Topic Maps & The Semantic Web

by James E. Harvey

Abstract

The Worldwide Web, based upon HTML, lacks sufficient metadata to portray information in relationship to its context. The amount of waste in terms of labor and time, and the weakness of search engines and HTML META tags are reviewed, as are the need to capture the taxonomies of human language and diction, collectively know as an "ontology," in a manner that can be consumed and processed by computers. This concept is popularly known as the *Semantic Web*, as envisioned by Tim Berners-Lee, the founder of the Worldwide Web Consortium. The Semantic Web is briefly introduced and then the concept of *Topic Maps* is introduced. Topic Maps provide a Web-friendly method of marking up ontologies and the XML version of Topic Maps (XTM) is discussed in detail and an example is provided. The paper concludes with an example of XTM and a discussion of its strengths and weakness.

Introduction

The Issue

Nobody would argue that the Internet has been of great benefit to individuals, corporations, and institutions, but the Internet as it is comes with a price: daily information overload containing a painful amount of irrelevant and therefore useless information.

The Web is broken and the problem continues to get worse with the passing of time. The statistics on information overload are becoming alarming:[1]

- The average Internet user will lose their place and begin an entirely different and totally unrelated search 9 times per day.
- The average Internet user will misfile 120 pages of information and miss-log 33 folders every month … more than one per day. As a result of hotlinking or "cyberjourneys," Internet users have become less analytical and display a greater frustration with compartmentalizing.
- On line "cybersoaring" with no apparent direction and no constructive research generates more than 400 hours of waste per employee/ per year.

According to the *2001 UCLA Internet Report*, the top reason why users started to use the Internet is to obtain information quickly, followed by work needs, and then access to e-mail, yet only 36.3 percent of users say that _*half*_ of online information is reliable and accurate. In 2001, 60 percent of all users consider the Internet to be a very important or extremely important source of information, up from 53.6 percent in 2000. Add those who say moderately important, and the total increases to 90.8 percent for 2001, up from 77.2 percent in 2000.[2] We are becoming dependent on the Internet and the Worldwide Web as our primary information resource yet is widely believed to be a highly fallible information source.

[1] *The is an Internet … E-mergency [Shocking Stats]*, by Michael Fortino, http://www.e-mergency.com/stats.html, The Fortino Group 2001.
[2] The UCLA Internet Report 2001 – "Surveying the Digital Future," http://ccp.ucla.edu/pdf/UCLA-Internet-Report-2001.pdf, UC Regents, November, 2001.

The Problem

The problem is that the underlying structure of the Web is insufficient to deal with the complexity of the global knowledge base and incapable of providing enough organization to deal wit the complexity of today's Worldwide Web, let alone the potential of what the Web could be.

Although HTML provides for META tags that can be used to provide a basis for search engine sorting and organization of results, the use of META tags is often promotional in nature and often misleading. META tags occur in the header section of an HTML page and may provide keyword and a description of the website for use by search engines and directories, as well as commands for use by Web spiders and search engines.

Example:

```
<HTML>
<HEAD>
<META name="description" content="This web page provides information on topic
maps and the semantic web, featuring and articles by James E. Harvey.">
<META name="Keywords" content= "web, web programming, topic maps, semantic
web, ontology, web ontologies, taxonomies, knowledge management, blind pink
polka dot ponies."
<META http-equiv="expires" content="Tue, 27 August 2002 12:00:00 GMT">
<META http-equiv="window-target" content="_top">
<META name="robots" content="All">
<META name="robots" content="Index">
</HEAD>
</BODY>
The article goes here.
</BODY>
</HTML>
```

In the above example META tags are used to tell search engines and Web spiders to index this web page and all sub-pages. It also communicates the expiration date of the information and instructs browsers to open the page in the top browser window. A description of the web page is provided along with key words for use by search engines and directories; however, the keywords do not have to make any sense at all. Our example includes "blind pink polka dot ponies" and that is perfectly legitimate HTML. The META tags of one manufacturer of desktop computers may include among its keywords the names of its competitors, or far worse yet, a porn site's META tag keywords may include "Teletubbies" or "Barney" — there is no "standard" context in which META tags originate or reside!

The eXtensible Markup Language (XML) offers the promise of more intelligent tagging, but XML alone lacks the consistency of application and design to help reduce information overload. Furthermore, even the best organized Web search engines and Internet directories are distorted by the realities of profit and loss. Although companies such as Overture and Inktomi provide technologies for improving search engine performance, web administrators must *actively market their Web pages to* major search engines and directories. Advice can be found for optimizing META tags for Web search engine capture and how to register with directories and search engines, but the Web *is not free*. Many search engines and web directories now offer *paid inclusion* or ask for a fee to be listed at the top of relevant search returns, and Yahoo now charges $299 to

$600 annually to be listed in its commercial directories.[3] The result is that Internet searches provide too many results for all searches and there is no context in which users can accurately determine the fit or relevance of their search results to their information needs.

