

# Beyond Web-log: Transform Blog into Personal Expertise and Social Network via myFOAF Support

Tse-Ming Tsai, Chia-Chun Shih, Jay Stu, and Wen-Nan Wang

Institute for Information Industry (III)  
8F., No.133, Sec. 4, Minsheng E. Rd., Songshan District,  
Taipei City 105, Taiwan  
+886-2-2713-9000 ext. 232

{eric, chiachun, jaxxstu, wennen }@iii.org.tw

Seng-cho T. Chou

National Taiwan University  
No. 1, Sec. 4, Roosevelt Road,  
Taipei, 10617 Taiwan  
+886-2-3366-1182

chou@ntu.edu.tw

## ABSTRACT

To utilize the dramatically increasing Blog contents and widely accepted web 2.0 services, this research proposes myFOAF which extends FOAF to model dynamic social contexts over the web. Based on the richer social semantic, we can generate the blogger's myFOAF profile, via the corresponding contents such as RSS, Blog content, tagging data, mashup service that retrieve from blogger's daily web behavior. Two main application models are addressed in the system. Firstly, the Personal Expertise Explorer can accumulate blogger's tagging ontology (myTag) to analyze the blogger's domain related expertise. The evaluation result shows that the average precision reaches around 77.6%. Secondly, the Personal Social Explorer gathers the blogger's friend information to form a personalized social network diagram which presents blogger's social activities in certain tag domain. With the extra semantic supported by myFOAF, blogger can easily explore his own implicit expertise and web social relationship with minimal efforts.

## Categories and Subject Descriptors

H.5.4 [Hypertext/Hypermedia]: Architectures, Navigation, Theory, User issues

## General Terms

Algorithms, Management

## Keywords

Blog, Web 2.0, FOAF, Expertise, Social Network, Semantic Web

## 1. INTRODUCTION

For years web has provided a successful information exchange platform for people to communicate with one another regardless of physical distance. Nowadays, one of the most popular web activities is Blogging – for everyone to express oneself, to publish information, to share knowledge, and to connect with others. Web 2.0, a more collaborative and social-related style form of using web, tends to dominate the way we create and deliver information. Blog is a website where entries are written in chronological order and displayed in reverse chronological order[1]. People use Blogs to share life, express opinions, introduce something cool, and also interact with friends [2]. The interaction may include comments, trackbacks, blogrolls, links, syndications and tag cloud, calendar, categories, and RSS subscription. These interactions are useful to understand how one communicates with others and could be used as critical clues to build personal profiles and social.

To some extent, Blog provides certain semi-structure content such as article (diary post), reader's comment (response), side bar (calendar, self-introduction, categories, friends, etc.), in different content part. In Web 2.0, blogger can further consume different information web services they want by mashing-up those widgets/gadgets. Some bloggers have even provided their own FOAF file. FOAF is an acronym for "Friend Of A Friend", an experimental project and vocabulary for the Semantic Web. It describes personal information which includes name, email address, and their friends within the form of XML and RDF. It allows software agent to process these descriptions to discover information about user and its community members. The FOAF vocabulary definition is written using a computer language (RDF/OWL) and there are some tools to create, present and assert those FOAF descriptions [3] [4] [5]. But the static description is not sufficient enough to analyze user's intension expertise and social relationships, especially under the web2.0-related dynamic applications. Therefore, how FOAF should be enhanced to provide user-centric social information becomes an important subject.

The research objectives are designed to utilize the popular Blog's content as the source of semantic over the web, to provide personal ontology web services, to analyze blogger's expertise according to tag usage, and to enable the ability to explore his up-to-date social context. Related works, myFOAF definition, system architecture, evaluation, discussion and conclusion are presented as follows.

## 2. RELATED WORKS

### 2.1 Social Network Analysis

Social network is a social structure made of nodes tied by one or more specific types of relations, such as values, visions, ideas, financial exchanges, friends, kinship, dislike, trade, web links, sexual relations, disease transmissions, or airline routes. A social network is a map of all of these relevant ties between the nodes. The network can also be used to determine the social capital of individual actors. These concepts are often displayed in a social network diagram, where nodes are the points and ties are the lines [6]. Social Network Analysis (SNA) provides some useful tools to look into a social system. A Blog could be a social platform, and SNA surely is helpful to analyze Blog communities. Through comments, trackbacks for each post, and blogrolls, syndication activities, bloggers build connections to others [7]. In most cases, these interactions leave messages in Blog content, so that we can track these interactions without user activity logs. Since different kinds of user interactions would imply different types and degrees

of relationships [8], we would identify possible types of relationships and map interactions to relationships with different degrees.

