everyone has access to it.

Personalization websites have been a great success over the internet in the past years. Sites like Amazon which provide a service that take into accounts the user interests is very helpful to its users and facilitates the site's use.

Web personalization can be defined as any action that adapts the information or services provided by a web site to an individual user. Principal elements of Web personalization include modeling of Web objects (pages, etc.) and subjects (users), categorization of objects and subjects, matching between and across objects and/or subjects, and determination of the set of actions to be recommended for personalization.

In a personalization system, the user model plays very important role because it represents the system's assimilation of the interaction and contains information of the user and the current context that can increase the system's ability to exhibit pragmatically correct behavior and, more generally, to engage in effective communication. The user model could be a static user model which is created and changed by the user and the system uses it for the personalization process. Or it could be a dynamic user model in which the initial user model could be created by the user and the system learns new preferences for that user by keeping track of the user behavior on the site. The system could also predict new user behavior based on previous users with similar preferences. The Model could also be entirely lead and updated by the system. In our case user model is created by user, changed by the user and the personalization process extract from it user preferences. It is also important that the system verifies its deductions were correct by taking feedback from the user (customer satisfaction feedback). The user model should assist the system on how to provide information for the user through a set of rules that describe such preferences along with the user goals and intensions; this is also combined with the preferred communication devices used. Our proposed system will show how user model could be used in the personalization process. Extracting new user preferences is done via Web Mining techniques, which is out of the scope of this work.

The semantic web [3] refers to web site that contains collections of semantically structured information. Ontology is commonly used for this task of structuring


doi:10.4304/jcp.5.10.1520-1526

knowledge, since they represent shared knowledge within a community. Ontology is a description of a domain in terms of entities and relationships that exist between those entities. Ontology provides a precise understanding of the domain, i.e. they provide an unambiguous and consistent interpretation of the domain in all scenarios. Also ontology enables web to share knowledge and provide reasoning support, i.e., verification of formalisms and rules (Checking for constraint violations in knowledge shared through the ontology can be automated). Ontology is used for the structuring, accessing, sharing and presentation of the knowledge.

The process of content management in semantic web is difficult, especially where content is continuously changing that is users continuously uploading, and removing pages from the web. Content management on semantic web basically consists of content editing and content presentation. The process of content editing on semantic web is usually performed collaboratively, normally with few resources (manpower, money). This supposes a great effort to maintain the semantic web and to integrate the information it contains (even if it is using ontology to structure it).

Content presentation is always a hard task, especially in knowledge-intensive web sites where content is continuously updated. To ameliorate the hard task of semantic web management, we need applications that automate these difficult processes (content provision and integration, content presentation, and content access). Furthermore, in these knowledge intensive websites we distinguish between content (knowledge asset) providers and content users. In an ontology-based knowledge websites, the editing of content mainly consists of editing concept instances, that is, inserting, updating and removing instances of ontology concepts. Content presentation in an ontology-based knowledge website mainly consists of showing the ontology concepts and their related instances, presenting the details of ontology instances and their relations with other instances, and allowing the navigation through these relations.

As we know personalization process provide required relevant information to user based on user preferences and these preferences are stored in the form of user model. In ontology-based websites the user preferences are taken from user as instance attribute values. In addition user can change the values of these instance attribute i.e. user preferences goes on changing time to time. So to keep up-to date web site content with user preferences we need a personalization system that provide content management facility along with the basic functionality of personalization. Consistency enforcement is also a major issue in this kind of personalization systems. In these types of systems, we need to check whether the underlying ontology concepts and their instances are consistent or not. Consistency in this type of systems means that there should not be any duplicity between different ontology concept instances. If duplicity exists then there is a possibility that for one particular preference user is getting different recommendations.

In this paper we proposed a web personalization system that makes use of ontologies for consistency enforcement and content editing. We have proposed an easy to use Web interface for inserting, updating and removing instances and their attributes in real-time. In addition, ontology-based user models are utilized to support personalized views.

The rest of the paper is structured as follows: In section 2 we examined the related work. In section 3 overview of the proposed architecture and functionality of the building blocks are given. Section 4 discusses the implementation detail. In section 5 results of the proposed system is evaluated and the results are discussed. Finally conclusion and future work is given in section 6.

## 2. RELATED WORK

Several research studies have focused on web personalization. Since a more detailed description of related research efforts is not possible here. We are giving here overview of some important initiatives in this area.

