International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 9, September- 2014

# Searching a Social Networking Site – Shift from Centralized Database to Metadata

Aamir Junaid Ahmad
Assistant Professor,
Department of Computer Science & Engineering,
Maulana Azad College of Engineering & Technology,
Patna, India

Sabina Priyadarshini
Assistant Professor,
Department of Information Technology
Birla Institute Of Technology, Mesra, India

***Abstract*— Social Networking is becoming part of every individual's life be it students, research scholars, professionals from different domains or retired people living at home with their grandchildren. Everyone wants to share their information with their friends. The sociological theory says that social interaction creates similarity and vice versa, interaction creates similarity. We often want to look for a person belonging to a particular group or characteristics (similar to our profile) but land up with either a lot of information or no information at all. We can connect only with people who are known to us by name or the people who have been suggested by the system. Here we present a method using Semantic Web to represent and process social network information so that user-queries can be better answered. We have done a survey to know what are the information that members in Social Networking sites usually keep in their profile and how concerned are they about the privacy of their data.**

***Keywords*— *FOAF, Resource Description Framework, Semantic Web, Social Networking, XML***

## INTRODUCTION

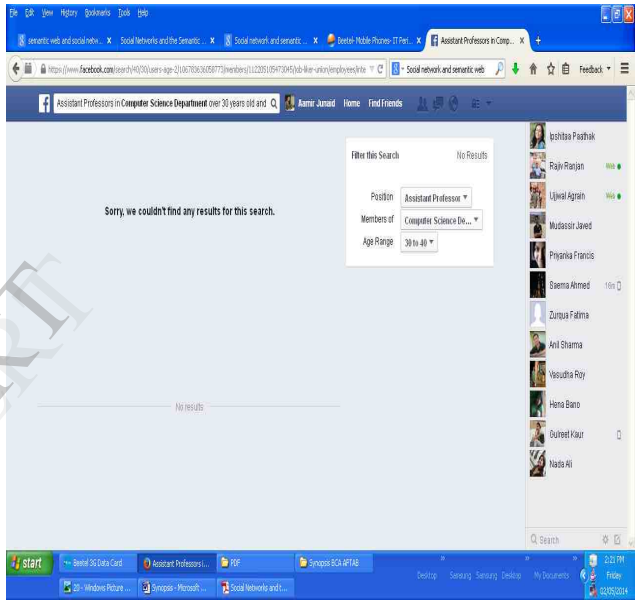
Social networking provides access to personal information sharing and online socialization. These sites allow members to create a profile with their basic information like contact details, education details, professional or work related details. The members can link to the profiles of their friends and invite others. They can browse the resulting social network in order to connect with common friends, find lost friends and to create new friendships based on shared interests.

Social Networking is an easy way for people to have a social connection with other people with similar interests. These sites serve communication purposes among special interest groups, but they do not have a proper searching option where we can look for people or groups having certain characteristics.

The following three searches in Facebook (02-May-2014) couldn't give any result

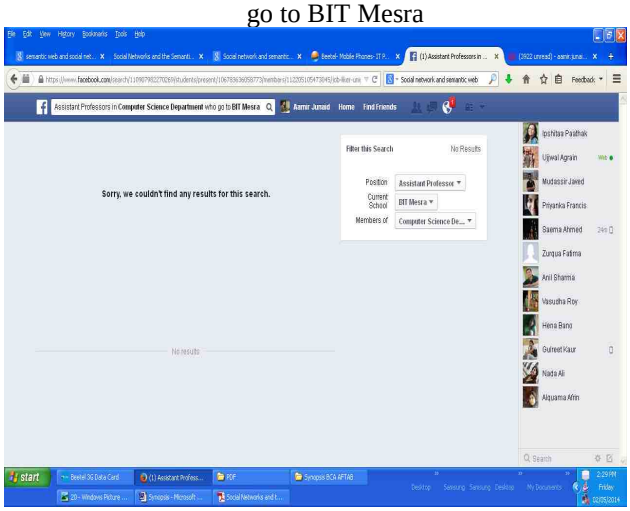
1. Assistant Professors in the Computer Science Department between 30 and 40 years of age
2. Assistant Professors in Computer Science Departme nt who go to BIT Mesra
3. Assistant Professors in the Computer Science department

Query 1. Assistant Professors in the Computer Science department between 30 and 40 years




The output is : Sorry we couldn't find any results for this search

2.Assistant Professors in Computer Science Department who go to BIT Mesra




The output is : Sorry we couldn't find any results for this search

3. Assistant Professors in the Computer Science department


All of these results are not as expected by the user.