## 2.2 Web 2.0 Personal Mashup Service

Web 2.0, a phrase coined by O'Reilly Media in 2003[9] and popularized by the first Web 2.0 conference in 2004, refers to a perceived second generation of web-based communities and hosted services - such as social-networking sites, wikis and folksonomies - which facilitate collaboration and sharing between users. These technologies can help users to contribute their own contents in the web 2.0 platform. Some web 2.0 applications that related to personal social behavior by the blog sidebar are:

- Mybloglog reveal the identity of visitors..
- del.icio.us is a social bookmarking website. It allows users to access their bookmarks everywhere via web, and most important is that its tagging system enable user associate each other
- YouTube is a video sharing website where users can upload/view/share video clips.
- Flickr is a photo sharing website and an online community platform.
- 43things is built on the principles of tagging, rather than creating explicit interpersonal links. Users can create accounts and list a number of goals or hopes.

These services reveal the user's preference and concerns, especially for those with tags support provide more clues for personal interesting and expertise mining. Once we have a conventional integration interface then we can link them all.

## 2.3 Collaborative Tagging

The trend of web 2.0 has introduced a new categorization method called folksonomy. Vander Wal coined the term folksonomy, and has defined it as: "A folksonomy is the result of personal free tagging of information and objects (anything with a URL) on the internet for one's own retrieval. The tagging is performed in a social environment (shared and open to all). The tagging action is done by the person consuming the information." [10] Tagging allows users to define their own terms to arrange items [11, 12, 13]. In Collaborative tagging environment, two different relations between any pair of tags can be defined by aggregating tagging data [14, 15]. First, two tags are in a "Co-Resource" relation if they are adopted by the same resource. This relation is stronger between tags with more shared resources. Second, two tags used by the same user are in a "Co-User" relation. While the "Co-Resource" relation is most appropriate for establishing a public concept hierarchy, the "Co-User" Relation is most suitable for establishing a private concept hierarchy.

## 2.4 Expertise

There are different definitions about "expertise" in previous works. We want to figure out how they define "expertise" and how they evaluate it.

Alan et al. [16] combines tagging and social networking to extract users' characteristics, especially for expertise location. It analyzes tagging behavior in a social bookmarking service and builds a social network around the clustered tag space. The clusters form different expertise areas. And a modified PageRank algorithm is performed to define the expert ranking. Zhang et al.

[17] wants to find optimal search algorithm like PageRank and HITS to find experts in social networks such as a programmer forum. The expertise here equals authorship in a linkage structure. Song et al. [18] tries to build a personal profile that contains all the research areas. They also build evolutionary expertise models using literature citations to match researchers' roadmap. Through the model, we can look into the relationships of researchers' master disciplines. It helps when searching for inter-discipline experts. The expertise here is the quality and quantity of publications. McDonald et al. [19] proposed a recommender system to recommend experts and show how it works and why it is beneficial. It typically performs in an enterprise and help users to address the experts in the organization. The expertise here is the working experience and employee's skills. Zhang et al. [20] wants to find an expert in a domain specific forum such like a java forum. The expertise is defined by the answering relationships in different posts. It proposes an ExpertiseRank algorithm to evaluate the expertise level to certain user. Those who answer more questions gain more credit of expertise he got. The expertise here is the points that a user gain from the askers. Campbell et al. [21] describes a way to extract social network from e-mail activity. The expertise topics are generated by the clustering of message content. The expertise level here is also determined by the linkage algorithm like HITS.

## 3. MYFOAF SPECIFICATION

Extending the static FOAF function, the myFOAF further integrate the dynamic information which the blogger may mash-up to enable the personal expertise and social network. Table 1 describes the myFOAF data types.

Table 1. myFOAF data type analysis

Data Type	Description
Static	The basic user information includes the Blog-related and FOAF information when blogger starts using the system.
Collection	The dynamic information includes the blogger interaction information or the information which blogger uses the myFOAF collection to gather web 2.0 widgets in personal ontology such as Blog, podcast, bookmarks in del.icio.us, photos in Flickr, videos in Youtube or blogroll and so on.
Artifact	This information comes from the blogger's social activity analysis. Personal Expertise Explorer shows the blogger's expert domains. Personal Social Explorer shows the blogger's social relationships or recommends him with some blogger he may want to know each other.

Table 2(a), 2(b), and 2(c) correspondingly shows the Static, Collection, and Artifact definition of myFOAF, the extended FOAF definition to support personal expertise and social network service..

Table 2(a). MYFOAF Static Information

myFOAF Classes	Description
Blog	The basic blog information such as profile, articles.

**Table 2(b). MYFOAF Collection Information**

myFOAF Classes	Description	
Blog Activities	Blogrolls	blogger's friends list
	MyTag	Blogger's tag usage (federal tag cloud)
	References	The URL links (content links) and Trackbacks
	Footprints	Exploring history lists of the blogs that the blogger visited.
	Comments	Message history which the blogger left on others' Blogs
	Subscriptions	Personal RSS feeds subscription list
Involved Web Services	Bookmark	Personal bookmark list in bookmark site. ie. del.icio.us
	Photo	Personal favorite list in photo site. ie. flickr
	Video	Personal favorite list in video site. ie. youtube
	Todo	To do list site. ie. 43thing
	...	...

