



Dissemination and Exploitation strategy

Distribution: Public

MedIEQ

Quality Labeling of Medical Web content using Multilingual
Information Extraction

National Centre for Scientific Research "Demokritos"
Teknillinen Korkeakoulu – Helsinki University of Technology
Universidad Nacional de Educacion a Distancia
Col.legi Oficial de Metges de Barcelona
Zentralstelle der deutschen Ärzteschaft zur Qualitätssicherung in der Medizin
Vysoka Skola Ekonomicka V Praze
I-Sieve Technologies Ltd

2005107 **Deliverable 17**

February 2007

Project ref. no.	2005107
Project acronym	<i>MedIEQ</i>
Project full title	<i>Quality Labeling of Medical Web content using Multilingual Information Extraction</i>

Security (distribution level)	<i>Public</i>
Contractual date of delivery	<i>31 December 2006</i>
Actual date of delivery	<i>16 February 2007</i>
Deliverable number	<i>D17</i>
Deliverable name	<i>Dissemination and Exploitation Strategy</i>
Type	<i>Report</i>
Status & version	<i>Final</i>
Number of pages	<i>15</i>
WP contributing to the deliverable	<i>WP2</i>
WP / Task responsible	<i>WMA</i>
Other contributors	<i>NCSR, AQuMed</i>
Author(s)	<i>Miquel A. Mayer, Angela Leis (WMA), Vangelis Karkaletsis, Ioannis Dermousis (NCSR) Dagmar Villaroel Gozales (AQuMed)</i>
EC Project Officer	<i>Artur Furtado</i>
Keywords	<i>Dissemination activities, papers, conferences, forums, seminars, exploitation strategy</i>
Abstract (for dissemination)	<i>The main goals of this activity are: define the possible strategies to guarantee the widest dissemination of the project results, define the possible targets groups and refine it constantly up to the end of the project.</i>

Table of Contents

Executive Summary	4
1. Introduction.....	5
2. Dissemination and Exploitation Strategy	6
3. Activities performed during the 1st year of the project	8
3.1. MedIEQ Website	8
3.2. Leaflets and Newsletter	8
3.3. Conferences, Workshops	8
3.4. Lectures	9
3.5. Advisory Committee	9
3.6. World Wide Web Consortium (W3C).....	10
4. Activities planned for the 2nd year of the project.....	12
4.1. Scheduled activities	12
4.2 Next potential conferences and activities	12

Executive Summary

This deliverable presents the details of the dissemination and exploitation strategy and the activities to be performed in the project. The general and specific objectives are presented. The specification of an early version of the dissemination and exploitation strategy is presented, defining all the possible targets groups and instruments to apply this strategy.

In this deliverable we describe the activities performed during the 1st year of the project that include: MedIEQ website, leaflets and newsletters, workshops and conferences, lectures, the Advisory Committee (AC) meeting and the World Wide Web Consortium (W3C). The AC is formed by different experts in the area and its members will be able to use and test several of the results in the project's development (see the "Advisory Committee Meeting: Compiled Report"). NCSR is actively involved as a sponsoring organisation in the W3C Incubator Group on Content Labels (WCL) and the results will be a relevant W3C recommendation.

The activities planned for the 2nd year of the project are described as well.

The dissemination strategy will be continuously refined until the end of the project.

1. Introduction

Deliverable D17 together with Deliverables D18 “Dissemination Workshop” and D19 “Exploitation plan”, make up WP2. WP2 will specify in detail the dissemination and exploitation strategy.

The main dissemination strategies are: papers submission to specialized journals, international and national conferences, the MedIEQ site, end-user and public events, press releases and the Info/Open Day workshop. Dissemination and exploitation strategy will be continuously reviewed and refined until the end of the project.

WMA and AQuMED are actively involved in WP2 to specify and to refine the exploitation and dissemination strategies, reviewing them continuously. Furthermore, they will develop its dissemination activities defined and will take over, with NCSR, the organization of the Workshop Info/Open Day. NCSR will take over the main role of organising the Workshop in the third year and. All the partners are actively involved in various dissemination activities (publications, demos, etc.).

Section 2 of D17 presents the objectives of the dissemination and exploitation strategy. Section 3 presents the relevant activities during the 1st year of the project whereas section 3 presents the activities planned for the 2nd year of the project (either scheduled ones or being examined).

2. Dissemination and Exploitation Strategy

There are two major mechanisms in medical quality labelling (see deliverables D4.1, D5). The first one is based on third party rating where the web site is assessed by a labelling agency and, if the criteria are met, a label is added to the web site. This is the model used by WMA. The second type of labelling mechanism examines medical web sites in specific thematic areas, characterizes them against certain criteria, filters some of them based on their characterization, and organizes the rest into web directories to facilitate access by health information consumers. This is the approach used by the Agency for Quality in Medicine (AQuMED).

