RDF/Topic Maps mapping and application framework
— Realization of Seamless Knowledge —

Motomu Naito
Knowledge Synergy Inc.
3-747-4, Kusunokidai, Tokorozawa,
Saitama 359-0037, Japan
http://www.knowledge-synergy.com
motom@green.ocn.ne.jp

ABSTRACT
The current Web deals with information resources, such as HTML pages and papers, so to speak the target of the Web is physical files. However, it is thought that what we need actually is not the information resource itself but the concept (= subject) or knowledge included in it. Several technical elements for dealing with the subject and for realizing the subject centric processing are becoming possible. This paper explains those technical elements and considers apply them to various domain.

Categories and Subject Descriptors

General Terms
Management, Documentation, Standardization

Keywords
Topic Maps, RDF, Published Subjects, Ontology, Remote Access Protocol, Query Language, Seamless Knowledge, Application Framework

1. INTRODUCTION
Many people are almost drowning to the information tsunami. And, it is anxious for realization of the mechanism for finding required information when needed. For that purpose, it is needed not only the information system that can handle physical files as the target objects, but also the system that can handle the subjects. Because of what we need actually is not the information resource itself but the concept (= subject) or knowledge included in it. And it is needed to realize the information system that can systemize and organize the information and knowledge based on the subjects.

Topic Maps and RDF are a technical element used as the base for realizing the subject centric processing system. RDF is created by W3C, on the other hand Topic Maps is created by ISO. Both Topic Maps and RDF consist of multiple standards. Topic map has the intention of supporting the construction of a highly efficient index to an information resource, in order to make information easy to find. On the other hand RDF has the intention of offering the metadata structuring the information resource and the base of logical inference for realization of the Semantic Web. In the recent circumstances of glut of information, there are infinite variety of the information to be required according to position, situation, people and so on. And the information to be required change according to view, timing, granularity and so on. Both standards family can be used in mutual complement. And they are expected to offer the means for people accessing requiring information when people require exactly.

In this paper, I explain the technologies including RDF and Topic Maps which can be used to realize collocation and organization of the information/knowledge based on the subjects (referred to as Seamless Knowledge). Those technologies are as follows:

- RDF and topic map
- Published Subjects
- Ontology
- Remote Access Protocol
- Query language

I also apply those technologies to some domain. Furthermore, I consider an application framework as general application structure.

2. RDF and Topic Maps
2.1 Standards family of RDF and Topic Maps
RDF and Topic Maps are technologies for describing semantically and structurally the subjects of information resources, and they are also information resources themselves. Both RDF and Topic Maps consist of multiple standards. They consist of the syntax for description, data model, restrictions language, query language, etc. Comparison of a RDF standard family and a Topic Maps standard family is shown in the following figure.

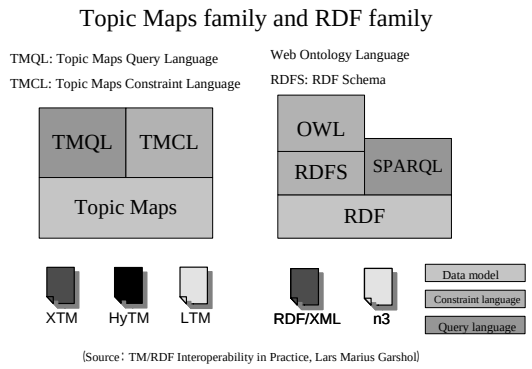

Figure 1. Standards family of RDF and Topic Maps

2.2 RDF to Topic Map Mapping

In SWBPD (Semantic Web Best Practices and Deployment Working Group) of W3C, RDF/Topic Maps Interoperability Task Force was established. Refer to http://www.w3.org/2001/sw/BestPractices/RDFTM/. And the draft of "A Survey of RDF/Topic Maps Interoperability Proposals" (29 March 2005) was published. Refer to http://www.w3.org/TR/2005/WD-rdftm-survey-20050329/.

As an example of the mapping method between RDF and Topic Maps, I introduce the method put forwarded by Lars Marius Garshol. It is the method of using a RDF vocabulary called RTM (RDF to topic maps mapping), and specifying the correspondence between RDF component and Topic Maps component. RTM consists of the RDF properties and resources which are shown below.

