OurDMOZ: A System for Personalizing the Web.

Dimitrios Pierrakos
Institute of Informatics & Telecomms.
NCSR "Demokritos"
Athens, Greece
dpie@iit.demokritos.gr

Georgios Paliouras
Institute of Informatics & Telecomms.
NCSR "Demokritos"
Athens, Greece
dpie@iit.demokritos.gr

Yannis Ioannidis
Dept. of Informatics & Telecomms.
University of Athens
Athens, Greece
yannis@di.uoa.gr

ABSTRACT

Community Web Directories constitute a form of personalization performed on Web directories, such as the Open Directory Project (ODP). They correspond to "segments" of the directory hierarchy, representing the interests and preferences of user communities and thus provide a personalized view of the Web. In this paper, we present OurDMOZ, a system that builds and maintains community Web directories by employing a Web usage mining framework. OurDMOZ, the prototype presented here, exploits Web directories to extend personalization to a larger part of the Web, outside the scope of a single Web site. OurDMOZ offers a variety of personalization functionalities including adaptive interfaces and Web page recommendations. An initial user evaluation of the system indicates the potential value of the enhanced personalized Web experience provided by OurDMOZ.

1. INTRODUCTION

Web Personalization, i.e., the task of making Web-based information systems adaptive to the needs and interests of individual users, or groups of users has received a lot of attention from the research community. A variety of services, such as Web page recommendation [6, 15], or adaptive Web sites [2], have been proposed in the literature. However, the majority of this work limits the personalization within the context of a single Web site. There is only a limited number of services, such as *StumbleUpon*[1], *My Web*[2], or *Genieo*[3], that offer personalization to the whole of the Web. Even, these services however, follow the approach of collaborative filtering, i.e., their recommendations are based on ratings of similar users, and therefore they suffer from several well-studied problems, such as cold-start and sparsity. The most important issue though is that personalization is restricted to what a finite number of users have seen and thus the coverage of the Web is limited.

[1]http://www.stumbleupon.com
[2]http://beta.bookmarks.yahoo.com/
[3]http://www.genieo.com/

The limitations of existing approaches motivated us to devise a new one that realizes Web-wide personalization, i.e., personalization that covers the Web, or to be more realistic a large part of it. In order to do this, we exploit Web directories, that attempt to organize Web content into thematic hierarchies. These hierarchies correspond to listings of topics, which are organized and overseen by humans. A Web directory allows users to find Web sites related to the topic they are interested in, by starting with broad categories and gradually narrowing down, choosing the category most related to their topics.

In previous work [12], we discussed the notion of *Community Web directory*, which is a personalized Web directory that corresponds to the interests of a user community. We also presented a Web Usage Mining framework to realize the personalization of Web directories. The framework has been used for the personalization of the Open Directory Project (ODP)[4], also known as DMOZ. In this work, we present *OurDMOZ*, a system that integrates and implements the various components of the proposed framework. In particular, OurDMOZ collects and processes usage data, maps the data onto the Web directory, uses machine learning techniques to extract the community models and finally builds the community Web directories.

OurDMOZ offers a number of novel personalization functionalities. First, a user can join a community either by specifying her preferences, or by using the system for some time and letting it decide on the most suitable community models. Thus, there is no requirement for personal information, or other private data, to be provided to the system. In addition, the assignment to a community can be either on a long-term or on a short-term basis, i.e., the user has the option to keep her community model across sessions or start with a fresh model each time.

The main contribution though of OurDMOZ is that it offers, through its Web application, a personalized view of DMOZ and through it a personalized view of the Web. In other words, the community Web directories support an adaptive interface, which can act as a starting point for navigating the Web. OurDMOZ also offers a recommendation service that suggests Web pages not within the scope of a particular Web site, but from the part of the Web that is covered by DMOZ. Being based on a thematic characterization of the Web, OurDMOZ supports associations between users and, as a consequence, recommendations, based on the semantics rather than the raw content of the Web pages.

[4]http://dmoz.org

The novel functionality provided by OurDMOZ has been evaluated by real users and the results indicate the potential benefits of the system and consequently of the concept of community Web directory to the end user.

The rest of this paper is organized as follows. Section 2 presents existing approaches to the construction of personalized Web directories. Section 3 introduces the knowledge discovery framework and describes the architecture of OurDMOZ. Section 4 provides results of the user evaluation study. Finally, section 5 summarizes the main conclusions of this work.

