A USER-SENSITIVE RESOURCE QUALITY ASSESSMENT APPROACH FOR HEALTH INFORMATION PORTALS

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

By

Jue Xie

(Bachelor of Information Management, Master of Information Technology)

Centre for Organisational and Social Informatics

Faculty of Information Technology

Monash University

December 2011

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Information quality control is a critical issue in online health information provision. As one of effective quality control approaches, metadata-driven health information portals provide direct access to the descriptions of selected online resources that are of perceived high quality and relevancy to targeted portal users. This is achieved via a review process manually undertaken by domain experts, who have expertise in both healthcare and information management areas. Due to the subjective, contextual, and dynamic nature of information quality, the labour-intensity of resource quality (RQ) assessment becomes a bottleneck, especially when taking account of diverse user needs and values. Determining the quality of online information resources is the interplay of domain experts, portal content management systems, and RQ assessment processes. Yet, how to support contextual value judgements on RQ from a user-sensitive viewpoint, have not been sufficiently addressed in the literature and practice. The emergence of socio-technical solutions is imperative to improve the scalability and finally the sustainability of RQ assessment processes.

This research endeavours to find new approaches that employ intelligent technologies to support decision-making processes of RQ assessment for health information portals. Using a socio-technical design science research approach, the research investigated RQ assessment issues through a user-sensitive systems development research process. It involved three interconnected research phases of concept building, system building, and system evaluation. As a result, a semi-automated and user-sensitive RQ assessment approach was proposed to standardise and facilitate decision-making processes on RQ.

The concept building research phase began with a comprehensive analysis of multi-disciplinary research literature, which investigated user-sensitive RQ assessment issues from theoretical, contextual, and technological perspectives. In addition, an exploratory case study of RQ assessment practices was conducted in the context of two metadata-driven health information portals in order to identify domain expert needs and corresponding design requirements of a RQ assessment approach. The conceptualisation of the approach encompassed a user-sensitive quality assessment framework and an intelligent quality tool. The framework defined the construct of RQ as a composition of *Reliability* and *Relevancy* in the healthcare domain. The measure of *Reliability* dimension was defined using an attribute-based approach.

In the system building research phase, the feasibility of the proposed approach was tested through the design and development of the Domain Expert Dashboard (DED) prototype system. Machine learning techniques were applied to implement an intelligent system feature.

Several other system features were also developed as part of the quality tool in order to meet the decision support needs of domain experts.

In the system evaluation research phase, a functional test and a usefulness and usability study were conducted to assess the effects of the DED prototype system on the RQ decision-making processes and outcomes against a multi-criteria evaluation framework. The results demonstrated that both the processes and the outcomes of RQ assessment were improved through the use of the prototype system.

This research makes significant theoretical, methodological, and practical contributions. The proposed user-sensitive RQ assessment framework integrates and extends the context-based information quality assessment theory. The framework measure perceived RQ as a relative and aggregated construct by specifying and mapping resource attributes to the characteristics of user information needs and quality perceptions. A more generic method is also proposed for developing domain-specific and user-sensitive RQ assessment metrics for other domains. Moreover, the study adapts the canonical artefact-centric design science research framework in a socio-technical context. Mixed methodologies have been employed to conceptualise and evaluate a socio-technical solution. Theory-building has played a central role, which informed the concept-building of design artefacts. The practical contributions include the conceptual architecture and the instantiation of a quality tool using intelligent technologies. The DED prototype system provides greater functionality to support domain experts making contextual value judgements on RQ, demonstrating how intelligent learning techniques can be applied to describe the quality attributes of online resources. The system has been integrated as part of an operational health information portal, the Breast Cancer Knowledge Online (BCKOnline) portal.

Keywords: information quality, resource quality assessment, design science, intelligent tool, metadata generation, content management system, health information portal

COPYRIGHT NOTICE

DECLARATION

I certify that this thesis, except with the Research Graduate School Committee's approval, contains no material which has been accepted for the award of any other degree or diploma in any university or other institution; and affirm that to the best of my knowledge, the thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

[redacted]

Jue Xie December 2011

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Prof. Frada Burstein, Prof. Sue McKemmish, and Dr. Joanne Evans, for their guidance, support, and encouragement throughout my candidature. Their dedication and mentoring have made my candidature an enjoyable, fruitful, and inspiring journey.

This research was partly funded by an ARC Discovery grant, the Smart Information Portals (SIP) project. My thanks to members of the SIP research team for the inspiring discussions and feedback on my work. Also my thanks to people who participated in the case study interviews and the evaluation of the prototype system.

I would also like to thank the staff and fellow research students at Caulfield School of IT, for their great company and help, particularly to Dr. Maria Indrawan, Dr. Campbell Wilson, and Mr. Ben Goodrich for their useful feedback on the technical component of this research. Thanks to Dr. Viranga Ratnaike, Dr. Pari Haghighi, and Mr. Rob Gray who helped with the copy editing of this thesis. Thanks also to Kate Lazarenko and Ambica Dattakumar for their friendship and encouragement. Those jokes and coffees constitute a wonderful piece of memory that I cherish within this academic journey.

Special thanks to my family – Mum, Dad, my beloved husband Alex, and our sweetest daughter Louise, for their love, patience, understanding, and great support throughout this journey. To Mum especially – thanks for helping me looking after Louise since the first day she was born, and taking care of us during all those sleepless nights.

Finally, to my Grandfather, whose spirit empowered me to cope with all kinds of difficulties and challenges to pursue the dream that I inherited from the family.

LIST OF PUBLICATIONS

Xie J. and Burstein F. (2011), Using machine learning to support resource quality assessment: An adaptive attribute-based approach for health information portals, in Xu, J., Yu, G., Zhou, S. and Unland, R. (eds.) *Database Systems for Advanced Applications*, vol. 6637, Springer Berlin/Heidelberg. Paper presented at the 4th International Workshop on Data Quality in Integration System (DQIS'2011), Hong Kong, China, pp. 526-537.

Xie J. (2009), Sustaining quality assessment processes in user-centred health information portals, in proceedings of the 15th Americas Conference on Information Systems (AMCIS09), San Francisco, California, USA, 6-9 August, 2009, Paper 189, http://aisel.aisnet.org/amcis2009/189.

Evans J., Manaszewicz R. and Xie J. (2009), The role of domain expertise in smart, user-sensitive, health information portal, in proceedings of the 42^{nd} Hawaii International Conference on System Sciences (HICSS-42), Hawaii, USA, 5-8 January, 2009, pp. 1-10.

Xie J., Burstein F. and Evans J. (2008), Intelligent support for resource quality evaluation and description in health information portals, in proceedings of the National Conference for the Health Informatics Society of Australia (HIC'08), Melbourne, Australia, 31 August - 2 September, 2008, Chapter 41.

ACRONYMS AND ABBREVIATIONS

AGLS: Australian Government Locator Service

AJAX: Asynchronous JavaScript and XML

ARFF: Attribute-Relation File Format

CIHI: Canadian Institute for Health Information

CMS: Content Management System

DC: Dublin Core

DCMI: Dublin Core Metadata Initiative

DED: Domain Expert Dashboard

DQ: Data Quality

HIDDE: Health Information Disclosure, Description and Evaluation Language

HONcode: Health On the Net code of conduct

IDSS: Intelligent Decision Support System

IE: Information Extraction

IM: Information Management

IQ: Information Quality

IR: Information Retrieval

IS: Information Systems

ML: Machine Learning

NLP: Natural Language Processing

RQ: Resource Quality

SIP: Smart Information Portals

SMO: Sequential Minimal Optimization

SVM: Support Vector Machine

WEKA: Waikato Environment for Knowledge Analysis

ZeroR: Zero-attribute-Rule

Chapter 1

1 Introduction

This thesis presents how user-sensitive design principles and intelligent technologies can be applied to conceptualise and implement a resource quality assessment approach for metadata-driven health information portals. The research was motivated by the needs of scaling and sustaining labour intensive resource quality assessment processes for quality-assured portal content management. This chapter begins with the introduction of the problems emerged in quality online information provision. Research questions and objectives are presented. It then introduces research background with key concepts that will be used throughout this thesis, followed by a brief description of the employed research approach. The research context, scope, limitations, significance and contributions are discussed before an outline of the thesis structure.

1.1 Problem Domain and Motivation

Information overload and information quality are two major concerns in online health information provision (Benigeri and Pluye 2003; Hall and Walton 2004). Due to the enormous volume and varying quality of information available on the Internet, finding reliable, relevant and useful online information can be problematic and time-consuming for health information consumers (Hall and Walton 2004). The recognition of such problems has a long history (Hopkins 1995) and the following guideline was given for designing user-oriented and value-added information systems: "the major problems in future information systems will revolve around the processes of reducing the amount and raising the quality of information brought to the attention of the user" (Taylor 1986, p. 58). Although modern web information systems, such as search engines and web information portals, provide online solutions to address these problems, managing the volume and quality of health information being delivered via the Internet still remains a challenging task.

In the field of consumer health informatics, which studies the distribution of health information on the web for patients and lay consumers, quality control of online health information has been raised as a critical issue (Eysenbach 2000). It becomes increasingly important to assure the quality of information that consumers retrieve from the Internet, especially when the information is used to make health related decisions (Malhotra et al. 2003). Having third-party evaluators is an effective quality control approach to protect information consumers from fraudulent or harmful health information on the Internet (Eysenbach 2000; Eysenbach and

Diepgen 1998). This approach has been implemented in health information portals (also known as filtered gateways), which provide direct access to descriptions of selected online resources that meet an explicit set of quality criteria. Metadata is usually one underlying mechanism used to describe those selected web resources that are external to a portal (Asprey and Middleton 2003; Candan et al. 2001; Moura et al. 1998; Özel et al. 2004). Based on standardised metadata schemas, descriptive metadata records are generated and preserved in a portal's metadata repository, and are indexed for later retrieval. Portals that use a metadata approach to facilitate indexing, navigation, filtering, and value-added information provision are called metadata-driven (Moura et al. 2002).

Labour-intensity of resource quality assessment is a bottleneck for achieving quality-assured online health information provision via the portal approach. For metadata-driven health information portals, both the quality appraisal of online resources and the creation of resource metadata records require intensive involvement of domain experts. Domain experts are responsible for managing quality-assured portal content through resource evaluation, selection, and description processes (McKemmish et al. 2009). Due to the subjective and contextual nature of information quality (Strong et al. 1997; Wang and Strong 1996), it is not surprising that resource quality assessment in those portals is still a manual process. This makes the development and maintenance of portal content very expensive and time-consuming (Eysenbach and Diepgen 1998).

The decision-support needs of domain experts in quality assessment have not been sufficiently addressed in the literature and practice. At present, the assistance, which domain experts received for making value judgements, is limited to resource assessment criteria (selection guidelines) and resource description schemes developed by individual portals. Existing portal content management systems only assist the description and publishing of resources after they have been evaluated and selected. Those systems do not support decision-making processes of resource quality assessment. To date, neither a unified assessment framework nor standardised assessment procedures exist, which makes it difficult to scale the resource quality assessment processes and assure the consistency of resource quality assessment outcomes, particularly when multiple domain experts undertake the tasks.

Furthermore, McKemmish et al. (2009) have recently proposed a user-sensitive design methodology, which aligns all aspects of system design with the values, needs, and expectations of users. The adoption of the user-sensitive design philosophy requires a portal to identify, select, and describe online information resources from the perspective of portal users. Resource quality in this sense needs to be assessed in response to the needs and values of individual users.

This user-oriented view is in line with the widely accepted view of information quality as a relative and contextual concept (Bevan 1999; Eppler 2006; Evans and Lindsay 2005; Shanks and Corbitt 1999; Strong et al. 1997). It also brings new questions on how to conceptualise resource quality and how to systematically assess it in a user-sensitive manner. The emergence of socio-technical solutions, which employ intelligent technologies, is imperative to conceptualise and support user-sensitive resource quality assessment.

1.2 Research Questions and Objectives

The purpose of this research is to define the construct of resource quality and to explore ways of assessing it from a user-sensitive viewpoint in a healthcare domain. In order to address the perceived problems of quality-assured portal content management, and to enrich the understanding of resource quality concept and its assessment aligned with user needs and values, this research was conducted to address the following central question:

How to conceptualise and support user-sensitive resource quality assessment for metadata-driven health information portals?

In order to answer this research question, two main objectives were addressed in this research. Each objective is presented below together with the related sub-questions of this research.

Objective one: To conceptualise resource quality and its assessment taking account of the information needs and quality perceptions of health information portal users. In order to achieve this objective, it is necessary to address the following sub-questions:

1. How to define resource quality in the context of health information portals?
2. How can existing information quality assessment theories, principles and approaches be extended and adapted to conceptualise resource quality assessment from a user-sensitive viewpoint?

Objective two: To support domain experts with user-sensitive resource quality assessment using intelligent technologies. The following sub-questions were formulated accordingly:

3. What kinds of domain expertise are required in performing user-sensitive resource quality assessment in the context of metadata-driven health information portals?
4. What tasks and activities are involved in user-sensitive resource quality assessment processes?
5. What are the needs of domain experts with regard to intelligent support?

6. How can intelligent technologies be applied to support user-sensitive resource quality assessment in metadata-driven health information portals?

This research endeavours to develop the concepts of resource quality and its assessment from a user-sensitive viewpoint. The research also commits to finding new approaches, which employ intelligent technologies, to tackle the challenges domain experts encounter with resource quality assessment in a portal context. In the context of this research, technologies that can help domain experts in solving complex decision-making problems are regarded as intelligent technologies (Kreinovich et al. 2004). Varying types of intelligent technologies have been used to capture different aspects of expert decision-making, such as probabilistic techniques, fuzzy logics, neural networks, classifiers and statistical learning. Amongst those, this research focuses on the intelligent learning technologies that use known examples of expert decisions to train computer systems to make same decisions.

1.3 Research Background

Consumer health informatics is a branch of health informatics that is concerned with issues of analysing, modelling and integrating consumer information needs and preferences into information management systems, as well as developing and evaluating methods and applications to support consumers in accessing and using health information (Eysenbach 2000). The field is making rapid advances in understanding the acquisition, storage, retrieval, and application of health and medical information in ways that are expected to transform the quality and safety of healthcare nationally and globally in the 21st century (Maheu et al. 2001).

During the last decade, the Internet became increasingly popular for health information consumers to find information about health problems, self-management of chronic diseases, or disease prevention to aid their health-related decision-making (Burstein et al. 2006). Quality control of online health information is therefore emerging from consumer health informatics as a research field to study the determinants and distribution of health information and misinformation on the web (Eysenbach 2002; Eysenbach 2005a). In order to protect health information consumers from fraudulent or harmful online information, it is suggested that quality assurance on the Internet can be achieved by the following four means (Eysenbach 2000):

- educating information consumers;
- encouraging self-labelling and self-regulation of information providers;
- having third-parties evaluate and rate information; and
- enforcing compliance with criteria.

The above approaches indicate that quality control of online health information can be achieved either by information providers in the information production stage, or by third parties and information consumers in the information retrieval stage. This research draws attention to approaches that address quality control issues at the information retrieval level. The web information portal approach is of particular interest to this study. The approach mediates information providers and consumers by offering a gateway to virtual repositories of selected online information (Anderson et al. 2003; Manaszewicz et al. 2002).

For health information portals, quality assessment of online information resources plays a significant role in portal content management processes. In order to get an overall opinion about the quality of an online resource, a wide range of factors need to be collected, assessed and then considered in conjunction (Intute: Health and Life Sciences 2007). It is still a manual process that undertaken by domain experts, who follow specific quality assessment criteria or resource selection guidelines.

1.3.1 Quality-Assured Portal Content Management

Health information portals that are investigated in this research basically provide review and search services to information consumers. How satisfying these services are to portal users depends on the quality of portal content (Calero et al. 2008; Caro et al. 2006).

Portal content in general refers to data or information that a portal contains, generates, retrieves and presents to its users (Addey et al. 2002). Information portals may contain metadata descriptions of external resources or preserve the full content of the selected resources in varying formats, e.g. documents, images, audio or video materials. In general, health information consumers use health information portals as entry points to access quality-assured online information. Recent research proposes that a portal with extended functionality, can act as an intelligent decision support system (IDSS) to meet specific knowledge needs of information consumers (Burstein et al. 2006). For a portal that acts as an IDSS, its content may also include new knowledge discovered from existing data, and stored in a knowledge repository. In this case, the types of data included in a portal's knowledge repository can include resource metadata records, metadata schemas, value encoding schemas, use logs, user surveys, and user feedback (Evans et al. 2009). In this thesis, portal content mainly refers to selected web-based information resources, e.g. websites or individual web pages that are external to a portal, and their internal metadata representations, i.e. resource metadata records.

Lewis (1997) represents the scope of a management task as a model of three components: people, process and system. This research applies this simple model to illustrate the three determinants of portal content quality, specifically the quality of resources accessible via a portal. As depicted in Figure 1.1, resource quality assessment is regarded as the interplay of domain experts, resource quality assessment processes, and the content management system of a portal.

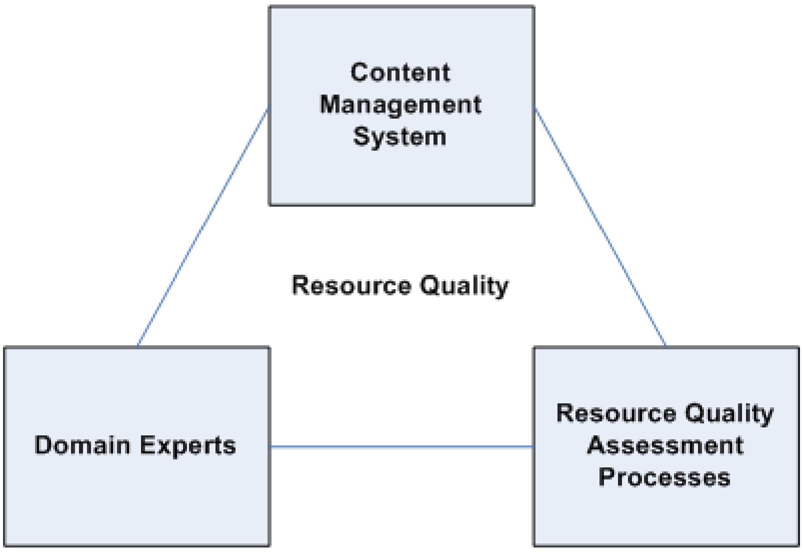


Figure 1.1 Three determinants of resource quality

The concepts of resource quality, domain experts, resource quality assessment processes, and content management system are discussed below.

Resource quality

In the context of this research, the information entities for quality assessment are web-based information resources. The study investigates the construct and assessment of resource quality (RQ), which represents a focal concept different to the notion of information quality in conventional information or data quality research. In this thesis, information quality is regarded as an attribute of a resource, whilst RQ is regarded as a relationship between a resource and a user. From this perspective, RQ is defined and assessed in terms of its relevancy and reliability in relation to user needs (information needs) and values (quality perceptions). For this reason, this research proposes the concept of RQ in the context of health information portals using the following definition:

Resource quality is the extent to which information contained in a web-based resource meets the information needs and quality perceptions of individual users.

Following this definition, individual portal users are regarded as the final judge on the quality of resources they retrieve from a portal. What domain experts assess in practice is perceived resource quality based on the knowledge of user needs and values. The role of domain experts is described next.

Domain experts

In the literature, a human intermediary in online information provision is called the 'subject specialist' (Ragetli 2005), the 'information professional' (Cooke et al. 1996), the 'metadata expert' (Greenberg et al. 2006), or as named in this thesis, the 'domain expert'. Domain experts play the role of a subject and metadata expert within the library profession in developing and maintaining quality-assured portal content. They identify and select online resources according to the resources' compliance with a portal's resource selection criteria, and create rich descriptions of selected resources based on the use of standardised metadata. Domain experts also need to continuously maintain the quality of portal content in order to reflect the dynamics in online information and user information needs (Stvilia and Gasser 2008).

Resource quality assessment processes

Broadly speaking, RQ assessment encompasses any process involved in achieving quality assurance in a portal. In the case of quality-assured information portals, the processes, which an online information resource goes through, for being included in a portal, are regarded as quality assessment processes. They include resource selection and description undertaken by domain experts. However, unlike a 'workflow', which defines the rules and specific tasks to enforce a routine process (Addey, Ellis et al. 2002), resource quality assessment processes involve the highly dynamic and discretionary activities of domain experts. Therefore, RQ assessment processes are difficult to break down into explicit and systematic quality assessment steps or procedures.

Content management system

While content management means a set of human processes for controlling data, a content management system (CMS) provides a software solution to facilitate the human control of the data. In the context of web information portals, a CMS is designed to create, publish and maintain the data records of a portal. In the case of metadata-driven web information portals, data records specifically refer to metadata descriptions of external online resources.

According to Arthur (2006), content management systems are primarily repositories that allow varying degrees of collaboration and management of portals in the creation and distribution process. Thus, a highly interactive user interface is required between distributed Internet resources and a centralised data repository. The literature suggests that the functionality of a CMS should be defined in relation to the user needs and the life cycle of the portal content to be managed (Addey et al. 2002; Browning and Lowndes 2001). As people are always drivers of content management processes, the implementation of a CMS has to take account of the human aspect (Addey et al. 2002).

In some cases workflow is suggested as one of the core functions of a CMS to ensure that the content "goes through an assessment, review or quality assurance process" (Browning and Lowndes 2001 p. 7). In practice, using a separate workflow system is popular to standardise human procedures for content management. However, the design or development of a workflow system is outside the scope of this research.

This section introduced the three determinants of RQ, namely domain experts, RQ assessment processes, and CMS, in the context of quality assured health information portals. In order to investigate the construct of resource quality and systematic ways to assess it, it is necessary to explore the theoretical basis and literature warrant for research quality found in information quality research literature in the information systems discipline, and in emerging quality control research in the field of consumer health informatics.

1.3.2 Information Quality and Quality Assessment Approaches

Defined as 'fitness for purpose/use' (Wang and Strong 1996), information quality (often interchangeable with data quality) is widely recognised as a multi-dimensional and contextual concept (Pipino et al. 2002; Strong et al. 1997). Information quality (IQ) needs to be defined and assessed for a specific reason. IQ can be assessed in a context of intended use of information (Katerattanakul and Siau 1999), or as what Shanks and Corbitt (1999) advocate in a context of its generation. Further, IQ has a dynamic nature as the context in which information is collected may change over time (Lee et al. 2004) and the perceptions of people who use the information may also vary (Strong et al. 1997).

In the literature, in order to facilitate quality measures, numerous IQ dimensions have been identified. Various definitions have been provided for their application in different contexts (Wang et al. 1995b). Taxonomies for classifying quality dimensions are also proposed (Eppler et al. 2003; Naumann 2002; Redman 1992; Shanks and Corbitt 1999; Wand and Wang 1996; Wang and Strong 1996). The most cited one is Wang and Strong's (1996) four category taxonomy, which defines quality information as intrinsically sound, contextually appropriate, clearly presented, and accessible to users. RQ as defined in this research only concerns the content quality of online information resources. Thus, the representational and accessibility quality dimensions are not considered when constructing RQ.

In addition to the constructs of IQ, a considerable number of quality assessment approaches have been proposed across multiple application domains. For instance, a considerable amount of effort has been spent on conceptual frameworks and methodologies for measuring and improving the quality of information or data in organisational database and information systems. Some of the significant works include the categorical IQ framework (Wang and Strong 1996), the semiotic data quality framework (Shanks and Corbitt 1999), the total data quality management methodology (Lee et al. 2006), the AIM quality methodology (Lee et al. 2002), and the IQ measurement methodology (Eppler and Muenzenmayer 2002a). Meanwhile in the field of consumer health informatics, extensive research has been done to assist information consumers in assessing health information on the Internet. Various standards, guidelines and quality tools have been developed, such as codes of conduct (HONcode 2007), quality checklists (DISCERN 2008), numerical rating instruments (Bomba and Land 2004; Gagliardi and Jadad 2002) and scoring systems (Currò et al. 2004).

Research on automated quality assessment for evaluating online health information resources (Griffiths et al. 2005; Wang and Liu 2007) is relatively new in the field of consumer health informatics. However, in the field of IQ research, the potential use of automation tools to assess the quality of online resources has been explored in a number of studies (Civan and Pratt 2006; Currò et al. 2004; Griffiths et al. 2005; Knight and Burn 2005; Price and Hersh 1999; Shankar and Watts 2003; Zhu and Gauch 2000). They all attempt to numerically rank web resources according to the level of compliance with specific evaluation criteria. A number of computational quality measures are developed based upon various quality metrics. A few ad hoc algorithms are developed with a high level of subjectivity to measure a subset of quality dimensions, such as completeness (Cappiello et al. 2004; Naumann 2002).

The above discussion indicates that a quality assessment approach for evaluating online information consists of a quality assessment framework, which defines the construct of IQ and its measures, and corresponding quality appraisal methods, either qualitative or computational for generating quality values. The adoption of user-sensitive design philosophy implies that understanding the information context, i.e. the needs and values of information consumers, is a prerequisite for defining the construct of RQ, and the framework for RQ assessment. However, existing IQ constructs, frameworks, and assessment approaches do not adequately address the subjective, contextual, and dynamic nature of IQ. They value the user context in which IQ should be assessed, but do not provide operational solutions on how to incorporate the user context in quality measurement. Hence, existing IQ constructs or assessment frameworks are insufficient to address user-sensitive RQ assessment issues for health information portal. It is necessary to explore how to incorporate user needs and values to make contextual value judgements on RQ in the context of health information portals.

1.3.3 Adoption of User-Sensitive Design for Resource Quality Assessment

Quality-assured information portals provide a promising solution to tackle both the quality and information overload problems in online information provision. Moreover, information portals can act as intelligent decision support systems (IDSS), given tailored information being provided to meet the knowledge and decision support needs of users (Burstein et al., 2006). For a portal that acts as an IDSS, understanding portal users and their information needs becomes a priority. What kind of information will portal users find interesting and useful? What factors will affect their quality perceptions and their use of retrieved information? Research in the relevant literature has revolved around how to incorporate users in the portal design processes and how to manage the portal on an ongoing basis from the perspectives of users (Zaphiris et al. 2006). One example is the Breast Cancer Knowledge Online portal (BCKOnline 2009), which explored the meaning of user-sensitive information provision for the breast cancer community. The resultant user-sensitive information portal is proposed as a promising model to reflect diverse user information needs and sensibilities (Burstein et al. 2006). It identifies, addresses, and reflects user information needs and values at all portal development stages.

The BCKOnline experience established the following set of principles for the design and development of user-sensitive health information portals (Evans et al. 2009, pp. 2-3):

- information is essential for informed decision-making and hence individual autonomy;
- information which leads to patient empowerment can also contribute to improved health outcomes;
- knowledge is contextual – its value to the individual is a dynamic interplay of personal, social, psychological, ethnic and cultural factors; and
- perspectives of people with direct and/or personal experience of a disease provide a valuable information resource and insight which both complements and enhances the scientific/biomedical view of the disease treatment, management and research.

The above principles require the development of socio-technical RQ assessment solutions for the context of health information portals. The implementation of such a user-sensitive approach involves user-centred design activities, such as understanding user needs, defining the intended context of use, specifying user requirements, producing design solutions, and evaluating designs against requirements from a user perspective (Bevan 1999).

In summary, in light of the user-sensitive design principles, this research values the important role the portal users play in constructing and assessing resource quality. It endeavours to adapt and integrate existing information quality constructs and assessment approaches, to conceptualise a user-sensitive RQ assessment approach that systematically measures RQ as the relationship between a resource and a user.

1.4 Research Approach

This thesis explores the research questions from a socio-technical point of view. The design science research paradigm in the information systems discipline (Hevner and Chatterjee 2010; Hevner et al. 2004) was adapted to a more suitable socio-technical research framework. Within this adapted design science research framework, mixed methodologies were employed in order to develop socio-technical solutions. Both the user-sensitive design methodology (McKemmish et al. 2009) and the systems development research methodology (Baskerville et al. 2009; Burstein 2002; Nunamaker et al. 1991) were adopted and integrated to govern a user-sensitive systems development process. The process consisted of three phases, namely concept building, system building, and system evaluation.

This research committed to contributing both new field knowledge and practice improvements through the three inter-connected research phases. The first concept building phase aimed at understanding the problem domain, articulating the design requirements, building theories that contribute to the conceptualisation of design artefacts, and finally conceptualising the RQ construct and user-sensitive RQ assessment solutions. Two health information portals provided a real-world context for investigating the practical problems associated with RQ assessment. These portals are the Breast Cancer Knowledge Online (BCKOnline) portal (www.bckonline.monash.edu.au) and the Heart Health Online (HHOnline) portal (www.sip.infotech.monash.edu.au/heart-portal/). These two portals represent the application of a metadata approach and user-sensitive design methodology in portal design and development. The RQ assessment processes, CMS systems, existing portal data, and development documentation of these two portals were analysed to obtain an in-depth understanding of what user-sensitive RQ assessment means in this specific context, what domain experts need for RQ assessment, and how to provide the required support. In addition, the BCKOnline portal also provided a concrete environment for prototyping the proposed RQ assessment solution.

In the system building phase, the feasibility of the proposed RQ assessment solution was tested through system prototyping. A resultant prototype system was developed as a concept demonstrator. The final research phase evaluated the utility and effects of the prototype system on the decision-making processes and outcomes of RQ assessment. Based on reflection on the systems development process and experience, the outcomes were consolidated to refine the proposed RQ assessment solution, as well as to suggest a transferable RQ assessment method for other domains.

1.5 Research Context

This research is related to a larger ARC funded Discovery project Smart Information Portals (SIP). In the SIP project, a metadata approach and user-sensitive design principles were adopted in developing the concept of a smart and user-sensitive information portal that (Burstein et al. 2006; McKemmish et al. 2009):

- provides a gateway to virtual, distributed knowledge repositories of relevant information;
- facilitates adaptivity and personalisation;
- provides additional value-added information to allow a user to make an informed decision about the quality of the information resource; and
- provides transparency in terms of portal features and functionalities.

The SIP project was identified as a suitable context for this PhD study to explore the research questions. The research draws attention to the quality assessment challenges associated with smart information portals and explores the issue from a user-sensitive point of view in a healthcare domain. Information needs and quality perceptions of health information consumers play a vital role in the conceptualisation of RQ and its assessment approach. The understanding of needs and values of health portal users was informed by the existing consumer health informatics literature and findings from the SIP project.

1.6 Significance and Contributions

This is the first major attempt in the fields of IQ research and consumer health informatics to study how user-sensitive design principles and intelligent technologies can be applied to conceptualise and implement a RQ assessment approach for metadata-driven health information portals. The research also aims to improve the scalability and sustainability of quality-assured content management for health information portals. The proposed approach was prototyped in a real-world context, which demonstrated the feasibility and efficacy of such an intelligent approach to supporting user-sensitive RQ assessment.

This research makes significant theoretical, methodological, and practical contributions. The central contribution of this study is that it defines what user-sensitive RQ assessment is, and how to support RQ assessment using intelligent technologies, at both conceptual and practical levels.

Contribution to the IQ theory: This research provides a user-sensitive RQ assessment framework, which extends the context-based IQ assessment theory (Strong et al. 1997; Wang and Strong 1996). The proposed framework integrates intrinsic and contextual IQ dimensions to construct the content quality of online health information. In this framework, the consensus understanding of the information needs and quality perceptions of health information consumers are characterised. The user model further informs the resource attribution. Perceived RQ is then measured as a relative and aggregated construct by specifying and mapping resource attributes to the user characteristics. By this means, contextual value judgements are standardised by explicit RQ assessment constructs, which reflect individual user circumstances.

Contribution to the design science research: This research addresses the problem of user-sensitive RQ assessment in health information portals. The artefact-centric design science research framework has been adapted to fit this socio-technical context using mixed methodologies. The user-sensitive design methodology and the systems development research methodology are integrated. The major adaptation to the canonical design science research framework is that this research involves two kinds of conceptualisation. One is to conceptualise the research problem, i.e. to investigate RQ assessment issues in the user context, and to develop a user-sensitive RQ assessment framework. The other one is to conceptualise socio-technical design artefacts based on the articulated new concepts.

Contribution to the quality control practice for online health information: This study elicits the design requirements of a quality tool for supporting user-sensitive RQ assessment in health information portals. The comprehensive quality indicator analysis and the instantiation of quality assessment metrics inform the design of semi-automated tools for facilitating resource attribute description. The developed prototype system provides greater functionality to support domain experts making informed decisions on RQ. The system demonstrates how intelligent learning techniques can be applied to describe the quality attributes of online resources. The system has been integrated as part of an operational health information portal.

1.7 Scope and Limitations

This research investigated the user-sensitive RQ assessment issue in the context of metadata-driven health information portals. It did not aim to capture general design requirements for portals developed using other approaches. Obviously, there are many other health information portals and medical websites, with different scopes and different approaches to delivering quality information. Large-scale health portals, such as Better Health Channel (www.betterhealth.vic.gov.au) and HealthInsite (www.healthinsite.gov.au), provide quality online information of multiple healthcare and disease domain targeting at a wide range of audiences.

These portals did not provide the same level of access, thus were not investigated in-depth. Due to a limited access to health information portals and eligible participants, a convenience sample (Marshall 1996) was used in both the exploratory case study and the prototype system evaluation.

In this research, the analyses of RQ assessment processes and domain expert needs were based on the data collected from the BCKOnline and HHOnline portals. They both are disease-specific and community-based health information portals. The selected two metadata-driven and user-sensitive portals are relatively small scale compared to the large scale health information portals, but they provided a rich context for the exploratory study. There is a small group of domain experts associated with these two portals. Most of them participated in the case study and the prototype system evaluation conducted by this research. The recruited domain expert participants were specialised in different areas. Rich data were collected that enabled the comparison of perspectives from different domain expertise backgrounds. The constructive data collection procedures and instruments ensured that the quality of collected data was satisfactory.

Limitations of this research mainly lie in the work conducted in the system building and system evaluation phases. The conceptualisation of RQ assessment was concerned with issues of *Reliability* and *Relevancy*. However, constrained by the limited work a PhD thesis research could undertake, only one of these two aspects was selected for prototyping and evaluation. As a result, the construct and measure of the *Reliability* dimension was explored in depth while further study of the *Relevancy* dimension of RQ is a matter for future research.

1.8 Thesis Outline

The structure of this thesis is depicted in Figure 1.2. The diagram illustrates the relationships between the eight thesis chapters and their role in addressing the central research question. A brief overview of the remaining chapters is also provided.

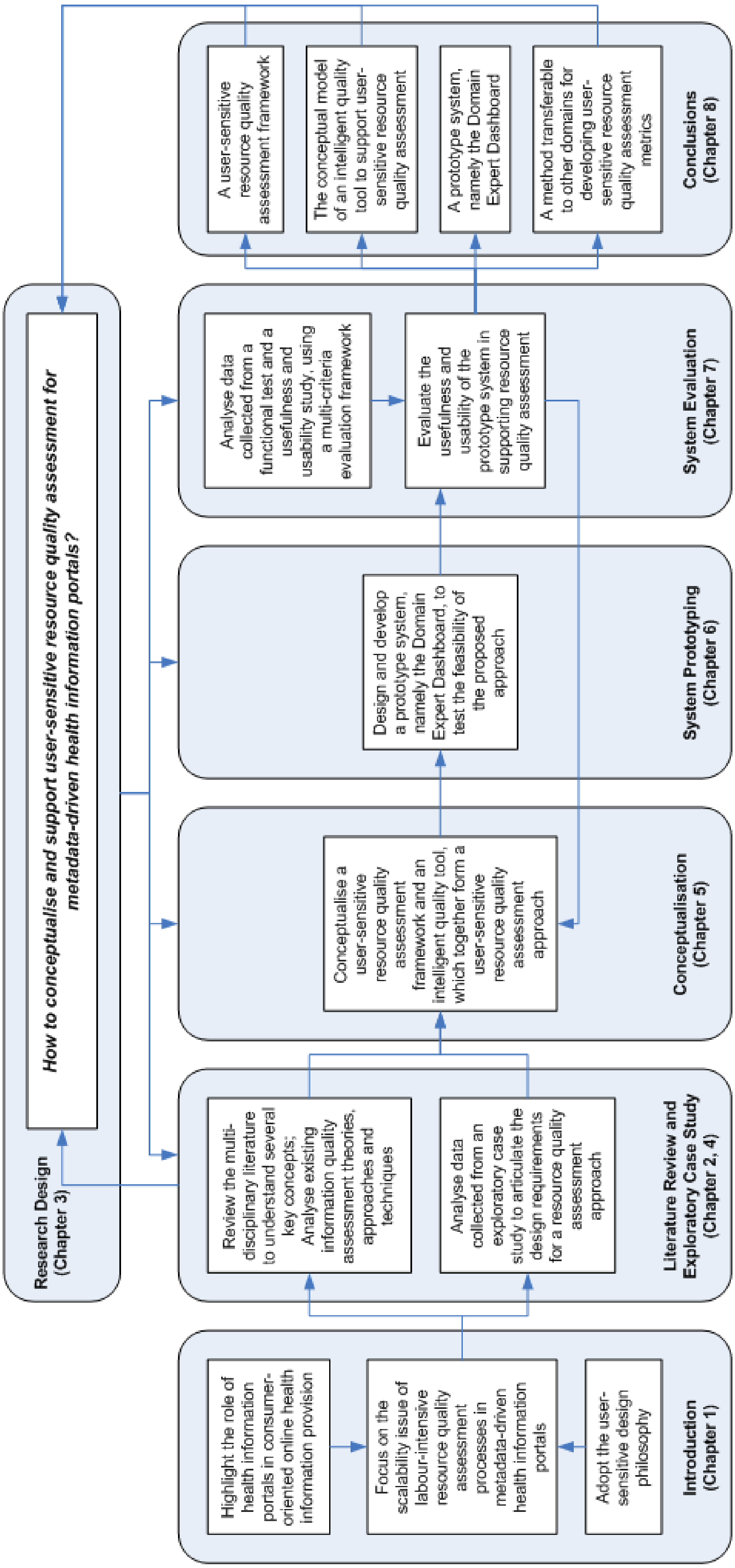


Figure 1.2 Thesis outline

In **Chapter 2**, multi-disciplinary concepts, constructs, methods and techniques are analysed and synthesised in relation to information quality and its assessment. The chapter reviews the relevant literature from theoretical, contextual, and technological perspectives. The analysis indicates that existing information quality constructs and assessment solutions do not sufficiently address RQ assessment requirements in the context of health information portals.

Chapter 3 presents how this research was conducted and evaluated within a socio-technical design science research framework. Mixed research methodologies were employed to deliver both new field knowledge and practice improvement. The user-sensitive design methodology and the systems development research methodology were combined to approach the research problem. The concept of user-sensitive RQ assessment was investigated through a use-sensitive systems development process of concept building, system building, and system evaluation.

Chapter 4 describes a qualitative and exploratory case study, in which multi-faceted domain expertise, RQ assessment processes, and domain expert needs are analysed and articulated in order to elicit the functional requirements of a portal CMS, as well as the design requirements of a user-sensitive RQ assessment approach.

Chapter 5 presents the development of a conceptual model of a user-sensitive RQ assessment framework for health information portals. Based on the proposed quality framework, an intelligent quality tool is conceptualised to support the decision-making processes undertaken by domain experts on RQ. The chapter also explores available intelligent technologies for their application in providing required functionality to support RQ assessment.

Chapter 6 investigates the feasibility of the proposed user-sensitive RQ assessment approach through system prototyping. The BCKOnline portal provides a live portal environment for implementing and evaluating a prototype system. The chapter introduces the resulting prototype system, namely the Domain Expert Dashboard. The system is characterised by a number of enhanced system functions and intelligent features to assist domain experts in RQ assessment. It also demonstrates how machine learning techniques can be used to describe RQ.

In **Chapter 7**, the usefulness and usability of the developed prototype system on RQ assessment processes and outcomes are evaluated. The evaluation consists of an internal functional test and a usefulness and usability study with domain experts. The evaluation results indicate areas for refining the tool architecture.

Chapter 8 reviews the overall socio-technical design science research process undertaken to investigate the user-sensitive RQ assessment issue for the healthcare domain. The chapter summarises how the research findings and outcomes address the research questions. Based on a consolidation of the research experience, a transferable method is proposed for developing user-sensitive RQ assessment metrics for other domains. In the chapter, the theoretical, methodological, and practical contributions of this research are also discussed. The adaptation of Hevner et al.'s (2004) seven design science research guidelines to this research is elaborated, which justifies the quality of this applied, socio-technical design science research. This thesis ends with suggestions relating to directions for future research.

1.9 Chapter Summary

This research was motivated by addressing labour intensive RQ assessment issues in the quality control practice for online health information. User-sensitive RQ assessment is the focal concept of this thesis. The central research question was how to conceptualise and support user-sensitive RQ assessment for metadata-driven health information portals. The question reflected two major objectives of this research, which were to conceptualise RQ and its assessment taking account of user information needs and quality perceptions, and to support user-sensitive RQ assessment with the use of intelligent technologies.

The chapter introduced the literature background, outlining the key concepts that will be used all through this thesis. The research approach was briefly discussed, in terms of how the research questions were explored via a user-sensitive systems development research process, within an adapted socio-technical design science research framework. It then defined the research context, scope, and limitations before a summary of research significance and contributions. Finally, the organisation of thesis chapters was presented, which illustrated the whole research process and the role each chapter plays in addressing the research questions.

C h a p t e r   2

2 Literature Review

This chapter presents an analytical review of the literature, considering various disciplines that are relevant to user-sensitive resource quality assessment issues investigated in this research. First, the chapter introduces the scope of the literature analysis from the theoretical, contextual, and technological perspectives. It reviews information quality theories and existing quality assessment approaches. Existing information quality constructs, frameworks, methodologies, and measurement techniques are synthesised for understanding the concepts of information quality and quality assessment. The chapter then discusses the quality issue of online health information, and introduces quality control perspectives and approaches in the field of consumer health informatics. Finally, the essential role of metadata in providing quality information in the context of web information portals is introduced before the chapter summary.

2.1 Scope of the Literature Analysis

This research was driven by the analysis of four streams of the literature, namely:

1. **Information Quality:** The field which provides the theoretical foundation for defining, constructing, and measuring information quality, as well as for developing quality solutions to address the specific quality assessment issue concerned by this research.
2. **Consumer Health Informatics:** A branch of medical informatics that analyses, models, and integrates consumer information needs and preferences into health information systems (Eysenbach 2000). It provides a rich environment for investigating the quality assessment issue of online health information from a consumer-centric point of view.
3. **Digital Library:** A branch of library and information science that comprehensively collects, manages, preserves, and delivers rich digital resources of measurable quality according to codified policies (Candela et al. 2007). The field studies the role that human curators play in selecting, cataloguing, and describing digital resources. It is also the field where multi-functional metadata are heavily used to manage resource collections.
4. **Web Information Retrieval:** A field concerned with searching for web-based information resources, and information within or about resources. It is the field from which online search solutions are derived, and also a rich area for the application of metadata and emerging intelligent technologies.

The selection of the research fields for literature review was based on the following consideration. First, this research recognises that information quality literature provides the theoretical foundation and methodologies for understanding information quality and generic requirements for developing quality assessment solutions. Second, the field of consumer health informatics helps this research to understand the problems, approaches, and stakeholders of consumer-oriented online health information provision. It defines the healthcare context, in which the research questions are formulated and explored. Finally, the literature of digital library and web information retrieval describes the technological aspect of the research problem. The attention was drawn on the essential role metadata play in providing quality online information in the context of web information portals. The following sections of this chapter analyse the relevant literature from the theoretical, contextual, and technological aspects respectively.

2.2 Information Quality Assessment

Information or data quality research is a well-established field in the Information Systems (IS) discipline. The field provides theories for constructing and assessing information quality in a systematic way. The difficulties and challenges associated with information quality assessment and improvement have been identified and described extensively (Lee et al. 2006; Madnick et al. 2009). Numerous solutions have been proposed to solve various quality problems in different application domains, such as organisational database, management information systems, and web systems (English 2002; Eppler and Muenzenmayer 2002a; Lee et al. 2002). The main purpose of this investigation is to obtain an in-depth understanding of information quality and the ways for assessing it. The development of new quality assessment solutions needs to build on, or extend existing quality constructs, frameworks, and assessment approaches. Therefore, in order to understand what user-sensitive resource quality assessment means for metadata-driven health information portals, it is important to define information quality per se, and discuss quality assessment frameworks, methods, and techniques.

2.2.1 Information Quality

Information quality and data quality are two terms used interchangeably when information or data quality problems are discussed in various research settings. Although the meaning of information is often richer than raw data, for the purpose of conciseness, this thesis uses information quality (IQ) to describe all quality issues and corresponding research literature associated with either data or information in their original works. IQ generally means the degree of its usefulness or "fit for use" (Wang and Strong 1996), which implies that information considered appropriate for one use may not contain enough attributes to satisfy another use (Tayi and Ballou 1998). Likewise, Wand and Wang (1996) define quality as mapping the state of an information system to the state of a real world. Strong et al. (1997) also contend that IQ cannot be assessed independent of the user context.

The notion of IQ as "fitness for use/purpose" introduces a task-oriented view of IQ, i.e. the ability of a piece of information to meet the requirements of a specific task at hand. On the other hand, Redman (1992) formally defines quality by the following statement: "A product, service, or datum X is of higher quality than product, service, or datum Y if X meets customer needs better than Y" (p. 5). This definition forms the foundation of a user-oriented view, which places great emphasis on user needs, expectations or preferences for quality control and management (Bevan 1999; Evans and Lindsay 2005). In a more recent work, Eppler (2006) adopts these two views, and describes IQ as a duality concept that needs to meet both the user expectations (subjective) and activity requirements (objective). The consensus of these definitions is that IQ is a contextual concept and needs to be assessed for a purpose, for example to meet specific information or decision support needs of users.

The IQ definitions discussed above reveal the contextual nature of IQ, but in order to perform the IQ assessment, the construct of IQ needs to be explicitly defined. It is widely recognised that IQ is a multi-dimensional and in some cases hierarchical concept (Klein 2001; Lee et al. 2002; Wang et al. 1995a). IQ dimensions or criteria are usually used to describe components of an IQ construct. Following the definition of Wang and Strong (1996), IQ dimensions are a set of quality attributes that represent a single aspect or construct of IQ. A number of IQ dimensions have been defined and categorised, and used in operational quality measurement models in different research settings (Batini 2006; Cappiello et al. 2003b; Naumann 2002; Wang and Strong 1996). IQ dimensions such as *Accuracy*, *Precision*, *Completeness*, *Currency*, *Trustworthiness*, and the *Reputation* of data source, are widely covered in theoretical and experimental studies (Berti-Équille 2007). However, there is always a terminology issue when comparing the labels and definitions of published IQ dimensions in different IQ measurement models. Due to the lack of sub-dimensional descriptions, two IQ dimensions with the same label could possibly describe different aspects of IQ or be associated with different quality characteristics.

Moreover, IQ is also recognised as a domain-specific concept. There is no 'gold' quality measurement model, and not all identified quality dimensions can be applied across multiple application domains (Gertz et al. 2004). It is suggested that "a detailed characterization of data quality should take into account the peculiarities of the specific domain" (Batini 2006, p. 222).

2.2.2 IQ Assessment Frameworks

In the literature, issues involved in quality assessment have been discussed extensively across multiple domains. A considerable number of IQ assessment frameworks have been proposed (Eppler and Wittig 2000; Madnick et al. 2009; Wang et al. 1995b). According to Eppler and Wittig (2000), an IQ framework needs to provide a systematic and concise set of criteria for IQ assessment. The framework also needs to help with the identification of IQ problems, and define constructs to enable IQ measurement. A comprehensive definition is given by Stvilia et al. (2007, p. 1722), who define an IQ assessment framework as "a multidimensional structure consisting of general concepts, relations, classifications, and methodologies that could serve as a resource and guide for developing context-specific IQ measurement models". In this section, IQ frameworks developed for organisational information systems or the web environment are synthesised. The purpose is to understand IQ constructs, identify common IQ dimensions, and compare IQ assessment metrics.

