# MEDFINDER

## *Using Semantic Web, Web 2.0 and Geolocation Methods to Develop a Decision Support System to Locate Doctors*

Alejandro Rodríguez, Jesús Fernandez, Enrique Jimenez, Myriam Mencke
Mateusz Radzimski, Juan Miguel Gómez
*Department of Computer Science, Universidad Carlos III de Madrid*
*Avenida de la Universidad 30, 28911, Leganés, Madrid, Spain*

Giner Alor-Hernandez, Rubén Posada-Gómez
*Division of Research and Postgraduate Studies, Instituto Tecnológico de Orizaba, Mexico*

Keywords: Differential Diagnosis, e-Science, Ontologies, Reasoning, Web 2.0, Semantic Web, Geolocation.

Abstract: Currently the introduction of new technologies for efficient medical diagnosis and treatment is having a profound social and organizational impact, initiating the need to exploit the latent potential of novel methods. A requirement has emerged to maximally exploit the fusion of new technologies for more efficient patient care than is presently available. This scenario motivated the objective of the current paper, combining an existing medical diagnosis system which uses semantic technology and probabilistic techniques with Web 2.0 and geolocation methods to develop a system to locate the most appropriate doctor for a patient. Results of an initial prototype implementation are promising.



## 1 INTRODUCTION

In today's IT landscape, a multitude of systems exist which permit a user to perform medical queries based on a series of factors, such as symptoms, laboratory tests, and medicines, among others. Computations made based on these factors are generally capable of providing a relatively precise diagnosis using a formal reasoning process, and in some cases, the diagnostic may be accompanied by an associated probability of the diagnosis outcome. However, until now, the possibilities which such a system could offer have not been fully exploited. A series of technologies could be added to enhance these types of systems in order to provide much more complex and complete functions. The objective of this paper is to describe a framework for the implementation of such a system, by combining an existing system which performs medical diagnosis using the methods just described with new technologies. This forms the basis for the construction of an architecture which relates both technologies.

The initial application used is an expert system which enables doctors to perform medical diagnosis based on determined parameters, which uses Semantic Web technologies for this purpose (Rodriguez, 2008).

An additional component will be introduced to the system, where once the diagnosis has been determined and the diagnosed illness has been classified as the most probable, and the illness has been verified as correct by a medical expert, a database of medical experts will be consulted. The objective of this step is to locate a series of experts which fulfil certain criteria in relation to the illness diagnosed. This database will interact as a positioning system, which among various other functions, is able to obtain the shortest distance between the expert and the patient, thereby determining the optimal route for the patient to reach the location of the expert. The system is used to realize this process in the fastest and least disruptive form available.



RodrÃguez A., Fernandez J., Jimenez E., Mencke M., Radzimski M., GÃşmez J., Alor-Hernandez G. and Posada-GÃşmez R.
MEDFINDER - Using Semantic Web, Web 2.0 and Geolocation Methods to Develop a Decision Support System to Locate Doctors.
DOI: 10.5220/0001831402880293
In *Proceedings of the Fifth International Conference on Web Information Systems and Technologies (WEBIST 2009)*, page
ISBN: 978-989-8111-81-4

In the motivating scenario section we will shortly explain the main reason for this research. In the MEDFINDER section, we will deeply talk about the system, explaining its architecture: the ODDx system, that is the system diagnosis provider; the Sorter subsystem, which will categorize the diseases into speciality groups; the Expert database, that contains all the information about medical experts that can deal with the speciality specified by the Sorter; the Web 2.0 Feedback system, that allows to change the expert punctuation; and the Geolocation system, which will locate the nearest doctor and design the best route to meet him. In the related work section we will introduce some systems that use similar techniques and, in the final chapters, we will talk about the findings and conclusions we came to, and the future work needed. Finally, in the acknowledgements section we will mention the main projects in which we are currently involved, and that allow us to make a better job.