2. RELATED WORK

A number of studies exploit Web directories to achieve a form of personalization. In [4], users build their profiles by specifying a set of categories from the DMOZ hierarchy, while automated profile construction is proposed in [13, 10, 11, 8]. These profiles are typically used for personalized Web search, while the directory itself is not personalized. The personalization of Web directories is mainly represented by services such as Yahoo![5] and Excite[6], which support the manual selection of interesting categories by the user.

An early approach to automate this process, was the Montage system [1], which was used to create personalized portals, consisting primarily of links to the Web pages that a particular user has visited, while also organizing the links into thematic categories according to the ODP directory. A related technique for mobile portal personalization was presented in [14], where the portal structure was adapted to the preferences of users. In [3], a Web directory was used as a "reference" ontology and the Web pages navigated by a user were mapped onto this ontology, using document classification techniques. In this manner, a personalized view of the ontology was obtained. Finally, in [7], the similarity between users, based on navigation data within the DMOZ, was used to create clusters of DMOZ categories. These clusters were further exploited to recommend shortcuts within the Web directory.

Our work differs from the above-cited approaches in several respects. First, instead of simply using the Web directory for personalization, it also personalizes the directory itself. Compared to existing approaches to directory personalization, it focuses on aggregate or collaborative user models such as user communities, rather than content selection for single users. User community modeling is considered more appropriate at that scale, since it is very difficult to acquire accurate personal information for each user. Unlike most existing approaches, OurDMOZ does not require a small set of predefined thematic categories, which could complicate the construction of rich hierarchical models. Finally, the work presented in [7], which is closest to ours is limited to the recommendation of short navigation paths in the ODP hierarchy, rather than the personalization of the whole Web directory structure. Moreover, OurDMOZ does not assume that usage data are collected from the navigation of users within the Web directory. Thus, its applicability to independent services, such as a Web portal, is more straightforward.

Compared to our earlier work on this topic [12], in this paper we present an integrated system and a Web application, based on the knowledge discovery framework for community Web directories. To our knowledge this is the first system to construct aggregate user models, i.e., communities, using navigational data from the whole Web. The Web application provides personalized access to the Web directories through an adaptive user interface and recommendations that cover a broad area of the Web, i.e., not within the limits of a Web site. Moreover, we present a real user evaluation of the system and we compare it with the results that we obtained from "in vitro" experiments.

[5]http://yahoo.com

[6]http://www.excite.com

3. SYSTEM ARCHITECTURE

The methodology proposed in this work for the construction of community Web directories, is based on Web Usage Mining and results in the construction of user models. User communities are formed using data collected from Web proxies as users browse the Web. The goal is to identify interesting behavioral patterns in the collected usage data and construct community Web directories based on those patterns. The process of getting from the data to the community Web directories is realized by the following steps: *Usage Data Preparation* which comprises the collection and cleaning of the usage data, as well as the identification of user sessions, *Web Directory Initialization* which provides the mapping of Web pages on the categories of a Web directory, *Discovery of Community Models*, i.e., the main process of discovering the user models from data, using machine learning techniques and *Community Web Directory Construction* where the user models are mapped onto the Web directory.

These basic components have been integrated into a prototype system named OurDMOZ. Based on the constructed directories, OurDMOZ offers personalized access to DMOZ through its Web application. OurDMOZ integrates off-line and on-line modules. The off-line modules deal with the data preparation, Web directory initialization and knowledge discovery. In other words, they correspond to the basic components of the community Web directory construction process. The on-line modules realize Web personalization. These modules are responsible for the user interaction, collection of real-time user data, visualization of the community Web directories, and calculation of recommendations. The system can construct personalized directories for any Web directory. However, we have chosen to incorporate and exploit the structure and the content of DMOZ. The reason is that DMOZ is the largest Web directory, in terms of topic coverage, and is widely used, among others by the Google Directory. A pictorial view of the system's architecture is given in Figure 1 and discussed in detail below.