## Methodology

The limitation in the present networking sites is the centralized storage of user profiles. The information is bound to the site in which the profile was created. It can only be accessed in the way the database owner has developed the site. This problem has been addressed with the use of Semantic Web technology. The Semantic Web can be used to improve the process of searching the relevant information from the vast store of continually updated information, as well as emergent data about individual. It is an effort to enhance current web so that computers can process the information, interpret and connect it, to help users to find required information.

The objective of this research is to use Semantic Web technology to represent and process social network information (user profiles and friendship network) in such a way that it can be processed by machine. The RDF[1] and FOAF[2], Web Semantic technologies[3] can be used for such representation.

The "FOAF" (friend-of-a-friend) initiative undertaken by the World Wide Web Consortium (W3C) has focused on developing ways to describe the properties of people such as date of birth, age, real name, nick name, contact details, professional details, educational details and their social relationships as expressed through their interests, people known, liking etc. Socially, the most important property in FOAF is foaf:knows which can be used in forming a FOAF-web.

Some of the most common properties are the following:

1) foaf: name - Full name of the person.
2) foaf:firstname - Individual's first name.
3) foaf: surname - Individual's family name.
4) foaf: nick - provides a nickname for a particular person.
5) foaf: homepage - specifies a link to the persons homepage, e.g. http://www.macet.net.in/aamir/
6) foaf: phone - Individuals phone number.
7) foaf: school - Link to the school homepage this individual is studying at.
8) foaf: gender - provides gender information
9) foaf: knows - specifies a person that this person knows.
10) foaf: depiction - points to a URL of an image of the person in question.

Table 1

FOAF profiles are created and controlled by the individual user and shared in a distributed fashion. Much like the way web pages are linked to each other by anchors, these profiles link to the profiles of friends by using the *knows* and *seeAlso* properties, creating the FOAF-web.



FOAF IS AN RDF (RESOURCE DESCRIPTION FRAMEWORK) VOCABULARY. THE GOAL OF FOAF IS TO PROVIDE A MEANS FOR DESCRIBING PEOPLE AND THEIR PERSONAL INFORMATION SUCH AS NAME, E-MAIL ADDRESS AS WELL AS SOCIAL LINKAGE TO THEIR FRIENDS BY USING XML AND RDF.
The first challenge in using FOAF for displaying information is the availability of basic user information on the social networking site. The second important issue is users' privacy in making their information public in the FOAF format. Since some information in a person's FOAF document may include the private information which the user doesn't want to make public, as the metadata extracted from the web will be given some sort of publicity. It is therefore important to know what are the information which the users usually keep in their profile and secondly how concerned are users about the privacy of the following information they submit on social networking sites[4]. To reach to the conclusion a survey was conducted in August 2014.

Survey Methodology: A questionnaire was prepared and distributed through

- Email
- Online Form
- Distribution in Classroom and
- Distribution in market

Total number of individuals who returned or gave response is 2120
The individuals range from 18 years old to retired persons.

Question: Please indicate what information you include on your social networking sites

From the above survey(Table 1) it is clear that a user usually keep those information in their profiles which can be easily converted to FOAF properties discussed above.

| Sl.No. | User Information | Size | Percentage |
|---|---|---|---|
| 1 | Real name | 1295 | 71% |
| 2 | Nick name | 510 | 28% |
| 3 | Gender | 1623 | 89% |
| 4 | Date of Birth | 875 | 48% |
| 5 | Email address | 875 | 48% |
| 6 | Mobile number | 291 | 16% |
| 7 | Home town/city | 1185 | 65% |
| 8 | Photos of yourself | 1368 | 75% |
| 9 | Photos of others | 1076 | 59% |
| 10 | Political views | 218 | 12% |
| 11 | People You Know | 1696 | 93% |
| 12 | Sexual orientation | 565 | 31% |
| 13 | Relationship status | 1276 | 70% |
| 14 | School/Education related | 1568 | 86% |
| 16 | Your Profession/Work related | 1495 | 82% |
| 17 | Your Company(Organisation) | 1423 | 78% |
| | Total/ respondents 1824 | | |

The second challenge is to know how concerned the users are about the privacy of their information. To check this another question was used in the survey.

