A Multiple-Ontology Search Interface for Retrieval of Clinical Guidelines

Robert Moskovitch and Yuval Shahar

Medical Informatics Research Center,
Department of Information Systems Engineering, Ben Gurion University,
Beer Sheva, 84105, Israel
Email: robertmo@bgumail.bgu.ac.il

Abstract. A major problem in the effective use of clinical guidelines is fast and accurate access at the point of care. Thus, we are developing a digital electronic guideline library (DeGeL) and a set of tools for incremental conversion of free-text guidelines into increasingly machine-comprehensible representations, which support automated application. Even if guidelines are represented in electronic fashion, care providers need to be able to quickly retrieve the guidelines that best fit the clinical situation at hand. We describe **Vaidurya**[11], a search and retrieval engine that exploits the hybrid nature of guideline representation in the DeGeL architecture. Vaidurya can use not only free-text keywords, but also multiple semantic indices along which the guidelines are classified, and the mark up of guidelines in DeGeL, using the semantic roles of **one or more** guideline-representation **languages**. Vaidurya offers a wide variety of querying options, in order to enable different types of users to query the guideline library in a manner that is both efficient and user friendly. We describe the customizable query interface, in which each user can create their own personal query interface.

1. Introduction

Clinical Practice Guidelines (CPGs) are a powerful method for standardizing the quality of medical care [1]. CPGs are a set of schematic plans for management of patients who have a particular clinical condition (e.g., insulin-dependent diabetes). Unfortunately, most clinical guidelines are represented in free text, whether in paper or in an electronic format. Paper-based guidelines are relatively inaccessible to care providers at the point of care, while the free-text electronic format provides little support for automated retrieval of the guidelines potentially most applicable to the patient at hand, especially when the user is not an expert in the relevant clinical domain, and no support for automated application.

2. CPGs Representation Formats and Ontologies

Several clinical practice guideline representation formats, or ***ontologies*** (a knowledge-base schema that includes a set of concept types, their properties, and the relations among them), have been proposed in order to represent clinical guidelines in a structured or even machine-comprehensible fashion. Examples include ONCOCIN [2], EON [3], Asgaard (Asbru) [4], PRO*forma* [5], and GLIF [6]. Each ontology specializes at different kind of representation, for example GEM specializes in documentation representation while Asbru specializes in the automated executable representation [10].

3. The DeGeL Architecture

Each CPG in the DeGeL library eventually undergoes a *semantic mark up* process, in which the CPG's text is labeled by one of the available ontologies implemented in DeGeL (such as Asbru or GEM). The mark up is made by dragging a text segment from a source CPG to a relevant element in the chosen ontology. The text can then be further modified or extended. The markup is performed mainly by a domain expert. Domain experts are organized by groups, with a corresponding permission model for search and editing

[9]. The markup process's main goal is to assist the knowledge engineer in transforming the guideline into a formal machine-comprehensible format, such as full Asbru. These segments of the text representing the relevant context are good for accurate retrieval operations. The DeGeL library represents guidelines using a meta-ontology format [8]. Ontology independent elements, such as documentary details and semantic classification indices, are common to every guideline, regardless of the lower-level ontology, such as Asbru, used to present the guideline's details.
To classify guidelines, mainly for purposes of efficient retrieval, seven semantic axes are implemented in DeGeL. Each Axis represents a major clinical aspect. Axes include (1) symptoms and signs (e.g., hypertension), (2) diagnostic findings (e.g., blood cells count, electrocardiogram), (3) disorders (e.g., ischemic heart disease, malignant neoplasm), (4) treatment (e.g., antibiotic therapy , abdominal surgery), (5) body systems and regions, or a relevant CPG classification (6) guideline types (e.g., screening, prevention), and (7) guideline specialties (e.g., radiology, internal medicine). Each Axis is implemented as a hierarchical tree of sub axes. Each CPG is indexed along one or more semantic axes, such as Disorders (e.g., malignant skin melanoma), Guideline specialties (e.g., oncology), etc.