Billions and Billions …

How extensive is the web organization problem? Even on the best of web search engines the number of results that are produced are typically too numerous to be useful, and can be erroneous to the subject of the search. Here's an example of an increasingly specialized web search conducted on Google (regarded by many as one of the best search engines on the Web):

Search	Results
HTML	352,000,000
HTML+4.0	2,730,000
HTML+4.0+Meta tag	59,100
HTML+4.0+Meta tag Optimization	3,890
HTML+4.0+Meta tag Optimization Usage	1,810
HTML+4.0+Meta tag Optimization Usage Dreamweaver	107
HTML+4.0+Meta tag Optimization Usage Dreamweaver+3.0	38
"Using HTML 4.0 meta tags with Dreamweaver 3.0"	0

Even knowing specifically what you want is not helpful. In the above example, the "Using HTML 4.0 meta tags with Dreamweaver 3.0" search produced no results and *missed* 38 helpful sites! The problem is matching an idea to a phrase is very difficult. Furthermore, if you know the proper name of something specific, you are just as likely to get information overload. For instance, I you wished to find information on the "Digital Millennium Copyright Act" you'll get 150,000 returns on Google … you'll even get 11,300 with a misspelling in the proper title! ("Digital Millenium Copyright Act") For the non-programmer population of Internet users, the effect can be daunting, rendering the Web nearly useless for ordinary usage (or more likely providing yet another source time consuming of frustration.) Here are some Google search results for more ordinary searches:

Search	Results
How to write a simple will	3,100,000
US Federal Banking Laws	381,000
Maryland Accountants	44,000
Comparison of Refrigerators	30,600
Comparison Notebooks Pentium+III	12,700
"Fine art paper"	4,960
"Treating asthma"	4,940
"Job Listings" + Nurse + Maryland	2,700
"Fixing leaky faucets"	746

[3] *Yahoo Now Charging Annual Listing Fee*, by Danny Sullivan, http://www.searchenginewatch.com/sereport/02/01-yahoo.html , The Search Engine Report, January 7, 2002.

Furthermore, there are huge differences in the performance of search engines. Below are the results of conducting the same search, (for “Comparison Notebooks Pentium+III”). Fewer results do not necessarily mean better results. For instance, our top performer (MonsterCrawler.com) included among its top ten returns for “Semantic Web,” a site titled, ‘RhymeZone rhyming dictionary and thesaurus.”

Search for Comparison Notebooks Pentium+III	Results
Alltheweb	19,917
Lycos	19,503
Google	12,700
MSN Search	4,582
AOL	4,425
AltaVista	2,784
MonsterCrawler	32
Looksmart	32 plus sponsored sites
Askjeeves & Yahoo	*Do not provide a count … and there appeared to be no end to clicking on “next”!*

Beyond Web Surfing

The Web’s lack of structure and organization is more than just a nuisance to the casual Web surfer, it posses other problems as well, including:

- Information on the Web isn’t functional — you cannot move information from a website to your PDA, or calendar, or Rolodex, or spreadsheet … an so on, without laborious cutting and pasting,
- The web is not conducive to enterprise application integration — Web pages and HTML are an end point and cannot be used to incorporate information with other enterprise applications … beyond a simple link, there is not integrating with web sources, only with underlying database, (*which leads us to…),* and
- The Web is not conducive to database applications — Although XML provides are great intermediary between databases and Web servers, that is the extent of it. Once pages are served up in HTML, the underlying associations and rules of the parent database are lost. This poses obstacles for e-commerce between organizations and in some cases, between customer and supplier.

The fundamental problem is that the Worldwide Web (and HTML) was designed for human consumption and to support free and unrestrictive associations — it was never meant to be *processed* or used by computers. Without advancing the underlying structure of the Worldwide Web, the Web will become and increasingly incredible and failed, yet enormously important source of information. The rather human organization of information on the Worldwide Web is very difficult to relate to the databases and information stores computers use to process information.

The Semantic Web

Cup holders and Dixie Cups

"Civil War" … "Love" … "Cup holders" Words have meaning to humans, and these meanings have associations and attributes, many of which are far too subtle for conventional computer logic. To U.S. citizens "The Civil War" and "civil war" bring to mind two very different sets of images, thoughts, and meanings. Even a single word such as "love" can invoke a myriad of ideas, memories, images, and thoughts. Psychologist would say that we anchor our self-perception to the world around us by the attributes we affix to the images, ideas, people, and thoughts that float around in our minds. "Float" is inappropriate, more specifically; we tend to organize things into boxes and cabinets (or in classic rhetoric, "rooms"), but one box may be in several cabinets, and each cabinet may open to several rooms — the human mind is not bound by physical laws, just the brain is.

"Cup holders" and "Dixie Cups" are containers, and one may hold the other. O.K., Computers can handle that. But you may also associate "Cup holders" and "Dixie Cups" with "drinks," "liquid," "cars," "picnics," "high chairs," "the Confederate South," "paper," "plastic," "inexpensive," "wax," "at the grocery store," or even a particular date you had with "Linda" in your first year of college … and she was in your calculus I class. *Put that in your computer and compute!*

Thankfully, nobody (I hope) is trying to file away sketches of your life into a public library; however, beyond the simplest levels, our collective knowledge of all things is organized in our minds and societies in the same comfortable and computer-unfriendly manner.

