HIC 2008 Australia's Health Informatics Conference
Editor: Heather Grain
Publisher: Health Informatics Society of Australia Ltd (HISA)
ISBN 978 0 9805520 0 3

Intelligent Support for Resource Quality Evaluation and Description in Health Information Portals

Jue Xie, Frada Burstein, Joanne Evans

Centre for Organisational and Social Informatics, Monash University, Melbourne, Victoria

Abstract

Quality control is a critical issue in online health information provision, where domain experts play an important role in evaluating and selecting quality and relevant resources to meet health consumers' individual knowledge needs. This research aims for a semi-automated approach to tackle challenges domain experts encounter in resource quality evaluation and description in the context of health information portals. This paper reports on the first stage of this research in which the domain experts' needs for intelligent supports are investigated. From this analysis, the architecture of a domain expert dashboard is proposed, where quality framework is distinguished among a suite of tools.

Objective:

This research aims to identify the need for intelligent supports to aid domain experts who perform information quality control through the process of resource identification, selection and description in health information portals.

Background:

In the field of consumer health informatics, the quality control of health information on the Internet has been raised as a critical issue (Eysenbach 2000), especially when consumers use information retrieved from the Web to make health related decisions (Burstein et al. 2005).

According to Eysenbach and Diepgen (1998), information quality can be controlled or monitored either at the time of production or by means of filtering published information against a set of quality criteria. The quality control approach used in the production stage is to develop ethical standards or codes of conduct for encouraging the self-regulation of health information providers (Risk and Dzenowagis 2001). For the quality management of published information on the Internet, two approaches have been discussed as "upstream filtering" by third parties versus "downstream filtering" by information consumers (Eysenbach and Diepgen 1998).

From the literature it can be seen that a lot of effort has been directed towards developing "downstream filtering" approaches to support health information consumers evaluating, filtering or rating online health information through the use of some technological means, such as quality checklists (DISCERN 2008), numerical rating tools (Gagliardi and Jadad 2002; Bomba 2005), scoring systems (Currò et al. 2004), quality seals and certifications (HONcode 2007; TRUSTe 2008). However, as suggested by Ziebland's (2004 p.1784) empirical study, "it is unreasonable to expect patients to be able to evaluate the vast quantities of health information they find on the web". Research has also shown that many people are unprepared for taking such responsibilities (Metzger 2007).

On the other hand, the "upstream filtering" approach is to have third parties, usually professional human reviewers mediating between information providers and information consumers. Such an idea has been implemented as filtered gateways or health information portals which provide consumers with only evaluated information resources that meet an explicit set of quality criteria for resource selection and inclusion (Burstein et al. 2005). The term domain expert is used for those evaluating, selecting and describing resources for inclusion in the portal.

HIC 2008 Australia's Health Informatics Conference
Editor: Heather Grain
Publisher: Health Informatics Society of Australia Ltd (HISA)
ISBN 978 0 9805520 0 3

The role of a domain expert requires a very good understanding of the specific problem domain, along with intimate knowledge of the information needs of portal users. These domain experts evaluate the quality of online resources and their relevance to the intended users by interpreting descriptive evaluation guidelines based on their individual domain knowledge, expertise and experience. Due to the subjective nature of information quality and its assessment, it is not surprising that resource quality evaluation and description in practice rely on manual procedures. Controlling the accuracy and the consistency of evaluation results drawn by different evaluators becomes a key concern. The emergence of new solutions, such as supporting these procedures with intelligent tools, is imperative to improve quality control outcomes.

Compared to the extensive research efforts made towards information providers and information consumers, domain experts' needs and their collaboration with intelligent systems in resource quality evaluation and description have not been sufficiently addressed. This research will identify the domain experts' needs for intelligent supports and propose the architecture of a domain expert interface to a portal's knowledge repository to aid in resource quality evaluation and description. The extent to which intelligent technologies can assist in assessing information for better quality control will then be studied through prototype implementation and evaluation.

Automated quality assessment methods: what are in use?

Research on automated quality assessment for online resources is relatively new in the literature. Automated quality assessment methods have been developed mostly aimed at assisting consumers during their information seeking activities. For example, Healthfinder.gov lists a number of rating instruments as recommended evaluation tools for consumers. However, Jadad and Gagliardi's (1998) study argued the validity and reliability of those instruments in measuring the quality of online health information.

Meanwhile a number of studies try to auto-detect widely used quality indicators in published guidelines, as listed in Table 1. For instance, works on automatically detecting technical quality criteria of Web-based health information have been published recently as their first step to assist consumers evaluating online health information (Wang and Liu 2007). Indirect quality indicators for evaluating websites, such as daily visits, updating frequency and the quantification of inbound links, have also been explored (Hernández-Borges et al. 1999).

