Abstract

The need for accessing electronic information anywhere, anytime is rapidly growing in the current age. The Semantic Web is a relatively new technology to help navigate medical information easily and efficiently process this information by machines, which is eventually read by humans. In my paper, I will discuss the current state of semantic web and its relationship to health related areas, such as accessing and the analyzing of medical information. Examples of semantic web applications and discussion of future work will also be given.

Keywords: Semantic Web, Medical information, BioInformatics, Ontology, Semantic

1.0 Introduction

The development of the Semantic Web is part of a phenomenon called the e-Health Revolution in which “paper” medical records are transformed into a system, related to the Semantic Web [1] should be in place to understand what is meant by the keywords that we use in searching and to relate that to the web pages that are available on the Internet. The semantic web can further aid in searching by filtering the user’s search results and thus returning. With the use of the semantic web, it will be possible for the consumers to access their records, find relevant information about health benefits, physicians, employment and much more. The Internet makes information easily accessible virtually anywhere in the world and is responsible for making the world flatter as time evolves. Due to the unorganized structure of information on the Internet, systems are needed to provide organization.

- Links to data sources.
- Direct access to information without having to go through an intermediate source.
- More efficient measures of ambiguity.

Based on the rapid growth of electronic data, systems are needed in which embedded indices exist, which serves as a platform for more precise data. The figure below illustrates the information environment and its various roles [2]:

![Figure 1. Stages of Information (Musgrave, 2004)](image)

Standards exist to make data between medical information easier [1].

A vocabulary technical committee has been formed that provides repositories to maintain coded vocabulary for various domains. This allows interoperability between various domains. Currently, interoperability between different domains is practically non-existent. Semantic web enables this capability.

Semantic web searches are also used for wireless devices using WXML and micro browsers. Since the medical staff and the consumers are increasingly using wireless devices such as
wireless phones, PDAs and others, semantic web can reduce time and amount of information [3].

2.0 Semantic Web and its features

Scientists and researchers are currently looking at ways in which Semantic Web technology can help in evaluating and analyzing different kinds of data. The director of a web conference suggested that “life scientists immediately start adopting a Semantic Web approach with their data [4] He further explained how the Semantic Web can be useful to their data since the field deals with a huge amount of heterogeneous data types and formats to solve a multi-faceted problem. A company called Partners HealthCare System is developing a Semantic Web-based decision support system. Resource Description Framework (RDF), a standard for Semantic Web, is used by Partners to make computer models process electronic medical records more efficiently. Once the data set is in RDF format, the Semantic Web Rules Language can be used to write decision statements for evaluating patterns among patients. The Web Rules Language can further be used in helping to develop a patient diagnostic test.

Many databases and SQL queries can do the same thing as the Semantic Web allows for, but with Semantic Web the coding language is very flexible, accurate and precise for searching for various things, such as drug discovery. It is stated that “there is a critical need to develop an informatics and knowledge model across the drug [development] pipeline” test. [4]. Since the field of drug discovery is volatile, the Semantic Web is a great tool for this purpose. With Semantic Web technology, more sophisticated and powerful approaches can be used with vast amount of data. It is said that “with the Semantic Web, you publish meaning, not just data” test [4].

Words, such as homophones, can be tagged with code (XML) to enable computers to find synonyms for the underlying word. The computer filters, categorizes and searches the information sources before presenting the web content for humans [1]. To access the information from text based web content and databases, applications have to incorporate intelligence to process the information. In the next step, data is tagged in XML so that it can be accessed by different applications in a single domain, for example, XML in the health care system.

The difference between XML and semantic web is that XML uses metadata but semantic web adds meaning and relationship between the metadata. The next step for the tagged data is to classify it in a hierarchical taxonomy. This data can be related to each other and combined by using simple relationships between categories in the taxonomy [5].

New data can also be inferred from existing data by following logical rules. An example of the application of semantic web is HP Lab semantic blogging demonstrator with emphasis on bibliography. A key word search can yield results in RDF or N3 formats. The meta data that results from this is fairly standard for a blog: they are

- Creator
- Title
- Description

The metadata can be edited and new blog entries can be added [6].

2.1 Resource Description Format (RDF)

The goal of semantic web research is to produce a machine-readable language, such as Resource Descriptor Framework (RDF), to describe and query Web resources more efficiently. Furthermore, RDF mainly describes innate objects and the relationships between them. This means that it will be extremely easy to reuse RDF information for various mobile devices. People have to publish web data in RDF format in order to take advantage of the Semantic Web. RDF systems take user data as input and store it in RDF format [5].