"Object-Oriented" (OO) for a programmer may bring to mind programming languages such C# or Java, or modeling and analysis methods that use the Universal Modeling Language (UML). The programmer may think of books she's read on the subject or classes on the subject that she attended. If combined with "leader," or "pioneer," or "expert," she may think of Grady Booch, Ivar Jacobson, or James Rumbaugh. Combined with "experience" or "history," she may recall projects she's worked on, certain customers, lessons learned, successes, and failures. The greater here experience and expertise in the OO subject area, the more "boxes," relationships, and attributes she is likely to have. The same is true of any area of expertise or industry: medicine, surgery, law, automotive repair, politics, electrical engineering, ethnomethologicalsociophyschology, and so on. And the further you drill down into a subject area, the further you get into topic-specific diction, definitions, and away from your desktop dictionary.

If you were to collect all the words that have meaning within an subject area (diction) and organize them in a way that parallels the thinking of that subject's experts and institutions (taxonomy) you would have a "ontology" of the subject area.

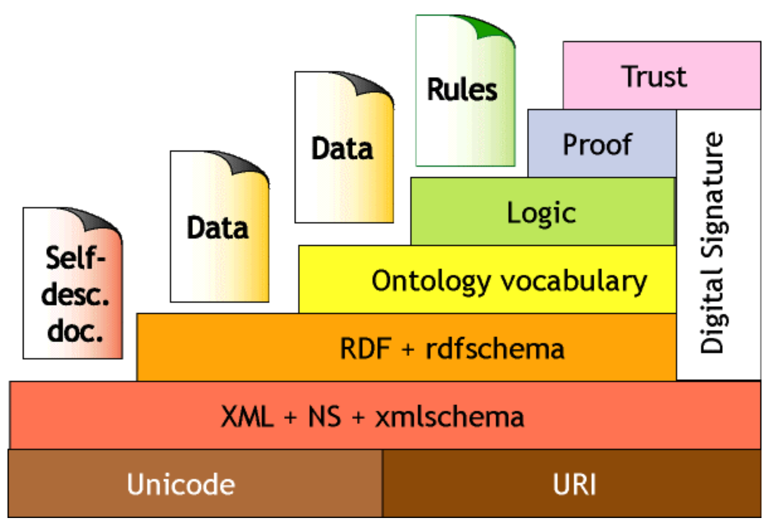
Stuffing Ontologies into Relational Databases

If the Web is going to provide long-term benefit as a functional global information resource for humans, we have to figure out how to organize and provide ontologies of human knowledge to computers in a form that is both flexible and adaptive (e.g., human knowledge, and hence ontologies, is not static), and useful for computing. This is the vision that Tim Berners-Lee provided when he introduced the concept of the "Semantic Web" in late 1998.

The concept of the Semantic Web is a net generation worldwide web in which information is provided in context, and that context can be processed by computers, which implies its structure can be can be organized in such a way that it can be moved rationally in and out of data stores and databases. For instance, your database of "friends" and "co-workers" may have some overlap, and communications about them depend on the context in that you may have different email addresses and phone numbers for individuals and what is used depend on the context (e.g., "friend or co-worker). Likewise, a medical web site may provide different types of information, and respond to searches differently, depending on whether the current visitor is a doctor or a patient. According to Berners-Lee, the Semantic Web would have several layers:

1. The Semantic Web must be "smarter" than HTML, and use XML and XML schema, and Uniform Resource Identifiers (URIs) and XML Namespaces to separate XML tags in one document but from different schema, as well as relating those tags to there source schema,[4]
2. The Resource Description Framework (RDF) is used to represent metadata in such a way as to no loose its meaning when exchanged between systems,
3. The ontologies themselves, and
4. A method of communicating the logic of the ontology's structure, and a method of validating the interpretation of ontologies.

This "stack" is depicted in the graphic to the right.[5] According to Tim-Bernners Lee, "We are not inventing relational models for data, or query systems or rule-based systems. We are just *webizing* them. We are just allowing them to work together in a decentralized system — without a human having to custom handcraft every connection."[6]



The Semantic Web Stack

RDF is a key component to the Semantic Web that provides, "a system of machine-processable identifiers that allows us to identify a subject, object, or predicate in a statement without any possibility of confusion with a similar-looking identifier that might be used by someone else on the Web … [in a] … machine-processable format for representing these statements and exchanging them between machines."[7]

---

[4] A useful practice is to create a Web page to describe the markup language (and the intended meaning of the tags) and use the URL of that Web page as the URI for its namespace.

[5] *Semantic Web - XML2000*, by Tim Berners-Lee, http://www.w3.org/2000/Talks/1206-xml2k-tbl/, Graphic Communications Association, XML 2000 Proceedings, December, 2000.