3.1 OurDMOZ Off-Line Modules

The Off-Line part of OurDMOZ has been presented in detail in [12]. Thus, we only briefly summarized it here:

Data Preparation involves the assembly of the various usage data into a consistent, integrated and concise view. Potential sources of such data are cache proxy servers, such as those operated by ISP's, large enterprisers or universities.
Web Directory Initialization associates the users' browsing data with the Web directory. The association step involves the mapping of users' data, i.e., the Web pages, onto the Web directory and is realized using the automated page classification method described in [5]. The Web pages are classified onto the DMOZ hierarchy using cosine similarity. Note that a Web page may belong in more than one DMOZ

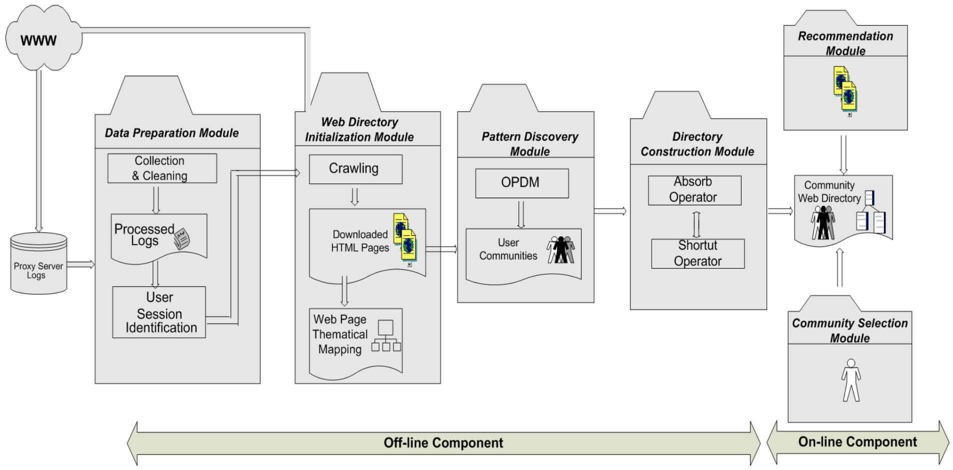

Figure 1: OurDMOZ architecture.

category. However, for reasons of simplicity, we assign the Web page to a single category, maximizing similarity.

Pattern Discovery. We employ unsupervised learning to discover patterns of interest, i.e., usage patterns that occur in data and represent the browsing preferences of community members. Each community model discovered contains a subset of the categories of the Web directory and is subsequently exploited to construct the community Web directory. For the extraction of community models, the pattern discovery module of OurDMOZ employs the *Objective Probabilistic Directory Miner (OPDM)* algorithm that achieved very good performance in our previous study [12].

OPDM is based on a probabilistic latent factor model, where the latent factors are responsible for associations between users. The method used by OPDM for the identification of latent factors in data is the PLSA method [9], which is supported by a strong statistical model. For each latent factor, the selected categories are used to construct a new Web directory. This corresponds to a topic tree, representing the community model, i.e., usage patterns that occur due to the latent factors in the data. The advantage of this approach is that it allows us to model more effectively the complex multi-dimensional preferences of users.

A number of categories from the initial Web directory are pruned, resulting in a reduced directory, named community Web directory. Each category retained combines two features: (a) it has a distinctive role in the original directory and (b) it is important for the users of the community. This approach has a number of advantages. First is the obvious shrinkage of the initial Web directory, which is directly related to the interests of the user community, ignoring all other categories that are irrelevant. Second, the selected approach allows us to construct overlapping patterns, i.e. a category might belong to more than one community directory, i.e. affected by more than one latent factor.

Directory Construction. The construction of useful community Web directories needs to go beyond the selection of categories by the pattern discovery algorithms. Further processing is required to improve the structure of the directory and this is achieved by the operators *Shortcut* and *Absorb* of the directory construction module. The the first operator creates "shortcuts" from the parent to the leaf nodes, whenever a category has a single descendant node. The Absorb operator applies to categories that became leaves in the community Web directory, which they were not leaves in the initial Web directory. Since all of their descendant categories are excluded from the community Web directory, they absorb their Web pages. This operator ensures that no information is lost, even when the "original" leaves are not included in the community Web directory. In the case though, where at least one descendant leaf is included in the community models, this operator is not applied, assuming that the users are not interested in the other leaf categories.

3.2 OurDMOZ On-Line Modules

Community Selection assigns a user to a particular community Web directory. OurDMOZ does not keep any personal data and this assignment is realized either by a semi or a fully-automated approach. The semi-automated process is based on the selection of interest terms by the user. The system generates automatically a set of terms that describe best the categories of each community Web directory. The union of these sets is presented to the user who can select a subset. Subsequently, OurDMOZ identifies and assigns the user to the community Web directory whose categories best match these terms, using cosine similarity.

