



INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

**Analysing Search in Inclusive Social Network
Services**

Júlio César dos Reis Rodrigo Bonacin
Maria Cecília Calani Baranauskas

Technical Report - IC-10-02 - Relatório Técnico

January - 2010 - Janeiro

The contents of this report are the sole responsibility of the authors.
O conteúdo do presente relatório é de única responsabilidade dos autores.

Analysing Search in Inclusive Social Network Services

Júlio César dos Reis*[†] Rodrigo Bonacin[†] Maria Cecília Calani Baranauskas*

Abstract

Social Network Services are a reality nowadays. These systems represent a propitious virtual environment for user interaction and communication, and an opportunity for people to share information and knowledge. Considering socio-economical aspects, these systems should provide inclusive access for all. The purpose of Inclusive Social Network Services (ISN) is to create situations where users' diversity is respected and the access difficulties minimized. The use of search engines is the principal alternative to find and to access information generated in digital media. Therefore is important that all people have the possibility to recover information in a natural way, with results that make sense to them. This research report investigates the search process in an ISN based on the 8th Semio-Participatory Workshop of the e-Cidadania project¹. The activity involved the observation of prospective users of an ISN in a set of search scenarios; the aim was to observe how users behave and their difficulties with the searching mechanism. We point out preliminary results that can lead to better search engines from an inclusive perspective.

1 Introduction

With the advance in Web technology a new software category appeared: the social software. This software introduced new opportunities for sharing information and knowledge among individuals. In this category are the Social Network Services (SNS). These systems primarily allow individuals to share their interests and activities, constituting communities with common interests. The social networks represent an opportunity to interaction and access to information and knowledge through the Web.

The Brazilian society faces today a situation characterized by enormous differences regarding socio-economics, culture as well as access to technology and knowledge [1]. According to the 4th Grand Challenge of the Brazilian Computing Society (SBC)², there exists no previous experiences to get inspiration from; the challenge is unique: to let the Brazilian citizens benefit from Information and Communication Technology, via user interfaces thus promoting the process of the constitution of a fairer society, embracing different competencies and needs among people [2].

*Institute of Computing, University of Campinas, 13081-970 Campinas, SP – Brazil

[†]CTI, MCT, Rodovia Dom Pedro I, km 143,6, 13069-901 Campinas, SP - Brazil

¹A FAPESP-Microsoft Research Institute funded Project.

²www.sbc.org.br

This report is a result from research activities in the context of the e-Cidadania project [1]. This project proposes solutions to deal with barriers that the majority of Brazilian citizens face to effectively use the Internet. The e-Cidadania aims at transforming Social Networking Systems in an engine for digital inclusion and citizenship. The networks with such characteristics can be understood as “Inclusive Social Network Services” (ISN). The main idea of the e-Cidadania is to consider the diversity of users and to take the digitally excluded into account (including illiterate and people with special needs). The Project investigates mainly the relationship that people establish in their informal communities organized around some special interests, how they use societal artifacts, including computational technology [4]. E-Cidadania is also about designing computational systems for ordinary people, especially those unfamiliar with the Information and Communication Technology (ICT), the vast majority of the Brazilian population. In order to realize the concept of ISN the information and knowledge must be effectively accessed and shared. For that it is necessary a common understanding of the parties and more efficient and accurate techniques to recover and access information.

This Technical Report investigates searching in ISN. This topic is important given that search engines are primary mechanisms to enable the access and use of information generated in Social Networks. Besides, many difficulties encountered by users (information overload and problems of context) can not be resolved by syntactic and usual search engines. Therefore is extremely important to seek for new mechanisms for information retrieval, taking into account the vocabulary and meanings created, shared and used by Social Network users. A search engine must be inclusive too, and reflect the “semantic reality” of the context, contributing to the accomplishment of the participatory and universal access to information. As a first step, this report shows the results of search activities in ISN conducted at the 8th Semio-Participatory Workshop of the e-Cidadania project. The goal of these activities was to observe a set of search scenarios with potential users of an ISN, to understand how these users make sense of a search mechanism.