# 2 MOTIVATING SCENARIO

The choice of a medical professional by an expert system to cure or treat a specific disease within a particular field is not a trivial subject. Generally, a person in urgent need of medical treatment arrives at emergency services where a qualified medical professional is assigned to attend to the case, however, on certain occasions, the patient is able to select the doctor due to circumstances such as having access to private healthcare. One of the objectives of the current system is to construct a link between the patient and doctor in such a form that the patient can access the professional aptitudes of diverse doctors as a function of his/her necessities, taking into account particular variables, such as the distance which separates both.

A different scenario with rather interesting application is the case of the contracting of specialists by clinics, private or public. In this case, the system would be a modification of the system previously described, where the component which makes the diagnostic would essentially be dispensed with, given that in large part it is no longer necessary. In this scenario, the person responsible for contracting specialists already knows the branch of medicine in which each specialist is qualified, therefore the diagnostic component is no longer fundamental. Once the specialization is known, the rest of the process is carried out similarly to that detailed before, given that the system is responsible for searching for the specialist or expert which best fits the data provided.

Additionally, this technology could be incorporated into many further scenarios unrelated to medicine, but where the entire system could be implemented with some modifications. For example, implementing the system for the detection of mechanical failures in a garage, where the software process is responsible for detecting the fault or breakdown, and the system locates the most appropriate expert for repairing the said defect. Additionally, in the case of replacement of parts, the closest part sellers with the best quality and price could be located by means of the system.

# 3 MEDFINDER

Given the proposed scenarios and the problems to resolve, the subsequent step is to describe the approximation for the solution of the problems. This solution is based on previous work described and developed in another existing solution, in particular, (Rodriguez, 2007), which is an existing medical differential diagnosis system based on an ontology, logical inference and probabilistic techniques. Additionally, in what follows, the remainder of the subsystems and parts involved in the elaboration of the framework will be explained.

## 3.1 System Architecture

The current structure of MedFinder is based on the construction of various components which interchange certain information among each other. The conjunction of these systems permits that the entire set of components functions correctly, and the necessary data can be obtained to achieve the desired outcome.

A diagram is shown below which displays the general structure of the system and the relationships between its elements. In the subsequent section, the internal functioning of each of the elements will be described, as well as their behavior with regard to the information interchanged with the rest of the components in the system. As additional detail for the understanding of the diagram, consider the following elements:

***[I]***: Represents the information which one subsystem sends to another.

***[A]***: Represents the action which one component performs on another component. This may include the transfer of information or not.

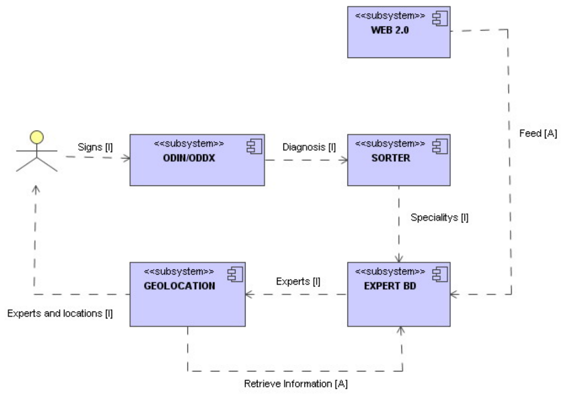


Figure 1: System Architecture.

### 3.1.1 ODDX

The first component of the system with which the user interacts is ODDX, as can be observed in the diagram. Viewed globally, it is in fact the initial and only part which involves user interaction. It is also this component which is the front end of the expert system that realizes the diagnostic for the user. The fundamental difference of this system compared with the previous system is that the architecture is extended considerably when compared with the existing architecture.

ODDX is an expert system which was capable of inferring medical diagnoses based on a given number of parameters, without any further functions. However, in the new architecture, this component is no more than the beginning of an entire process which not only diagnoses the illness, but also provides the user with much more information related to the diagnosis.

Therefore, the architecture of ODDX has not changed, however, a number of variations have been made.

