Beyond the Social Search: Personalizing the Semantic Search in Social Networks

Júlio Cesar dos Reis[1,2], Rodrigo Bonacin[2,3], and M. Cecília C. Baranauskas[1]

[1] Departament of Information System at Institute of Computing
University of Campinas (UNICAMP) - 13083-970, Campinas, SP, Brazil
[2] Center for Information Technology Renato Archer (CTI) - Rodovia Dom Pedro I, km 143,6, 13069-901, Campinas, SP, Brazil
[3] Faculty of Campo Limpo Paulista (FACCAMP) - 13231-230, Rua Guatemala, nº 167 Jardim América, Campo Limpo Paulista, SP, Brazil
{julio.reis,rodrigo.bonacin}@cti.gov.br, cecilia@ic.unicamp.br

Abstract. Nowadays search mechanisms are decisive to people successfully access relevant and valuable information. In the current Web environment, people connected to each other through Social Network Services (SNSs) bring new challenges for search approaches. Recent literature in the SNS field shows expressive evolution but the approaches for searching still utilize mechanisms that just consider the syntactical processing of the information. Based on socio-cultural aspects of a SNS, we propose a novel approach that enables to provide personalized semantic search. This approach makes use of ontologies to represent the shared meanings, and consequently to present better organized, ranked and adapted search results for each SNS user. The paper presents the way the proposed mechanism was implemented and points out examples that illustrate the possibilities raised by the approach.

Keywords: Semantic Search, Web ontology, Search Personalization, Organizational Semiotics.

1 Introduction

SNSs play a vital role in the Web environment nowadays. Since their introduction, Social Networks such as *MySpace, Facebook,* and *Orkut* have attracted millions of users, many of whom have integrated these sites into their daily practices [1]. In SNSs people interact making connections and constituting communities; they produce information, communicate, collaborate, and share their objectives, interests, activities and practices. SNSs emerged as a propitious virtual space to access and share information, representing an opportunity to share experiences and knowledge [2].

The access to the digital content generated in these applications depends on resources such as search mechanisms. In order to deal with the limitations of traditional search mechanisms, literature has proposed a set of approaches and techniques over the years, such as: semantic search, personalized and adaptive search, personalized information filtering, collaborative information seeking, as well social

A.A. Ozok and P. Zaphiris (Eds.): Online Communities, HCII 2011, LNCS 6778, pp. 345–354, 2011.

search. Some works have proposed to enrich the search mechanisms using users' preferences and characteristics, while others proposed to create ontology-based search engines. However, usually such solutions are not integrated, loosing information about how to make a bridge between user's preferences and characteristics to the semantic models used. Moreover, most of the search solutions for SNSs are directed to the search of users or experts grounded on friendship and reputation relationships. Certainly these solutions for social search are important; nevertheless they can be improved by providing a personalized semantic search for the content produced in the SNS. Besides friendship relationships, the shared contents and their meaning should be considered by the search mechanism.

While SNSs key technological features are fairly consistent, the cultures that emerge around them are varied [1]. Actually, people in SNSs share much more than just friendship relationships; they also share culture, language and meaning. People behave according to commitments, linguistic constraints, culture and other social aspects. Therefore, the main challenge in this context is to have methods that can deal with these aspects, being able to model the shared meanings from the SNS, and to explore it to improve personalized search results. The current search solutions for SNS do not take into account the intrinsic semantic aspects that come from the network's activity itself (*i.e.* they do not consider the meanings of words used by people in the SNS). Usually semantic search solutions are based on computational representation models disconnected of the reality (*i.e.* they do not include the social actors that interpret the concepts). Thus, a search mechanism capable of "interpreting" the shared meanings used by people in the SNS is necessary. For that, we need methods and approaches that allow to capture and to represent social and cultural aspects from the SNS such as their colloquial language.