This Technical Report is organized as follows: Section 2 presents the concept of ISN and its importance for the universal access to information; Section 3 presents the search scenarios and the analysis made on them; Section 4 presents a discussion about the results and finally Section 5 concludes.

2 Inclusive Social Network Services and the Universal Access to Information

According to Boyd & Ellison [5] since the introduction of the Social Networking sites such as MySpace³, Facebook⁴, Orkut⁵ and others, these have attracted millions of users and many of them have integrated these sites into their daily practices. Today there are hundreds of these sites with various technology supports including a wide range of interests and practices. Although their main technological characteristics are fairly consistent, the cultures that

³www.myspace.com

⁴www.facebook.com

⁵www.orkut.com

emerge around these social networks are varied. Some sites attend to different audiences, while others attract people based on common language, racial identity, gender, religion or nationality, among other topics.

The Social Networking sites are Web-based services that allow individuals: 1) to construct a public or semi-public profile within the limits of the system, 2) to articulate a list of other users with whom they share connections, and 3) to view and to browse their list of connections and those built by other users within the system. The nature and description of these connections (the type of relationship) can vary from site to site [5].

Online Social Networks or “communities of members” have great relevance in the Web as users spend much time of navigation on them. As shown by the data presented by Nielsen [6]: (1) Brazil is the country with the largest number of Internet users using social networking sites; (2) 80% of Brazilian Internet users are linked to sites of “community of members” - blogs and Social Networks, (3) Brazilian Internet users are those that most spend time in this type of site. Equivalent to 1 to every 4 'minutes of surfing in the Internet, (4) In Spain 75% of Internet users use social networks; in Italy 73% of users access social networks and in Japan 70% of users communicate in social networks. Also according to Nielsen [6], social networks are more popular than e-mail, with 66.8% of global reach. Around the world it represents the fourth most used resource at the Internet and 85.2% of penetration are in the portals and communities of general interest. Additionally, 85.9% of internet users use search engines, which is one of the most wanted activity.

Despite these great numbers and the success of Social Network sites among Internet users, there are yet many people without access to Internet and consequently without opportunities to access information and knowledge. According to Baranauskas [1] we live in a scenario of vast differences regarding socio-economic, cultural, which impact in the access to technology and knowledge. Social Indicators shown by the PNAD⁶ (National Survey by Household Sample) produced by IBGE (Brazilian Institute of Geography and Statistics) points out that in 2008, 65% of the population did not have access to Internet. The IBGE [7] also presents that 14.3 million of Brazilians or 10% of the population is illiterate (considering people over 15 years old).

In addition, important data from Ministry of Education in Brazil [8] reveal that about 30 million of Brazilians are functionally illiterate, defined as the population with over 15 years old and less than 4 years of schooling (21.6% of the population). Using a broader concept of functional illiteracy, according to a survey from Paulo Montenegro held in 2007 [9] (entity created by IBOPE⁷) and by the Non-Governmental Organization (NGO) “Ação Educativa⁸”, the majority (64%) of Brazilians between 15 and 64 years who have more than 4 years of schooling reach no more than the degree of rudimentary literacy, that is, they have no more than the ability to locate explicit information in short texts or make simple math, but they are not able to understand longer texts. And even worse: 12% can be considered absolute illiterates in terms of reading / writing, failing to even decode words and phrases. Besides, only 6% of them use computers, but 52% say that they read newspapers and 48% read magazines. From our population recorded by the last Census of 2000, we

⁶<http://www.ibge.gov.br/home/estatistica/populacao/acesoainternet2008>

⁷<http://www.ibope.com.br>

⁸Approximated Translation to English: “Educational Action”

have 24.5 million of people with some type of disability⁹ (14.5% of the population). This data illustrates only part of the challenge that we face in terms of designing systems that should include all these users. In this context it has become a major concern how to allow unrestricted access to online content available from Social Network Systems to all people in a more “natural” and efficient way.

