# Enhancing Semantic Web Services Discovery Using Similarity of Contextual Profile

Zahira Chouiref
University Akli Mohand Oulhadj
Bouira, Algeria
Email: zahirachouiref@gmail.com

Abdelkader Belkhir
University of Sciences and Technology
Houari Boumediene, Algiers, Algeria
Email: kaderbelkhir@hotmail.com

Allel Hadjali
Laboratory of Informatique and Automation
Systems - ENSMA, Poitiers, France
Email: allel.hadjali@ensma.fr

***Abstract*—Due to the proliferation of Web services and to the complexity and diversity of users' needs, new efficient mechanisms to automatically discover Web services are strongly required and desired. Recent advances have enabled the supply/consumption of services by suppliers/users of different categories. Every user has some particular interests and preferences when searching appropriate services on the Web. Supporting the "service profile" in the specification of Web services becomes then a paramount important issue. It can especially help applicants and Web services providers to produce accurate descriptions and thus improve the degree of relevance of the responses returned in the process of Web services discovery. In this paper, we present a novel method in the view of users requirements for Web services discovery. Especially, we propose (i) a framework to model the end user and the Web service contextual profile for best search, (ii) as well as a solution of semantic Web services discovery based on contextual profile similarity. This similarity is performed using a hybrid similarity measure we have defined.**

***Keywords*—Semantic Web Services; Context; Profile; Semantic Web Services Discovery; Hybrid Similarity Measure.**

## I. INTRODUCTION

The evolution of the IT (Information Technology) tools led to the development of new paradigms that describe interactions that exist over IT applications such as SOA (Service-Oriented Architecture) [15]. Service orientation is a promising paradigm for offering and consuming functionalities within and across organizations in the Web. Indeed, SWS (Semantic Web Services) allow a homogeneous use of heterogeneous software components deployed in large networks and in particular the Internet.

Several studies have already been done in the area of SWS discovery which is the process permitting to find the most suitable set of Web services that fulfill some functionality. Some of the discovery approaches are syntax based [3][5], while other are semantic based [7][8][9]. Despite the simplicity and facility of their implementation, syntactic approaches exhibit some serious limitations. The predominant problem is the restrictions posed by keywords matching that do not allow retrieval of SWS with similar functionalities; two WSDL (Web Services Description Language) descriptions can be used to describe the same service but with different words. However, when modeling Web services using ontologies, the semantic representation of concepts and their relations can be exploited and thus semantic matching to be performed. To enable SWS discovery, the first thing we need is a semantic service description language OWL-S (Ontology Web Language for Services) [14]. This development is significant since it seems to be able to approach certain insufficiencies of the syntactic approaches and tackle some of the UDDI (Universal Description, Discovery and Integration) register inadequacies.

However, with the exponential growth of SWS [17], the diversity of users and the conditions under which they access Web services, finding the relevant SWS that fit the needs and the context of the user is becoming a challenging task. With Web 2.0 applications and particularly e-business and e-commerce applications, SWS discovery is becoming much more important in a Web context. Our approach aims at solving queries using the profile of users from a contextual informational view of the Web where users have several features such as the client terminal, the client preferences, his/her location, etc. All these parameters form a particular context of user called the *contextual profile*.

The notions of context and profile have been the subjects of many works. Several meanings of such notions exist in the literature. For instance, and as pointed out in [21], "*A profile is defined by a set of attributes, possibly organized into abstract entities, whose values can be user-defined or dynamically derived from user behavior. A profile is supposed to characterize user domain of interest and all his specific features that help the information system to deliver the most relevant data in the right form at the right place and the right moment*". Since generally a person has several interests, user profiles should be defined to represent the various interests of the user. In [22], a user profile is proposed to represent the distinct interests related to a user. Various definitions of the context are also given and summarized in [10]. Brown et al. [23] define the context as being information about *location*, *the identity of people in close proximity*, *physical conditions*. Ryan et al. [24] add to this definition the notion of *time*. In [25], four categories of context {*location, identity (user), activity (state), time*} are identified as the more important parameters in practice. As stated previously, it is not easy to give a complete definition for a context. In fact, the notion of context is not universal but relative to some situation and application domain [23].

