# An architecture based on semantic weblogs for exploring the Web of People

Tuan-Anh Ta[1], Jean-Marc Saglio[1], Michel Plu[2]

[1] Ecole Nationale Superieure des Telecommunications
46 rue Barrault, 75634 Paris Cedex, France
{ta, saglio}@enst.fr
[2] France Telecom R&D
2 avenue Pierre Marzin, F-22307 Lannion Cedex, France
michel.plu@rd.francetelecom.com

**Abstract.** This paper presents an architecture that helps users to collect and qualify meta-information about web resources within a weblog and then share them with a group of people. Our architecture is a peer-to-peer network in which each peer contains a personal knowledge base represented in RDF. The designed weblog model is sufficiently robust, and the semantics of our data model is detailed to allow integration in our system of other Semantic Web applications as well as existing weblogs. We construct a "web of people" exchange protocol built on two primary services: the first one to query distributed personal knowledge bases, and the second one to enable user awareness based on metadata exchanges in asynchronous mode.

## 1 Introduction

Recent publications show that collaboration is an important factor for quality in a new information society [1]. Users of the WWW participate more and more often in virtual communities that support them in sharing knowledge and reusing each other's best practice. Plu et al. [2] have proposed a new paradigm for the web usage that they called the Web of People. They supposed that users rather than individually searching and collecting information, share and offer their comments about interesting web resources. Within such a framework, users have to manage themselves a knowledge base as a kind of personal memory that helps them to re-find their favorite visited resources. By sharing such meta-information with other each, a user becomes an autonomous and responsible mediator of information.

Compared to other information retrieval and management systems, our system is constructed on three principles: manual instead of automated collaborative information filtering, distributed knowledge instead of shared centralized knowledge, and a person-centric approach instead of a community centric approach.

**Manual collaborative information filtering.** Users in the Web of People will do much of the collecting, filtering, qualifying and organizing of the information. They annotate resources of interest by meta-information in their knowledge base. Instead of

relying only on web search engines or other automated recommender systems they exchange their collected meta-information to help each other in the community to find out needed information. Thus a user can perform efficiently a search over high quality pre-filtered information in the system. Our approach is to use the distributed intelligence of users instead of a centralized automated system.

**Structured and distributed knowledge.** In contrast to the centralized indexing approach of search engines, our system relies on distributed knowledge bases that provide structured annotations of web resources. Since such structured metadata is accessible by machine, we can enable useful automation or automated assistance services. The Semantic Web provides a set of standards and tools that directly address the issues of access to structured, shared metadata. In particular it defines a common data model RDF [3] that supports free structured metadata for web resources.

**Person-centric approach.** Some large organizations spend a long time developing standard terminologies that can then appear cumbersome and inflexible when applied to a given specific personal problem. Instead of using a common ontology, users can define their own terminologies. We support them in linking their own terms to existing ontologies which might be adapted for reuse. This bottom-up approach allows us to make better trade-offs between flexibility and reuse of ontologies.

These guidelines of the web of people are quite far different from traditional approaches for building information sharing systems. For example such traditional systems often use a shared common and centralized ontology to manage their information. The Web of People follows an autonomous and asynchronous bottom up approach for building the distributed ontologies it uses as its backbone for managing information. This approach is much more related to the new paradigm of semantic community portals which are using semantic web technologies where multiple ontologies can be managed separately but integrated using semantic links between their elements [4].

One key idea in the Web of People is to help users to build those links between their ontologies by relating their personal knowledge base with those of their acquaintances in order to support them to share information more efficiently. By doing so, users are creating a meta-view of the Web where recommended resources navigate from users to users instead of having users navigating from a resource to another one using the hyperlinks of the WWW [2].

In order to support this distributed and bottom up approach to build such a meta-information structure, a reference architecture and common protocols have to be defined which are the main topics of this paper.

The proposed architecture provides a common model for the representation of personal knowledge and is largely inspired by the development of Weblogs[1] and reuses many of their concepts. In fact, Weblogs are often used in order to publish on distrib-

[1] For on definition of Weblog see : http://en.wikipedia.org/wiki/Weblog

uted servers some personal observations about any resources, thus providing an interesting source of metadata about those resources. These observations are published and read by others who might also cite them in their personal new observations which are then also published and read by others. Weblogs are also highly connected and so users can easily navigate through Weblogs belonging to multiple users.