2.2.1 RTM properties

(1) rtm:maps-to property

It defines the mapping with a RDF property and a Topic Maps component.

(2) rtm:type property

It specifies the type of the Topic Maps component created by mapping.

(3) rtm:inch-scope property

It specifies the scope of the Topic Maps component created by mapping.

(4) rtm:subject-role property

In case a RDF statement is mapped to the "association" which is one of the Topic Maps components, it specifies the role type which correspond the subject of the RDF statement in the "association".

(5) rtm:object-role property

In case a RDF statement is mapped to the "association" which is one of the Topic Maps components, it specifies the role type which correspond the object of the RDF statement in the "association".

2.2.2 RTM resources

The resources specify the Topic Maps components which are mapped to in the RDF to Topic Maps mapping. There are the following resources.

(1) rtm:basename

(2) rtm:occurrence

(3) rtm:association

(4) rtm:instance-of

(5) rtm:subject-identifier

(6) rtm:subject-locator

(7) rtm:source-locator

3. Published Subjects

It is thought what we need actually is not the information resource itself but the subject included there. Published Subjects is a mechanism which enable person and computer to identify subjects (topics). And it is permanently published on networks and is aimed at making easy share/exchange Topic Maps. Recently the target of the Published Subjects is not only making easy to merge between Topic Maps, but also making to enable interoperability between OWL/RDF and Topic Maps. Subject indicators are the information resources which describe subjects, and the subjects can be identified by using unique URI or IRI. Subject indicator which published is called Published Subject Indicator (PSI). An example of PSI is shown below. An information resource which describes dolphin, sea animal related to whale is a Subject Indicator of "dolphin". It indicates dolphin real world thing in computer. If it is located in "http:www.knowledge-synergy.com/PSI/dolphin", we can use the address to indicate dolphin. And we can make clear what subject we are talking about.

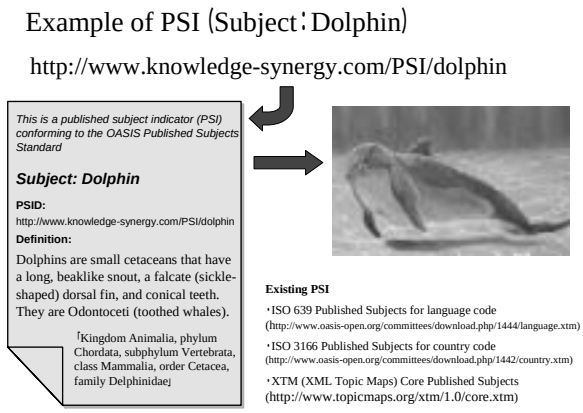

Figure 2. Example of PSI (Published Subjects Indicator)

4. Ontology

Although there are various definitions what ontology is, I define it as "definition and organization of concepts and relationship between concepts". There are also several classification about ontology itself.

(1) Example of classification 1

(a) Upper ontology

- limited to general, abstract concepts

- usually this means entity types, relationship types, and property types

- sometimes also especially prominent instances

- some attempts have been made at creating standardized upper ontologies

(b) Lower ontology

- basically a full ontology with instances minus the upper ontology

(2) Example 2 of a classification

(a) Upper ontology

(b) Domain ontology

- It is targeted certain domain, such as a plant, medical care, law etc.

(c) Task ontology

- It is targeted certain process to solve problem, such as diagnosis, a design, and study support.

Recently various ontologies were created or have been creating. The examples of the ontologies which seem to be able to use with application are shown below.

(1) UNSPSC (Universal Standard Products and Services Classification

- The product classification system of an electronic catalog

(2) SWEET (Semantic Web for Earth and Environmental Terminology)

- Ontology about earth environment

(3) ISO 12207 SLCP (Software Life Cycle Process)

- The definition of the software life cycle process

(4) ISO19115 Geographic information -- Metadata

- geographic information

These ontologies becomes possible to use as Published Subjects, if a unique identifier (URI or IRI) is given for each items in the ontologies and are published on the network permanently.

