# Managing Knowledge using a Semantic-Network

**Yoshitaka Kuwata and Masashi Yatsu**

Laboratory for Information Technology
NTT DATA CORPORATION
66-2 Horikawa-cho, Saiwai-ku, Kawasaki-shi, Kanagawa, 210 JAPAN
{kuwata, yatsu}@lit.rd.nttdata.co.jp

**Abstract**

It has become very important for advanced organizations to make the best use of information gathered from database in companies and from the Internet. There are three stages in the information life cycle; Finding, Organizing and Sharing. Many technologies have been developed for the finding stage. On the other hand, no concrete organizing and sharing technologies exist to manage the information found.

In this paper, we focus on the management of information sharing among group members. We propose to use a semantic-network for organizing information that has been gathered. We show an example of how to manage URL information in order to share it among small groups. As the system is under development, only a partial evaluation of the method is given.

## Introduction

Many data mining and search technologies have been developed to find information from various kinds of databases, including open databases like the Internet. However, no concrete management technology exists to manage the information found and make the best use of it in groups in an organization.

It is becoming a critical issue for advanced organizations to make the best use of information gathered from their database and from the Internet. Information is used to predict future trends and to make very important decision. To use information, itt must be organized first. Well organized information is regarded as 'company knowledge'.

For example, if we want to survey *'agent technology'* on the Internet, we need to retrieve information from the Internet. Search services and index services can be used for this purpose. To make use of the information gathered, we need to organize it as 'knowledge'. In the case of *'agent technology'*, we need to classify the technology and check the state of each class to write survey reports. The readers of the reports understand only when the information is organized.

Figure 1.1 illustrates our model of the information life cycle. There are three stages in the cycle; Finding, Organizing, and Sharing stages. In the first stage, information is found by humans. Next, it is organized and stored as knowledge to be referred. Then the knowledge is shared by members until it becomes obsolete.

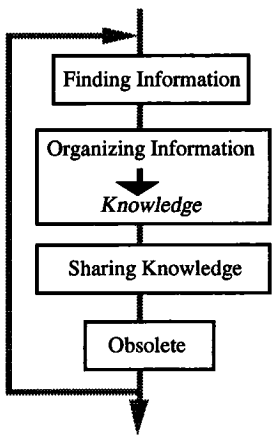


**Figure 1.1 Information Life Cycle**

Organizing information at the 2nd stage turns it into knowledge. This is essential to make the best use of the information.

For the finding stage, many commercial search and index services, such as 'Yahoo'[1] and 'Lycos'[2] are available on the Internet. The search services use natural language processing technology to find information. On the other hand, index services are maintained by humans by hand. We can find information by using these services.

In order to organize information gathered from the Internet, a mechanism called Bookmark or Hotlist is used in commercial web browsers. Bookmark and Hotlist allow us to organize information in a tree hierarchy. However, they are designed for personal use. In order to organize and share the information among groups, we need to reorganize the information. As members have many view points or perspectives and hierarchy can

[1]Yahoo: http://www.yahoo.com/
[2]Lycos:http://www.lycos.com/

represent only one aspect, hierarchy is not necessarily the best representation of knowledge.

In the following section, we will introduce a semantic-network for the management of information. We will also discuss what we need in order to share information among groups. In section three, URL management is discussed as an application of a semantic-network. In section four, a sample implementation of a URL management system named 'Knowledge Organizer' is shown. Current status and future work is described in the last section.

## Using Semantic-Network as knowledge representation for sharing

Semantic-networks are widely used to represent structured knowledge. An example of recent work is found in [Fujiwara et. al. '96], in which semantic-network is used to describe concepts.