4. The Vaidurya Search Engine

A major current focus is on the retrieval of CPGs as a valuable tool to improve the adoption and integration of CPGs at the point of care, as part of the evidence based medicine approach. Electronic CPG repositories, such as the National Guideline Clearinghouse (NGC) [7] provide access to electronic guidelines in a free-text or semi-structured format.
Recently we introduced **Vaidurya**[11], a powerful search and retrieval tool that uses three kinds of search: (1) free-text search, using standard key terms; (2) *concept-based search*, which we call also *external search*, which uses a semantic –axes structure that indexes guidelines, as occur in the DeGeL library (and to some extent, in the NGC repository), and (3) *context-sensitive search*, which we call also *internal search*, which exploits the semantic markup performed on guidelines in the DeGeL library or any other markup tool. Internal search focuses on searching for key terms only in the context of the text that exists within the scope of a particular semantic knowledge role. (A natural implementation is searching for text within specific predefined XML elements, although other representation formats are potentially possible) and also elements that were represented as attribute values.

4.1. The Vaidurya Search and Retrieval Model

There are three kind of CPGs representations at the DeGeL library, Source guidelines (GLS), Marked up guidelines (GLM) and Asbru guidelines. Search is implemented, currently, for two representations. GLS – an entity representing guidelines at their Full Text representations as appeared at the origin where the guideline was downloaded or uploaded from, using the DeGeL source ontology. The Source ontology contains elements representing the metadata on the CPG and documentation details. GLM – an entity representing guidelines that were marked up, using the URUZ tool [10,11], represented by the DeGeL hybrid ontology, this ontology includes the markup ontology that the guidelines was marked up to and also metadata and documentation elements referring to the DeGeL markup documentation.

4.2. Concept Based Search in Vaidurya

Multiple digital libraries are indexed in a hierarchical structure; examples include the known web portal Yahoo and the NGC library [7]. These sites allow browsing through the categories in the hierarchical structure, starting from the root of the categories tree, and ending at the most specific category, located at the leaves. The NGC web site has two main Axes based on MeSH[12] : Disease/Condition and Treatment/Intervention. Thus, part of the Vaidurya search model includes optional specification of one or more semantic axes, or sub axes, and logical operators defining the relations between the axes. The *concept search query* is thus a collection of constraints represented by chosen sub axes and the logical relations between them: conjunction or disjunction.

4.3. Context Sensitive Search in Vaidurya

The context sensitive search exploits the markup process that a CPG goes through in URUZ. Each marked up CPG in DeGeL is represented within the target ontology selected for it by the URUZ user, implemented as an XML structure that depends on the structure of the target ontology. CPGs that were marked at other tools and other ontologies can be indexed into Vaidurya and retrieved as well.

4.4. Vaidurya Search Query

CPGs are marked up at different ontologies and classified along the DeGeL axes , thus Vaidurya was built as an infrastructure that allows the storage for **any kind** of guideline ontology or axis representation for a guideline. Vaidurya Search Query supports any kind of query at any ontology and any axes. A Search Query is a set of queries or constraints obtained on a group of chosen elements from a

specific ontology and sub axes from one or more axes, this combination describes the group or single searched guideline. Generally in Search and retrieval tools such a query would be under *Advanced Search* mode, in Vaidurya this is the search we encourage our users to perform – that kind of search specifying the requested CPGs should bring more accurate results. Each element in the *search query* is searched by its *search type* character, for example a *Free Text* element will be queried by keywords while an *Integer search type* element will be represented by a numeric constraint.

5. The Vaidurya Search Interface Problem

Using the Search Interface at each kind of search, GLS or GLM, offers the user a wide variety of extended attributes at the varying ontologies elements where each element can be queried, including classification queries. Such a detailed Search Query Interface enforces a long interface. While offering a very specific search it is very complicated and hard to use especially for users who aren't familiar with the DeGeL structure, ontologies and axes. Figure 1 shows a screenshot of a full search interface for the Source ontology (GLS), this interface includes only the Source ontology without the axes. The Marked Up Search offers a search in a longer and more detailed ontology using the DeGeL Hybrid Ontology and the marked up ontology (Asbru, GEM..), adding the axes makes it even more complicated.