**Table 2(c). MYFOAF Artifact Information**

myFOAF Classes	Description
Expertise	Personal expertise degree analysis
Familiarity	The foaf:knows denotes familiarity degree and myTagCluster.

Figure 1 shows the myFOAF example based on the specification. The **collection** example comes from the tags when the blogger creates his account in tag supported web 2.0 sites. As blogger inputs his ID, our system can use the RSS subscription or web service link to gather his tags for up-to-date contents. In Figure 1, for example, the blogger tagged the picture with "sunset, colorful" in the flickr account.

The **Artifact** information comes from the analyzed blogger's social activities. The personal expertise explorer analyzes the blogger's myTag to compute the blogger's expertise degree in his myConceptSpace. The myFOAF:expertise represents the blogger's expertise information.

Some previous studies extend the FOAF to define relationships between users; however, the closeness of the relationship was never evaluated and quantified [22]. In our study, the personal social explorer computes the blogger's **familiarity** degrees with his friends. As demonstrated in Figure 1, the blogger has relationship score of 0.3 with "jaxx" in the "web2.0" domain.

```

- <foaf:Person rdf:ID="me">
  <foaf:name>ccshih</foaf:name>
  <foaf:mbox_sha1sum>f9621b5e7a4374407f7ae64a1da0ef3b33bb02d0</foaf:mbox_sha1sum>
  <foaf:weblog rdf:resource="http://blog.iii.org.tw/ccshih" />
- <myFOAF:references>
  <foaf:Document dc:title="iii" rdf:resource="http://www.iii.org.tw" dc:publisher="acespot" myFOAF:tag="iii,information technology" />
</myFOAF:references>
- <myFOAF:mytag>
  <myFOAF:familiarity dc:title="web2.0">10</myFOAF:familiarity>
  <!-- <myFOAF:familiarity dc:title="ads">10</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="social networking">6</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="social computing">4</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="mashup">4</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="blog">4</myFOAF:familiarity -->
</myFOAF:mytag>
- <myFOAF:footprints>
  <foaf:Document dc:title="acespot's blog" dc:publisher="acespot" rdf:resource="http://blog.iii.org.tw/acespot" />
</myFOAF:footprints>
- <myFOAF:collections>
  <foaf:Document dc:title="The Friend of a Friend (FOAF) project"
  rdf:resource="http://del.icio.us/url/e42cc29284bc88b6a4df7e432acf2d32" myFOAF:tag="foaf, semantic web" />
  <foaf:Document dc:title="Sunset" rdf:resource="http://www.flickr.com/photos/gardinergirl/634849037/in/pool-90631221@N00/"
  dc:publisher="gardinergirl" myFOAF:tag="sunset, colorful" />
  <foaf:Document dc:title="HK Bus uncle" rdf:resource="www.youtube.com/watch?v=RSHziqJWycM" dc:publisher="afewor" myFOAF:tag="hk,
  interesting" />
- <myFOAF:expertise>
  <myFOAF:familiarity dc:title="ads">0.93</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="ads2.0">0.83</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="social networking">0.78</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="blog">0.48</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="mashup">0.44</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="social computing">0.39</myFOAF:familiarity>
  <myFOAF:familiarity dc:title="web2.0">0.35</myFOAF:familiarity>
</myFOAF:expertise>
- <foaf:knows>
  - <foaf:Person>
    <foaf:name>jaxx</foaf:name>
    <rdfs:seeAlso rdf:resource="http://blog.iii.org.tw/jaxx" />
    <myFOAF:familiarity dc:title="web2.0">0.3</myFOAF:familiarity>
  </foaf:Person>
  </foaf:knows>
  - <foaf:knows>
    - <foaf:Person>
      <foaf:name>joseph</foaf:name>
      <rdfs:seeAlso rdf:resource="http://blog.iii.org.tw/joseph" />
      <myFOAF:familiarity dc:title="web2.0">0.1</myFOAF:familiarity>
    </foaf:Person>
  </foaf:knows>
- </foaf:Person>

```

**Figure 1. myFOAF sample**

## 4. SYSTEM ARCHITECTURE

The system architecture is illustrated in Figure 2.

After blogger registered his static and collection information in the web form, the Content Aggregation and Processing service generates the primer input to be collected for processing such as Blog content, FOAF information, corresponding involved web service contents (RSS, tagging data, bookmark, photo, video, 43things and so on). We can trace its sharing community topology and take into consideration.