More formally, let $\mathbf{G} = \{G'_1, G'_2, \ldots, G'_N\}$ be the set of N community Web directories, where $G'_i = \{c_1, c_2, \ldots, c_n\}, c_i$ a category, and T=$\{\tau_1,\ \tau_2,\ldots,\tau_m\}$ the union of sets of the m most frequent terms in each of the N community directories. Each community directory G'_i is represented as a vector $\vec{G'_i} = \{w_{i1}, w_{i2}, \ldots, w_{im}\}$, where $w_{ik} \in \{1, 0\}$, depending on whether the term t_k belongs to the set of frequent terms of G'_i or not. By selecting a subset of the terms in T, a user can be represented as a vector $\vec{u_j}$=$\{q_{j1},, q_{j2}, \ldots,q_{jm}\}$, where $q_{jk} \in \{1, 0\}$. Thus, the user u_j can be assigned to a particular community Web directory G'_i using cosine similarity as follows: $u_j \leftarrow G'_i \in \mathbf{G} : G'_i = \arg\max(sim(G'_i, u_j))$.

The alternative approach of fully-automated community selection determines the most suitable community Web directory for a particular user, based on the user's navigation in the system. This is realized by identifying the directory that contains the majority of the categories browsed by the user. Subsequently, the system informs the user that her interests match those of a community and the user can either choose the personalized directory, or continue browsing. The longer the user browses the directory, the more accurate the selection of a community directory becomes. More formally let $B = \{c_1, c_2, \ldots, c_\nu\}$ be the set of the categories browsed by a user u_j. The user u_j can be assigned to a particular community Web directory G'_i as follows: $u_j \leftarrow G'_i \in \mathbf{G} : G'_i = \arg\max |G'_i \bigcap B|$.

Community selection can be based either on the *short-term* or on the *long-term* model of a user. In the short-term case, the system adapts its personalization services to the current preferences of the user, which are determined either by the specification of new terms by the user, or by following the user's current browsing behavior. The short-term model does not require user registration to the system. It exploits only the current session's browsing behavior. In the long-term scenario, the system offers the option to store the community model that a particular user has been assigned to. This operation though, requires user registration to the system.

Recommendation in OurDMOZ exploits the community models to recommend Web pages. First, the system stores the Web pages that are accessed by the members of each community, while they navigate through the community Web directory. Whenever a community member requests recommendations, OurDMOZ delivers a set of the stored Web pages that have been viewed by the rest of the community and are novel to the user. For each request a maximum of 10 new Web pages are recommended to a user. More formally, if $A_{G'}$ is the set of Web pages browsed by users belonging to community G' and $B_u \subseteq A_{G'}$ is the set of Web pages browsed by user $u \in G'$, then the pages recommended to u will be a set $C \subseteq A_{G'}$ such that $C = A_{G'} \setminus B_u$.

Additionally, the user is offered a rating option for each recommended Web page. The rating scheme follows a five-level Likert scale from "Strongly dislike" to "Strongly like". The Web pages that receive very low rating from community users, are ignored by the recommendation process, while pages with high ratings are recommended first in the list. In contrast to other approaches, such as stumbleuppon.com, the recommendations offered by OurDMOZ are driven by the thematic structure of the Web directory and are justified semantically. More importantly, given that the recommendations are based on a community model, they typically cover a broad area of Web topics, whilst at the same time they are thematically cohesive. The recommendation and rating screen of OurDMOZ is shown in figure 2.

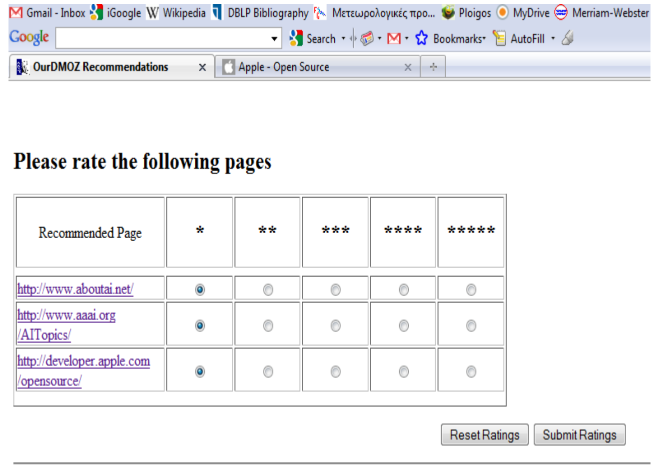

Figure 2: OurDMOZ Recommendation Sample..