Quality indicator detecting methods	For evaluating quality dimension(s) of
Website Usage analysis and citation analysis (Cui 1999; Hernández-Borges et al. 1999)	Reliability
AQA (Automated Quality Assessment) Procedure (Griffiths et al. 2005)	Accuracy
Automatic indicator detection tool (AIDT) (Wang and Liu 2007)	Authority, source, currency, content, disclosure, interactivity, and commercialization

Table 1. *Sample quality indicator detecting methods*

Apart from that, studies have been conducted to find the correlation between the technical IQ criteria and the content IQ criteria in order to predict the content quality, such as accuracy and completeness (Eysenbach et al. 2002; Martin-Facklam et al. 2002). Moreover, for assessing the content accuracy, Griffiths et al. (2005) proposed an automated procedure to identify whether a health website has evidence-based information, but limited in the domain of depression websites.

HIC 2008 Australia's Health Informatics Conference
Editor: Heather Grain
Publisher: Health Informatics Society of Australia Ltd (HISA)
ISBN 978 0 9805520 0 3

Intelligent support: what can be provided?

Cooke (2001) outlines a number of strategies for locating quality information resources on the Internet. An examination of these strategies suggests the following areas have the potential to be supported by intelligent tools, namely:

- using searching facilities to improve information retrieval,
- assessing the quality of an information source, and
- using and collecting quality indicators.

Hence the application of intelligent techniques based on modern information retrieval methods (Baeza-Yates and Ribeiro-Neto 1999; Kobayashi and Takeda 2000; Laender et al. 2002; Chowdhury 2004), such as Web data mining, natural language processing and machine learning techniques, to determine resource quality values merits investigation.

The study context: the BCKOnline portal

BCKOnline (www.bckonline.monash.edu.au) is an intelligent, user-sensitive information portal developed by a Monash University led research team to meet the diverse knowledge needs of the Victorian breast cancer community (Fisher et al. 2002). From the portal's user information needs analysis, the quality of online information resources, in terms of reliability, relevance and timing issues, has been identified as one of major concerns of intended portal users (Fisher et al. 2002). To reflect such a concern, the portal respects its users as final judges of the information quality and recognises the quality of an information resource as the relationship between the resource and a user rather than an attribute of the resource itself. As a result, the portal provides a metadata-based narrative description about the information quality to help its users to identify the reliability, provenance, authority and quality of the information resource based on their individual value systems (Burstein et al. 2005). However, the challenge of sustainability and scalability of manual quality assessment by domain experts in building and maintaining the portal's knowledge repository has been identified. New concerns, such as having quality assessment procedures supported by intelligent tools, have been raised in order to improve the information quality management in a portal context.

Method:

A study of manual procedures involving the domain experts was conducted to identify requirements for intelligent supports to aid domain experts in resource quality evaluation and description. This helped to:

- understand the roles of domain expertise – i.e. first hand experience and extensive knowledge of the medical, supportive and psychosocial information needs of the community associated with a particular disease or medical condition – in information resource quality evaluation and description procedures;
- identify the parts of the procedures that could be automated and/or augmented by intelligent technologies to improve their efficiency and effectiveness.

Discussions were drawn on three main issues, namely:

1. The BCKOnline user-centred resource categories and selection criteria: quality information resources of 'medical', 'supportive' or 'personal' category have been selected for resource inclusion based on results of an extensive user needs analysis of the breast cancer community. "Resources were deemed to meet a majority of the criteria designated for each category" (BCKOnline 2005).

2. The BCKOnline metadata schema: the schema has been designed to include user-centric resource descriptors, i.e. target audience. The quality element and related encoding scheme capture values of information quality assessment undertaken by domain experts.

HIC 2008 Australia's Health Informatics Conference
Editor: Heather Grain
Publisher: Health Informatics Society of Australia Ltd (HISA)
ISBN 978 0 9805520 0 3

3. The BCKOnline content management system: a content management system has been developed to capture resource identification, selection and description processes to build and maintain the portal's knowledge repository.