### Definition

Semantic-networks are consist of nodes and arcs; A node $N_1$ represents one concept and $R_{1 \to 2}(L)$ represents an arc from node $N_1$ to $N_2$ with label $L$, which indicates a relationship between the concepts. A set of nodes, labels, and arcs represents a set of knowledge in a semantic-network. A knowledge $K$ is given by the following equation;

$$K = \{\{N_1 \cdots N_k\}, \{L_1 \cdots L_l\}, \{\bigcup R_{i \to j}(\otimes)\}\}, \qquad (2.1)$$

where $k$ is the number of concepts, $l$ is a number of labels and $\otimes$ represents any arc between node $i$ and $j$.

Note that it is possible to have more than two arcs with different labels between a pair of nodes.

Figure 2 is a small example of knowledge in a semantic-network.

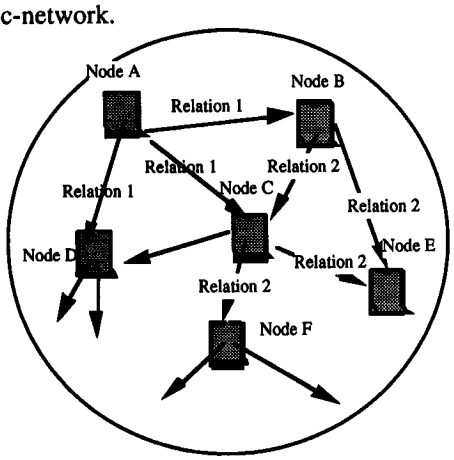


**Figure 2.1 An Example of Knowledge in Semantic-Network**

A tuple with a set of nodes, a set of labels, and a set of arcs makes knowledge in a semantic-network.

### A View Point

We introduce a *view point* in the semantic-network in order to express a current interest of users. A view point is a set of a node, arcs and a number. A view point $V_1$ is defined as follows;

$$V_1 = \{N_c, \{\bigcup^{\forall} R\}, STEP\}, \qquad (2.2)$$

where $N_c$ is a current node we are interested in, $\{\bigcup^{\forall} R\}$ is a set of labels we are interested in, and $STEP$ is a depth to see.

The view point works as a filter in a semantic-network. By disabling labels other than the view point, we can define thehorizon of our view and filter out other information from our view.

### Merging two Semantic-networks

It is possible to share sets of knowledge represented in semantic-networks by merging them into a sets of knowledge. Given two knowledge sets, $K_1$ and $K_2$, we can merge them by adding extra labels and arcs. This is represented by equation 2.5.

$$K_1 = \{\{N_1 \cdots N_{k1}\}, \{L_1 \cdots L_{l1}\}, \{\bigcup R_{i \to j}(\otimes)\}\} \qquad (2.3)$$

$$K_2 = \{\{N_1 \cdots N_{k2}\}, \{L_1 \cdots L_{l2}\}, \{\bigcup R_{i \to j}(\otimes)\}\} \qquad (2.4)$$

$$\text{and } L_{new} = \{L_1 \cdots L_l\}, R_{new} = \{\bigcup R_{i \to j}(\otimes)\}\} \qquad (2.5)$$

produces the following new knowledge set $K_{1 \cup 2}$

$$K_{1 \cup 2} = \{\{N_1 \cdots N_{k1}\} \bigcup \{N_1 \cdots N_{k2}\}\}, \{\{L_1 \cdots L_{l1}\} \bigcup \{L_1 \cdots L_{l2}\} \bigcup \{L_1 \cdots L_{new}\}\}, \{\bigcup R_{i \to j}(\otimes)\}\} \qquad (2.6)$$

Figure 2.2 shows an image of the two merging knowledge sets.

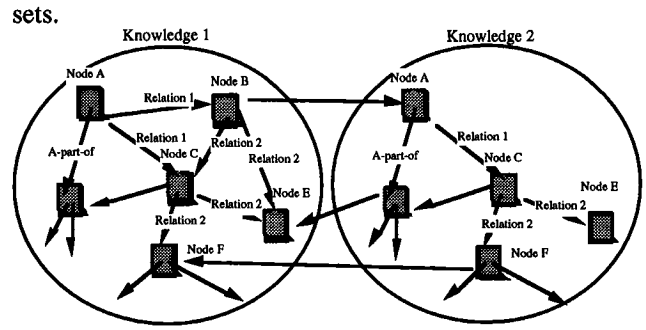


**Figure 2.2 Merging Two Knowledge Sets**
Knowledge 1 and Knowledge 2 can be merged by using cross-arcs with labels.

In other words, in order to refer to knowledge gathered by other members, arcs with special labels can be used to link to other members' nodes. We can distinguish these arcs from usual arcs, by adding a special label to the arcs, although it is not necessary to do so.