4. EVALUATION

The evaluation of OurDMOZ was performed in two stages. First we wanted to evaluate the effectiveness of the discovered models, i.e., to measure the potential benefit from the community Web directories. We call this *"in vitro"* evaluation, since it does not involve assessment of the system by real users. This evaluation is similar to the one performed and discussed in [12], but it was applied to a different dataset. The second stage involved the actual user evaluation, where OurDMOZ was given to a set of users, who interacted with the system and used its personalization functionalities. The evaluation procedure is described in the following subsections.

4.1 "In vitro" evaluation

For this evaluation, we used the log files recorded in the cache proxy logs of an Information Systems department of a Greek university[7]. The log files recorded the outgoing browsing behavior of users for a period of six months. Data cleaning was performed and the remaining data, i.e., 36,459 Web pages, were downloaded locally using a Web crawler. The Web pages were then mapped onto the upper levels of DMOZ. The upper five levels of DMOZ comprise 59,863 categories, as described in [5]. Using these data, we employed the OPDM algorithm to build the community models with 10 latent factors. This number has been estimated based on past experiments and the number of users that have participated into the experiments.

The "in vitro" evaluation investigates the effectiveness of our approach along two dimensions: *Coverage* and *User Gain*. Coverage corresponds to the predictiveness of our model, i.e., the number of Web pages that interest the users and are covered by the community directories. User gain is an estimate of the actual gain that a user would have by following the community Web directory, instead of the initial Web directory, to get to the desired Web page.

Typically, there is a trade-off between coverage and user gain and thus it is interesting to analyze the interaction of the two measures it and identify good operating points. The common choice for such tradeoff studies is the Receiver Operating Characteristics (ROC) curves that have been used extensively in evaluating diagnostic systems. Adapting the idea of ROC curves to our measures, we plot coverage against (1-User Gain). We name this plot a "trade-off" curve since we are not measuring exactly sensitivity and specificity, as commonly done in ROC analysis.

In Figure 3 we present the trade-off curve that is generated by measuring coverage and user gain for different values of the parameters of OPDM. The optimal position is the top-left corner, where coverage and user gain reach their maximum values. From this figure we can conclude that there is a significant user gain (0,35-0,45) while coverage remains above 0,80. This result provides an initial indication of the benefits of personalization.

[7]The Department of Archives and Library Science of the Ionian University, Greece.

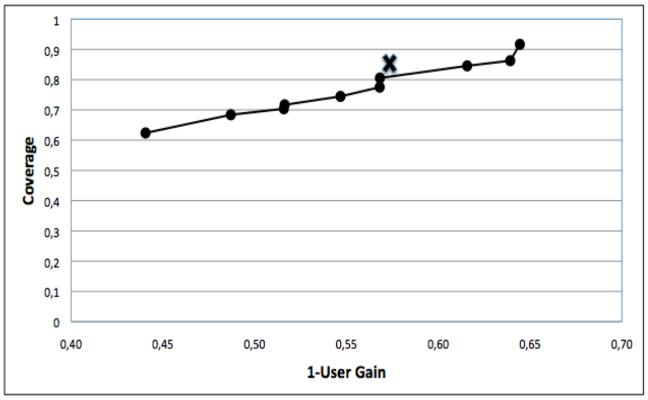

Figure 3: Coverage/User Gain Trade-off, for different values of the OPDM parameters.

4.2 User evaluation

User evaluation of the proposed personalization approach was achieved through OurDMOZ. We asked 79 postgraduate students from Departments of Information Systems, 28 men and 51 women, to evaluate the system. The community Web directories used for the evaluation are those that performed well in terms of coverage and user gain in the "in vitro" evaluation, specifically user gain higher than 0,43 and coverage larger than 0,81 (point X in figure 3). The evaluation involved three different scenarios, as described below.