At the Model and Algorithm part, Personal Ontology contains blogger's static personal information and dynamic accumulated MyTags. MyTags stores the tagging data from different resource collection services such as Blog tagcloud, de.licio.us, youTube, flickr etc.

Universal Tag Ontology provides detailed information about degree of strength amongst tag relationships [23], which conducted by analyzing tag co-occurrence in de.licio.us via long-term RSS subscription. The Universal Tag Ontology contains a large folksonomy and the relation weights between relevant tags. We construct those massive and collative tags as a concept space database and compute their semantic folksonomy.

Based on the architecture, we can then 1) calculate every blogger's expertise degree according to his tagging behavior and other blogger's responses and 2) analyze his personal social network and dynamic social context according to everyday blogging and surfing log we can find. This information clues in blogger's social activities such as blogroll, inner text reference, track back, comments, RSS subscription etc.

As for the upper layer application part, based upon the tagging semantic and social behavior infrastructure, we can then build the interesting applications such as Personal Expertise Explorer and Personal Social Explorer.

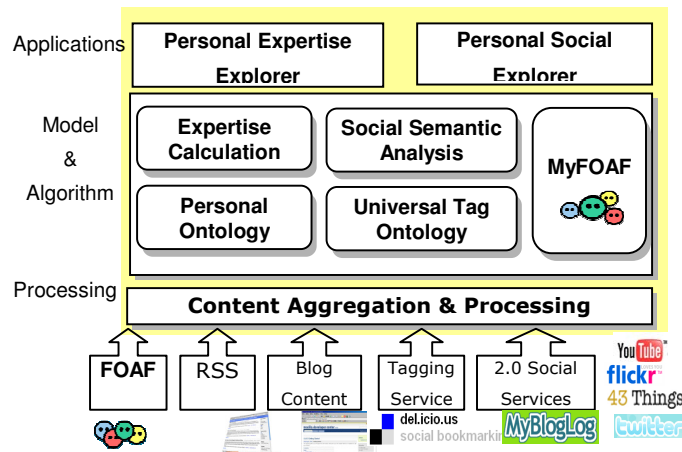


Figure 2. The System Architecture

## 5. APPLICATION SCENARIO

Assume the Blog Service Provider (BSP) provides the **static** Blog information adaptively, the blogger's job is to provide his Blog URL, then we can get the **collection** information by analyzing the sidebar widgets/web services or the RSS feed. Then we can compute the **artifact** information the Blog information can store blogger's myFOAF information. Three access steps showed in Figure 3:

1) blogger registers in our service website and fills out the related static information which includes blogger's Blog URL and related FOAF basic information.

2) blogger fills out his Blog related widgets, RSS subscriptions or web services links to involve the web 2.0 services to our system collections such as de.licio.us, bookmark, 43things, flickr and so on.

3) Generating the corresponding JavaScript codes of Personal Expertise Explorer and Personal Social Explorer then the blogger can store them to his Blog.

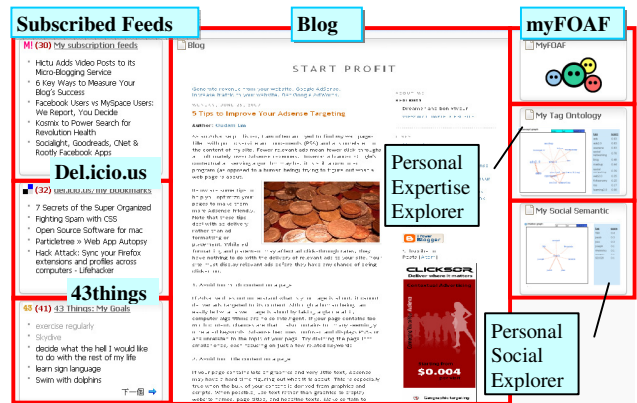


Figure 3. Application User Interface with myFOAF

### 5.1 Personal Expertise Explorer

In the web2.0 application, users can construct his tags in each site, and user can use web2.0's API to subscribe the web service. But the problem is that if user has the account at de.licio.us, 43things and flickr, he has to customize his tags in each site. To solve this problem, we propose the personal expertise explorer to construct personal used tags stored in personal ontology. The interface will get the user's de.licio.us tags, his 43things or his flickr URL through the RSS subscription information. And user's tags information will be integrated as personal MyTags which is stored in the personal ontology. And the personal expertise explore will analyze the blogger's domain expertise according to the personal myTags. Figure 4 shows the personal expertise explorer interface. In this example, this blogger is the expert in the "web2.0" domain, and those related tags the blogger used contain "ads", "ads2.0", "openpne", "social networking" and so on.