Electronic resources, such as files must be identified by a resource ID (URI), which can be globally identified. Slow capture of statements in a formal way allows slow aggregation of a knowledge base. The following table gives the RDF format:

<table>
<thead>
<tr>
<th>Language</th>
<th>Subject</th>
<th>Predicate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object oriented</td>
<td>Class</td>
<td></td>
</tr>
<tr>
<td>Property</td>
<td>Value</td>
<td></td>
</tr>
</tbody>
</table>
Using URIs in a document codes the triples of RDF and concepts are tied to a definition, which can be found by everyone on the web [7]. The goal is to make the computers understand the data. The statements can be illustrated in an RDF graph. The following XML based syntax and RDF graph of Dr. Eric Miller is given below [8]:

```xml
<?xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">
  <contact:Person rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
    <contact:personalTitle>Dr.</contact:personalTitle>
  </contact:Person>
</rdf:RDF>
```

2.2 Ontology

This is an area of knowledge (domain) that can be described as concepts and words such as, important things in the domain, their properties, and relationships among the things (rules). Ontology can also be defined as semantic similarity of terms of taxonomies [7].

Medical terms are common ontologies that are used by different applications. As a result of this, they should not be specific to any one client application but should be shared and reused among several applications. They can be shared and reused via a terminology server, which is a special kind of server used for recovery of similar ideas, and making them available for databases. The existence of different, sophistical medical domain ontologies is very important if used in any Semantic Web application and thus can help to save time and resources from searching [9].

An analyst states “the World Wide Web, as we know it today, is mostly unstructured content” [1]. Furthermore, the Semantic Web would basically allow search engines to become more specific by instilling more meaning and structure into words [1]. Nevertheless, some people argue that the term “semantics” is misleading due to the fact that the term semantics is usually used to define meaning via human interaction, and Artificial Intelligence is the area of research that is primarily concerned with machine replication of biological terms.

3.0 Semantic Web and Health Care

As health care can be intertwined with economic and environmental data found in databases, the semantic web could search for the location of sick people and data that is linked to environmental and economic factors that may have caused the illness. The Internet Relay Chat (IRC), an online tool used by semantic web community to chat online with each other, is a great device for real-time CHI (Consumer Health Information) exchange. In addition to this, users can use discussion groups for further exchange [10].

Examples
1. HealthCyberMap, a Semantic Web portal project, categorizes Internet health information...
resources in such a way that retrieval and navigation is easy to do. Health Information is categorized via health information visualizations using metadata, clinical codes and Geographic Information Systems [9].

An example of the HealthCyberMap categorization is given below:

![HealthCyberMap](image)

**Figure III.** BodyViewer...Human body maps with semantic web can help in visually browsing clinical information [9].

Clinical codes to describe medical Web resources can help improve metadata quality and offer better categorization of medical resources. In addition, clinical codes can create semantic relationships between similar medical content in databases. Two people, Appleyard and Malet, stated that the assimilation of the Unified Medical Language System (UMLS) and the Systematised Nomenclature of Medicine (SNOMED) into metatags will improve the quality of information in electronic medical systems. A whole electronic healthcare vocabulary should contain synonyms for different diseases and other medical terms, and should be able to interact with higher level meanings. A tool called Protégé-2000 contains modeling templates used to provide expression and clarity to the user [9].

This tool (Protégé 2000) had many features, including support for UMLS. This is shown in Figure IV:

![Protégé-2000](image)

**Figure IV.** Snapshot of Protégé-2000 using UMLS [9].

2. An another example of the use of semantic web in healthcare systems is the WebDG system. The authors for the Indiana social service administration designed this system for low-income citizens, seniors and disabled people. WebDG is written in web service description language. Two ontologies are defined for future semantic web services, category and type. An operation consists of two elements, one from each ontologies. An element has three attributes:

- **Name** (such as healthcare, adoption, nutrition for category ontology and eligibility, counseling and mentoring for type ontology)
- **Synonyms** (healthcare has a synonym as medical)
- **Specialization** (a set of categories).

It is important to maintain compatibility between type and category. Some composition rules are devised for that purpose. According to these rules, the compositibility operation compares the domains or categories of interest and chooses the optimized pair. One example could be the combination of...
healthcare and adoption from each domain. Composition soundness is implemented next, which measures added value by the combination of these two operations. The E-government provides this value by some stored templates. For example, a search for pregnancy not only would give a list of healthcare providers, but would give information on nutrition, housing, legal services and employment. This is a system for consumers on Medicaid to search for government medical help. Patient privacy requirements are also met in the system implementation [11].

Figure V shows the various parts of the systems. The client query scheme was developed for the database systems.

3. Patterns can be used to determine the semantic interoperability among healthcare systems. EbXML registry semantic constructs can be used to analyze these patterns. A patient’s health data can be located in many different locations. It is this reason that physicians are not able to retrieve patient data easily. Semantic interoperability means that common information between two or more systems are understood by a higher domain level. This can be reached through the development of a single information standard for healthcare. In healthcare, information sharing is common and thus, different archetypes (reusable expression of a distinct concept) are used. The language used for defining archetypes is the Archetype Definition Language (ADL) [8].