Preliminary experiment made on a small group of varying users. We evaluated the Query Interface with expert physicians familiar with DeGeL, expert physicians who weren't familiar with DeGeL, physicians, Med-school students, nurses and regular users (with no clinical background). All the users reported that the Query Interface is too detailed and too long. We had learned that each kind of user had different interests in clinical guidelines and different information needs, also each kind of group has different level of clinical knowledge in general and different kind of experience with clinical practice guidelines particularly. Another factor is their familiarity and experience with DeGeL and CPGs ontologies. We noticed that each group used different kind of ontology elements when querying with correlation to their knowledge and experience in clinical guidelines, for example regular users would have search only using simple elements like Source Full Text and Title while expert physicians would query also the classifications axes and internal marked up contents like Entry Conditions.

All the group agreed that to perform a search they prefer an interface offering only a small subset of the ontology elements, elements that they will use most of the time, based on their knowledge. Thus we decided to implement a customized query interface in which a user can customize a query interface based on his personal search preferences.

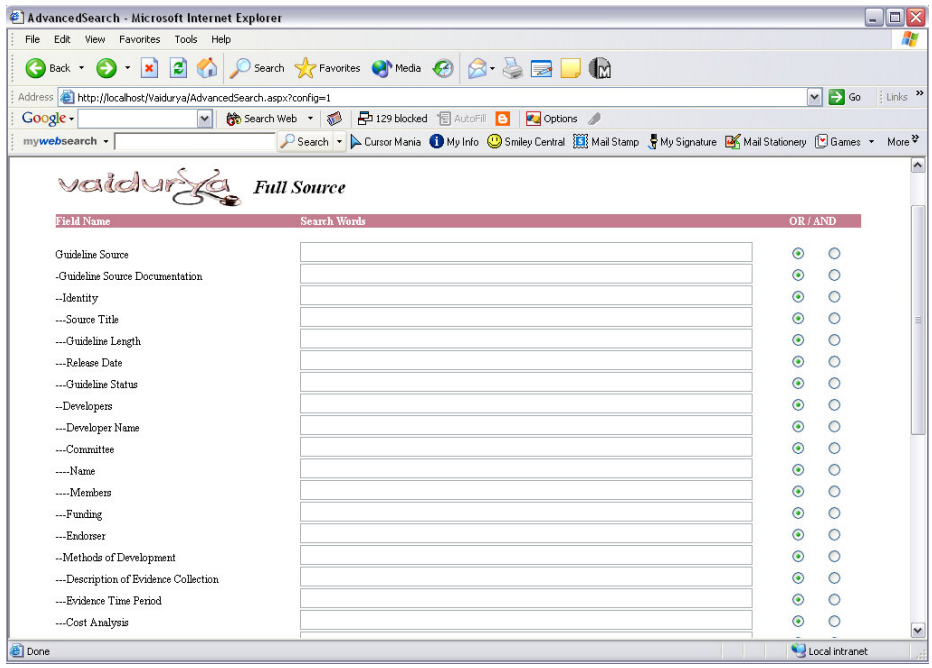

Figure 1 – Part of the Full Source Ontology Query Interface which is even shorter than the Markup Query Interface demonstrates the problem. This Figure shows element of free text, other type of elements can be queried as well

6. The Vaidurya Customized Query Interface (VCQI)

We are introducing here a Customized Query Interface (VCQI). The VCQI is saved by a given title and can be used in the future also by other users. Imagine a department in hospital, such department clinical crew would be interested probably at the same group of guidelines which is classified along one or more DeGeL axes representing their specific clinical interests, using this interface a user at this department will be able to create a VCQI customized to his colleagues needs, from that point each user at the department can use that VCQI and any other external users that it fits to their search needs.

6.1. The Vaidurya VCQI Designer

The design of a VCQI is made using a visual designer. Figure 2 shows the VCQI Designer which helps the user to create a VCQI in 3 simple steps in the same Web Form.

1. The user gives the VCQI a title, the VCQI title should describe the users it is created too.
2. The user chooses the kind of search Source Ontology (GLS) or Markup (GLM), the relevant ontology will be displayed according to kind of search chosen.
   Incase Markup (GLM) was chosen a list box will appear offering the user the markup ontologies implemented in DeGeL. Select the chosen elements by clicking on them and then clicking on the 'Add' button on right, the chosen elements will appear at the 'Chosen Field List' list box. The chosen elements order can be changed using the button on the right.
3. The user can set default values.(Optional)
   by clicking on the relevant element at the 'Chosen Field List' list box and entering a value at the Field Properties 'Value' text box.
   Clicking on the 'constant' check box won't allow changing the default value when performing the search.

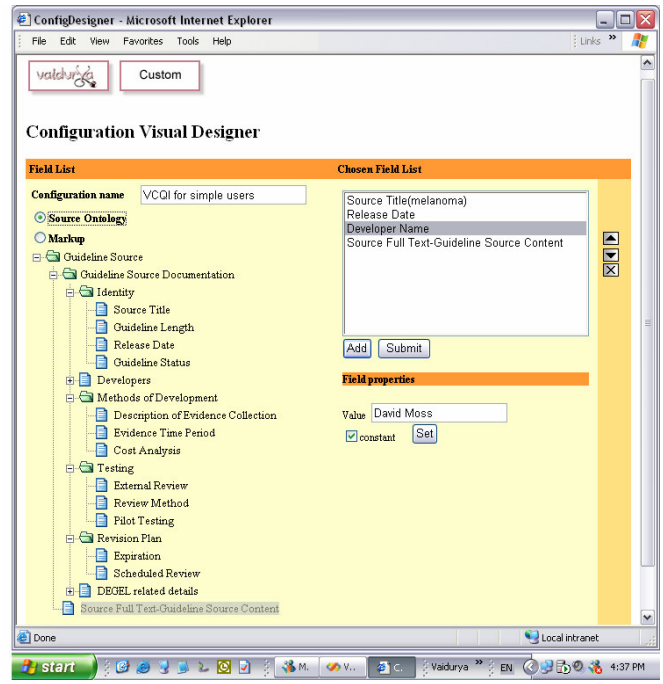

Figure 2. demonstrates a design of a VCQI called 'VCQI for simple users' created from the Source Ontology (GLS).

At Figure 2 the chosen fields are : 'Source Title', 'melanoma' set as default value, signed by parenthesis, 'Release Date', 'Developer name' 'David Moss' set as constant default value and signed by parenthesis and 'c' mentioning it is constant and 'Source Full Text-Guideline Source Content'. The VCQI design shown at Figure 2 will produce eventually the VCQI shown in Figure 3. The search results are displayed in **VisiGuide**[8], *VisiGuide* a browsing and visualization enables users to browse the set of returned results as well as visualize each returned CPG. *VisiGuide* organizes CPGs along the semantic axes distinguishing between the axes chosen at the query and these which were not but which where originally used to classify retrieved CPG.

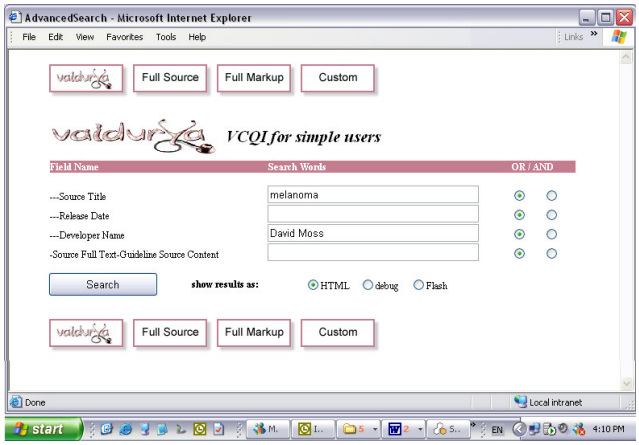

Figure 3 – the VCQI designed in Figure 2

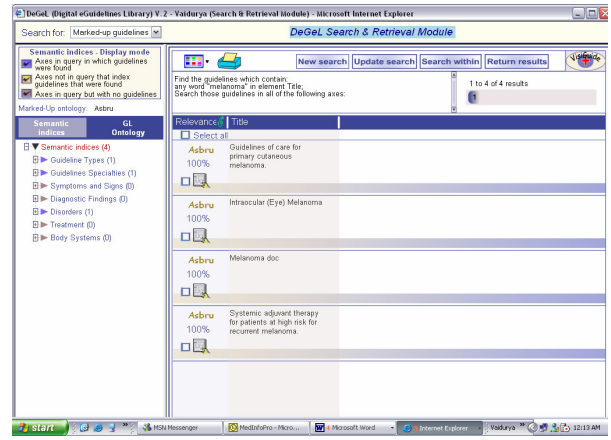

Figure 4 – Results presented at the VisiGuide

7. Future Work

We are working on a Template Based Query Interface where a user will design a form of a hybrid interface built from Free Text and the ontologies elements, such an interface will wrap the ontologies elements with text that will be self explanatory, thus the user will have to enter his requests at the proper place.