One common usage of Weblogs is to syndicate contents from multiple sources using a specific tool providing an aggregated view of new messages coming from multiple Weblogs. The protocols presented in this paper aims to support this process of navigation through multiple Weblogs and to produce aggregated views of these distributed knowledge base. Two protocols are complementary defined. One is for synchronous exploration of the web of people and the second is asynchronous in order to develop user awareness of the evolution of the web of people. They are composed of a reference data model syntactically based on RDF, a query language, and a message based protocol which will be presented in following sections. The next section will present the global architecture. The last sections relate this work to other's and conclude.

## 2. The general architecture of the Web of People

To keep in memory addresses of favorite resources we often use the bookmark functionality of web browsers. Unfortunately, there are still inconveniences of this tool such as the difficulty for retrieving information, and the availability of this precious data from any internet device. Nowadays, some users prefer Weblog [5] to present and archive their observations on favorite resources. A weblog is defined in a draft of the Oxford English Dictionary as "a frequently updated web site consisting of personal observations, excerpts from other sources, etc., typically run by a single person, and usually with hyperlinks to other sites; an online journal or diary". The personal observations are formatted as individual posts of the weblog. Each post contains typically a title, user comments and hyperlinks to relevant resources. We can visualize the posts by theme and/or in reverse chronological order. We call a topic a category for indexing posts.

In [1] Plu et al., have defined a general architecture for a Web of People on which each user has their own active personal “Web of People” server (Figure 1). The server comprises of an http server to communicate as a peer of the network. A web-based application on this server provides the user with the functionalities to manage his or her personal taxonomy and information reviews (i.e., observation posts in the weblog) within a knowledge base. The user can share his or her reviewed information with his or her friends so that they can access and read it via their active server. In the general architecture of the Web of People, a strategy for controlling information exchanges defined through trusted relationships was proposed. The accessibility of reviews, and thus the exchange of (meta-)information between users depends on a distribution list attributed to each topic. It groups together users allowed to access all information having a review indexed by the topic.

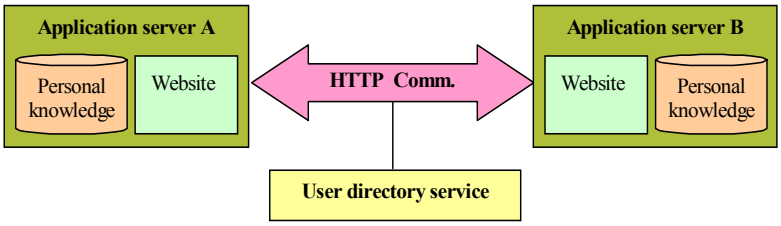


**Figure 1**: The general architecture for the Web of People in SoMeONe

In order to manage unique user identifiers and the associated addresses of their personal site, a central server offering the user's directory service is installed. If for the moment this service is centralized in one server, we imagine that it can be easily distributed following the traditional DNS approach of the internet (cf. IETF RFC 1035). Each time a user installs a new personal site, he must declare his name and the location of his or her server to the central server. From then, an entry for his or her name exists in the user directory managed by the central server.

Furthermore, some mediation services can be integrated in order to provide users with automated assistance services such as aggregating similar topics of different personal knowledge bases.

In our architecture (Figure 2), we separate the communication from the presentation layer, i.e., the http protocol. Thus, the Web of People relies on a peer-to-peer network so that searching information is more efficient. Each active personal server is now composed of a Weblog, a communicator and a personal knowledge base (PKB). The PKB is used to manage the user's taxonomy and collected (meta-) information in a persistent memory. The weblog is the HTML view of the PKB allowing the user to publish his or her personal knowledge on the traditional Web. The communicator aims to support communication protocols between user peers and/or a peer with the central server. Although there is an existence of the central server in the architecture, it is always a pure peer-to-peer network. No index data for searching are stored in the central server. It only plays a role of managing user identifiers that is closely related to the one of DNS server in the WWW.

In order to develop a reference architecture for the Web of People, we define in this paper the following components:

(1) a weblog-based data model that allows users to capture their observations about web resources and organize them within a personal taxonomy - we use RDF as the basis of data representation which helps the system to be easily integrated with other semantic web applications and existing weblogs;

(2) a model for querying distributed personal knowledge bases in the Web of People - this feature enables a user to access/reuse the (meta-)information of users belonging to his her trusted relationships network, and

(3) a messaging-based protocol which allows users to exchange their (meta-) information in a "push" mode for improving user awareness.

