Adaptive Searching Mechanisms for a Cardiology Information Retrieval System

S Chessa[1], E de la Vega[1], C Vera[1], MT Arredondo[2],
M García[1], A Blanco[1], R de las Heras[1]

[1]SIEMENS, SA, Madrid, Spain
[2]Universidad Politecnica de Madrid, Madrid, Spain

Abstract

Medical knowledge is inherently complex and uncertain. Medical experts may provide different interpretations for symptoms since all of them also depend on a given context and most of them are established by statistical utilization. It is necessary a whole knowledge baggage in order to understand and take care of patients with cardiovascular diseases adequately. For this reason, the NOESIS system, an Integrated Project (IP) partially funded by the EU, has been conceived aiming at being a valuable instrument for cardiologic medical information retrieval from heterogeneous, distributed medical databases that mediates medical decision of critical health conditions. The adaptation features that are supported in the NOESIS platform generate a personalized searching process for the users depending on the information stored in their personal profiles.

The objective of this paper is the description of a new adaptive searching mechanism developed using innovative technologies in order to obtain, use and manipulate medical information.

1. Introduction

Medical knowledge in the cardiology domain is inherently complex. Everyday medical professionals require quick and easy guidance to get through complex clinical cases and to answer complicated scientific questions. They also need to have acute information that assists them in the difficult procedure of patient's diagnosis.

The NOESIS [1] project, an Integrated Project partially funded by the EU, aims to help health professionals in answering their everyday clinical and scientific questions, providing a solution that combines both clinical and scientific data according to the user profile. The project objective is to develop an intelligent environment that enables ubiquitous management of citizens' health status and assists health professionals in coping with some major challenges like risk management and the integration into clinical practice of advances in evidence based medicine.

NOESIS is oriented to the cardiovascular medical field. Its primary stakeholders are Cardiologists, Cardiothoracic surgeons and General Practitioners. Additionally, other health care professionals, like nurses, medical researchers and medical students, could benefit from the use of the platform for scientific and training purposes.

NOESIS is the first integrated system that manages patients' clinical data and scientific literature, in parallel, relevant to the case under observation. Moreover, the decision support framework, provided by NOESIS, allows a preliminary identification of the patient diagnosis as well as proposes a list of the appropriated treatment. Furthermore, the clinician is advised on the possible complications caused by the disease or the treatment, the so called adverse effects.

All these outstanding NOESIS capacities are based on a unique system design supported by knowledge management and uploading capabilities, information retrieval functionalities and a novel decision support mechanism.

This paper presents the adaptive searching mechanism as the core of the NOESIS information retrieval system.

2. Methods

The searching mechanisms developed are based in the Adaptive Hypermedia (AH) Systems [2]. The AH is a recently research branch in the Hypertext area and User Modeling [2]. The AH Systems adapt the presentation, organization and management of the information in an autonomous way for every user. It is based on user profiles and context-aware applications.

The design of this adaptive searching process has been divided in two central elements: Profiling System and Search Engine.

User profiling in the NOESIS system permits the storage of user related information like demographics, preferences and personal interests in medical areas, in order to achieve highly personalised system behaviour. It is based on the idea of automated detection of the user's

behaviour.

The main profiling procedure consists of the segmentation of the suitable roles used to identify the users. Over this segmentation, different layers of user's characterization are presented to refine the user personal definition.

The dynamic profiling is achieved by means of two techniques: 1) using advanced methods of capturing the user's activity; and 2) discovering the information items that attract the user's attention by means of attentive agents. Acoustical information, cameras supporting visual analysis of the user's gaze and other means like interactions with software or devices, constitute the major inputs that attentive agents can capture. This way the user dynamic profile is acquired dynamically and automatically with minimal user input.

More specifically, the profiling mechanism initially retrieves information from each user's session, as the web pages visited by the users, and particularly, the information items within these pages that have been identified by the attentive agents as relevant to them. Using these two information sources, the user is comprehensively described into the system and the information retrieval functionality adapts the search mechanisms to each user preferences. User session patterns are created using this information and sent to the Profiling System for storage in the users' dynamic profile. These profiles consist in a network of terms or URLs that are generated, if the user does not have the dynamic profile, or dynamically updated whenever a new pattern arrives.
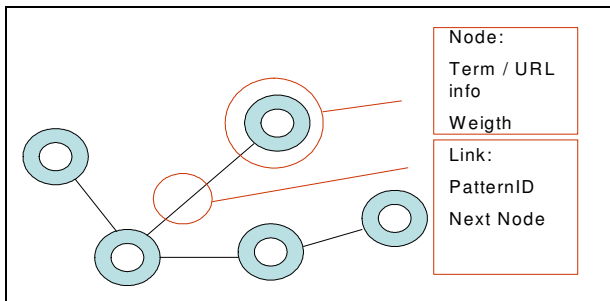