5. Remote Access Protocol

More and more Topic Maps and RDF are publishing on the network. It is natural that people want to exchange, update, merge and filter fragments of those. If we could realize those operations, those information resources would become more useful. It is necessary to specify the protocol for remote access of these resources.

An example of remote access of Topic Maps is shown below. Figure 3 shows "Lars Marius Garshol" topic and its related information in certain FOAF RDF file which is mapped to Topic Maps in server "poivre". Figure 4 shows "Lars Marius Garshol" topic and its related information in certain XTM (XML Topic Maps) file in server "pepper". Figure 5 shows "Lars Marius Garshol" topic and its related information which are result of merge of 1st RDF and 2nd XTM.

Figure 3. Topic Map on server “proive” (before merge)

Figure 4. Topic Map on server “pepper”

Figure 5. Topic Map on server "proive" (after merge)

Using remote access protocol, the Topic Map in serve "pepper" was accessed and merged into the topic Map in the server "poivre". In results, related information of "Lars Marius Garshol" topic in server "poivre" became much richer than before merge.

6. Query Language

As there is SQL in a relational database, standardization of a query language to the graph structure data model like RDF and a Topic Maps is advancing. SPARQL (SPARQL Protocol And RDF Query Languqge) is being created by W3C and TMQL (Topic Maps Query Language) is being creaed by ISO. Although it is a basic need to query and update Topic Maps and RDF, it is thought that a little more time is needed to complete the work. In the meantime, we can use a query language called tolog produced by Ontopia. The tolog has already implemented in tow or more tools.

7. Application Framework

By using the technical elements described above, it becomes possible to connect information and knowledge on the network seamlessly and organize and navigate them based on the subjects. Those technical elements have potential to realize collocation and one stop shopping according to the subjects. The image of identifying subjects and exchanging and merging Topic Maps and RDF based on the subjects is shown below.

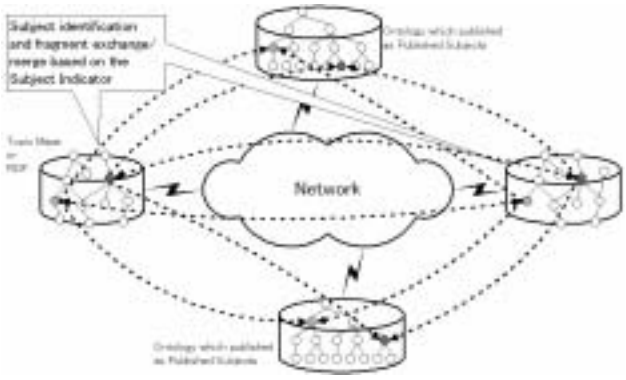

Figure 6. Subject identification and fragment exchange/merge based on the Subject Indicators

Moreover, in order to realize the above-mentioned knowledge share structure generally, an application framework is considered. There would be many applications which have same structure and functions. So the framework would be applicable to many applications. The image of the application framework is shown below.

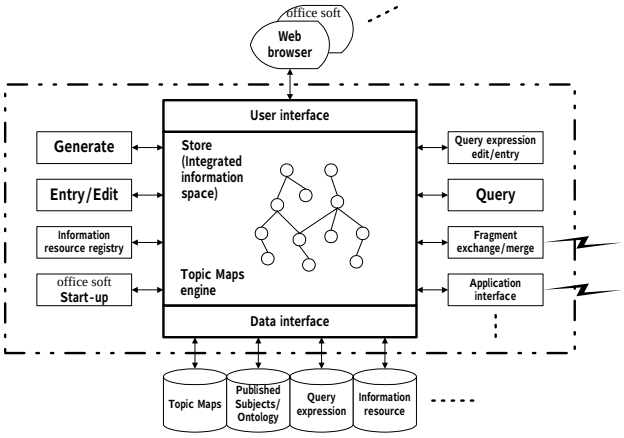

Figure 7. Application Framework

7.1 Example of Application 1. FOAF and personal data

Using FOAF (Friend Of A Friend) RDF data and Topic Maps for each person, the Social Networking which have rich content of person and one stop shopping system of the person can be made.