2.2.2.1 IQ frameworks for organisational information systems

Traditional IQ research originated from the field of organisational systems. There appears to be a wealth of IQ literature in organisational systems. Numerous IQ assessment frameworks, conceptual models, and methodologies are proposed in order to comprehensively assess, benchmark, and improve IQ in organisational databases or management information systems, such as the InfoQual framework (Price and Shanks 2004; Price and Shanks 2005a), the Comprehensive Data Quality (CDQ) methodology (Batini et al. 2006), and the AIM Quality (AIMQ) methodology (Lee et al. 2002). Quality dimensions such as accuracy, precision, completeness, currency, trustworthiness, non-duplication, and data source reputation are widely covered throughout the theoretical and experimental studies in the literature (Berti-Équille 2007). Table 2.1 summarises major IQ frameworks frequently cited in the organisational IQ literature from 1992 until recently. Terms used in the table are drawn from the original works.

Table 2.1 Summary of IQ frameworks for organisational information systems

Author and Year	**Summary of the framework**	**Construct**	**Application Domain**
(Zeist and Hendriks 1996)	Extended ISO model: 6 quality characteristics; 32 sub-characteristics	<u>Functionality</u> Suitability, Accuracy, Interoperability, Compliance, Security, Traceability <u>Reliability</u> Maturity, Recoverability, Availability, Degradability, Fault tolerance <u>Efficiency</u> Time behaviour, Resource behaviour <u>Usability</u>	Software quality

		Understandability, Learnability, Operability, Luxury, Clarity, Helpfulness, Explicitness, Customisability, User-friendliness Maintainability Analysability, Changeability, Stability, Testability, Manageability, Reusability Portability Adaptability, Conformance, Replaceability, Installability	
(Wang and Strong 1996)	TDQM (Total Data Quality Management): 4 categories; 16 dimensions	Intrinsic Accuracy, Objectivity, Believability, Reputation Accessibility Accessibility, Security Contextual Relevancy, Value-Added, Timeliness, Completeness, Amount of Info Representational Interpretability, Ease of Understanding, Concise Representation, Consistent Representation	DQ in organisational information systems
(Shanks and Corbitt 1999)	Semiotics-based DQ framework: 4 semiotic descriptions; 4 goals of DQ; 11 dimensions	Syntactic: Consistent Well-defined/formal syntax Semantic: Complete and Accurate Comprehensive, Unambiguous, Meaningful, Correct Pragmatic: Usable and Useful Timely, Concise, Easily Accessed, Reputable Social: Shared understanding of meaning Understood, Awareness of Bias	DQ in organisational information systems
(Dedeke 2000)	Conceptual framework for measuring IS quality: 5 quality categories; 28 dimensions	Ergonomic Quality Confortability, Learnability, Visual signals, Audio signals Accessibility Quality Technical access, System availability, Technical security, Data accessibility, Data sharing, Data convertibility, Ease of Navigation Transactional Quality Controllability, Error tolerance, Adaptability, System feedback, Efficiency, Responsiveness Contextual Quality Value added, Relevancy, Timeliness, Completeness, Appropriate data Representation Quality Interpretability, Consistency, Conciseness, Structure, Readability, Contrast	Quality measure of information systems
(Leung 2001)	Model adapted from Zeist & Hendriks's (1996) extended ISO Model: 5 quality characteristics; 25 sub-	Functionality Accuracy, Security, Traceability Reliability Maturity, Recoverability, Availability, Degradability, Fault tolerance Efficiency Time behaviour, Resource behaviour Usability	Software quality for Intranet applications

	characteristics	Understandability, Learnability, Operability, Luxury, Clarity, Helpfulness, Explicitness, Customisability, User- friendliness <u>Maintainability</u> Analysability, Changeability, Stability, Testability, Manageability, Reusability	
(Naumann 2002)	Classification of IQ metadata criteria: 3 assessment classes; 22 IQ criteria	<u>Subject-Criteria</u> Believability, Concise representation, Interpretability, Relevancy, Reputation, Understandability, Value-Added <u>Object-Criteria</u> Completeness, Customer Support, Documentation, Objectivity, Price, Reliability, Security, Timeliness, Verifiability <u>Process-Criteria</u> Accuracy, Amount of data, Availability, Consistent representation, Latency, Response time	IQ in integrated information systems
(Kahn et al. 2002)	PSP/IQ model that maps IQ dimensions to product and service performance: 2 quality types; 4 IQ classifications; 16 IQ dimensions	<u>Product Quality - Sound Information</u> Free-of-Error, Concise, Representation, Completeness, Consistent Representation <u>Product Quality - Useful Information</u> Appropriate Amount, Relevancy, Understandability, Interpretability, Objectivity <u>Service Quality - Dependable Information</u> Timeliness, Security <u>Service Quality - Useable Information</u> Believability, Accessibility, Ease of Manipulation, Reputation, Value-Added	Product and service quality
(Cappiello et al. 2003a)	Personalised data quality assessment model that incorporates user requirements to measure data relevance: 4 categories; 23 dimensions	<u>Subject</u> Interpretability, Ease of understanding, Concise Representation, Accessibility <u>Object</u> Believability, Accuracy, Objectivity, Reputation, Representational consistency, Internal Consistency, Data Completeness <u>Architectural</u> Availability, Responsiveness, Source Availability, Source Responsiveness <u>Process</u> Relevancy, Timeliness, Appropriate amount of data, Process Completeness, Value-added, Access Security, history, Cost	DQ in cooperative information systems
(Price and Shanks 2005b)	A refined semiotic IQ framework, based on Price and Shanks (2004): 3 categories; 16 criteria	<u>Syntactic (based on rule conformance)</u> Conforming to metadata <u>Semantic (based on external correspondence)</u> Mapped completely, Mapped unambiguously, Phenomena mapped correctly, Properties mapped correctly, Mapped consistently, Mapped meaningfully <u>Pragmatic (use-based consumer perspective)</u> Accessible, Suitably presented, Flexibly presented, Timely, Understandable, Secure, Type-sufficient, Allowing access to relevant	IQ in organisational information systems

		metadata, Perceptions of syntactic and semantic criteria defined earlier	
(Stvilia et al. 2007)	A general IQ assessment framework: 3 categories; 22 dimensions	<u>Intrinsic</u> Accuracy/Validity, Cohesiveness, Complexity, Semantic Consistency, Structural Consistency, Currency, Informativeness/Redundancy, Naturalness, Precision/Completeness <u>Relational/Contextual</u> Accuracy, Accessibility, Complexity, Naturalness, Informativeness/Redundancy, Relevance (Aboutness), Precision/Completeness, Security, Semantic Consistency, Structural Consistency, Verifiability, Volatility <u>Reputational</u> Authority	IQ in organisational information systems
(Berti-Équille 2007)	A classification of DQ dimensions: 4 categories; 46 dimensions	<u>Quality of the management of data by the system</u> Accessibility, Ease of maintenance, Reliability <u>Quality of the representation of data in the system</u> Conformance to schema, Appropriate presentation, Clarity <u>Intrinsic DQ</u> Accuracy, Uniqueness, Consistency <u>Relative DQ</u> User preferences (dependent on the user); Criticality, Conformance to business rules (dependence on the application); Variability, Volatility, Freshness (time-dependent); Data source reputation, Verifiability (dependent on a given knowledge-state)	DQ in data warehousing

Amongst these IQ frameworks, Wang and Strong's (1996) four-category conceptual framework is the most influential IQ framework that has been widely applied and adapted across multiple domains. The framework was developed based on a comprehensive survey study with information consumers. Four categories are proposed to group IQ dimensions according to their nature in regards to IQ:

- Intrinsic IQ denotes that information has quality in its own right;
- Contextual IQ highlights the requirement that IQ must be considered within the context of the task at hand;
- Accessibility IQ emphasises that information must be easily accessible and secure;
- Representational IQ comprises aspects related to the format and meaning of information.

It can be seen that the intrinsic and contextual IQ dimensions are related to the information per se and its usefulness, whilst the representational and accessibility IQ dimensions emphasise the importance of the system that conveys information.

According to this four-category conceptual framework, the authors define high-quality information/data as "intrinsically good, contextually appropriate for the task, clearly represented, and accessible to the data consumer" (p. 22). Over years, the constructs of these categories have been tested and adapted in a number of studies. The original and the exemplary adapted constructs of the four IQ categories are outlined in Table 2.2.

Table 2.2 Original and adapted constructs of Wang and Strong's (1996) four IQ categories

Category	**Wang and Strong (1996)** (the original construct)	**Knight (2009)** (adapted for the context of web information use)	**Katerattanakul and Siau (1999)** (adapted for measuring web-based IQ)	**Ying and Al-Hakim (2010)** (adapted for measuring IQ in ubiquitous healthcare services)
Intrinsic	Accuracy; Objectivity; Believability; Reputation	Accuracy; Objectivity; Believability; Reliability	Accuracy of the content; Accurate, workable, and relevant hyperlinks	Believability; Reputation; Traceability
Contextual	Relevancy; Value-added; Timeliness; Completeness; Appropriate amount of data	Relevancy; Currency; Uniqueness; Scope/Depth	Provision of author's information	Relevancy; Timeliness; Completeness; Appropriate amount of data
Representa-tional	Interoperability; Ease of understanding; Concise representation; Representational consistency	Completeness; Understandability; Conciseness; Consistency	Organisation, visual settings, typographical features, and consistency; Vividness and attractiveness; Confusion of the content	Interoperability; Ease of understanding; Concise/Consistent
Accessibility	Accessibility; Access security	Accessibility; Security; Usability; Efficiency	Navigational tools provided	Accessibility; Security

Based on a comprehensive investigation on users' attitudes and perceptions of IQ in the context of web-based information interaction and retrieval (information use), Knight (2009) proposes a conceptual model that confirms Wang and Strong's (1996) four IQ categories, but modifies the original model by adding or removing IQ dimensions. The intrinsic IQ dimension of *Reputation*, which is considered as implying information integrity (Bovee et al. 2003), but lacking a construct for quantifiable measure, has been replaced by *Reliability*. The accessibility IQ category includes two additional dimensions of *Usability* and *Efficiency*. Moreover, the adapted model also moves the *Completeness* dimension from the contextual IQ category into the representational IQ category. Representational IQ in this model has been defined as "the tangible representation of interaction between information and information producer, and the cognitive interaction between information and information receiver" (Knight 2009, p. 272).

Katerattanakul & Siau (1999) adapt the four IQ categories in the context of IQ measure for individual websites. IQ dimensions in Wang and Strong's original model are instantiated for their application in the web context. The intrinsic category refers to accurate and free-of-error page content, as well as accurate, workable, and relevant hyperlinks on a webpage. The contextual IQ category is concerned with the provision of authorship information. Representational IQ refers to a number of structural and visual features of a webpage, while accessibility IQ concerns the use of navigation tools to browse a website.

For the application in the healthcare domain, Ying and Al-Hakim (2010) also adapt the four IQ categories, but introduce much less variation to the original model. In order to define a quantifiable measure for each IQ dimension, intrinsic IQ dimensions of *Accuracy* and *Objectivity* are replaced by *Traceability*, and contextual IQ dimension *Value-added* is removed from the adaptive model.

The IQ studies discussed above illustrate the application of Wang and Strong's (1996) four IQ categories in addressing quality assessment issues for different domains. Similarly, these categories can be used to inform the selection of quality dimensions for assessing web-based information resources in a portal context.

Furthermore, according to Knight (2009), intrinsic IQ is the most important category to user IQ perceptions, followed by representational IQ, and then contextual IQ. Accessibility IQ is the category least concerned by users. This finding indicates that the value of the contextual IQ aspect is less recognised by users, comparing to the value of the representational IQ. However, information being well presented does not necessarily mean that it has good quality. On the contrary, if users have difficulty in judging the relevance or the timeliness of the information, or totally ignore these aspects, the information quite likely will be misused. Therefore, it is necessary to both define and disclose contextual IQ measure to users. Considering information's actual usefulness and fitness to user needs, the intrinsic and contextual IQ dimensions are regarded more important than the representational and accessibility IQ dimensions.

2.2.2.2 IQ frameworks for the web

Apart from the conventional organisation environment, the quality of online information has been raised as a critical issue, considering the increasing coverage and popularity of the Internet (Alexander and Tate 1999; Eppler and Muenzenmayer 2002a; Rieh 2002). For instance, IQ is regarded as a key factor that can influence consumers' use of online services delivered via e-commerce websites (Barnes and Vidgen 2001; Barnes and Vidgen 2002). Due to the dynamic and heterogeneous nature of the Internet, IQ associated with online information has introduced new issues and challenges for IQ research to investigate. Typical IQ problems in the web environment include outdated, inconsistent or inaccessible information, obsolete web links, and web design issues. As web information systems deal with the integration of semi-structured data from different sources, information retrieved from those systems is often found to be irrelevant, replicated, outdated, or conflicting. In order to address these specific IQ issues on the web, traditional IQ assessment frameworks are adapted to this new environment and in tandem with this change. New frameworks and models are proposed. Table 2.3 summarises some major works of IQ assessment in the context of the web, using the terms drawn from the original works.

Table 2.3 Summary of IQ frameworks for the web

Author and Year	Summary of the framework	Construct	Application Domain
(Alexander and Tate 1999)	Applying quality framework to the web: 6 criteria	<u>Authority</u> Validated information, Author is visible <u>Accuracy</u> Reliable, Free of errors <u>Objectivity</u> Presented without personal biases <u>Currency</u> Content up-to-date <u>Orientation</u> Clear target audience <u>Navigation</u> Intuitive design	IQ on the web
(Katerattanakul and Siau 1999; Katerattanakul and Siau 2002)	Adapted from Strong et al. (1997) to measure IQ of websites from the perspective of web users: 4 quality	<u>Intrinsic IQ</u> Accuracy and errors of the content; Accurate, workable, and relevant hyperlinks <u>Contextual IQ</u> Provision of author's information <u>Representational IQ</u> Organisation, visual settings, typographical features, and consistency; Vividness and	IQ of personal websites and e-commerce

	categories; 7 dimensions	attractiveness; Confusion of the content Accessibility IQ Navigational tools provided	
(Zhu and Gauch 2000)	Data quality metrics for centralised/distributed information retrieval: 6 assessment classes with corresponding quality metrics	Currency Measured as the time stamp of the last modification of the document Availability calculated as the number of broken links on a page divided by the total numbers of links it contains Information-to-noise ratio Computed as the total length of the tokens after pre-processing divided by the size of the document Authority based on the Yahoo Internet Life (YIL) reviews, which assigns a score, ranging from 2 to 4, to a reviewed site Popularity Number of links pointing to a webpage, used to measure the popularity of the webpage Cohesiveness Determined by how closely related the major topics in the webpage are	IQ for web information retrieval
(Barnes and Vidgen 2000; Barnes and Vidgen 2001; Barnes and Vidgen 2002)	WebQual instrument for assessing website quality. Initial version: 5 categories; 10 dimensions 24 questions Refined version: 3 categories 22 questions	Tangibles Aesthetics, Navigation Reliability Reliability, Competence Responsiveness Responsiveness, Access Assurance Credibility, Security Empathy Communication, Understanding the individual The refined WebQual 4.0 instrument contains 22 questions grouped in 3 categories: Usability, Information, Service Interaction	Quality of website
(Klein 2001; Klein 2002)	IQ dimensions selected from Wang and Strong's (1996) categorical IQ model: 5 IQ dimensions with corresponding preliminary factors	Accuracy Discrepancy, Timeliness, Source/Author, Bias/Intentionally false information Completeness Lack of depth, Technical problems, Missing desired information, Incomplete when compared with other sites, Lack of breadth Relevance Irrelevant hits when searching, Bias, Too broad, Purpose of website Timeliness Information is not current, Technical problems, Publication date is unknown Amount of Data Too much or too little information, or information is unavailable	IQ on the web

(Eppler and Muenzenmayer 2002b)	IQM (Information Quality Measurement) methodology: 2 quality types; 4 categories; 16 dimensions	Content Quality: Relevant Information Comprehensive, Accurate, Clear, Applicable Content Quality: Sound Information Concise, Consistent, Correct, Current Media Quality: Optimized Process Convenient, Timely, Traceable, Interactive Media Quality: Reliable Infrastructure Accessible, Secure, Maintainable, Fast	IQ of web pages
(Liu and Huang 2005)	A framework for credibility assessment: 6 evaluation criteria	Source, Content, Format and presentation, Currency, Accuracy, Speed	IQ on the web
(Custard and Sumner 2005)	A model for computing RQ: 5 categories; 16 quality indicators	Provenance Cognitive authority, Site domain Description Element count, Description length, Metadata currency Content Resource currency, Advertising, Alignment, Word count, Image count, Link count, Multimedia Social authority Google's PageRank, Annotations Availability Cost, Functionality	RQ in educational digital libraries
(Caro et al. 2006)	The PDQM (Portal Data Quality Model) for assessing the portal data, based on (Strong et al. 1997): 11 web portal functionalities; 6 categories of data consumer expectations	Web Portal Functionalities Data Points and Integration, Taxonomy, Search Capabilities, Help Features, Content Management, Processes and Actions, Communication and Collaboration, Personalization, Presentation, Administration, Security Data Consumer Expectations Privacy, Content, Quality of Values, Presentation, Improvement, Commitment	Quality of web portals
(Knight 2009)	CC/LC (Combined Conceptual Life-Cycle) model of IQ, adapted from Wang and Strong's (1996) categorical IQ model: 4 quality categories; 16 dimensions	Intrinsic IQ Accuracy, Objectivity, Believability, Reliability Contextual IQ Relevancy, Currency, Uniqueness, Scope/Depth Representational IQ Completeness, Understandability, Conciseness, Consistency Accessibility IQ Accessibility, Security, Usability, Efficiency	IQ for web information retrieval

Although multiple frameworks adopt different set of IQ dimensions, when looking closely, the terminology issue is detected. It is noticed that two IQ dimensions defined in different frameworks, even termed differently, can have a similar construct for actual measure. On the other hand, the same IQ dimension can be measured differently.

For instance, Alexander and Tate (1999) propose an IQ framework for the web context, which includes six dimensions of *Authority*, *Accuracy*, *Objectivity*, *Currency*, *Orientation*, and *Navigation*. *Authority* refers to information being validated and the author of the webpage being visible. *Accuracy* refers to a webpage being reliable and free of errors. *Currency* refers to the web content being up-to-date. Likewise, Klein (2002) proposes a framework of five IQ dimensions, including *Accuracy*, *Amount of data*, *Completeness*, *Relevance*, and *Timeliness*. In this framework, *Accuracy* is defined as the source and author of the information on a webpage being available. It uses a measure similar to the *Authority* definition in Alexander and Tate's (1999) framework. Besides, Klein defines *Timeliness* as the information on a webpage being current, or the publishing date of a webpage being disclosed. This definition is more related to the dimension of *Currency*.

The terminology issue implies that certain IQ dimensions have been used interchangeably in the literature. More likely, those IQ dimensions are not mutually exclusive to each other, or can be represented in a hierarchy. The analysis also indicates that without an explicit definition of sub-dimensions or measurable construct, the meaning of an IQ dimension can be ambiguous. Thus, the IQ dimensions and their constructs can be easily misinterpreted or mismeasured.

2.2.2.3 Common IQ dimensions

IQ frameworks as outlined in Table 2.1 and Table 2.3 present diverse IQ constructs defined for serving different assessment purposes in targeted application domains. Based on these frameworks, a set of common dimensions is identified. These dimensions are worthy of consideration for development of new IQ constructs that are used to evaluate online health information. These common IQ dimensions are listed in Table 2.4 with their definitions provided.

Table 2.4 Common information quality dimensions

Dimension	**Definitions** [1] adapted from Wang & Strong (1996) [2] adapted from Knight and Burn (2005)
Accessibility (Availability)	extent to which information is available, downloadable, or easily and quickly retrievable.[1]
Accuracy (Evidence-based)	extent to which information is correct, reliable and certified free of error. [1]
Appropriateness	extent to which the quantity or volume of available data is appropriate. [1]

(Amount of data)	
Believability (Credibility)	extent to which information is regarded as true and credible. [1]
Completeness	extent to which information is not missing and is sufficient. [1]
Consistency	extent to which information is presented in the same format. [1]
Concise	extent to which information is compactly represented without being overwhelming (i.e. brief in presentation, yet complete and to the point).[1]
Efficiency	extent to which data are able to quickly meet the information needs. [1]
Navigation	extent to which data are easily found and linked to. [2]
Objectivity	extent to which information is unbiased, unprejudiced and impartial.[1]
Relevancy	extent to which information is applicable and helpful.[1]
Reliability	extent to which information is correct and reliable.[1]
Reputation	extent to which information is highly regarded in terms of source or content. [1]
Security	extent to which access to information is restricted appropriately to maintain its security.[1]
Timeliness (Currency)	extent to which the information is sufficiently up-to-date.[1]
Usability	extent to which information is clear, and easily used.[2]
Understandability	extent to which data are clear without ambiguity, and easily comprehended.[1]
Usefulness	extent to which information is applicable and helpful.[1]
Value-added	extent to which information is beneficial, and provides advantages from its use.[1]

In the context of web information use, Knight (2009) investigates the user's perception in regards to the importance of these IQ dimensions. It is concluded that the surveyed user group values the intrinsic IQ dimensions the most. These intrinsic IQ dimensions include *Reliability*, *Objectivity*, *Accuracy*, and *Believability*. Representational and accessibility IQ dimensions, such as *Usability*, *Uniqueness*, *Accessibility*, and *Security*, are ranked the lowest. This research does not examine each common criterion in detail, but selects a set of exemplary IQ dimensions for their importance and relevance to health information consumers.

Reliability is named as the most important intrinsic IQ dimension for users (Knight 2009). It has been connected to past experience, and regarded as one of the trust-related dimensions (Bailey et al. 2001). The construct of *Reliability* implies other IQ dimensions such as *Authorship* (Civan and Pratt 2006), *Reputation* (Keast et al. 2001), and *Authority* (Conrad et al. 2008; Lankes 2008; Rieh 2002). Besides, according to Knight (2009), *Reliability* also denotes the presence of IQ dimensions, such as *Objectivity*, *Accuracy* and *Believability*, as "without these characteristics, information would be considered, by the discerning recipient, to be unreliable" (p. 243).

Accuracy is regarded as one of the major intrinsic attributes of information. In the context of organisational databases, *Accuracy* is often used synonymously with data quality. In the context of the web, *Accuracy* is just one aspect of the overall IQ, but has the same importance as for conventional information systems. Incorrect information would be useless and harmful to decision makers (Shanks and Price 2008).

Believability describes the credibility of information, and is used interchangeably with the *Credibility* dimension. According to Naumann (2002), the main source for measuring *Believability* is the identity of the author or the information creator, which refers to the characteristic of *Authorship*. The dimension is also intrinsically linked to *Accuracy* and *Objectivity* (Michnik and Lo 2009), which constitute the concept of information *Reliability*.

Currency is the degree to which information is up-to-date, relative to the information task being performed. The dimension is usually measured by identifying dates of creation, modification, or last updating. However, according to McKemmish et al. (2009), these fields do not fully capture the concept of *Currency*. For instance, in the context of online health information provision, some materials might be years old, but their efficacy as treatment protocols, or their representation as best practice may still apply. Also, materials about personal reactions to the experience of serious illness can be timeless. Hence, the authors argue that the concept of *Currency* is not adequately covered by any of existing constructs.

Relevancy in general means the relation of information to the matter at hand. Given the abundant information available online, *Relevancy* or *Relevance* (which is a more popular term in the field of information retrieval), has been regarded as one of the most important criteria for web information retrieval (Chowdhury 1999a; Klein 2001; Mizzaro 1997; Price and Shanks 2005b). However, in comparison to the other dimensions, the factor of *Relevancy* is rarely considered by the literature when assessing IQ for users (Knight 2009).

Completeness is a dimension that receives a less important weight in the context of the web. In contrast to the relatively stable status of data in organisational information systems, incomplete information is much more common on the web. Based on Knight's (2009) survey with users, although users frequently encounter information that does not appear to be complete, they have a greater tolerance to the situation. The study also reveals that *Objectivity* is the dimension least concerned by the surveyed user group. This finding demonstrates the users' high degree of tolerance for biased information.

2.2.2.4 IQ metrics

As discussed in the previous section, IQ dimensions define the quality aspects of an information entity for assessment. However, in order to perform the actual quality measure, the dimension "needs to be grounded meaningfully in measurable attributes of the entity" (Stvilia et al. 2007, p.

1722). Therefore usable IQ metrics have been developed on an ad hoc basis to solve specific quality problems in organisational or web settings (Eppler and Muenzenmayer 2002a; Naumann 2002; Naumann and Rolker 2000; Zhu and Gauch 2000).

For instance, principle IQ measurement functions, such as simple ratio, min or max operation and weighted average, are provided to guide objective IQ measurement (Pipino et al. 2002). Another example is found in the quality metrics proposed by Zhu and Gauch (2000), for quantifying the quality of information retrieved from the web. The below six quality metrics are defined for measuring IQ dimensions such as *Authority*, *Currency*, and *Availability*.

- Availability metric: is defined as the number of broken links divided by the total number of links on a webpage;
- Authority metric: refers to the score assigned to a reviewed website;
- Currency metric: refers to the last modification time stamp of a website;
- Information-to-noise ratio: refers to the total length of content tokens after pre-processing divided by the size of a webpage
- Cohesiveness: refers to the extent to which the major topics of a website are related;
- Popularity metric: refers to the number of links to a webpage.

In addition, comprehensive analysis has been undertaken to investigate relationships between IQ dimensions and IQ indicators. For instance, based on the analysis of website structure, Stvilia et al. (2009) propose five constructs of IQ markers (indicators), namely baseline, authorship, IQ assurance process, verifiability, and content ownership. The authors find that "the Accuracy construct of the model combined both the Accuracy criterion itself and the trust-related criteria of Credibility and Reliability" (p. 1788). This finding is drawn on the data collected by surveying a convenience sample of 108 health information consumers, and interviewing a sample of 20 survey participants. The finding indicates that consumers may use the markers of *Reliability* to assess *Accuracy* indirectly or to assess quality in general (Bailey et al. 2001). The proposed quality criteria constructs, marker constructs, and relationships between these quality criteria and markers are presented in Figure 2.1.

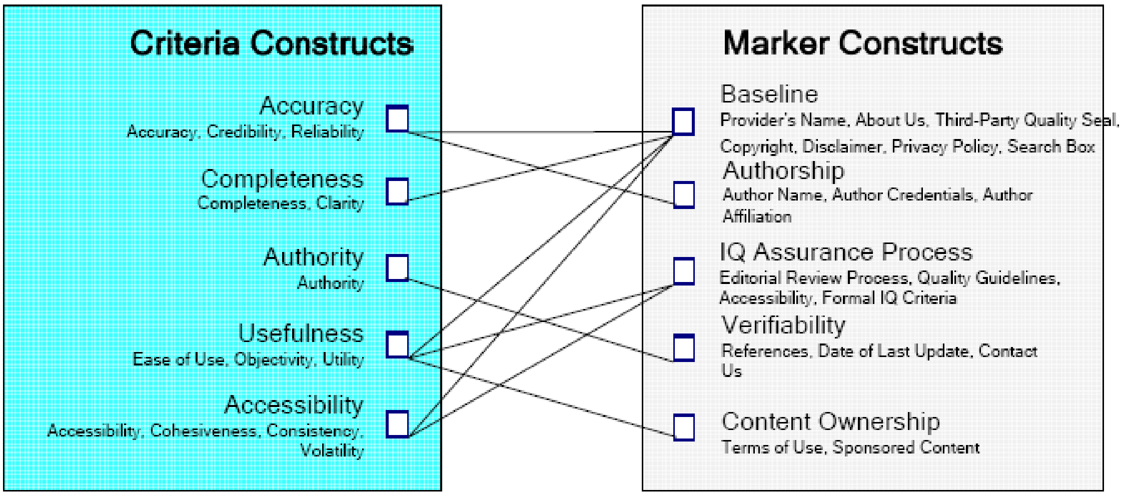


Figure 2.1 Mapping IQ criteria to webpage markers (Stvilia et al. 2009, p. 1789)

In another study, Naumann (2002) analyses five general-purpose metadata models for web sources in order to examine their applicability for IQ assessment. Those metadata models include (p. 46-47):

- Dublin Core: a metadata element set, intended to facilitate the discovery of electronic resources (DCMI 2011).
- STARTS: the Stanford Proposal for Internet Meta-Searching (STARTS) proposes a list of required metadata fields for documents (Gravano et al. 1997).
- Z39.50 (BIB-1): an ANSI and ISO standard that describes the communication between a client and a metadata server mainly with respect to searching. The Attribute Set BIB-1 describes bibliographic metadata and is made up of 100 attributes (Z39.50 1995).
- GILS: The Profile Global Information Locator Service (GILS) is not only a means to describe books or datasets, but also to provide data about people, events, meetings, artefacts, rocks etc. (GILS 2000).
- DIF: The Directory Interchange Format (DIF) was originally developed to make scientific, US-governmental catalogues describing interoperable data groups (DIF 2010).

Figure 2.2 shows the mapping between IQ criteria and a set of metadata attributes synthesised from the above five metadata models.

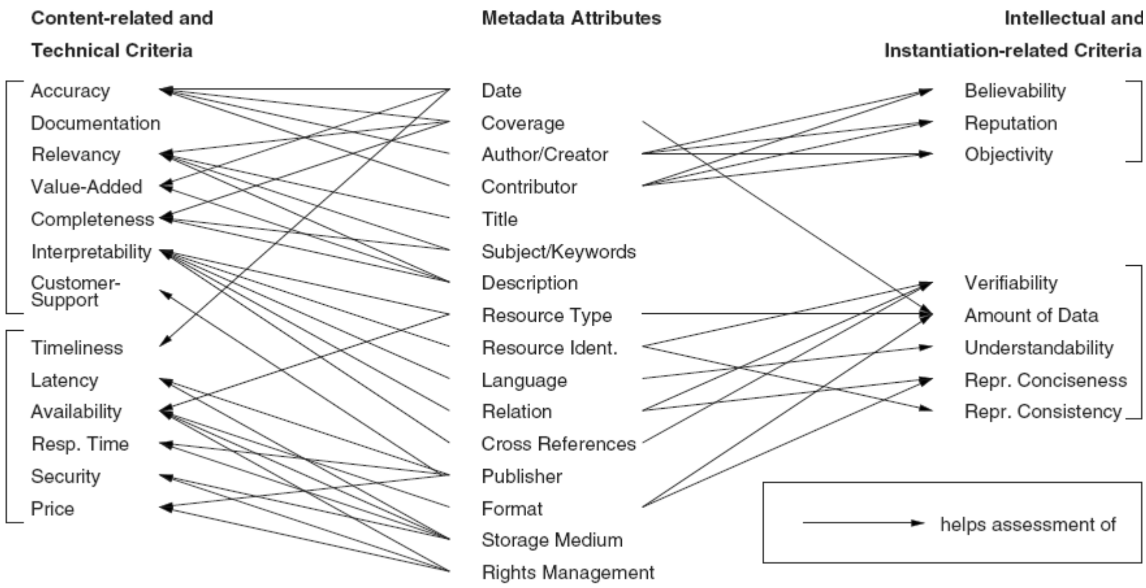


Figure 2.2 Mapping IQ criteria to general-purpose metadata attributes (Naumann 2002, p. 49)

The analysis concludes that in general, existing metadata attributes for resource discovery on the web do not directly satisfy the needs of IQ measure (Naumann and Rolker 1999). However, some metadata attributes may help to determine IQ criteria scores. For instance, *Relevancy* is linked to the metadata attributes of Coverage, Title, Subject/Keywords, and Description, based on the consideration that a document is relevant to the query if the query terms appear often and/or in prominent positions in the document. Another example is *Believability*, which is linked to Author/Creator, Contributor, and Publisher metadata attributes. This association reflects the similar construct of *Authorship*, which is defined as an IQ marker of *Accuracy* in Stvilia et al.'s model (see Figure 2.1).

It is worth mentioning the *Verifiability* criteria, the construct of which is also investigated by Stvilia et al. (2009). Naumann links *Verifiability* to the existence of Resource identifier, Relation, and Cross references, but the author claims that the content of these metadata attributes do not directly contribute to *Verifiability*. The construct of *Verifiability* is categorised by Stvilia et al. (2009) as an IQ marker associated with the *Authority* criteria.

The following section discusses IQ assessment methods, tools, and techniques that have been utilised in assessing online information.

2.2.3 Quality Assessment Methods and Techniques for Web-based Information

In the context of web information systems, various IQ assessment methods and techniques are proposed to capture and analyse the state of IQ in different web application settings, with varying purposes. It is necessary for this research to investigate whether existing IQ assessment approaches, particularly those developed for the selection of online resources (Cappiello et al. 2003b; Naumann et al. 1998), can be applied or adapted to support resource quality assessment for health information portals. In Table 2.5, pervious IQ assessment studies in the literature are categorised according to different technical methods they employ.

Table 2.5 Methods and techniques for assessing web-based information

Methods and techniques	IQ assessment purpose and web application domain	Related studies
Evaluation instruments, e.g. questionnaire	Charactering and measuring the quality of individual websites, e.g. electronic commerce websites	(Katerattanakul and Siau 2002)
Web application tools, e.g. site analyser, traffic analyser, web mining tools, sever and network monitoring, user feedback	Evaluating the quality of web pages and user satisfaction	(SortSite 2008)
Ranking or scoring systems	Evaluating and selecting web resources as external information sources of a data warehouse	(Zhu and Buchmann 2002)
	Selecting web resources for access according to their IQ measures	(Naumann 2002)
	Enhancing web information retrieval by incorporating IQ measures	(Knight and Burn 2005)
Applying metadata models for IQ assessment	Mapping IQ criteria and metadata models to facilitate IQ assessment of web resources	(Naumann 2002; Naumann and Rolker 1999)
	Measuring the quality and relevance of web information for preservation. Metadata are used to capture the quality status of web pages both at creation and update time.	(Cappiello et al. 2003b)
	Computing the values of IQ dimensions, authoritativeness in particular, from metadata values to recommend web pages	(Barros et al. 2008)
Applying intelligent techniques, e.g. machine learning and fuzzy logics for IQ assessment	Using machine learning to automatically classify digital resources into quality brands so as to support value judgements	(Custard and Sumner 2005) (Stvilia et al. 2009)
	Using machine learning to automatically assess RQ for educational digital libraries	(Bethard et al. 2009)
	Evaluating the IQ of websites by fuzzy computing with words and generating the linguistic recommendations	(Herrera-Viedma et al. 2003)
	Using fuzzy logics to capture humans' imprecise knowledge for predicting the IQ of retrieved web pages	(Barros et al. 2008)

Amongst these studies, IQ research conducted for online resource selection and quality information retrieval is considered as the most relevant to this research, therefore is further discussed. The application of quality metadata models and intelligent techniques (machine learning techniques in particular) to IQ assessment is also highlighted.

Naumann et al. (1998) first introduced the use of IQ measures to tackle resource selection problems in the context of integrated systems. The proposed "Data Envelopment Analysis" method utilised a computational model, which initially contained the three IQ criteria of *Understandability*, *Extent*, and *Availability*. Based on this preliminary work, Naumann (2002) later proposed an IQ scoring approach for resource ranking and selection. The multi-criterion ranking model extended the previous model by adding more IQ dimensions, such as *completeness*. The IQ criteria are measured in different ranges, units, and scales, with weighted importance values assigned by users. All criterion scores are finally aggregated to an overall quality score to determine the ranking of a resource.

Likewise, Cappiello et al. (2003b) propose a computational model for measuring quality of web-based information resources. The work breaks down a webpage into meaningful objects and evaluates the quality of these components for the purpose of supporting archivists in the selection of web pages being preserved. The overall quality of a webpage is aggregated from the measurement of multiple quality dimensions, including *Accuracy, Currency, Completeness, Consistency, Volatility, Authenticity*, and *Credibility*. The quality assessment results are preserved in a metadata repository named a "Quality Factory".

Another study proposes a similar approach to quantify the quality of a webpage, but the measure is independent of the content (Kc 2009). A set of high impact quality features are selected and grouped into two categories: one subset of link-based features that are based on information about the web pages before the page content is retrieved; another subset of page-based features that are related to the webpage. These features are namely: Spelling accuracy, Document size, Existence of references, Existence of authorship, Non-spam probability, Grammar correctness, and Correctness of content (p. 159). As a result, these features provide quantitative IQ parameters to construct a quality-aware crawler.

Apart from the ranking or scoring systems, machine learning techniques are utilised for resource evaluation and selection in the domain of educational digital libraries (Custard and Sumner 2005). Based on the quality metrics developed for 16 quality indicators, this study successfully classifies digital resources into quality bands and suggests the development of cognitive tools to support library developers and library users assessing the quality of educational resources and collections more effectively and efficiently (Sumner and Marlino 2004).

Finally, it is necessary to introduce the web application tools, which assess the IQ of an individual website or webpage from a technical perspective. Among those automation tools, a comprehensive one is SortSite (www.powermapper.com), a standards-based website checking tool, which can automatically evaluate the overall quality of a website or a webpage from a technical perspective. The tool analyses a number of quality aspects, including implementation accuracy, browser compatibility, accessibility, usability, legal compliance, page validation against W3C standards, and compliance with leading search engine guidelines such as Google, Yahoo and MSN (SortSite 2008). However, it is not clear whether those technical aspects can be used to indicate the web content quality.

Although it is widely recognised in the literature that a certain level of subjectivity in IQ assessment is unavoidable, existing research mostly measures objective IQ dimensions. Very few quality assessment algorithms incorporate the quality perceptions of individual consumers. There are some problems with the application of those computational IQ assessment models in the context of user-sensitive health information portals. First of all, some IQ dimensions, such as *believability* and *reputation*, are highly subjective and some writers claim that it is not possible to define objective measures for them (Naumann, 2002). Second, the validity of proposed algorithms is limited to the nature of the quality issues they are aiming to tackle. Usually, the development of an algorithm is highly dependent on its specific problem domain. Although some quality assessment frameworks claim to incorporate a user quality perspective, it is not easy to implement algorithms that tie the assessment phase to the information requirements of individual users. Computational algorithms for data quality measures need to consider not only the values of data, but also the context in which the data are used. Quality assessment in a user-sensitive information portal is highly context-based. The quality of an online resource in terms of fitness for purpose can only be measured in the context of an actual consumer's information needs. For this reason, models or tools developed and tested by these studies cannot be applied directly to replace the role that domain expertise plays in quality assessment in a portal scenario.

2.3 Quality Control of Online Health Information

Quality control of online health information is an emerging research discipline and methodology that studies the determinants and distribution of health information and misinformation on the web (Eysenbach 2002; Eysenbach et al. 2002). In this section, the quality issues of online health information and the existing quality control approaches are introduced with a focus on the role that quality-assured web information portals play in protecting consumers from fraudulent or harmful health information on the Internet.

2.3.1 Quality of Online Health Information

Health informatics (also called medical informatics) is a discipline at the intersection of health science, information science and computer science that "concerns itself with the cognitive, information processing, and communication tasks of medical practice, education, and research" (Greenes and Shortliffe 1990, p. 1114). The field endeavours to improve healthcare by using information technologies in innovative ways to collect, process, retrieve, analyse and manage health information (Hersh 2002). Consumer health informatics is a branch of the field concerned with several issues. Those include analysing, modelling and integrating consumer information needs and preferences into information management systems, as well as developing and evaluating methods and applications to support consumers in accessing and using health information (Eysenbach 2000).

The quality issue of online health information is specifically defined as "the degree to which web health information positively affects a user's health outcomes, quality of life, or disease-specific clinical end points" (Risk et al. 2002, p.2714). This notion implies that the quality of an online health information resource to an individual consumer is ultimately determined by the outcome of using the information. Previous research has shown that information leading to "consumer empowerment" can positively affect the decision-making outcomes of health portal users (McKemmish et al. 2009). The study demonstrates that providing value-added information about RQ is an effective approach to assist users in making informed healthcare decisions.

The web is regarded as an ever-expending virtual information system and a convenient communication channel for people to easily access and publish all sorts of information. Publishing information on the Internet used to require expert knowledge of web content engineering, e.g. HTML skills. However, the increasing adoption of web 2.0 technologies in the development of modern web applications greatly facilitates end-user contributions to web content. For instance, websites such as Facebook (www.facebook.com), Flickr (www.flickr.com), YouTube (www.youtube.com), and Wikipedia (www.wikipedia.org) encourage people publishing their own content, and commenting on content provided by others. Internet user communities have benefited from web 2.0 technologies, which support live interactions among users for communication and knowledge sharing. As a result, massive user-generated information is delivered on the Internet with varying quality and value to other information consumers. Everyone becomes his or her own medium. In this regard, information and knowledge become part of dynamic web content that only users can control (O'Reilly and Battelle 2009).

2.3.2 Quality Control Perspectives

According to Eysenbach (2005a), the design and development of consumer health information websites involve a number of stakeholders, i.e. health information providers, website developers, end-users, policy makers, and third-party experts such as annotators, certifiers, librarians, reviewers or evaluators. Therefore, there will be various purposes or perspectives when discussing IQ assessment (Eysenbach 2005a). Different types of stakeholders and their evaluation objectives are listed in Table 2.6.

Table 2.6 Quality control perspectives in the healthcare domain, based on Eysenbach (2005a)

Stakeholders	IQ assessment is aiming at
Developers	Continuously improving the health information published online in the form of web services and web sites
Researchers	Generating evidence to inform users, developers, and policy makers about the quality of online health information
Third parties	Guiding users to trusted online health information
End-users/consumers	Being able to evaluate online health information

As one of the third parties for quality control of online health information, web portals play a critical role in defining, filtering and delivering quality information to end-users. However, IQ assessment in the context of consumer-oriented web portals goes beyond finding trusted information sources for portal users, to determining how to make such quality assessment processes transparent to users. The diversity of user needs and values is respected in this way, making users more capable of evaluating the reliability and fitness of portal-recommended resources based on their individual situations. In this sense, quality assessment in web portals has to incorporate users' quality perceptions when defining what can be called quality or trustworthy information.

2.3.3 Quality Control Approaches

In the field of consumer health informatics, quality control of online health information has been raised as a critical issue (Eysenbach 2000). Eysenbach points out the following four fields, in which quality assurance can be achieved to protect consumers from fraudulent or harmful information on the Internet (p. 1715).

- educating information consumers,
- encouraging self-labelling and self- regulation of information providers,

- having third-parties evaluate and rate information, and
- enforcing compliance with criteria.

The quality control approaches for online health information are depicted in Figure 2.3.

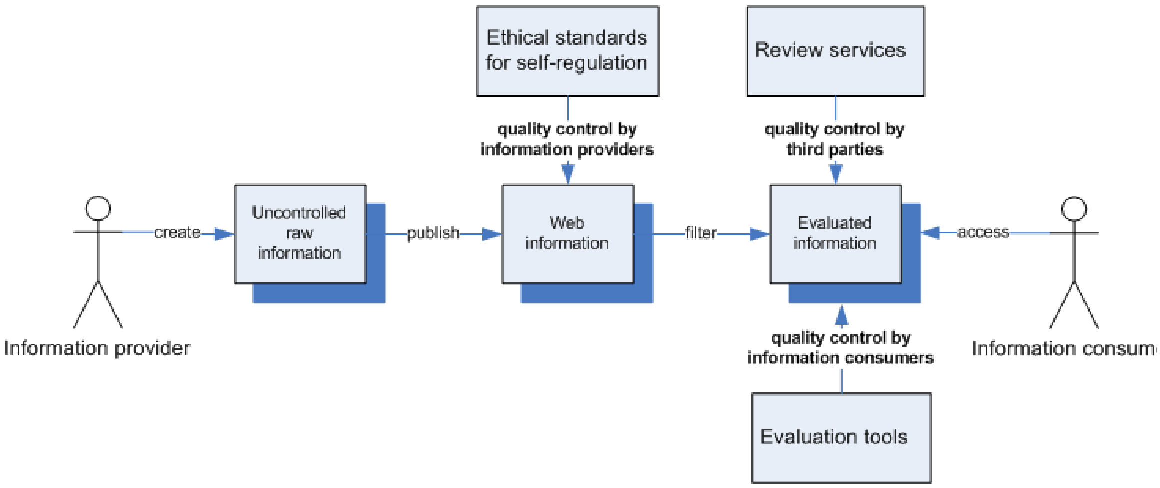


Figure 2.3 Quality control approaches for online health information, based on Eysenbach (2000)

A number of quality initiatives have been developed in the above fields to control or monitor the quality of online health information. The most popular ones are those quality initiatives which have developed ethical standards for information providers (Risk and Dzenowagis 2001) and educational tools for information consumers (Wilson 2002). For instance, considerable effort has been made to develop consumers' critical evaluation skills, through the use of quality checklists, rating instruments, scoring systems, quality seals and certifications (Bomba 2005; Gagliardi and Jadad 2002; Wilson 2002). However, critical evaluation skills are essential before lay users can use those tools correctly. One empirical study points out that "it is unreasonable to expect patients to be able to evaluate the vast quantities of health information they find on the web" (Ziebland 2004p.1784). The reality demonstrates that users are unprepared for taking such responsibilities (Metzger 2007).

Instead of shifting the burden of understanding complex quality criteria to consumers, having third parties mediate between information providers and consumers is one way to achieve quality-assured information provision. Such an approach has been implemented in filtered gateways or web information portals, which provide consumers with access to online information resources that meet an explicit set of quality criteria (Anderson et al. 2003; Manaszewicz et al. 2002). Cooke et al. (1996) consider the OMNI gateway to be an exemplar of this type of quality portal. The OMNI gateway (Organising Medical Networked Information) was initiated as the health and medicine section of the Resource Discovery Network (RDN) funded by UK government, and is now part of the Intute suite of subject gateways (Intute: Health and Life Sciences 2007). As claimed by the project's coordinator and architect Frank Norman, the portal brought librarians' skills to deal with the tangled web (Wickham 2006).

An EU-founded semantic web project MedCERTAIN has also developed a standardised and health domain specific quality framework for resource evaluation, selection and description. The objective was to encourage and enable communication and collaboration among various quality initiatives to finally create a network of trust for health information on the web (Eysenbach 2005b). The framework was delivered in a form of metadata vocabulary termed HIDDEL, Health Information Disclosure, Description and Evaluation Language (HIDDEL 2003; MedCIRCLE Collaboration 2002). It was built on top of W3C's technical standard PICS: Platform for Internet Content Selection (PICS 2008).

In the practice of health information portals, quality assurance is usually achieved by restricting the set of information sources to trusted websites. Information is filtered by the perceived authoritativeness of its source to assure the reliability of search results. The source of information is sometimes categorised to enable a user to make an explicit choice of preferred information source for each search. This approach provides a set of authenticated and high quality results for certain types of searches (Eysenbach 2001).

2.3.4 Quality Criteria Models

In the field of health informatics, more than a hundred quality criteria have been published either in scientific journals or on websites for the evaluation of online health information (Eysenbach et al. 2002; Risk and Dzenowagis 2001). For instance, one study systematically reviewed published explicit criteria for evaluating health related websites and identified over 165 individual criteria (Kim et al. 1999). The identified quality criteria have been categorised in relatively simple ways, such as objective versus subjective, or technical versus content (Eysenbach 2005b; Eysenbach et al. 2002; Kim et al. 1999). This research looked at 14 quality initiatives, which provided criteria models in the form of quality standard or code of conduct that were mostly cited in the field of consumer health informatics (see **Appendix A**). Major criteria published by these quality initiatives are categorised in three groups, namely content, privacy (confidentiality), and web design (technology aspects), as presented in Table 2.7.