The basic components of ODDX are the following:

- **Probabilistic.** The probabilistic system is the system that is responsible for managing and/or calculating the probabilities of every diagnostic inferred. Every disease that is diagnosed (one or more) from a group of indications has its own probability of happening.
- **Data Loading.** This is the engine which performs data loading from the ontology. It employs the Jena API to read the ontology file, which is an open source Java programming environment for Semantic Web applications, and supports the use of various languages, such as RDF, RDFS, OWL and SPARQL.
- **Combination System.** The combination system computes all of the diagnostic combinations possible which may be the result of the interaction of drugs. Basically it is a method which allows, given a patient with a group of indications and associated drugs, the calculation of the possible interactions which may be caused by drugs.
- **Inference.** The inference engine is the main engine of application, because it is a principal constituent of the system which really enables the diagnoses to be made. This engine requires access to the knowledge base with the diseases, symptoms, etc., and at the same time needs access to the knowledge base that contains inference rules.
- **Search.** This component takes the form of a search engine, which realizes SPARQL queries to the ontology to consult all of the data stored it. This permits fast access to all of the data stored within the ontology.

### 3.1.2 Sorter

The second component involved in the entire process has been entitled Sorter. This is essentially a system which is capable of generalizing or specifying. It receives an illness as input, which is assigned an ICD code (International Classification of Diseases by the World Health Organization), and by means of this code, the system is capable of classifying this illness within one of the multiple specialisation of the ICD's own classifications.

The implementation of this part of the system may be viewed as both a client and server at the same time. On the one side it acts as a server to receive information (the ICD code of the illness), which is provided by the client, and at the same time behaves as a client at the moment it sends its output (the specialization) to the subsequent subsystem.

Whether the use of this subsystem is dependent upon the final implementation is optional. This is because ODDX in its current state does not implement this classification component, which would be easily implementable and in fact a more efficient solution, given that this service could be discarded, thus speeding up the final solution.

Regarding the technologies used to implement this facet, the principal constituent proposed is an ontology, which allows the classification of the illness and its categorization within a speciality. For the present framework, it is proposed that the system uses the same ontology which the ODDX system uses, given that this ontology classifies the illnesses in distinct superclasses. In a specific level of these superclasses, they represent precisely the specialization. Due to the design of the ontology, the classification of the illness is thus a rather simple process.

### 3.1.3 Expert Database

The objective of the database of experts is precisely to provide a database which contains the most relevant data relating to an expert. The database contains a series of data, which are listed below:

- **Latitude Coordinate.** Coordinate to locate the expert, principal location.
- **Longitude Coordinate:** Coordinate to locate the expert, principal location
- **Expert Specialization Code.** Code(s) representing the distinct specializations of the expert.
- **Possible Alternative Coordinates.** Reference to possible alternative codes to locate the expert. These coordinates could also be linked to a particular specialization.
- **Patient Scores.** Scores which the patients can assign to the expert once having received the treatment. Further details will be provided below.
- **Expert Scores.** Scores of other experts in the same matter.
- **Other Data.** Other series of data such as name, usual address, telephone, email, treatment times, and observations, among others.

A diagram of the Entity-Relations in the database is provided below, followed by a description of the relations.

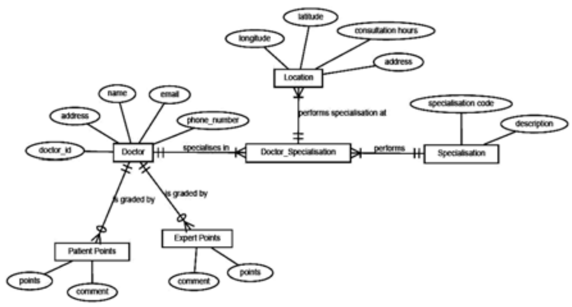


Figure 2: Entity-Relation Schematic Representation.

The ER Diagram description is provided above.

Entities:

**Doctor.** The main entity containing information about medicine doctors, such as name, address, contact and other data.

**Patient Points.** Patient Point entity represent each mark given by patient to score a doctor.

**Expert Points.** Export Point entity, similar to Patient Points, is a representation of scores given by doctors/experts to other doctors.

**Location.** Exact location of the doctor's/expert's office. Contains the exact geolocation data (longitude and latitude), address and hours of consultation.

**Specialization.** Represents all available experts' specializations, which are described by unified specialization code.

**Doctor_Specialization.** Additional entity which represent each doctor's/expert's specialization. As one expert may perform more than one specialization, each in different location. This entity uses foreign keys from Doctor and Specialization entities.