[6] *Business Model for the Semantic Web: Enterprise Application Integration and other stories*, by Tim Berners-Lee, http://www.w3.org/DesignIssues/Business.html, first written October 25, 2001, last change: October 29, 2001/10/2001.

RDF relates well to relational databases in that has RDF “nodes” that are akin to database records, RDF “propertyType” that is akin to database table columns, and “values” that are akin to cell values in a database. RDF communicates this type of information by grouping information in triplicates that provide a “subject,” “predicate” and “object.” For instance, in the following example …

```
<http://www.colorture.com/index.html>
<http://www.colorture.com/terms/address>
"1501 Grant Avenue, Bedford, Massachusetts 01730"
```

The page index.html contains the address "1501 Grant Avenue, Bedford, Massachusetts 01730" and the meaning of an “address” is defined in the www.colorture.com/terms/address directory.

The Semantic Web standards are still under development and there is much to do. RDF has been around for a while and is designed to work well with databases, but it is still lacking in its ability to capture the complexity of human ontologies. Berners-Lee recognized this himself and suggested that RDF and another XML technology, *Topic Maps,* needed to converge in order to make the Semanitc Web real.[8]

In Steps Topic Maps — Well, Sort of steps in …

Topic Maps properly is the *ISO/IEC 13250* standard. It is built on the Standard Generalized Markup Language (SGML) and HyTime, (another SGML related standard that provides mechanisms for extending SGML for managing hypermedia and multimedia data types), standards and it provides an extraordinarily flexible means for defining topics and the relationship between topics. Unfortunately, SGML’s flexibility is beyond the processing capacity of *all* Web browsers; in fact, XML itself is the technical community’s application of SGML for Web applications.

Even before Topic Maps was published in 1999 a consortium of Web experts, including some of the very same people who wrote and edited the ISO/IEC 1350, began working on an XML version of Topic Maps intended for Web applications known as “XML Topic Maps” or “XTM.” XTM 1.0, published in 2001, provides an extendable mechanism for capturing ontologies and applying them to HTML files. From here forward, “Topic Maps” and “XTM” may be used interchangeably, but we will only be discussing XTM and not its ISO standard SGML-based parent.

Key Concepts of Topic Maps

Topic Maps allow users to convey knowledge about their Web resources through a superimposed “map” of the resource’s subjects and the relationships between those subjects. This “map” is constructed with *topics, associations*, and *occurrences*. Topics represent real-world names for subjects such as “baseball” or “sports.” Topics can also have *occurrences*, which in the case of “baseball” may include computer and non-computer instances were “baseball happens,” such as a listing of regional little leagues or the Major League Baseball website; hence, a *topic* can be either a name that has an associated subject or a thing that embodies a particular subject. Another way of putting this is to say that topic subjects and occurrences can be *addressable* and *non-addressable*.

[7] *RDF Primer*, W3C Working Draft 19 March 2002, Worldwide Web Consortium, http://www.w3.org/TR/2002/WD-rdf-primer-20020319/.
[8] *The Semantic Web*, by Tim Berners-Lee and Ralph Swick, http://www.w3.org/2000/Talks/0516-sweb-tbl/all, W3C MIT/LCS, WWW9 Amsterdam, May 16, 2000.

An addressable subject or occurrence is one that has an associated URL, URI, file location, or other data item that can be used by a computer. However, many subjects and occurrences are non-addressable. Non-addressable subjects may be contained in media (occurrences) that is not accessible by computers (e.g., books, newspapers, etc.), and can also be concepts, ideas, and so forth.

Topics can define more than one subject as well, which is helpful in dealing with synonyms and multi-language applications. For instance, a single topic may be assigned to both “automotive” and “automobile,” or a single topic may encompass the subjects of “baseball” and “el béisbol.”

Topics can also participate relationships, called *associations*, in which they play roles as *members*. For instance, “baseball” could be a *member* of “sports.” Occurrences may also have relationships to subjects, such as “discussed-in, “mentioned-in,” or “depicted-in.”

Topic Maps have some other key properties. A Topic Map is a document (contained within your HTML header section or by itself) that is encoded in XML. A Topic Map may assign several topics to a subject, and multiple Topic Maps can be merged into one Topic Map. The ability to merge topic maps is important to the management of information across enterprises and may be of particular importance to search engines that are developed to make use of topic maps. When merging two topic maps there are several possible outcomes. For instance, if two topics A and B) are to be merged, the results may include:

- A single topic M exists,
- The set of name characteristics of M is equal to the union of the set of name characteristics of A and B,
- The set of subject indicators of M is equal to the union of the set of subject indicators of A and B,
- The addressable subject of M is equal to the addressable subject of either A or B,
- M replaces A and B as a player of any roles played in associations in the topic map,
- The set of occurrence assignments of M is equal to the union of occurrence assignments of A and B, and
- A and B no longer exist.