Table 2.7 Summary of quality criteria for evaluating online health information

Category	Quality criterion	Definition adapted from Cooke (2001), Anderson et al. (2003)
Content	Authorship	This includes the creator and publisher of the material and in some guidelines the 'authority' of the author – i.e. professional, lay, hospital etc. Generally taken to mean the need for credentialing.
	Accuracy	What 'evidence' is available to support the claims? The criterion here ranges from case studies to evidence-based medicine and the obligation to inform users as to which level of evidence has been used. Language defining this category is extremely vague.
	Source	Need for clear citations and/or links to cited evidence.
	Balance	This attempts to assess whether material is objective, emotional, biased, or simply pushing one line or product.
	Currency	Date of creation, posting, amendment to material(s). This criteria may also apply to the source of information if on another website.
	Review process	What is the editorial policy of the site?
	Purpose	Overall objective of site and material should be evident, i.e. is the aim to sell, promote, or simply educational. The material is then judged according to whether it fulfils its specific purpose.
	Caveats	Sites need clear indications that information is not medical advice etc. or itemised information as to sponsorships and how this affects editorial policies, coverage, etc.
	Comprehen-siveness	How extensive is the information given? Does it include areas of disagreement/lack of evidence?
	Further-reading	For some models the addition of 'reading lists' becomes an indication of 'quality' alerting the user to other sources of relevance
	Language	Whilst most sites argue for 'understandable' information which caters to a specific audience, only DISCERN specifically endorses the use of the Flesch measure. Basically the argument is to provide content that caters to a lay audience.
	Audience	This element ties back into language and purpose. Most criteria simply assume the 'consumer' or 'physician' audience.
	Copyright	Is there a clear indication of copyright obligations and possible access restrictions?
Privacy (Confi-dentiality)	Feedback mechanisms	The inclusion of addresses (mail and email) for complaints, queries, as part of 'accountability' mechanisms.
	Cookies	What is site's policy on collection, use, and administration of cookies? This should be clear to users and some codes call for user control – either 'opt out' or 'opt in'.
	Advertising	This subcategory includes alerting users to funding, clearly differentiating advertising material from editorial or educational matter; impact of advertising on editorial integrity, etc.
	Personal data collection	Involves alerting a user to what specific information is collected, how this will be used, and to whom it might be passed on. Several codes insist on explicit 'informed consent' – others for the 'opt out' option only.
	Disclosure	A wide category, which covers everything from sponsorship, personal bias/expertise, author/website affiliations.

Web design (Technology aspects)	Metadata	How this will be used, how updated, to what it will link
	Links	Are 'back-links' provided? Is the user aware that he is entering a foreign site? How well are the links maintained? Have links been through a review process? Accredited?
	Navigability	How 'easy' is it to use the website? Are directions clear and intuitive? This subcategory largely focuses on web design features.
	Searchability	Does the site have an efficient search engine?
	Software	What software is available to users? How will multi-media requirements impact on users?

There is no gold standard for measuring the quality of online health information (Gagliardi and Jadad 2002). The following quote illustrates the difficulty in assessing the quality of online information in a given context:

> "Quality assessment is not a straightforward procedure involving an identification of the presence or absence of different features or facilities. Instead, quality assessment is a complex process involving consideration of a wide range of interrelated issues that are of varying importance depending upon the nature of the source and the needs of the user…Due to the complex nature of quality assessment, it would not be possible to provide a straightforward list of criteria." (Cooke 2001, p. 13)

Although IQ criteria are not defined in a consistent manner, a number of consensus criteria for evaluating online health information have been identified by a number of studies. One example is Silberg et al.'s (1997) four principled quality criteria, namely *Authorship*, *Attribution*, *Disclosure*, and *Currency*. They have been widely incorporated in many later developed quality criteria models. Additional dimensions such as *Accuracy*, *Completeness*, *Authority*, *Readability*, and *Design* are also popular in many models (Eysenbach et al. 2002; Kim et al. 1999). As suggested by Kim et al. (1999, p. 647), it is necessary to identify a "clear, simple set of consensus criteria that the general public can understand and use". Consequently, a relatively simple set of criteria using the mnemonic "CREDIBLE" is suggested by Eysenbach (2002, p. 765) in order to train users to locate and assess online health information:

- Current and frequently updated
- References cited
- Explicit purpose and intentions of the site
- Disclosure of developers and sponsors
- Interests disclosed and not influencing objectivity (e.g. financial interests)
- Balanced content, lists advantages and disadvantages

- Labelled with metadata
- Evidence-level indicated

However, according to Eppler and Wittig (2000), quality criteria models proposed for the application in health informatics are not regarded as IQ frameworks, as they simply list criteria or guidelines without conceptual insights and problem solving capacities.

2.3.5 Quality Assessment Tools for Online Health Information

Quality assessment of online health information is recognised as a complex and challenging task. Quality assessment tools, such as checklists and rating instruments are available online to help professional evaluators with the critical evaluation of Internet-based resources. According to Merrill (1999), there are four checklists for librarians to assess the quality of online information. These include Katz's (1997) points for evaluating print sources and electronic databases, Tate and Alexander's (1996) checklist for web pages, Infofilter project's review template (Collins 1997), and genre categories (Boese and Howe 2005). Those rating instruments include Bomba and Land's (2004) Consumer Health Website Rating Index, and Aladwani and Palvia's (2002) 25-item instrument, which measures web quality in the aspects of specific content, content quality, appearance, and technical adequacy. However, it was pointed out that the validity and reliability of those consumer-oriented rating instruments in measuring the quality of online health information remained questionable (Jadad and Gagliardi 1998). Only a few of those published rating instruments were found to be functional after three years (Gagliardi and Jadad 2002).

Research on automated quality assessment of online health information is relatively new in the literature. Eysenbach and Diepgen (1998) propose a method using a metadata-based filtering tool. They suggest that by providing both self-labelled and third party annotated metadata, a tool residing on a user's browser can automatically filter information of good quality. However, because of the lack of popularity of resource descriptive metadata among information providers, and the immaturity of quality metadata models (Naumann and Rolker 1999), this approach was only empirically evaluated at the conceptual level (Eysenbach 2005b).

As a different approach, a number of studies try to auto-detect widely used quality indicators that are defined in published guidelines. Research has been conducted to find the correlation between the objective or technical IQ criteria and the content IQ criteria in order to predict the content quality, such as accuracy and completeness (Eysenbach et al. 2002). For instance, one analytical study identifies an association between content accuracy and citing references as well as an absence of financial interest (Martin-Facklam et al. 2002). A recently published work automatically detects indicators of technical quality criteria as a first step to assist users evaluating online health information (Wang and Liu 2007). Likewise, Griffiths et al. (2005) propose a procedure to automatically assess the content accuracy by checking evidence-based information in depression websites. In another study, the DISCERN score (DISCERN 2008) and Google PageRank (Google 2011) are proposed as indicators of the content quality of websites (Griffiths and Christensen 2005). Moreover, indirect quality indicators such as usage statistics, have also been explored for their usefulness in detecting quality medical websites (Hernández-Borges et al. 1999). The underlying assumption of the above studies is that by knowing the technical or objective attributes of a web resource, the content quality of the resource can be potentially predictable.

2.3.6 Quality Indicator Analysis

In the literature, numerous quality indicators have been identified and categorised according to their use in constructing quality metrics. For instance, Eysenbach and Diepgen (1998) classify quality indicators into direct and indirect groups, while Wang and Liu (2007) consider two categories of technical and non-technical indicators. Those automatically measurable or detectable quality indicators are used to compute quality metrics. Table 2.8 lists quality indicators that have been used by previous studies in the literature to assess quality dimensions.

Table 2.8 Quality indicator analysis

Quality dimension	IQ indicators	Automation method	Studies
Reliability	Web statistics, cybermetric indicators	Web link analysis and web usage statistics, e.g. Google PageRank	(Cui 1999; Hernández-Borges et al. 1999) (Aguillo 2000; Eysenbach and Diepgen 1998)
Accuracy	Information provider or third-party generated metadata (e.g. DC.Date, DC.Creator, DC.Publisher, DC.Contributor)	Metadata extractor	(HIDDEL 2003; IntuteIntute: Health and Life Sciences 2007; McKemmish et al. 2009; Naumann 2002; Stvilia et al. 2009)
	Evidence-based	AQA (Automated Quality Assessment) Procedure	(Griffiths and Christensen 2005; Griffiths et al. 2005)
	Scoring system (e.g. DISCERN, NetScoring, Quick)	Scoring system	(Griffiths and Christensen 2005)
	Disclosure information in a website's 'Disclaimer Statement'	HTML parser	(Griffiths and Christensen 2005)

Credibility	Information provider or third-party generated metadata (e.g. DC.Creator, DC.Publisher, DC.Contributor)	Metadata extractor	(HIDDEL 2003; IntuteIntute: Health and Life Sciences 2007; McKemmish et al. 2009; Naumann 2002; Stvilia et al. 2009)
	Author's name, credentials and affiliation, Reference, Copyright, Editorial review process, Advertising policy, Disclaimer, Statement of purpose, Privacy policy, Sponsorship, Name and type of the provider, Search, Contact us, Site map, Payment information	Automatic indicator detection tool (AIDT) from web locations of text, image or hyperlink	(Wang and Liu 2007)
	Third party labels or seal of certification (e.g. eTrust, HON Code of Conduct seal)	HTML parser	(Civan and Pratt 2006)
	Web statistics, cybermetric indicators (e.g., usage analysis, web link analysis	Google toolbar rating, Google PageRank	(Aguillo 2000; Cui 1999; Currò et al. 2004; Eysenbach and Diepgen 1998; Frické et al. 2005; Griffiths and Christensen 2005; Hernández-Borges et al. 1999)
	Key terms revealed in the web content (e.g. 'miracle cure', 'visa' etc.)	Site search	(Price and Hersh 1999)
	Site URL (e.g., .gov, .org, or .com)	URL inspector	(Civan and Pratt 2006; Price and Hersh 1999)
Currency	Information provider or third-party generated metadata (e.g. DC.Date)	Metadata extractor	(HIDDEL 2003; IntuteIntute: Health and Life Sciences 2007; McKemmish et al. 2009; Naumann 2002; Stvilia et al. 2009)
	Date of creation, Date of last update	AIDT tool	(Wang and Liu 2007)
Usability	Contact details (e.g. email address)	HTML parser	(Currò et al. 2004)

Additionally, in order to identify the relationships between IQ dimensions and metadata elements, an analysis was carried out by mapping elements of existing metadata models for resource description, including Dublin Core (www.dublincore.org), AGLS (www.agls.gov.au) and BCKOnline (McKemmish et al. 2004) to IQ dimensions defined in Naumann's (2002) assessment-oriented criteria model, Health on the Net codes of conduct (2007), and Bomba's (2005) criteria model for health websites (see **Appendix B**).

The quality indicator analysis presents a highly dynamic world of online information provision in terms of where and how to locate quality indicator values. For those direct quality indicators, their values are embedded in the body text of an online resource. In addition, indirect quality indicators, such as those resulting from web link analysis and web popularity analysis, can be generated or measured. They can provide insights that go beyond human cognitive abilities (Custard and Sumner 2005). However, as pointed by Eysenbach and Diepgen (1998), the relationship of those technical or indirect indicators to the content quality of information resource remains questionable.

2.4 Metadata-driven Web Information Portals

Information portals or filtered gateways are web applications that give users immediate access to the descriptions of online resources, and control over when and how they share information (Eboueya and Uden 2007). As the prime aim of an information portal is to provide a specific type of information, information portals are studied as a category in their own right. However, they are not mutually exclusive to the other portal types, such as horizontal or vertical portals, personal portal, and community portals (Tatnall 2005). Usually, web information portals themselves are not information creators, but provide gateways to collections of external online resources. These mostly include websites and individual web pages that can be easily accessed. In some cases, a portal also contains deep web information, which is accessible via the web but requires activation of certain programs, e.g. PDF, WORD documents, or multimedia files. The level of quality control distinguishes web information portals from search engines. Instead of indexing the entire web for general users, information portals only cover the part of the web that is within specified subject areas and of perceived usefulness to targeted user communities.

2.4.1 Web Portal Systems

A web information portal can be described as a web information retrieval (IR) system. Chowdhury (2004, p.4) categorises all tasks carried out by an IR system into two major groups – a) subject or content analysis "related to the analysis, organization and storage of information", and b) the process of search and retrieval that analyses "users' queries, creation of a search formula, the actual searching, and retrieval of information". According to Lancaster (1979), an IR system comprises six sub-systems, namely document, indexing, vocabulary, searching, matching, and user interface sub-systems. Recent advances in web IR research acknowledge the importance of having a usage analysis mechanism in order to better understand user behaviours (Srivastava et al. 2000). For this reason, a modern web portal system should also include a sub-system for analysing user behaviours. Figure 2.4 illustrates the conceptual model of a web information portal.

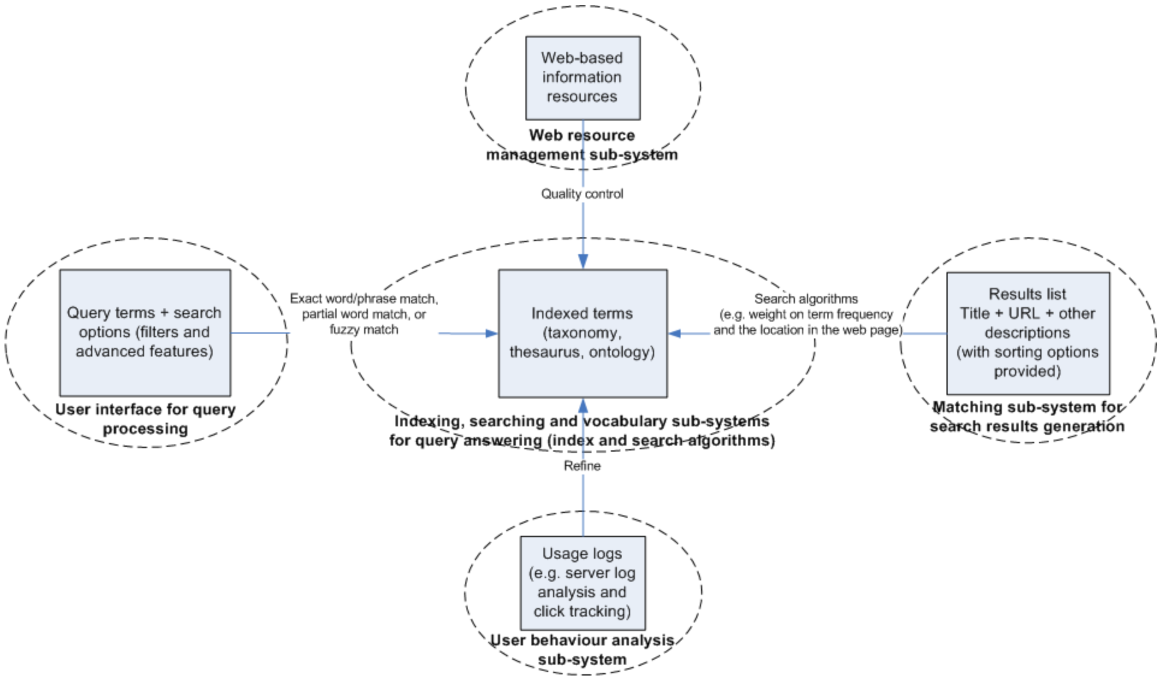


Figure 2.4 Conceptual model of a web information portal, based on Baeza-Yates and Ribeiro-Neto (1999)

Web resource management sub-system: based on a collection development and management plan, a group of resources are selected and preserved in a database system. In metadata-based web portals, this involves all evaluative works in relation to the identification, selection, and description of resources that have perceived high quality and relevancy to the targeted portal users.

Indexing, searching, and vocabulary sub-systems for query answering: In order to speed up searching, a search indexer needs to be used to build a data structure for pre-selected terms, which can be used to refer to the content of a resource. Usually general portals treat all words in a resource as index terms. Instead, metadata-based portals index selected metadata fields in order to reduce the noise and improve the precision of retrieved results. On the other hand, as user queries do not use the same words that can precisely match the indexed terms, a vocabulary list is required to define synonyms and related terms. It can be a taxonomy, which formally defines the classification for a given set of objects in a hierarchical tree structure. In some sophisticated systems, ontologies are used beyond the flat vocabulary list or the hierarchical taxonomy to enable reasoning amongst indexed resources. An ontology is defined as a data model that represents a set of concepts within a domain and the relationships between those concepts, so ontologies can be used to reason about the objects within that domain (Berners-Lee et al. 2001). By using ontologies, resources can be inferred from the others based on their represented and inter-connected topics at the semantic level.

User interface for query processing: the end-user interface of a portal usually provides both topic navigation and search options. A user query is a combination of word or phrase, and any options, filters, or advanced features supported by the portal. In some sophisticated systems, user queries are processed to extract concepts using linguistic analysis or they are expanded using pre-modelled user profiles. These user profiles can reduce the ambiguity and improve the relevancy of search results.

Matching sub-system for search results generation: In order to answer a user query, a portal system generates a result list by checking its index, and saving the matching resource entries. The portal sorts the search results usually according to its relevance algorithms, and displays the results in web pages. If a portal uses a metadata mechanism to describe its included resources, search results can then provide rich resource descriptions in addition to basic information of URLs and titles.

User behaviour analysis sub-system: in order to understand user information needs and monitor the usage of a portal, a server log analysis is usually performed to examine user queries, session times, and search satisfactoriness (Srivastava et al. 2000). Moreover, tracking user activities, such as capturing their clicks on search results, is also useful to identify the most popular items so as to fine-tune the ranking algorithms.

2.4.2 The Metadata Mechanism

In the context of web information portals, metadata can play an important role in the retrieval, management, preservation and use of online information resources that are external to the portals. The idea of characterising and labelling multiple aspects of information resources originated in the use of metadata in traditional library catalogues (Hudgins et al. 1999). Metadata have been widely used to aid resource description and discovery in various application domains, such as digital libraries, archives, and the web. This section focuses on the application of metadata in the context of web information portals and the challenges of generating valued-added metadata.

2.4.2.1 Metadata definition

Although metadata in general means data that describes data or information, the specific meaning of metadata depends on the context and the purpose of its use (Haynes 2004). For instance, the library community defines metadata in terms of controlling access within the information retrieval purpose (Hudgins et al. 1999). In the context of text-based information retrieval, metadata is described as anything about the documents being searched that goes beyond the words they contain (Bray et al. 2006). Another definition combines these views and describes metadata as "structured data about resources that can be used to help support a wide range of operations. These might include, for example, resource description and discovery, the management of information resources and their long-term preservation" (Day 2001, p. 11). Reflecting on the usage of metadata, Haynes (2004) gives a comprehensive definition describing metadata in the context of information retrieval and management as follows:

> "Metadata is data that describes the content, format or attributes of a data record or information resource. It can be used to describe highly structured resources or unstructured information such as text documents. Metadata can be applied to description of: electronic resources; digital data (including digital images); and to printed documents such as books, journal and reports. Metadata can be embedded with the information resource (as is often the case with web resources) or it can be held separately in a database." (p. 8)

However, the above definition does not recognise the essential role that metadata plays in understanding the provenance, relevance, and quality of information, which is captured by Jeffery (2000) in the web context. The role of metadata in finding quality online information is recognised by a number of researchers (Henninger 2008; McKemmish et al. 2009). This view is also adopted in understanding the significance of metadata in the context of web information portals.

2.4.2.2 Metadata standards and models

It is recommended that description of web resources follows the International standards defined by metadata initiatives in order to facilitate the exchange, sharing and reuse of generated metadata (Haynes 2004). The W3C Consortium (www.w3.org) is one of the most well known metadata initiatives that define the International standards in the area of web information retrieval. The Dublin Core Metadata Initiative (DCMI 2011) is another example. Since mid-1990s, the DCMI has established a standard for categorising and describing web content in order to enhance web information retrieval. A core set of metadata elements is defined for web resources' self-description so as to make them more accessible to users and intelligent information discovery systems. Dublin Core is now an international standard (ISO 15836:2009) and the basis of e-government metadata standards, such as AGLS (www.agls.gov.au) and eGMS (www.esd.org.uk).

2.4.2.3 Application of metadata in web portal systems

According to Haynes's (2004, pp. 15-17) model, metadata in a portal context is mainly used for resource description, information retrieval, and the management of information resources. A quality-assured web portal is described as "a selection of resources which meet quality criteria and to display a rich description of these resources with standards-based metadata" (Koch 2000, p. 24). In this sense, a web portal can be regarded as a virtual collection of information resources that are from various external sources on the Internet. After being fed into a portal, the web resources are transformed to corresponding metadata records, which are indexed for subsequent retrieval.

Having metadata brings many advantages as they augment web information with semantics to help with creating filters, and work with navigators. Nowadays not only data but also the semantics of the data have become search targets. Metadata become particularly useful in this sense. For instance, metadata are used to describe the quality of retrieved online information, which empower the decision-making abilities of information consumers (McKemmish et al. 2009).

However, the drawback is that the generation of value-added metadata, such as those describing RQ, goes beyond the words that original web resources contain (Moura et al. 1998). Assigning values to the metadata fields, such as information provenance and targeted audiences, requires intensive human inputs. Passin (2004) points out that using self-describing metadata for web searches has the following limitations (p. 124):

- non-standard metadata may be ineffective;
- incorrect metadata inevitably provides worse results than no metadata;
- expensive to create;
- leave many untouched by semantic mark-up;
- value of classification beyond simple keywords hasn't been demonstrated; and
- the potential for spoofing and distortion.

As the development and implementation of a mature metadata model needs to involve appropriate domain expertise (Sokvitne 2000), using metadata to facilitate resource description, discovery and management brings new challenges and bottlenecks to the quality control of portal content. The quality issues are no longer limited to the portal included external resources, but also lie in the internal metadata representations of those resources. It is envisaged that technological means, which can automate the metadata generation procedures, will be helpful to improve the quality of both portal included web resources and their corresponding metadata records.

2.4.2.4 Metadata generation

As one crucial procedure in the metadata lifecycle (Liu 2007), metadata generation is regarded as the bottleneck for content management in metadata-driven web information portals. Three main parties contribute to the creation of metadata for describing information resources, including resource creators (information provider), professional metadata creators, and intelligent web applications such as web crawlers (Liu 2007). Based on the level of human involvement, metadata are generated using the following three approaches (p. 114):

- manually created or enhanced by resource creators on the basis of automatically extracted descriptions;
- manually created by professional metadata creators, who are familiar with the format and cataloguing rules, thus achieving high quality and consistent results; or
- automatically created by a gathering process involving web crawlers.

Recent advances in web technologies enable web users to tag online resources with terms that might be useful to the others. Specialists in a given knowledge domain are also invited to annotate resources in shared web systems. These perspectives are not captured in the above view of metadata creators. Moreover, in the scenario of metadata generation by third parties, human professional expertise and inputs are highly regarded for the production of standards-based or schema-specific metadata (Liu 2007). Automated systems are considered not capable of producing subject metadata descriptions, which require the understanding of semantics of the resource content. Although the value of detailed metadata descriptions is recognised in improving searching precision, the tradeoffs of the approach are the higher investment in creation of metadata and the difficulty it brings to promote the metadata value consistency (Duval et al. 2002).

Automated metadata generation tools and techniques

For the purpose of this research, solutions developed for automating the metadata generation procedure are highlighted from related works in the literature, as discussed next.

Paynter (2005) describes the development of metadata assignment tools to support automatic record creation for virtual libraries, metadata repositories, and digital libraries, with particular reference to library-standard metadata. The study develops different automated metadata generation algorithms. Besides, different metrics and tools are used to evaluate different metadata fields, including Title, Creator, Keyphrase, Description, Library of Congress Subject Headings, and Category. The evaluation tools are based on and informed by the metadata created and maintained by librarian experts at the INFOMINE (infomine.ucr.edu) project.

In another study, Greenberg (2004) explores the capabilities of two Dublin Core automatic metadata generation applications, Klarity and DC-dot, by using a sample of 29 resources obtained from National Institute of Environmental Health Sciences (NIEHS). Klarity (no longer available from archive.klarity.com.au) was a commercial metadata generator that used text processing algorithms (document categorisation and learning) to assign values to the Keywords and Description metadata elements. The tool also harvested the value of Identifier from the browser address, and the value of Title from the resource source code. DC-dot (www.ukoln.ac.uk/metadata/dcdot/) is another metadata generator that uses harvesting algorithms to automatically assign values to multiple metadata elements. DC-dot harvests the value of Identifier from the browser address, and the values of Title, Type, Keywords, and Description from META tags. When META tags are not available, the value of Keywords is then generated by analysing hyperlinks and presentation encoding. Besides, the tool also harvests values of the Type, Format, and Data metadata elements from file properties.

Three metadata professionals evaluated metadata values generated by the two tools (Greenberg 2004). The overall accuracy results - 25 of 29 for Klarity, 22 of 29 for DC-dot, demonstrate the usefulness of extraction processing algorithms in automating metadata generation. The study also found that harvesting metadata from META tags created by humans positively impacts on automatic metadata generation. Based on these findings, the author suggests that integrating extraction algorithms with harvesting methods is an optimal approach to automating metadata generation.

In addition to the above automated metadata generation tools, Cardinaels et al. (2005) develop a framework to automate metadata assignment, namely the Simple Indexing Interface (SII). Besides, text summarisation is another active area that applies multiple techniques for automating text extraction and abstraction (Mani 2006). A detailed analysis of these automated metadata assignment studies can be found in **Appendix C.2**, which outlines the common metadata elements, the sources for generating their values, and automated metadata generation methods applied in relevant studies.

Value-added metadata generation

Choosing the metadata generation approach should respect the complexity of metadata components. According to Passin (2004, p. 114), metadata about any document or resource can be categorised into three groups, which are:

- Explicit metadata, e.g. author, keywords of its content, publisher
- Metadata published separately, e.g. annotations

- Implicit metadata that can be inferred, e.g. information quality

According to the above categorisation, metadata about resource quality are implicit metadata that can be inferred by ontology and rule-based methods (Hatala and Richards 2003) together with human involvement. From a metadata expert's point of view, a well-designed automatic metadata generation application should be interactive enough to allow a human to have ultimate control over what is generated (Greenberg et al. 2006). Metadata experts prefer to have automatic algorithms execute metadata suggestions for them to evaluate, edit and make the final decision. They believe human incorporation and control can significantly improve the quality of automatically generated metadata.

Resource descriptive metadata deliver value-added information to online information consumers but require intensive human inputs. The challenge lies in how to generate metadata more accurately, consistently and efficiently beyond the traditional manual approach. There is no easy answer to this question but recent advances of natural language processing, data mining, machine learning, and pattern recognition algorithms can contribute to automated metadata generation (Corcho 2006). The application of these techniques is discussed next.

2.4.3 Quality Control of Portal Content

The level of quality control differs in web information portals from that in generic search engines. As opposed to search engines, which aim to index the entire web, information portals only cover the part of the web that is of perceived high quality and relevancy to portal users.

Quality-assured health information portals provide a mediated search environment to information consumers. Domain experts are employed to evaluate the appropriateness of identified external resources for their inclusion in a portal. Candidate information resources are assessed and selected for their perceived high quality and relevancy to the information needs of targeted portal users (McKemmish et al. 2009). The quality of information being delivered via a portal is controlled and determined by domain experts, who make value judgements based on their expert knowledge of the healthcare domain. It is necessary to clarify what constitutes portal content and what are the determinants of portal content quality. These concepts are introduced in this section.

Quality of information in web IR systems is a more critical issue than it is in conventional text retrieval systems. On the Internet, information is published without peer-review processes. Web IR systems therefore have to deal with uncontrolled information sources with varying quality. Although in IR systems intelligent techniques have been widely applied in automating indexing, cataloguing and abstracting (Chowdhury 2004), their application in assessing RQ yet has not been adequately explored.

The Information Systems literature suggests that information quality and user satisfaction are two crucial dimensions to measure the success of an information system (DeLone and McLean 1992). This is particularly true when evaluating health information portals that act as quality intermediaries. Information portals fundamentally provide search and review services to information consumers. Their internal quality assurance processes determine the kind and the quality level of content to be presented to users. If users keep using and returning to a portal, the portal is then regarded as a competitive site of higher quality. For this reason, quality control of portal content needs to be emphasised as the most critical issue in order to achieve greater user satisfaction and final portal success.

In the literature, a number of models have been proposed to evaluate the quality of web portals. Portal content quality has been regarded as one criterion. For instance, one study examines a number of general or subject-specific portals which provide review services to Internet-based information resources (Cooke et al. 1996). The comparison is drawn on four categories including resource selection, the level of description, subject classification and organisation, and the evaluation criteria. Similarly, another study compares 25 health and medical portals from the design, content, administrative, and quality control perspectives (Anderson et al. 1999). A recent study proposes a generic portal quality model, which defines the quality of portal data as one of the most important characteristics of a web portal (Moraga et al. 2007). These studies all place great emphasis on the quality control issue of portal content.

According to Cooke (2001), types of information resources available on the Internet can be categorised into the following three groups:

- text-based content, including organisational, commercial or personal websites and web pages, electronic journals and magazines, and other full-text documents (e.g. online accessible WORD or PDF documents);
- rich content, such as audio, video, and images; and
- specialised content, such as blogs, forums, newsgroups, mailing lists, alerting services, database, and FTP archives.

Cooke (2001) also suggests that the overall quality assessment should be drawn on the following three aspects:

- Is the source valuable and useful, and is the information contained in it valuable and useful?

- Are there any reviews available, or is the site included in any databases of high-quality materials?
- Is it possible to elicit comments from someone who has used the source or who uses it regularly? What is their overall impression of the source?

Linking to the overall assessment of RQ in a portal context, the suggestions imply that the decision of accepting or rejection is determined by the overall quality conclusion. That's why the knowledge of the existing contents of the repository and previous curation experience are useful to support quality judgements. Feedback from both end users and domain experts can be used to elicit such comments. The selection of IQ metrics needs to consider the type of online resources and the specific characteristics each resource type might have. Assessing the quality of text-based web content is of major concern to this research.

Moreover, metadata about the quality of online information resources are value-added information. They go beyond the words those resources contain. Thus, metadata about information quality are more expensive to create. The question is how to implement such metadata more accurately, efficiently, and effectively beyond the traditional manual approach. There is no easy answer to this question but in the literature, intelligent technologies such as information extraction and machine learning (Corcho 2006) are adopted in experimental research activities about automatic metadata generation. However, more questions are raised, such as what kinds of intelligent techniques are applicable to support RQ description and how?

2.5 Discussion

Over the last decade, IQ research has been conducted primarily in the context of organisational database and management information systems, aimed at developing generic or specific solutions for solving quality assessment and management problems in varying application domains (Madnick et al. 2009). Health information portals provide specific application scenarios and settings for this research to investigate user-sensitive RQ assessment issues for achieving quality online information provision.

IQ is recognised as a contextual concept, determined by its generation process or its intended use (Klein 2002; Shanks and Corbitt 1999; Strong et al. 1997). For this reason, defining IQ per se is insufficient without defining the context in which the information will be used (Shankar and Watts 2003). The synthesis of IQ frameworks in both organisational and web contexts provide a general understanding of IQ constructs and assessment. As pointed out by Knight (2009), IQ frameworks are proposed for investigating quality issues from two major perspectives: either from an information generation/production perspective, or from an information retrieval/use perspective. In the case of IQ study from the information use perspective, Knight further states that (p. 279) "… user contextual IQ value judgements about the information are not so much governed by the actual characteristics of the information, but according to how well the information fits the user's need".

This research adopts the definition of RQ as "a component of a relationship between user and resource, rather than an appraisal of the resource alone" (Anderson et al. 2003, p.6). For domain experts, the quality of an online resource is assessed for its perceived fitness to satisfy the information needs and expectations of targeted portal users. A quality framework, which underpins domain experts' quality assessment processes, should therefore reflect the diverse quality perceptions and information needs of targeted portal users.

Although existing IQ frameworks provide theoretical insights for conceptualising RQ and its assessment, they do not offer pragmatic solutions to construct and assess RQ as the relationship between a resource and a user. Some of the existing IQ constructs propose assessment of IQ in the context of user needs and quality perceptions. For instance, the 'Contextual' category of Wang and Strong's (1996) framework recognises the user context and measures IQ as fitness to the user's task at hand. The 'Empathy' category of the WebQual instrument (Barnes and Vidgen 2001; Barnes and Vidgen 2002) and the 'Pragmatic' category of the InfoQual framework (Price and Shanks 2005a) both address the use-based consumer perspective. Although these frameworks address the user dimension, they do not explain how to perform the contextual quality assessment. Moreover, these works recognise that the attributes and dimensions used to assess IQ may vary due to the change of the context, but do not state how to reflect the dynamics of the context. Thus, it is imperative to explore how the user context can be modelled and incorporated to make contextual value judgements. This research fills this gap by conceptualising user-sensitive RQ assessment and developing a pragmatic RQ assessment approach in the user context.

It needs to be clarified that this research aims to develop an RQ construct comprehensive enough to address the healthcare domain, as well as to develop pragmatic solutions to support RQ assessment in a user context. The literature analysis was done to critically evaluate existing IQ constructs and assessment approaches. Although IQ has been well defined as a multi-dimensional concept by existing frameworks, its sub-dimensions or operational measurement models are rarely provided, which makes its actual measure a difficult task (Batini 2006; Pipino et al. 2002; Strong et al. 1997; Wang and Strong 1996). Moreover, the contextual nature of IQ indicates that general IQ solutions might not be applicable to address the user-sensitive requirement. It is argued that an IQ construct and its operational measurement model need to reflect specific quality requirements defined by its context (Bevan 1999; Eppler 2006; Evans and Lindsay 2005). Due to the lack of operational measurement models that can assess IQ in a user-sensitive manner, none of existing IQ frameworks or assessment approaches can be used directly to address user-sensitive RQ assessment issues for the healthcare domain. In order to fill this gap, this research synthesised existing IQ studies in the literature, and proposes a user-sensitive RQ framework, which integrates and extends existing IQ constructs. The framework not only defines RQ as a collective concept of a set of quality dimensions applicable to the healthcare domain, but also defines sub-dimensions and operational measurement metrics that fully reflect quality concerns of health information consumers.

Moreover, as pointed out by Eppler and Wittig (2000), existing IQ frameworks have strength either in their analytic dimension with thorough definitions, or in their pragmatic dimension offering concise sets of criteria and facilitating tools. Rarely, are they both theoretical and practical. It is suggested that new IQ frameworks need to be developed that provide the following criteria: 1) support systematic structures, 2) offer concise criteria sets, 3) show interdependencies between included criteria, 4) contain problem areas and indicators, and 5) include tools developed based on the frameworks.

2.6 Chapter Summary

This chapter presented a comprehensive analysis of multi-disciplinary research literature, which investigated RQ assessment issues from the theoretical, contextual, and technological perspectives. First, IQ research in the IS discipline provided useful theories for constructing and assessing IQ from the user's perspective. Second, the quality control issue for online health information in the field of consumer health informatics provided the specific context for exploring the research questions. Finally, in order to understand issues involved in achieving the quality assurance of portal content, the application of metadata in web information portals was reviewed. Drawing from these analyses, it was concluded that existing IQ constructs or assessment frameworks have not adequately addressed the concept of user-sensitive RQ assessment. It is imperative to develop an operational quality framework that integrates and extends the context-based IQ framework by incorporating user information needs and quality perceptions.

Chapter 3

3 Research Design

This chapter presents the overall design of the research. It describes how user-sensitive research quality assessment was investigated through a socio-technical design science research approach. The chapter starts with a brief introduction to the research questions. It is followed by the rationale for conducting this research using an adapted design science approach and mixed methodologies. The chapter then introduces the user-sensitive systems development process of this research, which has the centre role played by theory building. The process encompasses three research phases: concept building, system building, and system evaluation. For each phase, the research activities and related data collection methods and instruments are elaborated. The chapter also clarifies the relationship between this PhD study and the related Smart Information Portals project. Finally, the chapter discusses how Hevner et al.'s (2004) framework for design science research has been adapted and applied in this thesis research.

3.1 Research Questions

The previous chapter highlighted the role that metadata-driven health information portals play in quality online health information provision. For those portals, quality assurance of portal content is achieved via resource quality assessment processes undertaken by domain experts. The adoption of a user-sensitive design philosophy requires domain experts to evaluate the quality of online resources from the perspective of portal users (McKemmish et al. 2009). Domain experts assess the quality and relevance of an online information resource in terms of whether the resource addresses the information needs and quality perceptions of targeted portal users. This user-oriented view is in line with the canonical view of information quality (IQ) as 'fitness for purpose/use' (Wang and Strong 1996). However, existing multi-dimensional IQ constructs and assessment frameworks do not provide operational solutions to assess resource quality (RQ) in a user-sensitive manner. Moreover, due to the subjective and contextual nature of IQ and the dynamics of user needs and values, existing quantitative IQ assessment approaches are not suitable for addressing user-sensitive RQ assessment issues for health information portals. Thus, it is imperative to explore new approaches that can standardise and support the qualitative measurement of RQ, taking account of user information needs and quality perceptions.

This research is concerned with user-sensitive resource quality assessment in the context of metadata-driven health information portals. The research aimed at answering the following central question:

How to conceptualise and support user-sensitive resource quality assessment for metadata-driven health information portals?

The overall research question is further divided into six sub-questions with a number of research activities proposed accordingly, as listed below:

1. How to define resource quality in the context of health information portals?
2. How can existing information quality assessment theories, principles and approaches be extended and adapted to conceptualise resource quality assessment from a user-sensitive viewpoint?

In order to address the above two sub-questions, it is necessary to:

- understand existing definitions of information quality and its assessment, by reviewing associated theories, methodologies and existing quality assessment approaches;
- understand user-sensitive design philosophy and its application in the design and development of web information portals;
- understand metadata and its use in web information portals;
- understand quality control principles for online health information in the field of consumer health informatics; and
- define the construct of RQ, and conceptualise a framework for standardising user-sensitive resource quality assessment based on the comprehensive analysis of the above research fields.

3. What kinds of domain expertise are required in performing user-sensitive resource quality assessment in the context of metadata-driven health information portals?
4. What tasks and activities are involved in user-sensitive resource quality assessment processes?
5. What are the needs of domain experts with regard to intelligent support?

In order to address the above three sub-questions, it is necessary to:

- analyse the quality appraisal activities of domain experts, the types of value judgements they make on RQ, and the kinds of domain expertise required in making quality-associated decisions;

- understand the problems and difficulties domain experts encounter with RQ assessment and their needs for intelligent support;
- identify what parts of RQ assessment processes can be automated or augmented by intelligent technologies; and
- elicit design requirements for an RQ assessment solution and the functional requirements of a portal content management system to address the identified needs, based on the resulting articulation of domain expertise, domain expert processes and needs.

6. How can intelligent technologies be applied to support user-sensitive resource quality assessment in metadata-driven health information portals?

In order to address the above, final sub-question, it is necessary to:

- identify applicable intelligent technologies that can provide the required functionality;
- design and develop a prototype system that implements the proposed quality framework and applies intelligent techniques to support RQ assessment; and
- evaluate the prototype system with domain experts for its utility and efficacy in supporting RQ assessment.

This chapter elaborates how these research activities are justified and organised, within a use-sensitive systems development research methodology, and in a socio-technical design science research framework.

3.2 Research Framework and Methodologies

This research is concerned with finding approaches that employ intelligent technologies to scale manual RQ assessment processes in practice. It is multi-disciplinary and applied research committed to studying 'the effective design, delivery, use and impact of information technology' (Keen 1987 p.3) through practical problem solving in a real-world scenario. In this section, the design science research paradigm is discussed and compared with other research paradigms in the Information Systems (IS) discipline. A socio-technical design science approach is justified as a more suitable way to explore the research questions of this study. This section also discusses the user-sensitive design methodology and the systems development research methodology. These methodologies were integrated to govern a user-sensitive systems development process, which put emphasis on theorising as the main purpose of the research.

3.2.1 Socio-Technical Design Science Research Framework

Design science research is more often situated within a positivist paradigm than an interpretivist one, especially when concerned with solving technical rather than socio-technical problems (Iivari and Venable 2009). However, the design science research approach can also be combined with interpretivist approaches when the focus is on human rather than machine contexts. In this section, the design science research framework is integrated with a user-sensitive information systems development methodology to approach the research problem in a human context. This adapted framework enables the development of socio-technical solutions (Bostrom and Heinen 1977a; Bostrom and Heinen 1977b) instead of purely technological ones.

3.2.1.1 Design science paradigm in information systems research

Research in the IS discipline is about explaining, describing, predicting, and testing phenomena emerging from the interaction of people, organisations and technologies (Hevner et al. 2004; Klein and Myers 1999). While traditional IS research is concerned with building and testing theories, based on phenomena occurring in the application of information technology, a new design science research paradigm is emerging, to solve fundamental problems of IS through the creation of new and innovative artefacts (March and Smith 1995).

Weber (1987) discussed the potential power of what he termed an emerging IS research paradigm to articulate design principles independently from the technologies used via explanations of the behaviour of discrete artefacts. March and Smith (1995) later defined design artefacts more broadly as representational constructs, models, methods and instantiations. Simon (1996) classifies design science as one of the below four research types in scientific disciplines.

- Natural sciences – which study phenomena occurring in the world.
- Social sciences – which study structural level processes of a social system and its impact on social processes and social organisation.
- Behavioural sciences – which study the decision processes and communication strategies within and between organisms in a social system.
- Design science or science of the artificial – which claims that all or part of the phenomena may be created artificially, and thus study artificial objects or phenomena designed to meet certain goals.

Simon's (1996) definition shows that problem solving, through the creation of innovative artefacts, distinguishes design science from the others. Hevner et al. (2004) further examine the design science paradigm in the IS discipline, and compare with the behavioural science paradigm. Table 3.1 summarises the nature, characteristics, purpose, strength and limitation of these two types of research.

Table 3.1 Design science research versus behavioural science research, based on Hevner et al. (2004) and March and Smith (1995)

	Design Science	Behavioural Science
Nature	A problem-solving paradigm with roots in engineering and the sciences of the artificial.	A theory development paradigm with roots in natural science research methods.
Characteristics	IT artefacts are the outputs, defined as constructs, models, methods and instantiations. These are designed to meet the identified needs.	IT artefacts, mostly system instantiations, are the objects of studies, while the outputs are principled theories to predict or explain phenomena related to the identified needs.
Purpose	Seeks "what is effective" through the building and evaluation of new and innovative artefacts to extend the boundaries of human and organisational capabilities.	Seeks "what is true" by developing and justifying theories that explain or predict human or organisational behaviour.
Strength	Achieves higher relevance and is good for problem solving.	Achieves higher rigour and is good for theory development.
Limitation	Lack of ground theory.	Passive with technology.

Hevner et al. (2004) agree with Walls et al. (1992) that "design is both a process (set of activities) and a product (artefact) – a verb and a noun" (p. 78). The authors contend that design science and behavioural science are two distinct but complementary paradigms that should be combined to better achieve the relevance and effectiveness of IS research. As a result, a hybrid framework is proposed for IS research, which combines March and Smith's (1995) two-stage design processes of build and evaluate. They also propose seven general guidelines for evaluating the quality of design science research, which have been widely adopted in conducting design science research in the IS discipline. These guidelines are interpreted as the following.

- Design as an artefact – whether the output is an innovative artefact and how novel it is.
- Problem relevance – how to demonstrate the novelty of the problem.
- Design evaluation – how useful the artefact is.
- Research contributions – how novel the technology-based solution is.
- Research rigour – whether sufficient information is given for others to build the artefact.
- Design as a search process – how the solution is selected and whether alternative solutions are examined.
- Communication of research – whether the proposed solution can be replicated, or the artefact can be built again by others, to answer the problem of target audiences.

Hevner (2007) later identified the following three closely related cycles of activities in a research project, which clearly position and differentiate design science from other research paradigms.

- The Relevance Cycle inputs requirements from the contextual environment into the research, and introduces the research artefacts into environmental field testing.
- The Rigour Cycle provides grounding theories and methods along with domain experience and expertise, from the foundation knowledge base, into the research. It adds new knowledge, generated by the research, to the growing knowledge base.
- The central Design Cycle supports a tighter loop of research activity for the construction and evaluation of design artefacts and processes.

The adaptation of Hevner et al.'s (2004) design science research framework to this study is discussed in Section 3.5 of this chapter.

As discussed in this section, both design and artefacts are fundamental to the IS discipline (Hevner et al. 2004; March and Storey 2008). The design science paradigm pulls design and artefacts together to address "important unsolved problems in unique or innovative ways or solved problems in more effective or efficient ways" (Hevner et al. 2004, p. 81). The processes of design, which include multiple iterations of artefact building and evaluation, characterise design science as a rigorous and effective paradigm for IS research. In the following section, the selection of a design science research approach to guide this research is justified as more appropriate than an action research approach for this study.

3.2.1.2 Relevance of design science to the central research question

The central research question of this thesis implies two major research tasks: firstly acquisition of understandings and knowledge of the problem domain; secondly the development and evaluation of a new socio-technical solution to improve current practice. Two research paradigms: design science and action research (Avison et al. 1999), could have been applied to address this problem domain. Both are fundamentally problem-solving paradigms (Baskerville et al. 2009), sharing substantial similarities (Cole et al. 2005; Järvine 2007; Livari and Venable 2009). Thus it is difficult to distinguish one from the other. Peffers et al. (2007) suggest that the clearest distinction between them is in their conceptual origins, which lead to a difference in their central components. Design science research focuses on the designed artefact and the proof of its usefulness, while the focal point of action research in IS lies in the organisational context and the active search for solutions to problems. However, Peffers et al. (2007) further claim that "the search for a designed artefact could be presented as action research" (p. 72). Purao (2002) points out another difference between these two approaches relating to theory building. In design science research, theory building is "in the world of signs with a view to bringing the realization an artefact, leading to creation of knowledge and normative theories" (p. 26), whereas in action research, "it is the organizational setting, leading to theorizing using organizational metaphors" (p. 26).

As discussed above, action research as a problem-solving approach can be an alternative or complementary paradigm to designing artefacts. Action research achieves problem solving through social and organisational change. It requires researchers' active participation in a specific organisational setting and high interaction with business processes (Baskerville and Wood-Harper 1996; McKay and Marshall 2001). However, the purpose of this study is not to change the RQ assessment processes employed in practice, but to understand the processes with a user-sensitive viewpoint and then to scale the processes with a socio-technical solution. The investigation of user-sensitive RQ assessment issues and the exploration of socio-technical solutions do not require interactive participation of the researcher in the practice. Based on these considerations, the design science research paradigm is regarded as the more suitable research framework to guide the overall design of this research.

3.2.1.3 User-sensitive design methodology

User-sensitive design is both a philosophy and a methodology that integrates user information needs analysis (Wilson 1981; Wilson 1994), user-centred design (Norman and Draper 1986; Preece et al. 2002) and value-sensitive (Friedman et al. 2006) approaches, and systems development research techniques within an interpretivist framework (McKemmish et al. 2009). The philosophical view of user-sensitive places user's needs and values at the heart of system design and development.

The user-sensitive design philosophy is informed by the theory of user-centred design, which recognises the importance of incorporating user needs and interests in systems design process (Norman and Draper 1986). As a multi-disciplinary activity, user-centred design incorporates human factors, ergonomics knowledge and techniques with the objective of enhancing effectiveness and productivity, improving human working conditions, and counteracting the possible adverse effects of use on human health, safety and performance (Bevan 1999). It is suggested that four user-centred design activities need to take place at all stages during a project. These include:

- to understand and specify the context of use;
- specify the user and organisational requirements;
- produce design solutions; and

- evaluate designs against requirements.

The understanding of value sensitivity is informed by the theory grounded value-sensitive design approach, which draws attention to the social and cultural dimensions of systems design. According to Friedman (1996), value-sensitive design focuses primarily on addressing moral values of users, such as privacy, trust and autonomy.

Research involved in the development of the Breast Cancer Knowledge Online portal adopts the user-sensitive design philosophy to deliver personalised information of perceived high quality and relevancy to meet diverse information needs of the breast cancer user community (Fisher et al. 2002; McKemmish et al. 2009). According to McKemmish (2009), there are a number of values or principles underpinning the user-sensitive design philosophy in a portal development context (p. 1795), shown as the following.

- Users are not homogeneous. This principle represents a diversity of information needs and preferences.
- Knowledge is contextual. That is, its value to the individual is a dynamic interplay of personal, social, psychological, educational, and cultural factors.
- Quality, relevance, reliability, and trustworthiness are contingent, not absolute. They are not attributes of an information resource per se, but of the relationship between the resource, an information seeker, the information seeker's needs, and the values and life experiences they bring to their assessment of what is relevant, reliable, and trustworthy for their purposes at the time of search. The role of the portal is partly to provide them with sufficient information about a resource to make an assessment about its 'fitness for purpose'.
- The perspectives of users with direct and/or personal experience of certain disease provide a valuable information resource and insight, which both complement and enhance the scientific/biomedical view of the disease treatment, management, and research.
- Members of the user community who participated in the portal development are regarded as partners rather than participants.