**Challenges.** The problem of interest is that the search engine produces several results in response to a user query,

 ISBN: 978-1-61208-280-6 

with the consequence that the truly relevant results are often missed.

- In order to face this unsatisfactory situation, one of the suggested solutions leverages information related to user profile. Indeed, the results produced by a search engine for a given query are not the same, dependent on the user profile and the context in which the user made the request. The main idea is to understand what the profile information is depending on a user context during a session: How do results depend on profile and change in context? The service context can group the *service localization (geographical restriction)*, *the implementation cost*, *the QoS (Quality of Service) parameters*, etc. The user context can be formed of his/her *localization*, his/her *devices*, etc. Since the user profile can vary during a session, the system must be able to adapt it in order to select services according to the new context. "*A contextual profile of an end user consists of a list of concepts which specifies the user domain of interest, personal information (age, sex, etc.), data quality, data on behavior, preferences and security, devices, purpose in a given temporal context, social context, spatial context and informative context*". Personal characteristics can strongly influence the interaction with a system and security can distinguish the user from the others, in a given context. Suppose a doctor is looking for a hotel that is in close proximity to the airport, with shuttle service, and with at least one conference room that has Internet connection and a 3D printer. The doctor is performing his query via the smartphone. Current search engines provide a list of all the hotels, but the question remains: which one to choose? End user context consists of his/her device (smartphone) from which we can deduce its navigator, operating system, location and the type of printer that will be used (wifi printer), etc. The end user is considered as guest of the secured system; therefore, the access to the printer and Internet is controlled. According to the user profile (guest or administrator), the user should have received a confidential message regarding the access rights to the resources. This is done to identify the sender so that the system can determine the user's access rights. His/Her profile consists of his/her profession (doctor) and his/her preferences (hotel with conference room, Internet, 3D printer, close to the airport and shuttle service, etc).
- In addition, the methods available in the UDDI publication do not contain a formal model describing the context of services use. Therefore, services discovery could not be achieved efficiently without considering their contextual profile. When a user requests a Web services discovery system, (s)he would have services tailored to his/her context and profile.

**Contributions.** Our paper specifically includes the following contributions:

- Combining context and profile models in one model called *Contextual Profile Model* that takes advantages of both models. The proposed model provides relevant and adapted results to the user requirements. It also allows the representation of all information that characterizes both the user and the service.
- We allow discovering suitable SWS based on the calculation of similarity between the user profile and the service profile according to a given context (the same query issued by different users may have different results as it is evaluated using different profiles in different contexts).

In the remainder of this paper, we first define in Section II an overview of existing efforts towards the SWS discovery and then we provide a critical analysis of these approaches. In Section III, we set up the contextual profile model firstly, so as to present a formalism of service and user in details; then, we present SWS discovery architecture. Finally, Section IV depicts a conclusion with the main research directions.

## II. Related Word and Discussion

Our work can be positioned in the new interdisciplinary area of Web Science which is the effort to bridge and formalize the semantic and technical aspects of the World Wide Web.

Previous works [7][8][9] have focused mainly on providing means to describe the functionality of a SWS and to allow a very expressive language for querying services. However, none of the works discusses in depth the concept of the profile and how a publisher/requester should provide context data about his/her services. The information should have more user context centric presentation in the discovery system. For instance, multiple Web services with similar functionalities are often available, best service(s) among them should be selected. To achieve this goal, one way is to use context [20] and QoS parameters which, generally, include performance, usability, safety, cost, etc.