Topics are organized within a topic map by classes and instances. Classes of topics serve to organize instances of topics within like groupings. For instance, “baseball,” “hockey,” and “soccer” can be members of the super class “sports.” There may be super-classes and subclasses. For instance, “sports” may have the subclasses of “major league,” “minor league,” and “little league.” As such, they may share assigned characteristics such as a single association to the subject “rules of baseball,” but then may have their own scope and variations in further subordinated classes or subjects.

Finally, topics can have a *scope* and scopes can vary depending on context. For instance, the topic of “baseball” could have a scope that limits it to Major League Baseball, or the history of baseball, or regional little league baseball; hence, variations in the subject’s ontology can be accounted for depending on the context the topic map encompasses.

The Simplicity of Topic Maps

This may all sound terribly complex, but compared to its ISO parent standard or other XML specifications, XTM is relatively simple. All of the above-mentioned functions are handled rather economically by XTM — there are only nineteen element types or tags in the entire XTM 1.0 specification. The following are the 19 tags (with brief definitions) in a hierarchical organization:

- **<topicMap>:** Topic Map document element
  - **<scope>:** Reference to Topic(s) that comprise the Scope
  - **<topic>**: Topic element
    - **<baseName>**: Base Name of a Topic
    - **<baseNameString>**: Base Name String container
      - **<variant>**: Alternate forms of Base Name
      - **<variantName>**: Container for Variant Name
        - **<parameters>**: Processing context for Variant
    - **<subjectIdentity>**: Subject reified (e.g., represented or identified) by a Topic
    - **<instanceOf>**: Points to a Topic representing a class
    - **<association>**: Topic Association
      - **<topicRef>**: Reference to a Topic element
      - **<subjectIndicatorRef>**: Reference to a Subject Indicator
      - **<member>**: Member in Topic Association
      - **<roleSpec>**: Points to a Topic serving as an Association Role
  - **<occurrence>**: Resources regarded as an Occurrence
    - **<resourceRef>**: Reference to a Resource
    - **<resourceData>**: Container for Resource data
  - **<mergeMap>**: Merge with another Topic Map

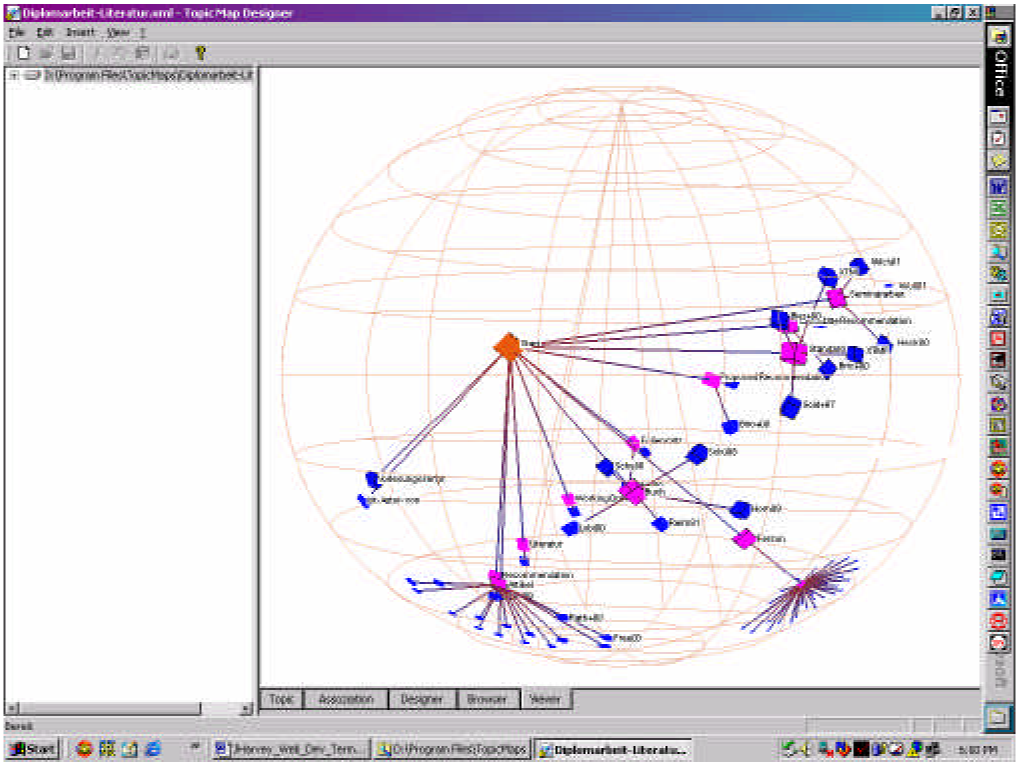
Putting Topic Maps to Work

Having XTM provides a standard for organizing and capturing human ontologies in computer form. Combined with RDF and the concept of the Semantic Web, you can now see how the problem of capturing the fee association of the Web in a computer processable could be solved, but you also must now understand how daunting the task is. Although, XTM is simple enough to hand code, the sheer magnitude of human ontologies and topic specific diction makes the prospect of creating topics maps by hand a scary proposition. There are, however, several tools that have already entered the market. Topic map and semantic web tools include knowledge "engines," topic map editors, visual semantic management tools, and even specialty items such as software that will take an current XML document type definition or schema and generate a topic map automatically. A complete list of available topic map and semantic web specialty software tools can be found in Appendix B below.