The above principles imply the importance of the user-sensitivity when developing RQ assessment solutions for health information portals. The principles require an online resource being assessed for its intended audience, i.e. based on its fitness to perceived information needs and values of different types of portal users. Besides, the level of resource quality description and the level of quality compliance are another two issues to be considered in resource evaluation and description. It is reasonable to infer that the implementation of such a user-sensitive approach will involve user-centred design activities, such as understanding user needs and values, defining the intended context of use, specifying user requirements, producing design solutions, and evaluating designs against requirements from a user perspective (Bevan 1999).

3.2.1.4 Systems development research methodology

As well as the selection and justification of a research framework for guiding the overall research design, it is necessary to justify the selection of a research methodology for actual problem solving. A research methodology is defined as a combination of process, methods and tools (Nunamaker et al. 1991). Research methods, for problem solving in IS research, normally include survey, case study, experimental design and systems development (Williamson 2002). Methods such as case studies, surveys, laboratory and field experiments are empirical approaches to capture the social sides of IS research (Benbasat and Zmud 1999). However, when comprehensive understandings of a problem domain are not obtained, irrelevant questions or meaningless hypotheses might be used to collect data, which may lead to invalid conclusions. The systems development methodology, on the contrary, follows an engineering approach to illustrate or test a proposed theory (Burstein and Gregor 1999). Hevner and Chatterjee (2010) describe a systems (software) development methodology as "a set of principles, models, and techniques for effectively creating software artefacts at different stages of development (e.g., requirements, design, implementation, testing, and deployment" (p. 100).

Within the systems development methodology, a design artefact is prototyped that provides a solution to a particular problem. However, the methodology is not about the creation of an artefact per se. It is about conceptualising a problem, testing the comprised concepts, and contributing to the related theories. It well underpins design science research activities for requirements elicitation, artefacts prototyping, consolidation and theorisation of research findings. Thus, the methodology is employed by this research as a problem solving approach.

From a methodological point of view, Nunamaker et al. (1991) state that the systems development process consists of the following five stages: construct a conceptual framework, develop a system architecture, analyse and design the system, build the (prototype) system, and finally observe and evaluate the system. Burstein (2002) further refines this process in an iterative systems development research cycle that comprises three major steps: concept building, system building and system evaluation. The first step emphasises on the concept that the system has to illustrate. The system implementation (development of a prototype system) is invoked when existing systems are not capable of demonstrating the features of the concept under investigation. Finally, the evaluation of the created system is performed, from the perspective of addressing the research questions set up in the concept building step. The exit point in the system development iterations would occur when the research questions are satisfactorily resolved. This research followed this three-step systems development research process.

3.2.1.5 The role of theorising in design science research

The role of theory and theorising in design science research in the IS discipline continues to be a debatable subject. How should theory inform design science research and should it be instrumental in developing and refining theory (Kuechler and Vaishnavi 2008).

Theorising initially was not regarded as part of design science (Hevner et al. 2004; March and Smith 1995). However, Bertelsen (2000) argues that theories are design artefacts, "guiding the designer and helping him assess the situations and keep the goals in mind, to tools mediating the achievement of specific results" (p. 20). Purao (2002) also categorises the significant outputs of design research as building of theories "articulating operational principles (knowledge) necessary to realize these, and demonstrating both in the form of artefacts as situated implementations" (p. 24).

In the case of design science research that employs the systems development research methodology, the role of theorising is paramount – it is considered to be the underlining philosophy that justifies the value and the validity of both system development process and outcomes. According to Nunamaker et al. (1991), the system building itself does not constitute research. Rather, it is essential to prove underlying theories. Many researchers embrace this view (Burstein 2002; Venable 2006; Weber 1987). Burstein (2002) states that the purpose of developing systems as proof-of-concept does not lie in the concepts by themselves, but the new knowledge they provide in relation to some theories. Weber (1987) also contends that the value and quality of the artefact design presented in IS research must be justified in the context of some theory of information systems. It is pointed out that theory building occurs throughout the whole systems development process and as a research output illustrated by a system (Burstein 2002).

Nunamaker et al. (1991) propose a multi-methodological IS research cycle, which employs four research strategies, including systems development, theory building, experimentation and observation. They argue that the systems development methodology represents a central part of the IS research cycle. The Centre for the Management of Information (CMI), University of Arizona, has revised this systematic approach in order to represent a generic research process (the CMI Arizona Research Model, cited in Burstein 2002, p. 151). The central part of this revised model is described as a 'new concept', which evolves through loops of prototype building and testing. Prototyping in this case, is regarded as a mechanism for theory testing and refinement. It also reflects back to the new concept under investigation.

3.2.2 User-Sensitive Systems Development Research Process

Venable (2006) argues that systems development is one subset of design science. The author proposes an activity framework for design science research that advocates theorising as the main research purpose. The framework suggests theory building as the central research activity interacting with problem diagnosis, technology design, and technology evaluation. Adapted from this model, Figure 3.1 below illustrates how this thesis addresses the problem of user-sensitive RQ assessment in health information portals, using socio-technical design and evaluation approaches, and the central role is played by theory building.

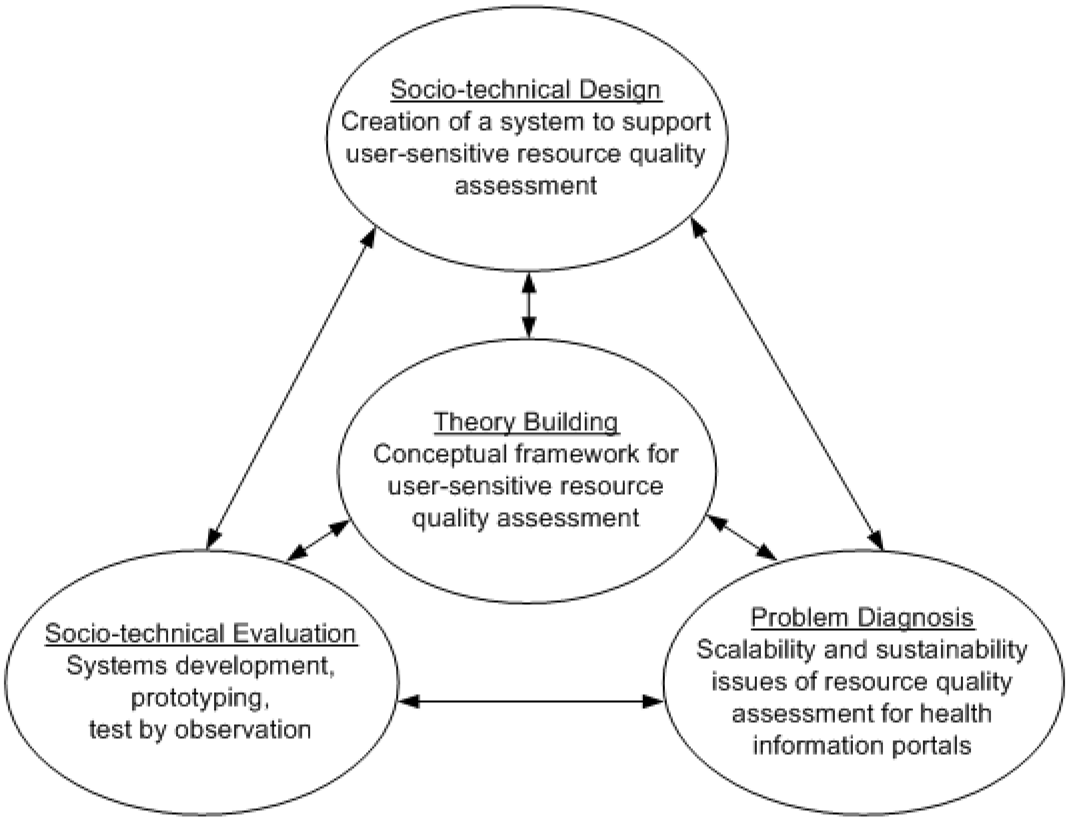


Figure 3.1 The role of theorising in this research, adapted from (Venable 2006, p. 17)

In this research, the central activity is the conceptualisation of user-sensitive RQ assessment. This concept building activity is regarded as an integral part of theory building. New knowledge is generated from the concept modelling process, and contributes to the theories underpinning the new concept. The emerging new field knowledge is tested and refined, via prototype building and evaluation in natural settings.

As can be seen from Figure 3.1, the focal point of this research is the conceptualisation of user-sensitive RQ assessment, which was developed and refined via a user-sensitive systems development research process. The process included three research phases, namely concept building, system building, and system development. Detailed research design of each phase is presented in the following section.

3.3 Research Design

3.3.1 The Related Smart Information Portals Project

As mentioned in the first chapter, this research is related to a larger Monash University led research project Smart Information Portals (SIP), which was funded by an Australian Research Council (ARC) Discovery Grant (2006-2009). The SIP project aimed at developing and evaluating approaches that used intelligent features to support user-centred quality online information provision to better meet the information and decision support needs of health information communities (SIP 2006).

The SIP project drew on the results of the Breast Cancer Knowledge Online (BCKOnline) project, which conceptualised and developed a user-sensitive information portal to meet the diverse knowledge needs of the Victorian breast cancer community. The BCKOnline project (BCKOnline 2009) was funded through an ARC Linkage Grant (2002-2003) with support from the government initiative BreastCare Victoria, and the industry partner Breast Cancer Action Group (www.bcag.org.au). The BCKOnline project investigated and analysed diverse information needs amongst women with breast cancer and their families. The project identified and analysed knowledge resources available at the time, and designed a scheme for the description of those resources in order to support differentiated online information access.

The SIP project adopted a user-sensitive design methodology and metadata-driven approaches, to build a conceptual model for smart information portals. A number of SIP concepts were defined and explored, including knowledge repository management, adaptivity and personalisation, smart information retrieval, quality control and metadata generation. This thesis addresses related quality control issues and the generation of metadata about quality. Relationships between this PhD study and the SIP project are discussed further in Section 3.4.

3.3.2 Research Design Overview

According to the three-step system development research procedure (Burstein 2002) proposed on the basis of Nunamaker et al. (1991), this study was conducted in three research phases, including concept building, system building and system evaluation. Figure 3.2 illustrates major research activities undertaken in each phase and their relationships.

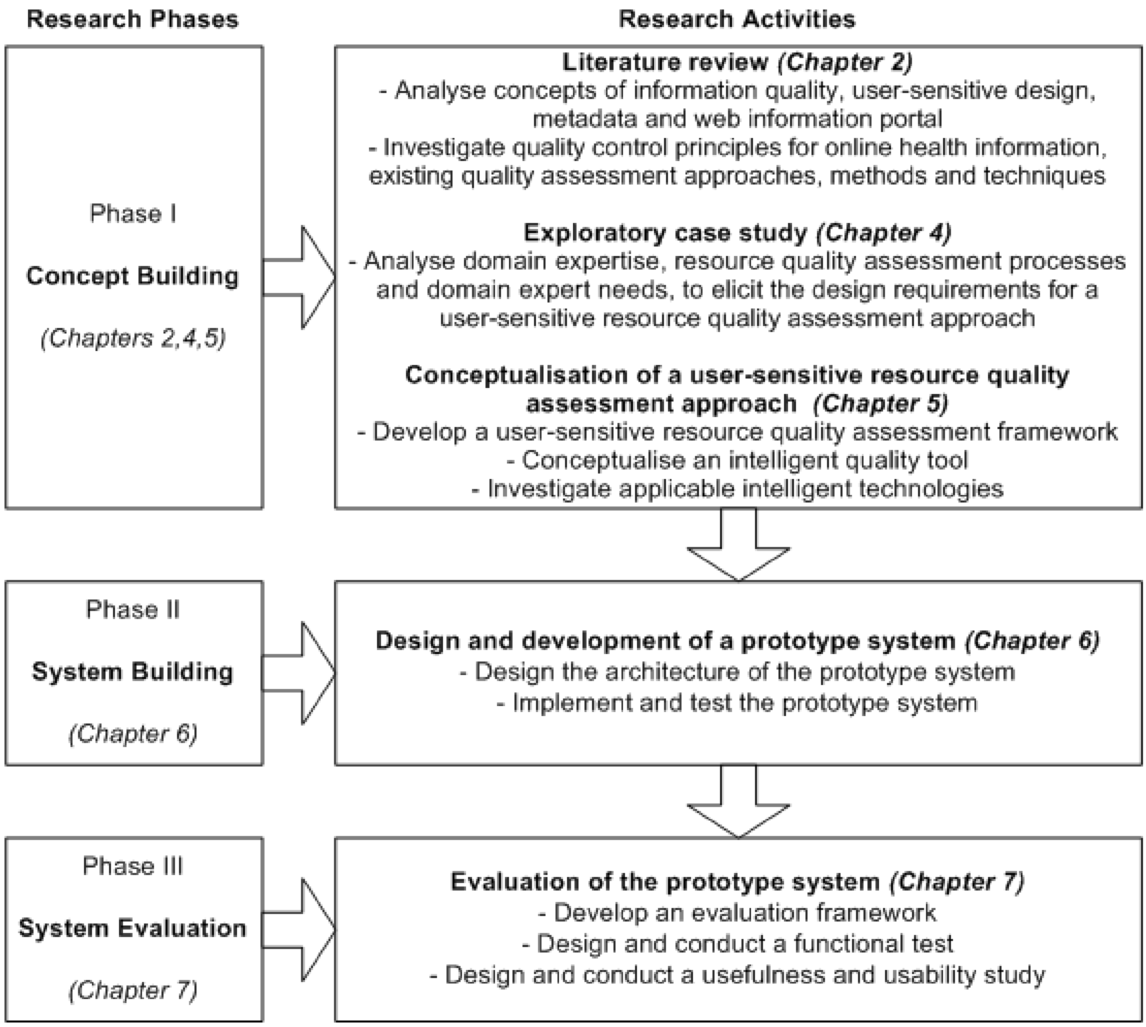


Figure 3.2 Research design overview

From Figure 3.2, it can be seen that each research phase involved a number of research activities. These research activities are introduced in the following sections.

- Literature review and analysis (see Section 3.3.3.1): justification of literature relevancy.
- Exploratory case study (see Section 3.3.3.2): justification of data collection methods and instruments.
- Conceptualisation of a user-sensitive resource quality assessment approach (see Section 3.3.3.3): justification of conceptual analysis process and techniques.
- Design and development of a prototype system (see Section 3.3.4): justification of the system building process.
- Development of an evaluation framework (see Section 3.3.5.1): justification of evaluation criteria selection.

- Design and development of a functional test (see Section 3.3.5.2): justification of data collection methods and instruments.
- Design and development of a usefulness and usability study (see Section 3.3.5.3): justification of data collection methods and instruments.

The data collection procedures and data analysis results, of these research activities, are elaborated in great detail in the following chapters of this thesis (See chapter numbers in Figure 3.2).

3.3.3 Concept Building Research Phase

In light of Hevner et al.'s (2004, p. 80) IS research framework, in order to justify the rigour and relevance of IS research, both the knowledge base (theories) and the environment (problem domain) of the research project need to be investigated in depth and breadth. As a result, three major research tasks were undertaken in this stage, including a review and analysis of the relevant literature, which has been extensively discussed in Chapter 2, an exploratory case study of RQ assessment practice, and a conceptualisation of a new RQ assessment solution that takes account of user information needs and quality perceptions for health information portals.

3.3.3.1 Literature review and analysis

As presented in the previous chapter, a literature warrant analysis was conducted to establish the theoretical grounds for this study. A comprehensive analysis of multi-disciplinary research literature investigated the user-sensitive RQ assessment issue from a theoretical, contextual and technological perspective. Information quality and its assessment provided useful theories for constructing and measuring quality. The quality control issues and principles in the consumer health informatics field defined the specific context for this research. The literature of digital library and web information portals was also reviewed, to seek existing metadata-based technical solutions, which might be applicable to supporting process automation.

3.3.3.2 Exploratory case study

In order to obtain a better understanding of user-sensitive RQ assessment in a real-world setting, an exploratory case study was developed and conducted in this research solely by me. The study investigated two metadata-driven and user-sensitive health information portals: the BCKOnline portal and the Heart Health Online (HHOnline) portal. The two portals provided a rich context to explore the following questions:

- What kinds of domain expertise are required for RQ assessment?
- What tasks and domain expert activities are involved in RQ assessment?

- What problems do domain experts encounter with RQ assessment? What are potential solutions to tackle the problems?
- What part of the processes can be automated? Where is human intervention needed?

A part of data collection and data analysis of this case study was conducted through a joint domain expertise study with the SIP project team. Semi-structured interviews were used for data collection, while the analysis of the data was shared. I identified relevant interview questions, conducted interviews, and transcribed all interview discussions. The collected data was analysed with a shared interest to articulate the role domain expertise plays in smart, user-sensitive health information portals (Evans et al. 2009). In addition to the domain expertise analysis, this PhD research further independently analysed the interview data independently with the focus drawn to the user-sensitive RQ assessment issues. Additional data was also collected to complement the interview data in order to analyse the RQ assessment processes, domain expert needs, and the design requirements of RQ assessment solutions. Data collection methods and instruments utilised in this case study are introduced below.

Selection of participants

This case study aims to investigate the user-sensitive RQ assessment issues in the context of metadata-driven health information portals. In order to obtain an in-depth understanding of domain expert activities involved in assessing resource quality and their needs for intelligent support, those who had practical experience in health portal content management were interviewed. Eligible candidates had to satisfy a number of selection criteria, including knowledge of health informatics, experience of using metadata for resource classification and description, as well as a good understanding of health information consumers. Unavoidably, these selection criteria narrowed down the number of eligible candidates. However, the quality and depth of the collected data was the major concern of this qualitative study, rather than the sample size or the saturation in data collection.

According to Crouch and McKenzie (2006), qualitative studies are concerned with meaning instead of making generalised hypothesis statements. Mason (2010) also argues that more data does not necessarily lead to more information. Therefore, for this study, the main concern was to assure that the data was collected from actual problem owners, who were able to provide first-hand insights on user-sensitive RQ assessment issues. Due to a limited access to health information portals and eligible candidates, the study used a relatively small number of participants, who were working for two metadata-driven and user-sensitive health information portals: BCKOnline and HHOnline portals. It is argued that these three participants are representative enough to illustrate the domain expert needs for intelligent support and design requirements of RQ assessment approaches.

Data collection methods

Data collection methods included semi-structured interviews together with cognitive walkthroughs of portal content management processes. Data were collected individually from three domain expert participants, who had been working for the BCKOnline and HHOnline portals. The case study analysed portal development documentation, including resource description schemas and resource evaluation criteria. The two web portal systems and existing portal data were also analysed in order to obtain a better understanding of associated issues, for instance, to identify the gap between the decision-support needs of domain experts and the existing functionality of portal content management systems.

Data collection instrument

Researchers involved in the joint domain expertise study created an interview protocol that contained two sets of open-ended questions. The first part of the interview reflected on the role domain experts play in portal design and development. Questions were formulated to elicit domain expert tasks and activities, in relation to resource identification, selection and description.

The second part of the interview took place in the context of a walkthrough of resource identification, selection and description exercises with the portal's domain expert interface and metadata schema. A set of resources was determined (from the first part of the interview) and the completion of the metadata schema for each resource was undertaken. Questions were asked regarding the sourcing of metadata values and the decision-making processes of the domain expert in completing those values. Opinions on the usability of the interface and the metadata schema and what might be done to improve it were sought. This interview instrument is provided in **Appendix G**.

Benbasat et al. (1987) proposes eleven criteria for conducting a case study in information systems research. Table 3.2 summarises how these key characteristics of case study research were addressed by this study.

Table 3.2 Addressing Benbasat et al.'s (1987, p. 371) key characteristics of case study research

Key characteristics of case studies	Application in this case study
1. Phenomenon is examined in a natural setting.	Data collection occurred in multiple locations (computer labs and meeting rooms) at Monash university.
2. Data are collected by multiple means.	Data were collected through semi-structured interviews, walkthrough observations, and examination of portal documentation.
3. One or few entities (person, group or organisation) are examined.	Formal contacts with domain experts who worked for the BCKOnline and HHOnline portals; Informal contacts with the portal developers.
4. The complexity of the unit is studied intensively.	This case study was concerned with the tasks and activities involved in RQ assessment, as well as problems domain experts encountered and their needs for intelligent support. Therefore, a number of aspects of the two portals were investigated, including the role of domain experts, RQ assessment processes and portal content management systems.
5. Case studies are more suitable for the exploration, classification and hypothesis development stages of the knowledge building process. The investigator should have a receptive attitude towards exploration.	The objectives of this case study were: first, to understand how user-sensitive RQ assessment was carried out in practice, and then to elicit design requirements for a new solution to support user-sensitive RQ assessment.
6. No experimental controls or manipulation are involved.	These were no experimental controls or manipulations involved.
7. The investigator may not specify the set of independent and dependant variables in advance.	Independent and dependant variables were not specified in advance.
8. The results derived depend heavily on the integrative powers of the investigator.	Data were collected by using multiple techniques, including semi-structured interviews with domain experts, walkthrough observations, documentation analysis, and portal systems and data examination. Each set of data disclosed one aspect of the RQ assessment issue being investigated. They were synthesised to form a complete view of what user-sensitive design meant to RQ assessment.
9. Changes in site selection and data collection methods could take place as the investigator develops new hypotheses.	The type of data to be collected was extended when the concepts of RQ and RQ assessment were further clarified during the research
10. Case research is useful in the study of "how" and "why" questions because these deal with operational links to be tracked over time rather than with frequency or incidence.	The exploration of how the user-sensitive design philosophy impacts on the RQ assessment practice was the major concern of this case study. Quality assurance in a portal is fundamentally achieved by the user-sensitive RQ assessment processes undertaken by domain experts.
11. Focus is on contemporary events.	The focus of this case study was to provide systematic and intelligent RQ assessment solutions to standardize and scale the manual process.

Although this case study was conducted jointly with the SIP project, it collected and analysed data to serve the purposes of this research, i.e. to articulate the key concepts of RQ assessment processes, domain expert needs, and design requirements of RQ assessment solutions.

3.3.3.3 Resource quality assessment solution conceptualisation

Based on the outcomes of the literature review and the case study, a user-sensitive quality assessment solution was conceptualised using the method described in Wang et al. (1995). Wang et al. (1995) systematically analyse IQ research, in the field of organisational databases, using a set of International standards for quality management systems (ISO 9000 to ISO 9004 inclusive). The authors conclude that the development of data quality management solutions must establish quality requirements for data products, and translate them into technical specifications for data manufacturing systems. Wang et al. (1995) identify the following three components of a data quality management solution (p. 625):

- definition of data quality dimensions and measurement of their values with metrics and models;
- analysis and design of data quality aspects; and
- design of data manufacturing systems that incorporate data quality aspects.

The above three elements can be described as quality characterisation, data modelling, and system design and integration, which in general define how to assess IQ in a systematic way. Accordingly, for this research, the conceptualisation of a user-sensitive RQ assessment solution should incorporate the development of a RQ assessment framework, a resource description scheme that defines resource quality attributes, and a corresponding quality tool to assist the user-sensitive RQ assessment in the context of health information portals. The existing BCKOnline user-centred resource description scheme was considered as a suitable data model to constitute the solution. Therefore, this research focused on the design and development of the other two components, which were the quality framework and the quality tool.

Resource quality assessment framework development process

Building frameworks is regarded as a legitimate approach to theory building (Porter 1991). Porter describes the aim of frameworks as follows:

> "Frameworks identify the relevant variables and the questions which the user must answer in order to develop conclusions tailored to a particular industry and company. […] The theory embodied in frameworks is contained in the choice of included variables, the way variables are organized, the interactions among the variables, and the way in which alternative patterns of variables and company choices affect outcomes" (p. 98).

In addition, although not all the interactions among the many variables in the frameworks can be rigorously drawn, the frameworks seek to help the analyst to better understand the problem and develop better solutions (Porter 1991). Derived from this view, this research defines a framework as an operational structure that encompasses the concepts and problem-solving solutions to illustrate embedded theories. In this sense, the development of RQ assessment framework is recognised as a means to conceptualise user-sensitive RQ assessment for achieving both theory building and practice improvement.

Development of the user-sensitive RQ assessment framework adapted a method proposed by the Canadian Institute of Health Information (CIHI), which evaluates data quality using a four-level conceptual model (Long and Seko 2002). As depicted in Figure 3.3, at the foundation level, various quality criteria must be analysed from published quality frameworks or criteria models. These criteria then need to be aggregated into the second level of quality characteristics. At the third level, based on the identified quality characteristic, a set of quality dimensions can be defined for the overall quality evaluation at the fourth level. As a result, the CIHI framework derives five quality dimensions (*Accuracy*, *Timeliness*, *Comparability*, *Usability* and *Relevance*) from 19 quality characteristics and 61 quality criteria (CIHI 2009). The five quality dimensions are eventually measured and aggregated into one overall evaluation value, using computational algorithms.

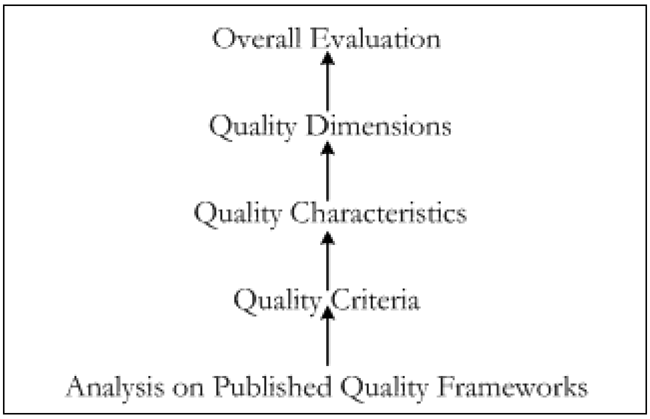


Figure 3.3 Conceptual model of the four-level CIHI data quality framework, adapted from Long and Seko (2002)

The CIHI data quality framework is designed for application in a specific database context and to address the needs of a specific group of database users (CIHI 2009, p. 2). So the framework as originally developed is not suitable for addressing the RQ assessment issue in a portal context. However, the method can be adapted to develop a user-sensitive RQ assessment framework for health information portals using the following process:

1. Analysis of published IQ assessment frameworks in the domains of organisational and web information systems, and consumer health informatics (the 1st level, see Sections 2.2 and 2.3).
2. Analysis of quality criteria for the evaluation of online health information from the user's perspective (the 2nd level, see Section 5.2.1).
3. Selection and definition of quality dimensions in a hierarchy (the 3rd level, see Section 5.2.2).
4. Definition of assessment metrics for measuring quality dimensions, and the aggregation of assessment values to an overall qualitative measure of RQ (the 4th level, see Section 5.3).

It needs to be clarified that this research assesses RQ dimensions and aggregates evaluation results to an overall quality appraisal via a qualitative approach. This approach is in contrast to the computational approach CIHI employs to quantify the overall quality measure at the fourth level of the model.

The third component of the quality management solution is a quality data manufacture system that incorporates data quality aspects (Wang et al. 1995). What is required in this research is a quality tool, as part of a portal content management system, to implement the quality framework delivered from the process described above. The design of the conceptual model for the quality tool needs to draw on the design requirements elicited from the case study. In addition, available intelligent technologies needed to be investigated for their application in providing the required functionality.

3.3.4 System Building Research Phase

The second phase of this research employed a system prototyping approach to verify the feasibility of the proposed RQ assessment solution.

The system building phase contained a number of research activities associated with the design and development of a prototype system, namely the Domain Expert Dashboard (DED). These activities are outlined below:

- identify system functionalities in response to the design requirements;
- design system architecture, describing the components and their interrelationships;

- define processes to carry out system functions;
- identify alternative solutions, and select a solution for the system;
- develop and test the system, to verify the feasibility of the proposed solution; and
- gain insights about the problems and the system complexity.

Details of the system building process and the developed DED are presented in Chapter 6.

3.3.5 System Evaluation Research Phase

Evaluation of design artefacts is a key component of the systems development research process, for the further refinement of the design. According to Hevner and Chatterjee (2010), design research has two specific goals: to evaluate the design of an artefact, and to establish the utility of the artefact in an application field. It is crucial for design science research to evaluate not only the design of an artefact, but also the actual utility and efficacy of the design artefact in solving a real problem (Hevner et al. 2004). For this research, the artefact evaluation aimed at testing and demonstrating the feasibility, utility, and efficacy of the proposed solution for addressing the scalability and sustainability issues of RQ assessment. It also sought to identify areas for further improvement so as to increase the rigour and validity of the proposed RQ assessment solution.

Pries-Heje et al. (2008) suggest that a design artefact can be evaluated via an 'ex ante' approach prior to its construction, or via an 'ex post' approach after the artefact is created. They propose a strategic framework to analyse both the design process and design artefacts by questioning "*when* evaluation takes place, *what* is actually evaluated, and *how* it is evaluated" (p. 260). This research delivered three design artefacts: a user-sensitive RQ assessment framework; the conceptual architecture of an intelligent quality tool that implemented the framework; and a prototype system (the DED) that instantiated the proposed quality tool in a real-world scenario. 'Ex post' artefact evaluation was conducted on the utility and efficacy of the DED prototype system. As a result, a functional test was carried out to assess the implemented intelligent feature, via a series of machine learning experiments. Effects of the DED, on RQ assessment processes and outcomes, were empirically evaluated with domain experts via a usefulness and usability study.

3.3.5.1 Evaluation framework development

For the purpose of evaluating the utility and efficacy of the proposed solution in improving RQ assessment, it is necessary to define how the potential improvement in RQ assessment can be measured. This research evaluated the effects of the DED prototype system on both the decision-making processes and the outcomes of RQ assessment.

Both the usefulness and usability of the prototype system were of concern to this evaluation. The prototype system was designed and developed to support RQ assessment, by providing greater functionality and improved usability. It is believed that poor usability will undermine the usefulness of the system functionalities being evaluated, and thereby negatively influence the overall user satisfaction.

Usefulness is mainly concerned with the functions and features of a system. *Usefulness* and *Ease of use* are regarded as significant determinants to the use of a technology (Bagozzi et al. 1992; Davis 1989). Davis (1989) defines *Perceived Usefulness* as "the degree to which a person believes that using a particular system would enhance his or her job performance" (p. 320). *Perceived Ease of use* is defined as "the degree to which a person believes that using a particular system would be free from effort" (p. 320).

Usability on the other hand, concerns whether or not users are able to use system functions and features. The international standard ISO 9241-11:1998 provides guidance on *Usability,* and defines it as: "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO/IEC 1998). In alignment with this definition, John and Marks (1997) use the following three factors to measure the *Usability* of a system.

- Effectiveness: the extent to which users achieve the intended goals of use of the overall system (can users achieve what they need to do).
- Efficiency: the resources (e.g. time, costs, or effort) that are required to achieve the intended goals (how much effort do users have to spend).
- Satisfaction: the extent to which users find the overall system is acceptable (how do users feel about their interaction with the system).

Criteria utilised for evaluating RQ assessment processes and outcomes were derived from the literature of decision support systems (DSS), in particular the Analytic Hierarchy Process (AHP) method. The AHP method combines both process-oriented and outcome-oriented measures for assessing DSS effectiveness (Forgionne 1999). Based on Forgionne's integrated AHP model, Phillips-Wren et al. (2006) propose a multi-criteria model that includes both process and outcome measures for the evaluation of intelligent decision-making support systems (i-DMSS). Six criteria from Phillips-Wren et al.'s (2006) evaluation model were considered relevant to measure the effects of the DED prototype system, on RQ assessment processes and outcomes (see Table 3.3).

Table 3.3 Decision-making process and outcome measures, derived from Phillips-Wren et al.'s (2006, p.17) evaluation model

Measures	Evaluation Criterion
Process - Effectiveness	
- *Usefulness*	System functions and features are helpful to achieve what need to be done
Process - Efficiency	
- *Ease of use*	More efficient
- *Time and effort*	Faster decision
Process - Satisfaction	
- *User* (domain expert) *satisfaction*	The overall system is usable and supports decision-making and learning
Outcome - Quality	
- *Accuracy*	Machine suggested values are accurate and reliable
- *Consistency*	Comparison to expert opinion

As can be seen from Table 3.3, the process measure linked system Effectiveness to *Usefulness*, system Efficiency to *Ease of use* and *Time and effort*, and *Satisfaction* to user (domain expert) perceptions on the overall usefulness and usability of a system to fulfil required tasks. The outcome measure was drawn on the predicting accuracy of the intelligent feature, and the consistency of RQ assessment results.

Based on these selected evaluation criteria, an evaluation framework was developed for guiding the prototype system evaluation phase of this research. Figure 3.4 illustrates how the selected evaluation criteria are addressed by employing multiple evaluation methods.

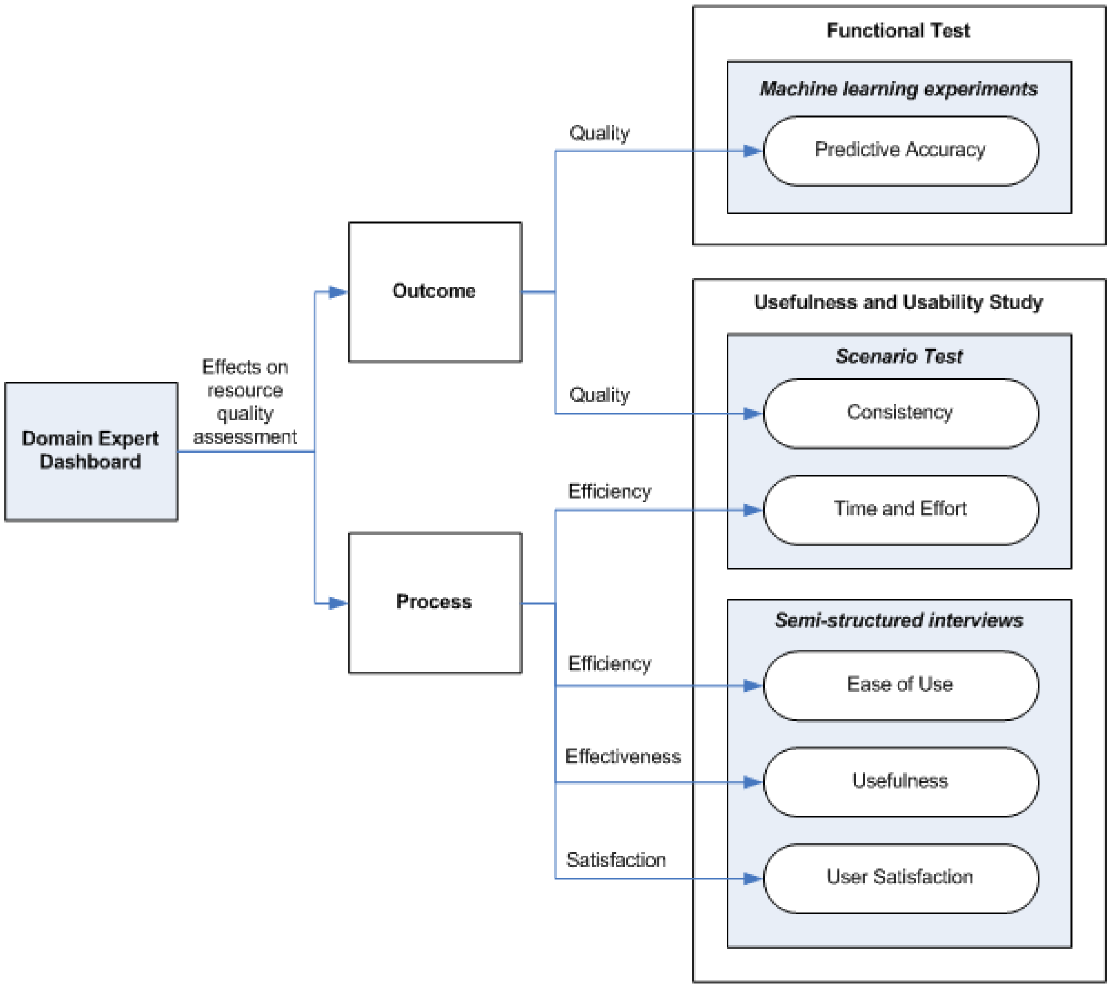


Figure 3.4 Framework for evaluating the prototype system

In this framework, the utility of the DED prototype system, in supporting and facilitating RQ assessment, is evaluated through both process-oriented and outcome-oriented measures. The evaluation is concerned with whether the system enhances the efficiency and effectiveness of the quality assessment processes, and whether the quality of assessment outcomes is improved.

According to the multi-criteria evaluation framework, different approaches were applied to conduct the evaluation according to the nature of the criterion being assessed. Two evaluation processes were undertaken, including an internal functional test and a usability and usefulness study with domain expert participants. An internal functional test was conducted for measuring the *Predictive accuracy* of the implemented intelligent feature via a series of machine learning experiments. The internal evaluation focused on the functional mode of operation, i.e. whether the system executed the intelligent feature in an appropriate way. Machine learning experiments provided statistical evidence that the implemented intelligent feature was functioning. The *Consistency*, *Time and effort*, *Ease of use*, *Usefulness*, and *User satisfaction* aspects of the DED were empirically evaluated via a usefulness and usability study with domain expert participants. The design of these two evaluation studies is further discussed below.

3.3.5.2 Functional test

The internal functional evaluation aimed at testing whether the implemented intelligent feature was functioning and executed in an appropriate way. The functional test is usually conducted by running a series of designed test cases. In this case, a series of machine learning experiments was carried out to ensure that the accuracy and reliability of machine-generated values were satisfactory. Without this type of test, the operation of the intelligent feature itself would be questionable, and would affect the potential usefulness of the feature in supporting decision-making on RQ.

3.3.5.3 Usefulness and usability study

The main purpose of the DED prototype system evaluation was to demonstrate the feasibility of the proposed semi-automated RQ assessment approach. In addition to the functional test, the DED prototype system was also empirically evaluated by domain expert participants via a usefulness and usability study. In this study, the evaluation data were collected through scenario tests and semi-structured interviews.

Selection of evaluation data collection methods

According to Zaphiris et al. (2006), there are a number of evaluation methods that are applicable for knowledge elicitation and usability evaluation via user assessments. These include interviewing, surveys, focus groups, observation, paper prototyping, cognitive walkthrough, and heuristic evaluation. Likewise, Brender (2006) investigates the evaluation methods for health informatics and identifies a few popular methods, including field study, focus group interview, think aloud, cognitive walkthrough, questionnaire, and interview. Amongst these methods, interviews are recommended as "particularly suited for elucidation of individuals' opinions, attitudes, and perceptions regarding phenomena and observations" (p. 66). Based on the exploratory nature and the need to capture user-sensitive characteristics, the usefulness and usability study mainly employed semi-structured interviews. In addition, in order to compare the decision-making processes and outcomes with and without the use of the intelligent feature, scenario tests were also conducted.

Selection of participants

As the design requirements for the DED prototype system were elicited from the exploratory case study of two real portal systems, it was considered that those who had participated in the case study, or had worked on the two portals, were ideal candidates to evaluate the usefulness and usability of the prototype system. Moreover, the domain expertise study articulated the attributes of domain expertise, and proposed a quadrant diagram to measure the level of expertise in one dimension against the other. The diagram provided an instrument for selection of DED evaluation participants. It was considered that the Novice/Novice quadrant described lay users rather than eligible domain experts. As a result, three domain expert participants were selected, each representing one quadrant in the domain expertise diagram.

Scenario test instrument

The scenario test was designed to measure the *Consistency* of RQ assessment outcomes and the *Time and effort* domain experts spent in making value judgements. Domain expert participants were asked to undertake tasks requested by the scenario test before participating in the semi-structured interview. About half an hour was required to complete the scenario test.

A scenario test instrument (see **Appendix H**) was developed to guide the participant to complete tasks in two different scenarios: with or without metadata value suggestions. In total, the participant was asked to describe quality attributes for four online resources, using the encodings defined by the BCKOnline metadata schema.

The four resources used by the scenario test were carefully selected from the BCKOnline metadata repository. The following criteria were used for selecting resources:

- use four new resources that were fully described (their metadata records must be complete), but not be published or indexed by the portal;
- select one resource from a popular, large-scale health website owned by an organisation or a commercial body;
- select one resource from a personal website;
- select one resource from a website that was known by the portal, i.e. one or more existing resources published or indexed by the portal were from the same website;
- select one resource from a new website that none of existing resources published or indexed by the portal were from the same website.

In the first scenario, the participant was asked to assign ten quality attribute values for two resources, without referring to metadata value suggestions. In the second scenario, the participant was asked to assign ten quality attribute values for another two resources, with value suggestions generated by the machine. The instrument collected two types of data, including quality attribute values assigned by the participant and the time spent for completing each task. An introduction sheet of DED functions and features was attached to the scenario test instrument as the background information.

Semi-structured interview instrument

Semi-structured interviews were conducted as part of the usefulness and usability study, in order to collect participants' opinions about their use experience of the prototype system for RQ assessment. An interview instrument (see **Appendix I**) was created accordingly.

The instrument contained two parts. The first part was an introduction to the system functions and features being evaluated. The second part consisted of twelve interview questions grouped into three sections. The first four questions were designed for profiling a participant, focusing on the portal content management experience and the level of domain expertise. The fifth question was concerned with the *Ease of use*, *Usefulness* and *User satisfaction* of multiple system functions and features developed for assisting RQ assessment. Such a multi-criteria and multi-target evaluation was organised in a questionnaire. The last section included seven open-ended conclusive questions regarding the strengths and weaknesses of the prototype system. Participants' use experience of the intelligent feature gained from the scenario test, was also covered during the interview discussion.

As introduced above, in order to evaluate the prototype system in a constructive way, a questionnaire was used to complement the open-ended interview questions. Individual participants were asked to complete the questionnaire during their own interview sessions. Although the questionnaire technique is usually not used as part of interviews, it is necessary for this evaluation study to have some control over how and when the questionnaire was answered. This study recruited a limited number of participants, thus it was necessary to ensure the quality of collected data. Having the questionnaire answered during the interview not only made the participants being more conscientious with their responses, but also enabled me to acquire extra information about how and why the participants come up with their answers.

The questionnaire was designed and developed based on Fisher et al. (2004)'s usability instrument for evaluating websites. The instrument has been extended and validated via the use of a factor analysis (Fisher et al. 2008). Questions proposed by this instrument (Fisher 2009) were examined for their application in this evaluation study. The original instrument includes a number of questions on topics of navigation, information display, information quality, ease of use, system features, and user satisfaction. This research consolidated original questions to fewer but more explicit and concise questionnaire items, which better addressed the purpose of this evaluation study. As a result, 17 questions were developed. These include eight questions evaluating individual data management functions, six questions assessing intelligent features, and three questions concerning with overall user satisfaction. This research also associated the 17 questions to the three evaluation criteria, and categorised all questions correspondingly. By this means, the Ease of Use, Usefulness, and Use satisfaction aspects of the DED prototype system were evaluated explicitly. Table 3.4 outlines these questionnaire items and the corresponding evaluation criteria they address.

Table 3.4 Summary of questionnaire items

Item code	Questionnaire item	Evaluation criterion
E1	It is quick and flexible to open or switch between data management views	Efficiency – Ease of use
E2	It is easy to browse data.	
E3	It is easy and flexible to add or edit data in data entry forms.	
E4	It is easy, flexible and safe to delete data records.	
E5	It is easy and flexible to sort data.	
E6	It is easy and flexible to retrieve data.	
E7	I can use the above features without written instructions.	
E8	I can use it successfully and can recover from mistakes quickly and easily.	
U1	Data visualisation in the workbench view is useful for monitoring resource distribution and value consistency.	Effectiveness – Usefulness
U2	The URL checking and reporting feature is helpful.	
U3	Grouping metadata in different Tabs helped me to better describe a resource in the data entry form.	
U4	The metadata value suggestion feature is helpful.	
U5	By using the above mentioned features, I felt more confident about the quality of metadata values I assigned.	
U6	I found the above mentioned features would help me to be more productive and consistent.	
S1	I am satisfied with the system for data management.	Satisfaction – User satisfaction
S2	The system works the way I want it to work.	
S3	I prefer to use the system comparing to the other similar systems.	

In the questionnaire, each item was measured by a five-point Likert scale (i.e. strongly disagree, disagree, neither agree nor disagree, agree, strongly agree). According to Hinkin (1995), a five-point Likert-type scale is adequate for the use in this study. Further details on prototype system evaluation are presented in Chapter 7.

It needs to be clarified that this Usefulness and Usability study only employed qualitative data analysis techniques, although the data collection instruments included a questionnaire component. Ratings on questionnaire items were not used for a typical quantitative analysis. Instead, they indicated the strength of feelings the participants had on aspects that were evaluated. The participants were also asked to explain their ratings during the interviews. This combined data enabled the evaluation of the usefulness and usability of the system functions and features, as well as the identification of areas for design improvement.

3.3.6 Summary of Employed Methods and Techniques

In summary, this research invoked mixed data collection methods and data analysis techniques, to serve the purposes of design artefact conceptualisation, construction and evaluation. Table 3.5 summarises the major activities of this research, and the related data collection methods and data analysis techniques that have been discussed in this chapter. The table also references the chapters which elaborate each research activity, its data analysis procedures and results.

Table 3.5 Summary of major research activities, methods and techniques

Major research activities	Related data collection methods or data analysis techniques	Chapter in this thesis
Literature analysis	An analytical review over relevant concepts, constructs, theories, principles, methodologies and technologies	Chapter 2
Exploratory case study	Semi-structured interviews with domain experts; analysis of portal design and development documentation, the web portal systems, and existing portal data	Chapter 4
Conceptualisation of a user-sensitive resource quality assessment approach	Concept building based on literature review and requirements engineering	Chapter 5
Design and development of a prototype system	System prototyping	Chapter 6
Evaluation of the prototype system against the proposed evaluation framework	Functional (black box) test Usefulness and usability study via scenario tests and semi-structured interviews with domain experts	Chapter 7

In addition to these research activities, experience gained from the whole socio-technical design science research process was consolidated for further refinement. Such a reflection and future research directions are discussed in the last chapter of this thesis (Chapter 8).

3.4 Relationships between this PhD Study and the SIP Project

This section aims to clarify the relationships between this PhD study and the SIP project. As mentioned previously in Section 3.3.1, this research was related to the SIP project. It needs to be clarified that this research was conducted as an independent project that had a clear boundary with SIP research activities and outcomes. This PhD study adopted the user-sensitive design philosophy advocated by the SIP project. The research questions were formulated to explore what user-sensitive design meant to RQ assessment in the context of health information portals. This study contributed to the conceptualisation and refinement of the SIP architecture by modelling, prototyping, and evaluating a user-sensitive RQ assessment approach. The relationships between the research activities of this PhD study and the SIP project are illustrated in Figure 3.5.