This paper brings a semantic search solution for SNS able to deliver personalized semantic search results not just focused on people, but also on the content created by them in the SNS. Based on a semantic model created from the SNS contents, the results are adapted according to the user that performs the search. Since search results personalization are very dependent on who performs the search, the proposed solution considers that the meaning of the words depends on who produces the information and also who searches for it. Consequently, the solution is based on a semantic model that links the concepts meanings from who produces and who searches the information. In order to reach it, the proposal uses a Semiotic-based approach [4] in the modeling process resulting on Web ontologies (WOs) that enable search results more adequate for users of that context. This represents a novel approach for ontology-based search in SNS, grounded on semantic modeling aspects and techniques that can reach more adequate and personalized search results. The solution is implemented and illustrated in the *VilanaRede*[1] SNS.

The paper is organized as follows: Section 2 reviews proposals for Web search strategies and situate them into SNS context. Section 3 describes the methods used in a Semiotic-based approach to personalize semantic search results in SNS; Section 4 presents the design and implementation issues of the search mechanism developed at *VilanaRede,* illustrating the search strategies with examples as well the software architecture used to explore the semantic aspects of the semantic model. Finally, we

[1] www.vilanarede.org.br

conclude the paper with a brief discussion, including a summary and directions for further work in Section 5.

2 Reviewing Search Approaches and Strategies Situated in SNS

Literature has presented a number of approaches and solutions to improve search mechanisms (both in the Web and also in SNS) based on various strategies, such as: recommendation, trust, as well as semantic and personalization aspects. In this section we briefly explain and review each of these.

2.1 Social Search

Usually the studies of SNS search focus on search for SNS users. The work conducted by Yu & Singh [5] proposes a solution to find among participants of a SNS, specialists in some domains. Their proposal makes use of multi-agent system techniques. Based on users' information regarding their social network, artificial agents find people based on reputation. Vieira *et al.* [6] argue that the friendship relationship is an important evidence to find people in a SNS, and the search results could be modeled in function of a friendship graph. Algorithms in graphs try to find the shortest path to friends, aiming at ordering the users' profile during the search. More recently, other approaches have proposed to use aspects of recommendation as well as trust to find items in SNS. For instance, Gürsel & Sen [7] propose a search system to find items based on social network concept. This system uses a framework based on artificial agents that mine the group of relationships of a user on the network. The objective is to facilitate the search for items of interest through the use of social connections. As a consequence, it is possible to rank the search results highlighting the items of preference recently posted by friends of the SNS, which will be of particular interest to a user. Other approaches stress the use of cognitive authority from the SNS for information retrieval in folksonomy-based systems [8], in order to reach information quality.

2.2 Semantic Search

The goal of optimizing search results has motivated research in the semantic search field. The need to retrieve information semantically enriched gave rise to an increasing interest in research on ontologies. According to Wei *et al.* [8], semantic search extends the scope of traditional information retrieval (IR) paradigms from mere document or people retrieval to entity and knowledge retrieval, improving the conventional IR methods by looking at a different perspective, *i.e.* the meaning of words.

In order to provide search mechanisms with semantic features, it is necessary to interpret a knowledge model about the domain. The knowledge can be represented by means of ontologies. A wide adopted definition for ontology in Computer Science is done by Gruber [10]. The ontologies are usually described by computational ontology languages such as *Web Ontology Language* (OWL) [14]. Works such as Heflin & Hendler [11], Guha *et al.* [12] among others have introduced the idea of using

ontologies by semantic search mechanisms. Considering the SNS context, Choudhari *et al.* [13] describe an architecture to develop semantic search for SNS. However, instead of using ontologies they used an approach based on clustering of terms semantically related to perform the semantic search.