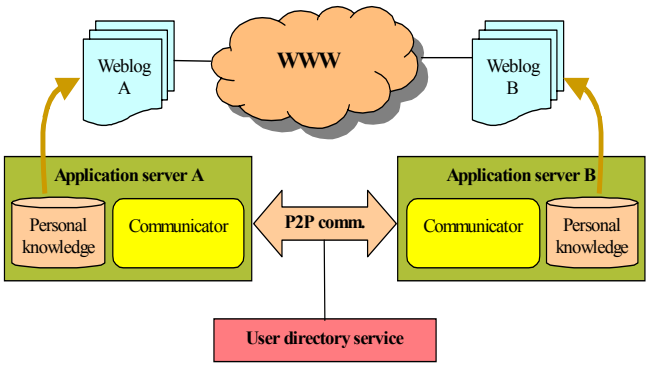


**Figure 2**: The peer-to-peer architecture for the Web of People

## 3. The semantic weblog data model

A PKB includes a personal taxonomy and individual posts. The personal taxonomy is modeled as a hierarchy of topics so that each post can be indexed by one or more topic. We need to identify individual posts and topics as personal resources having unique URI. The weblog is an HTML view of this PKB. This view is composed of URLs (i.e., web pages) automatically associated to each URI of personal resources. In the weblog the URL for each individual entry is often referred to as a "permalink".

We used RDF as the basis to describe weblog resources. Our Weblog model is inspired from the bookmark schema developed in the Annotea project [9], but we adapt this original schema to distributed semantic weblogs.

The global schema of the Web of People is presented Figure 3. In order to be exhaustive the next sections describes precisely each elements of this schema and its associated semantics. Those specifications can be browsed easily according to specific interests. The main contributions of this schema in order to build semantic portals are précised in section 3.3. Thus, readers may focus their attention to this specific section and find complementary information in the others. The name space of this schema is: http://purl.org/webop/1.0/.

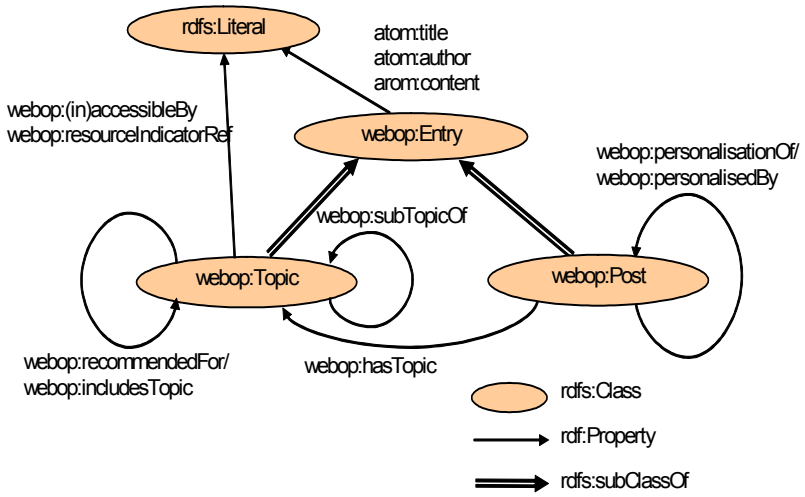


**Figure 3**: The schema of the weblog model

### 3.1 Instances and Classes of the Webop schema

Using RDF we needed to define a URI scheme (conforming to the RFC2369 [7]) for defining identifiers of managed resources. Each managed resource is an instance of a class of the schema. Given a name of each managed resource called uid, the URI must follow the syntax "urn:webop:" + uid where "webop" is the scheme name used for the Web of People. The unique uid of a user is controlled and managed by the user's directory service of the architecture (cf. Figure 2). We reserve urn:webop:all as a special user identifier to indicate the whole list of users in the Web of People.

**webop:Entry.** A user's weblog base is defined as a finite set of personal resources corresponding to the posts and the topics created by the user. *webop:Entry* is an abstract class. Each personal resource must be typed either as topic or post by assigning it a *rdf:type* property indicating *webop:Topic* or *webop:Post*, respectively. In order to enable distinguishable location for personal resources, each of them must have a unique URI as "urn:webop:" + rid + "$" + uid where rid is a unique local name attributed to the personal resource and managed locally by the personal server of the user. The resource descriptions are always found in the owner's base. For example, all the URIs ending with "$user1" must be reserved for personal resources in the PKB of a user named user1. This mechanism is very useful for the efficient querying of personal resources in distributed bases.