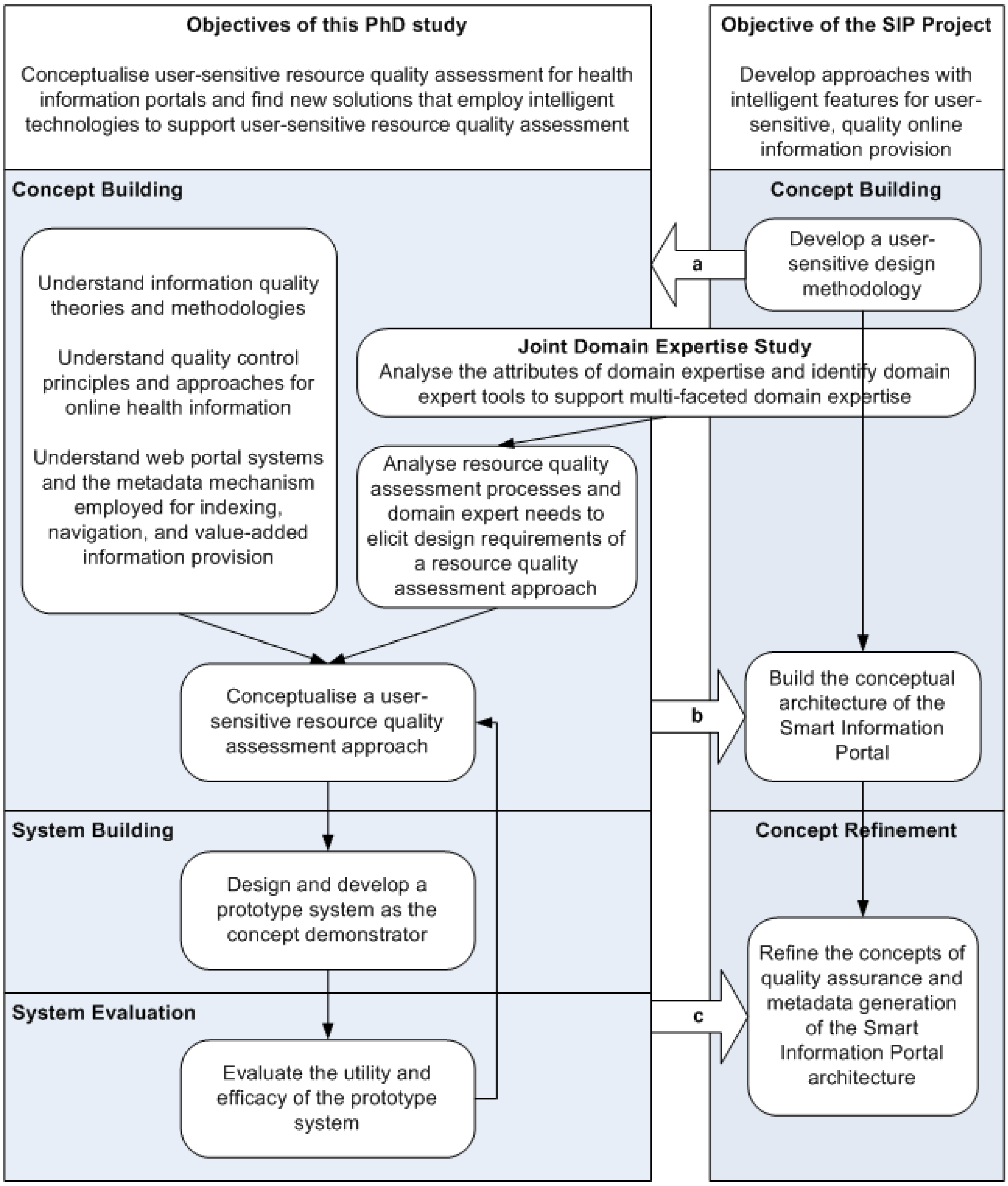


Figure 3.5 Relationships between this PhD study and the SIP project

The SIP project was conducted in two research phases, including concept building and concept refinement. The concept building phase utilised various data collection and analysis methods, such as literature review and case studies, to build the conceptual architecture of SIP. The concept refinement phase consolidated the research findings and experience to refine the constructs of SIP, and to provide SIP design principles and implementation guidelines.

As depicted in Figure 3.5, informed by the user-sensitive design methodology developed within the SIP project, this research investigated the concept of RQ from the perspective of portal users, and explored solutions to support user-sensitive RQ assessment for the healthcare domain. Drawing on the comprehensive review of the relevant literature, a case study was conducted jointly with the SIP project. The case study investigated the attributes of domain expertise and conceptualised a set of domain expert tools to support the articulated domain expertise, which addressed the common interests of the two projects. The PhD research focussed on aspects of the case study relating to RQ assessment practice and their implications for the design requirements of a RQ assessment solution. Therefore, data collected from the joint case study was analysed independently to serve this purpose. The resultant user-sensitive RQ assessment framework provided a conceptual construct of the quality control issue, for smart information portals. It also provided the conceptual model of an intelligent quality tool. The quality framework together with the quality tool was proposed, as one of domain expert tools for the SIP architecture. Furthermore, the feasibility, utility, and efficacy of the proposed RQ assessment approach were tested through the development and evaluation of a prototype system. Although SIP did not employ a systems development research approach, the developed system can be used to construct a SIP prototype. Table 3.6 below describes the SIP inputs to this PhD study as well as the contributions of this PhD study to the SIP project.

Table 3.6 Relationships between this research and the SIP project

<table>
<tr><th colspan="2">Input to this PhD study: the user-sensitive design methodology and the SIP approach</th></tr>
<tr><td colspan="2">a. The user-sensitive design methodology and principles for the design and development of smart information portals</td></tr>
<tr><th colspan="2">The joint domain expertise study</th></tr>
<tr><td>Joint research activities:
▪ Collect data via semi-structured interviews with domain experts;
▪ Analyse portal design and development documentation, including metadata schemas and resource evaluation guidelines;
▪ Articulate attributes of domain expertise;
▪ Identify domain expert tools to support the multi-faceted domain expertise;
▪ Develop quadrant diagrams to represent and measure aspects of domain expertise.</td><td>Independent research activities:
▪ Analyse portal systems and existing portal data;
▪ Analyse quality appraisal tasks, activities, and problems;
▪ Articulate RQ assessment processes;
▪ Identify domain expert needs for intelligent support;
▪ Elicit design requirements for a RQ assessment approach, which forms part of domain expert tools.</td></tr>
<tr><th colspan="2">Contributions of this PhD study to the SIP architecture</th></tr>
<tr><td colspan="2">b. Conceptualisation of the user-sensitive RQ assessment issue and the functional requirements for supporting RQ assessment in the context of SIP.
c. Further refinements of the SIP functional requirements and architecture, by modelling and prototyping a new solution, for standardising and facilitating user-sensitive RQ assessment.</td></tr>
</table>

It needs to be further clarified that a number of research outcomes resulted from the joint domain expertise study. Although this research contributed to the construction of these joint research outputs, they are not claimed as independent deliverables of this research. These include the articulation of multi-faceted domain expertise attributes for effective content management, two quadrant diagrams for representing and measuring the aspects of domain expertise, and a set of domain expert tools proposed for supporting multi-faceted domain expertise.

3.5 Adaptation of Hevner et al.'s Framework for Design Science Research

In summary, this research aimed to build and refine a user-sensitive RQ assessment approach through artefact design, construction and evaluation. Figure 3.6 illustrates how this research was conducted using a combination of Hevner et al.'s (2004 p. 80) framework for design science research and user-sensitive information systems development methodology.

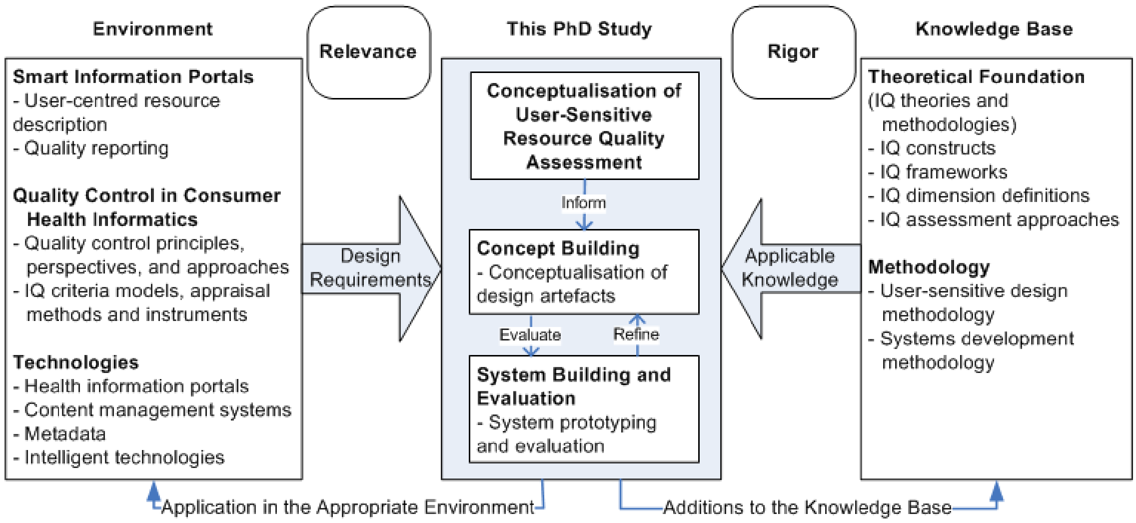


Figure 3.6 Adaptation of Hevner et al.'s (2004) framework for design science research

In this diagram, the environment partition reflects the components of the phenomena being investigated. A combination of health information portals, quality control issues and technologies provides a rich context for design requirements elicitation, which connects and informs the conceptualisation of use-sensitive RQ assessment, as well as the conceptualisation, construction, and evaluation of corresponding solutions. As opposed to the artefact-centric approach, which only concerns the conceptualisation of design artefacts, the adapted framework encompassed two kinds of conceptualisation. One was the theory-building relating to the concept of RQ assessment that incorporated user needs and values. This informed the development of a user-sensitive RQ assessment framework. The other one was the concept-building relating to the design and development of an intelligent quality tool and its prototype.

Further, there are also theories and methodologies, serving as a solid knowledge base for both problem solving and theory building.

Figure 3.6 illustrates how this research brings new knowledge to the information quality theories and the design science research approach, and how the research advances the quality control practice in the field of consumer health informatics. The adaptation of Hevner et al.'s (2004) seven design science research guidelines to this research, is summarised in the final chapter of this thesis (See Section 8.5), which demonstrates how this research achieved the goals through a socio-technical design science research approach.

3.6 Chapter Summary

This chapter described the research design for this multi-disciplinary and applied socio-technical design science research. The chapter introduced the design science paradigm in IS research, and justified its relevance to addressing the central question concerned by this research. It explored how a socio-technical design science research approach was integrated with mixed methodologies. Three research phases of a user-sensitive systems development process were elaborated, namely concept building, system building, and system evaluation. A number of key research activities, related data collection methods and instruments were described. The chapter then explained the relationships between this PhD study and the SIP project. Finally, how this research adapted Hevner et al.'s (2004) design science research framework was discussed.

The following chapters of this thesis are organised according to the three conducted research phases. In light of the findings drawn from Chapters 2, Chapters 4 and 5 elaborate the conceptualisation process of a user-sensitive RQ assessment solution, which encompasses a user-sensitive RQ assessment framework and an intelligent quality tool. Chapter 6 describes the system prototyping procedures and outcomes. Chapter 7 presents the evaluation results of the prototype system. Finally, Chapter 8 synthesises the research outcomes and draws final conclusions.

C h a p t e r   4

4 Exploratory Case Study

In this chapter, domain expert activities and needs for resource quality assessment are analysed, in order to elicit design requirements for standardising and facilitating user-sensitive resource quality assessment. The analysis is based on data collected from a case study that was jointly conducted with the Smart Information Portals project. The chapter begins with an introduction of the two investigated metadata-driven and user-sensitive health information portals. Quality appraisal tasks and domain expert activities are analysed next, in order to articulate what constitutes resource quality assessment processes. The needs of domain experts for intelligent support are identified via the analysis of problems they encounter with resource quality assessment. The chapter then introduces the research outcomes of this study, and proposes a semi-automated quality assessment solution to address the identified domain expert needs. Finally, the chapter summarises the design requirements and implications for developing the semi-automated resource quality assessment approach.

4.1 Introduction

This case study is aimed at understanding resource quality (RQ) assessment practices, which adopt the user-sensitive design philosophy to achieve quality-assured portal content management. It is necessary to understand the quality appraisal tasks and domain expert activities involved, in assessing RQ and managing portal content. Generally, there is a lack of access to content management documentation or content management systems (CMS) that are currently employed by health information portals. The available information is limited to the documentation, relating to collection development, published by a few portals from their home sites. For instance, Intute gateway (www.intute.ac.uk), an online subject service provider for academic users, provides a set of guidelines to explain and delineate the requirements or policies for collection development, evaluation and cataloguing. These documents define the rules for action. They are insufficient for researchers to obtain an intimate knowledge of domain expert quality appraisal activities and the problems they encounter with RQ assessment. In this chapter, RQ assessment processes and domain expert needs are analysed via a qualitative and reflective case study.

4.1.1 The Joint Domain Expertise Study

As introduced in the previous chapter (see Section 3.3.1), this research is related to the Smart Information Portals (SIP) project. The case study was conducted jointly with the SIP project team. Two metadata-driven and user-sensitive health information systems were examined in-depth. Data collected through the case study was analysed with a shared interest to articulate the multi-dimensional domain expertise required in building and sustaining smart, user-sensitive health information portals (Evans et al. 2009). In addition to the domain expertise analysis, this PhD study further analysed the interview data independently with the focus drawn to user-sensitive RQ assessment issues. Moreover, additional data was collected to complement the interview data, such as analysis of portal development documentation, review of actual portal systems, and examination of existing portal data (see Data collection methods on page 73).

The case study conducted a series of semi-structured interviews with three domain experts, who were working for the BCKOnline and HHOnline portals. Working profiles of the interviewees are described in Table 4.1 below.

Table 4.1 Profiles of domain expert interviewees

Expert interviewee code	Was working for	Experience in the role of domain expert	No. of resources described	Level of domain expertise
E1	BCKOnline portal	4 years	1500+	Expert in information management (IM) and the breast cancer disease domain
E2	HHOnline portal	3-6 months	150+	Novice in IM but expert in heart disease domain
E3	HHOnline portal	6-12 months	400+	Novice in IM but expert in heart disease domain

The semi-structured interviews were conducted with these domain expert participants individually. As introduced in Chapter 3, the interview questions covered a number of topics, reflecting on the role domain expertise plays in portal design, development, and management. This included the description of resource identification, selection and description processes, the decision-making processes, and the technical infrastructure (see **Appendix G**). The cognitive walkthrough technique was also utilised in these interviews, which stepped through the resource identification, selection, and description processes with individual interviewees. The focus was on the domain knowledge and skills applied in making value judgements. In addition, metadata schemas of the two portals and corresponding portal development documentation were analysed to complement the analysis drawn from the interview data.

In this case study, I identified the relevant questions to be included in the interview instrument, observed domain expert behaviours in portal content management, conducted semi-structured interviews, and transcribed all interview discussions. This research analysed the interview data with a view to identify quality appraisal tasks, activities, problems, and needs, in order to articulate user-sensitive RQ assessment issues. The metadata schemas of the two portals were also analysed from the perspective of how they handle and represent the contingency of the user context. The analysis focused on user-sensitive resource description elements, such as *Subject*, *Audience* (user profile), and *Quality*. I also examined the portal resource assessment criteria, existing portal data and portal content management systems, from a user-sensitive design perspective.

4.1.2 The BCKOnline and HHOnline Portals

This section introduces the two metadata-driven and user-sensitive health information portals that were investigated in this case study.

The Breast Cancer Knowledge Online (BCKOnline) portal (www.bckonline.monash.edu.au) was initially launched in 2004 as a research prototype, of an Australian Research Council funded project (BCKOnline 2009). Over a number of years the BCKOnline portal has evolved to a web portal that serves a community of women with breast cancer, their family, friends and caregivers. In the last two years, the portal has attracted more than 70,000 visits (see Figure 4.1).

Figure 4.1 Homepage of the BCKOnline portal

The Heart Health Online (HHOnline) portal (www.sip.infotech.monash.edu.au/heart-portal/) is a research prototype system, funded by the Australian national depression initiative 'beyondblue' (www.beyondblue.org.au). The homepage of the HHOnline portal is presented in Figure 4.2. The portal was built to assist General Practitioners to manage depression and related illnesses, in people with coronary heart disease (HHOnline 2007).

Figure 4.2 Homepage of the HHOnline portal

Both the BCKOnline and HHOnline portals are of unique value to their user communities, as they deliver relevant, reliable and timely information through flexible and customised search strategies. They both provide three search options, allowing users to decide how they want to look for information specific to their situations. Those options include a simple search, a search on a list of topics that cover disease trajectory and treatment options, and a personalised search that allows user profiling. For instance, the portal matches a user's profile with the information resources specifically selected and described in the portal, and retrieves information relevant to that person's profile.

For both the BCKOnline and HHOnline portals, domain experts and end-users (health information consumers) were identified as two types of portal users, as depicted in Figure 4.3. The portal user interface consists of two sub-systems: a domain expert interface and an end-user interface. These two sub-systems both dynamically interact with a centralised metadata repository.

Figure 4.3 Two types of portal users

The two portals respect a portal end-user as the final judge of RQ, in terms of what is relevant to the user's information need and what is reliable, based on the user's value system. For this reason, in the two portals, online resources are categorised, evaluated, selected, and described in a way to reflect perceived user needs and values. User-sensitive resource assessment criteria and description schemes were developed to provide value-added information to portal end-users, in order to empower them to make their own quality judgements (McKemmish et al. 2009).

4.2 Analysis of Resource Quality Assessment Processes

According to Addey et al. (2002), content management processes can be simply described as passing "a document through several discrete stages of editing and approval on its way from creation to publication" (p. 138). Informed by this definition and based on the interview discussions, quality-assured content management tasks associated with the BCKOnline and HHOnline portal practices, can be categorised into the following two series of steps:

- new content development steps: identify, evaluate, select, and describe new resources ;
- existing content maintenance steps: approve, publish, and review portal included resources.

In this section, the analysis is focused on the steps for developing new portal content. The purpose is to identify the parts, which require value judgements by domain experts, as well as those, which may be automated or augmented by intelligent technologies.

For the BCKOnline and HHOnline portals, resources are assessed for inclusion via a series of review processes undertaken by domain experts. First, candidate online resources are identified based on domain experts' prior knowledge of valuable websites or sources resulting from their own searching. Second, the candidate resources are categorised according to the portal's resource selection guidelines. Based on the type of the resources, different quality assessment criteria are used to evaluate their perceived usefulness to portal users. Then, facilitated by the portal CMS, the selected resources are fully described according to the portal's resource description scheme. Corresponding metadata records are created for the selected resources, and stored in the portal's metadata repository. These metadata descriptions capture the outcomes of value judgements made by domain experts, revealing the likely relevance and perceived quality of resources to their intended audience. Finally, complete and approved resource metadata records are indexed by the portal's search engine, for later retrieval. Figure 4.4 below illustrates these processes.

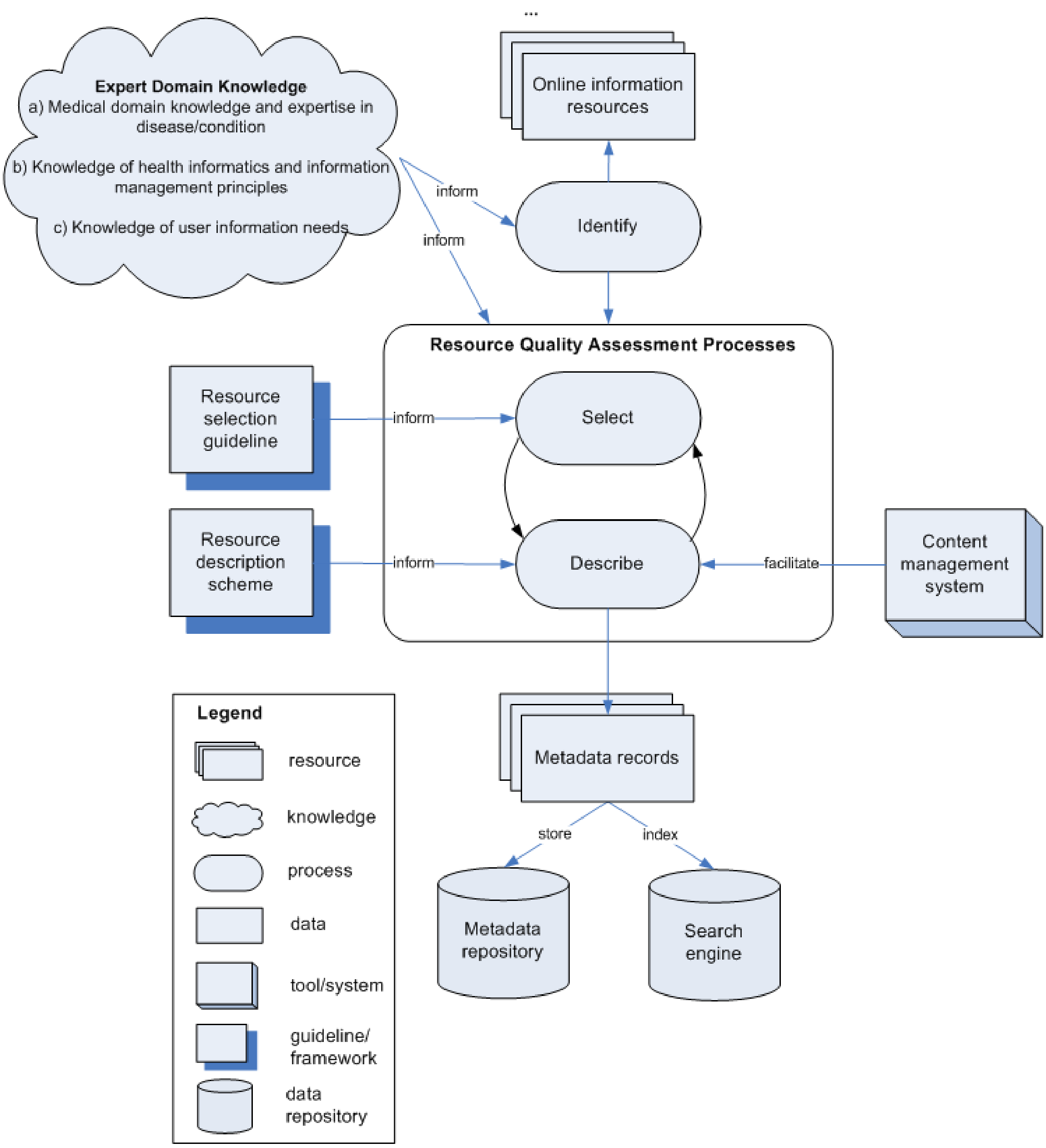


Figure 4.4 Existing RQ assessment processes

As depicted in Figure 4.4, decisions made by domain experts mainly determine whether to include candidate resources in a portal, and how to describe their attributes for later retrieval. These decisions are made based on the candidate resources' compliance to the portal's resource assessment and selection criteria, and the domain experts' understanding of portal targeted users and their information needs. However the involved value judgements are made during the whole processes, and not in a linear order. Candidate resources are categorised, characterised, and evaluated during the selection process. The resource metadata records of selected resources capture the value judgements made in these intermediate steps. Domain experts assess the selected resources continually, until they are fully described according to the portal's metadata schema. In this sense, resources selected for description are not fully assessed. Therefore, the selected resources can still be rejected during the description process.

In the following sub-sections, the resource selection and description processes, employed in developing quality-assured and user-sensitive portal content, are further discussed. The BCKOnline portal is used as an example to illustrate how the user perspective is incorporated in these processes. Content management systems of the two portals are compared. The comparison reveals the deficiency of the systems in supporting the decision-making processes.

4.2.1 User-Sensitive Resource Selection

In the resource selection process, resource assessment and selection criteria are adapted to the type of candidate resources. In the case of the BCKOnline portal, all identified online resources are classified into three categories, namely *Medical*, *Supportive* and *Personal*, to deal with the diverse needs of the breast cancer population (Williamson and Manaszewicz 2003). The ways resources are classified are described below in Table 4.2.

Table 4.2 BCKOnline category scheme (McKemmish et al. 2004, p.23)

Category	Scope	Example
Medical	Resources related to the treatment and management of disease, such as: treatment options; clinical trial reports; drug news.	Scientific papers; Clinical Practice Guidelines; Clinical trial information and/or results; Drug reports; Treatment reviews; Research media releases.
Supportive	Psychosocial outcomes which include: the effect of the disease on the woman and her family; aspects of social and psychological functioning, either on a temporary basis or as the long-term consequences of the disease; facilitative information which may include addresses of support groups, government assistance information; local council contacts for home help.	Information regarding financial, emotional support, practical assistance; Information on various therapies designed to ameliorate the physical and psychological impact of disease and treatment; Case studies.
Personal	These may overlap with the psychosocial, but may be different and generally embody the value systems, in the life of the individual. These outcomes are often based on reflection of the experience of disease, including its treatment rigours, and incorporate the stories and personal reflections of other individuals. The 'voice' of the health professional would also be included in this category.	Women's/families' and professional reflection or appraisal of the individual's experience of breast cancer.

The BCKOnline portal developed the resource quality assessment criteria based on existing quality control codes, principles, and guidelines for online health information (Anderson et al. 2003). Within each of the *Medical*, *Supportive*, and *Personal* categories, different quality criteria are used to assess the quality of resources, for their inclusion in the portal. This is presented in Table 4.3 below.

Table 4.3 BCKOnline resource quality assessment criteria, adapted from Burstein et al. (2005, p. 6)

Category	Mandatory quality criteria	Recommended quality criteria
Medical	Authorship; Currency; Evidence-based	Review process; Referenced (citations)
Supportive	Authorship; Currency; Purpose	Referenced (citations)
Personal	Authorship; Exemplifying range of views	Relevance to Australian audience

These mandatory or recommended quality criteria are defined as the following:

- **Authorship**: This includes both the creator and publisher of the material, and should clearly establish the 'authority' of the author. Generally taken to mean the need for credibility.
- **Currency**: Dates of creation, posting and amendment to resources are provided. May also apply to the source of information, if on another website.
- **Evidence-based**: Adherence to recommended strategies as published in a variety of Australian National Health Medical Research Council (NHMRC) Clinical Practice Guidelines for breast cancer.
- **Review-process**: Editorial policy of a website is provided, indicating whether the resource has gone through a peer review or editorial review process. List of names and expertise may be included.
- **Referenced**: Citations and/or links to cited evidence may be provided.
- **Purpose**: Overall objective of site and material is evident – i.e. educational, promotional, commercial.
- **Relevance**: Material which may have specific cultural, social, geographic or ethnic applicability to the Australian context.
- **Exemplifying range of views**: Resources encompass the broad spectrum of opinion, on a particular issue.

The assessment and selection of online resources also consider the need to include information that is of potential relevance to the Australian context, of interesting topics highly demanded by portal users, and presented in alternative formats. In resource quality assessment processes, the following values are also recognised:

- the importance of personal meaning and individual values of the potential user;
- the 'value' of diversity, in the sense of not privileging information according to source or type or content – these 'decisions' are to be based on user preference and choice; and
- the recognition of the often ambivalent nature of the information seeking process for the individual with a serious illness.

In summary, the design of the resource categories, and the resource assessment and selection criteria, combines user requirements with resource evaluation and selection practices. To some extent, the criteria regulate the decision-making of domain experts from the perspective of portal users.

4.2.2 User-Sensitive Resource Description

Instead of preserving and indexing the full-text content of external resources, both the BCKOnline and the HHOnline portals generate and index metadata descriptions of selected resources, and preserve those records in metadata repositories.

In the case of the BCKOnline portal, the BCKOnline metadata schema was specifically designed to build a full descriptive record of selected external resources. Domain experts, who had the user-sensitive perspective, were involved in the development of the schema. As a result, the schema adapted and extended the AGLS metadata standard (formerly known as the Australian Government Locator Service and the AGLS Metadata Element Set) to provide user-sensitive resource description and cataloguing. The AGLS metadata Standard extended the Dublin Core set of metadata elements by introducing new elements such as *Audience*. The BCKOnline metadata schema adopted sixteen AGLS metadata elements and added a new *Quality* element (McKemmish et al. 2004). For all these seventeen elements, qualifiers and encoding schemes were defined accordingly. The HHOnline portal employed the BCKOnline metadata schema for resource description, but with some minor modifications to suit the specific needs of the heart disease community. A summary of the BCKOnline metadata elements, and the qualifiers and encodings adapted for the HHOnline portal, are presented in **Appendix C.1**.

In the BCKOnline metadata schema, resource category information is captured by the *Type* element to indicate the medical, supportive or personal nature of the resource. Moreover, the schema defines user-sensitive resource descriptors, such as *Audience* and *Quality* metadata elements. The *Audience* element is adapted from the AGLS metadata schema, but uses new qualifiers to describe the target audience. The element encodes user-profiling information, including age group and disease stage (specific to the BCKOnline portal), disease status (specific to the HHOnline portal), audience type, information preference and locality. The innovative *Quality* element and its corresponding qualifiers describe information about the quality aspect of a resource. The element addresses the user community's information quality concerns on reliability and trustworthiness.

These special metadata elements enable user-sensitive information retrieval and quality reporting in search results, where the portal users' individual quality perceptions and decision-making capacities are fully respected. Facilitated by the *Audience* and the *Type* elements, both portals enable users to build profiles in the search interface, to retrieve information of more relevance. Meanwhile, information about reliability, provenance, authority, and timeliness of resources is captured by the *Quality* element, based on which narrative quality reports are generated for retrieved resources. By these means, the two portals disclose the explicit quality criteria used in resource assessment and selection, to address user concerns about the quality of information retrieved from the portal. This kind of value-added information empowers users to judge the quality or the fitness of the resources, based on their individual information needs and value systems (McKemmish et al. 2009).

4.2.3 Portal Content Management System

As discussed in Chapter 2, a CMS basically does three things (Addey et al. 2002, p. 12):

- asset management: organising units of content;
- transformation: presenting the content; and
- publishing: delivering the content to audience.

In the context of the BCKOnline and HHOnline portals, the asset management functionality of portal CMS is of most relevance to resource selection and description processes. A number of functional modules are involved to provide the functionality. The modules include metadata schema, authoring interface, site structure, resource management, link management, user management and glossary management. This research inspected and compared these aspects of the two portals' content management systems. The results are summarised in Table 4.4.

Table 4.4 Comparison of portal CMS

Functional module	BCKOnline portal	HHOnline portal
Metadata schema	The BCKOnline metadata schema	Adapted from the BCKOnline metadata schema
Authoring interface	Metadata fields are distributed in multiple web pages	Metadata fields are organised in one flat webpage
Site structure	Side menu	Top menu
Resource management	Add/Edit/Delete; Searching is only available on title; Sorting is not available	Add/Edit/Delete; Searching is not available; Allow multi-dimensional sorting
Link management	Not available	Not available
User management	Add/Edit/Delete; Accessible from the main menu;	Add/Edit/Delete; Accessible from the main menu
Glossary management	Add/Edit/Delete; Not linked to the authoring interface	Add/Edit/Delete; Linked to the authoring interface

As was discussed in the previous sections, resource selection and description processes involve significant value judgements on the user-sensitive attributes of a resource, which are captured by the BCKOnline metadata elements, such as *Audience* and *Quality*. By examining the two portals' functional modules, it is concluded that neither of the two content management systems provided assistance to the decision-making processes on RQ.

4.2.4 Resource Quality Assessment Processes

The above analysis of portal content development processes indicates that RQ assessment mainly involves two interconnected and intertwined processes: resource selection and resource description. When selecting a resource, a domain expert may already have a mental idea of the description and how it may meet user profiles and quality criteria. When describing a resource, particularly when completing the quality report, the domain expert is verifying their selection decision.

Decisions such as whether or not to accept a candidate resource are made based on multi-dimensional domain expertise that may require the involvement of more than one domain expert. The portal content management systems are designed mainly for authoring metadata records, as specified in descriptive metadata schemas. They capture decision-making outcomes and enable the creation, storage, indexing, and publishing of resource metadata records. However, the systems do not provide functionality to support or facilitate decision-making processes on RQ.

Findings from the process and CMS analysis raise questions, such as the following. Which parts of the quality assessment processes can be automated or augmented with intelligent tools, in order to improve the scalability and ultimately the sustainability of these processes? Where is human intervention required? What are the requirements of an intelligent quality assessment solution, to support the capacity of domain experts, and to facilitate the collaborations between domain experts and the system? Answers to these questions are explored in the next section, via an analysis of domain expert activities.

4.3 Analysis of Domain Expert Activities and Needs

The analysis of domain expert activities and needs is based on the data collected from the semi-structured interviews, conducted in the joint case study. In this section, the collected interview data are analysed for the purpose of identifying quality appraisal activities of domain experts and the techniques applied in performing resource identification, selection and description tasks. The analysis leads to an in-depth understanding of RQ assessment problems and corresponding domain expert needs for intelligent support.

4.3.1 Domain Expert Activities for Quality Assessment

Domain expert activities, in relation to RQ assessment, are synthesised from an analysis of the interview transcripts. Table 4.5 outlines the quality appraisal activities and techniques employed in the resource identification, selection and description processes.

Table 4.5 Summary of quality appraisal activities

Process	**Quality appraisal activities of domain expert interviewees** (Code of domain expert interviewee, who explicitly mentioned the activity)
Resource Identification	Start with search engines and mining their websites for generic or relevant information (E2, E3)
	Routinely visit key medical information channels and own bookmarks, e.g. PubMed database, official site of Food and Drug Administration (E1, E2, E3)
	Search in patient discussion forums and bulletin boards for topics that are relevant to a patient community (E1, E3)
	Use personal communication to deal with individual queries (E1)
	Search in subscribed medical journals, conference proceedings or RSS feeds (E1, E2)
	Use the search facility provided by a website to find topics and relevant resources (E1, E2, E3)
	Map identified resources to user information needs (E1, E3)
	Use personalised search in the portal's front-end to find information gaps (E2)
Resource Selection	Select for some particular categories (E1, E2, E3)
	Refer to disease trajectory (E1)
	Refer to a self devised map, mapping identified resources to a topic list (E3)
	Consider query terms from a user perspective (E2)

	Deal with repetition (E1, E2, E3)
	Determine the level of granularity at which a resource is assessed for inclusion, e.g. linking to a website or a specific webpage within the site (E1, E2, E3)
Resource Description	Use a web browser's built-in search facility to find metadata information (E1, E2, E3)
	Check other pages of the website, e.g. Disclaimers, Copyrights, and About us, for finding information on review process (E1, E2, E3)
	Inspect the domain name of a URL for credential information (E3)
	Contact information provider when certain information cannot be found from the site (E2)
	Refer to the list of all glossary terms when assigning subject terms (E1, E2, E3)
	Copy and paste terms and pay special attention to capitalisation and spelling (E1, E3)
	Use personalised search to validate or revise own coding for resource description (E2)
	Look for what has been assigned in the past, e.g. check previous credentials for creator, contributor and publisher for consistency (E3)

As illustrated by Table 4.5, resource content quality, its intended audience, and topical relevance to the disease trajectory defined by a portal, are major concerns of domain experts when they assess the fitness or usefulness of an information resource for its inclusion in a portal. Implicitly, these value judgements are informed by the information analysed from the resource's original source, e.g. the collected quality evidence and implicit relationships across the analysed resource attributes. Decisions are also informed by previous value judgements made on similar resources.

Lay expertise is the key to user-sensitive, especially in the BCKOnline context. The role of domain expert as patient advocate is reflected in resource identification, selection and description processes. Domain experts are required to understand user information needs, the types and formats of information user required, and the information gaps. As required by user-sensitive resource description, domain experts make extensive use of profiled user information needs. Resources, in this sense, are selected and described to cover a wide range of information needs defined by user profiles.

Moreover, medical expertise is required in the domain expert role for the assessment of clinical trials reports, evidence-based medicine, as well as the awareness of controversial or unproven medical treatments. In addition, domain experts continually monitor a variety of channels to identify resources of interest. They also bring to the processes an awareness of the various sources of resources, and their strengths, weaknesses and limitations.

Besides, information management skills are also required in the role. Apart from the management of metadata records, domain experts need to be aware of the spectrum of disease trajectory that needs to be covered. As mentioned in Table 4.5, domain experts create topic maps for existing resources to identify the topic gap in the repository. They also check the resource coverage by using the personalised search function provided in the end-user interface.

In summary, the RQ assessment processes encompass not only basic content management activities, such as adding, editing, deleting, indexing and publishing metadata records, but also categorising and charactering resources, and profiling their target audience. Decisions made via these activities are informed by multi-dimensional domain expertise. Decisions such as whether or not to include a new resource or to publish a fully described resource are made based on the combination of knowledge, skills, and experience of individual domain experts. These decisions are assisted by a portal's resource assessment criteria and resource description metadata schema. Due to the different expertise levels, domain experts have, collaborations among multiple domain experts are sometimes required to make value judgements regarding the same resource.

4.3.2 Domain Expert Needs for Intelligent Support

In this section, the kinds of intelligent supports required by domain experts are analysed in order to inform the design of a tool-assisted RQ assessment solution. In Table 4.6, problems associated with RQ assessment, domain expert needs, and corresponding solutions are analysed.

Table 4.6 Analysis of RQ assessment problems and associated domain expert needs

RQ assessment problems	Needs	Solutions
Problem of articulating and implementing resource selection guidelines	Awareness of relationships between user characteristics and resource attributes, for achieving user-centred resource evaluation and selection	Devise an operational and user-sensitive RQ assessment framework
Selecting and describing repetitive or duplicated resources	Awareness of the data already in the portal when describing new resources. Requires much more specific searching when the number of described resources grows	Provide flexible internal search facility, URL check, resource content similarity check
Slowed down by repetition. Have difficulties in finding topics or audience requirements	Awareness of gaps between user information needs and the available information provided by the portal	Analyse server log, track user behaviours, and provide feedback channels in end-user interface
In some cases, have difficulties in finding information such as author, publisher etc., which are important to assess RQ	Getting access to information beyond the cognitive ability of domain experts	Provide automated metadata generation tools that harvest any existing metadata in the source; could also populate fields like identifier, title etc. to create a stub record, or reuse previously assigned metadata
Determining the currency of resources is problematic as new resources might alter the	Perpetual monitoring of published resources at any stage when new research findings are released	Review or re-assess the currency of resources over the space of a year

currency values for other resources		
Lack of support for consistency, metadata quality and efficiency	Improving consistency of resource descriptions, e.g. classification of authorship, even for solo domain expert; supporting collaboration and learning amongst multiple domain experts; capacity to service implicit practices and dynamic decision-making; capacity to allow for domain experts to interact with one another; ability to see examples from the data repository to help in the description process and decision-making	Reverse-engineering as part of the learning process – assess how another domain expert described a resource

As illustrated by Table 4.6, once a cache of candidate resources is identified, problems emerge relating to the articulation of selection criteria. This then leads to the problems of quality assessment, the description of quality attributes, and the consistency of value judgements.

Firstly, there is a need to exploit a systematic approach to associate resource attributes, required by the resource description scheme, to the decision-making procedures of resource assessment and selection. The comprehensive BCKOnline metadata schema is regarded as the conceptual representation of user-sensitive resource description. The process of metadata value generation is also a decision-making process on resource quality. Two types of information are gathered in this process: the attributes of the resource that indicate its intrinsic quality, and the contextual information in which the value judgements are made. Therefore, it is necessary to develop a quality framework to aid in resource evaluation, selection and description. There is also a need to design and prototype aspects of the domain expert interface, to build and monitor domain expertise.

Secondly, domain experts attempt to select resources that healthcare consumers need. For this reason, initial and continuing feedback on user information needs is highly demanded. There must be a mechanism to identify the gaps between included resources and user information needs.

Thirdly, there is a need for the exploitation of relationships among metadata descriptions of resources. For example, the authorship elements such as creator, publisher, and contributors, have relevance to the credential aspect of resource quality. This kind of relationship can ensure consistent quality reporting. Mining the metadata repository and providing feedback to the domain expert can also achieve better consistency and quality of resource identification, selection and description.

Finally, due to the dynamics of online information, resource descriptions need to be updated on a regular basis. For instance, corresponding resource metadata fields need to be constantly monitored to reflect changing URLs, updates to clinical trials data, and impacts on the currency of information.

However, the proposed solutions may still not be able to address some of the problems, such as the creation of abstracts in lay language, based on the anticipated portal user reaction to the information. The following quotation describes such a concern of the domain expert interviewee:

> "It would have been very, very easy to simply rely on MESH subject headings, but the 'end product' would have been unusable. Hence, since I knew the language that women used, these often found their way into BCKOnline – not merely the colloquial, but the abbreviations, etc." – E1

It is recognised that the ultimate challenge is to keep the user needs and values continually in mind, and try to incorporate these elements throughout the entire process. Based on the above analysis of quality appraisal activities and problems, domain expert needs are summarised below:

- gain an overview knowledge of the existing portal data, e.g. resource distribution;
- avoid duplication of effort;
- make consistent value judgements;
- collaborate with other domain experts;
- learn from decisions made by other domain experts;
- check the status of portal content on a regular basis (remove dead links, update URLs, replace existing resources with better ones on the same topic);
- monitor user behaviours to learn of changes in information needs.

Two types of intelligent support are required to address the domain expert needs: firstly, to support the making of informed decisions; and secondly, to support collaboration and learning. These needs are further discussed next.

Needs for making informed decisions

In the previous section, multi-faceted domain expertise required in RQ assessment processes was discussed. In addition to these expert domain knowledge and skills, the analysis of domain expert activities and needs indicates that knowledge of existing portal data helps domain experts make informed decisions on RQ. A number of IM skills are required to build and maintain the knowledge on existing portal data, as listed in Table 4.7.

Table 4.7 Required knowledge and skills for making informed decisions

Required knowledge on existing portal content	Required information management skills
▪ Knowledge about existing metadata values in a portal's metadata repository, in order to reuse metadata, and improve value consistency; ▪ Knowledge about portal included resources, e.g. topic coverage, resource distribution, most credible sources/websites etc., in order to avoid redundancy and duplication of effort; ▪ Knowledge about topic priorities in order to guide resource identification and selection.	▪ Management of metadata records and glossary terms; ▪ Implementation of resource assessment and selection criteria and metadata standards; ▪ Information searching, navigation, and categorisation.

Needs for collaboration and learning

The analysis reveals another domain expert need for collaboration and learning. Collaboration amongst multiple domain experts is of particular importance for sustaining quality-assured portal content management. The domain expert interviewees highlighted problems associated with having multiple domain experts involved in the development and maintenance of portal content. On one hand, an individual domain expert may not necessarily possess all facets of domain expertise. In some cases when domain experts find themselves lacking expertise or confidence in certain areas, they may want to leave that part of the work to others, or wish to learn from past experience. On the other hand, all the interviewees wanted to produce work consistently and collectively with the others. For instance, they would check how a resource from the same source was previously categorised and described.

The analysis also found that user quality experience was disconnected from RQ assessment processes. The profiling of user information needs was done before the design and development of a portal's resource assessment criteria and resource description scheme. The investigated portals did not provide means or channels to monitor changes of user information needs over a period of time. Thus, domain experts were not able to adapt the selected information to continuously reflect the dynamics of user information needs.

For the above reasons, providing a collaborative and interactive learning environment is necessary, in order to foster the development of user (patient community) advocacy, medical expertise of diseases and conditions, as well as IM skills. For instance, information about previous quality assessment decisions on similar resources needs to be automatically generated and provided by the portal CMS system. By this means, a domain expert's medical knowledge about disease conditions and IM skills can be quickly improved.

4.4 Analysis of Portal Included Resources and Their Quality Descriptions

Analysis, of RQ assessment processes and domain expert needs, indicates that domain experts experienced the most difficulty in describing resource target audience and the quality attributes defined by the portals' metadata schemas. As stated in Chapter 2, previous quality indicator studies in the literature indicate that a considerable amount of information can be retained at the website level, e.g. the ownership and sponsorship of the information provider. If a portal contains a number of resources from the same site, it is reasonable to assume that certain parts of their metadata descriptions share common values. In order to identify the sources of resources included in the BCKOnline and HHOnline portals, published resources of these two portals were analysed. The analysis results are presented in Table 4.8.

Table 4.8 Overview of resource distribution

Measure	BCKOnline Portal	HHOnline Portal
No. of published resources in total	871	441
No. of resources from the top 10 websites	356	224
Percentage of resources from the top 10 websites	40.87%	50.79%
No. of source websites in total	223	123
Top 10 websites locality	2/10 are Australian	8/10 are Australian
Top 10 websites ownership	9/10 are Organisational	10/10 are Organisational

The results indicate that although the source websites of the two portals showed a good variety (BCKOnline: 223; HHOnline: 123), a number of major websites were highly favoured by domain experts. In the case of BCKOnline, more than 40% of portal included resources were from 10 key source websites. In the case of HHOnline, the share of resources from key websites was even higher, to more than 50%. Amongst those 10 key sources of each portal, organisational portals were dominant. According to the locality of those key sources, the HHOnline portal was more localised than the BCKOnline portal. Detailed information about this resource distribution analysis is provided in **Appendices D.1 and D.2**.

Since a majority of resources included in a portal can from the same website, it is necessary to investigate where to find relevant information at the website level, to inform RQ assessment. For this reason, a sample of 25 healthcare websites was examined. The sample consisted of:

- the top 10 websites, which provided more than 40% resources to the BCKOnline portal;
- the top 10 websites, which provided more than 50% resources to the HHOnline portal; and
- five other popular healthcare information portals or search engines.

The ownership of these websites ranges from organisation, institution, journal, commercial body and private group. Each website was checked for its information provision in the following categories: 'Copyright', 'Disclaimer', 'Privacy', 'About us', 'Contact us', 'Linking policy', and 'Quality seal'. Detailed results can be found in **Appendix D.3**.

It was found that all those websites provided 'About us' and 'Copyright' information, as well as the 'Contact us' channel. The 'Disclaimer' statement and 'Privacy' policy were missing in a few sites. In contrast, not many websites described their 'Linking policy', and only nine websites, either institutional or commercial, were certified for their compliance with HONcode's (2007) eight quality principles. Among those nine websites, five Australian websites were also approved by HealthInsite, an Australian government initiative, which itself has been accredited by HONcode since February 2002.

Based on the above preliminary analysis results, it is concluded that values of quality indicators, in particular those associated with the credibility of online resources, can be detected from the standard information sections provided at the website level. The remaining problem is whether information provision patterns can be found to facilitate the indicator quality value generation, and what intelligent technologies are applicable to automate the value detection, extraction and generation processes.

4.5 Findings and Outcomes of the Case Study

4.5.1 Joint Research Outputs

This case study delivered a number of joint research outputs in regard to the investigation of the role domain expertise plays in portal design and development. The research outcomes achieved jointly with the SIP project include:

- articulation of domain expertise attributes and domain expert information seeking behaviours;
- development of an instrument, in a form of two quadrant diagrams, to represent and measure domain expertise; and
- conceptualisation of the domain expert tools, to support the identified multi-faceted domain expertise.

These three joint research outcomes are introduced next respectively.

4.5.1.1 *Attributes of domain expertise*

In SIP, domain experts are defined as people with first-hand experience and extensive knowledge of the medical, supportive and psychosocial information needs of the user community (2006). The role of domain expert is responsible for building up and maintaining a metadata repository through resource identification, selection and description processes (Evans et al. 2009). As a result of the domain expertise study, the multi-faceted domain expertise required in these processes was identified. This is described in Table 4.9.

Table 4.9 Attributes of domain expertise, adapted from Evans et al. (2009, p. 5)

Attribute	Description
Medical and lay experience	The domain expert processes are all based on an intimate knowledge of the specific disease domain experience from both medical and lay perspectives.
User advocacy	The domain expert is an advocate for the portal users. In both cases of the BCKOnline and HHOnline portals, resource identification, selection and description are all based on an intimate knowledge of the disease experience.
Awareness of the kinds and styles of information most valued by the portal's user community	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.
Connection into a range of information networks	The domain expert comes to the portal, already connected into a range of information networks in their area of expertise.
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 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.

In summary, the role of a domain expert requires a good understanding of the specific disease and intimate knowledge of information needs of portal users. They evaluate the perceived usefulness of online information resources to prospective portal users, based on their medical domain knowledge, information management expertise and lay user experience (Evans et al. 2009).

4.5.1.2 *Domain expertise quadrant diagram*

In order to represent the interplay between the attributes of domain expertise as identified above, quadrant diagrams are used to plot domain expertise across multiple dimensions (Evans et al. 2009). One of the proposed diagrams is shown in Figure 4.5, which plots information management expertise against combined medical and lay expertise.

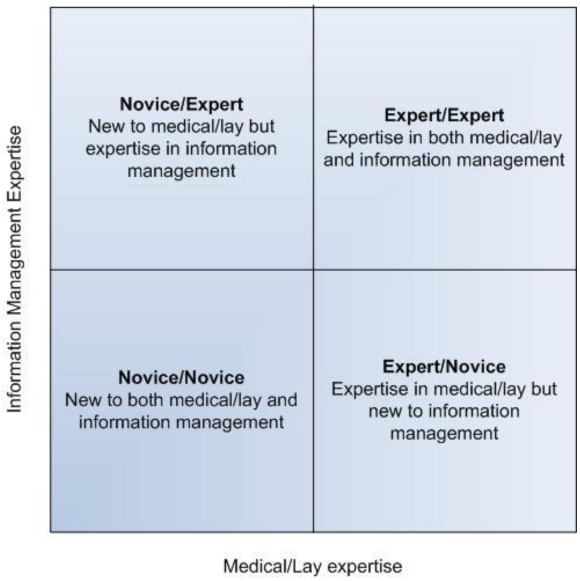


Figure 4.5 Domain expertise quadrant diagram, adapted from Evans et al. (2009, p. 6)

This diagram is particularly useful to describe and measure the areas of domain expertise that individuals are good at or lacking in. The diagram was used to select and profile domain expert participants in the exploratory case study and to select and profile those involved in the system evaluation study conducted by this research.