Figure 1. Elements of the dynamic networks

The user dynamic profile is a network of terms or URLs that contains the information of his sessions and permits the system to personalise his queries (Figure 2).
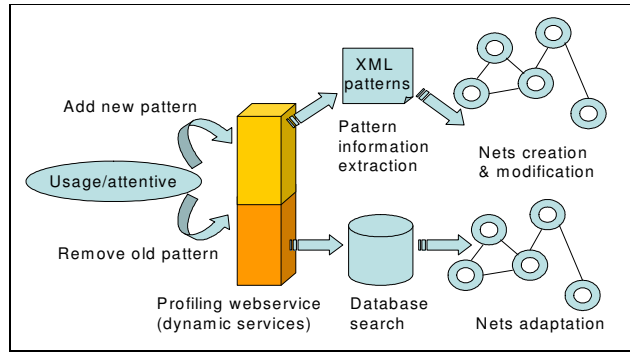

Figure 2. Queries personalization

Whenever a pattern is estimated as outdated, it is removed from the system. Additionally the system is able to expand the user queries by means of the Search Engine. The algorithm implemented for expanding the user's queries is mainly based on the following:

- Find whether the list of terms belongs to a single pattern or not.

- Complete the query with the pattern.

- Try to find a "logical" path between terms of the network.

- Maximize the weight.
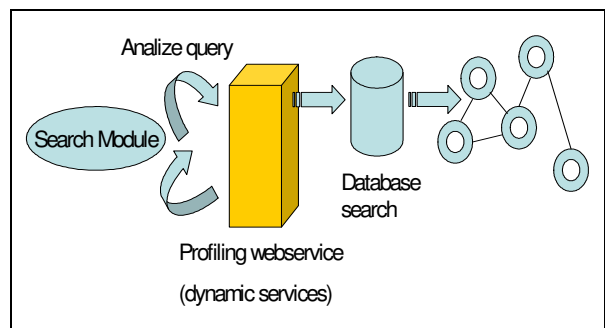
- Minimize the jumps.

- Avoid loops.

Figure 3. Dynamic profile query expansion

The technologies and the reason for their use in the development of the profiling system, after having tested some alternatives, are the following:

- XML: for its portability.

- XML native database eXist [3]: for its facilities to manage XML files.

- Web Services [4]: for both their accessibility and integration facility.

- JAVA [5]: Java is a portable programming language ideal for the development of network applications.

- XML parser: JDom [6] and SAX [7].

The searching process that takes place within the NOESIS platform is governed by the Search Engine. The profiling utilizes the user data to build its own dynamic net, personalized according to the user preferences.

The Search Engine has the purpose of improving the queries introduced by the users of the platform, creating an advanced query that will be the base for the personalization of the results presentation. This feature aims at the management of the retrieved results in response to the previous query delivery. The results are ranked and divided calculating the relevance to the user. This module allows the integration between the information retrieval system and the graphic user interface. The Search Engine is the component that exchanges the information relevant to the user with the other components of the platform. As in the profiling system, the technologies used in the development of this module are Java Applets [5] and Web Services technologies [4].

3. Results

The System developed for the Profiling Module is a XML multi-platform access information server. It has a multi-access structure implemented with XML to be managed through the Internet.

The general configuration of the system consists of three modules:

- A Storage Module: Responsible for the storage of all the information related to a user.

- A Management Module: Responsible for the management of the information stored in the profiling system. It basically manages the profiles upgrading, the reports analysis and the access to

specific user information.

- A Communication Module: Responsible for the transmission and reception of all the information stored in the system.
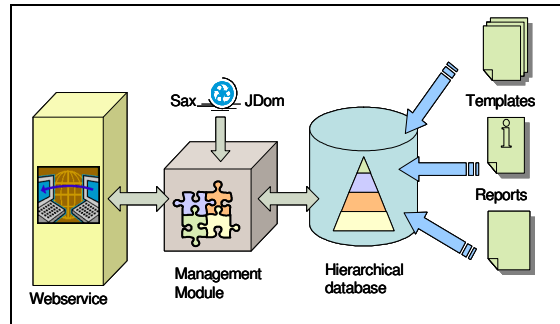

Figure 4. General structure of the Profiling system

The Search Engine has been divided in two services according to the type of data that the module manages in the initial query. These are the textual querying service and the images querying service. Both services are mainly determined by the same procedure. However, the way in which the information included in the queries is processed and the modules involved in these operations is different.