The main and more common problems that these mechanisms face are related to the lack of machine processable labels that can be identified and parsed by search engines or web browsers, as well as the need for continuous review and control of already characterised (accredited or classified) web sites and the identification of ones that have not been characterised yet, tasks that currently require a huge amount of human effort. WMA, as third-party accreditation system, for instance, periodically reviews manually the accredited web sites to renew the corresponding quality labels.

Labelling authorities will be equipped with technologies that support the monitoring of already labelled sites as well as the detection of unlabelled ones. This requires the use in practice of

- web content analysis technologies, such as crawling for detecting medical web sites, spidering for locating inside those sites web pages relevant to the labelling criteria examined, and
- information extraction technologies for acquiring data from the located web pages that correspond to the labelling criteria, and which will be either compared to existing labelling data or will be stored in order to be validated and enriched by the labelling experts.

For more details on these technologies see deliverables D7.1 “Web Content collection toolkit” and D12, where the MedIEQ integration architecture (AQUA system) is described.

The MedIEQ tools and technologies are expected to have a significant impact on medical quality labelling assisting the work of labelling experts, increasing the number of labelled medical sites across Europe and their effective monitoring, and thus improving the quality health knowledge disseminated through the Web. These technologies will be evaluated according to the evaluation strategy presented in deliverable D15. The results of this evaluation will be shared with other Labelling Agencies and with the experts of the Advisory Committee that was created in MedIEQ as another instrument to increase the dissemination of project results.

General objectives of the dissemination plan are:

- Dissemination and exploitation of project results
- Contribution to international standardisation bodies (W3C semantic web initiatives)

Specific objectives of the dissemination plan are:

- To publicize the development of the MedIEQ tools, encouraging as many labelling schemes as possible to participate and to engage relevant policy makers.
- To encourage the practice of quality labelling agencies.

- To develop partnerships for all the applications related to MedIEQ.
- To form a user group with professionals from medical associations, labelling organizations and consumer associations.
- Contribution to international standardization bodies (W3C semantic web and content labelling initiatives)
- To exploit the links of the research partners with other researchers in the area of semantic web techniques.

The dissemination strategy will be continuously refined until the end of the project taking into account the technical progress, as well as the inputs of the potential users. WMA, in cooperation with AQuMED and NCSR, will be redefining continuously the details of the dissemination strategy. We emphasize the role of the project website which we expect to become a focus of attention among public health practitioners and information seekers.

The following groups will be targeted for dissemination purposes:

- Rating and labelling initiatives:
 - Trustmak schemes and Medical gateways as Health on the Net Foundation (HON) or CiSMeF
- Health content providers In order to create awareness and technical interest, explain benefits
- Search engines
- Industry associations
- Potential End Users as labelling authorities, universities and other information society professionals as documentary professionals
- World Wide Web Consortium (W3C)

In the context of the project dissemination strategy the following instruments will be exploited:

- MedIEQ website
- Attending/organizing forums, seminars and workshops
- Speaking demonstrations
- General and individual meetings with target groups
- Press releases
- Information leaflets
- Papers in scientific journals
- National and International medical and informatics conferences.
- Tools developed in QUATRO EU project for reading RDF labels: LADI meta-search functionality and ViQ browser plug-in.

3. Activities performed during the 1st year of the project

The following dissemination activities have been done during 2006 by the project partners:

3.1. *MedIEQ Website*

The project web site was set up by NCSR at <http://www.medieq.org> (the domain name was purchased for duration of 3 years).

The public area of the project site contains information on the project objectives and partners as well as a set of the public documents produced so far (project leaflets, press releases in 8 languages, papers accepted for publication, presentations in conferences).

A link is also provided to the relevant recently ended project QUATRO (DG INFSO, Safer Internet programme) and the tools developed (QUATRO tools will be exploited in MedIEQ for demonstrating project results to the visitors of labelled web sites).

Links will also be added to relevant labelling activities either on-going or recently completed.

All the website partners put a link in their websites.

The project logo was finalized updating the one used in the proposal (this was set up by NCSR) following WMA's suggestions.

3.2. *Leaflets and Newsletter*

- Project leaflets in English have been produced and distributed to several events.
- Newsletter addressed to more than 25,000 medical doctors. "El projecte internacional MedIEQ (in Catalan)". July-October 2006.