**webop:Topic.** A topic resource defines an informal category for indexing posts. Topics may have subtopics and may refer to terminologies in more complex ontologies (see sub section 3.4).

**webop:Post.** A post is a user observation about one or several web resources. It contains typically a title, a content and links to the observed resources.

### 3.2 Metadata properties

Using RDF, any kind of properties can be defined to any managed resources in the Web of People. Those properties can be defined within any already defined schema using a namespace. For example to define some metadata about entries like title or others, the traditional Dublin Core namespace [13] can be used but the new ATOM schema can be also found more relevant for weblogs [14]. Some properties of the Atom schema are essentials for a post which are:

- *atom:link*, the permalink of the resource in the HTML view and
- *atom:created*, the created date
- *atom:modified*, the last modified date
- *atom:issued*, the time that the resource was issued
- *atom:author*, the identifier of the owner. For each personal resource, a *atom:author* property is automatically created that conforms with the specified syntax for personal resource URIs, e.g., *urn:webop:MusicResources$user1* having *urn:webop:user1* as the author.
- *atom:content*, conveying the content of the post. A post can contain one or more content elements.

To describe users the FOAF (Friend Of A Friend) namespace [15] can be also used.

We impose that a property can only be defined by the owner of the entry defined in the domain of the property. This applies also for the properties defined below.

### 3.3 Webop properties

Those properties let users to index their post and control their accessibility in order to improve management facilities and automate their sharing.

**webop:subTopicOf.** Many topics can be created for a personal taxonomy, and there are subsumption relationships between them. Given two topics $a$ and $b$, $a$ is subsumed by $b$, denoted by $a \underline{\pi} b$, if $a$ has a *webop:subTopicOf* property with the value of $b$. *webop:subTopicOf* is transitive and reflexive. We impose that topics a and b belong to the same user.

**webop:hasTopic.** Individual posts are associated with topics of the taxonomy via *webop:hasTopic* properties. We denote $I(a)$ an interpretation of topic $a$ which consists of a set of posts indexed with the topic. This interpretation is valid when for every topic $b$, if $a \underline{\pi} b$ then $I(a) \subseteq I(b)$. The PKB must be built on valid interpretation, i.e., if $x$ is a post associated with $a$, $x$ is also associated with $b$.

Figure 4 shows a sample of PKBs. In this figure, posts are represented as triangle and their membership to the interpretation of a topic is indicated by a line. Bold-arrows indicate subsumption between topics. For readability, we do not represent the entire topic subsumption relation but its transitive reduction, in which reflexive and transitive arrows are suppressed. We also suppress the lines representing transitive interpretation. For instance, the post entitled "r1" belongs to the interpretation of the topic Rock as well as MusicResources.

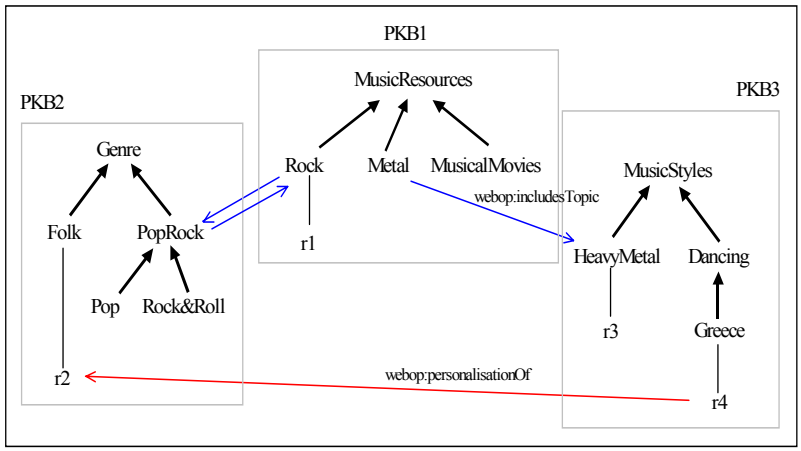


**Figure 4:** A sample of PKBs

**webop:resourceRef.** A post is considered to represent a user's observation on one or several web resources. Such entries allow the user to re-find collected resources through annotations created by him or her. We use *webop:resourceRef* property to associate a post with referred resources. A web resource is typically addressed by an URL.