We can also extend the concept of our view points, which is represented by equation 2.2 by adding a new label and arc set into the original view point.

There is a trade off between flexibility and maintainability of the knowledge sets. Although we can add arcs and labels that links to other members' knowledge independent from other members' knowledge, the new labels can make knowledge inconsistent. For example, if two knowledge sets are built from different perspective, or ontology, it is necessary to re-organize them for merging them in order to keep the whole knowledge set consistent. Thus, when we are planning to integrate more than one knowledge sets, but need to organize information by each members, it is a good idea to have a set of standard labels and to share it in the group. Having a standard label set makes it easy to keep knowledge set consistent.

### Hyper text vs. Semantic-network vs. hierarchy representation

A semantic-network is more flexible than hierarchy representation, as it is a directed graph rather than a tree. Hypertext is considered to be a kind of semantic-network without explicit labels. By adding explicit labels to links in hyper text, the relationships become clearer.

## Sharing URL information in Special Interesting Groups (SIG)

We use an information management method using a semantic-network for URL management. URL management problems have the following characteristics.

(1) URL is considered as a hyper text, which has links to other URL. However, there are no explicit labels or meanings in these links.

(2) Information on the Internet is updated frequently.

(3) There is no mechanism to notify a change in these links. Thus users can know the change only when they visit.

There are many commercial index services and search services on the Internet. The index service is an effort to make a general purpose yellow pages service in which the index is maintained manually. The search service is based on a natural language processing technology, especially full-text search. Programs called 'robots' are used to gather information from the Internet. The information is indexed and scored for the search services.

Both index and search services are intended for a general purpose service. If we have a special interest in 'agent technologies', for example, we need to maintain the information by ourselves, although we can use these services to find new information. Although search services can be used to find URL which include word 'agent', the information found is not always related in our interest. We will find many unrelated information such as 'secret agent 007'. On the other hands, index services can have a category which is matched to our interest by chance. However, we cannot expect index service to provide very detailed category and up to date information, as they are for general interest, not for special interest, and maintained by humans.

Thus it is very important to gather and organize information from a users point of view in group information management, not a general point of view.

We are focusing on URL information management among small groups, because it is a good example of information management that need information sharing. Figure 3.1 shows information sharing in small groups.

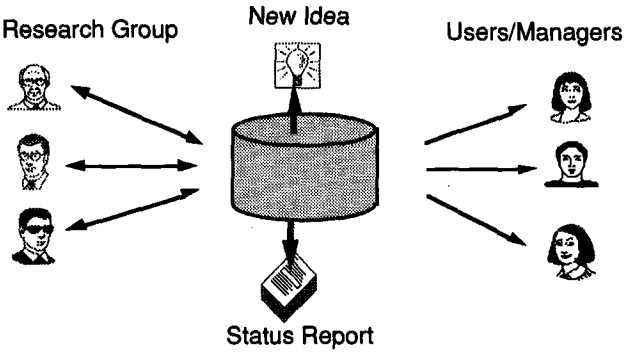


**Figure 3.1 Information Sharing among SIG**
Information is organized as knowledge and shared among special interest groups.

## Implementation of Knowledge Organizer

We implemented a group information management system named 'Knowledge Organizer' in order to evaluate the knowledge management and sharing method. The system is based on a semantic-network to represent URL information. The knowledge organizer consists of three components; a) a URL analyzer b) a URL management Server, and c) a semantic-network browser. Figure 4.1 shows the structure of the knowledge organizer.