2.3 Personalized and Adaptive Search

There are also a number of studies for personalizing Web search that develop strategies of adaption of the search results according to user's characteristics. Many works claim that the users' interests, preferences and need are an essential element in identifying the user's context for the search. Various approaches employ user modeling components based on the users' interactions, such as: previous queries or Web navigation activities history (*e.g.* visited pages). In this context, some works analyze implicit information about the user's interests [*e.g.* 22, 21, 22, 24]. However, each one tackles a single facet of personalization, either biasing search results to general user-interests [22] or respecting the current search session's context [21], or using both strategies [*e.g.,* 24]. Some personalization techniques not only consider a single user, but also take into account the actions of a surrounding group of users [*e.g*., 22] through a collaborative filtering approach. Another alternative to address personalization is the use of ontologies. The proposals include the (semi) automatic generation of semantically enriched ontology-based user profiles. Ontology-based personalized search approaches [*e.g.* 25] involve building models of user context as ontological profiles, by assigning implicitly derived interest scores to existing concepts in the domain ontology. Moreover, recent literature has presented adaptive ontologies process for domain evolution [*e.g.,* 25] in adaptive ontology-based Web search.

3 A Semiotic-Based Approach for Search Personalization in SNS

Semiotics [15], the ancient doctrine of signs, leads us to a precise understanding of information as various properties of signs. Signs are simple entities easy to deal within the inter-subjective domain. Anything standing for another thing or used to signify something is an example of sign: words, traffic lights, diagrams, a wave of hand or a facial expression. From a Semiotic point of view, people communicate in SNS by sharing signs, and such signs are socially constructed and shared over interaction. In this sense, we have to understand the signs and how they influence the communication process. By studying the use of signs in the organized behaviour, Organisational Semiotics (OS) [3] theory and methods support us in understanding and modeling social aspects in SNS.

More adequate search mechanisms for SNS should take into account the user's language aspects and their applied meanings aiming to deliver search results that make sense to them [16]. For that, meanings must be represented as they are used in the SNS context; thus, computational models (*e.g.* ontologies) capable of representing the semantic of the language that come out from the SNS are necessary. The proposed method, developed to construct more representative WOs that come out from SNS content, is based on the Semantic Analysis Method (SAM) [8]. SAM delineates the

area of concern of a social context and identifies the basic patterns of behaviour (affordances) of their agents [3]. It produces a stable Ontology of the context. This Ontology describes the semantic aspects of the signs shared in a SNS.

A WO that represents the semantics of the SNS application should contain the agents that interpret the socially shared signs. This approach incorporates and takes to WOs concerns, representations arising from a Semiotic perspective [4]. In this process theoretical and methodological concepts of SAM are used in conjunction with technologies from the Semantic Web (SW) initiative to describe computationally tractable ontologies using OWL. The proposed method is based on a subjectivist paradigm [3], which understands reality as a social construction based on the behaviour of agents participating on it, instead of an objectivist paradigm which presupposes that there exists a world independent of the observer.

SAM uses an agent-in-action ontology, determining the underlying semantics of a social context and the relationship between the human agents and their patterns of behaviour [3]. SAM supports the analysis, specification and representation of a social system assisting users or problem-owners in eliciting and representing their meanings in a formal and precise semantic model - the Ontology Chart (OC). OC is a graphical representation of a conceptual model that describes a view of responsible agents in the focal domain, including their pattern of behavior named affordances and the ontological dependences between them [3].

The most important concepts in SAM are Affordance and Agent. Affordance is a concept introduced by Gibson [17] that can be used to express the invariant repertories of behaviour of an organism made available by some combined structure of the organism and its environment. In SAM the concept introduced by Gibson was extended by Stamper [18] to include invariants of behavior in the social world; affordances are social constructs in a certain social context. The Agent is a special kind of affordance, which can be defined as something that performs responsible behaviour. An agent can be an individual person, a cultural group, a language community, a society, etc. (an employee, a department, an organization, etc.).

SAM provides a different and independent view of the social context under study addressing issues that are not represented in any traditional WO. Moreover, SAM allows us to discover the agents and their respective pattern of behaviour (affordances) within a SNS. In the process to generate OWL ontologies from the SNS content, the OCs inform the SNS semantic search mechanism. Figure 1 illustrates the method to create OWL ontology from SNS content through SAM. The transformation from OC to OWL is carried out by an assisted process [19]. A case study applying the entire process based on real contents of a SNS is described in [20].