Therefore, the design for inclusion involves considering all the diversity of users and contexts in the development of information systems in general, and their interfaces in particular. There is an urgent need of methods that are based on a deeper understanding of the concept of inclusion and of the differences [1]. Thus, it is extremely important to create more sophisticated methods to allow the effective access and use of information conveyed in digital media for all. This could benefit from the ISN concept.

According to Hayashi [4], the complexity of developing a system to support ISN that make sense to these users requires a socio-technical vision of the problem and a participatory and inclusive approach for the proposal of solutions. Hence, it is extremely important to create spaces and means for people be included in a natural way to the access and use of technology and its digital content. According to Baranauskas [1] solutions for digital inclusion should be treated as instruments of deep transformation in society; among the possible solutions are ISN. We understand ISN as a “virtual communication space” based on the concept of social networks, which is inclusive and allows the community to share knowledge about the community know-how. This space has to be generalized in order to facilitate “exchange” (of knowledge, goods and services) in accordance with the collaborative (project team, partners, community) system conception. Some definitions of the concept of ISN that have been written by potential users, people with different levels of education and experience [4] are as follows:

1. *A system that gathers subjects, ideas, information that are of interest for the diverse and different users that exist in our society. And that this information exchange and this interaction among users should benefit all communities.*
2. *A group of people that interacts sharing different elements without discriminating participants, that is, when we mention “inclusive” that means that everyone is part of that network and that the network has a common objective.*
3. *Inclusive social networks are a space - not necessarily physical, that makes it possible the exchange among people. These exchanges can be of information, products, etc.*
4. *An Inclusive social network is a way to “connect” common interests of people/communities without offering barriers to the participation of the people that are part of the network (or that can become part of it), without forgetting that there are rules for participating/acting.*

It is worth to mention that in an ISN there are not target users, but all users are relevant and should be included without discrimination. Thus, there may be users without skills to handle certain technological features of the system and consequently without knowledge

⁹<http://mj.gov.br>

to find information that they need in the system. In addition, those users most often use colloquial terms to express themselves through the system. For example, they may use the term “*postinho*”¹⁰ instead of “Basic Health Unit” (formal). They use terms that make sense to them, but in fact, these expressions semantically mean the same. So when someone is trying to retrieve information from the ISN, these factors must be taken into account by the search engine. On the other hand, when a user searches for something in a non-formal or not refined way and the same concept but in its formal way (cult) is returned, it is an opportunity for learning.

Thus, we should seek for a computational search solution that takes into account the meaning that is adopted or emerges in the context of use of that network; i.e. the meaning that people bring to the network, and those that are constructed by using the system over time (through interaction). This may facilitate and provide better access to the content generated by users of the network. Results from the e-Cidadania workshop point out in this direction, as detailed in the next section.

3 Analysing Search Scenarios of an Inclusive Social Network

From a practical point of view, the e-Cidadania project resulted on a system called “Vilanarede”¹¹. This system represents an opportunity to work with a set of users with typical characteristics of the Brazilian population. Baranauskas *et al.* [1] describe it as a Scenario* (Star Scenario) that tries to represent the diversity of users of the Brazilian context. From the computational point of view, the Vilanarede system provides basic functionalities for this research.

As a direct activity of the project, we have conducted the 8th Semio-Participatory Workshop, which took place at the “*Centro de Referência da Juventude*”¹² / “*Casa Brasil*”, which is a telecenter located at ‘Vila União’ neighborhood, of the Campinas city, State of São Paulo - Brazil. In this workshop we developed an activity related to search in ISN. The purpose of this activity was to explore a set of search scenarios with potential users of an ISN - the Vilanarede system. The objective was to observe some major points including: (1) how would the users react to the search engine, which key-words would they use? (2) Would they have some difficulty in completing the proposed scenarios? And (3) What would be their impressions about the search results. We aimed at analyzing the reason and importance of new and better search mechanisms for ISN.