When we find interesting information on the Internet or on Intranets, we can store the information like URL, the title of the page, comments, date and so on, on the URL management server so that the information can be shared with other members. If we want to describe the relationship with other information in the semantic-network, we can add a directed link with a meaning label.

Users can look at the information and their relationships via a semantic-network browser. Users have their own knowledge set, which can have links to other members knowledge sets. If users find related information in other members knowledge, they can make a link to the node. This mechanism allows users to share

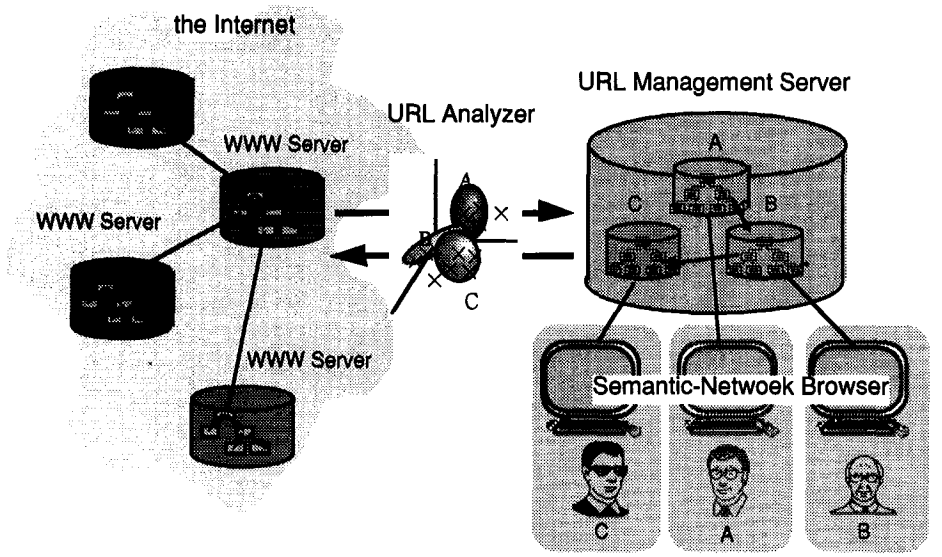


**Figure 4.1 Knowledge Organizer**

There are three main components in a Knowledge Organizer; a URL Management Server, a URL Browser and a URL Analyzer. All components are implemented as cgi scripts in a WWW server by Perl script language.

knowledge. The whole semantic-network in the URL management server is regarded as a group knowledge.

Because the URL management server is implemented as typical web server, users can browse semantic-networks by using typical WWW browsers, such like Netscape Navigator[3].

The URL analyzer is used to analyze relationships between categories in semantic-network and/or www index pages. It helps to visualize the difference between two categories of knowledge by analyzing the common URL information included. The details of the category analyzing and merging method used in the URL analyzer is described in [Nonaka and Kuwata '96].

Figure 4.2 is a snapshot of the browser screen of a knowledge organizer.

We use the concept of 'view point' to represent a semantic-network. The buttons at the top of the screen are used for controlling. The top half of the screen shows information of the current node; information about the URL of NTT DATA CORP. in this case. The title of the node, user id, updated date, entry date and comments are shown in this field.

The bottom half of the screen show the relations to other nodes within the steps specified as a view. We use a hierarchy representation for the semantic-network, because it is easy to understand; the indents of the list are used to represent steps to the nodes. For example, *the First Financial Systems Division* is *a-part-of Financial Systems Sector,* which is also *a-part-of* NTT DATA CORP. Thus the First Financial Systems Division is two steps from our current node. Two kinds of links, *part-of* and *affiliate*, are shown within two steps of the current