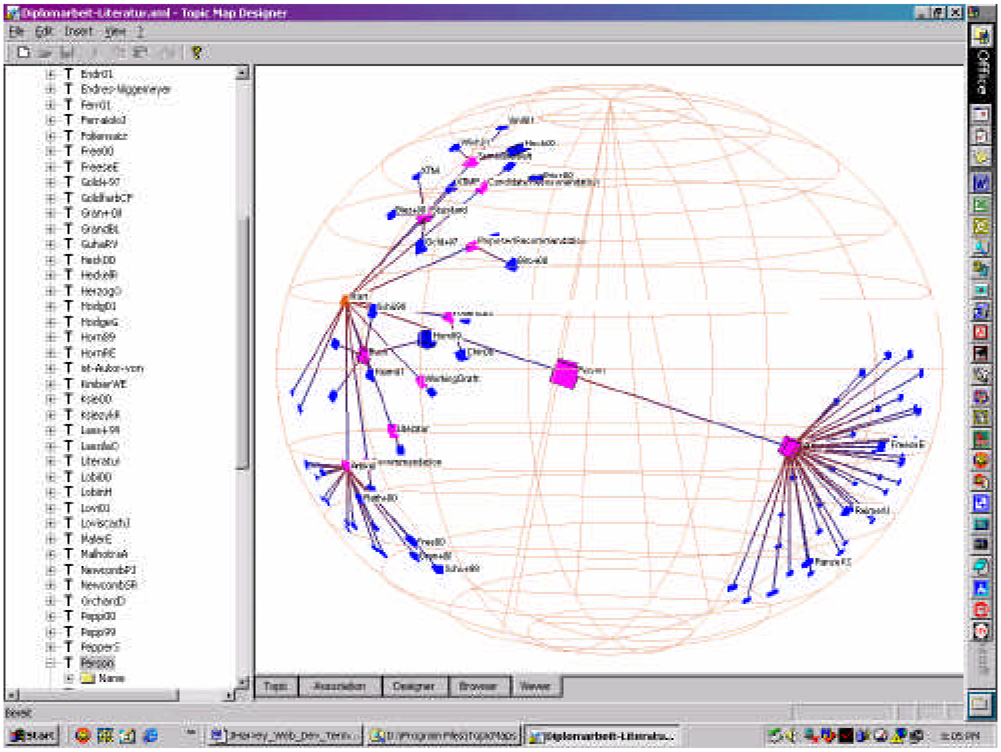
An Example

The following is a graphical view of a Topic Map that captures the work of XTM.org, the consortium that created XML Topic maps. The base name for the highest-level topic of this topic map is "Topic." This view (and subsequent views) of this topic map was created with the free Topic Map Designer 1.1, by Ronald Heckel. Topic Map Designer includes two sample Topic

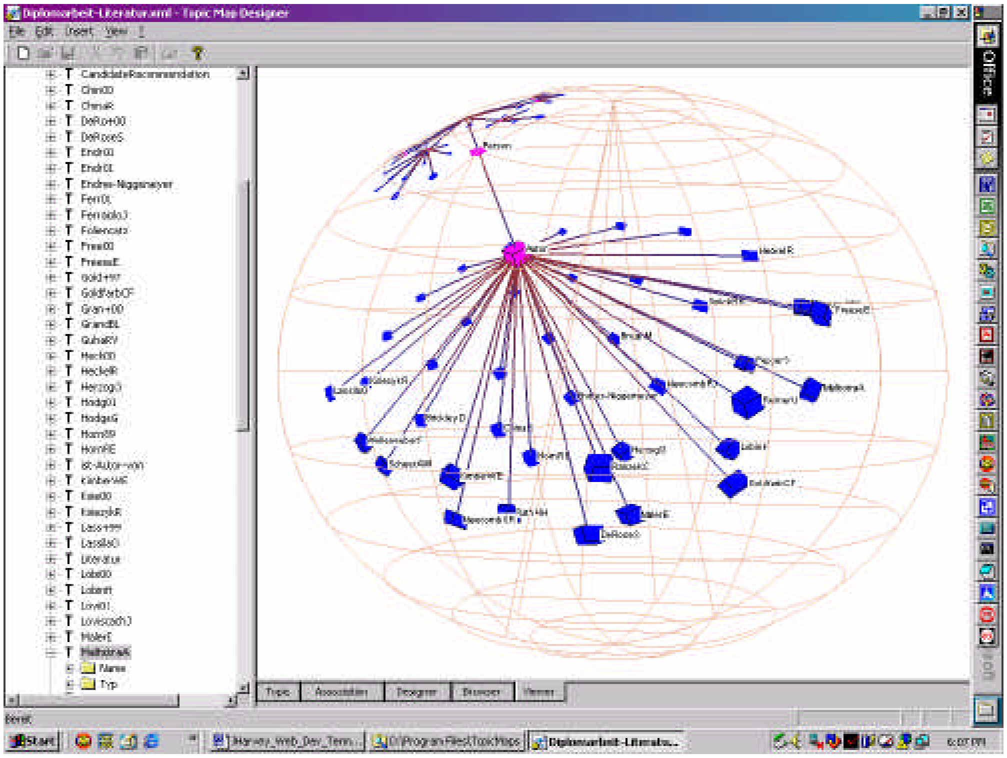
Maps, including this one. The first view shows the highest level of XML.org topics, including committees, standards, recommendations, and person.