3.3. *Conferences, Workshops*

- J. Kosek, P. Nalenka, "Relaxed—on the Way Towards True Validation of Compound Documents", Proceedings of 15th International World Wide Web Conference (WWW'2006), 23-26 May 2006, Edinburgh, Scotland.
- Mayer MA, Karkaletsis V, Stamatakis K, Leis A, Villarroel D, Thomeczek C et al. MedIEQ – Quality Labelling of Medical Web Content Using Multilingual Information Extraction. En: Medical and Care Compunetics 3. L. Bos et al. (Eds). IOS Press. Proc ICMCC Event 2006: 183-190.
- J. Kosek, M. Labsky, J. Nemrava, M. Ruzicka, V. Svatek, Projekt MedIEQ: hodnocení zdravotnických webových zdrojů s využitím extrakce informací (in Czech). In: Datakon 2006, Proceedings of the Annual Database Conference, October 2006, Brno, Czech Republic, 267-270.
- V. Karkaletsis, K. Stamatakis, V. Metsis, V. Redoumi, D. Tsarouhas, "Health-related Web Content: quality labelling mechanisms and the MedIEQ approach", Proceedings

of the 4th International Conference on Information Communication Technologies in Health (ICICTH-2006), July 13-15, 2006, Samos Island, Greece (<http://www.ineag.gr/ICICTH/index.html>)

- Mayer MA, Leis A, Karkaletsis V, Vilarroel D. El proyecto europeo MedIEQ (Quality Labelling of Medical Web content using Multilingual Information Extraction): la Web Semántica al servicio de los usuarios de salud (in Spanish). In: Nuevas interfaces centradas en el usuario: tendencias en la organización de contenidos, documentos y bibliotecas. Proceedings of the VIII Congress of Information Management. Spanish Scientific and Documentary Information (SEDIC), Madrid (Spain): 43-51. October 2006.
- Mayer MA, Leis A, Ruiz P, Karkaletsis V, Stamatakis K. MedIEQ: metadatos y sistemas de extracción semántica de información sanitaria en Internet y su aplicación en estrategias de calidad (in Spanish). In: Las tecnologías de la información. Herramienta para la mejora de la calidad asistencial. Proceedings of the XI National Conference of Medical Informatics. Informed 2006, Murcia (Spain): 67-72. November 2006.

3.4. Lectures

- V. Karkaletsis, invited speaker on the Public Health Programme at the eHealth 2006 High Level Conference, Session “Global trends and perspectives”, Malaga, 12 May, 2006 (<http://www.ehealthconference2006.org/>).
- V. Karkaletsis, invited speaker (lecture title “Quality Labeling of Web Content”) at the 3rd IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI 2006), 9 June, Athens, Greece (<http://www.icsd.aegean.gr/aiai2006/>)
- V. Karkaletsis, presentation of the MedIEQ project at the Joint Meeting of the DG SANCO Health Systems Working Party and DG INFSO eHealth Working Group, 22 June, Luxembourg.
- MA Mayer, invited talk at the Workshop “La acreditación de calidad de los sitios web de salud” at the Autònoma University of Barcelona. InCom-UAB. Barcelona, Spain, 9 October 2006.

3.5. Advisory Committee

The advisory committee (AC) formed involves known experts in the area:

- Phil Archer (Internet Content Rating Association – ICRA, UK)
- Celia Boyer (Health On the Net Foundation – HON, Switzerland)
- Stefan Darmoni (Catalog and Index of French-speaking Medical Sites – CISMEF, France)
- Sabine Laversin (Haute Autorite de Sante – HAS, France)
- Petra Wilson (Cisco Systems, Belgium)

We consider the AC as another instrument to increase the dissemination of project results. The AC members will be able to use and test several of the results in the project’s course,

providing their feedback to the consortium but also presenting and discussing their views outside the consortium. An AC meeting with the project consortium and the project officer was held at the context of the 3rd plenary project meeting, at Helsinki, October 24, 2006 (see “Advisory Committee Meeting: compiled report”).

3.6. World Wide Web Consortium (W3C)

NCSR is actively involved as a sponsoring organisation in the W3C Incubator Group on Content Labels (WCL) (<http://www.w3.org/2005/Incubator/wcl/>). P. Nasikas from NCSR attended the group meeting in Edinburgh in the context of the 15th International World Wide Web Conference (WWW'2006), held on 23-26 May 2006, in Edinburgh, Scotland. Dan Brickley, subcontractor to NCSR, is also an active member of the same group. He also attended the group meeting. The aim is to follow in MedIEQ the data model for content labels to be proposed by the group.

WCL members are currently working for the formation of a W3C working group, named Protocol for Web Description Resources (POWDER) Working Group, that aims to exploit WCL results towards a relevant W3C recommendation (see http://www.w3.org/2006/12/powder_charter.html).