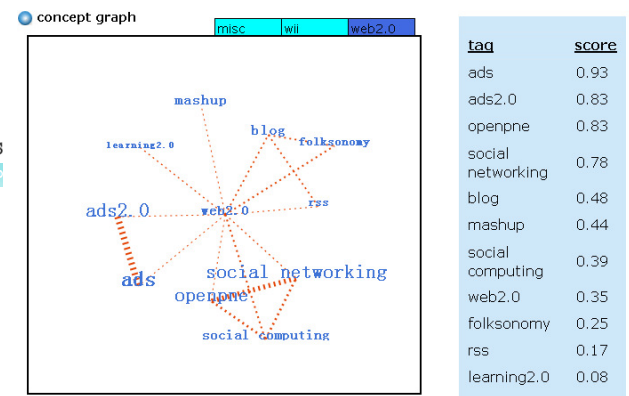


Figure 4. Personal Expertise Explorer

The personal expertise explorer conducts the application by the following data sets:

1. **myTags:** It defines the personal tags used in users' Blog. It can represent that the blogger is familiar with these domains or the blogger may have interests in them.

2. A refined **Collaborative Content Sharing Module** is proposed to enhance the user experience on content sharing. The concept space brings both functions of tag recommendation and concept based search [23]. This research applies the blogger's myTag dataset to generate the myTaxCluster and myConceptSpace.

The myConceptSpace is the left diagram of the Figure 4 which computes the MyTags' relationships between each other from the Universal Tag Ontology and constructs the personal tag cloud. If the blogger would like to tag something to the collection in other service like Del.icio.us, 43Things, or Flickr, the blogger could see myConceptSpace to choose appropriate tag. The thickness of the line represents the relationship of tags such as "web2.0" in Figure 4.

The collaborative tagging system will generate a hierarchy concept space. The tags can be clustered into fewer tags which called myTagCluster. This indicates that the blogger has some expertise or interests in some domains. In the upper side of Figure 4, the blogger has following domain interest tags: "Web2.0", "wii" and "misc".

3. **Expertise Degree:** This represents the blogger's expertise degree in different tag concepts. The higher score represents the blogger is more knowledgeable/recognized in corresponding tag, and the word size will also show relatively larger in the diagram of myConceptSpace. Expertise Degree is computed from the blogger's myTag Usage and the interactions from others shown in Figure 5. The Content Repository represents whole bloggers' Blog in their BSP.

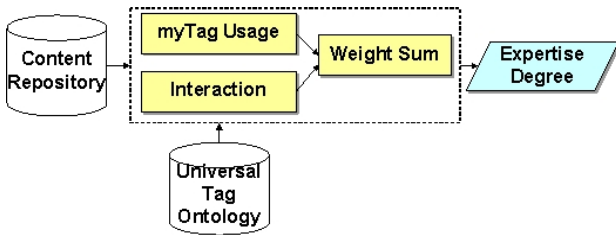


Figure 5. Expertise Degree computation flow

The score is affected from the following parameters:

Blogger's **myTag Usage (Su)**: The used tag shows the blogger's expertise or interests in the tags. But the frequently used tag introduces that the blogger is the expert of this tag domain. Where  $f(t_n, u)$  denotes tag<sub>n</sub> usage of user<sub>n</sub>,  $f(t_n)$  denotes tag<sub>n</sub> usage of all users and  $Su \in [0,1]$  defines each normalized tag usage from 0 to 1.

$$Su = f(t_1, u) / f(t_1), f(t_2, u) / f(t_2), \dots, f(t_n, u) / f(t_n) \dots \dots (1)$$

**Interaction:** These variables come from the interactions of other Internet viewers, which includes the Blog page views (footprint), RSS subscription counts, references (content links and trackbacks), comments and blogroll. The blogroll variable denotes the number which other bloggers add the blogger to their friends list. The more friends list where the blogger is added; the blogger must be the expert in some tag domain. With the more interactions from other Internet viewers, we can conclude that the blogger has the higher expertise in that tag domain.

If user1 has some interaction (B) to user2, we can define the relation (R) as the following formula. Where T(A) denotes the tag set of article A, which is the computation resources to analyze the relation within the criteria of the interaction (B).

$$R = (User_1, User_2, B, T(A)) \dots \dots \dots (2)$$

Then the relation (R) will be decomposed into the Atomic Relations (AR). This information will help to construct the interaction relation diagram. Where t denotes a tag in T(A) and  $w \in [0,1]$  denotes the weight of this AR in R.

$$AR = (User_1, User_2, B, t, w) \dots \dots \dots (3)$$

The HITS algorithm [24] is applied to calculate the hub value and authority value. User relationship graphs are constructed for each tag and the calculated authority values of nodes (bloggers) are taken as the blogger's Expertise Degree. Each user (User<sub>i</sub>) will get an interaction vector (Iu). Where (User<sub>i</sub>, t<sub>n</sub>) is the denotes the weight value of tag<sub>n</sub> about user<sub>i</sub>. If user<sub>i</sub> has no value in tag<sub>i</sub>, the value of (User<sub>i</sub>, t<sub>n</sub>) will be 0.

$$Iu = ((User_i, t_1), (User_i, t_2), \dots, (User_i, t_n)) \dots \dots \dots (4)$$

Then we analyze whole blogger's myTag and posted Blog to get the Expertise Degree (Eu) of User (User<sub>i</sub>) where  $\alpha$  denotes an experience parameter.

$$Eu = \alpha * Su + (1 - \alpha) Iu \dots \dots \dots (5)$$

At last, myConceptSpace shown in Figure 4 presents the blogger's expertise degree, where  $\alpha$  is an empirically tuned parameter.