The illustration of this language is shown in the following figure, which shows a section of the “Complete Blood Count:”

```
OBSERVATION[at1000.1] matches {  
  code matches {ac00041} -- complete blood count} 
  data matches { 
    LIST.S[at1001] matches {-- battery 
      items cardinality matches (0..+) \subseteq { 
        ELEMENT[at1002.1] } matches {-- hemoglobin
          name matches { 
            CODED_TEXT matches { 
              code matches {ac0002} -- hemoglobin
              value matches { 
                QUANTITY matches { 
                  value matches (0..1000)
                  unit matches (g/l|g/dl|..)}}} 
          } 
    } 
  } 
  ELEMENT[at1002.2] occurs occurrence matches (0..1) matches {-- hematocrit
    name matches { 
      CODED_TEXT matches { 
        code matches {ac00041} -- hematocrit
        value matches { 
          QUANTITY matches { 
            value matches (0..100)
            unit matches ("%")} 
        } } 
    } 
  ELEMENT[at1002.3] occurs occurrence matches (0..1) matches {-- platelet count
    name matches { 
      CODED_TEXT matches { 
        code matches {ac00058} -- platelet count
        value matches { 
          QUANTITY matches { 
            value matches (0..100000)
            unit matches ("/mm³")} 
        } } 
    } 
  } 
```

Figure VI. ADL definition of Blood Count [8].

In order for various healthcare systems to share information with each other, they need to incorporate different archetypes and the meaning related to them. In the figure below, archetypes and ontology plays a major role in establishing communication between healthcare databases/system [8].

![Figure VII. Archetype Metadata][8]
Templates, such as the one below, are created for the purpose of querying healthcare systems through different domains:

![Image of a brain template](image)

**Figure VIII.** Brain Template [8].

4. Another example of semantic web application for health care systems is the effort to convert some biological ontologies (GO and MGED) into the web ontology language (OWL). Public biological databases like Uniport are being converted into RDF. OWL is one level of abstraction above RDF. It goes beyond XML and RDF to represent machine interpretable content on the web. OWL adds more vocabulary to its domain. One application of OWL is in the use of BLAST search [12].

The paper describes the interoperability, accessibility and interaction between different domains by the use of Owl and medical ontologies. The author focuses on the application of the e-business model to the e-Health systems involving the interaction of independent resources from different domains in different entities [8]. The goal of e-Health is given so that the patients can access their records from anywhere, anytime and “clinicians have access to critical health care information when treatment decisions are being made.” He applies this to e-prescription and outlines a system where the ebXML registry is translated into an OWL repository. OWL, the web ontology language provides another layer of abstraction over RDF to objectively connect the different RDF ontologies. OWL explains the domains in terms of classes and properties and connects different domains and entities subject to some rules.

### 4.0 Discussion

In this paper, I have described four healthcare systems used in conjunction with the Semantic Web. HealthCyberMap is a Semantic Web project that uses all kinds of techniques, from visualization to tactile, for the purpose of easy information retrieval and manipulation. WebDG is a Web Service Management System to allow users to perform services offered by the government, such as in health care. EbXML and archetypes are used to provide semantic interoperability, or information sharing, among healthcare systems. OWL is a web ontology that is slightly above RDF from an abstraction point of view.

Not all of the examples I mentioned are being implemented today. Some of the systems are still being researched or in the review stage. WebDG, EbXML and archetypes are implemented as real systems and are important components to make semantic interoperability possible. HealthCyberMap is still a semantic project that will hopefully, lead to the development of a Semantic Web Portal. Some of its elements are functional, such as BodyViewer (the application I talked about earlier), while other components are still under research. OWL is an ontology that is used to define classes, subclasses and their properties, and connects different ontologies containing metadata. The Semantic Web has grown and it has the capability to provide a way that different healthcare systems communicate with each other.

### 4.1 Future Work

Currently, there is no strong interoperability among health care systems and even if they were, those systems would be proprietary. However, OWL and various archetypes provide some communication among healthcare systems. The health information systems that are proprietary may serve only a few departments within a healthcare company or institution, and a patient’s health data may be in different locations that are not interoperable. The semantic Web may help to solve this problem by providing meaning to data and utilizing metadata for communicating among
different systems. More standards will be available for the future. Work is also in progress to develop more sophisticated ontologies and to convert databases into metadata [13].

5.0 Conclusion:

The goal of Semantic Web is to make information retrieval more user friendly. As the Internet is growing rapidly in this modern era, information is a vital tool to have. Semantic Web for healthcare systems helps interoperability among healthcare systems. The Semantic Web also allows for more specific searching, and can be used to search for health-related information. Words will have meaning assigned to them and synonyms can aid in searching.

Acknowledgements

I like to thank Dr. T. Dennison for encouraging me to take a course on Consumer Health Information for which I wrote this paper. I would also like to thank my parents and Ms. Andrea Scala for proofreading my paper.

6.0 References: