

HealthCyberMap: Mapping the Health Cyberspace Using Hypermedia GIS and Clinical Codes

PhD Research Project

Maged Nabih Kamel Boulos

MBBCh, MSc (Derm & Vener), MSc (Medical Informatics)

1 Summary

“The application of GIS is limited only by the imagination of those who use it.”

—Jack Dangermond, President, ESRI

HealthCyberMap (<<http://healthcybermap.semanticweb.org>>) is a Semantic Web service for healthcare professionals and librarians, patients and the public in general that aims at mapping parts of medical/ health information resources in cyberspace in novel ways to improve their retrieval and navigation. The Semantic Web (<<http://www.w3.org/2001/sw/>> and <<http://www.semanticweb.org>>) aims to be the next-generation World Wide Web by giving machine-readable semantics and context to the currently presentation-based Web pages.

HealthCyberMap features an unconventional use of GIS (Geographic Information Systems) to map conceptual spaces occupied by collections of medical/ health information resources. Besides mapping the semantic and non-geographical aspects of these resources using suitable spatial metaphors, HealthCyberMap also collects and maps the geographical provenance of these resources. Some of HealthCyberMap Web interfaces are visual (maps for browsing resources by clinical/ health topic, by provenance and by type), while others are textual (multilingual interfaces for browsing resources by language, and a directory of topical resource categories, besides HealthCyberMap Semantic Subject Search Engine that goes beyond conventional free-text and keyword-based search engines, and supports synonyms, disease variants, subtypes, as well as some semantic relationships between terms).

HealthCyberMap adopts a clinical metadata framework built upon a clinical coding scheme (vocabulary or ontology—ICD-9-CM clinical classification in the current pilot service). Clinical coding schemes serve as a reliable common backbone for topical resource indexing, automated topical classification, topical visualisation and navigation of coded resource pools (using suitable metaphors), and enhanced information retrieval and linking. A resource metadata base based on Dublin Core metadata set with HealthCyberMap’s own extensions holds information about selected high-quality resources. HealthCyberMap then uses GIS spatialisation methods to generate interactive navigational cybermaps from the metadata base. These visual cybermaps are based on

familiar metaphors for image-word association to give users a broad overview and understanding of what is available in this complex conceptual space of medical/ health Internet resources and help them navigate it more efficiently and effectively.

HealthCyberMap cybermaps can be considered as semantically-spatialised, ontology-based browsing views of the underlying resource metadata base. Using a clinical coding scheme as a metric for spatialisation (“semantic distance”) is unique to HealthCyberMap and is very much suited for the semantic categorisation and navigation of medical/ health Internet information resources. HealthCyberMap also introduces a useful form of cyberspatial analysis for the detection of topical coverage gaps in its resource pool using choropleth (shaded) maps of human body systems. The project features a cost-effective method for serving Web hypermaps with dynamic metadata base drill-down functionality. It also demonstrates the feasibility of Electronic Patient Record to Online Information Services (like HealthCyberMap) Problem to Knowledge Linking using clinical codes as crisp problem-knowledge linkers or knowledge hooks.

The Semantic Subject Search Engine queries the same HealthCyberMap resource metadata base. Explicit concepts in resource metadata map onto a brokering domain ontology (ICD-9-CM) allowing the search engine to infer implicit meanings (synonyms and semantic relationships) not directly mentioned in either the resource or its metadata. Similarly, user queries would map to the same ontology allowing the search engine to infer the implicit semantics of user queries and use them to optimise retrieval.

A formative evaluation study of HealthCyberMap pilot service using an online user evaluation questionnaire, in addition to analysis of HealthCyberMap server transaction log, has been conducted during the period from 18 April 2002 to 1 June 2002 with very encouraging results. This two-method evaluation approach was guided by methodologies described in NIH[†] Web Site Evaluation and Performance Measures Toolkit among other resources.

Many exciting future possibilities have been also investigated by the author, including the further development HealthCyberMap as a customisable, location-based medical/ health information service.

Keywords

Semantic Web, Resource Description Framework (RDF), Metadata, Dublin Core, Ontologies, Controlled Vocabularies, Clinical Codes, Problem to Knowledge Linking, Clinical Decision Support, Geographic Information Systems (GIS), Cartography, Cyberspace, Cybergeography, Cyberspatial Analysis, Medical/ Health Information Spaces, Spatialisation, Spatial/ Visual

[†] National Institutes of Health, US

Metaphors, Internet/ Web Resource Visualisation, Web Browsing/ Navigation, Cybermaps, Web Maps, Hypermaps, Hypermedia, Digital Libraries, Information Retrieval, Medical Search Engines, Location-based Services, Evaluation of Online Information Services, Server Transaction Log Analysis

2 Aims and Objectives

Background and Motivation

The Internet is a system with unimaginable complexity. It has been said the Web is its own map. But as we surf the Web, we can only appreciate a very small part of it at any given time. We cannot for example figure out the relations of the page/ site we are visiting to the rest of the Web, how it measures compared to the rest of the Web, or how the rest of the Web looks like. We can have difficulties finding our way back to a resource we visited just few hours ago. We cannot easily locate medical/ health Web resources covering narrower, broader, related or similar topics as the one we are currently looking at, and except when we have no particular goal(s) in mind, we cannot plan our Web journeys ahead. We need a map or set of maps for this purpose. Ideally, we should be also able to link our health-related and clinical questions to such maps, and use them (or the underlying information) to find the appropriate answers to these questions in the right place and at the right time.