4. Activities planned for the 2nd year of the project

At the beginning of the 2nd year of the project, the partners have already scheduled certain dissemination activities, for which they are currently working on, and they are exchanging views for more activities in the coming months.

4.1. *Scheduled activities*

- Conference. 11th National congress of Internet, Telecommunications and Information Society. It will take place in Malaga (Spain), March 14-16. Accepted.
- Preparation of a paper on MedIEQ content collection methodology, to be submitted at the 11th Conference on Artificial Intelligence in Medicine (AIME-2007), Amsterdam, July 7.
- Conference. Mednet 2007, Leipzig (Germany).
- Organise an initial demo of AQUA during the "eHealth week 2007: From Strategies to Applications", April 16-20, 2007, in Berlin.
- Invited talk of M.A. Mayer. Information Resources Group (IRG) Workshop at Health Technology Assessment International (HTAi) in Barcelona. June 2007.

4.2 *Next potential conferences and activities*

- **Annual Healthcare Computing conference**
URL: <http://www.health-informatics.org/>
Location: Harrogate, United Kingdom
Date: March 19-21, 2007
- **WEBIST 2007 (the 3rd International Conference on Web Information Systems and Technologies)**
URL: <http://www.webist.org/>
Location: Barcelona
Date: March 3-6, 2007
- **2007 Government & Health Technologies Conference and Expo**
URL: <http://gov.wowgao.com/>
Location: Canada
Date: April 17 -18, 2007
- **eHealth Conference 2007**
URL: <http://ehealth.gvg-koeln.de>
Location: Berlin (Germany)
Date: April 16-20, 2007
- **Medical Informatics in Enlarged Europe**
URL: <http://www.hdmi.hr/stc2007/>
Location: The Brijuni Islands, Croatia
Date: 30th May to 1st June 2007

- **The 3rd Annual Government Health IT Conference & Exhibition**
URL: <http://www.e-gov.com/EventOverview.aspx?Event=GHIT07&NoCache=632926163938526983>
Location: Ronald Reagan Building, Washington, DC
Date: June 14-15, 2007
- **The international Council on medical & care compunetics (ICMCC 2007)**
URL: <http://2007.icmcc.org/>
Location: The Netherlands
Date: June 8-10, 2007.
- **ACL 2007**
URL: <http://ufal.mff.cuni.cz/acl2007/>
Location: Prague, Czech Republic
Date: June 23-30, 2007
- **4th European Semantic Web Conference**
URL: <http://www.eswc2007.org/>
Location: Innsbruck, Austria
Date: June 3-6, 2007
- **Health Technology Assessment International (HTAi)**
URL: <http://www.htai.org/>
Location: Barcelona, Spain
Date: June 17-20, 2007
- **The International Symposium on Health Information Management Research (iSHIMR 2007)**
URL: <http://dagda.shef.ac.uk/ishimr07/index.html>
Location: Sheffield, United Kingdom
Date: July 18-20, 2007
- **11th Conference on Artificial Intelligence in Medicine (AIME 07)**
URL: <http://www.aimedicine.eu/AIME07/index.jsp>
Location: Amsterdam, The Netherlands
Date: July 7-11, 2007
- **MedINFO 2007**
URL: <http://www.medinfo2007.org/>
Location: Australia
Date: August 20 - 24, 2007
- **18th European Conference on Machine Learning**
URL: <http://www.ecmlpkdd2007.org/>
Location: Poland
Date: September 17-21, 2007
- **The 12th World Congress on Medicine in Internet (MedNet 2007)**
URL: <http://mednet2007.com/content/>
Location: Leipzig, Germany
Date: October 7-10, 2007.
- **The World of Health IT (focuses on Innovation and Deployment)**
URL: <http://www.worldofhealthit.org/>

Location: Vienna, Austria
Date: October 22-25, 2007

- **The 6th International Semantic Web Conference**

URL: <http://iswc2007.semanticweb.org/>
Location: Busan, Korea
Date: November 11-15, 2007

Paper topics proposed for the above conferences/workshops include:

- General MedIEQ system architecture. (Differentiation from previous papers: more technical details.)
- Information extraction methods used, problems we face during the processing of the web pages (e.g. new technologies like AJAX etc.)
- Methods used for the classification of Web Pages in various categories. Performance of these methods. Cross-language evaluation
- Contribution of semantic web technologies in web site labelling. Which features of semantic web MedIEQ is using in the project. Problems we face.
- Current labelling techniques & labelling authorities.
- Current state of medical web content in the countries that participate in the project. Analysis of the corpus collected (Medical sites URLs).
- Focused crawling and corpus formation methodology.