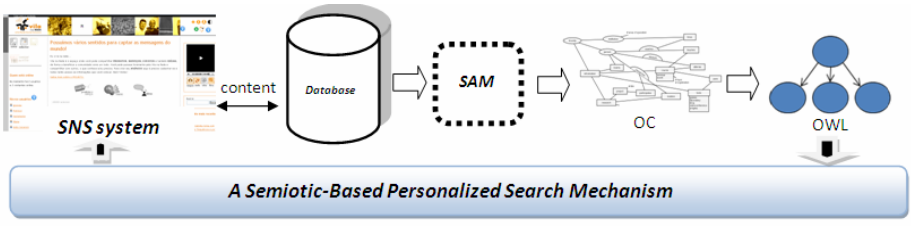

Fig. 1. Informing the Search Mechanism with WO built from SNS content through SAM [adapted from 16]

4 Providing Personalized Semantic Search Results in SNS

The agents modeled during the proposed method (that also composes the OWL ontology) are adequately associated to each SNS user. The key point for the proposed personalization strategy is to use information about the social agents as a context for the search. Thus, the possible meanings of a word are connected to agents in both situations, during the search time and during the processing of the search index of the contents generated by the SNS users. Thus, the disambiguation of possible polysemic words is led by the agents that interpret the terms. For example, considering a query as "Java", some users may be interested in documents dealing with the programming language, while others may want documents related to "coffee" or touristic information (Indonesian island). Therefore, since a SNS user may be connected to an agent "programmer", the meaning applied will be "programming language", as well whether the user is connected to an agent "tourist", the meaning applied would be other. The relationship between the users with the possible social agents can be reached following various strategies. First, the users can point out some agent that most fits their profile; in addition techniques can be applied trying to find the possible agents for a user based on their activity on the SNS (*e.g.* the contents posted). Moreover, the social network can also be used to point out such agents.

Based on the agents, semantic rules described by Semantic Web Rule Language[2] (SWRL) are created to infer the correct meanings. Such rules make a matching between a possible term and an agent inferring as meaning a class of the OWL ontology. *E.g.* if the agent is a 'programmer' and the term is 'java' then the meaning (class) is 'programming language'. These rules are used in both cases to classify search keywords during the search time, as well during the processing of the search index formed based on the SNS contents (*e.g.* announcements). Each of these is processed using the rules that classify each term of the content forming an inverted index search. In this index a term would have different classifications (classes) that point out to various contents in the SNS, since the same term can have different meanings treated by the rules based on the agents. For example, the combination "Java" (programming language) may point out to some content in the index, and another combination "Java" (tourism) may point out to other content.

This classification is used to organize the search results and to personalize it. Rather than providing the search results as a single ranked list, we propose a structured organization for the search results based on boxes (see Figure 2). In the User Interface (UI) for the search results, the semantic results are organized into two main boxes, embedding other boxes. The first box has the title 'Personalized' and the second has the title 'Other'; both have the same intern organization as presented in Figure 2. The internal boxes organize the semantic results according to the expansion of the queries in the ontology-based search. The words found during it are queried in the index to find semantic search results. The most relevant are placed right above the internal boxes. These results strictly match the keywords considering the semantic and syntactic criteria. The internal boxes may contain results with 'separated keyword', once the search engine could not find results containing exactly the combination of all keywords; there is also a box that keeps the synonymous (the synonymous terms are also modeled by the ontology); and a box with 'more specific

[2] http://www.w3.org/Submission/SWRL/

results' and 'more generic results' that contain results of generic-specific relation (is-a) from the OWL ontology; and also a box of 'related results' got from relationship between concepts in the ontology, for instance, ontological dependence relationships.