Results:

Multiple facets of required domain expertise are identified, as shown in Table 2:

Domain expertise	Description
Intimate domain knowledge and first hand experience	The domain expert is an advocate for the portal users. In the case of BCKOnline, resource identification, selection and description are all based on an intimate knowledge of the breast cancer experience.
Connection into information networks	The domain expert comes to the portal already connected into a range of information networks in their area of expertise.
Awareness of the kinds and styles of information most valued by portal users	Subject and tone are both important selection criteria for a resource. With the plethora of information available it is either their topic or take on the topic that informs their inclusion. In the case of BCKOnline, the desire is to include resources that are not 'patronising' and that feature 'lesser voices'. Local resources are also given preference.
Understanding of health informatics	The domain expert requires knowledge of the structure and nature of health information in their domain, along with the systems that produce it. The ability to critically appraise sources of healthcare information from the perspective of their portal users is also required. In the case of BCKOnline, appraisal is from a breast cancer patient or a healthcare consumer perspective.
Understanding of library and information management principles	The domain expert also requires an understanding of classification and indexing principles, particularly impacts on precision and recall. They should also have an appreciation of the change digital and networked technologies are having on traditional practices.

Table 2. *Domain expertise required in resource quality evaluation and description.*

The identified domain expertise can be supported with the following components which form the architecture for the domain expert interface to the portal's knowledge repository. This integrated system has been named a domain expert dashboard, with a quality framework tool identified as a key component to support domain experts in assessing the resource quality. This architecture is presented in Figure 1.

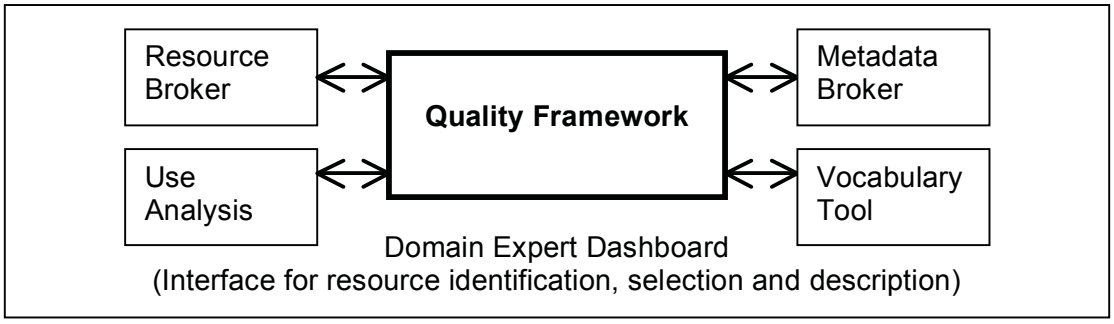

Figure 1. *Domain expert dashboard*

ISBN 978 0 9805520 0 3

A range of intelligent technologies can be utilised to implement the components of the architecture, with the following functionalities:

Quality framework supports domain experts in making decisions on information quality issues. It sets the quality selection criteria as well as provides intelligent services for performing resource quality evaluation and description.

Resource broker facilitates resource identification in terms of searching candidate resources of the portal's subject interests through monitoring appropriate information channels.

Metadata broker applies automated techniques to aid metadata generation against the portal's resource description metadata schemas and encoding systems.

Use analysis component monitors and analyses user behaviours and user information needs in order to refine both resource selection and resource description.

Vocabulary tool aids the development and maintenance of the portal's subject indexing vocabulary. It includes vocabularies of lay terms as well as medical jargon, and captures relationships among them.

Discussion:

The findings of the study presented in this paper reveal a need for intelligent supports for domain experts in their quality evaluation and description activities. These findings are based on the experience of developing the BCKOnline health information portal. Further implementation of the proposed architecture is being undertaken to explore the extent to which intelligent technologies can support domain experts in evaluation and description of information resource quality, and the impact this approach will have on information quality control outcomes.

Acknowledgement:

This research is supported in part by a Monash University Postgraduate Research Scholarship and an ARC Discovery Project grant DP0665353, Smart Information Portals: Meeting the knowledge and decision support needs of health care consumers for quality online information, Chief Investigators A/Prof. Frada Burstein, Prof. Sue McKemmish and A/Prof. Julie Fisher; Partner Investigator Prof. Jim Warren; Research Fellows: Dr Joanne Evans and Rosetta Manaszewicz; Research Assistants: Kate Lazarenko and Jue Xie.

References:

BCKOnline (2005), 'Breast Cancer Knowledge Online: How resources have been selected', viewed 4 March, 2008, <http://www.bckonline.monash.edu.au/search/howResources.jsp>.
Baeza-Yates R. and Ribeiro-Neto B. (1999), Modern information retrieval, ACM Press; Addison-Wesley, New York; Harlow, England.
Bomba D. (2005), 'Evaluating the quality of health web sites: Developing a validation method and rating instrument', in proceedings of the 38th Annual Hawaii System Sciences, Hawaii, USA, 3-6 January, 2005, pp. 139-148.
Burstein F., Mckemmish S., Fisher J., Manaszewicz R. and Malhotra P. (2005), 'A role for information portals as intelligent decision support systems: Breast cancer knowledge online experience', in Gupta J., Forgionne G. and Mora M. (eds.) Intelligent Decision-making Support Systems: Foundations, Applications and Challenges, Springer, New York, pp. 359-383.
Chowdhury G. G. (2004), Introduction to modern information retrieval, 2nd, Facet Publishing, London.
Cooke A. (2001), A guide to finding quality information on the Internet: selection and evaluation strategies, 2nd, Library Association, London.