One of the earliest attempts to take advantage of the information that can be gained through exploring a visitor's navigation through a Web site resulted in Letizia Lieberman [8], a client-site agent that monitors the user's browsing behavior and searches for potentially interesting pages for recommendations. The agent looks ahead at the neighboring pages using a best-first search augmented by heuristics inferring user interest, in as much as they're derived from the user's navigational behavior, and offers suggestions.

Perkowitz and Etzioni et al. [13] were the first to define the notion of adaptive Web sites as sites that semi automatically improve their organization and presentation by learning from visitor access patterns. The system they propose semi automatically modifies a Web site, allowing only nondestructive transformations. Therefore, nothing is deleted or altered; instead, new index pages containing collections of links to related but currently unlinked pages are added to the Web site. The more recent work is presented in [14].

One of the most popular systems from the early days of Web usage mining is Web Watcher Joachims et al. [7]. The idea is to create a tour guide agent that provides navigation hints to the user through a given Web collection, based on its knowledge of the user's interests, the location and relevance of various items in the location, as well as the way in which other users have interacted with the collection in the past. The system starts by profiling the user, acquiring information about her interests. Each time the user requests a page, this information is routed through a proxy server in order to easily track the user session across the Web site and any links believed to be of interest for the user are

highlighted. Its strategy for giving advice is learned from feedback from earlier tours.

Cingil et al. [4] provides a broader view of personalization, through the use of various W3C standards. They describe how standards such as XML, RDF, can be used to create personalization applications. In this architecture, a log of the user's navigation history is created as a "user agent" at the client site gathers click stream information about the user.

Coenen et al. [5] propose a framework for self-adaptive Web sites, taking into account the site structure except for the site usage. The proposed approach is based on the fact that the methods used in Web usage mining produce recommendations including links that don't exist in the original site structure, resulting in the violation of the beliefs of the site designer and the possibility of making the visitor gets lost following conceptual but not real links. Therefore, they suggest that any strategic adaptations based on the discovery of frequent item sets, sequences, and clusters should be made offline and the site structure should be revised.

The most advanced system is the Web Personalizer, proposed in Mobasher et al. [9]. Web Personalizer provides a framework for mining Web log files to discover knowledge for the provision of recommendations to current users based on their browsing similarities to previous users. It relies solely on anonymous usage data provided by logs and the hypertext structure of a site. After data gathering and preprocessing (converting the usage, content, and structure information contained in the various data sources into various data abstractions), data mining techniques such as association rules, sequential pattern discovery, clustering, and classification are applied, in order to discover interesting usage patterns. The results are then used for the creation of aggregated usage profiles, in order to create decision rules. The recommendation engine matches each user's activity against these profiles and provides him with a list of recommended hypertext links.

Another approach proposed by Mobasher et al. [2] presents a scalable framework for Web personalization based on association rule mining from click stream data. This framework includes an efficient data structure for storing frequent item sets combined with a recommendation algorithm which allows for the generation of recommendations without first generating all association rules from item sets. We have also studied the impact of using multiple support levels for different types of page views, as well as the use of varying-sized user histories on the precision and coverage of the generated recommendations. The results show that the proposed framework can provide an effective alternative to standard collaborative filtering mechanism for personalization.

Personalization based on Web usage mining can enhance the effectiveness and scalability of collaborative filtering. However, without semantic knowledge about the underlying domain, such systems cannot recommend different types of complex objects based in their underlying properties and attributes. Middelton and Shadbolt et al. [16] explore ontological approach to recommender systems. Ontological user profiles allow inference to be employed, allowing interests to be discovered that were not directly observed in the user's behavior. Once profiles are represented using an ontology, they can communicate with other ontology which share similar concepts. One advantage of using an ontological user profile is that the profiles themselves can be visualized. The ontological representation allows users to provide feedback on their own profiles, which is used to significantly improve profile accuracy.

Dai and Mobasher et al. [6] explore various approaches for integrating semantic knowledge into the personalization process based on Web usage mining. We have considered approaches based on the extraction of semantic features from the textual content contained in a site and their integration with Web usage mining tasks and personalization both in the pre-mining and the post-mining phases of the process. We have also presented a framework for Web personalization based on full integration of domain ontology and usage patterns.