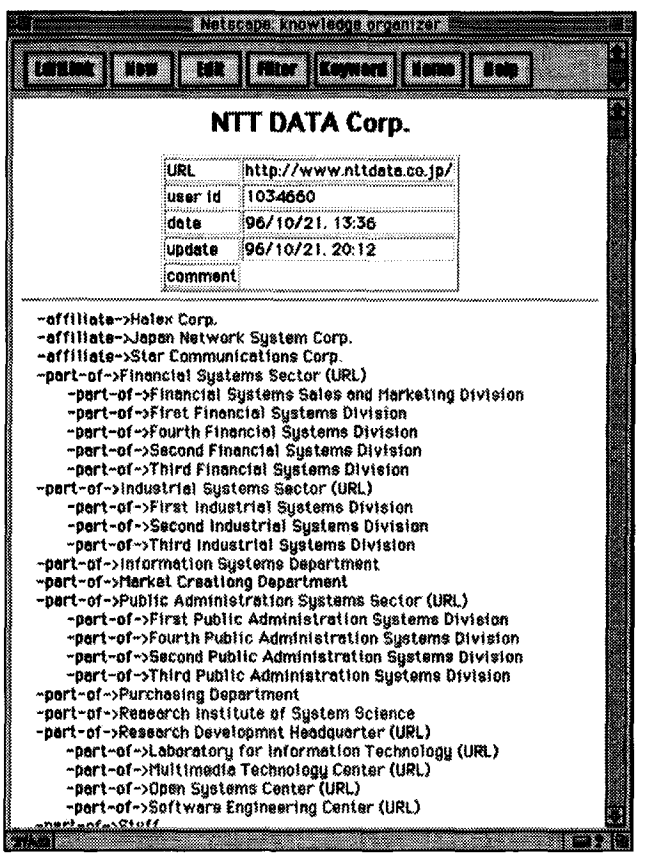


**Figure 4.2 Screen Image of Knowledge Organizer**

Snapshot of 'View Mode' is shown. By using Buttons on the top of the screen, we can control browser.

[3] Netscape Communications Corp.: http://home.netscape.com/

node. By clicking nodes shown in this list, we can move our current node to these nodes. We can also jump to the URL page directory by clicking the URL field of the nodes.

A filter button at the top of the screen is used to change the current view point; we can change steps and arc sets.

By changing our current node and our view, we can explore the semantic-network using this viewer. The edit button is used to transfer to the edit mode, which allows us to edit nodes and links.

## Current Status and Future Work

The knowledge organizer is currently at the first evaluation stage . A small group has been using the system on a daily bases for about five months. The following are findings from their experience.

(1) The Knowledge Organizer is useful as a personal tool. It allows users to organize their information and show it to other members.
(2) However, it is not easy for members to keep knowledge up to date, especially information gathered from the Internet because it can change very rapidly. Users need to update knowledge frequently. A mechanism to monitor URL in the knowledge is necessity.
(3) As we didn't provide any standard label sets for arcs, the members create their own. This makes the Knowledge Organizer very flexible but difficult to integrate knowledge sets and keep consistency between knowledge sets gathered. We are planning to build standard label sets for organize.

We are also planning to do off-line experiments in order to find out best knowledge representation for sharing; We are going to compare semantic-network and other knowledge representations in these experiments.

## References


Usama Fayyad et. al, From Data Mining to Knowledge Discovery in Databases, *AI magazine*, Vol. 17, No.3, 1996:37-54

Yuzuru Fujiwara, Ye Kiu and Jingjuan Lai, The Conceptual Structure Model for Description of Semantic Meaning, Information Processing Society of Japan SIG Notes Vol. 96 No. 88, 1996

Yoshitaka Kuwata, et. al, Organizing URL by using semantic-network, part 1, In Proceedings of the 53th Information Processing Society of Japan, 3-169, 1996

Satoru Nonaka and Yoshitaka Kuwata, A Category-Merging Method of the Information on the World-Wide Web, In Proceeding of the 1996 IEICE General Conference, D-143

Masashi Yatsu and Yoshitaka Kuwata, Organizing URL by using semantic-network, part 2, In Proceedings of the 53th Information Processing Society of Japan, 3-171, 1996