Relations:

Specializes_in (related entities: Doctor, Doctor_ Specialization) – an one-to-many relation describing doctor's specializations. Since some doctors/experts may perform more than one specialization, each one is represented by a distinct entry in Doctor_Specialization entity.

Performs_specialization_at (related entities: Location, Doctor_Specialization) – an one-to-many relation defining the locations (office, clinic) where certain doctor's specialization is performed.

Performs (related entities: Specialization, Doctor_Specialization) – an one-to-many relation describing the specializations that are performed by doctors/experts. It allows to map exact specializations to actual doctors' specializations.

Is_graded_by (related entities: Patient_Points, Doctor) – an one-to-many relation between doctor's scores and certain doctor. Each doctor is given many scores from patients. Relation between Expert_Points and Doctor is analogical.

Additionally, the system described above receives input from another subsystem, namely Web 2.0. Further details of this system will be provided below.

### 3.1.4 Web 2.0 Feedback System

The aim of the Web 2.0 system is to insert or update data relating to experts within the system. This system also permits the modification of the scores assigned to distinct experts, thereby facilitating the users of the system in the decision to choose one expert or another. This scoring may be divided into two distinct parts:

**Scores of Previous Patients.** These scores are based on the scores which patients that have been previously treated by a particular expert may assign to this expert.

**Scoring of other Experts.** This system is based on permitting other experts to assign scores to their colleagues, thereby improving the global vision of these by structuring it such that their knowledge is considered important and other professionals in the sector value their work.

### 3.1.5 Geolocation System

The geolocation system is the final constituent of the system. It is fundamentally a system which communicates with the database of experts to obtain certain data required to locate the expert. The essential data are the localization, longitude and latitude coordinates, however, other data may be required. The system should calculate a **real** route (realizable by car or on foot, not the distance between two points), which exists between the patient and expert(s).

There are already a number of tools in existence which achieve this aim, however, in the current framework, particular importance was given to the aspects high potency, ease of management and reliability. These features were selected having in mind that the system not only had to establish the physical location of the specialist, but also had to be capable of establishing a route between the patient's current position and the expert.

The API which permits these characteristics is proprietary of Google, *Google Maps*. This framework provides a library which allows the creation of maps of a determined location, establishing distinct tags, controls, and determining routes between two points. Another feature in favour of selecting this platform is that it allows the drawing up of a physical, political and hybrid view between both places (Bühler, 2006). This implies that the user has more visual references for the place at which he wants to arrive, and additionally, for intermediate points which he must pass to reach his destination (Muller, 2004).

However, the key strength of the framework is that it allows the calculation of the shortest route between two points with the simple process of introducing the coordinates of the user and the expert, which in this case can be obtained from ExpertDB. Additionally, a detailed description of the route is provided indicating specifically the route which the user should take, which turns he should take, and the public transport available in each of the streets on the route (Grabler, 2008). It also offers further information about the total length of the journey, and the approximate time it takes to complete the entire journey.

Another very interesting aspect is that this API offers the possibility to select how the journey should be realized, on foot or by car, displaying distinct alternatives according to the option chosen. For example, if the user wants to undertake the journey on foot, the system does not process or take into account restrictions or prohibited ways which would be encountered by a car.

## 4 RELATED WORK

In the domain of medical diagnosis systems, a myriad of approaches exist, which are comprised of various algorithmic techniques for automatic diagnosis that have been tested in research, as well as actual systems currently available for use. Approaches in research which apply the use of combined techniques such as the current one include neuro-fuzzy methods (Noy, 2005), the application of genetic algorithms (GAs) for rule selection (Ishibuchi, 1999), or the unification of genetic algorithms with fuzzy clustering techniques (Roubos and Setnes, 2000).