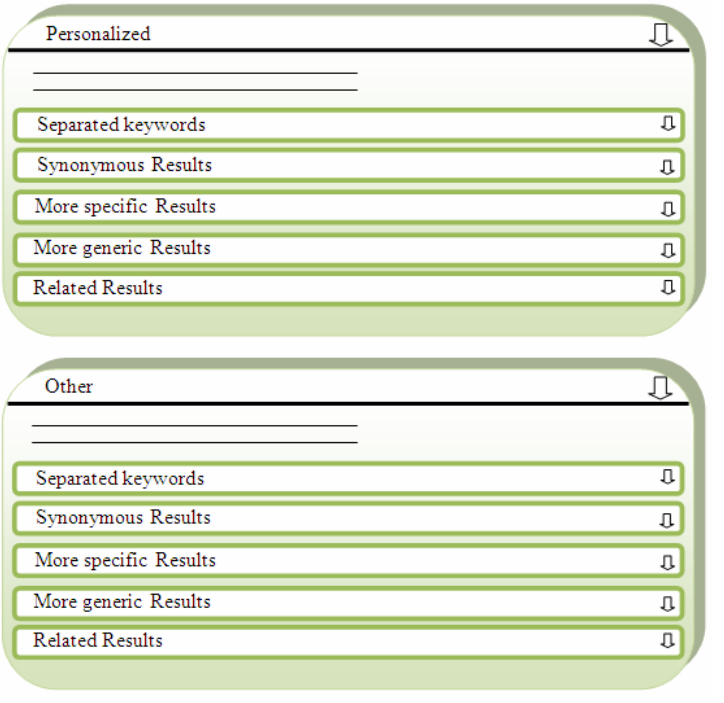

Fig. 2. Organization for the Personalized Semantic Search Results

For the "Java" example, supposing a "programmer agent" performing the search, all the results concerning "Java" (as programming language) as well as the synonymous and generic-specific results in the same context must be placed into the 'personalized' box and into its internal boxes. All other possible search results, with other meanings or without any recognized meaning are placed into the 'Other' box. Since both boxes and internal boxes results are hidden by default, the users can explore the results in a small initial list, and expand it as they like.

Additionally to this approach to search by the meanings, other search results items may also be ranked into the 'Personalized' box using an agents approach. Since the agents are connected to all SNS users, in this approach, the items that were posted by users that contain some agent equal to the agent of the user that performs the search are also considered as personalized results.

The proposed search mechanism was implemented at *VilanaRede* SNS. This SNS was developed as a product of *e-Cidadania's*[3] Project, with the objective of being

[3] www.nied.unicamp.br/ecidadania

accessible for all, including those users less familiar with technology or with low literacy levels. The OWL ontology that informs this search mechanism was built based on the *VilanaRede* contents. The *VilanaRede* system uses the Content Management System Drupal 5[4] to manage users' accounts, content types and other general functionalities; the proposed semantic search service is hosted in a JBoss Application Server[5] as an external and independent Web Service.

The implemented Web Service manages all the necessary information and provides all services to handle the search request, such as: the index processing, the adding of agents and user-agent relationship as well as the ontology management. The search service includes the search algorithm that provides the search results by lists. The organization of these lists is performed at the Drupal side. Figure 3 presents an example of the *VilanaRede* UI after the search request processing.