Maps are well known as powerful graphic tools that can be used to classify, visualise, communicate and navigate spatial and/ or spatialised relations (relations projected into some conceptual space for mapping) in worlds that are too large and too complex to be seen directly. Maps build on humans' powerful spatio-cognitive abilities. Good Web navigational maps should be also based on sound cartographic principles and variables (to ensure their spatial legibility and utility), and adapted to the unique nature of the Web.

GIS (Geographic Information Systems) take simple cartography one step further by providing contextual links between maps and underlying databases (where attributes of features on the maps are stored). On the Web these links can be implemented as sensitive clickable maps (hypermaps). The quality and utility of such maps will obviously depend on the quality of the information in the underlying databases.

Aims and Mission Statement

HealthCyberMap (<<http://healthcybermap.semanticweb.org>>) is a Web-based service that aims at mapping medical/ health information resources in cyberspace in unique and novel ways to deliver a semantically superior experience to resource providers and consumers (the general public, patients, healthcare professionals and librarians),

helping them plan, manage, search and navigate this complex “virtual” space more efficiently and effectively. This is achieved through semantic indexing, “intelligent” categorisation, and interactive hypermedia visualisation of the medical/ health information cyberspace using metadata, clinical codes and GIS technologies. HealthCyberMap only attempts to map *parts* of available health and healthcare Internet resources and services.

Objectives

(with measurable outcomes)

A navigational hypermap for browsing collections of Internet information resources should be ideally driven by an underlying metadata base that stores meta-information (information about information) describing these resources. HealthCyberMap comprises two main arms or layers. The top-level visualisation/ navigation arm (interface layer) is founded upon a robust semantic layer.

Semantic Arm Objectives

1. To define and model a suitable framework that tailors existing generic metadata standards/ recommendations to suit the description of medical/ health information resources.
2. To improve resource metadata quality and semantics through proper use of clinical codes from a controlled vocabulary as descriptors of medical/ health resource contents.
3. To propose and examine ways of reasoning semantically with resource metadata (including clinical codes) beyond simple queries, and combining metadata with other related and useful knowledge bases (ontologies or vocabularies) for the purpose of customising the language, interface and content of material delivered to consumers and offering them location-specific information.
4. To build a pilot metadata base describing a small part of medical/ health Internet resources based on the above framework, with the application of suitable resource quality benchmarking measures.
5. To develop a pilot Medical Semantic Subject Search Engine that outperforms conventional free-text and keyword-based search engines, and supports synonyms, disease variants, subtypes, as well as some semantic relationships between medical/ health terms.

6. To develop a pilot clinical Problem to Knowledge Linking solution based on clinical codes to contextually link the Electronic Patient Record to the resource metadata base.

Visualisation/ Navigation Arm Objectives

7. To find an effective semantic metric based on clinical codes that can serve as a basis for the topical mapping of medical/ health information resources in cyberspace.
8. To find and use suitable spatialisation methods and familiar metaphors to visually browse and navigate medical/ health information resources on the Internet in different (but complementary) ways based on clinical codes and other metadata elements in the resource metadata base. This objective will be complemented by carrying a review/ critique of some existing and related cybermapping projects.
9. To use GIS to automate the classification and generation of spatialised browsing views (navigational cybermaps) of the resource metadata base based on the above cybermapping methods.
10. To develop a method for identifying topical coverage gaps (infogaps) in mapped resources.
11. To develop a cost-effective solution for serving hypermaps with dynamic database links/ drill-down functionality on the Web and use it for publishing the GIS-generated navigational maps on the Web while maintaining their links with the underlying resource metadata base (these maps form the main part of HealthCyberMap pilot service on the Web). It should be also possible to use the same solution to publish other interactive GIS-driven maps on the Web, e.g., maps of real-world health problems.
12. To develop a strategy for maintaining the currency of the generated hypermaps and dealing with problems related to Web resource link persistence.

Evaluation Objectives

13. To identify the need and reasons for evaluating HealthCyberMap pilot service interfaces on the Web, and investigate and develop/ refine suitable methods for conducting this evaluation and analysing its results.
14. To conduct a small-scale formative evaluation of HealthCyberMap research pilot using suitable methods (developed in the above objective), analyse the results, and

identify any useful lessons that could be learned and possible future directions that could help improving HealthCyberMap if the service is further developed beyond the current project plan and timeline.