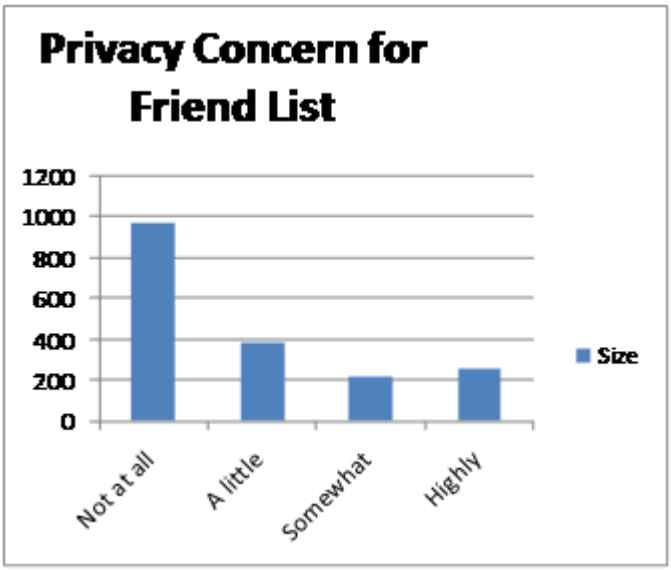
Question) How active are you in safeguarding (protecting) the following information submitted by you on social networking sites?

| | Not at all | A little | Somewhat | Highly |
|---|---|---|---|---|
| Friend List | ○ | ○ | ○ | ○ |
| Your contact details (email-id, mobile- no, Hometown, city etc) | ◉ | ○ | ○ | ○ |
| Education Detail | ○ | ○ | ○ | ○ |
| Date of Birth | ○ | ○ | ○ | ○ |
| Lifestyle related (eg. photos, views etc) | ○ | ○ | ○ | ○ |
| Professional / work related information | ○ | ○ | ○ | ○ |
| Relationship status / Sexual orientation | ○ | ○ | ○ | ○ |

The response submitted by 1824 users are as follows

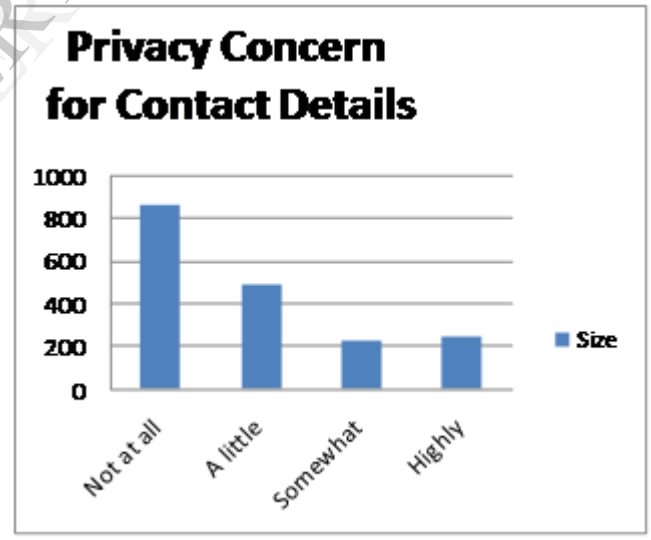
Privacy Concern for Friend List

| Sample Size | | Percentage |
|---|---|---|
| Not at all | 967 | 53% |
| A little | 383 | 21% |
| Somewhat | 218 | 12% |
| Highly | 256 | 14% |
| Total 1824 | | 100% |