8. Acknowledgments

This research is supported partly by the NIH/NLM award LM-06806. We thank our colleagues at Standford University Drs. Mary Goldstein, Susana Martins, Soroka Hospital ,Prof Eitan Lunenfeld, were extremely helpful in assessing and developing Vaidurya's query interfaces and E&C Medical Intelligence Inc. for their help in building a test collection. We would like to thank also Lior Bar-on and Chen Sason, undergraduate students who made their senior project at the Vaidurya research at the Department of Information Systems and all the physicians and users that gave their important comments.

9. References

[1] Grimshaw, J.M. and Russel, I.T. (1993). Effect of clinical guidelines on medical practice: A systematic review of rigorous evaluations. Lancet, 342: 1317–1322.

[2] Tu, S.W., Kahn, M.G., Musen, M.A., Ferguson, J.C., Shortliffe, E.H., and Fagan, L.M. (1989). Episodic Skeletal-plan refinement on temporal data. *Communications of ACM* 32: 1439–1455.

[3] Musen, M.A., Tu, S.W., Das, A.K., and Shahar, Y. (1996). EON: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Infor-mation Association 3*(6): 367–388.

[4] Shahar, Y., Miksch, S., and Johnson, P. (1998). The Asgaard project: A task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artificial Intelligence in Medicine,* 14: 29-51.

[5] Fox, J., Johns, N., and Rahmanzadeh, A. (1998). Disseminating medical Knowledge: the PROforma approach. *Artificial Intelligence in Medicine,* 14: 157-181.

[6] Peleg M, Boxwala A. A., Omolola O., Zeng Q., Tu, S.W, Lacson R., Bernstam, E., Ash, N., Mork, P., Ohno-Machado, L., Shortliffe, E.H., and Greenes, R.A. (2000). GLIF3: The Evolution of a Guideline Representation Format in *Overhage M.J., ed., Proceedings of the 2000 AMIA Annual Symposium (Los Angeles, CA, 2000),* Hanley & Belfus, Philadelphia.

[7] National Guideline Clearinghouse, www.ngc.org

[8] Shahar Y, Young O., Shalom E., Mayaffit A., Moskovitch R., Hessing A., and Galperin M. (2003a). DEGEL: A hybrid, multiple-ontology framework for specification and retrieval of clinical guidelines. *Proceedings of the 9th Conference on Artificial Intelligence in Medicine—Europe (AIME) '03,* Protaras, Cyprus.

[9] Shahar Y, Shalom E., Mayaffit A., Young O., Galperin M., Martins S.B., and Goldstein, M.K. (2003b). A distributed, collaborative, structuring model for a clinical-guideline digital-library. *Proceedings of the 2003 AMIA Annual Fall Symposium*, Washington, DC.

[10]Mor Peleg, Samson Tu, Jonathan Bury, Paulo Ciccarese, John Fox, Robert A. Greenes, Richard Hall, Peter D. Johnson, Neill Jones, Anand Kumar, Silvia Miksch, Silvana Quaglini, Andreas Seyfang, Edward H. Shortliffe, and Mario Stefanelli, "Comparing Computer-interpretable Guideline Models: A Case-study Approach". Journal of the American Medical Informatics Association, Vol. 10, No. 1, Jan-Feb 2003, pp. 52-68.

[11] Robert Moskovitch, B.Sc., Alon Hessing, B.sc., and Yuval Shahar, M.D., Ph.D, Vaidurya – A Concept-Based, Context-Sensitive Search Engine For Clinical Guidelines

[12]Shiffman RN, Karras BT, Agrawal A, Chen R, Marenco L, Nath S. GEM: a proposal for a more comprehensive guideline document model using XML. J Am Med Inform Assoc 2000; 7(5):488-498