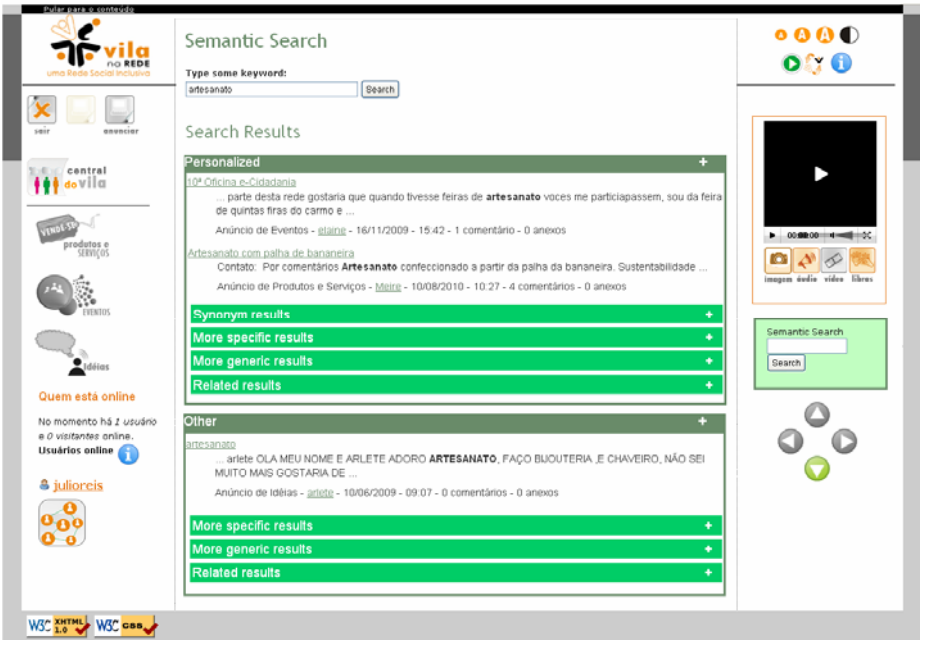

Fig. 3. *VilanaRede* User's Interface showing Personalized Semantic Search Results

5 Final Considerations

Most of the information available in SNSs are in a format that can not be directly interpreted by machines. Also, most of the search mechanisms have not been designed to interpret or to represent the meanings of the terms. They have been fundamentally constructed based on the approach "one fits all" and on keywords comparison with lexical-syntactic processing, *i.e.* do not considering semantic

[4] Drupal. `http://drupal.org`

knowledge. Consequently, the mechanisms do not have the capacity to differentiate the relevant search results from the semantics point of view. In this context, ontologies may be a fundamental cornerstone for designing useful semantic search services.

The personalization approaches for search described in literature are based on topics of preferences and on users' needs. These topics are discovered and modeled mainly by browse history and past user's search queries analysis. Even the proposals that consider semantic knowledge to create the user profile use a number of calculations and scores that do not model the users' language. In personalization approaches using the users' social network, the search results are mainly ranked with items viewed or highlighted by people that compose their personal network. We have observed that it is necessary a novel way to personalize and to organize the search results by taking into account the SNS semantic aspects. In our approach the users' social characteristics are considered in the adaptation of the semantic search. Results are given according to responsible social roles, *i.e.* the agents identified during the semantic modeling process by SAM (*e.g.* social identity of the users as their profession). The agents determine the meanings of the words used in the index and during the search.

Therefore, our approach provides means to discover as well to distinguish the meanings used by people at the SNS through the agents represented into OCs. Such agents are used by semantic rules described in SWRL in the ontology to get adequate meaning for the words. Based on that, we presented a semantic search mechanism implemented at *VilanaRede* SNS. This mechanism is able to personalize and better organize the semantic search results according to the user that performs the search, presenting first the most relevant and meaningful search results. The mechanism considers both lexical and conceptual aspects (*e.g.* generalization, specialization and related results). Further works involve conducting a case study with real users using the mechanism at *VilanaRede* in order to discuss experimental results.

References

1. Boyd, D.M.: Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication, 210–230 (2008)
2. Jamali, M., Abolhassani, H.: Different aspects of social network analysis. In: Proc. of 2006 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 66–72. IEEE Computer Society Press, Washington, DC (2006)
3. Liu, K.: Semiotics in information systems engineering. Cambridge University Press, Cambridge (2000)
4. Reis, J.C., Bonacin, R., Baranauskas, M.C.C.: A Semiotic-based Approach to the design of Web Ontologies. In: ICISO 2010, Reading, UK, pp. 60–67 (2010)
5. Yu, B., Singh, M.P.: Searching social networks. In: Proc. of the 2th International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 65–72. ACM Press, New York (2003)
6. Vieira, M.V., Fonseca, B.M., Damazio, R., Golgher, P.B., Reis, D.C., Neto, B.R.: Efficient Search Ranking in Social Networks. In: Proc. of the 16th ACM Conference on Information and Knowledge Management, Lisbon, Portugal, pp. 563–572 (2007)