4.5.1.3 Domain expert tools

In order to support multi-faceted domain expertise required in resource identification, selection and description, the architecture of an integrated domain expert interface was conceptualised. It is suggested that a range of intelligent technologies can be utilised to implement the components of the architecture, with functions supported by the following domain expert tools (Evans et al. 2009; Xie et al. 2008):

- *Quality Framework and Tool* supports domain experts in making decisions on quality issues. It sets the quality selection criteria and also provides intelligent services for performing RQ evaluation and 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.
- *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.

Amongst the above tools, the *Quality Framework and Tool* component needs to be highlighted. It interacts with the other domain expert tools for the purpose of standardising and facilitating RQ assessment processes. The component was proposed and conceptualised by this PhD study independently from the SIP project. The quality framework and corresponding quality tool constitute a semi-automated RQ assessment approach, which is discussed next.

4.5.2 A Semi-Automated Approach for RQ Assessment

The analysis of RQ assessment practices implies that quality assessment of external online resources plays a vital role in defining and delivering 'quality' resources endorsed by a portal. Quality assessment activities are assisted by mechanisms such as resource assessment and selection criteria and resource description scheme. However, having these mechanisms does not necessarily indicate the achievement of quality assurance of portal content. The real challenge lies in how to make the quality assessment activities in compliance with the resource selection criteria, how to assure the consistency and quality of user-centred resource descriptions, and how to make the quality assessment criteria transparent to portal end users.

In order to address these concerns, this research proposes a semi-automated quality assessment solution, which consists of a resource quality assessment framework and an intelligent quality tool that implements the quality framework and semi-automates RQ assessment processes. In this thesis, an intelligent quality tool refers to a computer system that employs intelligent technologies to support or facilitate decision-making processes on the quality of online information resources. As the quality tool is devised mainly for supporting resource quality assessment, it can also be termed as a quality assessment tool.

Figure 4.6 illustrates RQ assessment processes being assisted by an intelligent quality tool, which is integrated into a portal CMS.

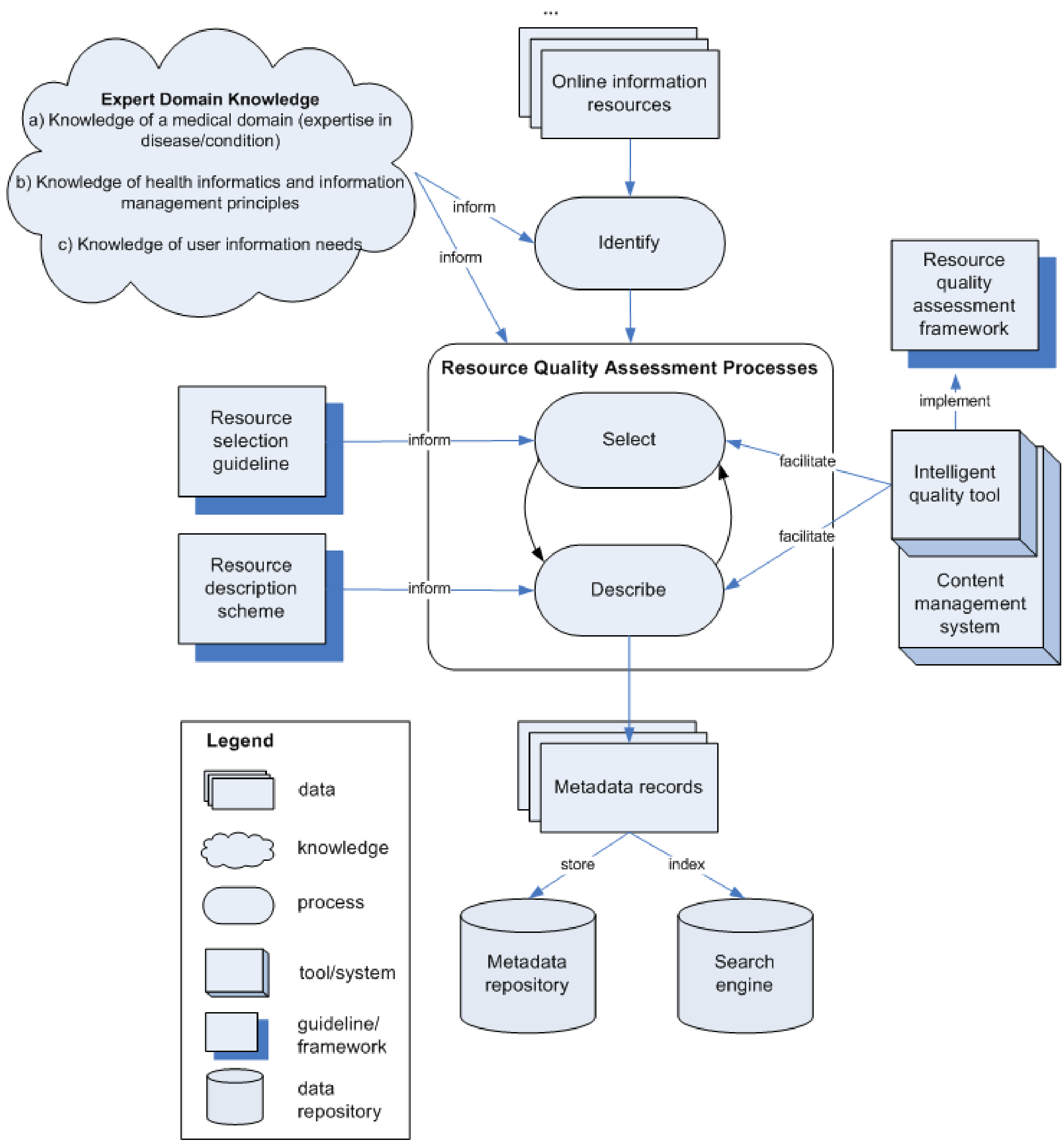


Figure 4.6 Tool-assisted RQ assessment processes

The quality framework and quality tool are proposed to support decision-making processes on RQ in the following two ways:

- to facilitate resource selection (whether to accept candidate resources for their inclusion in a portal's metadata repository) using the portal's resource assessment and selection criteria; and
- to facilitate resource description using the portal's metadata schema and corresponding encoding systems.

Quality-assured content development and maintenance in user-sensitive health information portals involve intensive expert judgements that cannot be replicated by automated tools.

Offering domain experts an intelligent tool, which takes account user needs and values in quality assessment, can be an optimal solution.

4.6 Design Requirements and Implications

The analysis of RQ assessment processes indicates that quality assurance of portal content is achieved at two stages. In the first stage, portal developers set the standards of quality information at the portal design and construction stage. Understanding perceived subject interests and quality perceptions of intended portal users is an essential task to define RQ at this stage. Information, such as what kind of information users find interesting and useful, and what factors affect the use of retrieved information, can be collected by conducting user needs studies or by analysing user search behaviours.

In the second stage, portal content quality is determined and maintained by domain experts on an ongoing basis. Quality control at this level involves significant human effort to assess RQ from a user perspective. This requires domain experts have not only the expert knowledge of a specific disease domain, but also the intimate knowledge of specific needs and preferences of targeted user community (Evans et al. 2009). However in many cases, such multi-dimensional domain expertise is not necessarily possessed and maintained by individual domain experts.

Besides, the analysis of domain expert activities implies that assessing the quality of online resources is a complex task full of challenges. The challenge firstly lies in understanding resource quality from the perspective of portal users. Secondly, it lacks an explicit and operational framework that standardises and facilitates the processes involved in assessing resource quality. Thirdly, a challenge lies in evaluating online health information resources, selecting and describing them to empower portal users to make individual value judgements. Finally, it is difficult to continuously monitor the quality of resources included in a portal in response to the dynamics in online information and user information needs.

Based on these analyses, it is concluded that RQ assessment processes mainly involve two interlinked and intertwined processes of resource selection and resource description. The RQ assessment is fundamentally a subjective issue that reflects the value system of individual domain experts. Decisions, such as whether to include a new resource or publish an existing resource, are made based on the combined knowledge, skills, and experience of individual domain experts. Although assisted by portal resource selection criteria and resource description schemes, domain experts tend to use heuristic techniques and tacit knowledge when making value judgements. It is found that different domain experts have their own decision patterns, search strategies, and quality appraisal techniques, which cannot be replicated by automation tools. This is in line with findings of Greenberg (2004) and Paynter (2005) that human experts seek decision-making support, thus prefer semi-automated approaches. On the other hand, existing portal CMS systems are mainly designed for organising and facilitating data creation and publication (Addey et al. 2002; Browning and Lowndes 2001). They do not provide extra functionality for supporting or facilitating quality assessment processes. It is argued that an intelligent quality tool, which is regulated by an ad-hoc RQ assessment framework, can fill this gap and assess RQ in a systematic, scalable, and sustainable manner.

The quality framework and tools were proposed to work with domain expertise, and support subjective aspects of RQ assessment. The quality tool was to be integrated into a portal CMS, as an enhanced domain expert interface to the portal's metadata repository. It is envisaged that the enhanced portal content management functionality can assist in building and maintaining required domain expertise. Based on the analysis of domain expert activities and needs for RQ assessment, design requirements of a quality framework, and functional requirements of a quality tool are summarised below.

Design requirements for a RQ assessment framework

- assess quality in the user context, i.e. user information needs and quality perceptions;
- define the construct of RQ from the user perspective;
- define operational assessment metrics, in a way to facilitate automation.

Functional requirements for a quality tool:

For developing new portal data:

- collect or create new information from an information resource's original site to support resource characterisation;
- support resource description through automation, e.g. pre-populate metadata values, or provide value suggestions for describing resources;
- help to build domain expertise, support learning, and facilitate collaboration among domain experts;
- provide information about automation tasks and their interrelationships;
- enhance the workflow and the consistency of RQ assessment procedures.

For maintaining existing portal data:

- check resource duplication;
- reduce the workload of maintenance;

- support strategic quality review, identifying problematic resources that require further review or resources that need culling, as they no longer meet the selection criteria.

In summary, both the BCKOnline and the HHOnline portals are constrained by the technological limitations available at the time of portal development. The emergence and maturation of web tools and related intellect technologies can satisfy the intelligent support needs in desired ways. A more intelligent and sophisticated portal CMS is required to provide greater functionality and better interface, to meet the requirements of domain experts. Being user-sensitive is always an important part, to facilitate the assessment, creation, and maintenance of portal content.

4.7 Chapter Summary

This chapter described an exploratory case study of user-sensitive RQ assessment practices. Through a close examination of domain expert activities in the context of two metadata-driven and user-sensitive health information portals, the kinds of decisions domain experts made and their interactions with portal content management systems were identified. It was found that domain experts evaluate the quality of online resources continually, in two connected and entwined processes of resource selection and description. Because of varying domain expertise backgrounds, different search strategies and quality appraisal techniques are utilised to make quality-associated decisions. These kinds of decisions cannot be fully replicated by automated tools. However, parts of the RQ assessment processes can be automated or augmented by incorporating a quality tool in the design of a portal CMS. Regulated by a RQ assessment framework, and tool was envisaged to employ intelligent technologies to assist domain experts in making informed decisions and building required domain expertise. Such a semi-automated approach can assist RQ assessment in a more systematic, scalable, and sustainable manner.

The major outputs of this chapter include the articulation of multi-dimensional domain expertise analysed jointly with SIP (Section 4.5.1), and the RQ assessment approach proposed by this PhD study (Section 4.5.2). Although the analysis of the collected data is shared, articulation of RQ assessment processes (Section 4.2), needs of domain experts for intelligent support (Section 4.3), quality description issues (Section 4.4), and corresponding design requirements of RQ assessment solutions (Section 4.6) are the outcomes of my independent work. The conceptualisation of a semi-automated RQ assessment approach, which includes a user-sensitive RQ assessment framework and an intelligent quality tool, is presented in the next chapter.

Chapter 5

5 Conceptualisation of a User-Sensitive Resource Quality Assessment Approach

This chapter presents the conceptualisation of a user-sensitive and semi-automated resource quality assessment approach for health information portals. The chapter begins with an overview of the conceptual development process, which encompasses the conceptualisation of two components: a user-sensitive resource quality assessment framework, and an intelligent quality tool. Then, it elaborates how a resource quality assessment framework was developed to address the diverse information needs and quality perceptions of portal users. An attribute-based approach was adapted to develop data metrics for assessing the *Reliability* dimension of resource quality. Finally, the architecture of an intelligent quality tool is conceptualised based on the proposed quality framework. Available intelligent technologies are also investigated for their application in providing required functionality.

5.1 Overview of the Conceptual Development Process

As discussed in the previous chapter, in order to support and sustain user-sensitive resource quality (RQ) assessment processes involved in the creation and maintenance of portal content, a semi-automated RQ assessment approach was proposed for health information portals. The new RQ assessment solution encompasses two components, namely:

- a user-sensitive RQ assessment framework; and
- an intelligent quality tool based on the proposed quality framework to support and facilitate decision-making processes on RQ.

The user-sensitive RQ assessment framework was conceptualised, using a method adapted from the approach to developing the Canadian Institute of Health Information (CIHI) four-level conceptual model (Long and Seko 2002). The construct of the proposed resource quality is in line with the information quality (IQ) literature. The design of the quality framework was informed by the context-based IQ assessment framework (Strong et al. 1997; Wang et al. 1995b; Wang and Strong 1996), the user-sensitive design principles (McKemmish et al. 2009), and the design requirements elicited from the case study of available practices (see Chapter 4).

Based on the quality framework and corresponding assessment metrics, the intelligent quality tool was conceptualised to address the domain expert needs identified in the case study (see

Chapter 4). Informed by the functional requirements elicited from the case study, and an examination of applicable intelligent technologies, the conceptual architecture of the quality tool was proposed.

5.2 Developing a User-Sensitive RQ Assessment Framework

In the context of health information portals, the quality of a web-based health information resource, in terms of its fitness to meet the needs and values of portal users, is highly subjective, and is ultimately determined by the individual user (McKemmish et al. 2009). General quality frameworks lack the capacity to address the information needs and quality perceptions of health portal users. Thus, they are not suitable for solving this domain-specific and user-sensitive RQ assessment problem. It is imperative to develop an operational quality framework, which incorporates user information needs and quality perceptions when assessing the quality of online health information resources. The framework needs to address the following four issues.

- The construct of RQ in terms of what quality dimensions are concerned by health information consumers and how to define them from the user perspective.
- The definition of attributes of resources relevant to RQ assessment.
- Characteristics of the context in which information contained in the resource will be used.
- Relationships between resource attributes and user characteristics.

In this section, the application of the user-sensitive design methodology in developing a RQ assessment framework is presented. The following questions are addressed in the framework development process.

- What quality aspects are of concern to health information consumers?
- How to model users and user information needs?
- How the user model can be used to characterise resources, so as to facilitate the user-sensitive quality assessment?

5.2.1 Modelling User Information Needs and Quality Perceptions

As discussed in the previous chapter, the most critical design requirement for developing a RQ framework is to select and define quality dimensions from the user's perspective. According to Strong et al. (1997, p. 104), "quality of data cannot be assessed independent of the people who use data". This view indicates that in a portal context, RQ assessment is highly dependent on the needs and values of portal users. Hence, before defining the resource quality aspects for assessment, it is necessary to first understand the information needs and quality concerns of health portal users.

User information needs and information search process are well described in the field of Information Retrieval (IR) (Cole 2011; Rose and Levinson 2004). In the context of web information retrieval, an online search starts from a user formulating an information request, which is represented by a set of query terms. The user search is followed by a decision-making process, judging on the accuracy, currency, and credibility of search results with reference to the user's information request. As the underlying goal of a user search varies, a user's quality concerns on retrieved information, in terms of what is accurate, current or useful, also varies even with the same query terms (Rose and Levinson 2004). It is the intention of a web search - the actual information needs underlying a user query - that determines the user's judgements in relation to the quality and relevance of retrieved information.

Although queries and behaviours of individual users at online search are highly dynamic and unpredictable, their information needs and quality perceptions in a specific domain can be identified and modelled by conducting comprehensive user needs studies. In the case of the Breast Cancer Knowledge Online (BCKOnline) portal, interviews and focus groups were used to understand user information needs and information seeking behaviours in the breast cancer domain (Williamson and Manaszewicz 2003). The study showed that breast cancer users required the retrieval of information not only relevant to their queries, but also concerned with their individual circumstances, information preferences, and quality perceptions. The resulting construct of user characteristics formed the basis for value-added information provision and quality reporting of breast cancer knowledge resources from a user-sensitive perspective (McKemmish et al. 2009). According to the BCKOnline user needs analysis, the information needs of health information consumers can be modelled as a composition of user circumstance (e.g. profiling a group of users who share common characteristics), information preference, and subject interest.

Moreover, based on the information needs analysis of breast cancer users, Anderson et al. (2003) suggest seven quality criteria as relevant to the quality concerns of health information consumers, namely *Credentials of resource producers*, *Review process*, *Citation of sources (Attribution)*, *Evidence-based*, *Purpose*, *Balance*, and *Currency*. These are the subset of content quality criteria for evaluating online health information, which have been outlined in Chapter 2 (see Table 2.7). Quality dimensions and sub-dimensions (quality characteristics), which best summarise these quality criteria can be defined. It needs to be clarified that this research was only concerned with the content quality criteria of online health information. According to Wang and Strong's (1996) four-category IQ framework, IQ dimensions of the intrinsic and contextual categories were considered relevant to the content quality of online health information resources. Representational and accessibility IQ dimensions thus are not discussed in this thesis.

It needs to be emphasised that the final judge of *Reliability* should be portal end-users. That is, whether a resource is reliable or not to an actual user is to be determined by the user's information needs and value system at the time of search. In order to assist portal end-users to evaluate the *Reliability* of retrieved information, domain experts provide factual information on the resource's *Accuracy*, *Currency* and *Credibility*.

5.2.2 Selecting and Defining RQ Dimensions

As discussed in Chapter 2, IQ is widely recognised as a multi-dimensional and in some cases hierarchical concept (Wang et al. 1995a; Wang and Strong 1996). There is no 'gold' framework that standardises the construct and measure of IQ. However, a number of common IQ dimensions were identified across multiple studies. Those intrinsic and contextual quality dimensions include *Reliability* (or *Reputation*), *Accuracy*, *Objectivity*, *Believability* (or *Credibility*), *Timeliness* (or *Currency*), *Completeness*, and *Relevancy* (or *Relevance*). Detailed discussions of these common IQ dimensions can be found in Chapter 2 (see Section 2.2.2.3).

According to Stvilia et al. (2009), seven IQ dimensions are relevant to healthcare IQ judgements, namely *Accuracy*, *Authority*, *Completeness*, *Currency*, *Objectivity*, *Relevancy*, and *Understandability*. Their relationships to the quality perceptions of health information consumers are examined by a number of studies in the field of consumer health informatics (Adams et al. 2006; Cline and Haynes 2001; Diaz et al. 2002; Eysenbach and Köhler 2002; Marshall and Williams 2006; Williamson and Manaszewicz 2003). Based on these studies, it is concluded that quality concerns of health information consumers are different to those of generic users. For instance, absolute accuracy is hard to measure and varies over time. Consensus might be a better word to describe accuracy of health information. Besides, these quality dimensions are not mutually exclusive to each other. For example, *Authority* is also interpreted as *Credibility* or *Reliability* (Lankes 2008). *Objectivity* is defined with reference to tone (biased or not), *Completeness* and *Comprehensiveness*, and *Objectivity* is also linked to *Accuracy* (Frické et al. 2005).

Based on the analysis of IQ dimensions proposed for the healthcare domain, and their relationships to the seven quality criteria proposed by Anderson et al.'s (2003), five IQ dimensions were selected to construct resource quality, including *Reliability*, *Accuracy*, *Credibility*, *Currency*, and *Relevancy*. Amongst these selected quality dimensions, *Reliability* was considered as the one that best describes the quality concern of health information consumers. Therefore, *Reliability* was used to denote the relation of a resource to the quality perceptions of its user. The concept of *Reliability* is represented as a combination of *Accuracy*, *Credibility*, and *Currency*. Meanwhile, *Relevancy* was regarded as the quality dimension that best denotes the relation of a resource to the information needs of a user. Definitions of the quality dimensions that are used to denote resource quality aspects for the healthcare domain are provided in Table 5.1.

Table 5.1 Resource quality dimensions for the healthcare domain

RQ Dimension	**Definition** derived from (McKemmish et al. 2009; McKemmish et al. 2004)
Reliability	Extent to which the resource and its source are regarded as true, credible, and up-to-date for the consumers' information needs.
Accuracy	Extent to which information contained in a resource is correct, certified as free of error, or conforms to common consensus in the field.
Credibility	Extent to which information contained in a resource is highly regarded in terms of its source or content.
Currency	Extent to which information contained in a resource is representative of up-to-date practice, views and/or wisdom on a particular topic.
Relevancy	Extent to which the resource is applicable and useful for the consumer's information needs.

A simple hierarchy, as presented in Figure 5.1, can illustrate RQ dimensions and their inter-relationships.

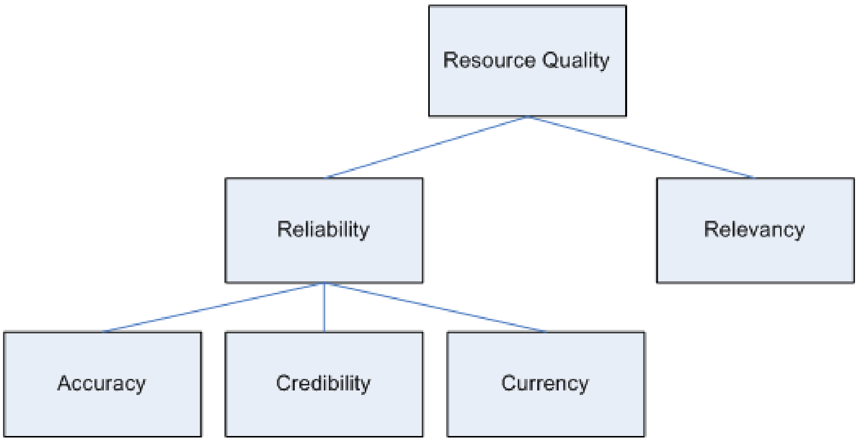


Figure 5.1 A hierarchy of RQ dimensions

The hierarchy of RQ dimensions and their inter-relationships form the basis for developing a user-sensitive RQ assessment framework, which is discussed next.

5.2.3 A User-sensitive RQ Assessment Framework

In this section, the conceptual model of a user-sensitive RQ assessment framework is presented. The model assesses the quality of web-based information resources in the context of the information needs and quality perceptions of health portal users. The development of the quality framework is based on a number of analyses, which include:

- analysis of relevant IQ literature, quality initiatives in health informatics, and IQ categories and dimensions that are relevant to the healthcare domain (as discussed in Chapter 2);
- analysis of the BCKOnline portal's resource assessment and selection criteria, and the BCKOnline metadata schema for user-sensitive resource description (as discussed in Chapter 4);
- analysis of domain expert quality appraisal activities, and expert guidance received from the case study (as discussed in Chapter 4); and
- analysis of user information needs and quality perceptions in the healthcare domain, the BCKOnline experience in particular (as discussed in previous sections of this chapter).

As a result, this research defines RQ as a composition of *Reliability* and *Relevancy*. *Reliability* is defined as the extent to which a resource and its source are regarded as true, credible, and up-to-date to meet user quality perceptions. *Accuracy*, *Credibility*, and *Currency* are regarded as sub-dimensions of *Reliability*. On the other hand, *Relevancy* is defined as the extent to which information contained in a resource is applicable and useful to meet user information needs.

The concept of *Relevancy* is derived from the notion of *Relevance* in the IR literature (Mizzaro 1997; Saracevic 2007a; Saracevic 2007b), which denotes *Relevance* as how well a retrieved document meets the information need of a user. In the context of health information portals, the quality of online resources is assessed with the aim to develop and maintain quality-assured resource collection for later retrieval. The measure of *Relevancy* in this scenario is related to, but is also beyond the perceived subject interests underling user queries. The measure of this dimension also needs to capture the context of search, such as user circumstances. Informed by the search behaviour of health portal users (Cline and Haynes 2001; Eysenbach and Köhler 2002; Williamson and Manaszewicz 2003; Zaphiris et al. 2006), and the information needs analysis conducted with the breast cancer community (Williamson and Manaszewicz 2003), *Relevancy* is defined as a composition of *User profile*, *Information preference*, and *Subject interest.*

The construct of RQ can therefore be denoted in the following formula:

$$\mathrm{RQ} = \mathrm{Ry}\ (\mathrm{Ac}, \mathrm{Cr}, \mathrm{Cu}) \cup \mathrm{Re}\ (\mathrm{Up}, \mathrm{Ip}, \mathrm{Si})$$

Where:

RQ : Resource Quality;
Ry: Reliability;
Ac: Accuracy;
Cr: Credibility;
Cu: Currency;
Re: Relevancy;
Up: User profile;
Ip: Information preference;
Si: Subject interest.

The above RQ construct informs the way to define user-sensitive resource attributes. In other words, it defines how a resource should be characterised to reflect the information needs and quality perceptions of individual users. Reflecting on the components of RQ, the following four categories of user-sensitive resource attributes were proposed to facilitate the RQ assessment:

- *Information quality* to address user quality concerns. This category describes the content quality aspects of a resource.
- *Target audience* in response to user profile. This category captures the profile information of the intended audience of a resource;
- *Type and format* in response to user information preference. This category describes the type of a resource and its presentation format;
- *Topic and content* in response to user subject interest. This category summarises the content of a resource and the subjects the resource is related to.

As discussed in Chapter 4, RQ assessment encompasses a series of decision-making processes on the rationale of resource attributes to perceived user needs and values. Value judgements made on RQ are based on domain experts' understanding of resource attributes, User characteristics, Resource Quality dimensions, and their relationships. The decision-making processes of user-sensitive RQ assessment were presented in the framework, as depicted in Figure 5.2.

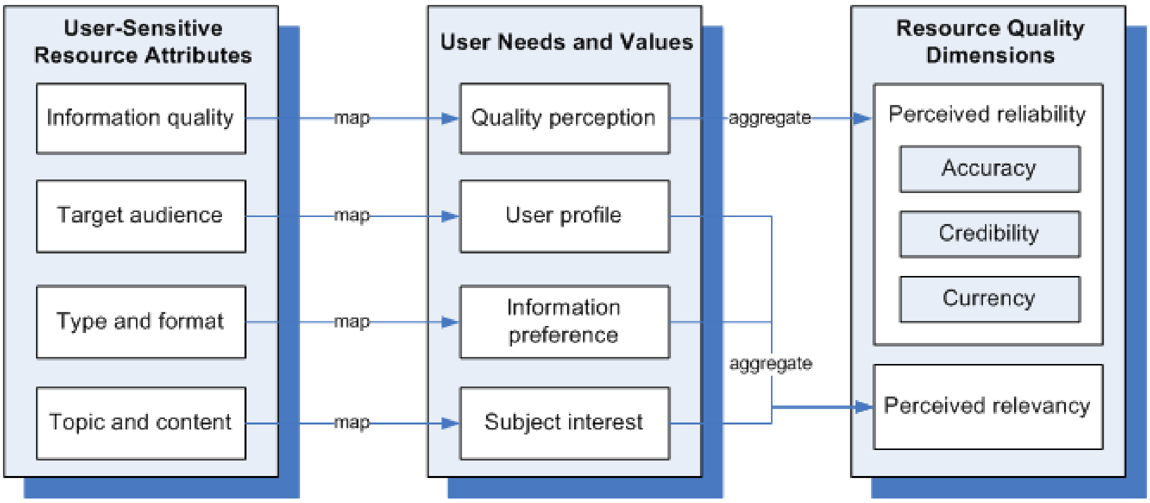


Figure 5.2 A user-sensitive RQ assessment framework for the healthcare domain

In the context of portal content management, domain experts assess the quality of a web-based information resource for its *Perceived Reliability* and *Perceived Relevancy*, based on a model of user information needs and quality perceptions. The value of *Perceived Reliability* is aggregated from a mapping of a resource's information quality attributes to the perceived user quality perceptions (e.g. whether the contained information is perceived as accurate, credible, and current). The value of *Perceived Relevancy* is aggregated from the mapping of three types of use-sensitive resource attributes to corresponding components defined by the user information needs model. As illustrated by Figure 5.2, the intended audience of a resource needs to be profiled in response to the user circumstances (e.g. the contained information is suitable for a group of users of certain disease stage and age group). Besides, the resource type (e.g. medical, supportive, or personal) and format (e.g. text, audio, or video) are described to address the specific information preference of users. Finally, the topic and content of a resource must meet the perceived subject interests of targeted portal users, which have been encoded in a portal's disease trajectory.

The proposed quality framework provides an operational mechanism for characterising and assessing the quality of online resources for their inclusion in a portal's metadata repository. The framework informs the RQ assessment processes undertaken by domain experts from a user-sensitive perspective. It also helps domain experts to systematically collect quality-associated data from different sources. They will then be able to make value judgements based on these data in a more consistent and efficient manner. Portal end-users will also benefit from the framework by knowing how the quality of a resource, in terms of its reliability and relevancy, is determined.

More importantly, such a framework represents expert knowledge of quality evaluation and description, and provides a specific quality assessment guideline for the design and development of an intelligent quality tool to assist the quality appraisal tasks. The framework facilitates the semi-automation of RQ assessment by:

- identifying quality dimensions and associated resource attributes;
- characterising user information needs and quality perceptions; and
- defining the relationships between resource attributes and user characteristics.

In this section, RQ is assessed as whether information contained in a web-based resource is perceived as accurate, credible, current, and relevant to the information preference and subject interests of profiled portal users. A hierarchy of RQ dimensions was presented, which conceptualised the quality concerns of health portal users. The design of the quality framework is in alignment with existing IQ constructs, quality assessment frameworks, as well as quality criteria models and guidelines proposed by national and international quality initiatives. The next section will describe the development process of assessment metrics using an adaptive attribute-based approach. The data metrics define the measure of the *Reliability* dimension of RQ.

5.3 Developing Quality Assessment Metrics

This section introduces an adaptive attribute-based approach for quality assessment. Based on an attribute-based data model, quality assessment metrics are instantiated for the healthcare domain. The metrics consist of subjective quality dimensions, their associated objective quality indicators, and collective quality attributes. The metrics provide a quality data model to semi-automate RQ assessment with the use of intelligent techniques.

5.3.1 An Attribute-based Approach

Wang et al. (1995a) propose the use of an attribute-based approach for data quality management. Their approach defines a data quality attribute as a collective term that refers to both subjective quality dimensions and objective quality indicators. The relationship among quality attributes, quality dimensions, and quality indicators is illustrated in Figure 5.3:

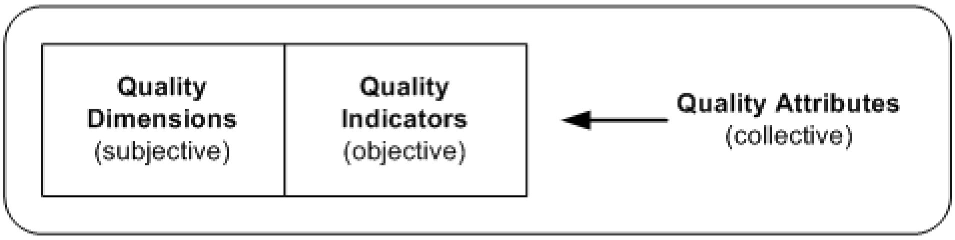


Figure 5.3 Relationship among quality attributes, quality dimensions and quality indicators, adapted from Wang et al. (1995a p. 369)

It is proposed that this approach can be utilised to facilitate the semi-automation of RQ assessment in a portal context. Adapted from Wang et al.'s (1995a, p. 354) definitions, the following terms are introduced for the assessment of RQ:

- *Quality indicators* provide objective information about the characteristics of a resource. A quality indicator is objective if it is generated using a well-defined and widely accepted measure.
- A *quality dimension* describes a qualitative or subjective single aspect of RQ, the value of which is based on the values of underlying quality indicators.
- A *quality attribute* is a collective term that refers to both quality dimensions and quality indicators. A quality attribute describes one quality aspect of a resource.

These terms define a data model to facilitate RQ assessment. The assessment consists of subjective RQ dimensions that are based on corresponding quality indicators of an online resource. These indicators can be collected from the resource's original site, or can be harvested from available metadata descriptions of the resource. Values of the collective RQ attributes can then be derived from these quality indicators using technological means. These quality attributes can assist domain experts to make informed value judgements on whether to include a resource in a portal. This model can be used to semi-automate quality assessment processes, as it explicitly defines the attributes to be taken into consideration when assessing RQ.

In the previous section, three sub-dimensions were defined to construct the *Reliability* dimension of RQ in the healthcare domain. As illustrated by Figure 5.4, based on Wang et al.'s (1995a) conceptual model (see Figure 5.3), an adaptive attribute-based data model was conceptualised to assess *Reliability*. In this model, each quality dimension is dependant on a set of corresponding quality indicators, which will be analysed in the following section.

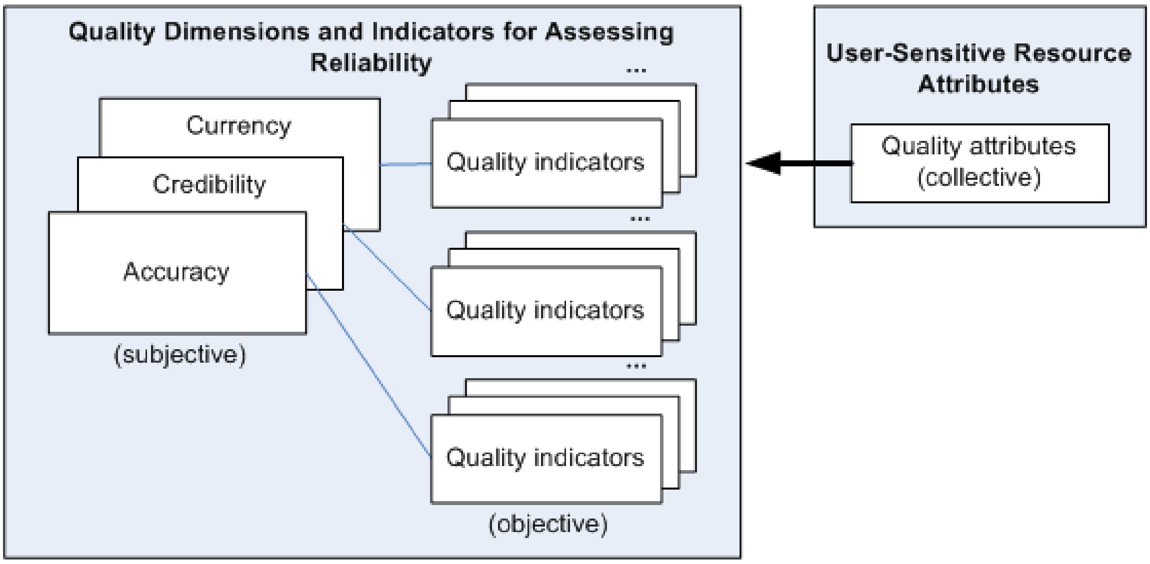


Figure 5.4 An attribute-based data model for assessing *Reliability*

Comparing to the proposed user-sensitive RQ assessment framework (see Figure 5.2), the user dimension is not explicitly represented by this model. However, it is argued that this model was devised for its potential to facilitate the semi-automation of RQ assessment. Besides, in a portal context, domain experts assess the quality of resources on the basis of their knowledge of the target audience and what is relevant to portal users. The quality attributes of a resource capture the results of value judgements made by domain experts. Therefore, to some extent, values of those attributes capture and encode domain experts' tacit knowledge of users and what users need to know to judge the reliability of the resource.

It needs to be clarified that although both *Reliability* and *Relevancy* were recognised as the determinants of RQ, this research was only concerned with the assessment of *Reliability*. How to measure the *Relevancy* dimension of RQ is outside the research scope.

5.3.2 Data Metrics for Assessing Reliability

In this section, the proposed attribute-based data model is instantiated to assess resource *Reliability* in the healthcare domain. According to the attribute-based data model introduced in the previous section, quality indicators are used to construct quality attribute values, which describe associated quality dimensions. As discussed in Section 5.2.2, the *Reliability* aspect of online health information resources can be represented by a metric of *Accuracy*, *Credibility* and *Currency* sub-dimensions. Having the quality dimension metric defined, the remaining issues are to identify the quality indicators associated with these three quality dimensions, and to define the quality attributes suitable for describing the reliability of online health information.

Quality indicators are regarded as the evidence base for the assessment of quality dimensions. Relationships between quality dimensions and various quality indicators have been discussed in Chapter 2. The analysis discussed what those quality indicators are and how to detect or generate values of those indicators using automated tools or techniques. Based on the quality indicator analysis, it is noticed that metadata elements defined by existing resource description metadata models have been used as quality indicators. In fact, most quality indicators are already captured by those metadata elements. In turn, values of those metadata elements can also be derived from other kinds of quality indicators. The difference is that metadata present concepts in a structured format whilst other types of quality indicators are presented in unstructured formats. Due to the value generation complexity of quality indicators and the abundant resource metadata descriptions available in a portal's metadata repository, using applicable metadata elements as quality indicators is considered as a feasible solution to semi-automate RQ assessment processes.

Furthermore, as discussed in Chapter 4, in order to make the resource assessment criteria transparent to portal end-users, the BCKOnline metadata schema defined an innovative *Quality* element that contained seven qualifiers and corresponding encoding systems. These qualifiers capture and encode the evaluation results of domain experts in resource metadata records. Based on these quality attributes, portal users can make their own value judgements on the reliability of resources retrieved from the portal. The seven BCKOnline quality qualifiers, namely *Credentials*, *Review process*, *Evidence basis*, *Attribution*, *Currency*, *Purpose*, and *Balance*, are selected to build a quality attribute metric.

Based on previous analysis on quality dimensions and their associated quality indicators (see Section 2.3.6), this research linked relevant quality indicators to the seven quality attributes. Detailed results are presented in **Appendix E**. Table 5.2 presents the resulting metrics of quality dimensions, quality indicators, and quality attributes for assessing *Reliability* in detail. In this table, the quality attribute metric is linked to quality indicator metrics that are associated to the *Accuracy*, *Credibility* and *Currency* quality dimensions. Two kinds of quality indicator metrics are provided. One metric is composed of 'unstructured' quality indicators that have been investigated in related studies. The other metric is composed of the 'structured' BCKOnline metadata elements (McKemmish et al. 2004).

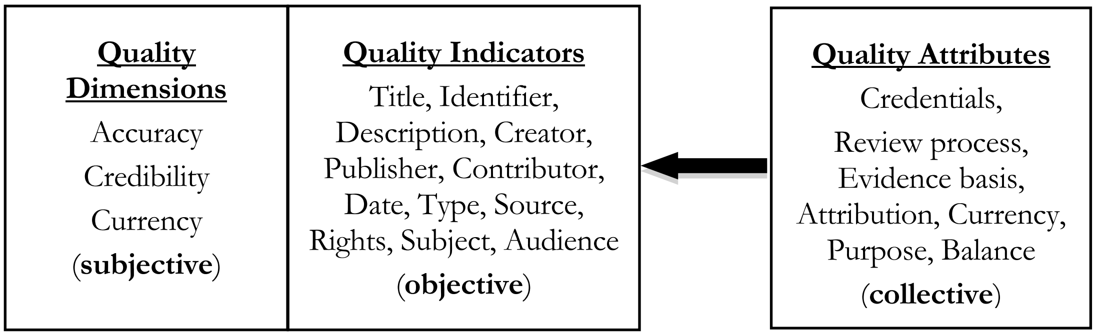
Table 5.2 Quality metrics instantiated for the healthcare domain

Metric type	Metric construct		
Quality Dimensions	*Reliability sub-dimensions*: Accuracy; Credibility; Currency		
Quality Attributes	*BCKOnline quality qualifiers*: Evidence basis; Balance; Credentials; Review process; Purpose; Attribution; Currency		
Structured Quality Indicators	*BCKOnline metadata elements*: DC.Title; DC.Identifier; DC.Description; DC.Creator; DC.Publisher; DC.Contributor; DC.Date; DC.Type; DC.Source; DC.Rights; DC.Subject; AGLS.Audience		
Unstructured Quality Indicators	*Quality indicators*	*Where to find and sample values*	*Techniques used in related studies*
	▪ Evidence-based ▪ Bias or potential conflicts of interest ▪ Authors, their affiliations and relevant credentials ▪ Ownership of the website ▪ Sponsorship of the website, contributors such as publishers, sponsors and developers ▪ Advertising, underwriting, commercial funding arrangements or support, or potential conflicts of interest, and arrangements in which links to other sites are posted as a result of financial considerations ▪ Use of user information (Privacy) ▪ References and sources for web contents ▪ Copyright information ▪ Editorial review process ▪ Web statistics ▪ Third-party labels or seal of certification ▪ Date when the actual resource was created ▪ Date when the actual resource was last updated ▪ Date on which the resource was made formally available in	DC, AGLS or HIDDEL metadata (e.g., Creator, Publisher, Contributor)	Metadata extractor
		Copyright, Disclaimer, Privacy, About us, Contact us, Linking policy of the website, Statement of purpose, Creator affiliation	HTML parser (Eysenbach 2005a; Eysenbach 2005b; HIDDEL 2003; Wang and Liu 2007)
		Third-party labels or seal of certification (e.g. HONcode)	Seal of certification (Price and Hersh 1999)
		Terms of credentials in the web content (e.g. "MD")	Search techniques and web data extraction tools (Griffiths and Christensen 2005; Griffiths et al. 2005; Naumann 2002)
		URL (e.g., .gov, .org or .com)	URL inspector (Price and Hersh 1999)
		Web statistics	Web link analysis and web usage statistics, e.g. AQA procedure, Google PageRank (Cui 1999; Griffiths and Christensen 2005; Griffiths et al. 2005;

	its current form ▪ Frequency of update and maintenance of site	Hernández-Borges et al. 1999; Price and Hersh 1999)

Based on the analysis results presented in Table 5.2, simple data metrics were abstracted to assess the *Reliability* dimension of RQ for the healthcare domain. The assessment metrics are depicted in Figure 5.5.

Figure 5.5 Data metrics for assessing the reliability of health information resources



Quality indicators defined in the above metrics are BCKOnline metadata elements derived from the Dublin Core (DCMI 2011) and AGLS (AGLS 2010) metadata models. For this reason, the data metrics are only applicable when the values of these metadata elements are available in portal metadata repositories, or can be harvested from original information sources. In other cases, website search, information extraction, or HTML parsing techniques need to be applied to generate their values. Indirect or more technical quality indicators, as have been discussed in Chapter 2, can also be used to assist decision-making on subjective quality dimensions. For instance, these can include third-party quality seals and certificates, Google PageRank, web link analysis, and web popularity analysis. However, it is questionable whether these quality indicators will positively or effectively influence the decision-making processes of individual domain experts. This issue has been addressed by previous quality indicator studies in the literature (Wang and Liu 2007), but is out of the scope of this research.

In this section, the measure of the *Reliability* dimension of RQ was defined by the attribute-based assessment metrics, which provided an operational structure to facilitate semi-automation. It is reasonable to conclude that the proposed user-sensitive RQ assessment framework and the corresponding data metrics for assessing *Reliability* met the design requirements elicited from the case study analysis (see Section 4.6).

5.4 Conceptualising an Intelligent Quality Tool

This section presents the conceptual design of a quality tool, which supports and facilitates resource quality assessment with the use of intelligent techniques. According to Wang et al. (1995b), in order to constitute an effective quality management solution, it is necessary to incorporate the devised quality framework and corresponding quality data structure into a data management system. In the context of health information portals, the purpose of having a quality framework is to standardise the user-sensitive RQ assessment in a systematic and consistent manner. The framework can only work as a guideline until it is encoded in a portal's content management system (CMS). For this reason, a quality tool was conceptualised based on the proposed user-sensitive RQ assessment framework. The quality framework together with the quality tool provides the architecture of a user-sensitive and semi-automated RQ assessment approach.

5.4.1 Addressing the Functional Requirements for a Quality Tool

As concluded from the case study analysis in Chapter 4, RQ assessment processes are composed of two highly dynamic and interconnected processes of resource selection and description. Together, these intertwined processes constitute a quality assurance mechanism over the portal content. Hence, the design of a quality tool should reflect the interactions between these two processes, and present the connections explicitly for domain experts. For instance, learning and building decision-making patterns from previous evaluation results is one way to strengthen the inter-relationships between resource selection and description.

As discussed in the preceding sections in this chapter, the final quality of a resource, in terms of *Reliability* and *Relevancy*, is highly dynamic and contextual, judged subjectively by its users. The quality attributes outlined in the proposed assessment metrics, describe the *Accuracy*, *Credibility*, and *Currency* aspects of RQ. These RQ descriptions empower portal users to make their own judgements on *Reliability*. It is proposed that a quality tool can support the decision-making processes of quality assessment by maximising the efficacy of quality attributes description. Specifically, the tool can tackle the bottleneck of detecting or generating quality indicators. This solution addresses part of the functional requirements elicited from the case study analysis for developing new portal data (see Section 4.6). In order to assist in selecting and describing new resources, the quality tool needs to provide the following two kinds of functionality.

Semi-automating the quality indicator detection and generation process

According to the way quality indicators are detected or generated, they are categorised into two groups: technical indicators and evaluative indicators. The technical quality indicators are machine detectable or measurable, whilst evaluative indicators involve human judgements. As indicated by the quality indicator analysis (see Section 2.3.6), different approaches can be used to deal with different types of quality indicators, such as harvesting existing metadata values (Greenberg 2004; Stvilia et al. 2009), parsing HTML tags (Civan and Pratt 2006; Currò et al. 2004; Griffiths and Christensen 2005; Wang and Liu 2007), measuring web links (Aguillo 2000; Cui 1999; Hernández-Borges et al. 1999), or extracting semantic entities from web context (Eysenbach 2005b). Besides, using a built-in search facility provided by major browsers is another useful method to locate information relevant to the quality indicator value generation (Price and Hersh 1999).

Semi-automating the quality attribute generation process

As discussed previously in Section 4.2, the metadata mechanism was employed by the two investigated health information portals to facilitate indexing, filtering, and value-added information provision. Domain experts' value judgements on resource attributes are captured in the two portals' metadata models for resource description. Thus, existing resource descriptive metadata values are regarded as evaluative quality indicators, from which quality attribute values can be derived. In the following two sections, available intelligent technologies are examined in terms of their ability to provide the proposed functionality.

5.4.2 Intelligent Technologies for Generating Quality Indicators

Online information resources are normally organised as hyperlinked web pages hosted on websites. As opposed to the structured data, information contained in web-based resources is not defined or organised using an explicit data schema. According to Hsu and Dung's (1998) categorisation for online information, most web content are semi-structured, as they contain natural language texts labelled by HTML tags. Although online information is increasingly presented in multi-media formats such as image, audio, and video clips, this research is only concerned with the application of intelligent technologies in extracting quality indicators from text-based web content.

There is a wide range of techniques and approaches that have been applied to extract syntactic entities or semantic concepts contained in text-based web content, such as statistical and rule-based natural language processing, HTML parsing, machine learning, and ontologies. The choice of approaches or techniques to extract information entities is highly dependent on the nature of information, and how the web content is organised and presented. HTML parsing and Information Extraction techniques are two popular ways to extract and process text-based information entities. They are considered suitable for detecting or generating quality indicator values from web pages, thus their application in addressing the problem is further discussed next.

As mentioned in the relevant literature, there are other types of intelligent techniques, which can contribute to quality indicator detection or generation. These include Z39.50 or SOAP for distributed searching, OAI protocol for metadata harvesting, algorithms of web link analysis, web popularity analysis, and user behaviour analysis. As the purpose of this study is not to compare different approaches or find the best approach, these techniques are not further discussed in this section.

Parsing HTML-based web pages

In the traditional syntax analysis field, parsing is defined as a computational process that converts individual sentences or connected texts to some representational structure useful for further processing (Warner 1987). HTML tags have been used by one study to detect possible locations of certain quality indicators, i.e. author, reference, update date etc. (Wang and Liu 2007). In the scenario of this research, parsing HTML-based web pages is one way to detect quality indicator values. Previous quality indicator studies in the literature (see Chapter 2) have proved that useful information can be extracted from different sections of a website, such as 'Copyright', 'Disclaimer', 'Privacy', 'About us', 'Contact us', 'Linking policy', and 'Quality seal'.