A task sheet was elaborated (see Appendix A) for each pair of participants of the Workshop. The task had 4 search scenarios with one more extra called Scenario X, and a form (see Appendix B) to an observer (researcher) of the activity. We had 7 pairs of people in total. An initial instruction about the activity was given for the participants. The pairs were formed by the users themselves, and in each scenario every pair should write the words used to make the search and the title of the announcements found from the result. As a direct artifact resulting of this activity, we have both the sheet tasks from the pairs and the

¹⁰in Portuguese

¹¹www.vilanarede.org.br

¹²Approximated translation to English: “Youth Reference Center”

observation forms filled out by the researchers. Besides, the activity was filmed and there was audio recording of each pair during the task execution. The 4 scenarios of the sheet task were:

- **Scenario 1:** Find announcements on how to popularize the ‘VilanaRede’.
- **Scenario 2:** Find announcements of mango (fruit) at ‘VilanaRede’.
- **Scenario 3:** Find announcements related to food at ‘VilanaRede’.
- **Scenario 4:** Check for some announcement related to religion combined with handi-craft at ‘VilanaRede’.

The goal of these scenarios was to verify whether users would have trouble in finding the announcements of products/services, events or ideas at the ‘VilanaRede’ system through a syntactic search engine (built primarily through comparisons of key-words and lexical-syntactic information processing, different from semantic search that take into account the meanings of the words in order to generate more relevant results). Thus, each scenario was designed to verify whether semantic capacity was needed for the search mechanism. The time for the completion of the scenarios was approximately 45 minutes. After the execution of the search scenarios, there was a discussion in order to verify the general impression from users about the activity. At this time, several interesting stories were collected.

In Scenario 1 (Find announcements on how to popularize the ‘VilanaRede’) we wanted to observe whether users would use synonymous of “popularize” to find the announcements. Some pairs had difficulty to understand the scenario, as well as difficulty in choosing the terms for the search. However some pairs linked the word “popularize” to “divulge” and quickly found related announcements. At this scenario one pair used some unusual key-words such as: “*boca-a-boca*”¹³ (word-of-mouth), “email”, “phone” and “posters”. By using the term “word-of-mouth” in order to find announcements about how to divulge the site, unusual results also appeared to users as an advertisement for “*Bife de casca de banana*” (steak of banana peel). It happened because in one of the comments of this announcement we find “I’m with water in my mouth” regarding to the announcements of “steak of banana peel”. However, this was an irrelevant result to the user. Phrases for search like “divulcation of the ‘VilanaRede’ ” or verbs such as “to popularize announcement” or “advertising Vila” were also used in this scenario.

In scenario 2 (Find announcements of mango (fruit) at ‘VilanaRede’) we wanted to verify if users would find any announcement related to mango (fruit) in the application. There was no announcement about the mango fruit in the system. However there was an announcement about *mangá* (cartoon) and it was written as misspelled of the word ‘*manga*’ (mango) in Portuguese. At this scenario, users mainly used the key-words: “mango (fruit)”, “fruit”, “mango”, “mango fruit”, “mango / fruit”. There were no findings related to mango (fruit), just with *mangá* (cartoon). Some pairs were uncertain if they would have to put the key-word “fruit” or not. Note that in a semantic search, by putting the key-word “fruit”,

¹³Is a popular expression used in Brazil that connote “passing of information from person to person” an approximated translation to English is: “word-of-mouth”

the application should return all the announcements about mango (fruit), in the case of announcements semantically related to fruit.

In the Scenario 3 (Find announcements related to food at ‘VilanaRede’) we attempted to explore whether users would use the key-word “food” in the search or if they would make a search for specific foods through the search engine. As a result, when users tried the key-word “food”, the system returned nothing. However there are several announcements on food in the system, such as: sale of “*salgadinhos*” (homemade snacks), “*pão de queijo*” (cheese bread) and others. Among the relevant considerations from the observers, during the execution of this scenario users said that the system should relate “*salgadinhos*” (homemade snacks), “*pão de queijo*” (cheese bread) and “*Bife de casca de banana*” (steak of banana peel) with the concept of food. And this makes sense because semantically all of these are food. During the discussion phase one of the users commented: “Using food is easier because it already covers everything,” i.e., all types of food that there is in the system. Another did the following statement: “For the site be more ‘lean’ and practical for those who are starting (in terms of computer literacy), like us, when we enter ‘food’, it should return a variety of foods due to our difficulty.” Yet the description of another user says: “Maybe to use food does not help in the search for something more specific, but if it is something that we have no knowledge of the domain, or we do not know what to look for, the tool would be useful and helpful.” The main key-words used in this scenario were “food”, “*comida caseira*” (homemade cooking), “food sale”, “*salgado*” (homemade snacks), “*salgadinhos*” (small homemade snacks), “*salgadinho frito*” (fried homemade snacks), pies, “*doces*” (sweets), “*pão-de-queijo*” (cheese bread), “*docinhos*” (small sweets), cake, *pastel* and “*brigadeiro*” (chocolate sweet). Note that users utilize several variations in words such as “homemade snack”, “little homemade snack” and “little homemade snacks”.