7. Gürsel, A., Sen, S.: Improving search in social networks by agent based mining. In: Proc. of the 21st International Jont Conference on Artifical Intelligence Table of Contents, Pasadena, California, USA, pp. 2034–2039 (2009)
8. Pereira, R., Silva, S.R.P.: The Use of Cognitive Authority for Information Retrieval in Folksonomy-Based Systems. In: ICWE 2008, Yorktown Heights, NJ, pp. 325–331 (2008)
9. Wei, W., Barnaghi, P.M., Bargiela, A.: Search with Meanings: An Overview of Semantic Search Systems. Int. Journal of Communications of SIWN 3, 76–82 (2008)
10. Gruber, T.R.: A translation approach to portable ontologies. Knowledge Acquisition 5(2), 199–220 (1993)
11. Heflin, J., Hendler, J.: Searching the web with SHOE. Artificial Intelligence for Web Search. In: AAAI Workshop, WS-00-01, pp. 35–40. AAAI Press, Menlo Park (2000)
12. Guha, R., McCool, R., Miller, E.: Semantic Search. In: Proceedings of the 12th International Conference on World Wide Web, Budapest, Hungary, pp. 700–709 (2003)
13. Choudhari, A., Jain, M., Sinharoy, A., Zhang, M.: SmartSeek: A Semantic Search in Social Networks (2008), http://www.cc.gatech.edu/projects/disl/courses/8803/2008/project/project_deliverables/group22/final_report.pdf (accessed, January 2011)
14. Web Ontology Language (OWL). W3C Recommendation (February 4, 2004), http://www.w3.org/TR/owl-features (accessed, January 2011)
15. Peirce, C.S.: Collected Papers. Harvard University Press, Cambridge, Mass (1931-1958)
16. dos Reis, J.C., Bonacin, R., Baranauskas, M.C.C.: Prospecting an Inclusive Search Mechanism for Social Network Services. In: Filipe, J., Cordeiro, J. (eds.) ICEIS 2010. Lecture Notes in Business Information Processing, vol. 73, pp. 555–570. Springer, Heidelberg (2011)
17. Gibson, J.J.: The Theory of Affordances. In: Shaw, R., Bransford, J. (eds.) Perceiving, Acting, and Knowing (1977)
18. Stamper, R.K.: Social Norms in requirements analysis – an outline of MEASUR. In: Jirotka, M., Goguen, J., Bickerton, M. (eds.) Requirements Engineering, Technical and Social Aspects. Academic Press, New York (1993)
19. Reis, J.C., Bonacin, R., Baranauskas, M.C.C.: From Ontology Charts to Web Ontologies: Heuristics and Transformation Rules (IC-11-02), IC/UNICAMP (2011)
20. Reis, J.C., Baranauskas, M.C.C., Bonacin, R.: Modeling Meanings from Inclusive Social Network Services (IC-11-03), IC/UNICAMP (2011)
21. Shen, X., Tan, B., Zhai, C.: Implicit user modeling for personalized search. In: CIKM (2005)
22. Teevan, J., Dumais, S.T., Horvitz, E.: Personalizing search via automated analysis of interests and activities. In: ACM SIGIR, Salvador, Brazil, pp. 449–456 (2005)
23. Sugiyama, K., Hatano, K., Yoshikawa, M.: Adaptive web search based on user profile constructed without any effort from users. In: WWW, New York, USA, pp. 675–684 (2004)
24. Elbassuoni, S., Luxenburger, J., Weikum, G.: Adaptive Personalization of Web Search. In: 1st WS on Web Information Seeking and Interaction, Amsterdam, The Netherlands (2007)
25. Sieg, A., Mobasher, B., Burke, R.: Web Search Personalization with Ontological User Profile. In: CIKM, Lisboa, Portugal, pp. 525–534 (2007)
26. Guelfi, N., Pruski, C., Reynaud, C.: Experimental Assessment of the TARGET Adaptive Ontology-based Web Search Framework. In: 10th International Conference on New Technologies of Distributed Systems, Tozeur, Tunisia, pp. 297–302 (2010)