# A Topic Map based Framework for Managing Medical Data in Dermatology Domain

Georgios Grammatikopoulos[1], Christos Manavis [2], Georgios Chaidemenos[3]
[1]Department of Esthetics, [2]Department of Electronics,
Technological Educational Institution of Thessaloniki
P.O BOX 14561, GR - 541 01, Thessaloniki , GREECE
[3]Hospital of skin and venereal diseases, Thessaloniki, Greece

*Abstract*: In the context of this paper we discuss the potential use of topic maps to manage medical information, make a brief review of related attempts and consequently propose a web-based framework that incorporates topic maps technology for managing medical data in the dermatology domain. A key factor, in order to profit from the application of web technology in healthcare, it is essential to represent the way medical data is expressed and perceived in electronic form. Topic Maps as a new technology of semantic web, seems to subscribe to the satisfaction of this requirement, as they were introduced in order to build a bridge between knowledge representation and information management.

*Keywords:* medical data, dermatology, semantic web, topic map, retrieval, query

## 1 Introduction

The increasing popularity of World Wide Web as an information service resulted to a proportional increase of the expectations and requirements towards the opportunities of applying new technologies in medical information systems [1]. Walter Fierz [2] remarks that the structured content and connectivity of medical information will lead to "Medical Data Web". Furthermore, he describes the principles on which this "Medical Data Web" must be based [3]. He notices that medical information resides in the connectivity of data. An isolated data element on its own has little information content. Semantic Web provides the conceptual and technical framework for a Medical Data Web. Information profiling is pivotal to the success of a Medical Data Web.

Topic map technology as a representative standard in the field of the Semantic Web, intends to organize information in such a way that it is optimized for navigation and knowledge representation [4,5]. Topic maps(TM) are suitable for large or dynamic information sources, better querying than a static index can be implemented and make the merging of diverse information resources possible, hence there are a lot of potential applications of this new technology [5,6]. TM seem to provide a reasonable way to structure medical data for improving the access to problem relevant pieces of knowledge information [7].

In the context of this paper we briefly discuss the potential use of TM to manage medical information and refer to recent research approaches in this field up to now. In the following sections a prototype system framework is proposed, that incorporates topic map technology for information management of medical data in dermatology domain, ongoing research, and corresponding technical issues.

## 2 Topic maps and medical data

According to Walter Fierz [2] the principles on which the "Medical Data Web" must be based [3] include

- the granularity of data elements
- the way to attach semantic information to the data elements, links and structures
- the storage of the data together with their structure and the connections between the data elements
- a query system for the extraction of the information contained within the structure and connectivity as well as from the data proper
- and the display of the query result in a way that structure and connectivity are intuitively and usefully expressed and can be stored again in a structured, machine-accessible way.

Topic map technology seems to meet these requirements. The international organization for standardization (ISO) of topic maps standard provides a

data model that allows the representation of arbitrary relationships between resources. Such relationships form the basis for a context sensitive search and accurate search results [11].

Topic Maps (TM) describe what an information set is about by formally declaring topics, by establishing typed relationships among topics, and by linking the relevant parts of the information set to the appropriate topics [4,5]. TM provide a more abstract layer of data modeling, storage and querying. The special characteristic of the TM model is the clear separation between the description of the information structure and the physical information resources [5,6].

One obvious question emerges and is concerned with the benefits of the potential use of TM to manage medical data instead of a relational database. Although a relational database can model just about anything, the problem is that in order to add or change properties and relationships between objects there is the need to change the database table structure continuously, and also keep on altering all SQL queries. By modeling various objects in a topic map, a potential user may keep on altering properties and relationships, without having to rewrite program code or queries. So if the potential users data is subject to continuous alterations, and utilizes enriched relationships, TM offer a viable solution to manage information [8].

One of the advantages of such a project is that the same collection of information can be used for more than one purpose or audience. TM allow users to create different personalized views of the same set of resources capturing viewpoints through the use of scopes, e.g., a view for consultant dermatologists and another one for general practitioners treating skin diseases [9]. Additionally, TM permit the representation of knowledge in an interchangeable form [5]. Thus TM-based applications promote reuse, sharing and interoperability of information resources and could support exchange of information with other health organizations [10]. Furthermore topic maps were designed from the start for ease of merging. The concept of subject identity and the ability to establish a topic's identity through a subject address and multiple subject indicators are key to this capability [5,6]. Hence they support merging of information with other health organizations [9,10].

The scope of this work is to propose a framework for managing medical data in the specific area of dermatology, and eventually exploit the aforementioned potential advantages of using this new technology for knowledge representation and information management of medical data in dermatology domain.

## 3 Related Research Approaches

De Bruijn and Martin [12] support that literature mining offers powerful methods to support knowledge discovery and the construction of topic maps and ontologies and review the recent developments in medical language processing. One of the advantages of such a project is that the same collection of information can be used for more than one purpose or audience. On top of a medical information system a Topic Map can be defined that will give precise information for both a cardiologist and a lung specialist. A number of topics will be equal for both – like blood vessels and cholesterol – but their relations with other topics can be quite different [13].

On the other hand, Beier and Tesche from the hyperCIS [14] have developed a front-end topic map application that works with text search engines in the background. More specifically, the access to relevant, up-to-date and reliable information is a time critical, but nevertheless very important task in the daily work of physicians and nurses. HyperCIS developed an intelligent information retrieval system (IRS) with a knowledge-guided user front-end and an automatic generation of search engine queries.

The medical knowledge of MeSH (Medical Subject Headings) classification was transferred into a Topic Map, the new XML-based knowledge representation standard (ISO 13250). A graphical user-interface allows the fast and associative browsing in networks of themes [15]. Each theme or topic, as a node of the topic map, contains information on title, synonyms, translations, definition, scope, sub- and super classes.

Health Level 7 (HL7) is an American organization working on medical information systems and connectivity and is recognized by the ANSI as a standards producing capable organization. One of the HL7 projects is the Kona project which intents to integrate all hospital filing systems across all the American hospitals and clinics into one comprehensive web XTM system [16].

## 4 Proposed TM-based system for dermatology

The key to our understanding, of the state of dermatology domain is the collection of accurate and relevant data. Such data should provide a current picture of the spectrum of dermatological data and provide clear view of areas requiring attention. However, until recently, current data collections were usually manual systems and too little information was routinely collected.

The scope of this work is the provision of an integrated system providing easy access to personal patient data, medical records and statistics in a dermatology domain and other related information on a large scale. All the patient data will be transformed to machine processable structured data, which then could be related to a knowledge base repository. Since this data will be machine processable the use of rules and queries will be able to deliver all the kind of clinical information necessary for an accurate diagnosis. Electronic systems based on structured data, (semantic data), could use this medical information, in order to build the potential of huge improvements in the quality and completeness of dermatological clinical data.

A knowledge base built upon semantic standards helps to ensure interoperability and platform independence. The use of metadata, which describes conceptually a data source represent the corner stone of the semantic information. The knowledge base repository will represent the semantic understanding of any particular subject {clinical data) as seen by subject experts. The main standards for metadata are mainly the RDF (proposed by W3C) and Topic Maps.

The aim of this work is to illustrate the way topic maps could be applied to manage medical data and discusses some potential benefits of encoding information using topic map technology. The proposed TM-based framework could be used for image indexing and retrieval, and will operate on a collection of dermatological images and associated health clinical data in order to develop a functional system for efficient management of medical data.

A Web-based distributed architecture is important, so that the resources can be accessed anytime, anywhere. Dermatologists may use the system for cases where they need the remote collaboration of other specialists. The data of such a system should be complete, comprehensive, consistent, reliable, and timely. All medical information acquired from a TM-based medical application should be available to authorized users anywhere and at any time.

The medical content may include different formats (text, images, forms). For reasons of interoperability, scalability and backwards compatibility the system should be open to the interconnection with different data sources. The medical content may come from different sources. This means that the source of the data might not only be hospitals or private clinics, but data from insurance organizations, other smaller units (such as hospital units), privately owned databases and any kind of potentially useful information source might be utilized.

Medical data that should be included in such a system can be summarized as follow:

- Report of incidents that were faced
- Photographs of incidents
- Background of patients
- Clinical picture
- Differential diagnosis
- Final diagnosis
- Development of disease
- References to international bibliography
- Treatment
- Statistical data (frequency of appearance of disease in the particular region of medical activity, results and development of incidents with regard to various type treatments, appearance of diseases regarding to age, profession, mortality etc).

Management and form of medical data will be proportional to the objective at which they are addressed, as:

**Doctors:** Complete briefing and exchange of opinions and information, including references to international bibliography, complete background of patients, photographs, clinical picture, differential diagnosis, comparison with other similar incidents from medical files, statistical data and ways of treatment.

**Patients:** General information are included with regard to diseases, precautions that should are received for the reject of deterioration of disease and information for the way of treatment of disease.

**Nursing personnel**: Briefing on the disease, rules of hospitalization of patient according to the requirements of disease, methods of precaution from likely infection or deterioration of disease.

**Auxiliary personnel of hospitals and clinics**: General information on the disease, ways of infection and precaution.

**Insurance companies**: Diagnosis and forecast of disease, statistical data.

Different “users” possess different requirements. This might include all the staff working in the health sector

(such as doctors, nurses, physicians, administrational staff) and other people such as researchers, insurance companies and the patients themselves. Effective and easy to use software tools are needed for different kinds of data to different kinds of users.

There must be standardized definition if a global system is to be launched. The SNOMED CT Core terminology provides a common language that enables a consistent way of capturing, sharing and aggregating health data across specialties and sites of care [17].

Finally, data security is an essential requirement in all health care applications. Developers of medical information systems should utilize the existing security development and evaluation methods to foresee as many of the technical and human factors that may endanger data security as possible and apply appropriate precautions [18].

## 5 System architecture

In order to achieve the above scope the following partial objectives are necessary:

- Creation of a user friendly database application for the medical units (easy data input or data retrieval even by a doctor with no IT background)
- Setting up of the equipment at the medical units (computers, databases, and classes/workshops given to the doctors about the use of the system)
- Using a software application tool for encoding medical data to topic map structure
- The same tool should offer enhanced navigational functionality within different representative forms
- as well as retrieval functionality, by providing enhanced features for employing rules and queries in TM.

We propose the following system architecture, which has arisen from the aforementioned functional requirements and the corresponding technical requirements (Figure 1).

Proposed TM-based system consist of four Levels:

- Resource Level contains all the medical data in dermatology domain. The medical content may include different formats (text, images, forms). For reasons of interoperability, scalability and backwards compatibility, the system should be open to the interconnection with different data sources.

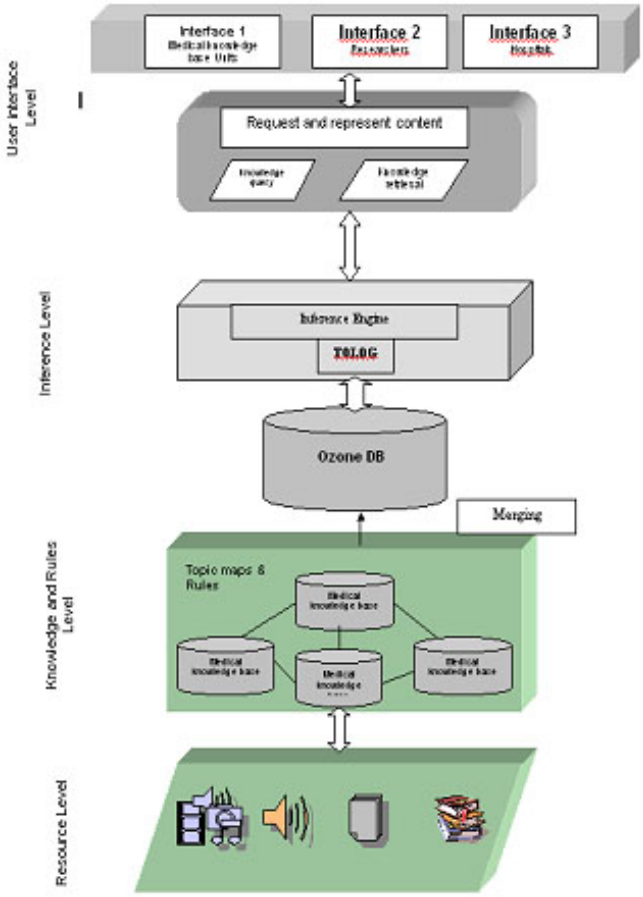


**Figure 1 : TM-based system architecture**

- Knowledge Level is the core of the system. It contains all topic maps that encode knowledge and semantically annotate resources. In this level, rules and queries are also preserved in order to be used for inference and querying.
- Inference Level is technically the most important, because it implements the semantic retrieval system. The inference engine reason over the knowledge level with respect to user requests, via TOLOG language.
- User-interface Layer implements the GUI Interface and enables user to communicate with the system.

Potential end users of the proposed system could be categorized to two different kinds. Those just browsing and using TM as a way to find information and those administrating a TM. For browsing there is a need of a software application that should be able to provide a good representation of topic maps structure and easy multi-modal access, navigation and retrieval of information. On the other hand, a TM administrator

needs an integrated software tool, which would enable efficient management of topic map.

A key factor in a successful appliance of topic maps is a software application tool, which provides an integrated environment for editing, viewing, managing and merging topic maps. For encoding medical data to topic map structure, an editor tool has been selected.

TM-Editor (TM-Ed) was designed and implemented aiming to reap TM capacities and eventually to support the use of TM technology [19,20]. TM-Ed (Figure 2) enables the programmatic creation, manipulation and medical data in an initiative way and additionally to query knowledge and retrieve the information they are looking for.

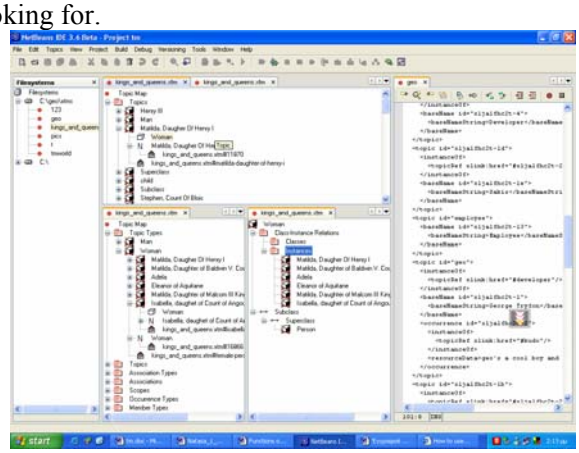

**Figure 2. Screenshot of TM-Ed**

TM-Ed supports creation and management of new topic maps or opening ones that already exist. It provides facilities for building topic map structures. In addition, TM-Ed enables powerful navigational functionality within different representative forms as well as retrieval functionality, by providing enhanced features for employing rules and queries in TM [20]. A key requirement in every database management system is an appropriate mechanism to query the data. In the same way as relational databases can be queried, a structured retrieval of Topic Map data, is offered using TOLOG language [21]. TM structure enables intelligent retrieval of information through the use of inference-based queries. TOLOG can query topic maps for topics of specific types, which participate in certain combinations of associations, and also supports inference-based rules.

Another key issue that it should be taken under consideration is the database that is going to be uses for data storage. TM4J supports two fully-featured backends - one based on storing the processed XTM file [22] in memory and one as persistent storage in an Ozone OODBMS back-end [23]. Medical data could be preserved in an XTM file [22]. But it has been realized that by adopting this solution, it would be difficult to have remote calls from client computers in order to update and manage the topic map, as it would be necessary to use ftp calls and change the txt file that holds the XTM files. Moreover, using a database technology could enforce the persistent of the system and enhance the potential size of medical records saved in a topic map repository.

Ozone (OODBMS)[23] is a fully featured, object-oriented database management system completely implemented in Java language and distributed under an open source license. Ozone is based on a central activation architecture. Objects never leave the server for any reason and client applications control their objects in the database via proxy objects. Administrators of the system are able to use all the features that TM-Ed offers, in order to create and manage topic map and store it to the ozone database. An easy to use graphical user interface will enable them to encode medical data to topic map structure.

By means of this tool, dermatologists can simultaneously locate, enter or update the corresponding medical data, which is subsequently stored in a TM-based database. On the other hand, users of the system will have limited access to TM-Ed features. TM-Ed provides the potential user of the proposed TM-based system the possibility to request and locate information, to navigate, browse and visualize and steps to be followed could be summarized as follows:

- Installation of the necessary software (short/medium term)
- Development of topic maps and rules
- Launch and use of the system (short/medium term)
- Evaluation of the system (short/medium term)

## 6 Conclusion

The scope of this work was to discuss the potential advantages of using topic maps in order to manage medical data, and consequently propose a functional TM-based framework for managing medical information, providing easy access to medical knowledge bases in order to enhance the level of communication and information provided on-line aimed at students, researchers, academics and practitioners in dermatology domain.

Topic Maps, as a representative standard in the field of the Semantic Web, seem to offer a feasible solution for managing medical data. Topic Maps can facilitate access to information effectively without changing the infrastructure or the databases, and support navigation and information retrieval in a more associative and intuitive way. Furthermore medical data could be more easily obtained, reused, exchanged and merged with information from other health organizations.

The future scope of this work is the developmemt and implementation of the proposed system, in order to produce a running application and not only a research prototype and consequently, to exploit and evaluate in practice the potential advantages that had been previously outlined, by incorporating topic map technology for medical data management.

*References*


1. Grótter, R., (2002 )Knowledge Media in Healthcare: Opportunities and Challenges University of St. Gallen, Switzerland, ISBN: 1-930708-13-0
2. Fierz,, W. (2002) "Information management driven by diagnostic patient data: right information for the right patient", Expert Review of Molecular Diagnostics 2(4),355-360
3. Fierz, W. Structured content and connectivity of medical information: the medical data Web, http://portal.acm.org/citation.cfm?id=779313&dl=ACM&coll=portal
4. ISO 13250, International Organization for Standardization, *ISO/IEC 13250, Information technology — SGML Applications — Topic Maps* (ISO, Geneva 2000)
5. Biezunsky, M., (1999). Topic Maps at a Glance, in: Proc. of XML Europe 99
6. Pepper, S. (2000) The TAO of Topic Maps. Proc. of XML Europe, Paris
7. Dr Jeremy Rogers, Medical Informatics Group, University of Manchester, www.cs.man.ac.uk/mig/education/tm.ppt
8. Easy topic maps homepage: http://easytopicmaps.com/index
9. Kamel Boulos MN, Roudsari AV, Carson ER. Towards a Semantic Medical Web: HealthCyberMap's Tool for Building an RDF Metadata Base of Health Information Resources Based on the Qualified Dublin Core Metadata Set. *Medical Science Monitor* 2002 Jul;8(7): MT124-36<URL: http://www.medscimonit.com/pub/vol_8/no_7/2615.pdf> [PubMed/MEDLINE]
10. Millar, D. (2001). Applying Topic Maps to the Classification of Health Interventions, XML Europe 2001, 21-25 May 2001, Internationales Congress Centrum (ICC),Berlin, Germany
11. Schweiger, Hoelzer, Rudolf, Rieger and Dudeck,2003 www.hb.se/bhs/webbredaktor/material/rolf/Examinationsuppgift_infsok.doc
12. De Bruijn and Martin (2002), "Getting to the (c)ore of knowledge: mining biomedical literature", http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Abstract&list_uids=12460628
13. Semantic layer creates an associative browsing system, http://www.diderottrack.nl/ en.articles. semantisch.html
14. Beier and Tesche, "Navigation and interaction in medical knowledge spaces using topic maps", Computer Assisted Radiology, 15th International Congress and Exhibition.
15. MESH subject, HOMEPAGE, www.nlm.nih.gov/mesh/meshhome.html
16. The HL7 Patient Record Proposal, http://www.hl7.org/
17. SNOMED Clinical Terms http://www.snomed.org/index.html
18. Niinimäki J., Savolainen M., Forsström J. J., Methodology for Security Development of an Electronic Prescription System, http://www.amia.org/pubs/symposia/D004712.PDF
19. Athanasios Hatzigaidas, Anastasia Papastergiou, George Tryfon , Zaharis Zaharias, "Topic Map Existing Tools: A Brief Review", ICTAMI 2004 (International Conference on Theory and Applications of Mathematics and Informatics) 16 –18 Σεπτεμβρίου 2004 Χαλκιδική, Full Paper
20. A. Hatzigaidas, A. Papastergiou, G. Tryfonand, A. Manitsaris: A Topic Map Editor and Navigation Tool. ICWI 2004: 359-365
21. Garshol, Lars Marius (2001). ,TOLOG, a topic map query language, http://www.ontopia.net/ [Moore, Graham, RDF and TopicMaps: An Exercise in Convergence, In *XML Europe 2001*, Berlin]
22. XTM 1.0 ,Pepper, S. and Moore, G (2001.): XML Topic Maps (XTM) 1.0 (TopicMaps.Org, http://www.topicmaps.org/xtm/1.0/
23. Ozone database homepage www.ozone-db.org/