Apart from the systems described above, more specialized systems are available, for example, those in which clustering techniques are used in the detection of epidemics (Cardoso, 1999), decision and action support systems in relation to illnesses according to region (Gosselin and Lebel, 2005), and systems which aid the differential diagnostic in "erythemato-squamous" form, incorporating classification algorithms for trees and other similar methods (Güvenir, and Emeksizb, 2000). Possibly the most interesting and most similar application in relation to the current system is that developed by (Faria, Fernandes and Perdigoto, 2008), which propose a specific type of monitoring for old people or people with mental illnesses such as Alzheimers, which using a mobile system ensures that the person can be geographically located in any place in the case that he needs help, or physically reached by someone else by means of a button which is designed to deal with such cases.

Besides these types of geolocation systems, currently there is a scare number of systems in medicine which provide this type of functionality. These types of systems are in fact more advanced in other domains, such as tourism. There are localization systems for destinations or touristic monuments to create personalized information for travelers (Kalczyński, 2001), and of course the well-known tools such as Google, Google Earth and Google Maps, as previously detailed. This makes the incorporation of such tools in medical applications a challenging future field.

# 5 CONCLUSIONS AND FUTURE RESEARCH

The current paper has outlined a system for recommendation of medical diagnostics which uses an ontology, logical inference, and in particular, introduces the feature of detection of medication that may interact with other medicines and other signs like allergies. An additional step which has been taken is to provide a framework for geolocation of doctors, an architecture which is currently under development.

A further research step which is envisaged is to also allow users to introduce other complementary parameters into the system to construct a more specific system that allows the user to obtain more precise results.

# ACKNOWLEDGEMENTS

This work is supported by the Spanish Ministry of Industry, Tourism, and Commerce under the project SONAR (TSI-340000-2007-212), GODO2 (TSI-020100-2008-564) and SONAR2 (TSI-020100-2008-665), under the PIBES project of the Spanish Committee of Education & Science (TEC2006-12365-C02-01) and the MID-CBR project of the Spanish Committee of Education & Science (TIN2006-15140-C03-02).

# REFERENCES

Bühler E.R., Localización Geoespacial con Google Maps. *Solo Programadores 143*, pp 36-41, 2006

Faria S., Fernandes T.R. and F.S. Perdigoto., Mobile web server for elderly people monitoring. *ISCE 2008, IEEE International Symposium on Consumer Electronics,* 2008

Gladys M., Casas C., Grau R. and A. Milagros., *Técnicas de clustering para la detección de epidemias,* 1999

Gosselin P., Lebel G., Rivest S. and D. Monique, The Integrated System for Public Health Monitoring of West Nile Virus (ISPHM-WNV): a real-time GIS for surveillance and decision-making. *International Journal of Health Geographics*., 2005

Grabler F., Agrawala M., Sumner Robert W., and M. Pauly., Automatic generation of tourist maps. *ACM Transaction on Graphics*, 27 (3), Article 100, 2008

Güvenir H.A. and N. Emeksizb.,An expert system for the differential diagnosis of erythemato-squamous diseases. *Expert Systems with Applications*, 18 (1), pp. 43-49., 2008

Ishibuchi H., Nakashima T., and T.Murata., Performance Evaluation of Fuzzy Classifier Systems for Multidimensional Pattern Classification Problems. *IEEE Trans.* SMC–B 29, pp. 601–618., 1999

Kalczyński P., Paprzycki M., Fiedorowiczl I., Abramowicz W. and C. Maria., Personalized traveller information system, *Proceedings of the 5th International Conference Human-Computer Interaction*, 2001

Muller Ross M., Kyusuk C., Croke K. G. and K. Mensah Edward., Geographic Information Systems in Public Health and Medicine. *Journal of Medical Systems* 28 (3), pp. 215-221., 2004

Rodriguez A. 2007. *ODDX: Ontology Based Differential Diagnosis*. *Final Year Project*, Computer Science Department, University of Oviedo, Oviedo, Spain.

Rodriguez A., Mencke M., Alor Hernandez G., Posada Gomez R. and J. M. Gomez., Medboli: Medical Diagnosis Based on Ontologies and Logical Inference. *The Third International Conference on Digital Society, ICDS 2009*, February 1 - 7, 2009, Cancun, Mexico (accepted, to appear)., 2008

Roubos J.A., M. Setnes., Compact Fuzzy Models through Complexity Reduction and Evolutionary Optimization. In *FUZZ-IEEE*, pp. 762-767., 2000