Those previous properties of our schema are respectively similar to the following Annotea properties: *subTopicOf*, *hasTopic*, and *recalls* [9]. We re-defined these vocabularies in our context of weblog data. Moreover, a post can have several *webop:resourceRef* properties while a bookmark may only have one *recalls* property.

**webop:accessibleBy.** The value of this property is a user identifier. A topic has this property if the author allows the user with this identifier to access it (i.e., read the descriptions and posts concerning the topic). One can read a post, and thus the post is accessible by a user, if he or she has access permission for at least one of the post's topics. In order to be similar to the http access control mechanism, some applications only allow authors to make accessibility a topic while all its super-topics are also accessible. On the other hand, some others enable the propagation of *webop:accessibleBy* to subtopics, e.g., in SoMeONe, an initial application for the Web of People [6].

**webop:inaccessibleBy.** If a user wants to give access of one of his or her topic to everyone except a particular user, for instance urn:webop:user1, he or she can put a *webop:accessibleBy* with the value of urn:webop:all, and *webop:inaccessibleBy* with the value of urn:webop:user1.

A topic is accessible by urn:webop:user1 if and only if it has a property *webop:accessibleBy* with the value of urn:webop:user1 or urn:webop:all and has not a property *webop:inaccessibleBy* with one of those two values. A post is accessible if one of its topics is accessible and none of them are inaccessible.

### 3.4 Properties to contribute to the Semantic Web

In order to contribute to the Semantic Web vision, we support users in linking their own terms to existing ontologies. For a minimal model we defined the following properties in order to have mappings between users' topics and concepts of shared ontologies. In a further version, more complex mappings as in [10] will be allowed by defining other properties (maybe with the help of OWL descriptions [12]). Those properties should contribute to a better automation for exchanging, aggregating, filtering, classifying information describe with personal taxonomies. This is a major contribution of our work since traditional approach is to build a community portal base on a common shared ontology.

**webop:subjectIndicatorRef.** Like the subject indicator concept of the XTM model [8] which is useful for mapping two different topic maps, we defined the *webop:subjectIndicatorRef* property that helps refer a topic to a concept of a more complex ontology (i.e., any URI that belongs to a large shared and accepted ontology). Two topics with the same subject indicator should be considered as being semantically similar. Some inferences can be made using this property but are application-dependent. It lets user's personal ontology to be related to more complete and global ontologies which will be available on the semantic web.

**webop:personalisationOf.** In practice, a user might wish to add his own metadata to a post already created by another user. To do so he or she should create a new post having a *webop:personalisationOf* property with the value of the personalized post's URI. This kind of link allows the user to prevent a duplication of effort in case the observation has been done by his or her colleague. In Figure 4 the personalisation arrow indicates that r4 is a personalization of r2, so that it copies metadata descriptions such as comments, resource links of the original post.

**webop:personnalizedBy** is the inverse relation and is created automatically by the server of the owner of the personalized post. This is assumed by the communication protocol defined in section 4. This back link relation is justified in [16] for ego network analysis.

**webop:includesTopic.** This property is essential to the bottom up process of the distributed ontologies creation. It let users to define that another user's topic is included in it's topic. It is semantically similar to the subtopic property but for two topics which do not belong to the same users.

**webop:recommendedFor** is another way to relate his/her topics with others one-means that a user think that his topic should be included in another user's topic. But only this other user can decide to create this inclusion relation In the Web of People, users are encouraged to create trust relationships on which mutual exchanges are allowed on users' similar topics. A user discovering that one topic *a* is recommended for one of his topics *b* is supposed create the *webop:includesTopic* property between *b* and *a* if he/she confirms this recommendation. Then a trust relationship is established between the users based on these topics and thus are considered to be equivalent. This property is also used in the communication protocol (see sect 4).

## 4. Communicating with peer users

In order to support personal knowledge exchanges in the Web of People, we design two application-independent services: one for accessing and exploring remote knowledge bases by querying and the other for users' awareness by sending and receiving posts based on a messaging protocol.

### 4.1 Querying personal knowledge bases for exploring the Web of People

Navigation seems to be a critical functionality for accessing information. The Web of People allows users to navigate the Web with an alternate hyperlink structure. Instead of following links in HTML pages, they will first navigate through topics in distributed personal knowledge bases, sometimes consulting referred resources in some posts of the currently viewed topic. Thus, the basic queries of the PKBs will be produced for navigation and visualization needs in each user personal weblog. Moreover this querying process has to support the need of Weblogs users to aggregate post from distributed PKBs. To aggregate those contents we require the query protocol to be able to manage distributed queries in several PKBs. We also added in the requirements of the querying protocol to support the discovery of new topics in the web of people.