However, transforming online information content from its web representation to a structural representation requires extensive human analysis over the entire structure of a website as well as the template the hosting web pages apply. The parsing algorithms and the resulting web document object models also need human evaluation for their accuracy and completeness (Ye and Chua 2006). Moreover, given the extensive adoption of web 2.0 technologies, information accessible via the Internet becomes more dynamic and subject to more frequent changes. How long a parsed model will be valid or up-to-date always remains questionable.

Information extraction from text-based web page content

Information Extraction (IE) is a subset of established Natural Language Processing (NLP), text mining, and knowledge discovery research. The term describes the intelligent techniques for extracting and processing useful information entities from natural language texts (Gaizauskas and Wilks 1998). IE techniques are used for various purposes, such as abstracting or summarising texts, populating databases, identifying key words for information retrieval, and classifying text items according to pre-defined categories (Chowdhury 2004). There is no doubt that IE techniques can be applied to extract text entities in relation to information quality. For instance, template mining is one NLP based IE technique that extracts data directly from text when certain data patterns can be recognised (Chowdhury 1999b). This technique has been successfully applied in extracting citation information from online articles (Ding et al. 1999). The extracted bibliographic information, such as author, title, keywords, and the like, indicates both the *Reliability* and *Relevancy* of information resources to portal targeted users. These studies demonstrate the feasibility of using IE techniques to detect quality indicators, particularly those technical ones.

Although technical quality indicators are machine detectable (Wang and Liu 2007), the real challenge lies in the construction of extraction rules and the conversion of extracted information into a more structured form. Besides, the approach is effective only if certain standardisation in the presentation can be detected. Unfortunately, due to the dynamic and distributed nature of online information, there is no unified template utilised by web content creators. Considering the health information on the Internet as an example, the layouts of health information resources vary largely from one website to another. Mining templates for several major information sources might be useful to facilitate the process, but it is simply not practical to create templates for all candidate resources that might be useful to portal users.

5.4.3 Intelligent Technologies for Describing Quality Attributes

In order to semi-automate the process of RQ attribute description, the focus is on how collected quality indicators can be automatically aggregated to describe quality attributes. As discussed in Chapter 4, resource evaluation and description processes involve significant and contingent value judgements of domain experts, thus the selection of applicable intelligent technologies is concerned with their capability to generate quality attribute values (which are outputs) from collected quality indicators (which are input parameters). The content similarity comparison is one technique to provide value suggestions for describing quality attributes of new resources. Learning decision-making patterns from existing resource descriptions to emulate domain experts' value judgements on new resources is another feasible way to semi-automate the quality attribute description process. Both these approaches are based on the reuse of previously assigned resource descriptive metadata, which are regarded as evaluative quality indicators generated by domain experts. Potential applications of these two approaches that describe the quality attributes are briefly discussed below.

Content similarity comparison

Modelling text document similarity, especially the semantic similarity between text documents is a research field for cognitive science (Lee et al. 2005). A number of text document similarity models have been developed, among which the most popular ones are word-based and n-gram measures. These algorithms can be used to measure the content similarity of web-based information resources.

The web content similarity comparison approach attempts to use previously assigned resource descriptive metadata to provide value suggestions for describing new resources. The underlying hypothesis is that if two web resources have high similarity in their full text content, their metadata will also share some common values. In a portal context, existing resource descriptive metadata values can be used to annotate new web resources if they contain similar content.

Data mining of decision patterns

Data mining is defined as the process of discovering useful patterns and relationships in large quantities of data (Seidman 2001). Witten and Frank (2005) describe data as "recorded facts" and information as "a set of patterns, or expectations, that underlie the data" (p. xxiii). Data mining is the extraction of implicit, previously unknown, and potentially useful information from data. Such information can bring new insights and knowledge for prediction on future data. Decision patterns are explicit descriptions of decision structures that are typically expressed as sets of rules understandable by humans. Strong patterns, if found from a dataset, can be generalised to predict new data. In a portal scenario, multi-faceted attributes, including quality attributes, are fully described and preserved for each resource included in a portal. Patterns of value judgements made on quality attributes can be mined from existing resource instances to predict quality attribute values of new resources. For instance, quality attributes defined in the BCKOnline metadata schema (McKemmish et al. 2004) capture domain experts' value judgments on RQ, thus are regarded as suitable concepts for learning.

5.4.4 Conceptual Model of an Intelligent Quality Tool

Based on the data metrics for assessing the *Reliability* dimension of health information resources, and the examination of applicable intelligent technologies, the architecture of an intelligent quality tool was conceptualised. The tool was devised to support and facilitate RQ assessment involved in developing new portal data, as well as in maintaining existing portal data. That is, to semi-automate processes for quality indicator generation, quality attribute description, and maintenance of quality attribute values. The functionality of the quality tool is defined as to:

- semi-automate quality indicator generation and quality attribute description, e.g. automatically collect authorship information to describe resource credentials;
- exploit connections to previous value judgements on resource quality, e.g. how resources from the same website were previously categorised and described; and
- report and alert problematic resources that have broken links, or duplicate/invalid attribute values.

The functionality is provided by two sub-systems of the quality tool:

- A sub-system external to a portal CMS: this sub-system has components to detect and generate quality indicators. This can be achieved by employing a number of intelligent techniques, such as harvesting metadata values if there are any, parsing the original HTML-based web resources, inspecting URL links, or analysing and extracting information entities from the web content. The sub-system can be implemented as a web browser extension.
- A sub-system integrated to a portal CMS: this sub-system has components to generate value suggestions to describe quality attributes, to create and manage resource metadata records, and to monitor portal included resources and their metadata records for reporting or alerting problematic resources. The sub-system can be implemented as the domain expert interface to a portal's metadata repository.

As discussed in Chapter 4, the case study analysis indicates that part of the system functionality for supporting RQ assessment can be achieved by having enhanced data management functions. Thus, the quality tool architecture incorporates components that encapsulate data management and analysis functionality, in addition to the components that apply intelligent technologies. The architecture of the conceptual quality tool is presented in Figure 5.6.

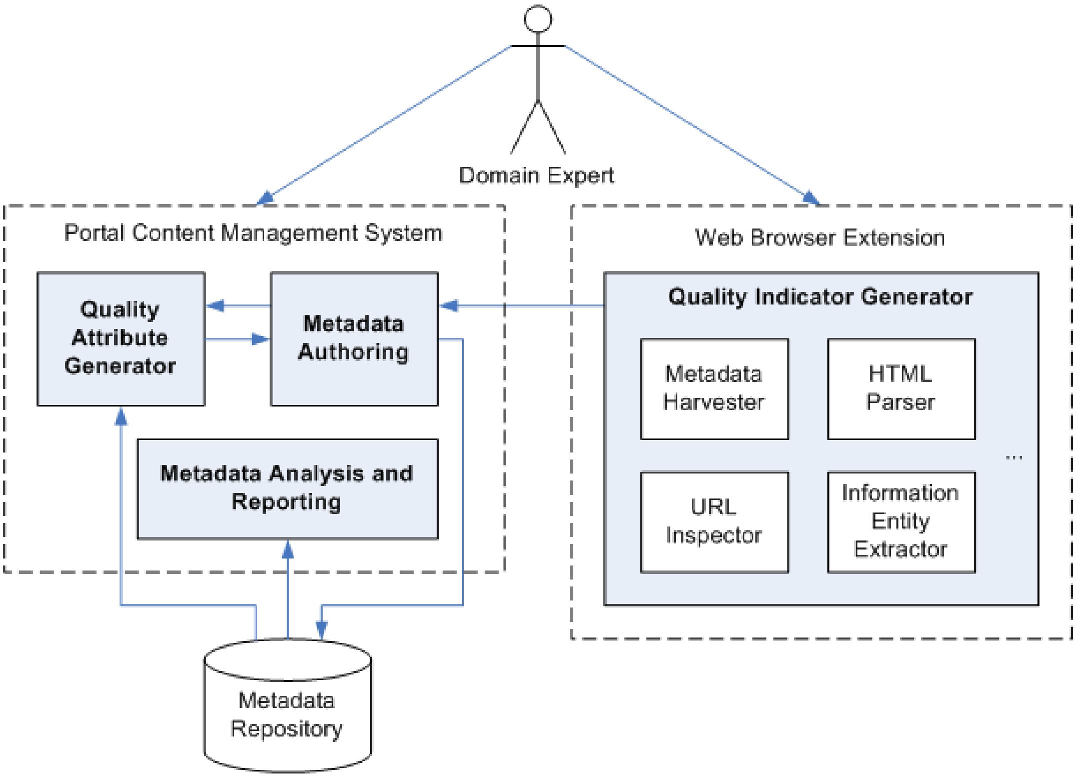


Figure 5.6 Conceptual model of an intelligent quality tool

As illustrated by Figure 5.6, the conceptual model of the quality tool consists of four key components, namely *Quality Indicator Generator*, *Quality Attribute Generator*, *Metadata Authoring*, and *Metadata Analysis and Reporting*. The four components and their interactions are introduced below:

- *Quality Indicator Generator*: it is designed as a component of the web browser extension. The component automatically detects available technical quality indicators associated with a new resource through the use of different technological means. The tool interacts with a portal CMS's metadata authoring tool, which provides domain experts with an interface for meta-tagging resources. Within the *Metadata Authoring* component, all collected quality indicators are further processed, either automatically by intelligent techniques, or manually by domain experts. Values of collected quality indicators are then transformed to structured resource descriptions, and preserved in a portal's metadata repository.
- *Quality Attribute Generator*: it is devised as a component of a portal's domain expert interface. The component employs intelligent techniques to exploit the relationships between quality indicators and quality attributes. Based on the described resources and their metadata records, this component can generate value suggestions for describing resource quality based on the mined decision patterns of domain experts. These suggestions are provided to the *Metadata Authoring* component for generating quality attribute values.

- *Metadata Authoring*: it is devised as a component of a portal's domain expert interface for managing resource metadata records. The basic data management functions include creating, editing, and deleting metadata records, as well as data browsing and searching. This component enables domain experts to assess and describe quality attributes based on values generated by the *Quality Attribute Generator* component, or quality indicators detected by the *Quality Indicator Generator* component. Besides, this component also provides created metadata values of a new resource to the *Quality Attribute Generator* to construct value suggestions for describing quality attributes of new resources.
- *Metadata Analysis and Reporting*: it is devised as a component of a portal's domain expert interface for analysing, monitoring, and alerting problematic resources and their metadata records. The component analyses resource distribution on different categories in order to identify whether a category has insufficient resources. It also picks up inconsistent metadata values, detects broken URL links, and generates reports for alerting detected problems.

The conceptualised quality tool supports RQ assessment processes by providing systematically constructed quality indicators and quality attribute descriptions for domain experts to make final value judgements on RQ. Based on machine generated information and multi-dimensional domain expertise, the overall quality of an online resource can be assessed systematically and consistently.

In the conceptual design of the *Quality Indicator Generator*, it is recognised that providing sufficient quality indicators can assist domain experts in making informed value judgements. The feasibility and efficacy of detecting technical quality indicators have been demonstrated in previous studies in the relevant literature (see discussions in Section 2.3.6 and Section 5.4.2). For this reason, how quality indicators can be specified, collected, aggregated, and presented in the *Quality Indicator Generator* component is not the focus of this research. The focus is on the feasibility and utility of the proposed sub-system of the domain expert interface, which consists of *Quality Attribute Generator*, *Metadata Authoring*, and *Metadata Analysis and Reporting* components. This research proposes the use of intelligent learning techniques to detect decision patterns from fully described resources in order to support domain experts making quality judgements on new resources.

5.5 Chapter Summary

In this chapter, a semi-automated approach was conceptualised to assess the quality of web-based information resources from a user-sensitive viewpoint. The approach encompassed a user-sensitive RQ assessment framework and an intelligent quality tool. The proposed framework defined how a web-based information resource could be characterised and assessed to reflect the profiled user information needs and quality perceptions. In this framework, information quality dimensions of *Reliability*, *Relevancy*, *Accuracy*, *Credibility*, and *Currency* were selected to construct RQ, because of their relevance and significance to address the quality concerns of health information consumers. RQ was defined as a composition of *Reliability* and *Relevancy*, while *Reliability* represented a combined concept of *Accuracy*, *Credibility*, and *Currency*. Data metrics for assessing the *Reliability* dimension of RQ was defined using an attribute-based approach. Based on the proposed quality framework and the examination of applicable intelligent technologies, the quality tool was conceptualised. It employed a range of intelligent techniques to support RQ assessment involved in developing new portal data as well as maintaining existing portal data.

Chapters 2, 4, 5 reported the research procedures and outcomes from the concept building research phase. In the next chapter, the system building research phase is described. The proposed quality tool is prototyped in the context of the BCKOnline portal.

Chapter 6

6 Design and Development of the Domain Expert Dashboard Prototype System

This chapter investigates the feasibility of the proposed semi-automated and user-sensitive resource quality assessment approach through system prototyping. A prototype system, namely the Domain Expert Dashboard, was designed and developed based on the conceptual model of the intelligent quality tool, which was discussed in the previous chapter. The Breast Cancer Knowledge Online portal provided a real-world context for building and testing of the prototype system. The chapter begins with an introduction of the architecture and functional design of the prototype system. It then describes the logical and technical components of the system. Finally, the resultant Domain Expert Dashboard is presented. The system is characterised by a number of enhanced functions and intelligent features for supporting RQ assessment.

6.1 System Design

This section introduces the architecture design of the Domain Expert Dashboard (DED). The DED provides an enhanced domain expert interface with intelligent features to better support decision-making processes associated with RQ assessment. The system was designed and developed to address the decision support needs of the domain expert, for RQ assessment as identified in Chapter 4. The system design was also based on the QFAT conceptual model as was discussed in Chapter 5.

The objective of system prototyping was to demonstrate the feasibility of the proposed semi-automated RQ assessment approach. As discussed in the previous chapter, an intelligent quality tool was conceptualised, to illustrate how intelligent technologies can be applied to semi-automate or augment RQ assessment processes. In this section, architecture of the DED is presented, followed by a description of key system functionality to support and facilitate RQ assessment.

6.1.1 Architecture of the Domain Expert Dashboard

As conceptualised in the previous chapter, the conceptual model of the intelligent quality tool consisted of two sub-systems. One sub-system supported informed decision-making by automatically detecting or generating quality indicator values. As the application of intelligent techniques in generating quality indicator values has been demonstrated in a number of previous studies in the literature (see Sections 2.3.6 and 2.4.2.4), this sub-system was not prototyped. The attention was drawn to the other sub-system, which was devised to support informed decision-making by generating quality attribute value suggestions, and monitoring the status of portal included resources and their metadata records. This sub-system served as an enhanced domain expert interface to a portal's metadata repository. As a result, a DED prototype system was developed as the enhanced domain expert interface to the metadata repository of the Breast Cancer Knowledge Online (BCKOnline) portal. Figure 6.1 illustrates the high level architecture of the DED system.

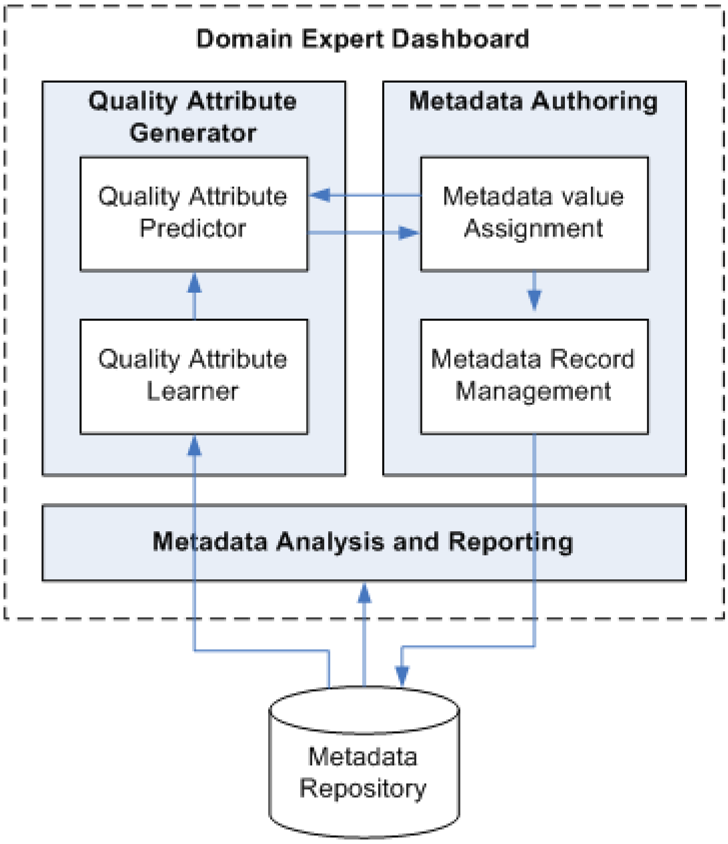


Figure 6.1 System architecture of the Domain Expert Dashboard

As can be seen from the DED architecture diagram, the system consists of the following components:

- *Quality Attribute Generator*: generates values suggestions for describing resource quality attributes. It is composed of a *Quality Attribute Learner* and a *Quality Attribute Predictor*, which provide the following functionality.
    - *Quality Attribute Learner*: transforms existing fully described resource metadata records to data attribute instances for machine learning; learns and builds prediction models for the quality attributes.

  - *Quality Attribute Predictor*: retrieves resource attribute descriptions of a new resource from the *Metadata Authoring* component, and generates value suggestions for describing the new resource's quality attributes, based on the prediction models provided by the *Quality Attribute Learner*.

- *Metadata Authoring*: represents the core metadata records management module. It provides a metadata-authoring interface for creating, editing, or deleting metadata records. It also presents value suggestions from the *Quality Attribute Predictor* to facilitate metadata generation. The functionality is provided by the following two components:
  - *Metadata Value Assignment*: deals with metadata value generation and updating; and presents value suggestions generated by the intelligent module, e.g. the *Quality Attribute Predictor*.
  - *Metadata Record Management*: provides basic metadata record management functionality, such as adding, editing, deleting, browsing, and searching metadata records.

- *Metadata Analysis and Reporting*: monitors and alerts the status of resource descriptions preserved in the portal's metadata repository; analyses resource distribution to detect resource insufficiency on certain categories; checks metadata value inconsistency; and reports URL availability.

In the DED, the intelligent system module for generating quality attribute value suggestions actively interacts with the module for metadata authoring and management. The intelligent module fetches newly generated metadata values to refine learning and facilitate prediction. Meanwhile, the status of created resource metadata records are maintained and monitored by the system model responsible for analysing and reporting detected problems. The DED key functionality provided by these logical modules is described next.

6.1.2 Functionality of the Domain Expert Dashboard

The DED provides domain experts flexible access to a range of data management services through an interactive web interface. The system features the following core functionality, which supports the decision-making processes on RQ.

- Browsing/navigation

This allows domain experts to quickly scan portal data, or to locate specific resources from the sorted resource list.

- Searching

This provides a powerful and flexible search function. Domain experts want to be able to retrieve resources that meet specific requirements, so the metadata elements used for describing resources need to be searchable for internal data management purposes.

- Collaboration

This provides a platform for supporting collaborative resource evaluation and description. Due to the complexity of RQ assessment, completing a resource metadata record may require inputs from multiple domain experts. It needs to be clarified that the system does not intend to provide intelligent support to a concurrent work process that involves multiple domain expert users. Instead, it is concerned with how the expert knowledge of existing portal data and user information needs can be transferred amongst multiple domain experts, so as to improve the quality and consistency of RQ assessment results. Moreover, the involvement of portal end-users in resource recommendation and description would be encouraged in order to sustain a portal in the long run. This requires collaborative work between domain experts and end-users to achieve quality-assured information provision.

- Metadata value generation

The conceptualised intelligent quality tool fits here to automate the RQ assessment processes. Various intelligent techniques can be applied to automatically generate metadata values for describing different aspects of an online resource, reducing the amount of time domain experts have to spend entering the resource into a portal. Based on the decision-making patterns encoded in existing resource metadata descriptions, the quality attributes and the overall quality appraisal of a resource can be derived from these generated metadata values. This functionality is particularly useful to novice domain experts, as they can build their expertise by learning from value judgements made by others. Meanwhile, the consistency of decision-making outcomes can also be improved.

- Metadata analysis and reporting

As the number of resources included in a portal grows, identifying and filling the gap of existing portal data becomes the priority task undertaken by domain experts. In order to avoid duplication of effort and reduce the workload of maintenance, it is necessary to provide domain experts an overview of existing portal metadata records, such as resource distribution under different categories and the availability of resource URLs. Both the status of portal included resources and the status of their corresponding metadata records need to be monitored and checked on a regular basis.

- Metadata record management

This focuses on the functionality for managing metadata descriptions of online resources included in a portal. Apart from basic data management functions, such as adding, editing, or deleting metadata records, enhanced data sorting and searching functions can greatly assist domain experts in finding previous value judgement results. Similar functionality can be applied to manage other types of portal data, such as user accounts, glossary terms, user feedbacks, user behaviours and usage data.

Reflecting on the domain expert needs and functional requirements identified from the case study analysis (see Chapter 4), the system functionality described above not only facilitates domain experts making informed decisions, but also supports their collaboration and learning. The intelligent metadata value suggestion functionality semi-automates the user-sensitive resource description process; the metadata analysis and reporting functionality informs user-sensitive resource selection and maintenance. Other kinds of system functionality also augment the decision-making processes on RQ.

6.2 System Development

Based on the system architecture, the logical and technical components were defined for system implementation. Constrained by the available time and resources, only one sub-system of the conceptualised quality tool was implemented. Nevertheless, the developed system provided sufficient evidence to demonstrate the feasibility of the proposed approach, and proved to be suitable for evaluating the usefulness and usability of the DED with domain expert users.

6.2.1 DED Logical Components

The DED prototype system adopted the client-server architecture, and followed the three-tier design pattern, which decoupled presentations (application), business rules, and data sources (Reese 2000). Applying this three-tier model, the user interface, functional process logics, and data storage and access, were developed and maintained as independent modules. In this research, the three tiers are named as "presentation layer", "service layer", and "data sources". Amongst these logical tiers, the service layer needs to be highlighted. In this layer, the system functionality defined in the previous section was realised using a service-based approach. Each unit of work, or a system function, was referred as a service. The functional process logics are presented as a collection of well-defined, self-contained, and loosely coupled services. These include search service, indexing service, user service, resource service, and suggestion service. Figure 6.2 illustrates the three-tier logical view of the DED, with resource service and suggestion service highlighted. The functionality encapsulated in these two components was the focus of the system prototyping. Therefore their implementation details are elaborated in this section.

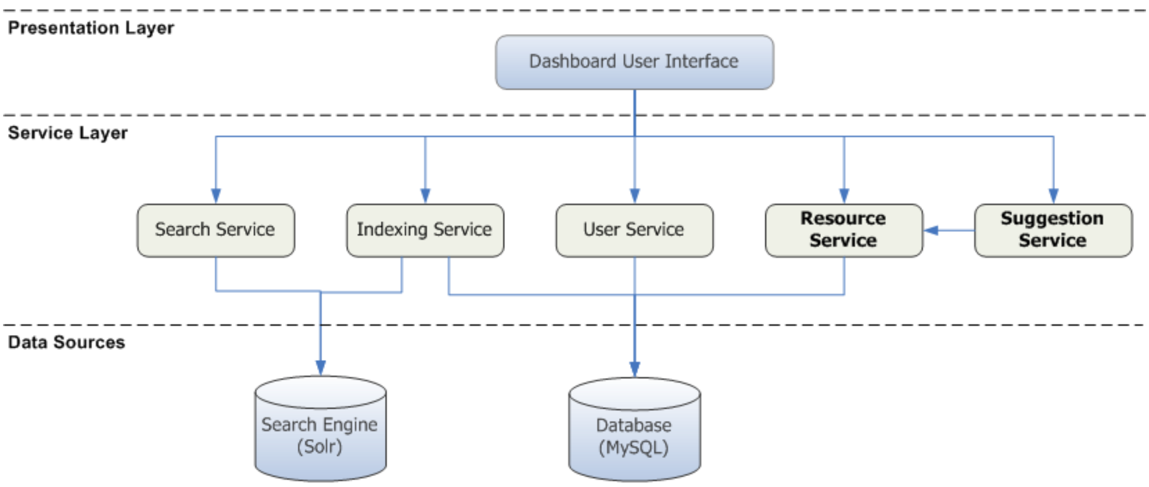


Figure 6.2 DED logical component view

Components of each layer are:

- Presentation layer

This layer was composed of two major components, including a dashboard user interface and the supporting controller that routs a request to a correct final destination.

- Service layer

The functionality provided by each service was encapsulated in a well-defined interface and corresponding implementation classes. Most of the system features are encapsulated in a set of services, namely:

  - resource service, which encapsulates basic data management functions, such as Add/Update/Delete resources and change resource status;
  - suggestion service, which generates metadata value suggestions for describing resources;
  - user service, which manages user accounts in relation to account authorisation and authentication;
  - indexing service, which indexes resources, and stores indexed terms for later retrieval; and
  - search service, which retrieves ranked resources based on the given query terms.

Each of the above services is bound to a particular application type, either locally or remotely. For instance, when a service is used locally, it is bound to a regular java object. When the service is invoked remotely, it is exposed as an XML web service or JSON service for the AJAX styled web application. All services are provided based on two components, which are:

- web support, which is a set of classes that supports html-based web application and AJAX styled remote invocation; and
- service support, which is a set of classes that exposes underlying features as a simple service interface.

As highlighted in the diagram, the development of the prototype system focused on the implementation of resource service and suggestion service, which are elaborated in the following sub-sections.

- Data sources

A Search Engine and an SQL database are two sources of data. The database is the primary data repository of resource metadata, glossary terms, user accounts and user feedback. The search engine then indexes the content within the database, and provides powerful full text search and attribute search.

The following sub-sections elaborate the implementation details of the suggestion service and the resource service.

6.2.2 Implementation of Resource Service

As introduced above, the suggestion service interacts with the resource service to retrieve resource metadata for generating prediction models. The resource service is another key component of the DED prototype system as it encapsulates all the other services for data management, data analysis, as well as URL status checking. As outlined in Table 6.1, a number of methods are defined in the ResourceService interface.

Table 6.1 Definition of resource service interface

```
public interface ResourceService {
   Long addResource(ResourceDto resource);
   void deleteResource(Long id);
   void updateResource(ResourceDto resource);
   void changeStatus(Long id, StatusType newStatus);
   void checkUrl(Long id);
   ResourceDto findResource(Long id);
   List<DataItem> findPropertyStat(String property);
   void indexAllResources();
```

```
    PaginatedList findResourceByStatus(String search, StatusType
    statusString, String orderBy, boolean isAscending, int offset, int
    len);
    Document findResourceAsXml(Long id);
}
```

The above code snippet describes how the basic system functions, such as add/delete/update resource and change resource status, are defined in the service interface. It also defines the data searching and multi-dimensional data sorting, and URL checking functions. In the next section, the resulting functions and features of the prototype system in relation to RQ assessment are introduced.

6.2.3 Implementation of Suggestion Service

This section introduces the process involved for implementing the suggestion service facilitated by the WEKA (Waikato Environment for Knowledge Analysis) data mining software (Hall et al. 2009). By using machine learning techniques, the decision patterns of domain experts were observed to best predict the information quality aspect of new resources, provided that other required descriptive metadata of the same resource are annotated beforehand. The whole experimental data mining process includes steps of data cleanup, selection of metadata elements for machine learning, selection of machine learning schema, selection of training and testing method, and model generation and application. The following sub-sections briefly introduce the experiment environment, and describe all experimental steps. Evaluation of the learning results was also carried out as part of the whole experimental process. The evaluation results are presented and discussed in Chapter 7.

Machine learning (ML) is defined as the acquisition of structured descriptions from examples that can be used for prediction, explanation and understanding (Witten and Frank 2005). The extracted patterns capture the implicit decision structure, and explain learning in an explicit way. Decision-making involving human judgement is a typical application field of ML. Both ML and statistics are data analysis techniques. ML techniques are used to learn concepts from instances and their attributes. Statistical tests are used to validate ML models and to evaluate ML algorithms. Each instance is characterised by the values of attributes that measure different aspects of the instance. It matches the case of this study perfectly.

6.2.3.1 Machine learning procedures

This research proposes the use of ML techniques to detect decision patterns from described resources so as to predict quality attribute values for new resources. This section introduces the ML procedures required for implementing the suggestion service. Three steps are involved when building prediction models:

1. selection of learning scheme;
2. selection of learning concepts and data attributes; and
3. data cleaning and transforming.

The WEKA data mining software and these ML steps are introduced next.

WEKA data mining software in Java

The open-sourced WEKA data mining software facilitated the implementation of using ML techniques to generate value suggestions. WEKA is written in Java, and distributed under the terms of the GNU license. WEKA provides a unified workbench, which contains a range of implementations of state-of-art ML algorithms and a comprehensive toolkit for data pre-processing and visualisation. Those built-in ML algorithms can be applied to any given dataset for undertaking standard data mining tasks, including classification, regression, clustering, association rule mining and attribute selection (Hall et al. 2009). It provides an API (Application Programming Interface), which allows direct use of WEKA implemented ML algorithms in other Java applications. Due to the extensive support WEKA provides for doing ML tasks, its latest 3.6 stable release was utilised for the experiments.

Selection of learning scheme

According to Witten and Frank (2005), there are four basic learning styles: classification, association, clustering, and numeric predication. The first two learning schemes are considered as applicable to solving the learning problem concerned by this research. Classification learning is also known as supervised learning as "the method operates under supervision by being provided with the actual outcome for each of the training examples" (Witten and Frank 2005, p. 43). Association rules usually involve nonnumeric data attributes, and differ from classification learning in two ways: 1) they can predict any data attribute not just the class, and 2) they can predict values for more than one data attribute at a time. For the purpose of supporting RQ assessment, the most straightforward way is to build classifiers for those RQ attributes. Mining association rules among RQ attributes and their corresponding data attributes also sounds promising, but is not considered for implementation.

To illustrate the use of ML techniques for describing RQ attributes, the classification approach was implemented for those RQ attributes measured in a nominal scale, i.e. it had a limited set of discrete values. The data source for building classifiers was resource metadata records retrieved from the BCKOnline portal. Those metadata records contained descriptions of RQ attributes, which were the targets for prediction. Also, in order to find the classification scheme that yielded satisfying prediction performance, different classification algorithms and data settings were tested via a series of ML experiments. These experimental results will be presented in Chapter 7.

Selection of concepts and data attributes for learning

The BCKOnline portal uses a comprehensive metadata schema (McKemmish et al. 2004) for describing external online resources suitable for inclusion in the portal. The schema defines a Quality element, which is composed of seven quality attributes. In Chapter 5 (Section 5.3.2), the relevance of these quality attributes to measuring the Reliability of health information has been justified. The remaining issue is to investigate whether the ML techniques are applicable to build classification models for all these quality attributes in order to support semi-automation. It is assumed that concepts learned from existing quality attribute values can be used to describe quality attributes of new resources. As a result, five quality attributes, namely *Attribution of sources*, *Balance*, *Publisher credentials*, *Purpose*, and *Review process*, were considered suitable for learning. Constrained by the available data, the other two quality attributes, namely *Currency* and *Evidence-based*, were excluded from learning. However, this does not mean that ML classification is not appropriate for these quality attributes in general. For the *Currency* attribute, all existing resources in the BCKOnline metadata repository were tagged as containing "current" information, which made the classification algorithms unable to discriminate "non-current" against "current". On the other hand, the *Evidence-based* attribute allowed multiple values, which brought more complexity in generating and evaluating learned classification models. To simplify the issue, this quality attribute was not selected. Table 6.2 below outlines the definitions and corresponding classes of the five selected quality attributes that served as the concepts for classification learning.

Table 6.2 Selected BCKOnline quality attributes for classification learning

Quality attribute (concept for learning)	**Definition** (McKemmish et al. 2004)	**Attribute values** (classes)
Attribution of sources	Whether or not the actual resource has a quality attribution, which clarifies the source of the information.	Yes, No.
Balance	What kind of issue is the actual resource, and is it noted.	Controversial Issue – Noted, Controversial Issue – Not Noted, Non-Controversial Issue.
Purpose	Describes the purpose for which this article was written.	Commercial, Discussion Forum, Educational/Informative, Reportage of Results, Review.
Publisher credentials	The authoritativeness and credibility of the individual or organisation responsible for the	Cancer Organisation, Clinician, Commercial Body, Consumer Group, Educational Institution, Government

	document.	Organisation, Lay Author, Researcher, Medical Organisation.
Review process	How the actual resource was reviewed.	Editorial Board, Peer Review Process, No Editorial/Peer Review Process.

Except the *Quality* element, which served as the 'concept' to learn, there were 16 metadata elements defined by the same schema. Those elements were evaluated for their appropriateness to serve as data attributes. As a result, eight elements were selected as data attributes for building classifiers; the other eight elements were not selected due to various reasons. Table 6.3 lists all the relevant BCKOnline metadata elements, amongst which *Title*, *Description*, *Creator*, *Publisher*, *Type*, *Rights*, *Subject* and *Audience* are noted as the selected ones for ML. The table also explains why the other eight elements were not selected.

Table 6.3 BCKOnline metadata elements

Element Name	**Definition** (McKemmish et al. 2004)	**Whether or not selected as a data attribute for learning**
Title	A name given to the resource	Selected as a data attribute
Identifier	An unambiguous and unique reference to the resource within a given context	Not selected as transforming a URL to word vectors generated too much noise
Description	A brief textual description of the content and purpose of the resource	Selected as a data attribute
Creator	The name of the person or organisation responsible for creating the resource content	Selected as a data attribute
Publisher	The entity responsible for making the resource available online	Selected as a data attribute
Contributor	The name of a person or organisation with an important contributory role in the creation of the resource content	Not selected as it had too many missing values
Date	A date associated with an event in the life of the resource	Not selected as it had too many missing values
Type	Types used to categorise the resource	Selected as a data attribute
Format	Format contains the description of the physical or virtual characteristics of the medium in which the resource is stored	Not selected as it was considered as irrelevant to the problem
Language	The language of the resource	Not selected as it had singleton value
Source	Information about a current resource and from where it is derived	Not selected as it had too many missing values
Rights	Information about rights held in and over the resource	Selected as a data attribute
Availability	How the resource can be obtained or accessed, or contact information	Not selected as it had too many missing values
Subject	Provides subject descriptors for the topic or content of the resource	Selected as a data attribute

Audience	The target audience of the resource	Selected as a data attribute
Quality	Information related to the quality of a resource	Selected as a data attribute
Relation	A reference to a related resource	Not selected as it had too many missing values

The purpose of using ML to implement the suggestion service was to demonstrate the feasibility of the proposed intelligent quality tool for describing RQ attributes. The implementation only concerned the quality attributes related to the *Reliability* dimension of RQ. Metadata elements, such as *Type*, *Format*, *Subject*, were considered as relating to the measure of the *Relevancy* dimension. Therefore they were not considered as the concepts to learn or predict. Nevertheless, it is believed that the same approach and techniques can be applied for predicting metadata elements other than *Quality*, given they have nominal values.

Data cleaning and transformation

Datasets used for building classifiers were retrieved from a single data source - the metadata repository of the BCKOnline portal. All retrieved resource metadata records need to be converted into a flat file that WEKA can understand or be able to parse. The file format is named ARFF (Attribute-Relation File Format), which represents a matrix of instances versus their attributes. An ARFF file has two distinct sections. The first section is the file header, which declares the name of the relation, a list of attributes that describe different aspects of the instance, and attribute types. The second section contains the actual data instances. Each instance is characterised by the values of attributes defined in the header. Header definition of the ARFF files used in both system prototyping and evaluation is provided in **Appendix J.1**.

As all datasets used for ML were from a single data source, data integration and aggregation were not issues to concern. However, data integrity soon emerged as a critical issue. The data integrity problem was due to the interface that domain experts used not supporting lookups and validation. If the value of a quality attribute was misspelled or incorrectly capitalised, an extra possible but unwanted value would be created for that attribute. Therefore, in order to assure the reliability of learned classifiers, before the dataset was fed into WEKA for ML, the internal data consistency was checked to eliminate inaccurate, inconsistent, or missing values.

Moreover, upon examination of the data format of the eight selected metadata elements for ML, three of them were categorical and allowed multiple values, including *Audience*, *Type,* and *Subject*. The other five elements allowed for free text entry, including *Title*, *Description*, *Creator*, *Publisher,* and *Right*. Different methods were used to transform these two groups of metadata elements to corresponding data attributes suitable for building classifiers. For those categorical elements that allow multiple values, the label of each category was transformed to a data attribute in a binary format. As a result, the applicable eight BCKOnline metadata elements were transformed into 34 data attributes for building classifiers. Later at the time of building classifiers, the WEKA built-in 'unsupervised.attribute.StringToVector' filter was utilised to tokenise the other five string-type elements to word vectors before being used for ML. All transformed word vectors were used as part of the data attributes. Details of filter settings and transformed ARFF file headers are provided in **Appendix J**. A screenshot of the parameter configuration of the 'StringToVector' filter in WEKA Explorer is provided. **Appendix J** also provides code snippets of ARFF file headers, showing how *Audience*, *Type,* and *Subject* elements are transformed into binary data attributes.

6.2.3.2 Suggestion service interface and components

Quality attribute prediction is achieved via the suggestion service. This service component interacts with the resource service and WEKA APIs to build WEKA classifiers for predicting quality attributes, including *Attribution of sources*, *Balance*, *Purpose*, *Publisher credentials*, and *Review process*. Table 6.4 below defines the kinds of suggestion services in the QualitySuggestionService interface. It takes a resource object, and predicts the corresponding quality attribute. It also provides a method, which takes a resource id and predicts values for all required quality attributes.

Table 6.4 Definition of quality suggestion service interface

```
public interface QualitySuggestionService {
  public String suggestAttributeOfSource(Resource resource);
  public String suggestBalance(Resource resource);
  public String suggestPurpose(Resource resource);
  public String suggestPublisherCredentials(Resource resource);
  public String suggestReviewProcess(Resource resource);
  public Map<String, String> suggestAll(long id);
}
```

Implementation details of WEKA-based suggestion service are depicted in Figure 6.3, which explains how WEKA classifiers are built for prediction.

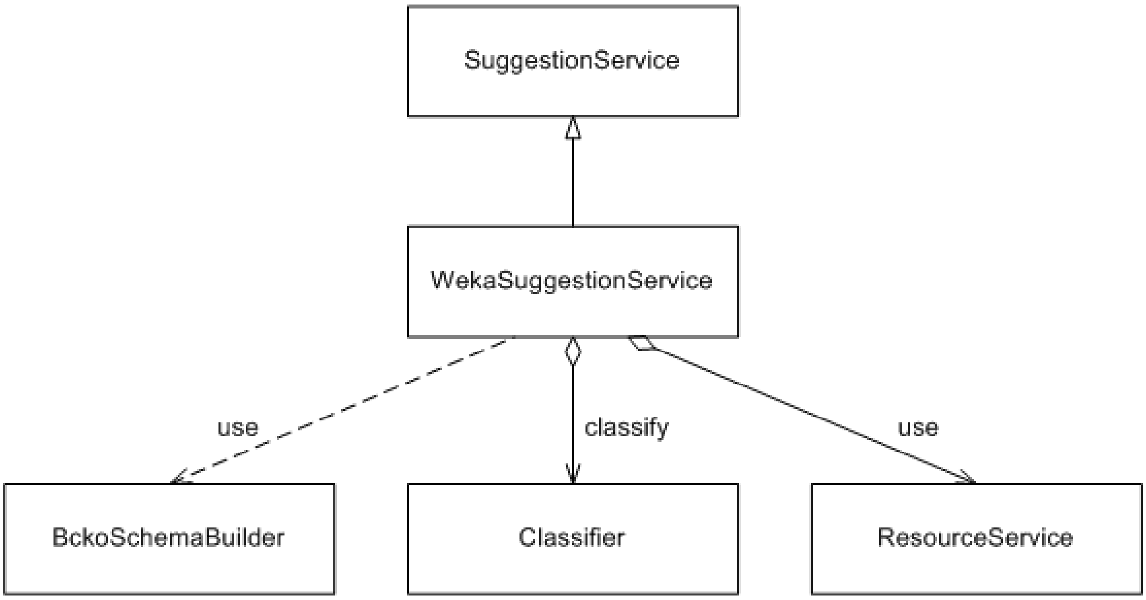


Figure 6.3 Component view of the suggestion service

The SuggestionService interface is implemented by WekaSuggestionService, which is a WEKA-based implementation. WekaSuggestionService uses ResourceService as a repository to retrieve resource metadata. It then fetches the resource metadata to BckoSchemaBuilder to create a data attribute schema that WEKA can understand. The schema is then used in WEKA to build multiple classifiers for attributes that need to be predicted.

6.2.4 DED Technical Components

The technical architecture of the prototype was based on the following considerations.

- A web based solution for collaboration - domain experts may reside in different locations. Therefore a platform needs to be provided to facilitate their collaboration.
- Cross platform deployment - it is necessary to consider a cross platform strategy to satisfy the dynamics of the deployment environment.
- Need to support full text and attribute based search - one fundamental issue of working with a large resource collection is to enable domain experts to quickly locate and filter resources.
- Need to conform to standard – there is no need to reinvent the wheel. Using a mature framework can reduce the workload. In this system, the popular Java EE standard was used.
- Scalability – it is not easy to estimate how the system will be used. To prepare for the future, it should be relatively easy to distribute the components into different physical environments.

The system contains four major components - a web application, an indexer, a search engine, and a database. The web application and the indexer run within the lightweight Tomcat application server. However, they can be hosted in any Java EE compatible application server, such as BEA WebLogic. The database and the search engine run in a separate server process. They communicate with the web application under the client/server architecture. The technical architecture of the prototype system is depicted in Figure 6.4.

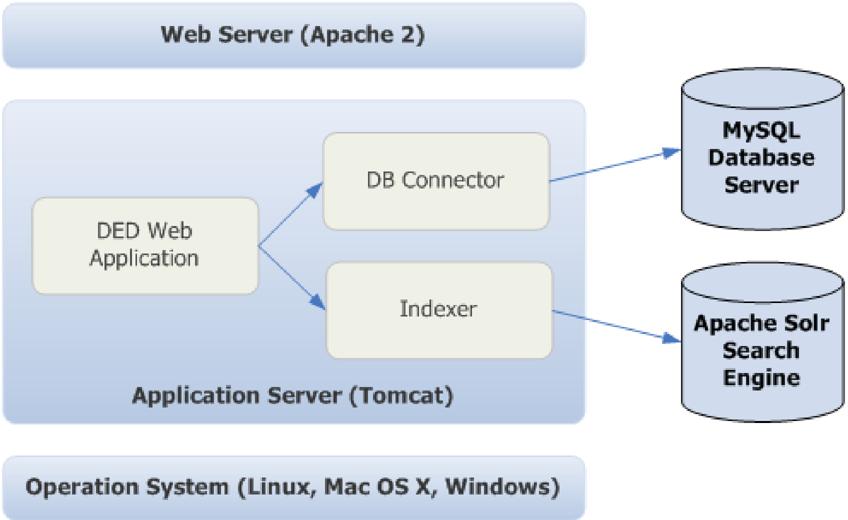


Figure 6.4 DED technical component view

The DED prototype system was built using open source and open standard technologies available under the GNU General Public License (the "GPL"). Table 6.5 specifies the software suite used to set up the development environment.

Table 6.5 Technical specification of the DED prototype system

Category	Organisation/Corporation	Product
Web Server	Apache Software Foundation	Apache
Application Server	Apache Software Foundation	Apache Tomcat
Database	Oracle Corporation	MySQL
Search Engine	Apache Software Foundation	Apache Solr
Programming Languages	Oracle Corporation	Java
	Mozilla Foundation	JavaScript
Operation System	Ubuntu Foundation	Ubuntu Lucid

DED is a web application programmed in Java and JavaScript languages using technologies such as Java EE and the Spring framework. The system connects to a MySQL database using JDBC, and uses Hibernate to map the database table to java objects. AJAX (Asynchronous JavaScript and XML) is also used, which allows data streams to be extracted from different sources, and decoupled from their presentations.

6.3 Enhanced Functions and Intelligent Features for Supporting RQ Assessment

In this section, the resulting DED prototype system is presented. The system advances the other similar systems by providing interactive data presentation, enhanced data management functions, and intelligent decision support features. Table 6.6 summarises the enhanced functions and features, which are provided by the system to better facilitate and support the decision-making processes of RQ assessment.

Table 6.6 Summary of enhanced system functions and features

Interactive data presentation	Enhanced data management functions	Intelligent decision support features
▪ Tab views for data management ▪ Data entry form ▪ Metadata grouping	▪ Multi-dimensional data sorting ▪ Data searching	▪ Data visualisation ▪ URL availability checking and reporting ▪ Metadata value suggestion for describing quality attributes

In the context of this research, features that can assist domain experts in making informed value judgments are regarded as intelligent features. For the DED prototype system, the data visualisation feature provides domain experts with an overview of resource distributions, while the URL availability checking and reporting feature constantly monitors the status of portal included resources. Both these features provide useful information that is not previously available to domain experts. This kind of information has been identified as critical to performing quality appraisals (see Section 4.3.2). As outlined in Table 6.6, implemented data visualisation, URL availability checking and reporting, and metadata value suggestion of the DED prototype system, are regarded as intelligent decision support features (Kreinovich et al. 2004). The following sub-sections illustrate the developed system from the data presentation, data management, and intelligent decision support aspects respectively.

6.3.1 Interactive Data Presentation

Enabled by the AJAX technology, the DED system provides an interactive, flexible, and efficient web interface for accessing and managing portal data. In order to facilitate the evaluation and description of selected online resource, the system provides two data management views including the workbench view and the resource management view. Moreover, a unified data entry form is used to consolidate the use experience of creating, editing and publishing resource metadata records. These data presentation modules are described as follows:

6.3.1.1 *Tab views for resource management*

The prototype system organises the management of different types of data in Tab views, such as the workbench view and the resource management view. In addition, the prototype system provides a workbench view, in which existing portal data are visualised in graphs. The workbench view will be introduced in a later section.

The resource management view provides an interface for browsing and managing resource metadata records. As illustrated by Figure 6.5, the feature is characterised by the following features:

- Resource status filter: a dropdown list is provided for filtering resources by their status;
- The width of each column is adjustable;
- Can directly go to any page of the list;
- The list content can be refreshed anytime by just pressing the Refresh button;
- Select a resource in the list to see its description underneath the list;
- Sorting data in ascending or descending order;
- Searching data on specific fields, e.g. Title or Id;
- Checking the status of one or multiple URL links at one time. The link status can be either VALID, INVALID or UNCHECKED.