With scenario 4 (Check for some announcement related to religion combined with handi-craft at ‘VilanaRede’) we aimed to determine which key-words users would use when looking for a specific announcement. One of the observers indicated that the pair found the “Saint Anthony” because they already knew that this announcement was in the system. The same was reported by several other observers. The vast majority of the pairs used the key-words: “homemade craft”, “Crafts saint”, “holiness”, “holy” and “saints”. But one of the pairs put key-words like: “Orisha”, “Orisha of cloth”, “religious” and “sculpture”. Users found the desired information successfully.

Several observers stated that the users utilize terms/concepts from their own language or routine to perform the search; examples can be seen as “*manga rosa*” (pink mango), “*manga coquinho*” (coconuts mango), “*tutu de feijão*” (tutu bean), “word-of-mouth” for the term “disclosure”, “small sweets”, “little homemade snack” and “Orisha”. Also on several occasions the pairs discussed before reaching an agreement on which word to use in the search.

Additionally, the results indicate that users from the context of this study had difficulty with the search button; i.e., they do not have a clear concept of “Searching” in an application on the Internet. Some users had no idea about the scope of the search. They did not know that the search is only for the announcements in the Vilanarede system. This stay explicit in the description from a user who said: “Look for *fondue* because it is something chic”. However another report from other user says: “*Fondue* is very chic, we do not have it here

in our network... we will not put it in the search because the network is ours, it is poor... it will not have *fondue ...*".

Even with this lack of sense about the scope of the search, one of the observers explains that users stay surprised with the power of the search, and they explored and tested it easily. This is confirmed with the users' behavior during the execution of the activity. Positive results were obtained from it. From the data filled by the observers' forms, it is possible to indicate that approximately 85% of the pairs felt comfortable during the task. Approximately 85% of the pairs performed the scenarios with independence. Around 71% of the pairs did not make many questions to the facilitator during the task. In the discussion after the activity, users said that they had to reason about the choice of words and test various alternatives.

An interesting fact reported by the observers is that sometimes users initiate the search by entering complete questions in natural language, or at least they think aloud in that way. This is confirmed during the scenario X. This scenario was described as follows: "Suppose you want to make a reservation for medical consultation at the local hospital and went to the 'VilanaRede' system to get information (e.g. phone of the hospital). How would you make the search of some announcement related to it?" In this scenario a few pairs used key-words such as: "Hospital", "Health Center", "phone of the health clinic", "scheduling of medical appointments." However, some pairs used sentences in natural language such as: "Can anyone tell me how to make a reservation for medical consultation at PUC¹⁴?" and also "What is the phone of the SUS¹⁵ for appointments?" Observers reported that after the trial of natural language, users started using terms and key-words, and sometimes they use a combination of more than one word. During the final discussion, after the execution of the scenarios, users explained that they had learned that a complete phrase usually "do not work", and frequently only one "right word" as said by an user, is enough to return satisfactory results.

4 Discussion: The Need for Semantics from the Social Network

One of the ways of accessing information in a network is through the use of search engines. Difficulties faced by users of ISN to get information requested can be explained due to: (1) overload of information present in the systems; (2) problems encountered by users related to the contextualization of meanings from terms used. This problem becomes even more critical when dealing with users' difficulty of handling digital artifacts. As an attempt to solve this problem, we have investigated approaches that can result in better and more appropriate search engines in ISN. The objective is to help users to access information more easily and naturally.