Dolog, Henze, Nejdl and Michael Sintek et al. [12] described an approach to bring personalization to the semantic web for the area of education and learning. We have shown how personalization functionalities can be embedded into semantic web services, supported by other services for retrieving learning resources or user information.

Eirinaki, Mvazirg, Varlamis et al. [10] presents SEWeP, a system that makes use of both the usage logs and the semantics of a Web site's content in order to personalize it. Web content is semantically annotated using a conceptual hierarchy (taxonomy).They introducing C-logs, an extended form of Web usage logs that encapsulates knowledge derived from the link semantics. C-logs are used as input to the Web usage mining process, resulting in a broader yet semantically focused set of recommendations.

A. Sieg, B. Mobasher, R.Burke et al. [1] Present an approach to personalized search that involves modeling the user context as ontological profiles by assigning implicitly derived interest scores to existing concepts in domain ontology. A spreading activation algorithm is used to maintain and incrementally update the interest scores based on the user's ongoing behavior. Experiments show that re-ranking the search results based on the interest scores and the semantic evidence captured in an ontological user profile enables an adaptive system to present the most relevant results to the user.

V. Wolowski, N. Ishikawa and H. Sumino et al. [15] a system was presented which uses machine learning, metadata, ontology and inference for realizing an improved process of content personalization. Through the

machine learning procedure ID3, it was possible to construct a user preference model and to generate metadata in an automatic manner.

Although there are various approaches to web personalization exists. But none of them took into consideration the provision for maintenance of contents in real time, which further helps in consistency enforcement on web using semantic web technologies.

## 3. PROPOSED ARCHITECTURE

The content personalization application autonomously provides user specific information and recommendations. The content management provides maintenance of contents in real time.

When developing a personalized website one should take into account that the personalization process depends greatly on the understanding of the user needs, goals and objectives. For personalizing any website we must have knowledge of two fundamental things: The data (contents of the website), the user model (user intensions, preferences and goals). Knowing the context or domain of the website is very important, as it guides the system to what the users most probably mean by their queries.

In [11] a combined approach to web personalization and content management is proposed. This combination results in better personalized contents and up-to-date information make sure that the user preferences and web contents are consistent with each other. In ontological structure based websites management of contents mainly consists of content editing and content presentation. Content editing means editing concept instances that is inserting, updating, and removing instances of ontology concepts. In this paper the implementation details of the proposed system is presented and the results of the implemented system is discussed.

The following functionality is provided through content editing:

**Instance creation and removal:** We can create instances of any concept in the ontology created, provided that user have enough permission to create such instances. The same occurs with the removal of concept instances.

**Instance editing:** To edit instances, we present the attributes of the concept from which the instance is instance of, as well as the attributes inherited through the concept hierarchy. User can also update the ontology concept instances.

**Consistency enforcement:** The system administrator can also check weather the underlying ontology concepts and their instances are consistent or not. Consistency enforcement is performed here to prevent inconsistencies arises from the instance editing.

Content presentation in an ontology-based website mainly consists of showing the ontology concepts and their related instances, presenting the details of ontology instances and their relations with other instances, and allowing the navigation through these ontology instances.

The proposed architecture is shown in Fig.1.

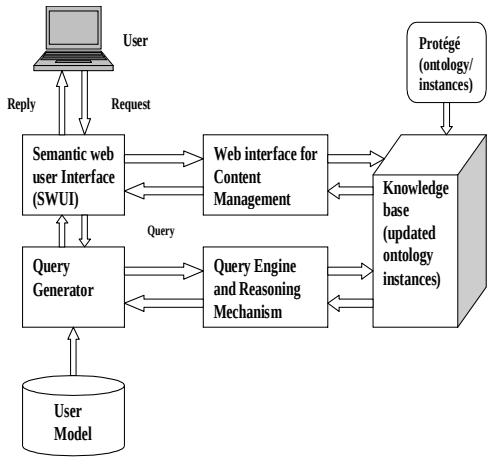


Figure 1: Proposed architecture

The explanation of different components of this architecture is as follows:

### 3.1 Knowledge Base

The backbone of this architecture is a knowledge base (KB), where the domain model ontologies/instances are stored. The KB can be connected with protégé ontology editor which allows the update of ontology concepts. After inserting, removing and updating instances the changes are stored in the knowledge base. The need of the web user for specific information is satisfied by querying the knowledge base.