To query our RDF models many RDF query engines have been developed [17]. Unfortunately, none of them support distributed queries. Edutella [18] is a recent research on peer-to-peer networking based on RDF. The Edutella Query Exchange Language is intended to be a standardized query exchange mechanism for RDF metadata stored in distributed RDF repositories. It provides the syntax and datalog semantics for an overall standard query interface across heterogonous peer repositories. Therefore we can use the Edutella query service for searching entries in weblog networks. However, the Edutella query language family RDF-QEL-i is not particularly

well suitable for navigating in the hierarchical nature of weblog entries. We need a query language expressive enough to allow users to browse their own weblog or others' one by choosing a topic of interest and then moving up or down. We believe a specific query processing, notably with an indexing on the hierarchies of weblog entries, could be more efficient than the Edutella general approach. Moreover, the new query can also help to facilitate the implementation of caching weblog metadata on browsing We also define some important feature not identified in the Edutella language.

Since the extensive description of our query language can be the focus of a specific paper we provide here only some ideas of the essential features. This query language is largely inspired by the LDAP protocol [18]. Each query includes a search scope and a filter. Intuitively, let first view the network of PKBs as a global graph of personal topics being linked by *webop:subTopicOf* and *webop:indudesTopic* properties and the inverse of the *webop:hasTopic* property. Then the result of a query will contain all the entries which satisfy the conditions specified in the filter part of the query and found by recursively following the links in the entries graphs of the PKBs according to the scope part of the query. A scope defines the initial topic where entries will be searched, the direction and the length of the links which can be followed.

To clarify how it works, let see an example using the PKBs in Figure 4. To simplify we consider that all topics and Post are accessible to all users. But in the general case, any entry in an answer of a query has to be accessible to the user issuing the query.The following query will aggregate all the topics and Posts created after the 1st of august 2004 which can be browsed from the topic MusicRessources of user PKB1.

```
GET webop:hasTopic, atom:title, webop:includesTopic, we-
bop:subTopicOf, atom:author, webop:ressourceRef
SCOPE sub(MusicRessources$PKB1)
FILTER (rdf:type=webop:Topic) or (rdf:type=webop:Post and
atom:created >= 2004-08-01)
```

The scope sub indicates a scanning of all entries descending the *webop:subTopicsOf webop:includesTopic* and the inverse of *webop:hasTopic* lattices. The filter is used to keep all the Topics and only the Post created after the 1st august 2004. Then an extract of the answer will be:

```
<webop:Topic rdf:about="urn:webop:Metal$PKB1">
<atom:title>Metal</atom:title>
<atom:author>urn:webop:PKB1<atom:author>
<webop:subTopicOf
rdf:resource="urn:webop:MusicResources$PKB1" />
<webop:includesTopic rdf:resource="urn:webop:HeavyMetal$PKB3"/>
</webop:Topic>
<webop:Topic rdf:about="urn:webop:HeavyMetal$PKB3">
<atom:title>Heavy Metal</atom:title>
<atom:author>urn:webop:PKB3<atom:author>
<webop:subTopicOf rdf:resource="urn:webop:MusicStyles$PKB3"/>
</webop:Topic >
<webop:Post rdf:about="urn:webop:r3$PKB3"
<atom:title>r3</atom:title>
<atom:author>urn:webop:PKB3<atom:author>
```

```
<webop:ressourceRef>http://www.hard_plus.com
</webop:ressourceRef>
<webop:hasTopic rdf:resource="urn:webop:HeavyMetal$PKB3"/>
</webop:Post>
```

The answer will also include the descriptions of the subtopics MusicalMovies$PKB1, Rock$PKB1, then it will also include the description of the post in PKB1 entitled "r1". But it will also include the included topics PopRock$PKB2 and its subtopics Pop$PKB2 and Rock&Roll$PKB2.

This example shows how the query language allows distributed queries in multiple distributed PKB and how it allow the discovery of new topics which can be found following newly created *webop:subTopicOf* or *webop:includesTopic* properties in remote PKBs. A single query in this language can collect many pieces of information as much as needed which can be after visualized in a user interface. The quantity of displayed information can then be controlled locally without new queries.