The studies in [13] focus on QoS in discovery systems. The service consumer searches UDDI registry for a specific service through discovery agent which helps to find best quality service from available services which satisfies QoS constraints and preferences of requesters. Context-Awareness as proposed in [11] performs the necessary changes in the service behavior and/or the data handled in order to adapt the service to the context of the each user. Rong et al. [12] suggest with an example that context should be domain oriented or problem oriented in Web services discovery system. They divide context in two categories as explicit and implicit, with Personal profile oriented context, Usage history oriented context, Process oriented context and other context. Chukmol et al. [6] propose the personal opinion on service functionality and quality or invocation cost should also be considered by collaborative tagging-based environment for Web services discovery. The study done in [4] has proposed a novel approach to enhance Web services discovery based on, among others, QoS, customer's preferences and past experiences. The work in [26] presents an alternative approach for supporting users in Web services discovery by implementing the implicit culture approach for recommending Web services to developers based

 ISBN: 978-1-61208-280-6 

on the history of decisions made by other developers with similar needs.

The limitation of these approaches is that they make the system architecture more complicated when new attributes and constraints are introduced. SWS properties include several parameters like the functional (*Input, Output, Preconditions, Effects, functionalities,* etc.) and non-functional parameters (*QoS, property identifies the technical standards or protocols for implementing services*, and *categorization*). However, the majority of suggested approaches focus only on some parameters: *QoS, localization, user behavior.* Moreover, few works took into account multiples qualitative and quantitative parameters to help users to find the best service during the discovery process. It is also noticed that the suggested semantic approaches are based on the same technique, which consists in calculating the semantic correspondence level between the functional and non-functional parameters of the services and those cited in the user request. One of the big problems of Web search is the definition of a correspondence function between the representation of the proposed service and the user request. In order to fill this gap, we propose a new approach to improve the automatic SWS discovery, based on the measure of similarity between the user contextual profile and the one of the available services.

Thus, Web services that fit better the profile and the context of end user are retrieved. In our approach, we also indicate the interest of the proposed similarity measure in order to sort the candidate services due to the profile and the context attributes.

## III. A Contextual Profile Similarity Based Semantic Web Services Discovery Approach

We present in this section the service, the user and the contextual profile formalism, as well as the proposed architecture. Finally, we show how a new similarity measure can be used to enhance the discovery process.

### A. Model of Service/User Contextual Profile

Different attempts have been done to collect and classify profile's information. Most of the profile categorization has been done by Amato et al. [2].

The information of the contextual profile can be static (personal data, etc), evolutionary (intellectual quality, preferences, etc) and temporary (localization, devices, etc). These pieces of information must be captured to match demands to offers of services, on the syntactic and semantic level in order to improve the relevance of answers during a discovery session. In Figure 1, we show the proposed model that contains several dimensions able to describe the most information characterizing a profile. This general structure takes the form of a tree that contains a hierarchy of concepts. Each concept is constituted of one or several sub-concepts, that contain to their turn one or several attributes. The structure thus defined is flexible in the sense that different features can be spread through the tree structure of the proposed description. It permits to model the user's contextual profile soliciting the service as well the SWS's contextual profile offered.