The following figures show graphical representation of RDF and Topic Maps. The left side figure shows the graphical representation of FOAF RDF data. The middle one shows the graphical representation of Topic Maps data of one person. The right side figure shows the result of merge of RDF data and Topic Maps data. The common nodes between RDF data and Topic Map data are merged (integrated) according to the Subject Indicator.

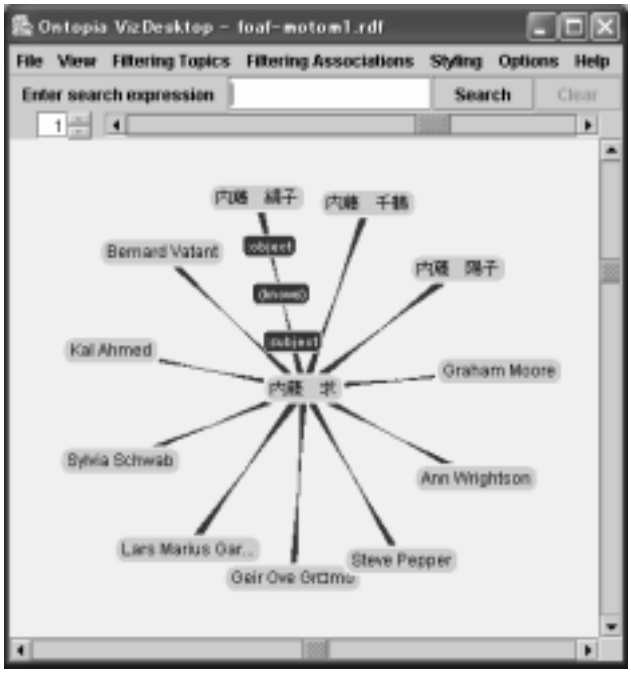

Figure 8. Graphical presentation of FOAF RDF data

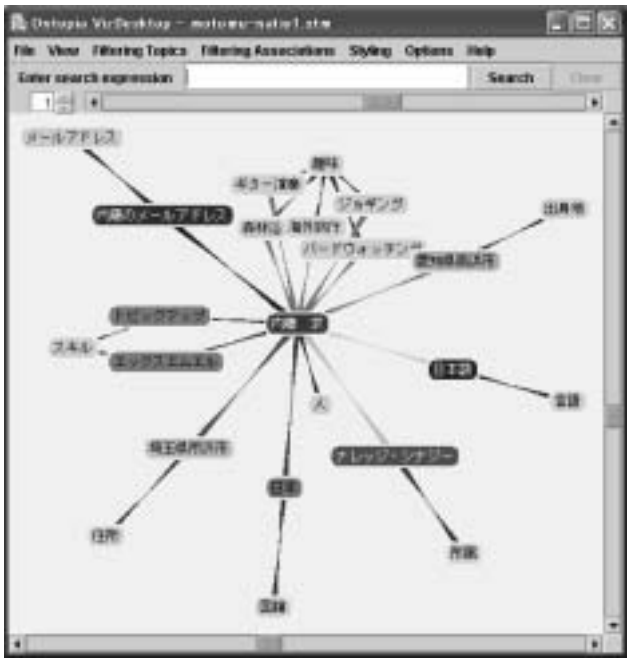

Figure 9. Graphical presentation of personal XTM data

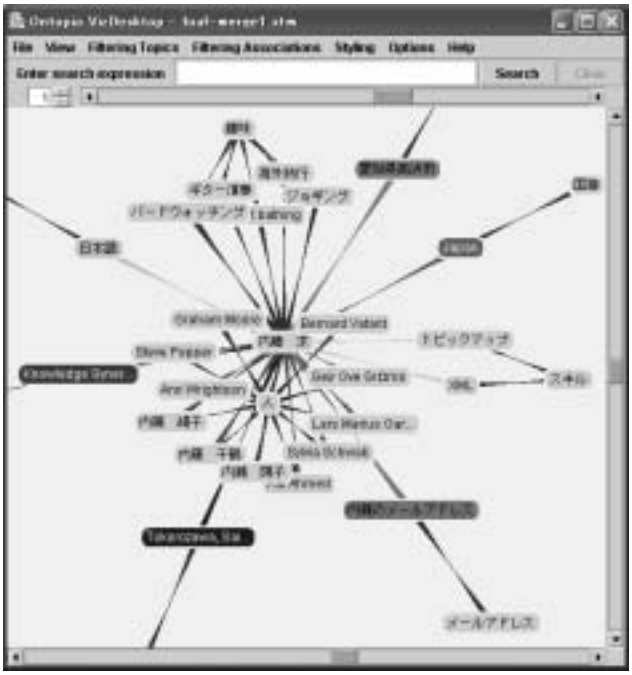

Figure 10. Graphical presentation of the result of merge

7.2 Example of Application 2. SLCP and document data

Using and merging Topic Maps for SLCP (Software Life Cycle Process), project, person etc and Dublin Core metadata (RDF data) of a document, navigation according to the various points of view are made possible. And they also make possible project management, content management etc from many viewpoints. The result of the merge of Topic Maps data and RDF data is shown below.

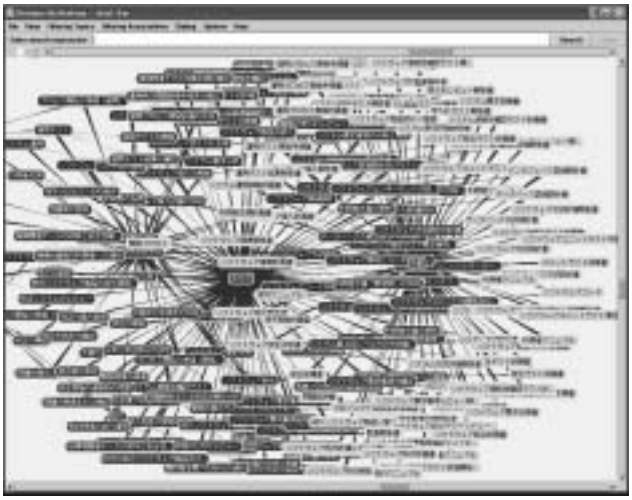

Figure 10. Graphical presentation of the result of SLCP related XTM data

8. Conclusion

I enumerate the technical elements explained above.

(1) RDF, topic map

- Information resource structured semantically

(2) Published Subjects

- A mechanism identify subjects globally

(3) Ontology

- Definition and organization of concepts and relationship between concepts