Evaluation Scenario 1 The purpose of this scenario was to obtain a comparative assessment of the personalized and the non-personalized versions of the Web directories. The users were divided into two groups and they were asked to search for a set of Web pages in OurDMOZ, given a short description of each page. Group A was asked to perform the task first using the common DMOZ and then using the personalized version. Group B did the opposite, in order to avoid the bias of having located the correct answer when using the personalized directory. The same target Web pages were used in all stages of the evaluation, in order to avoid accidentally setting an easier task on one or the other group.

The measures that we used for the evaluation were the average time that a user spends to identify the requested pages and the average number of clicks that she performs in order to arrive at them. For time measurements, we have identified and removed from cache proxy logs, very long periods of inactivity. When measuring clicks, we excluded backward references. The assignment of a user to a community has been performed using the semi-automated method, described in section 3.2. The results of the evaluation averaged over all users (groups A and B) and for all Web pages requested are presented in Table 1. From this table we can see that a user spends much less time, and fewer clicks to arrive at a requested page using the personalized version of the Web directory. In fact, on average the user spends almost one third of the time and clicks to find the Web page.

Evaluation Scenario 2 In this scenario, the users navigated through OurDMOZ and were assigned automatically to community Web directories, using the technique presented in section 3.2. Then, each user was presented with a set of Web pages, and was asked to rate them using the five-level Likert scale supported by the Recommendation module. The purpose of this scenario was to evaluate the recommendation functionality of the system, as well as the match of the community Web directories to users.

Table 1: Average time and number of clicks used to arrive at a Web page.

	Avg. Time (min)	Std. Dev.	Avg. Clicks	Std. Dev.
non-personalized	7.22	1.16	79.56	36.5
personalized	2.58	0.50	28.68	11.82

Table 2: Average Recommended Web Page Ratings.

	Avg. Rating	Std. Dev.
Baseline	2.56	0.30
Personalized	3.32	0.38

In Table 2 we present the average ratings obtained for the top-10 visited pages that have been recommended to community members, compared to the top-10 visited pages recommended to users, i.e. regardless of their community assignment. The latter set of pages has been used as a baseline. The averaging is over all communities and over all recommended Web pages. From this table, we conclude that the recommendations generated by the system receive higher rating values than the baseline recommendations. This is a confirmation that OurDMOZ identifies valid common preferences inside each community.

Evaluation Scenario 3 Similar to scenario 2, the users navigated through OurDMOZ and were assigned automatically to community Web directories, using the technique presented in section 3.2. Subsequently, they were asked to use OurDMOZ in a "free style" and fill in a small questionnaire. The questions were answered in a seven-level Likert scale from "Strongly disagree" to "Strongly agree", and evaluated the following factors:

1. How easy it is to use the system (*Ease of use*).
2. How easy it is to learn the system functionalities (*Learnability*).
3. How efficient is the organization of the personalized information in the system (*System's organization*).
4. How close are the community directories to the actual user preferences (*Fit of the model*).
5. How a user is assisted by the personalized system (*User Gain*).
6. What is the level of the user's satisfaction from the system (*Overall Satisfaction*).

Figure 4 presents the normalized results of the users' responses to the questionnaire. A number of interesting conclusions can be derived from this figure. First, in all assessment factors, OurDMOZ receives ratings above average. It is considered a particularly easy system to learn and use, offering well-organized personalized information. Furthermore, the community models, seem to match user preferences at a level of above 60%. By about the same percentage, users also believe that they benefit from the system. Given the difficulty of providing personalization and corresponding recommendations across the Web, these results are

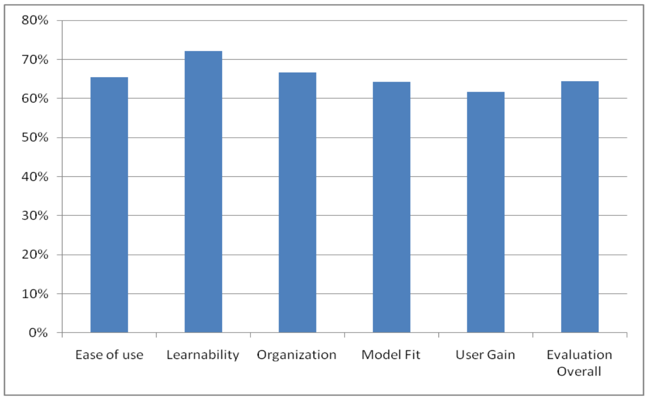

Figure 4: Questionnaire Results.

very encouraging. Finally, we can also conclude that the results for model fit are comparable to those obtained by the off-line evaluation of the system. In fact, user gain reached higher values compared to the off-line evaluation.