ISBN 978 0 9805520 0 3

Cui L. (1999), 'Rating health Web sites using the principles of citation analysis: a bibliometric approach', Journal of Medical Internet Research, vol. 1(1), no. e4.
Currò V., Buonuomo P. S., Onesimo R., De Rose P., Vituzzi A., Tanna G. L. D. and D'Atri A. (2004), 'A quality evaluation methodology of health web-pages for non-professionals', Medical Informatics & The Internet in Medicine, vol. 29, no. 2, pp. 95 - 107.
DISCERN (2008), 'The DISCERN instrument', viewed 7 April, 2008, <http://www.discern.org.uk/discern_instrument.php>.
Eysenbach G. (2000), 'Recent advances: Consumer health informatics', British Medical Journal, vol. 320, no. 7251, pp. 1713-1716.
Eysenbach G. and Diepgen T. L. (1998), 'Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information', British Medical Journal, vol. 317, no. 7171, pp. 1496-1502.
Eysenbach G., Powell J., Kuss O. and Sa E.-R. (2002), 'Empirical studies assessing the quality of health information for consumers on the World Wide Web: A systematic review', JAMA, vol. 287, no. 20, pp. 2691-2700.
Fisher J., Burstein F., McKemmish S., Manaszewicz R., Anderson J. and Williamson K. (2002), 'Building a user sensitive Intelligent portal to breast cancer knowledge to meet diverse information needs', paper presented to 13th Australasian Conference on Information Systems, Melbourne, Australia.
Gagliardi A. and Jadad A. R. (2002), 'Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination', British Medical Journal, vol. 324, no. 7337, pp. 569-573.
Griffiths K. M., Tang T. T., Hawking D. and Christensen H. (2005), 'Automated assessment of the quality of depression websites', Journal of Medical Internet Research, vol. 7, no. 5, pp. e59.
Hernández-Borges A. A., Macías-Cervi P., Gaspar-Guardado M. A., Torres-Álvarez DeArcaya M. L., Ana R.-R. and Jiménez-Sosa A. (1999), 'Can examination of WWW usage statistics and other indirect quality indicators help to distinguish the relative quality of medical websites?' Journal of Medical Internet Research, vol. 1(1), no. e1.
HONcode (2007), 'Operational definition of the HONcode principles', viewed 19 September, 2007, <https://www.hon.ch/HONcode/Guidelines/guidelines.html>.
Jadad A. R. and Gagliardi A. (1998), 'Rating health information on the Internet: navigating to knowledge or to Babel? ' JAMA, vol. 279, no. 8, pp. 611-614.
Kobayashi M. and Takeda K. (2000), 'Information retrieval on the web', ACM Computing Surveys, vol. 32, no. 2, pp. 144-173.
Laender A., Ribeiro-Neto B., Silva A. and Teixeira J. (2002), 'A brief survey of web data extraction tools', SIGMOD Record, vol. 31, no. 2, pp. 84-93.
Martin-Facklam M., Kostrzewa M., Schubert F., Gasse C. and Haefeli W. E. (2002), 'Quality markers of drug information on the internet: an evaluation of sites about St. John's Wort', The American Journal of Medicine, vol. 113, no. 9, pp. 740-745.
Metzger M. J. (2007), 'Making sense of credibility on the Web: models for evaluating online information and recommendations for future research', Journal of the American Society for Information Science and Technology, vol. 58, no. 13, pp. 2078-2091.
Risk A. and Dzenowagis J. (2001), 'Review of Internet health information quality initiatives', Journal of Medical Internet Research, vol. 3 no. 4, pp. e28.
TRUSTe (2008), viewed 7 April, 2008, <http://www.truste.org>.
Wang Y. and Liu Z. (2007), 'Automatic detecting indicators for quality of health information on the Web', International Journal of Medical Informatics, vol. 76, no. 8, pp. 575-582.
Ziebland S. (2004), 'The importance of being expert: the quest for cancer information on the Internet', Social Science & Medicine, vol. 59, no. 9, pp. 1783-1793.