Practical results show that the users' colloquial language should be considered during the development of more appropriate search engines. Individuals in a social network can create and use proper terms from a community. This occurs because individuals in the

¹⁴PUC¹⁴ is a hospital at Campinas city from the Pontifical Catholic University of Campinas

¹⁵Unified Health System in Brazil

social network communicate, exchange data and information creating and sharing meanings. Consequently, they also use their own terms and meanings in searching. Thus, it is important to propose methods and computational mechanisms to identify these meanings.

Therefore, it is necessary to understand and to construct computationally tractable models from the semantic point of view that come out from the network itself. Semantics here is understood as the interpretation of signs [10] by individuals and their association with real world elements. This interpretation is socially contextualized, i.e., individuals and communities may have different interpretations for the same sign and a sign may connote different meanings depending on the context applied.

In a social network the emergence of semantics is an ongoing process in which meanings and interpretations are constructed, used and shared through the system from the interactions and expressions of users in communication. So it is going to be necessary to create new methods, mechanisms and computational tools to deal with it.

Research in this aspect shows as a prime importance, with semantic models more faithful to the social network reality resulting in more precise search engines. They should take into account the meaning used and emerged in the specific social network, making sense for individuals of that context. New approaches to this representation should be investigated in order to provide capacity for "inclusive semantic searchers". New opportunities for the universal access to information could be created by analyzing how the semantics is developed in a social network, representing it and using this representation in a search engine. The results indicate that this could facilitate the interaction with computer systems, since they would provide access and knowledge sharing in a more effective way.

In the field of the "Semantic Web" there are several proposals and architectural approaches for semantic searching. Mangold [11] presents a survey of the main proposals in the literature, as well as a reflection on the open questions about semantic search. The work of Mangold [11] provides a starting point on the architectural decisions to elaborate a new search mechanism for our context. The intention is to develop an architectural strategy that best suits the needs and constraints of the context of our research for semantic search in ISN. In each of the categories described in the Mangold's work there is a need to rethink the solutions because we are in a different and sensible context. However, his research already points out some interesting ideas, for example, regarding the category of transparency presented in his work. This category indicates if the semantic capabilities of the system are transparent or not for users in their interaction with the application. According to Mangold [11] the more effective semantic search solution provides both transparency to novice users and interactive behavior for more experienced users.

One of the studies analyzed in Mangold [11] is a semantic search engine from the project TRUST of Amaral *et al.* [12]. The authors present a semantic engine of questions and answers, as well as its architecture, tools and processing mechanisms implemented for the module of the Portuguese language. The objective of this project is to develop a semantic and multi-language search engine capable of natural language processing. Additionally, there are other proposals for semantic search such as Celino [13], which provides an extensible framework of semantic search for the development of semantic search engines, which takes into account characteristics necessary in the search for multimedia content.

Regarding semantic search in social networks, we found the proposal of Choudhari *et*

al. [14]. They describe an architecture that proposes to develop a semantic search with an approach based on grouping of semantically related terms. If the user wants to make a search through a key-word, the search could be generalized to all the terms from the group. For example, assuming that “to swim”, “to run” and “sport” are in a group (the terms of the groups are created according to information from the user profiles), to make the semantic search would be equivalent to use any of these key-words. The system automatically would use all the key-words from this group to execute the search.

The proposal of Choudhari *et al.* [14] does not deal with issues of polysemy nor articulates ideas for solutions of other aspects also problematic in searching. The goal of their work was to develop an application (a plug-in) for the FaceBook and to test the system in this social network system. Their work also focus mainly on semantic search of people (users of the network): a user could search for people who love sports. So, having some information about any sport in the user profile: basketball, soccer, etc..., the search should return these people who are part of the user’s social network. Other study like Mika [17] is a relationship between social networks and Semantic Web technologies, such as computational ontologies. This is primarily concerned with social network analysis, and does not address the topic of semantic search into social networks.