### 3.2 User Model

The user model contains information on the goals, the needs, the preferences or intentions of the users. The user model stores the most relevant characteristic of the user in the context of the application domain. The user model should assist the system on how to provide information for the user through a set of rules that describe such preferences along with the user goals and intentions. In this case different user information is collected through user profiles. We stored the user information containing user preferences in the KB. This information describes relationship between the user model and the domain model concepts.

### 3.3 Query Generator

The user could enter his query in free text format and/or selects from existing options. On the other hand he could be navigating the system and looking for information on a certain topic. In either case this query is not in formal semantics format. The Semantic Web query generator transforms the query to the Semantic Web query based on its knowledge of the underlying domain ontology and application ontology which contains all the terms used within the application along with its semantics. The query is semantically rich, and retrieves data based on the existent relationships within the knowledge base.

### 3.4 Query Engine and Reasoning Mechanism

The reasoning mechanism would infer for new rules that come from ontologies and then it will evaluate the query. The query engine executes the query and returns the results. If the query results are satisfying to the system, it will continue with the next step, which is displaying these results. If on the other hand the query results are not satisfying (do not meet a certain threshold for the number of outputs or there were no results found) the system will give a message to user that no results found.

### 3.5 Content Management

In order to keep information up-to-date on the web, an interface is developed for content Management. Authorized users can access instances from their personalized homepages. Users can add, delete or change attributes from this interface. Changes are permanently saved to the Knowledge base. Consistency checking is also performed and presented to prevent inconsistencies.

## 4. IMPLEMENTATION DETAILS

The proposed architecture is being implemented using protégé ontology editor and java code. We have developed a web front-end, so that the users who do not want personalized information can navigate the web information using ontology concepts and their related instances. We have also developed a web interface through which system administrator can do content management. Consistency checking is also performed through this interface. The front end is a set of java servlets which query the KB over HTTP using SQL queries.

For implementation the web-site of university is chosen, which hosts various pages such as course information, faculty information, department information etc. The visitors are divided into two groups: student and teacher. To provide personalized information to users we used user profiles. User profile describes relationship between the user model and the domain model.

The Personalization starts when a user logged in into the web. The personalization is achieved here by modeling user profile as a set of concepts that has relationship to domain model concepts. A web interface is developed for users to maintain their profiles. At the time of login the user information is matched with the ontology instances stored in the knowledge base. Based on these matching personalized link is displayed in the user's homepage. The personalized home page guide user's to relevant information. If user is not satisfied with the provided links and want to access some more specific information. For this purpose information navigation through ontology is provided.

If the user wants to modify his or her profile then for this a web interface is provided. The user can access the instances from their personalized homepages. Then all the attributes are listed for update. Users can change attributes value from this interface. Multiple values can be changed at the same time. Changes are permanently saved to the knowledge base.

An interface for navigation is also provided. Ontology is used for the navigation as shown in figure 2. Users can access concept instances by clicking onto hyperlinks from the ontology hierarchy. All instances are queried from the KB and displayed on the right side. When a user clicks on an instance from the right side, a more detailed view is opened which contains all the information related to that particular instance.

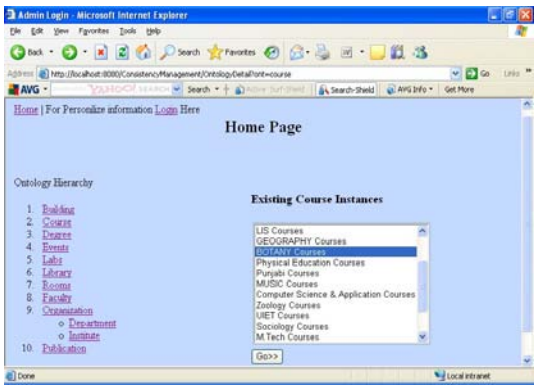


Figure 2: A view from semantic navigation

In order to keep website information up-to-date, we have developed a web interface for content management by authorized users. To simplify the insertion, deletion and updation of ontology instances, existing instance values are displayed in drop-down list. The system administrator can add, delete and update instance value from this web interface which is shown in figure 3. Changes are permanently saved to the KB and can be seen without restarting the server. Consistency checking is also performed and presented to prevent inconsistencies.