(4) Remote Access Protocol

- Protocol for exchange, merge and filter of fragment of Topic Maps and RDF on network

(5) Query language

- Query and updating of RDF and Topic Maps

By combining these technical elements organically, it is possible to organize the information/knowledge based on not the information resources itself but the subjects included in the resources. Thereby, the findability of information/knowledge is significantly improved. To apply those technologies to various problem domains, we could enjoy the merit of them with users.

9. ACKNOWLEDGMENTS

I would like to thank Steve Pepper, Sylvia Schwab and Lars Marius Garshol for their leads, suggestions, and constant support.

10. REFERENCES

[1] Lars Marius Garshol, Living with topic maps and RDF: Topic Maps, RDF,DAML,OIL, OWL,TMCL, http://www.ontopia.net/topicmaps/materials/tmrdf.html.

[2] Lars Marius Garshol, The RTM RDF to topic maps mapping : Definition and introduction, http://www.ontopia.net/topicmaps/materials/rdf2tm.html.

[3] Steve Pepper and Sylvia Schwab, Curing the Web's Identity Crisis : Subject Indicators for RDF, http://www.ontopia.net/topicmaps/materials/identitycrisis.html.

[4] Steve Pepper, Seamless Knowledge with TMRAP, http://www.ontopia.net/topicmaps/materials/Seamless Knowledge with TMRAP.ppt.

[5] Lars Marius Garshol, Metadata? Thesauri? Taxonomies? Topic Maps: Making sense of it all!, http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html.

[6] Riichiro Mizoguchi, オントロジー工学, Ohmsha, ISBN4-274-20017-5.