From these related works, along with the analyzed search scenarios and the final discussion from the 8th workshop of the e-Cidadania, we could identify and point out some factors that can contribute to the improvement of a search engine in an ISN. Based on the collected and observed data with the involved subjects, the experiment in the workshop shows that availability of a search with semantic emphasis will be more natural to people. Moreover, it was found that users use different terms that are specific to their everyday language (informal), and they have semantic relationships with several other cult terms (formal).

Therefore, this study confirms the importance of trying to use the users’ language (expressed in many moments in their network activity) for the development of computational semantic representation that will be used as basis to the search engine. This fact may improve search engine in the context of an ISN, because it will tend to return results that best suit the user profile. For this accomplishment new theoretical approaches must be employed. The use of Semiotic theory [10, 15, 16] may show great relevance in this sense, and can contribute to new methods for semantic modeling and innovative research results that best represent search engines.

5 Conclusions

Nowadays we face a scenario in Brazil and in other developing countries as well, that is characterized by a diversity of people, cultures and social problems. Computer Science should contribute with new methods and systems to provide unrestricted access to information and knowledge. Thus it is necessary to seek for solutions that support the creation of a digital culture respecting the diversity of our society. Based and inspired by this challenge, e-Cidadania project brings these concerns with the ‘Inclusive Social Network Services’ (ISN) concept. New search mechanisms can make a difference helping people effectively and naturally find and use information through an ISN. Consequently, this would be a way to promote the participatory and universal access to information.

Aiming to accomplish it, this report presented the results of an analysis about searching in ISN. We got important insights about how to improve a search mechanism considering aspects related to the digital and social inclusion. We could verify, along the experiment with real users, that semantic aspects can make the difference to user to find and access information. Also, it was verified that the current search mechanisms built primarily through comparisons of key-words and lexical-syntactic processing is not adequate for the ISN context.

There is not much research in literature related to search on social networking sites. It is therefore extremely important to develop new and more powerful mechanisms for information retrieval to really consider people’s diversity. Therefore, it is necessary to develop methods and tools to identifying and also computationally represent the semantics used in the network. These mechanisms should take into account the meanings (semantics) created, shared and used by people in the network, so they could provide more suitable search. It was also observed that the Semiotics theory can contribute mainly regarding semantic and modeling aspects. The next steps of this research involve to investigate ideas based on this theory and to propose better solutions to search engines for ISN.

Acknowledgments

This work is funded by Microsoft Research - FAPESP Institute for IT Research (grant #2007/54564-1) and also by CNPq/CTI (680.041/2006-0). The authors also thank colleagues from CTI, IC - Unicamp, NIED, InterHAD and Casa Brasil for insightful discussion.

References

- [1] Baranauskas, M.C.C. (2007): e-Cidadania: Systems and Methods for the Constitution of a Culture mediated by Information and Communication Technology. Proposal for the Microsoft Research-Fapesp Institute.
- [2] Baranauskas, M. C. C., and de Souza, C. S. (2006): Desafio 4 acesso participativo e universal do cidadão brasileiro ao conhecimento. *Computao Brasil VII*, 23, 7.
- [3] Baranauskas, M.C.C.; Hornung, H.H.; Martins, M. C. (2008): Design Socialmente Responsável: Desafios de Interface de Usuário no Contexto Brasileiro. In Proceedings of the 35o. Semish Seminário Integrado de Software e Hardware, XXVIII Congresso da Sociedade Brasileira de Computação SBC 2008.
- [4] Hayashi, E.C.S.; Neris, V.P.A.; Almeida, L.D.A.; Rodriguez, C.L.; Martins, M.C.; Baranauskas, M.C.C. (2008): Inclusive Social Networks: Clarifying Concepts and Prospecting Solutions for e-Cidadania. Technical Report IC-08-029, Institute of Computing, University of Campinas - Unicamp.
- [5] Boyd, D. M. & Ellison, N. B. (2008): Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*. Number 13. pp. 210-230.
- [6] Nielsen (2009): Global Faces and Networked Places. The Nielsen Company USA. http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/03/nielsen_globalfaces_mar09.pdf – Accessed July 2009.
- [7] IBGE (2008): Síntese de Indicadores Sociais – 2008. http://www.ibge.gov.br/home/presidencia/noticias/noticia_imprensa.php?id_noticia=1233 Accessed July 2009.
- [8] MEC (2007): <http://portal.mec.gov.br> – Accessed November 2009.
- [9] IPM (2007): Indicador Nacional de Alfabetismo Funcional. <http://www.smecc.salvador.ba.gov.br/site/documentos/espaco-virtual/espaco-dados-estatisticos/indicador%20de%20analfabetismo%20funcional%202007.pdf> – Accessed January 2010.
- [10] Peirce, C.S. (1931-1958): *Collected Papers*, Cambridge, Mass: Harvard University Press.
- [11] Mangold, C. (2007): ‘A survey and classification of semantic search approaches’, *Int. J. Metadata, Semantics and Ontology*, Vol. 2, No. 1, pp.23-34
- [12] Amaral, C.; Laurent, D.; Martins, A.; Mendes, A.; Pinto, C. (2004): ‘Design and implementation of a semantic search engine for Portuguese’, *Proceedings of 4th International Conference on Language Resources and Evaluation, LREC 2004*, Lisbon, Portugal, Vol. I, pp.247-250.