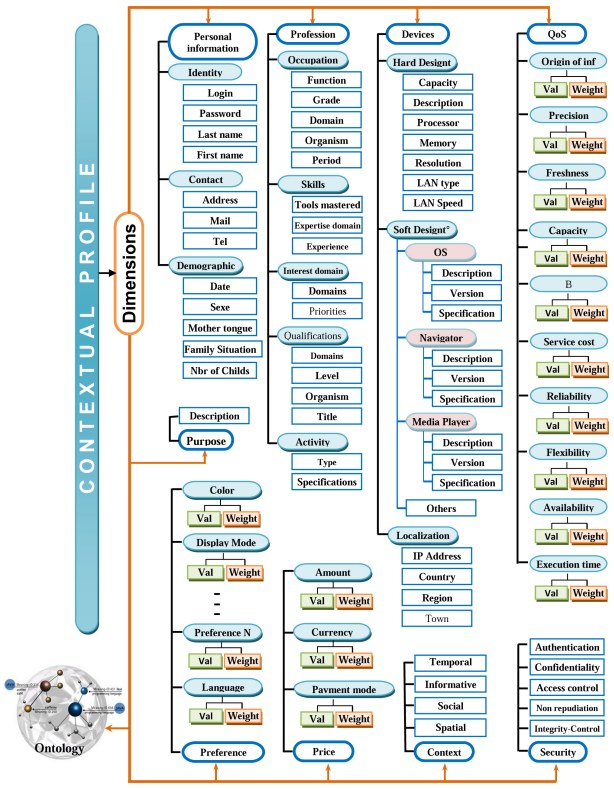


Fig. 1. Concepts and sub-concepts of the contextual profile of the service/user.

### B. Proposed Discovery Architecture

Our approach supported by the architecture depicted in Figure 2, is capable of integrating the contextual profile into the process of SWS discovery. It is composed of: *(i) basic elements of SOA* [15] namely: the service requester, the service provider and the service registry; *(ii) Interoperability Module* which comprises an administrator of the profile, an inter ontology similarity module and a contextual profile database, *(iii) Discovery and Selection Engine* which contains a contextual profile filtering module of Service/Request and a similarity measure treatment module.

### C. Semantic Web Service, Request and Contextual Profile Formalism

Formally, an SWS is defined as a quintuple $\{n_s, d_s, p, op, cpp\}$ such that:

- $n_s$ is the name of the SWS.
- $d_s$ is the functional description of the SWS.
- $p$ is the set of parameters describing the SWS.
- $op$ is the set of operations of the SWS.
- $cpp$ is the set of concepts constituting the contextual profile of provided service.

Operation Op is defined as a quintuple $\{n_o, d_o, i, o, pre, ef\}$ such that:

- $n_o$ is the name of the operation.
- $d_o$ is the functional description of the operation.
- $i = (i_1, \ldots i_n)$ is the set of input parameters of the operation

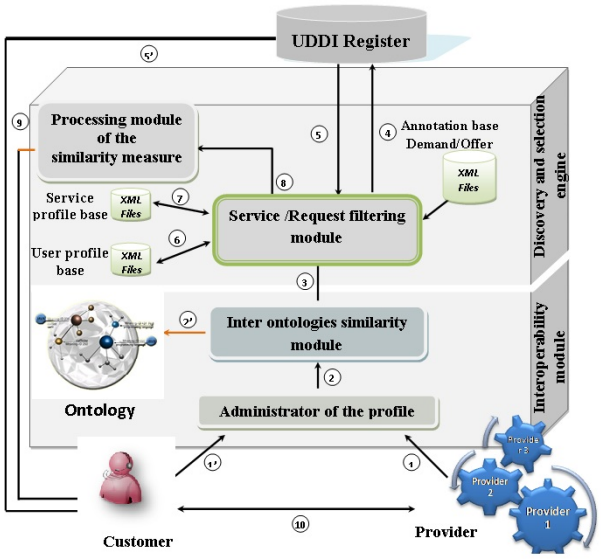


Fig. 2. Proposed architecture.

- $o = (o_1, \ldots o_n)$ is the set of output parameters of the operation.
- $prc$ specifies the preconditions of the service.
- $ef$ specifies the effects of the service.

Request R is a quintuple $\{n_r, i_r, o_r, cpr, s\}$ such that:

- $n_r$ is the name of the request.
- $i_r = (i_{r1}, \ldots i_{rn})$ is the set of input parameters of the request.
- $o_r = (o_{r1}, \ldots o_{rn})$ is the set of output parameters of the request.
- $cpr$ is the set of the concepts constituting the contextual profile of the required service (user contextual profile).
- $s$ $(0 < s < 1)$ is the interest score defined by the user.