The topic "Person" serves as a class that in turn owns the topic of "autor," which is German for "author."



Finally, we see that the subject “autor” has many subjects, which are the individuals that are authors.



Although this Top Map ends there, these individuals could be related to occurrences such as email addresses, or other classes of topics such as mailing address. The graphical view helps to provide clarity for the reader and is also an example of what it means for Topic Maps to provide computer processable ontologies. The following is an example of the code for the same topic map:

Example:

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<topicmap xmlns:xlink="http://www.w3.org/1999/xlink/namespace">
<topic id="Informationsressourcenpool" types="Fachbegriff">
   <topname>
        <basename>Informationsressourcenpool</basename>
        <dispname>beschreibt die Gesamtheit der in der Topic Map verwendeten Informationsressour-
        cen</dispname>
   </topname>
</topic>
<topic id="TopicType" types="Fachbegriff">
   <topname>
        <basename>TopicType</basename>
        <dispname>beschreibt den Typ des Topics im Sinne einer Klasse-Instanz-Relation; ein Typ verweist
        wieder auf ein Topic</dispname>
   </topname>
</topic>
<topic id="Standard">
   <topname>
```

```
            <basename>Standard</basename>
        </topname>
    </topic>
    …
```

The Strengths and Weaknesses of Topic Maps & Conclusion

Topics Maps does provide one way to add structural information to the Worldwide Web that is relatively easy to add. There are however, only a handful of tools available that make use of topic maps. Notable omissions include HTML browsers and Worldwide Web search tools. Topic Maps do not prevent users from creating their own structures and organization for information, which promulgates the flexibility and freedom that has made the Worldwide Web a success, but most likely multiple Topic Maps for the same subject will appear on the Web. Topic Maps provides for this eventuality by facilitating the merger of multiple Topic Maps; however, the ability of software and search engines to complete these merges is not proven. Furthermore, subjects and topics are likely to be defined and associated to other subjects and topics in redundant and conflicting ways. Just any two experts in a given field of study are likely to have disagreements, Topic Maps produced by any two experts in a given field are likely to conflict as well.

It is too early to say whether or not Topic Maps will be successful. In our review of Semantic Web and Topic Map resources and activities, we found that there is very little coordination between the two areas of standardization, yet they may be mutually dependent on each other for their success. The study of knowledge management and human ontologies is much older than either the Semantic Web or Topic Maps, and the participation of language specialists, social psychologist, cognitive psychologist, and librarians in the application of topic maps to centralized and publicly available Topic Maps may be critical to the success of the Semantic Web, but this creates a conflict with the decentralized spirit of the Worldwide Web. Fundamentally, the Worldwide Web must either mature and become reliable global information resource, or continue to be a broadly popular but problematically unreliable information resource. Whether or not the Semantic Web and Topic Maps is adopted, may be the answer to whether or not the Worldwide Web can mature into a reliable information resource.

About the Author

Jim Harvey obtained parallel degrees in Psychology and Sociology at SUNY Buffalo and was a certified crisis prevention and suicide counselor. A year spent at a computerized dating service provided the bridge between his education and career. He spent seven years a Volt Information Sciences as its business development manager and was responsible for establishing SGML, CD-ROM, digital printing, online systems, and data warehousing services and products at Volt in the late 1980s and early 1990s. From 1993 to 2001, he was VP of Spectrum Operations at Graphic Communications Association (now IDEAlliance), and in 2001 he established his own consultancy, Media4theWorld. Media4theWorld has no mission or business plan, as Jim has been too busy to get around to it. The Semantic Web concept makes him ‘all tingly.’

References:

- *Business Model for the Semantic Web: Enterprise Application Integration and other stories*, by Tim Berners-Lee, http://www.w3.org/DesignIssues/Business.html, first written October 25, 2001, last change: October 29, 2001/10/2001.