The language allows different scope similar to LDAP such as up, down for one step exploration, sub for recursive exploration and base for exploring only equivalent topics. We have also extended the language by allowing the definition of a maximum distance to explore from the initial topic.

### 4.2 Sending and receiving metadata for users' awareness

In addition to the query functionality as a pull mechanism to exchange metadata, we also support push exchanges of metadata through a messaging-based protocol. This protocol can provide users with awareness of new recommended posts/topics and also feedbacks from others about their contributions.

A message is sent each time a post or a topic is created. It is sent to each user who has the accessibility to the new entry. Here is an example of metadata sent from the user of PKB2 to the user of PKB3. The exchanged metadata can be considered as a message to notify PKB3 of new topic Folk in PKB2 which is recommended for topic Dancing in PKB3. The *webop:recommendedFor* property defined between topics or between a post and a topic is communicated in the message in order to precise to the recipient why he/she should be interested by the recommendation.

```
<webop:Topic rdf:about="urn:webop:Folk$PKB2">
  <webop:recommendedFor rdf:resource="urn:webop:Dancing_1$PKB3" />
</webop:Topic>
```

For messages related to new posts the property *webop:personalisationOf* a post is indicated to the owner of this post when it is created. The inverse *personalizedBy* can then be created in the recipient PKB.

For each recommendation message, a message in the form of metadata may be created to reply with acknowledgement content, which depends on how the recipient considered the recommendation. For example when receiving metadata of recommended topic Folk from PKB2, the user of PKB3 can formulate queries to discover more about it, e.g., the posts indexed with this topic. He then sends metadata to

PKB2 for an acknowledgement if he finds the topic interesting. The *webop:readBy* property indicates the resource being appreciated by the user. In order to express unappreciation, we can employ the *webop:ignoredBy* property. Here is an example of such message.

```
<webop:Topic rdf:about="urn:webop:Folk$PKB2">
  <webop:readBy>urn:webop:PKB3</webop:readBy>
</webop:Topic>
```

Making acknowledgement is not mandatory. But we are convinced that they are essential for encouraging the development of the Web of People. Our current experimentations of information sharing systems show that by receiving such feedback, users can see in their contributions to such a sharing network a way to satisfy their needs of social relationships**.**

## 6. Related work

An RDF schema for shared bookmarks has been contributed in [9]. This work is a part of the Annotea project that aims to provide a framework for rich communication about Web pages through shared annotations. Shared bookmarks can be stored as metadata in one or several user selected servers and retrieved by other users who have subscribed to and have permission to access those servers. In opposite this centralized approach, the Web of People enables distribution of PKBs in a peer-to-peer network.
Many personal knowledge sharing system have been also developed. The ePerson project [21] is most closely related to our work. The ePerson snippet manager is an application built upon the ePerson infrastructure that enables a virtual network of personal knowledge bases. Each user has a RDF knowledge base to store their user profile, preferences and the information snippets of interest to them. We have also already cited the Edutella framework [18] and compared some technical choices with our proposal.
But related to those systems our goal, and the focus of this paper, are to define a minimal architecture and standardized protocols in order to enable users to exchange metadata whatever the sharing knowledge system they use. Our proposal starts from a current Web usage where each user has a weblog for their personal memory. These users will exist and will contribute to the web of people as long as their weblog exists. Therefore we need to integrate legacy weblogs and create extensions to current weblogs software in order to have an application server compliant with our proposal for each user's weblog. Using RDF as the basic layer of our information model, the integration of existent weblogs should be really simple.
We are currently in the process of implanting a prototype of this architecture using open source software such as Jena [22], Jabber [23], and Tomcat. But this software will have to remain a simple tool for editing and publishing the user’s semantic weblog.

## 7. Conclusion

We have presented an architecture to enable the development of the Web of People where users can exchange metadata about WWW resources. This architecture is based on a semantic weblog model. Compared to the current weblog supports we first proposed a well defined semantic data model to represent personal knowledge. It provides accessibility management and links creation between personal and shared ontologies in order to facilitate common understanding. Our architecture relies on a query model suited for exploring distributed personal knowledge bases and navigating in the meta-data hierarchical structure of the Web of People. Finally we specified a messaging-based protocol to allow a push mechanism of metadata exchanges.