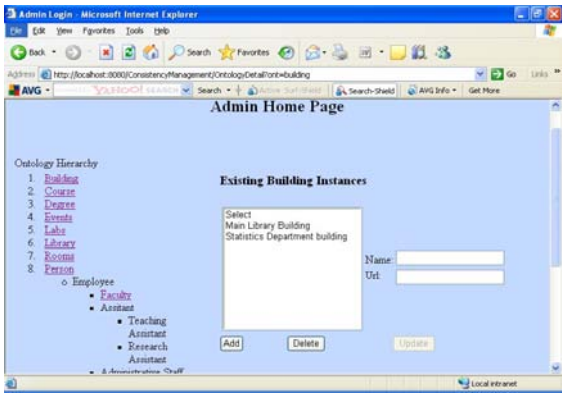


Figure 3: Web interface for content management

Next time whenever a user visits the website he will get up-to-date and consistent personalized information. The user model is also up-to-date in this case.

## 5. RESULT EVALUATION AND DISCUSSIONS

To evaluate the results of the system being implemented we carried out an experiment. We tested simple personalized system and ontological personalized system in terms of navigation links. For this experiment we have taken the websites of three different universities: Panjab University, Delhi University and Mumbai University. The

Experiment is performed by five users using both systems. During this experiment, five different users were asked to find the answers to five different questions. In this experiment no time limit was placed. For each user the total no. of links followed to find the correct answer is calculated. All users performed this task firstly on simple personalized web of different universities and then this task is performed on ontological personalized web of different universities. In this type of system our main aim is to reduce the no. of links followed to find particular information to as minimum as possible.

In this evaluation experiment, our main aim is to compare the performance of simple personalized system and ontological personalized system in terms of navigation links. Here we are comparing the results of simple system of Panjab University, Delhi University and Mumbai University with the results of ontological system of the same three universities. After comparing the results of different universities an average comparison between simple web of three universities and ontological web of three universities is done.

The results are illustrated in figure 4, 5, 6 and 7. The results show that users performed task better using ontological web comparing to simple web. Users found answer to questions more easily on ontological web. The main reason was the structure of information on ontological web. Ontology hierarchy provided a good structure for accessing web data. Also, users liked semantic links; they enable them to complete task easily.

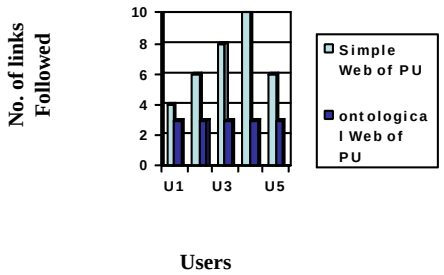


Figure 4: Results for Panjab University

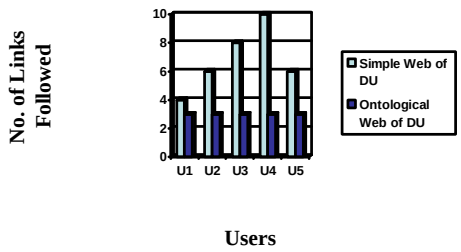


Figure 5: Results for Delhi University

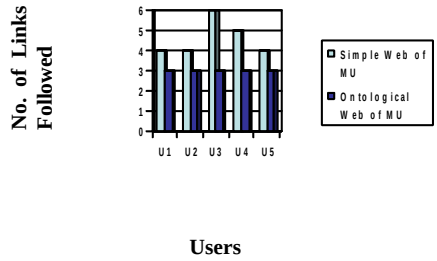


Figure 6: Results for Mumbai University

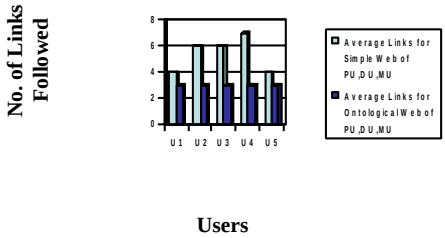


Figure 7: Average Results for PU, DU and MU

From this experiment we can say that the no. of links followed in simple personalized system is more than the links followed in ontological personalized system. This shows that the performance of ontological system is better than the simple system in terms of navigation.

From the results we can say that the ontological personalized system reduces the number of links when a user navigates the university system. As we know that lesser the number of links followed lesser will be the time to reach to a particular piece of information. So the ontological personalized system will give better results in terms of time constraint also.