The black box design of the module can be depicted (Figure 5) as a double pipe structure with basic inputs and advance outputs.
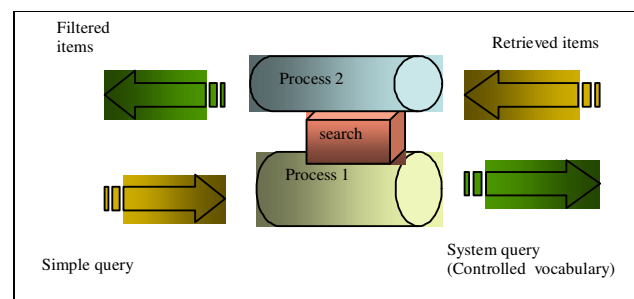

Figure 5. Search Module input/output

This search general process is divided in two different mechanisms:

- Query Improvement: It is the mechanism that generates a new query using as input the initial one, after a term selection made by the Natural Language Processing Module (NLP) [1], adding new features that will depend on the content of the query and the

User information.

- Information's relevance calculation: By means of this mechanism, along with the user profiles and the indexing information of the documents, the Search Engine will classify the results ordering them, depending on their relevance.

The query is expanded after the first intervention of the Profiling Module, and then, all the static information that the user stores within its profile can be used to manage the results already obtained.

The simple results (links) obtained from the query are complemented with additional features from other modules of the NOESIS system:

- Lists of results depending on the text document or image sources.

- Several buttons to communicate with other tools to annotate the information, to obtain recommended links and to mark a specific document or author.

- Access to some relevant metadata, like title, author, language, publication's date, etc.

- Possibility of implement new search processes using the information retrieved as origin for new queries.

Additionally, the information displayed on the screen and the set of functionalities offered to each user depends on the user type and his access rights.

4. Discussion and conclusions

The system has not been clinically tested yet. Up to now, all the tools have been tested during the project integration phase that has already proved the benefits of using web services technologies to integrate tools developed under different platforms. It has been also shown the efficiency in the tasks related to the profiling and searching processes described in this article.

The main conclusions derived from this research and the development done within this project can be summarized in the following points:

- The need of updated cardiology information from studies or patients clinical records to elaborate consistent assumptions and treatments regarding cardiac affections is a must.

- Save time gathering all this information is a way to improve the time per patient and consequently to improve the patient treatment and the quality of the attention. This is translated into the elimination of all the unnecessary information associated to a searching process and the adaptation of the information searched for a single user.

- Personalisation of information means to use profiling techniques and information to modify basic search processes and to create an adaptive search mechanism. This means the creation of a virtual user description based on user information and actions inside a system with the purpose of filtering a search process according this information.

Acknowledgements

References

[1] NOESIS project "Description of Work", IST 507960. European Commission Framework Programme.
[2] Hubertus Hohl, Heinz-Dieter B–cker, and Rul Gunzenh"user. "HYPADAPTER: An Adaptive Hypertext System for Exploratory Learning and Programming, User Modeling and User-Adapted Interaction, 6(2-3), 131-155, 1996.
[3] eXist, 2005: Open Source Native XML Database. http://exist.sourceforge.net/index.html
[4] Wolfgang Meier: Web, Web-Services, and Database Systems. NODe 2002 Web- and Database-Related Workshops - eXist: An Open Source Native XML Database.
[5] SUN Microsystem: http://www.sun.com/java.
[6] Parses, manipulates, and outputs XML using standard Java constructs. API is similar to DOM, but easier to use. [Open Source, BSD-like]: www.jdom.org/
[7] Rusty Harold, Elliotte, 2001. "Processing XML with Java: A Guide to SAX, DOM, JDOM, JAXP, and TrAX". Also available on the Web at the direction: http://www.cafeconleche.org/books/xmljava/

Address for correspondence

Salvatore Chessa

SIEMENS, S.A., HS

Ronda de Europa, 5 - 28760 –Tres Cantos –Madrid, Spain

noesis.es@siemens.com