Each parameter (*i*, *o*) in the list of (inputs, outputs) respectively is defined as a quadruple $\{n_t, v, t, c_o\}$ such that:

- $n_t$ is the name of the attribute.
- $v$ is the value of the attribute.
- $t$ is the data type of the attribute value.
- $c_o$ is the concept of the ontology which is connected to the attribute to improve its semantic quality. Some approaches use ontologies for the purpose of searching the context and the interests of the users [18][19].

Contextual profile is modeled through a set of concepts, noted $CP_i$ (contextual profile parameters). Each $CP_i$ is characterized by a set of values and weight associated with each parameter, $CP_i = (v_i, w_i)$.

### D. A New Similarity Measure

As we described above, the profile is modeled as multidimensional attributes through the tree structure of the description proposed in Figure 1. In general, there are two different approaches for expressing these attributes: a quantitative and a qualitative one. In our work, we propose to improve the similarity measure defined in [16] by combining syntactic with semantic similarity of each qualitative/quantitative attribute of request/service contextual profile. This hybrid similarity measure allows us to compute an overall score for each service retrieved and thus to recommend to the user the best services that match his/her profile. The similarity measure introduced can be formalized as follows:

- Let P be a set of objects' profile (users, documents, Web services, etc.). An object is described by m contextual characteristics $X = (CP_1, CP_2, \ldots CP_m)$.
- Let N be a set of services. Each service is composed of a set of contextual profile concept.
- Let $CP_x$, $CP_y$ be a contextual profile belonging to P for a given request and service respectively, and $(w_1, \ldots w_m)$ is a set of weights associated to each characteristic of $CP_x$, where $\sum w_i = 1$ ($w_i$ is a real number between 0 and 1).
- We define a threshold $s$ in order to present only services that have a rate of similarity with the user profile higher than the defined threshold .

The quantitative similarity measure $QSim$ [16] is defined as follows:
$QSim : P \times P \rightarrow [0, 1]$

$$QSim(x, y) = \frac{a \sum_{i=1}^{a} w_i \times sima(x_i, y_i)}{(a \sum_{i=1}^{a} w_i) + (b \sum_{i=1}^{b} w_{a+i}) + (c \sum_{i=1}^{c} w_{a+b+i})} \quad (1)$$

where $x_i$ and $y_i$ are numerical values.

As for our similarity measure $SimPro$, it is given by the following formula:
$SimPro : P \times P \rightarrow [0, 1]$

$$SimPro(CPx, CPy) = \frac{a}{a+b} \times \frac{a \sum_{i=1}^{i=a} w_i \times sima(CPx_i, CPy_i)}{(a \sum_{i=1}^{i=a} w_i) + (b \sum_{i=1}^{i=b} w_{a+i})} \quad (2)$$

- a is the set of common characteristics of $CP_x$ (requested by the client) and $CP_y$ (suggested by the service).
- b is the set of characteristics existing in $CP_x$ and not existing in $CP_y$ (requested by the client but not suggested by the service).
- $c$ is the set of characteristics existing in $CP_y$ and not existing in $CP_x$ (suggested by the service but not requested by the client).
- sima is the atomic similarity between each characteristic of $CP_x$ and $CP_y$. It is defined in a universe **U** and can be modeled as follows:

$sima : \mathbf{U} \times \mathbf{U} \rightarrow [0, 1]$

$$sima(CPx_i, CPy_i) = \begin{cases} 1 & if \quad c_1 \\ 0 & if \quad c_2 \\ \frac{min(CPx_i, CPy_i)}{max(CPx_i, CPy_i)} & if \quad c_3 \\ value \in [0, 1] & if \quad c_4 \end{cases} \quad (3)$$

With:

$c_1 \Leftrightarrow CPx_i = CPy_i$.
$c_2 \Leftrightarrow CPx_i = 0 \vee CPy_i = 0 \vee CPx_i = \perp \vee CPy_i = \perp$.
$c_3 \Leftrightarrow CPx_i, CPy_i$ are quantitative values.
$c_4 \Leftrightarrow CPx_i, CPy_i$ are qualitative values.