- [13] Celino, I.; Turati, A.; Valle, E. D.; D. Cerizza. (2007): Squiggle - a semantic search engine at work. In Proceedings of the 4th European Semantic Web Conference.
- [14] Choudhari, Ajay; Jain, Mohit; Sinharoy, Avik; Zhang, Min. (2008) Semantic Search in Social Networks. http://www.cc.gatech.edu/projects/disl/courses/8803/2008/project/project_deliverables/group22/proposal.pdf – Accessed July 2009.
- [15] Stamper, R. K. (2000): Information Systems as a Social Science: An Alternative to the FRISCO Formalism. In E. D. Falkenberg, K. Lyytinen, A. A. Verrijn-Stuart (eds) Information System Concepts: an Integrated Discipline Emerging, Kluwer Academic Publishers, USA, pp.1-51.
- [16] Kecheng, Liu. (2000): Semiotics in information systems engineering, Cambridge University Press.
- [17] Mika, Péter (2006) Social Networks and the Semantic Web. Academic Dissertation. Vrije Universiteit. Available at: <http://dare.uvu.vu.nl/bitstream/1871/13263/5/7915.pdf> – Accessed July 2009.

A Appendix A

Workshop 8 – e-Cidadania: Search Activity Task Sheet

Participants: _____

Scenario 1: Find announcements on how to popularize the 'VilanaRede'.

Please write here the words used to search
Please write here the results from the search

Scenario 2: Find announcements of mango (fruit) at 'VilanaRede'.

Please write here the words used to search
Please write here the results from the search

Figure 1: Task Sheet

Scenario 3: Find announcements related to food at 'VilanaRede'.

Please write here the words used to search
Please write here the results from the search

Scenario 4: Check for some announcement related to holiness combined with handicraft at the 'VilanaRede'.

Please write here the words used to search
Please write here the results from the search

Scenario X: Suppose you want to make a reservation for a medical consultation at the local hospital and went to the 'VilanaRede' system to get information (e.g. phone of the hospital). How would you make the search of some announcement related to it?

Please write here the possible words that you would use to search

Figure 2: Task Sheet

B Appendix B

Workshop 8 – e-Cidadania: Search Activity

Observer Form

Observer: _____

Participants: _____

General questions about the activity:

- What were the main difficulties encountered by users in making the search?

- Do you think that users used terms or concepts of their own daily language to perform the search? Could you exemplify?

Figure 3: Observer Form

- Final considerations/Additional observations

About the users' behavior, please answer:

Behavior	Answer
Did the user show to be confident during the tasks?	<input type="checkbox"/> Yes <input type="checkbox"/> No
Was the user independent on performing the tasks?	<input type="checkbox"/> Yes <input type="checkbox"/> No
Did the user ask a lot of questions to the observer?	<input type="checkbox"/> Yes <input type="checkbox"/> No

Figure 4: Observer Form