## 5.2 Personal Social Explorer

With the myFOAF information, the personal social explorer gathers the blogger's friend information and constructs his social network diagram shown in Figure 6. The diagram center is the blogger "ccshih" and the lines connect to all his friends. The line thickness presents the closeness of blogger between his friends. The more rough line represents the higher score which means whom is more close to the blogger. The upper side of Figure 6 shows the domains from personal expertise explorer. Since the blogger may know a lot of friends in different domains such as wii, web2.0, and all. (A domain is defined as a collection of tags filtered by Universal Tag Ontology) The right side lists the friends and corresponding familiarity score. The personal social explorer also lists the blogger's friend relationship with the social network diagram.

The blogger's friend relationship is calculated from the following design. Assume there are m users and n domains, we build a m\*m\*n 3-dimension matrix d in which, M is the number of users, N is the number of domains and  $D_{ijk} (0 \leq i \leq m, 0 \leq j \leq n, 0 \leq k \leq n, i \neq j)$  is the number of interactions from user i to user j related to domain k. The score represents the friend relationship familiarity degree between the blogger and his friend.

$$Score(User_a, User_b, Domain_c) = \frac{D_{abc}}{\arg \max_{0 \leq j < m} D_{ajc}} \dots (6)$$



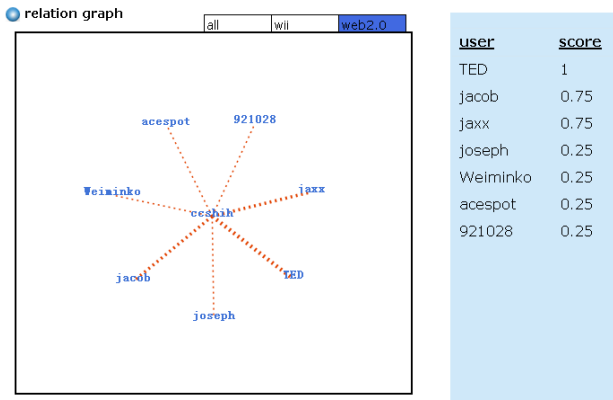


Figure 6. Personal Social Explorer for “ccshih” in “Web2.0” domain

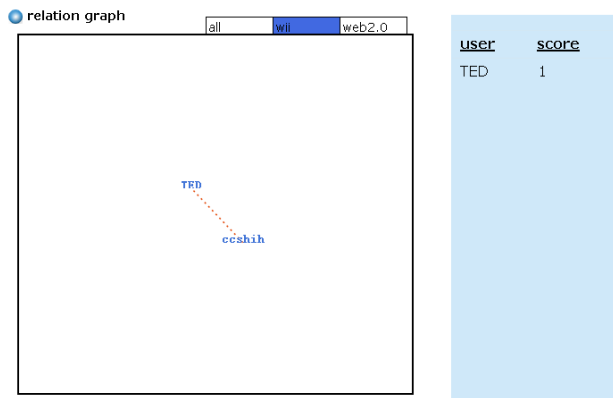


Figure 7. Personal Social Explorer for “ccshih” in “wii” domain

## 6. Evaluation

### 6.1 Experiment Setting

In this work, we collect 36,785 blog users and 500,224 articles as our experimental data from Roodo, a domestic BSP in Taiwan. Table 3 shows the categorical distribution of the data. After auto-tagging process, we generate 60,091 distinct tags from these articles. In average, each article is assigned 4.12 tags.

Table 3. myFOAF data type analysis in popular tags

Category		Category	
Diary	303905	Idol	4648
Non-classified	28054	Food	4439
Writing	27247	Sport	4322
Life	18396	Pet	4319
Travel	10712	Moblog	3904
Art/Design	10222	Job	3638
Music	9784	Finance/Investment	3273
Scholar	8562	Parent	3143
News	7210	Health/Beauty	2656
Reading	6755	Game	2493

Anime/Comic	6597	Management/Marketing	1851
Love	5796	Adult	1075
Movie/TV	5217	Fashion	1050
Hobbies	4962	Commerce/Auction	1037
Networking/Gadget	4957		

First, we apply the auto-tagging process referred in [26] to generate keywords/tags from articles and conduct an experiment to examine effectiveness of auto-tagging process for our dataset. We invited 8 experts to read those randomly selected auto-tagged articles and judge appropriateness of the assigned tags. Examiners are asked to classify each tag into three categories: adequate, inadequate, not related. Total of 1084 articles and 2018 machine-assigned tags were reviewed. Table 4 shows the detailed result.