## 6. CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed a web personalization system that makes use of ontologies for consistency enforcement and content editing. In the proposed architecture web personalization approach is combined with content editing which solves the problem of maintenance of contents in real time. This architecture also performs consistency checking between the web information and the ontologies underlying it. In conclusion this paper presented the features of the proposed system and our evaluation. The results show that users performed task better using ontological structure and their navigation was improved as compared to simple web. For future work we are planning to create the web services of the university system and providing personalized data on web.

## REFERENCES

[1] A. Sieg, B. Mobasher, R. Burke, “Learning Ontology - Based User Profiles: A Semantic Approach to Personalized Web Search”, In IEE E Intelligent Informatics Bulletin, Nov 2007, Vol. 8, No. 1, pp. 7-18.

[2] V. Aragao , A. Fernandes, C. A. Goble, "Towards an Architecture for Personalization and Adaptivity in Semantic Web", In the Proceedings of 3rd International Conference on Information Integration and Web-based Applications and Service (IIWAS 2001), LIAZ, Austrian Computer Society, Sept. 2001, pp. 139-149.

[3] T. Berners-Lee, J. Hendler and O. Lassila, "The Semantic Web", In: Scientific American , May 2001, Vol. 279, No. 5, pp. 34-43.

[4] I. Cingil, A. Dogac, and A. Azgin, "A Broader Approach to Personalization", Communications of the. ACM,, August 2000., Vol. 43, No. 8 , pp. 136–141.

[5] F. Coenen, G. Swinnen, K. Vanhoof and G. Wets, "A Framework for Self Adaptive Websites: Tactical versus Strategic Changes", In proceedings of WEBKDD'2000 Web Mining for ECommerce—Challenges and Opportunities, Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM Press, Boston, August 2000, pp. 75-81.

[6] H. Dai, B. Mobasher, "Integrating Semantic Knowledge with Web Usage Mining for Personalization", In Web Mining: Applications and Techniques, 2004, pp. 276-306.

[7] T. Joachims, D. Freitag and T. Mitchell, "WebWatcher: A Tour Guide for World Wide Web", In Proceedings of International Joint Conference on Artificial Intelligence IJCAI97, Nagoya, Japan, August 1997, Vol. 1, pp. 770-775.

[8] H. Liberman, "Letizia: An Agent that Assists Web Browsing", In Proceedings of Fourteenth International Joint Conference on Artificial Intelligence, Montreal, Quebec, Cana da, 1995, Vol. 1, pp. 924-929.

[9] B. Mobasher, R. Cooley and J. Srivastava, "Automatic Personalization Based on Web Usage Mining", Communication of the ACM, August 2000, Vol. 43, No. 8, pp. 142–151.

[10] M. Eirinaki, M. Vazirgiannis, I. Varlamis, "SEWeP: Using Site Semantics and Taxonomy to Enhance the Web Personalization process", In Proceedings of ninth ACM International Conference on Knowledge and Data Mining SIGKDD'03 , Washington DC, USA ,August 2003 , pp. 99-108.

[11] N. Goyal, M. Kumar, R. Vig, P. j. Kaur, "Consistency Management and Content Editing using Ontologies on Personalized Web", In proceedings of IEEE International Advanced Computing Conference (IACC 2009), Patiala, India, March 2009, pp. 1694-1696.

[12] P. Dolog, N. Henze, W. Nejdl, M. Sintek," Personalization in Distributed ELearning Environments", In Proceedings of WWW2004, New York, USA, 2004, pp. 170-179.

[13] M. Perkowitz, O. Etzioni, Towards "Adaptive Web Sites: Conceptual Framework and Case Study", In Artificial Intelligence, 2000a, Vol. 118, No. [1–2], pp. 245–275.

[14] M. Perkowitz, O. Etzioni , "Adaptive Web Site", Communication of the ACM, August 2000b, Vol. 43, No. 8, pp. 152-158.

[15] V. Wolowski, N. Ishikawa and H. Sumino, "Semantic Web Approach to Content Personalization", In Proceedings of International Conference on Mobile Ubiquitous Computing, System, Services and Technologies (UBICOMM2007), 2007, pp. 109-117.

[16] S.E. Middleton, N.R. Shadbolt, D.C. De Roure, "Ontological User Profiling in Recommender Systems", ACM Transactions on Information Systems (TOIS), Jan. 2004, Vol. 22, No.1, pp. 54-88.