- ISO/IEC 13250:2000 Topic Maps: Information Technology -- Document Description and Markup Languages, Michel Biezunski, Martin Bryan, Steven R. Newcomb, ed., 3 Dec 1999, http://www.y12.doe.gov/sgml/sc34/document/0129.pdf
- *Overcoming Information Overload* by Paul Krill, http://www.infoworld.com/articles/ca/xml/00/01/10/000110caoverload.xml, InfoWorld, January 7, 2000.
- *RDF Primer*, W3C Working Draft 19 March 2002, Worldwide Web Consortium, http://www.w3.org/TR/2002/WD-rdf-primer-20020319/.
- *Search Engine Relationship Chart ™*, (Cover image), http://www.bruceclay.com/web_rank.htm, Bruce Clay, LLC, July 21, 2002.
- *Semantic Web - XML2000*, by Tim Berners-Lee, http://www.w3.org/2000/Talks/1206-xml2k-tbl/, Graphic Communications Association, XML 2000 Proceedings, December, 2000.
- *The is an Internet … E-mergency [Shocking Stats]*, by Michael Fortino, http://www.e-mergency.com/stats.html, The Fortino Group 2001.
- *The Semantic Web*, by Tim Berners-Lee and Ralph Swick, http://www.w3.org/2000/Talks/0516-sweb-tbl/all, W3C MIT/LCS, WWW9 Amsterdam, May 16, 2000.
- *The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, by Tim Berners-Lee, James Hendler, and Ora Lassila, Scientific American, http://www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2, May 2001.
- The UCLA Internet Report 2001 – "Surveying the Digital Future," http://ccp.ucla.edu/pdf/UCLA-Internet-Report-2001.pdf, UC Regents, November 2001.
- *Yahoo Now Charging Annual Listing Fee*, by Danny Sullivan, http://www.searchenginewatch.com/sereport/02/01-yahoo.html, The Search Engine Report, January 7, 2002.
- What the Semantic Web can represent. By Tim Berners-Lee, http://www.w3.org/DesignIssues/RDFnot.html, September 1998.
- XML Topic Maps (XTM) 1.0, edited by Steve Pepper and Graham Moore, TopicMaps.Org, http://www.topicmaps.org/xtm/index.html#ref_iso13250, August 2001.

Appendix A — Semantic Web and Topic Map Tools

Source: http://business.semanticweb.org/, July 21, 2002

Company	Product	Product Category
AIdministrator www.aidministrator.nl	Sesame Spectacle	RDF(S) storage & retrieval Ontology-based information presentation
Applied Semantics www.appliedsemantics.com	Circa	Ontology-based automatic categorization
Cycorp www.cyc.com	Cyc Knowledge Server	Multi-contextual knowledge base / inference engine
DigitalOwl www.digitalowl.com	KineticEdge	Content Management / Publishing
Empolis www.empolis.co.uk	K42	Topic Map Server
Eprise www.eprise.com	Participant Server	Content Management
Epigraph www.epigraph.com	Xcellerant	Content Management / Ontology Management
forward look inc www.forwardlook.com	ContextStreams	Data Asset Management
GlobalWisdom www.globalwisdom.org	Bravo engine	Facilitated Ontology Construction / Dynamic Knowledge Engine
H5 Technologies www.h5technologies.com	H5 Atlas, H5 AutoTagger & H5 Syndica	Content categorization & omni-contextual knowledge-base/analytics
Infoloom www.infoloom.com	Topic Map Loom	Topic Map editor
Intellidimension www.intellidimension.com	RDF Gateway	RDF Data Management System
Inxight www.inxight.com	ThingFinder Server Star Tree Viewer	Content extraction Web content navigation
Knowledge Processors www.knowledgeprocessors.com	Universal Knowledge Processor	Dynamic taxonomies
Language And Computing www.landc.be	TeSSI	Ontology-based Information Management

Company	Product	Product Category
Mohomine www.mohomine.com	Several	Information extraction and classification
Mondeca www.mondeca.com	Several	Topic Maps to improve Content Management
Network Inference www.networkinference.com	Cerebra	Inference engine and tools
Ontopia www.ontopia.net	Topic Map Engine Topic Map Navigator	Topic Map Client and Server
Ontoprise www.ontoprise.de	Ontobroker	Inference Middleware
Persist www.persistag.com	Semantic Base	Knowledge Management System
Plugged In Software www.pisoftware.com	Tucana Knowledge Store & Tucana Metadata Extractor	Enterprise Distributed Metadata Management Suite
Profium www.profium.com	Smart Information Router (SIR)	Semantic Content Management based on RDF
R-Objects www.r-objects.com	Pepper pepper.r-objects.com	Personal Knowledge Management
Semio www.semio.com	SemioMap	Content Categorization and Indexing
Semtation www.semtation.com	SemTalk www.semtalk.com	RDFS editor based on Visio
Tarragon Consulting Corporation www.tgncorp.com	High-performance knowledge & content management systems	Custom systems design and development
TheBrain.com www.thebrain.com	TheBrain	Information organizer
Unicorn Solutions www.unicorn.com	Unicorn Coherence	Ontology modeling and data integration
Verity www.verity.com	K2	Business Portal Infrastructure
Voquette (formerly Taalee) www.taalee.com	Semantic Engine WorldModel	Knowledge-based Rich Media Content Management