Adequate	Inadequate	Not Related
90.24%	8.37%	1.39%

Table 4: Personal Social Explorer

The result shows that over 90% of assigned tags are recognized as adequate tags by examiners. The results prove that the program-generated tags could effectively represent the meaning of the corresponding articles.

The generated tags and the interactions between articles are fed as raw materials to calculate users' expertise degree of tags. Total of 168,944 expertise records (user-tag-score triple) are calculated with minimum score threshold = 0.001.

The verification process goes as follows: First, we choose 6 tags which appear frequently in our generated tag set: 棒球 (Baseball), 攝影 (Photograph), 日劇 (Japanese Drama), 樂生 (Loshen, a protest event in Taiwan), 法國旅遊 (French Travel), 韓國旅遊 (Korea Travel). Then we calculate top ten experts of the six tags and ask editors of Roodo to flag unqualified experts.

### 6.2 Evaluation Index

We use the precision concept in IR (information retrieval) to evaluate our results. The precision definition to our evaluation shows in equation 7.

$$Precision = \frac{| \{human - defined\ experts\} \cap \{recommended\ experts\} |}{| \{recommended\ experts\} |} \quad (7)$$

### 6.3 Experiment Results

The evaluation result is shown in Table 5. Besides photograph tag, Precision(n=3) is very significant, even average Precision(n=10) reaches 77.6%.

Table 5. myFOAF Experiment Results

	Precision (n=3)	Precision (n=5)	Precision (n=10)
樂生(Loshen)	100%	100%	100%
棒球(Baseball)	100%	100%	70%
法國旅遊 (French Travel)	100%	80%	70%
韓國旅遊 (Korea Travel)	100%	80%	70%
日劇(Japanese Drama)	100%	100%	78%
攝影(Photograph)	33%	40%	50%

## 7. DISCUSSION

The expert identification research has developed in many different applications and there are also different definitions of expertise. This study aims to discover the experts in an online community, namely in a blogosphere. We determine the expertise level by both semantic and interaction cues in blogs. Because of the rapid growing of the tagging behavior in web 2.0, we use the two cues under the tags. Both the two cues can relate to the tags to represent the characteristics of the blog posts and the bloggers. Generally speaking, tags present the domains and topics information to us and we are able to find the domain experts by tags. The result of our method is acceptable. The average precision is around 0.8 in our simple evaluation. It also shows that tag usage is more powerful. We hope we can work on evaluating a larger blogosphere to verify our method.

We leverage the current Blog content in minimum user effort to support social context. A typical static web form is provided to generate myFOAF profile with sidebar widget support. For the dynamic part we provide linking mechanism to mash-up personal 2.0 services for his/her whole view and up-to-date status of personal SNA. Unlike the disruptive approach of Semantic Web, the solution to extend FOAF and to integrate mash-up service is an evolutionary approach. We build the solution based on Web 2.0 so that any blogger can easily plug in. In addition, any new emerging service can add into myFOAF extension, as long as it provides certain meta semantic and related syndication mechanism such as RSS or Web Services. There is also a feasible way to connect with current architecture; for example, the SIOC (Semantically-Interlinked Online Communities) [27] links other online communities not only by blogosphere but also by collecting all the information from posts and replies. The integration of the SIOC with FOAF and myFOAF can be further studied to enrich the social relations and the expertise information for all web users.

## 8. CONCLUSION

Web 2.0 mostly concerns about the wisdom of the crowd and social collaboration, however we'd like to look back of individual need in personal social aspect. We propose a semantic infrastructure that intelligent agent can further calculate personal reputation, social content, and social capital. Based on the richer social semantic supported by myFOAF, we are able to accumulate blogger's tagging ontology and analyze his/her social network to enable a personalized social semantic service - without interfering web users. In the function of the personal social explorer, we propose the social network diagram to show the blogger's friend relationship in domain cluster. In previous research, we use traditional recommendation method to find someone with higher similarity to the blogger[25]. For our further study, we plan to focus on personalized content recommendation which include blogger, article, community, and finally the advertisement.

## 9. ACKNOWLEDGEMENT

This research was supported by the III Innovative and Prospective Technologies Project of Institute for Information Industry and sponsored by MOEA , ROC.