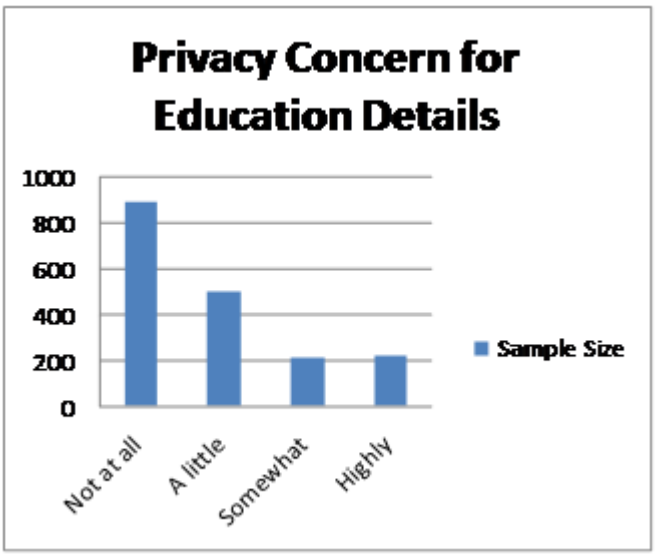

Privacy Concern for Contact Details

| Sample Size | | Percentage |
|---|---|---|
| Not at all | 863 | 47% |
| A little | 487 | 27% |
| Somewhat | 228 | 13% |
| Highly | 246 | 13% |
| Total 1824 | | 100% |



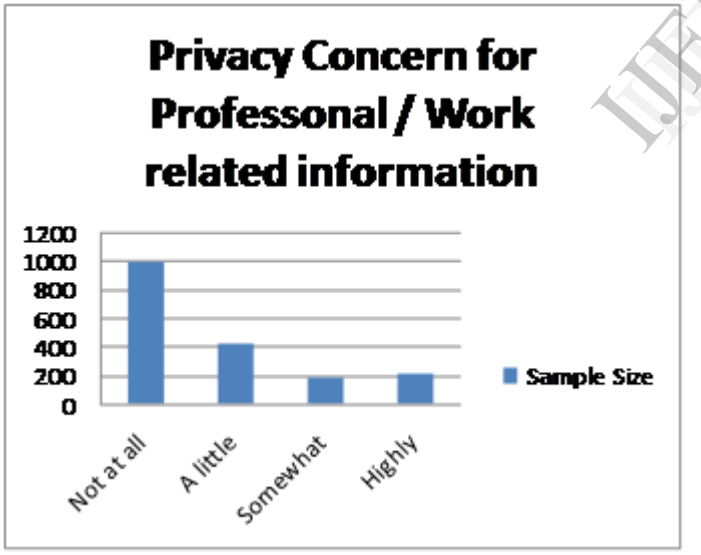

Privacy Concern for Education Details

| Sample Size | | Percentage |
|---|---|---|
| Not at all | 894 | 49% |
| A little | 496 | 27% |
| Somewhat | 213 | 12% |
| Highly | 221 | 12% |
| Total 1824 | | 100% |




Privacy Concern for Professional / work related information

| Sample Size | | Percentage |
|---|---|---|
| Not at all | 994 | 55% |
| A little | 426 | 23% |
| Somewhat | 183 | 10% |
| Highly | 221 | 12% |
| Total | 1824 | 100% |




From the above chart we can see that most of the users are not very concerned about their privacy related to the people they know, their contact details, their education detail as well as professional/work related information. Therefore we can publish these information in the form of metadata by extracting them from the Web without worrying about the privacy quotient. The social networking sites can even add this to their terms and conditions that the information will be published in the form of metadata which can be used in the searching process.

In addition to the privacy problem, another potential problem may be of copyrights. Some Social Networking Sites may prevent the utilization of their contents for revision or manipulation especially for commercial objectives. We have to consider this problem before we publish the data.

## CONCLUSION AND FUTURE WORK

Once we are sure that the required data is available in the database the Social Networking sites can create FOAF profile that contains user information including email address, location, interests, a list of friends, etc., and place that file in a web-accessible location. Users can edit their profiles using web-based forms, hence, most users may be completely unaware of the technology's existence and their use of it.

Now these FOAF files can be used in different ways to process and find answer to user queries. Some research are going on like Indexing FOAF files as well as Flink system for the extraction, aggregation and visualization of online social networks which employs semantic technology for reasoning with personal information extracted from a number of electronic information sources including web pages, emails, publication archives and FOAF profiles.

Thus we can say that FOAF profiles can make ways for answering user queries in a better way.



*Acknowledgment*

The authors would like to thank all faculty and students of Maulana Azad College of Engineering & Technology, Patna who have shown interest and supported in this research.

*References*

[1] www.obitko.com/tutrials/ontologies-semantic-web/resource-description-framework.html

[2] John C. Paolillo and Elijah Wright. The Challenges of FOAF Characterization.

[3] Peter Mika, Social Network and the Semantic Web

[4] John P. Scott. Social Network Analysis: A Handbook. Sage Publications