5. CONCLUSIONS

We presented a novel approach to the personalization task. In particular, we described OurDMOZ, a prototype that implements a Web usage mining methodology, for the construction of community directories. Community Web directories are an attempt to personalize the whole Web, based on the interests and preferences of user communities. OurDMOZ offers a variety of personalization functionalities including adaptive interfaces and Web page recommendations. The novelty of OurDMOZ focuses on the fact that personalization takes place across the Web and not in a single site. We also performed a user evaluation study, using qualitative and quantitative metrics, in order to assess: (a) whether the concept of community Web directories is beneficial to the end users and (b) the value of different functionalities. The results have shown that the users considered community Web directories as an interesting approach to Web personalization and they found OurDMOZ easy and helpful.

We hope that this paper will contribute to the effort of moving from Web site personalization, to Web-wide personalization. In this direction, several open issues remain. Document classification at the scale of a Web directory is a very challenging task. Moreover, the community selection and recommendation modules of OurDMOZ can be improved using more advanced techniques. Finally, additionally user studies can be performed to identify potential improvements to the system.

6. ACKNOWLEDGMENTS

We would like to thank Christos Papatheodorou and John Tsakonas, of the Ionian University for helping us with the user evaluation and analysis of the results. We would also like to thank the students who participated in the experiments.

7. REFERENCES

[1] C. R. Anderson and E. Horvitz. Web montage: A dynamic personalized start page. In *11th International Conference on World Wide Web*, pages 704–712, May 2002.

[2] P. Brusilovsky, A. Kobsa, and W. Nejdl, editors. *The Adaptive Web, Methods and Strategies of Web Personalization*, volume 4321 of *Lecture Notes in Computer Science*. Springer, 2007.

[3] J. Chaffee and S. Gauch. Personal ontologies for web navigation. In *9th International Conference on Information and Knowledge Management (CIKM '00)*, pages 227–234. ACM, 2000.

[4] P. A. Chirita, W. Nejdl, R. Paiu, and C. Kohlschütter. Using odp metadata to personalize search. In *28th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 178–185, 2005.

[5] C. Christophi, D. Zeinalipour-Yazti, M. D. Dikaiakos, and G. Paliouras. Automatically annotating the ODP Web taxonomy. In *11th Panhellenic Conference on Informatics (PCI 2007)*, Patras, 2007.

[6] W. Chu and S.-T. P. Park. Personalized recommendation on dynamic content using predictive bilinear models. In *WWW*, pages 691–700, 2009.

[7] T. Dalamagas, P. Bouros, T. Galanis, M. Eirinaki, and T. Sellis. Mining user navigation patterns for personalizing topic directories. In *9th annual ACM international workshop on Web information and data management*, pages 81–88, 2007.

[8] J. Garofalakis, T. Giannakoudi, and A. Vopi. Personalized web search by constructing semantic clusters of user profiles. In *12th International conference on Knowledge-Based Intelligent Information and Engineering Systems*, pages 238–247, 2008.

[9] T. Hofmann. Probabilistic latent semantic analysis. In *15th Conference on Uncertainty in Artificial Intelligence (UAI' 99)*, pages 289–296, 1999.

[10] Z. Ma, G. Pant, and O. R. L. Sheng. Interest-based personalized search. *ACM Transactions on Information Systems*, 25(1):5, 2007.

[11] T. Oishi, K. Yoshiaki, M. Tsunenori, H. Ryuzo, F. Hiroshi, and M. Koshimura. Personalized search using odp-based user profiles created from user bookmark. In *10th Pacific Rim International Conference on Artificial Intelligence*, pages 839–848, 2008.

[12] D. Pierrakos and G. Paliouras. Personalizing web directories with the aid of web usage data. *IEEE Transactions on Knowledge and Data Engineering*, 22:1331–1344, 2010.

[13] A. Sieg, B. Mobasher, and R. Burke. Ontological user profiles for representing context in web search. In *IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Workshops*, pages 91–94, 2007.

[14] B. Smyth and C. Cotter. Personalized adaptive navigation for mobile portals. In *15th European Conference on Artificial Intelligence*. IOS Press, 2002.

[15] G. Xu, Y. Zhang, and Y. Xun. Modeling user behaviour for web recommendation using lda model. In *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 529–532, 2008.