## 10. REFERENCES

[1] Wikipedia. Blog. <http://en.wikipedia.org/wiki/Blog>

- [2] C. Miller, D. Shepherd. Blogging as Social Action: A Genre Analysis of the Weblog. <http://blog.lib.umn.edu/blogosphere/>
- [3] Dan Brickley, Libby Miller. FOAF Vocabulary Specification 0.9-Namespace Document. Rehydrated Edition, <http://xmlns.com/foaf/spec/20070524.html>
- [4] Leigh Dodds. FOAF-a-Matic. <http://www.ldodds.com/foaf/foaf-a-matic>
- [5] FOAF Explorer, <http://xml.mfd-consult.dk/foaf/explorer/>
- [6] Wikipedia. Social Network Analysis. [http://en.wikipedia.org/wiki/Social\\_network#Social\\_network\\_analysis](http://en.wikipedia.org/wiki/Social_network#Social_network_analysis)
- [7] S. Herring, I. Kouper, J. Paolillo, and L. Scheidt. Conversations in the blogosphere: An analysis from the bottom up. the 38th International Conference on System Sciences (2004)
- [8] C. Marlow. Audience, structure and authority in the weblog community. papers presented at International Communication Association, 2004
- [9] Sara Winge. Controversy about our Web 2.0 service mark. [http://radar.oreilly.com/archives/2006/05/controversy\\_about\\_our\\_web\\_20\\_s.html](http://radar.oreilly.com/archives/2006/05/controversy_about_our_web_20_s.html)
- [10] Thomas Vander Wal. Folksonomy. presentation at Online Information Conference in London, <http://www.online-information.co.uk/o105/day2.html>
- [11] Marieke Guy, Emma Tonkin. Folksonomies: Tidying up Tags? D-Lib Magazine. vol. 12 no 1, <http://www.dlib.org/dlib/january06/guy/01guy.html>
- [12] Peter Merholz. Metadata for the Masses. <http://www.adaptivepath.com/publications/essays/archives/00361.php>
- [13] Luigi Canali, De Rossi. Folksonomies: Tags Strengths, Weaknesses And How To Make Them Work. [http://www.masternewmedia.org/news/2006/02/01/folksonomies\\_tags\\_strengths\\_weaknesses\\_and.htm](http://www.masternewmedia.org/news/2006/02/01/folksonomies_tags_strengths_weaknesses_and.htm)
- [14] Ulises Ali Mejias. Tag Literacy. [http://ideant.typepad.com/ideant/2005/04/tag\\_literacy.html](http://ideant.typepad.com/ideant/2005/04/tag_literacy.html)
- [15] Scott A. Golder, Bernardo A. Huberman. The Structure of Collaborative Tagging. Systems Tech Report Of Information Dynamics Lab, HP Labs, 2005
- [16] Alan, J. and Seligmann D. Collaborative tagging and expertise in the enterprise, WWW'2006
- [17] Zhang J. and Ackerman M. Searching for expertise in social networks: a simulation of potential strategies, GROUP'2005
- [18] Song X., Tseng B., Lin C.-Y., and Sun M.-T. ExpertiseNet: relational and evolutionary expert modeling, 10th International Conference on User Modeling, 2005
- [19] McDonald D. and Ackerman M. Expertise Recommender: a flexible recommendation system and architecture. Proceeding of the ACM 2000 Conference on Computer Supported Cooperative Work (CSCW'00), pp. 231-240, 2000
- [20] Jun Zhang Mark S. Ackerman and Lada Adamic. Expertise Networks in Online Communities: Structure and Algorithms, WWW'2007

- [21] Christopher S. Campbell, Paul P. Magilo, Alex Cozzi and Byron Dom. Expertise Identification using Email Communications, CIKM'2003
- [22] Ian Davis, Eric Vitiello Jr. RELATIONSHIP: A vocabulary for describing relationships between people. <http://vocab.org/relationship/>
- [23] W.-T. Hsieh, W.-S. Lai, S.-C. T Chou. A collaborative tagging system for learning resources sharing. Current Developments in Technology-Assisted Education. Nov. 2006, vol. 2 pp.1364-1368
- [24] Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. Journal of the ACM (JACM), vol. 46, no. 5, pp. 604 – 632, 1999
- [25] T.-M. Tsai, C.-C. Shih, S.-c. T. Chou. Personalized Blog Recommendation: Using the Value, Semantic, and Social Model. Proc. of Innovations in Information Technology. Dubai, pp. 1-5, 2006
- [26] Jay Stu, S.-C. Ben Chou, W.-T. Hsieh, C.-C. Shih. Methods and Systems to Recommend the Content Annotation. Proc. of 18th workshop on object-oriented technology and applications, Taiwan, 2007
- [27] A. Harth, J.G. Breslin, I. O'Murchu, S. Decker, Linking Semantically-Enabled Online Community Sites, Proceedings of the 1st Workshop on Friend of a Friend, Social Networking and the (Semantic) Web (FOAF Galway), Galway, Ireland, pp. 19-29, September 2004.