 ISBN: 978-1-61208-280-6 

if $CPx_i$ and $CPy_i$ are qualitative values, *sima* is calculated (in our example) by using the Jaccard coefficient [1].

$SimPro$ checks the properties of similarity measures defined in [16]. Let us note also that the most powerful strength of the quantitative/qualitative similarity measure proposed is its ability to take into account attributes of different natures (either qualitative or quantitative) both in the contextual profile of the request and the contextual profile of the Web services. Other approaches borrowed from information retrieval or domain ontology [27], can be used for measuring atomic similarity ($sima$) between concepts.

Note also that the performance of Web service may depend on the number of features it was mentioned by the user and published by the provider, the weighted atomic similarity of each request's characteristic and the quantity (a/a+b) that is in fact the average frequency of all mentioned attributes.

## IV. Conclusion and Future Work

We have proposed in this work a new approach to SWS discovery. The aim of this approach is to automatically discover relevant services based on measuring the similarity of user request and SWS using the contextual profile information during the search and selection steps. A quantitative and qualitative similarity measure is applied for the management of the contextual profiles. This hybrid similarity allows to retrieve SWS that better satisfy the user needs. We plan to conduct thorough experiments to study the effectiveness of the proposed formula for measuring similarity, to analyze the quality of SWS from a user point of view, and to consider the use of ontology for context and profile concepts. Another idea to improve the quality of the answers is to consider the parameters of user profile modeled using gradual concepts which can be represented thanks to fuzzy sets.

## References


[1] R. Real, "Tables of Significant Values of Jaccard's Index of Similarity," Misc. Zool, 22.1, 1999, pp. 29-40.

[2] G. Amato, U. Straccia, "User Profile Modeling and Applications to Digital Libraries," in Proceedings of the Third European Conference on Research and Advanced Technology for Digital Libraries, Paris, France, 1999, pp. 184-197.

[3] M. B. Blake and M. F. Nowlan, "A Web Service Recommender System Using Enhanced Syntactical Matching," IEEE International Conference in Web Services ICWS, USA, January.2007, pp.575-582.

[4] R. Benaboud, R. Maamri, and Z. Sahnoun, "User's Preferences and Experiences based Web Service Discovery using Ontologies," in Proceedings of the Fourth International Conference on Research Challenges in Information Science (RCIS), Nice, France, May.2010, pp. 121-126.

[5] P. Palathingal and S. Chandra, "Agent Approach for Service Discovery and Utilization," in Proceedings of the 37th Annual Hawaii Internationnal Conference on System Science HICSS-37, USA, January.2004.

[6] U. Chukmol, A. Benharkat, and Y. Amghar, "Enhancing Web Service Discovery by using Collaborative Tagging System," IEEE 4th International Conference on Next Generation Web Services Practices, October.2008, pp. 54-59.

[7] F. Nawaz, K. Qadir, and H. Farooq Ahmad, "SEMREG-Pro: A Semantic based Registry for Proactive Web Service Discovery using Publish Subscribe Model," Fourth International Conference on Semantics, Knowledge and Grid, IEEE Xplore, Pakistan, December.2008, pp. 301-308.

[8] C. Wu, E. Chang, and A. Aitken, "An Empirical Approach for Semantic Web Services Discovery," 19th Australian Conference on Software Engineering, IEEEXplore, Perth, March.2008, pp. 412-421.

[9] G. Wen-yue, Q. Hai-cheng, and C. Hong, "Semantic Web Service Discovery Algorithm and its Application on the Intelligent Automotive Manufacturing System," International Conference on Information Management and Engineering, IEEEXplore, China, April.2010, pp. 601-604.

[10] M. Bazire and P. Brezillon, "Understanding Context before using it," in CONTEXT'05 Proceedings of 5th International Conference on Modeling and Using Context, 2005, pp. 29-40.

