Chapter 7.9
Management of Medical Website Quality Labels via Web Mining

Vangelis Karkaletsis
National Center of Scientific Research "Demokritos", Greece

Konstantinos Stamatakis
National Center of Scientific Research "Demokritos", Greece

Pythagoras Karampiperis
National Center of Scientific Research "Demokritos", Greece

Martin Labský
University of Economics, Prague, Czech Republic

Marek Růžička
University of Economics, Prague, Czech Republic

Vojtěch Svátek
University of Economics, Prague, Czech Republic

Enrique Amigó Cabrera
ETSI Informática, UNED, Spain

Matti Pöllä
Helsinki University of Technology, Finland

Miquel Angel Mayer
Medical Association of Barcelona (COMB), Spain

Dagmar Villarroel Gonzales
Agency for Quality in Medicine (AquMed), Germany

ABSTRACT

The World Wide Web is an important channel of information exchange in many domains, including the medical one. The ever increasing amount of freely available healthcare-related information generates, on the one hand, excellent conditions for self-education of patients as well as physicians, but on the other hand, entails substantial risks if such information is trusted irrespective of low

competence or even bad intentions of its authors. This is why medical Web site certification, also called quality labeling, by renowned authorities is of high importance. In this respect, it recently became obvious that the labelling process could benefit from employment of Web mining and information extraction techniques, in combination with flexible methods of Web-based information management developed within the Semantic Web initiative. Achieving such synergy is the central issue in the MedIEQ project. The AQUA (Assisting Quality Assessment) system, developed within the MedIEQ project, aims to provide the infrastructure and the means to organize and support various aspects of the daily work of labelling experts.

INTRODUCTION

The number of health information websites and online services is increasing day by day. It is known that the quality of these websites is very variable and difficult to assess; we can find websites published by government institutions, consumer and scientific organizations, patients associations, personal sites, health provider institutions, commercial sites, etc. (Mayer et.al., 2005). On the other hand, patients continue to find new ways of reaching health information and more than four out of ten health information seekers say the material they find affects their decisions about their health (Eysenbach, 2000; Diaz et.al., 2002). However, it is difficult for health information consumers, such as the patients and the general public, to assess by themselves the quality of the information because they are not always familiar with the medical domains and vocabularies (Soualmia et.al., 2003).

Although there are divergent opinions about the need for certification of health websites and adoption by Internet users (HON, 2005), different organizations around the world are working on establishing standards of quality in the certifica-

tion of health-related web content (Winker et.al., 2000; Kohler et.al., 2002; Curro et.al., 2004; Mayer et.al., 2005). The European Council supported an initiative within eEurope 2002 to develop a core set of "Quality Criteria for Health Related Websites" (EC, 2002). The specific aim was to specify a commonly agreed set of simple quality criteria on which Member States, as well as public and private bodies, may build upon for developing mechanisms to help improving the quality of the content provided by health-related websites. These criteria should be applied in addition to relevant Community law. As a result, a core set of quality criteria was established. These criteria may be used as a basis in the development of user guides, voluntary codes of conduct, trust marks, certification systems, or any other initiative adopted by relevant parties, at European, national, regional or organizational level.

This stress on content quality evaluation contrasts with the fact that most of the current Web is still based on HTML, which only specifies how to layout the content of a web page addressing human readers. HTML as such cannot be exploited efficiently by information retrieval techniques in order to provide visitors with additional information on the websites' content. This "current web" must evolve in the next years, from a repository of human-understandable information, to a global knowledge repository, where information should be machine-readable and processable, enabling the use of advanced knowledge management technologies (Eysenbach, 2003). This change is based on the exploitation of *semantic web* technologies. The Semantic Web is "an extension of the current web in which information is given a well-defined meaning, better enabling computers and people to work in cooperation" based on metadata (i.e. semantic annotations of the web content) (Berners-Lee et.al., 2001). These metadata can be expressed in different ways using the Resource Description Framework (RDF) language. RDF is the key technology behind the Semantic Web, providing a means of expressing

Related Content

A Knowledge-Based Web Information System for the Fusion of Distributed Classifers
Grigorios Tsoumakas, Nick Bassiliadesand Ioannis Vlahavas (2004). *Web Information Systems (pp. 268-304).*
www.irma-international.org/chapter/knowledge-based-web-information-system/31128

Saving Face in Online Learning: New Directions in Teaching and E-Learning
Lena Paulo Kushnirand Kenneth Berry (2015). *Artificial Intelligence Technologies and the Evolution of Web 3.0 (pp. 167-187).*
www.irma-international.org/chapter/saving-face-in-online-learning/127288

Efficient Low-Power Compact Hardware Units for Real-Time Image Processing
Khaldoon M. Mhaidat, Mohammad I. Alaliand Inad A. Aljarrah (2014). *International Journal of Information Technology and Web Engineering (pp. 24-39).*
www.irma-international.org/article/efficient-low-power-compact-hardware-units-for-real-time-image-processing/124027

A Novel Approach to Design a Scalable Air Quality Monitoring System
Ivaylo Atanasov, Ventsislav Trifonov, Kamelia Nikolovaand Evelina Pencheva (2012). *International Journal of Information Technology and Web Engineering (pp. 37-54).*
www.irma-international.org/article/novel-approach-design-scalable-air/72991

'Talking Tools': Sloyd Processes Become Multimodal Stories with Smartphone Documentation
Annika Wiklund-Engblom, Kasper Hiltunen, Juha Hartvik, Mia Porko-Huddand Marléne Johansson (2016). *Web Design and Development: Concepts, Methodologies, Tools, and Applications (pp. 1770-1788).*
www.irma-international.org/chapter/talking-tools/137421