The promoted usage of the system is to facilitate and encourage the sharing of information between communities of practice, such as groups of researchers, amateur clubs or peer-to-peer ad hoc communities, etc. The services offered by the system propose a distributed alternative to the centralized search engine or web directory to discover and find web resources. In this approach, the selection of web resources as the result of a specific query is no longer decided by a single service provider. In opposite, it is the result of a distributed collaborative filtering process based on a community of relationships composed of persons carefully chosen by each user. In addition to a new search and discovery service for the WWW, this meta-web (i.e., a web about web resources) infrastructure can also largely contribute to the development of the Semantic Web. Indeed, it should stimulate the bottom-up production of personal ontologies related to each other, and enrich the Semantic Web with metadata using those ontologies. But the challenge is now to make the Web of People scale up with the size of the Web. That is why the integration of any software component able to manage meta-data about web resources is crucial. We proposed in this paper an architecture which can be a starting point to build a reference architecture for the integration in a web-like infrastructure of any weblogs and more generally any meta-data management software.

## Acknowledgement

This work is partly supported by France Telecom R&D with the contract number 42703042.

## References

1. Andreas Neus. Managing Information Quality in Virtual Communities of Practice. In The 6th International Conference on Information Quality at MIT, 2001.
2. M. Plu, L. Agosto, P. Bellec, W. Van De Velde. The Web of People: a dual view on the WWW. Twelfth International World Wide Web Conference, Budapest, 24-26 Mai, 2003.

3. Ora Lassila and Ralph R. Swick. Resource description framework (RDF) Model and syntax specification. Recommendation, W3C, February 1999.
4. SWAD-Europe deliverable 12.1.5:Semantic Portals-Requirements Specification http://www.w3.org/2001/sw/Europe/reports/requirements_demo_2/
5. Wikipedia encyclopedia – Weblog definition. http://en.wikipedia.org/wiki/Weblog
6. M. Plu, L. Agosto, L. Vignollet, P. Bellec. Someone: a cooperative system for personalized information system. In the proceedings of the ICEIS 2003 Conference, 2003.
7. T. Berners-Lee, R. Fielding, U.C. Irvine, L. Masinter. Uniform Resource Identifiers (URI): Generic Syntax. RFC 2396, August 1998.
8. S. Pepper, G. Moore. XML Topic Maps (XTM) 1.0. TopicMaps.org Specification, 2001.
9. M. R. Koivunen, R. Swick, J. Kahan, E. Prud'hommeaux. Annotea Shared Bookmarks. KCAP03 workshop October 25-26, 2003 at Sanibel, Florida.
10. François Goasdoué and Marie-Christine Rousset, Querying Distributed Data through Distributed Ontologies: a Simple but Scalable Approach, IEEE Intelligent Systems, 2003.
11. Tzitzikas, Y. and Meghini, C. Query evaluation in peer-to-peer networks of taxonomy-based sources. Proceedings of the 10th International Conference on Cooperative Information Systems, CoopIS'2003. Sicily, Italy.
12. D. L. McGuinness, F. v. Harmelen. OWL Web Ontology Language Overview. Recommendation, W3C, February 2004.
13. Dublic Core Initiative : http://dublincore.org/
14. AtomEnabled. http://www.atomenabled.org/
15. FOAF Vocabulary Specification : http://xmlns.com/foaf/0.1/
16. Ohmukai, I. et al., 2003. Egocentric Search Method for Authoring Support in Semantic Weblog. Workshop on Knowledge Markup and Semantic Annotation, 2003.
17. Maganaraki, A. et al., 2002. Ontology Storage and Querying. *Technical Report No 308*. ICS-FORTH
18. Nejdl, W. et al, 2002. Edutella: a p2p networking infrastructure based on rdf. *In Proceedings of the eleventh international conference on World Wide Web*. pages 604-615. ACM Press.
19. Wahl, M., Howes, T. and S. Kille. Lightweight Directory Access Protocol (v3). RFC 2251, July 1997.
20. Bouthors Vincent, Dedieu Olivier, 1999. Pharos, a collaborative infrastructure for web knowledge. In proceedings of the third European Conference on Research and Advanced Technology for Digital Libaries, ECDL'99. Lecture Notes in Computer Science, 1696, pp. 215-233, Springer-Verlag Inc..
21. Personal Information Infrastructure - ePerson. http://www.hpl.hp.com/semweb/e-person.htm
22. Jena 2 - A Semantic Web Framework. http://www.hpl.hp.com/semweb/jena2.htm
23. Jabber website. http://www.jabber