[11] M. Keidl and A.Kemper, "Toward Context/aware Adaptable Web Services," in 13th international World Wide Web Conference, USA, 2004, pp. 55-65.

[12] W. Rong and K. Liu, "A Survey of Context Aware Web Service Discovery: From User's Perspective," Fifth IEEE International Symposium on Service Oriented System Engineering, June.2010, pp. 15-22.

[13] T. Rajendran and P. Balasubramanie, "An Optimal Agent-Based Architecture for Dynamic Web Service Discovery with QoS," Second International Conference on Computing Communication and Networking Technologies, IEEEXplore, July.2010, pp. 1-7.

[14] D. Martin and all, "OWL-S: Semantic Markup for Web Services," W3C Member Submission (2004) http://www.w3.org/Submission/OWL-S, November.2004 [retrieved:April, 2013].

[15] D. Booth, and all, "Web Services Architecture," http ://www.w3.org/TR/2003/WD-ws-arch-20030808/, Aout.2003 [retrieved:April, 2013].

[16] A. Belkhirat, A. Belkhir, and A. Bouras, "A New Similarity Measure for the Profiles Management," in the Proceedings of the 13th International Conference on Computer Modeling and Simulation (UKSIM), Cambridgeshire United Kingdom, April.2011, pp. 255-259.

[17] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," Scientific American, vol. 284, no. 5, May.2001, pp. 34-43.

[18] C. Bizer, and all, "A Crystallization Point for the Web of Data," Web Semantics : Science, Services and Agents on the World Wide Web, vol. 7, n 3, September.2009, pp. 154-165.

[19] S. Calegari and G. Pasi, "Ontology-Based Information Behaviour to Improve Web Search," Future Internet, vol. 2, n 4, October.2010 pp. 533-558.

[20] M. Baldauf, S. Dustdar, and F. Rosenberg, "A Survey on Context-aware Systems," International Journal of Ad Hoc and Ubiquitous Computing , vol2 (4), June.2007, pp. 263-277.

[21] M. Bouzeghoub and D. Kostadinov, "Data Personalization: a Taxonomy of User Profiles Knowledge and a Profile Management Tool," Reports of the laboratory PRISM, http://www.researchgate.net/publication/228768551_Data_Personalization_a_Taxonomy_of_User_Profiles_Knowledge_and_a_Profile_Management_Tool/file/9fcfd50cf300649190.pdf, 2007 [retrieved:April, 2013].

[22] G. Bordogna and G. Pasi, "A Flexible Multi Criteria Information Filtering Model," Soft Comput. DOI: 10.1007/s00500-009-0476-3. January.2010, pp. 799-809 [retrieved:April, 2013].

[23] P. J. Brown, J. D. Bovey, and X. Chen, "Context-aware Applications: From the Laboratory to the Marketplace," IEEE Personal Communications, October.1997, pp. 58-64.

[24] N. Ryan, J. Pascoe, and D. Morse, "Enhanced Reality Fieldwork: the Context Aware Archaeological Assistant," Dingwall, L., S. Exon, V. Gaffney, S. Laflin and M. van Leusen (eds.), Archaeology in the Age of the Internet. CAA97. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 25th Anniversary Conference, University of Birmingham, (BAR International Series 750). Archaeopress, Oxford, April.1997, pp. 269-274.

[25] A. K. Dey, G. D. Abowd, and D. Salber, "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-aware Applications," Human-Computer Interaction Journal 16(2), December.2001, pp. 97-166.

[26] N. Kokash, A. Birukou, and V. D'Andrea, "Web Service Discovery Based on Past User Experience," in Proceedings of the International1 Conference on Business Information Systems (BIS), LNCS Volume 4439, 2007, pp. 95-107.

[27] Y. Tsai, S. Hwang, and Y. Tang, "A Hybrid Approach to Automatic Web Services Discovery," in International Joint Conference on Service Sciences, IEEEXplore, 2011, pp. 277-281.


 ISBN: 978-1-61208-280-6 