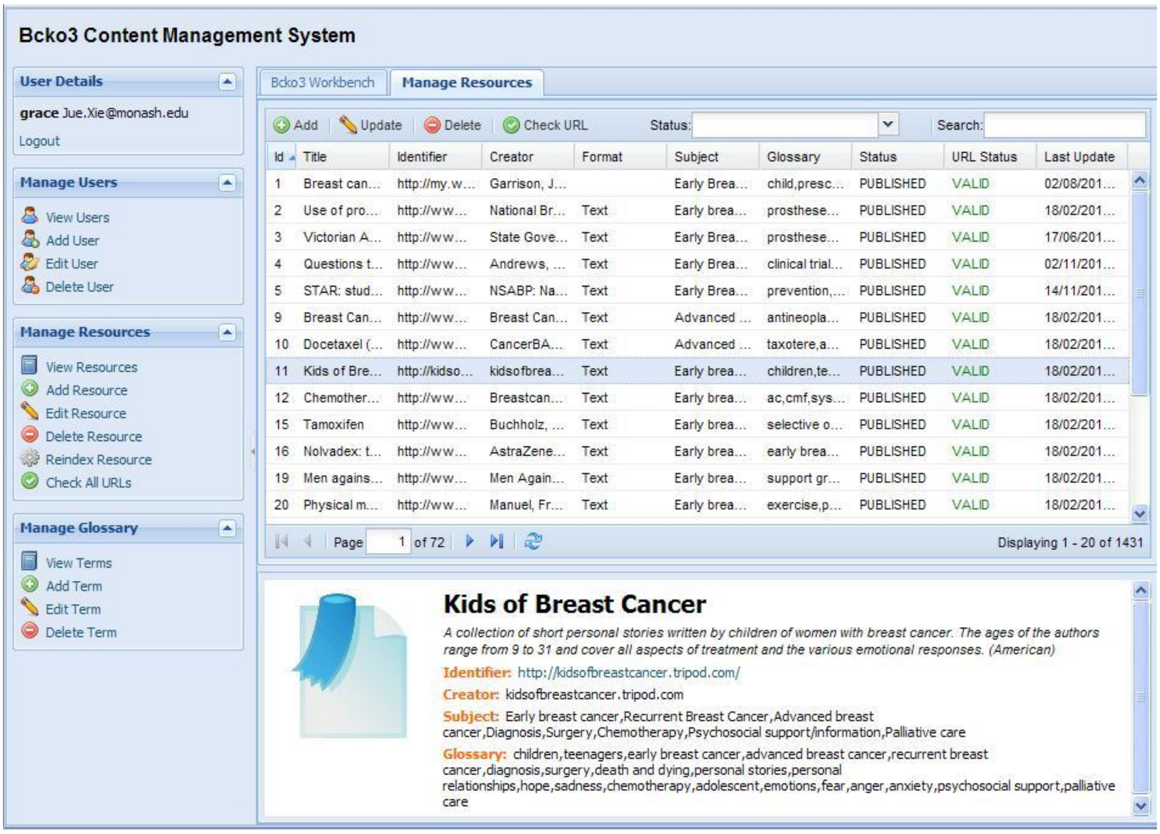


Figure 6.5 Resource list view

The resource management Tab displays resource metadata records in a configurable list view. The kinds of metadata fields shown in the list are subject to user configuration. When a record is selected in the list, full descriptions of the metadata record's key elements are presented underneath. A full record description can be viewed in a pop-up metadata entry form, as what is introduced next, by clicking the 'Update' option, or by double clicking the selected item in the list.

6.3.1.2 *Metadata entry form*

A unified data entry form is used to create, edit, or publish the metadata record of an online resource. Opening up the form is fast and flexible by left-clicking the 'Add/Update' option or by double-clicking the selected resource in the resource list view. Figure 6.6 shows the pop-up metadata entry form shown in front of the resource list view.

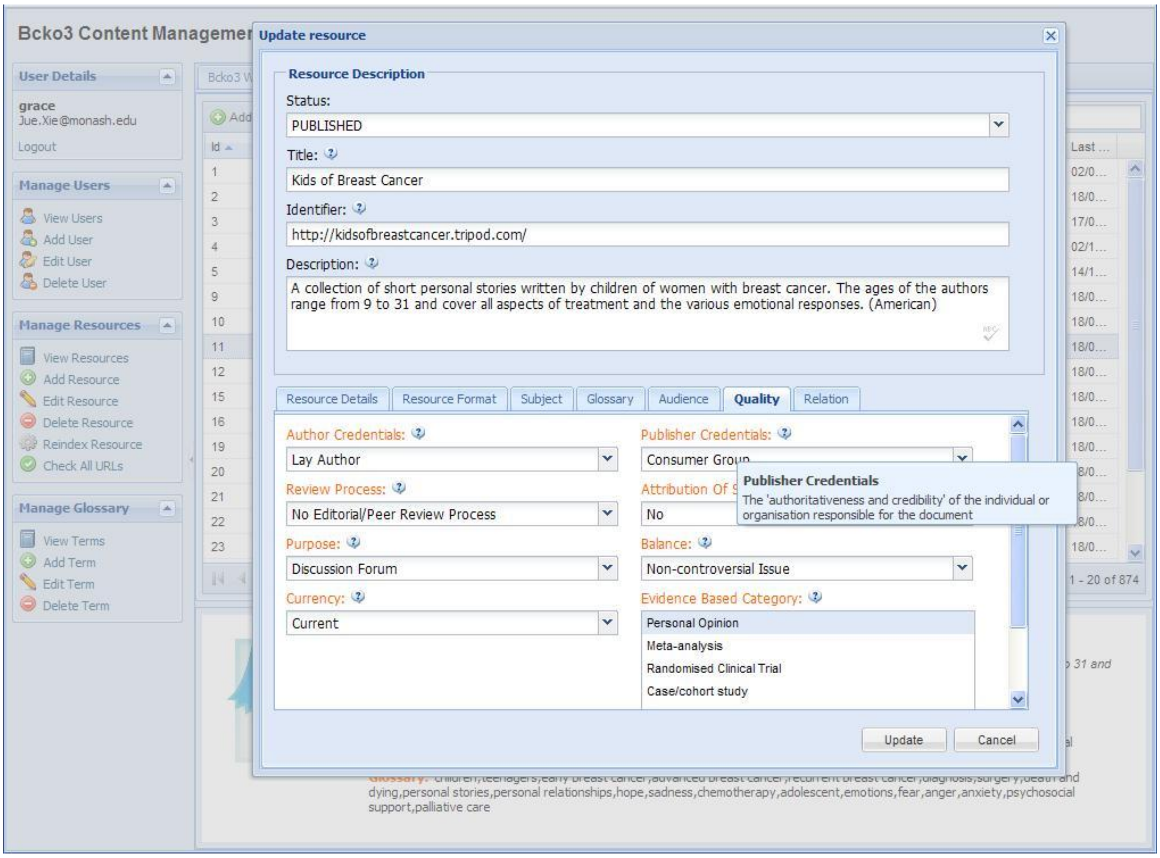


Figure 6.6 Metadata entry form

The metadata entry form provides a unified and concise metadata-authoring interface for adding or updating resource metadata descriptions. In order to display the comprehensive metadata model and all the defined metadata elements in one single form, different types of metadata are grouped and organised in their own Tabs. Users can easily switch between different metadata elements, while the status, title, URL, and description of the resource are always displayed above those metadata element Tabs.

6.3.1.3 Metadata grouping

In the data entry form, in order to facilitate the metadata assignment, metadata elements are organised in different Tabs according to the attributes of the resource they describe. As illustrated in Figure 6.6, basic descriptive metadata elements are presented in the top half of the form, while the other metadata elements are grouped and presented in their own Tabs. Metadata elements that have multiple attributes are also placed in separate Tabs, e.g. *Audience* and *Quality* element. Comparing to the traditional way that presents all metadata elements in a flat file, this design has obvious advantage when collaborations among multiple domain experts are required to complete a resource metadata record.

6.3.2 Enhanced Data Management Functions

In the DED prototype system, support for collaboration and learning is achieved by facilitating domain experts to get familiar with portal included resources, as well as to learn how other domain experts had described similar resources. As previously discussed in Section 4.3.2, the conducted exploratory case study highlights problems associated with having multiple domain experts involved in the development and maintenance of portal content. Domain experts want to avoid selecting repetitive or duplicated resources. They also want to learn from past experience so as to produce work consistently and collectively with others. Therefore, the system needs to provide support for domain experts to easily obtain an overview of portal included resources and to find relevant information from existing portal data. In order to address these issues, the DED prototype system provides a number of enhanced data management functions. For instance, the system provides multi-dimensional sorting and searching features, which assist domain experts in finding relevant information from previous decisions made by others. Such kind of information is of particular usefulness to make informed and consistent decisions on the quality of new resources.

6.3.2.1 Multi-dimensional data sorting

Sorting data in meaningful ways is a new feature to facilitate data analysis. The feature is implemented in the resource view, which displays a table of existing resource metadata. In the table, each row represents the internal metadata record of one resource; each column represents one metadata attribute of the resource. A domain expert can:

- sort text fields into alphabetical order; and
- sort numeric fields into numerical order.

Figure 6.7 illustrates an example of data sorting in the prototype system. In this diagram, published resources are sorted by their titles in an ascending alphabetic order.

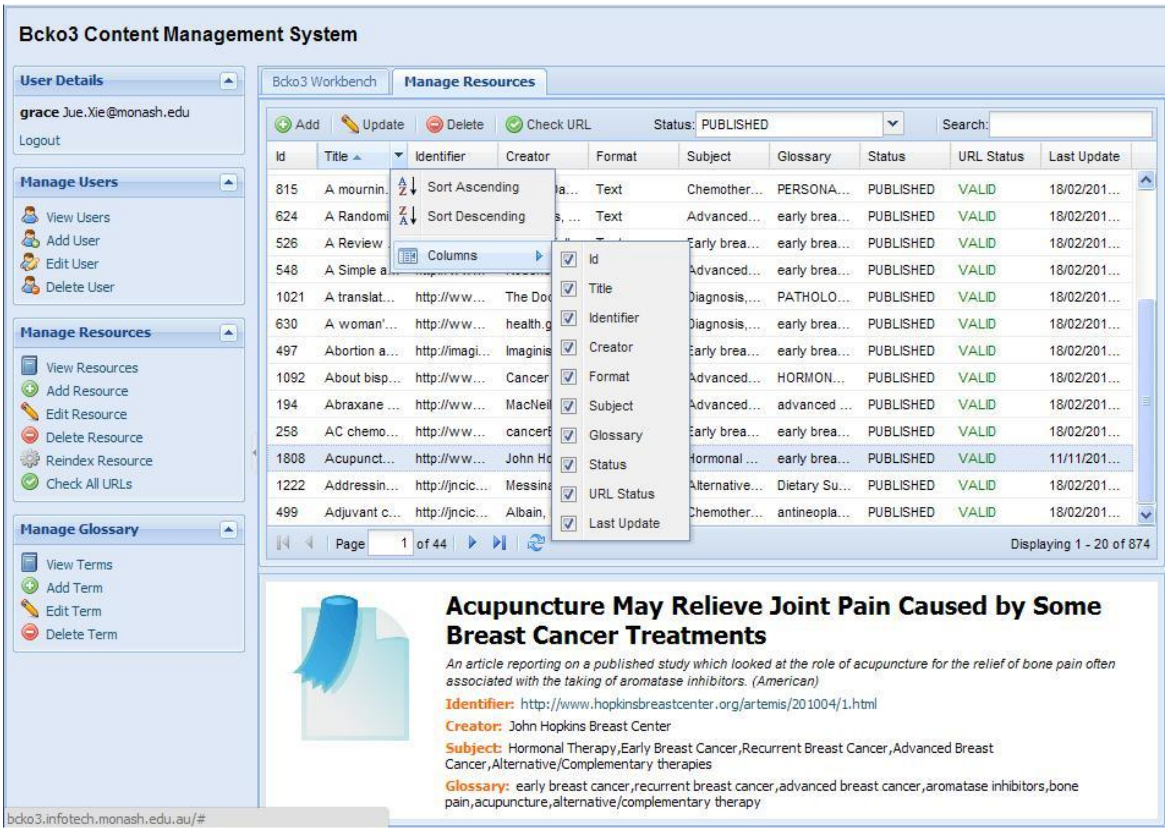


Figure 6.7 Sorting published resources

The multi-dimensional sorting feature addresses the domain expert needs for collaboration and learning. As mentioned by domain expert interviewees in the case study, they want flexible and easy navigation to existing resource metadata records by sorting those records in desired ways. This feature is particularly useful when domain experts want to gain a general view of existing portal data, to learn from previous value judgements, to locate incomplete resources ordered by certain attribute, or to group similar resources together when having difficulty in formulating specific search queries. This feature can be more powerful if used in combination with the data searching feature described next.

6.3.2.2 Data searching

Searching different fields of metadata records is another feature desired by domain experts. The feature can help domain experts to effectively retrieve specific data of interest. Such an internal search mechanism will serve the data management processes by improving the efficiency and effectiveness. Constrained by time, the comprehensive internal search functionality is not implemented. It is designed to support advanced querying as what is provided at the end-user interface. In this prototype system, the search facility is available on resource title and resource id only. Figure 6.8 presents the search result of published resources, which have the word "treatment" in their titles.

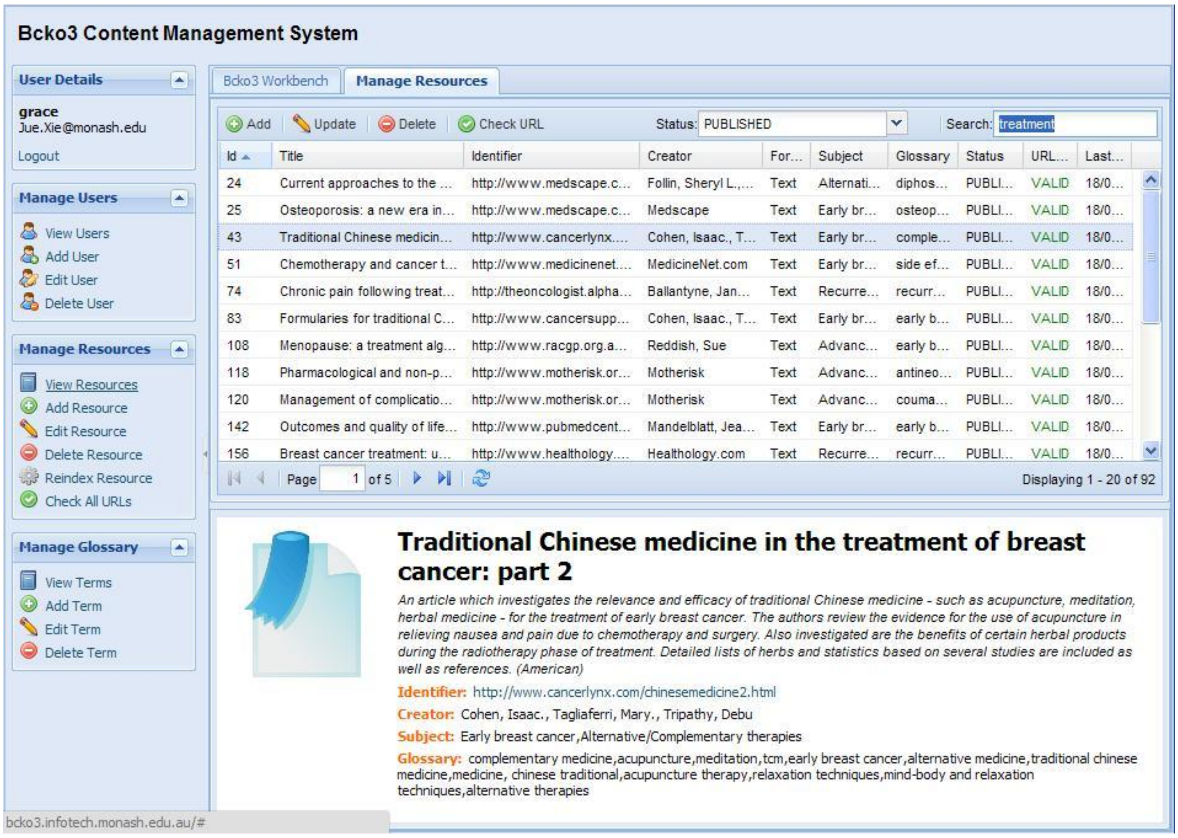


Figure 6.8 Searching published resources

The built-in search facility enables domain experts to easily retrieve a set of metadata records that share common attribute values or meet certain criteria, e.g. from the same website, created by the same author, or targeted at the same audience group. If used in combination with the multi-dimensional data-sorting feature, domain experts can easily order retrieved data, which may provide valuable insights about portal included resources, thus may assist in making value judgements on new resources.

6.3.3 Intelligent Decision Support Features

Intelligent features listed in Table 6.4 were specifically designed and implemented for their potential usefulness to support decision-making processes of RQ assessment. These features are described respectively in the following sub-sections.

6.3.3.1 Data visualisation

For the purpose of monitoring resource distribution and metadata value consistency, graphical data visualisation is provided in the form of pie chart and bar chart in the system's workbench view. These charts are rendered to visualise the existing portal data at a different granularity level of resource descriptions.

- Pie chart of portal included resources: resource status is colour coded representing PUBLISHED, UNAPPROVED or INCOMPLETE resources; the size of the colour-coded area in the chart represents the proportion of resources of a certain status.
- Bar chart of existing metadata values: colour represents the resource status; length represents the number of resources in a certain category.
- Pie chart of URL status: URL status is colour coded representing VALID, INVALID, or UNCHECKED resource URLs; the size of the colour-coded area in the chart represents the proportion of resource URLs of one status.

These data graphs enable domain experts to set a better knowledge of existing data to inform their decision-making. For instance, the graphs can help them to easily identify a lack of resources for certain categories or potential errors in the data by looking at outliers, as illustrated in Figure 6.9.

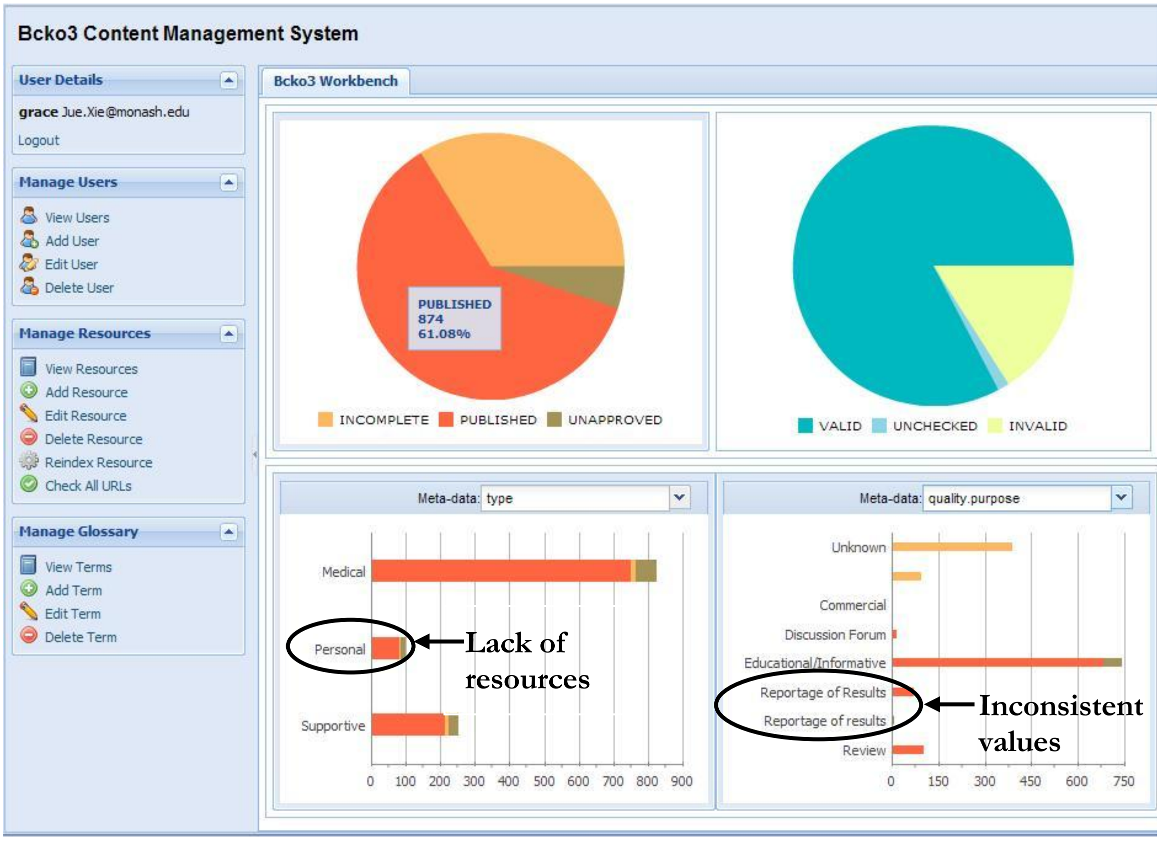


Figure 6.9 Data visualisation in the workbench view

In the above screenshot, the bottom right bar chart shows the inconsistent value of the quality attribute *Purpose* (shown as "quality.purpose"). A few UNAPPROVED resources had "Reportage of results" assigned instead of "Reportage of Results". On the other hand, the bottom left bar chart reveals that the *Personal* resource category is under supplied comparing to the other resource categories.

6.3.3.2 URL availability checking and reporting

Dealing with online information means dealing with dynamic information. Most of those URLs are not persistent links, and more likely they will be frequently changed over time. For portal end-users, they will be mostly frustrated if they retrieve resources from the portal that have broken links. Therefore, the availability of URLs preserved in the portal needs to be checked on a regular basis.

In DED, two options are provided to check the URL availability. As illustrated by Figure 6.10, the URL of one or multiple selected resources can be checked by choosing the option provided above the resource list view. Another option is provided in the main menu, which allows users to check the availability of all URLs at one time. The URL status of each resource, which has been introduced in the last section, is colour coded. This helps users to easily identify resources that have invalid links. Users can also sort resources by their URL status in order to retrieve all resources that have invalid or unchecked URLs.

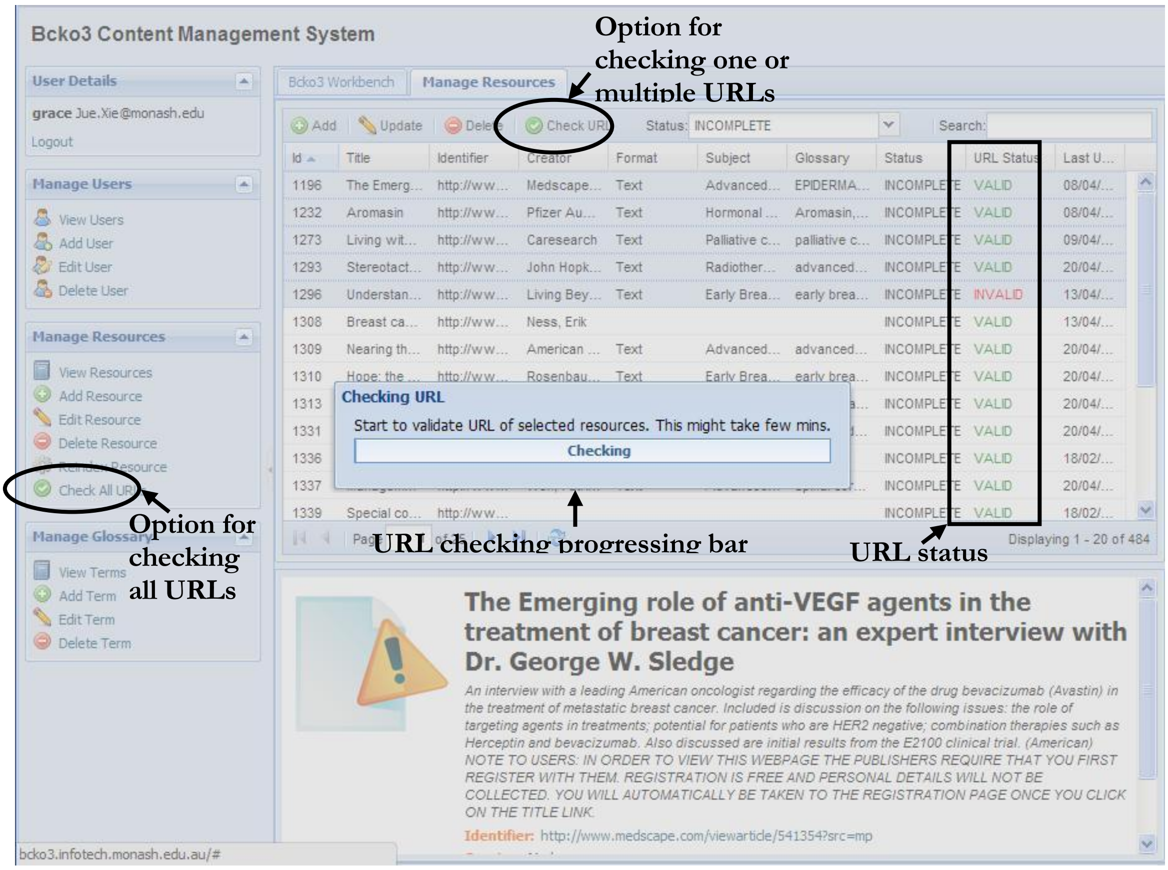


Figure 6.10 URL availability checking in progress

Although this feature allows users to quickly detect broken URLs, the system yet does not provide solutions to recover them. The difficulty is that a URL may become broken due to different reasons. For instance, a webpage may be relocated, redirected, or totally removed from the Internet. Domain expertise is required here to find a resolution to each detected problem.

Based on the URL status of all resources included in the portal, a Pie chart is generated illustrating the proportion of 'VALID', 'INVALID', or 'UNCHECKED' resources in the repository (see Figure 6.9). This has been introduced in the previous section. The Pie chart provides a means to visualise the URL availability report. It is envisaged that more sophisticated reports can be generated, which allow users to go to problematic resources directly from the chart.

6.3.3.3 *Metadata value suggestion for describing quality attributes*

Metadata value suggestion is an intelligent feature that resides in the Quality metadata Tab of the data entry form (see Figure 6.11). As introduced previously in Section 6.2.3, the feature provides value suggestions for describing five RQ attributes, namely *Attribution of sources*, *Balance*, *Publisher credentials*, *Purpose*, and *Review process*.

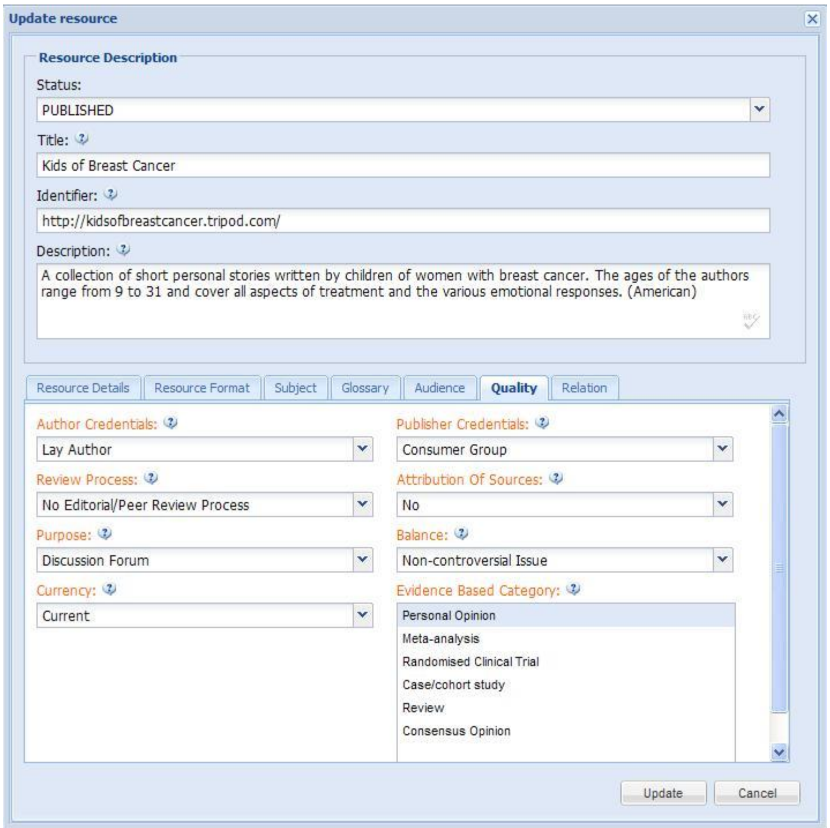


Figure 6.11 Quality metadata Tab in the data entry form

System prototyping of this research is aimed at testing the feasibility and usefulness of intelligent features in supporting RQ assessment. Attention therefore was drawn to the realisation of functionality. In this case, the realisation of intelligent learning and prediction to generate value suggestions for describing RQ attributes. As the later conducted system evaluation was not concerned with the usability of intelligent features, integrating this feature into the DED user interface was not necessary. For this reason, functionality of the metadata value suggestion feature was implemented at the service layer, but was not connected to the presentation layer (see Figure 6.2). The usefulness of this intelligent feature was evaluated separately from the system, assisted by a scenario test instrument (see **Appendix H**). Value suggestions used in scenario tests were generated by the prototype system, but via the use of command lines.

6.4 Chapter Summary

This chapter described the design and development of the Domain Expert Dashboard prototype system. The proposed intelligent quality tool was instantiated in the context of the BCKOnline portal. The system was implemented as an enhanced domain expert interface of the portal's content management system. The chapter described the DED architecture and functionality, followed by the implementation details of the system's logical and technical components. The resultant DED was characterised by a number of enhanced system functions and intelligent features, which were introduced respectively.

The DED prototype system illustrated the concept of supporting user-sensitive RQ assessment. It also demonstrated the feasibility of using intelligent technologies, such as the machine learning techniques, to semi-automate RQ assessment processes. In the next chapter, evaluation of the DED prototype system is elaborated, presenting the system's usefulness and usability in supporting and facilitating RQ assessment.

Chapter 7

7 Evaluation of the Domain Expert Dashboard Prototype System

In the previous chapter, the feasibility of the proposed semi-automated, user-sensitive resource quality assessment approach was demonstrated through system prototyping. The Domain Expert Dashboard (DED) prototype system was developed as an enhanced domain expert interface of the Breast Cancer Knowledge Online (BCKOnline) portal's content management system. This chapter presents an evaluation of the prototype system using the multi-criteria evaluation framework that was discussed in Chapter 3. It first reports a functional test of the implemented intelligent feature for describing resource quality attributes. The test measured the predictive accuracy of the feature via a series of machine learning experiments. The chapter then describes a usefulness and usability study, which evaluated the effects of the DED on the decision-making processes and outcomes of resource quality assessment. Finally, implications for design improvement are discussed before the chapter summary.

7.1 Functional Test

This section evaluates the intelligent metadata value suggestion feature, which was implemented for describing resource quality (RQ) attributes (see Section 6.2.3). According to the evaluation framework presented in Figure 3.4, the predictive accuracy of the intelligent feature was evaluated via an internal functional test. A series of machine learning (ML) experiments was carried out to assure the ML techniques worked in the desired way, and the quality of machine generated values was satisfactory (Xie and Burstein 2011). Otherwise, values suggested by the system feature for describing RQ attributes would be useless or misleading to domain experts.

In this functional test, the intrinsic validity of using ML techniques, for generating RQ attribute descriptions, was well measured by comparing the machine generated values to expert-assigned values. Statistical tests, of prediction performance of different classification schemes, were conducted using the WEKA workbench (Hall et al. 2009) to:

- compare the prediction performance of common ML schemes, and select a suitable learning algorithm for solving the specific learning problem of this research; and
- measure and evaluate the predictive accuracy of the selected learning algorithm, on the given datasets.

In the following sub-sections, the statistical evaluation method and the experimental datasets, employed in the test, are introduced. Predictive accuracy resulting from the test is also presented. It indicates the prediction performance of different learning schemes, as well as the level of confidence in using the learned models to predict new data.

7.1.1 Statistical Evaluation Method

In order to obtain reliable measurement results of predictive accuracy, the standard *stratified 10-fold cross-validation* evaluation method was employed. According to Witten and Frank (2005, p. 149), when the amount of data for training and testing is limited, the dataset cannot be assumed to have a normal distribution. Therefore, the sample used for training or testing might not be representative enough, or, on the contrary, over-representative. To mitigate any possible bias caused by a particular sample, standard statistical techniques of stratification (a primitive random sampling method) and cross-validation (a way to split data for training and testing) were adopted. A dataset was randomly divided into 10 folds, approximately of the same size. Then each data fold was used in turn for testing, and the remainder for training. For each experiment, the training and testing process was repeated 10 times with different random samples, and the overall estimate was averaged on all iterations.

7.1.2 Datasets for Experiments

All ML experiments conducted in this functional test used resources from the BCKOnline portal. Those included 780 indexed metadata records (published resources) and 100 unindexed metadata records (fully described but unpublished resources). These records were manually assigned by multiple domain experts. The retrieved resource metadata records were pre-processed and transformed, to create datasets that would be suitable for building classification models for five selected quality attributes, namely *Attribute of source*, *Balance*, *Credential publisher*, *Purpose*, and *Review process*. For each of these quality attributes, a dataset was specifically compiled for the training and testing of a classifier for the quality attribute. Each dataset was represented in a flat file, as a matrix of data instances versus 35 data attributes. These data attributes contained the quality attribute itself and the other 34 data attributes that were transformed from the BCKOnline metadata elements (see Section 6.2.3). Therefore, the only difference between these datasets was the last data attribute, which was the quality attribute to be classified.

7.1.3 Comparison of Prediction Performance

This section presents and summarises the evaluation results, of the prediction performance of different ML schemes for classification. The prediction performance of the classifier learned from a training set, was assessed by measuring the success rate on a testing set. This approach gives an objective measure of how well the concept has been learned and how confident the learned classification models can be, in predicting new data. (Witten and Frank 2005).

The experiments compared the prediction performance of the most common learning algorithms in solving the classification problems of quality attributes. These algorithms included Support vector machines, Bayesian networks, Decision trees, Rules, and Lazy algorithms (Witten and Frank 2005). Although finding the best possible classification method was not the objective of this study, experimenting with different classification algorithms helped to measure how well the concept could be learned, and by which means. The WEKA experimenter (Scuse and Reutemann 2008) provided a platform to perform the comparison, and generated the t-tests outputs as shown in Table 7.1 below.

Table 7.1 T-tests comparison of common ML methods for the problem

Dataset	SMO	NaiveBayes	J48	IB1	ZeroR
Attribute of source	83.20±3.84	85.38±2.82	83.20±3.70	75.26±3.54 •	61.80±0.35 •
Balance	78.59±4.52	77.69±5.17	72.18±4.95 •	78.33±3.39	65.38±0.00 •
Credential publisher	82.26±5.17	75.58±4.32 •	80.47±4.66	78.06±5.01 •	27.46±0.69 •
Purpose	89.38±3.47	75.70±2.90 •	85.17±1.98	85.03±2.48 •	84.02±0.98 •
Review process	84.61±3.96	71.27±5.02 •	81.28±2.29	81.80±2.34	83.21±0.33

• statistically significant degradation

SMO: Support vector machine
NaiveBayes : Bayesian network
J48: Decision tree
IB1: Association rule
ZeroR: Lazy algorithm

The numbers shown in Table 7.1 were average accuracies over 10 runs, with a random 90% train 10% test split +/- the standard deviation. The experiments used paired t-test, with a significance level of 0.05. All schemes were compared to the scheme in the first column. In this scenario, SMO (Sequential Minimal Optimisation) was selected as the representative algorithm of the SVMs (Support Vector Machines) learning method.

A statistically significant improvement means significantly higher accuracy than SVMs; a statistically significant degradation means significantly lower accuracy than SVMs. In Table 7.1, the absence of any 'statistical improvement' dots means that there were no datasets, where a classifier provided a significant improvement over SVMs, although for some datasets/classifiers the performance was neither significantly better nor worse.

The t-test results reveal that SVMs were the only classifiers that did not suffer significantly degraded accuracy, on at least one of the datasets. The performance on these datasets may be related to the large number of attribute values, which SVMs tend to handle much better than the other classifiers (Evangelista et al. 2006). The comparison to the ZeroR (Zero-attribute-Rule) algorithm is important as it indicates the classification success of the other algorithms. The algorithm does not count any data attribute, and simply classifies everything as belonging to the largest class. On a particular dataset, if no other algorithm is better than ZeroR, this indicates that the dataset is extremely hard to learn, and/or is highly imbalanced (Kubat and Matwin 1997).

7.1.4 Predictive Accuracy of SVMs

In this section, the performance of the SVM classification scheme, in predicting quality attribute values, is presented. Table 7.2 below summarises the learning results on a dataset, of 780 BCKOnline metadata records, by using the SMO classifier of SVMs. The evaluation results demonstrate that SVMs work on the learning problem in the desired way, and the quality (potential accuracy) of machine-generated values is satisfactory.

Table 7.2 Summary of learning accuracy using an SVM

Dataset (35 attributes)	Correctly Classified Instances	Incorrectly Classified Instances
Attribution of sources	658 (84.36%)	122 (15.64%)
Balance	611 (78.33%)	169 (21.67%)
Purpose	701 (89.87%)	79 (10.13%)
Publisher credentials	568 (72.82%)	212 (27.18%)
Review process	660 (84.62%)	120 (15.38%)

For a quality attributes, such as *Publisher credentials*, eliminating irrelevant data attributes from its dataset can generate results of much better accuracy for the classification problem. As denoted in Table 7.3, statistically significant improvement was received when using a dataset of four attributes, namely *Publisher*, *Creator*, *Rights* and *Publisher credentials*. The number of correctly classified instances was increased to 640 (82.05%) from previously 568 (72.82%). The latter number was generated by using a dataset of 35 data attributes.

Table 7.3 Comparison of predictive accuracy for *Publisher credentials*

Dataset	Correctly Classified Instances	Incorrectly Classified Instances
Publisher credentials (35 attributes)	568 (72.82%)	212 (27.18%)
Publisher credentials (4 attributes)	640 (82.05%) ◦	140 (17.95%)

◦ statistically significant improvement

Intuitively, this makes sense, since other attributes such as *Title*, *Description*, *Subject, Glossary* and *Audience* are probably irrelevant to the determination of a publisher's credentials. Similar cases may be found with the other quality attributes, where using a subset of data attributes will increase the predictive accuracy. Relationships between data attributes and a specific quality attribute are not explored in this thesis, as achieving better prediction performance is not the purpose for running this test. This is a potential area for future research.

In addition, evaluation results presented in Table 7.2 indicate that classifying the quality attribute *Purpose* achieved around 90% overall accuracy. It used an SVM, which was the highest amongst the five attributes. However, an overall accuracy can sometimes obscure detailed information about the class. If observing the confusion matrix and the accuracy detail by class, some learning problems can be detected. For instance, Table 7.4 below shows the confusion matrix of *Purpose*, which has overall accuracy broken down to recall and precision for each class. Here, recall refers to the proportion of relevant instances of the class, while precision refers to the proportion of accurately classified instances.

Table 7.4 Confusion matrix of quality attribute Purpose

A	b	c	d	e	← classified as	Recall	Precision
639	14	1	3	0 |	a = Educational/Informative	97.26%	92.07%
26	**53**	0	2	0 |	b = Review	65.43%	74.65%
12	0	**4**	0	0 |	c = Discussion Forum	25.00%	80.00%
14	4	0	**5**	0 |	d = Reportage of Results	21.74%	50.00%
3	0	0	0	**0** |	e = Commercial	0.00%	0.00%
					Weighted Average	**89.9%**	**88.4%**

The quality attribute *Purpose* has five pre-defined classes, namely *Educational/Informative*, *Review*, *Discussion Forum*, *Reportage of results*, and *Commercial*. For the class of *Educational/Informative*, both recall and precision were above 90%. In contrast, recall and precision values for the other classes were much lower, varying between 0% and 80%. These numbers imply that the accuracy drops when the class size gets smaller. As almost all of the training data belonged to the *Educational/Informative* class, the classifier favoured this value enormously. In the ML literature, a problem like this is referred to as "imbalanced", one that is known to be notoriously difficult to learn (Akbani et al. 2004; Kubat and Matwin 1997). Detailed experiment results, including confusion matrices of all quality attributes, are provided in **Appendix J.3**.

7.1.5 Prediction Performance on New Data

In addition to the ML experiments conducted by using published resources, the prediction performance of learned classifiers was further evaluated on a set of new resources, which were recently added to the BCKOnline portal. The five quality attribute classifiers learned by using the SMO algorithm, were applied on new resources to measure their prediction performance. 100 fully described new resources, which were not used for training classifiers, were used to compile five datasets for running the test. These 100 metadata records were randomly selected from the pool of 150 new resources of the BCKOnline portal. Table 7.5 summarises the predictive accuracy on these new data.

Table 7.5 Summary of predictive accuracy on new data

Dataset (35 attributes)	Correctly Classified Instances	Incorrectly Classified Instances
Attribution of sources	78 (78.00%)•	22 (22.00%)
Balance	81 (81.00%)	19 (19.00%)
Purpose	66 (66.00%)•	34 (34.00%)
Publisher credentials	55 (55.00%)•	45 (45.00%)
Review process	85 (85.00%)	15 (15.00%)

• statistically significant degradation

It can be seen that the classification accuracy of three quality attributes, namely *Attribution of sources*, *Purpose* and *Publisher credentials*, was degraded significantly using the new datasets. For instance, *Attribution of sources* degraded from 84.36% to 78%, *Purpose* from 89.87% to 66%, and *Publisher credentials* from 72.82% to 55%. This can be caused by a number of reasons. First, new domain experts might have been employed to select new resources and assign metadata values. Value judgements, made by a new or novice domain expert, can be highly inconsistent with the decisions made by the other experts. Second, new resources might be found from websites that are brand new to a portal. This may lead to the learned classifiers to be unsuitable for classifying new data. Finally, for a particular quality attribute, the classifier might be learned from an imbalanced training dataset, where instances of one class far outnumber the others. Thus, the classifier may not be capable of handling a new resource, if it belongs to a minority class.

7.1.6 Summary of Test Results

The ML experiment results indicate that SVM is a suitable classification method to solve the specific learning problem of this research. The achieved prediction performance on a set of published resources of the BCKOnline portal ranged from 73% to 90%, indicated the feasibility of using ML techniques, to generate value suggestions for describing RQ attributes. As the accuracy of machine-generated values is not definitive, this approach can only provide guidance to domain experts who are responsible for making the final value judgements on RQ attributes. The significant degradation of classification performance found on new data also strengthens the argument for semi-automation as a more suitable approach. The results imply that the classification models for prediction need to be fine-tuned continuously by incorporating expert decisions made on new resources. That is, in order to achieve higher accuracy or reliability of machine-generated values, it is necessary to build classification models from incremental data at the time of prediction.

7.2 Usefulness and Usability Study

In order to evaluate the DED prototype system in supporting and facilitating RQ assessment, a usefulness and usability study was conducted. Multiple criteria were used to assess both the decision-making processes and the outcomes of RQ assessment. Scenario tests and semi-structured interviews were conducted with domain expert participants, to evaluate the Efficiency (*Time and effort*, and *Ease of use*), Effectiveness (*Usefulness*) and Satisfaction (*User satisfaction*) with RQ assessment processes, as well as the Quality (*Consistency*) of RQ assessment outcomes. A scenario test instrument (see **Appendix H**) and a DED evaluation interview instrument (see **Appendix I**) were developed to facilitate the evaluation data collection. Chapter 3 had detailed the design and development of evaluation criteria and data collection instruments (see Section 3.3.5). This section justifies the selection of participants, and discusses data collection procedures, a pilot test, and data analysis results.

7.2.1 Selection of Domain Expert Participants

The evaluation study employed the domain expertise quadrant diagram (see Figure 4.5) for the selection of domain expert participants. As one of the key outputs of the joint domain expertise study, this 2×2 grid diagram was proposed for measuring multi-faceted domain expertise (Evans et al. 2009). Each quadrant of the diagram represents a different combination of information management (IM) expertise and medical/lay expertise in a disease domain. The same diagram was used for self-evaluation, during semi-structured interviews with individual domain expert participants.

This study involved three selected domain expert participants; each represents one quadrant in the domain expertise diagram. The remaining 'Novice/Novice' quadrant describes lay users. They are new to both medical and information management expertise, thus are not qualified to perform the domain expert role. The candidates were contacted by email, to determine their willingness to participate in this study. Once their participation was confirmed, a time and venue for each participant was organised. The scenario test and the interview commenced after each participant provided their consent, by signing the Consent Form after reading the Explanatory Statement (see **Appendix F**). Semi-structured interviews were recorded only if consent for recording was explicitly granted by the participant. Table 7.6 below summarises the domain expertise profiles of the three domain expert participants.

Table 7.6 Profiles of the domain expert participants

Domain expert participant	Experience in content management for health information portals	Level of experience in online search of health information	Level of experience in lay users of health information portals	Level of domain expertise
Medical and IM expert	More than 4 years	Very experienced	Very experienced	Expertise in both medical/lay and IM
IM expert	Less than 3 months	Very experienced	Limited experience	New to medical/lay; expertise in IM
Medical expert	Less than 3 months	Limited experience	Very experienced	New to IM; expertise in medical/lay

It is noted that the Medical and IM expert, who had more than 4 years content management experience for the BCKOnline portal, also participated in the joint domain expertise study (see Table 4.1). The other two recruited domain experts participated in this evaluation study only. The IM expert and the Medical expert both had less than 3 months experience in performing the domain expert role, and had limited experience in either medical/lay or IM area. Therefore, they are regarded as novice domain experts.

It needs to be clarified that the main purpose of this evaluation was to demonstrate the feasibility of the proposed semi-automated RQ assessment approach. As indicated by the domain expertise quadrant diagram, a sample size of minimum three participants is sufficient to serve the purpose of this study. After the system is refined based on the initial evaluation results, it is proposed to conduct a comprehensive system evaluation with more participants for future research (see Section 8.6).

7.2.2 Data Collection Procedures

The usefulness and usability study employed scenario tests and semi-structured interviews, to collect both quantitative and qualitative data from domain expert participants. Scenario tests were used to evaluate the intelligent metadata value suggestion feature. Semi-structured interviews were conducted to collect feedback on the usefulness and usability of DED functions and features.

Prior to participation in the interview sessions scheduled for individuals, the domain expert participants undertook the scenario test individually at times and locations of their own choice. In the scenario test, each domain expert participant was asked to complete required tasks in two scenarios: with and without using the intelligent metadata value suggestion feature. The first scenario was to describe quality attributes for two online information resources, without value suggestions. The second scenario was to describe quality attributes for the other two resources, by referring to machine generated value suggestions. Participants were asked to write down the assigned quality attribute values in the instrument, and record the time spent for completing each task. Additional information about the experience of using the intelligent feature was collected through a semi-structured interview with each participant.

Each one-to-one interview session comprised two parts. In the first half of the interview, the participant was given a brief introduction of the targeted system functions and features for evaluation. The introduction covered a number of topics, including the user interface, basic content management tasks, and the workflow to complete these tasks. Then, the participant walked through the system functions and features by performing several tasks. This process helped the participant to get familiar with the system, in order to assess whether or not the required tasks were well supported by the provided system functions and features. I wrote down the detected problems that might lead to design changes. The second half of the interview started from a questionnaire. The participant was asked to rate the *Ease of use*, *Usefulness* and *User satisfaction* aspects of the system, and was encouraged to provide additional explanations while filling in the questionnaire. Finally, the participant was asked a number of open-ended interview questions, which reflected on the overall use experience of the DED.

7.2.3 Pilot Test

In order to improve the effectiveness of the data collection, prior to conducting formal evaluation sessions, data collection procedures and instruments were pilot tested by a domain expert who had expertise in the IM area. The test was conducted as a combined session of three components, and lasted for 1.5 hours. The system evaluation session started from a 20-minute introductory tutorial of the DED prototype system. Then the participant undertook the scenario test, and took about 30 minutes to complete all required tasks. In the remaining 40 minutes of the session, the participant described personal experience and domain expertise, completed a questionnaire on the *Usefulness*, *Ease of use*, and *User satisfaction* aspects of the system, and answered follow up interview questions. The pilot test did not lead to significant revisions of the scenario test instrument, the questionnaire items, or the interview questions. However, the evaluation procedures were adjusted by separating the scenario test from the semi-structured interview. This enabled participants to complete the scenario test, at the time of their own selection before being interviewed. In order to assist participants in getting familiar with the DED, an introduction sheet of system functions and features was also created as part of the scenario test instrument (see **Appendix H.3**). The refinement of evaluation procedures improved the quality of evaluation data collected from the limited number of participants, and increased the validity of evaluation results drawn from the data.

7.2.4 Data Analysis and Results

In this section, the data collected from both the scenario tests and interviews are analysed using the evaluation framework (see Figure 3.4). That is, to evaluate whether the prototype system can improve the quality of RQ assessment outcomes, as well as the efficiency and effectiveness of RQ assessment processes. The analysis is also concerned with the overall user satisfaction in terms of whether the system facilitates RQ assessment tasks, increases decision-making confidence, and builds up domain expertise.

7.2.4.1 Quality of RQ assessment outcomes - Consistency

In the scenario test, data collected from each participant include 20 quality attribute values for describing four resources, and four corresponding time measures of the minutes the participant spent in describing each resource. The participant assigned these quality attribute values in two scenarios: with or without value suggestions. Altogether, data collected for each scenario include 30 values assigned by three participants (10 values each) and 10 values generated by the machine. Based on these data, the *Consistency* of RQ attribute values is measured and compared, which indicates the effects of the metadata value suggestion feature on the quality of RQ assessment outcomes.

For each scenario, the 10 values assigned by each participant or generated by the machine, were cross-validated based on the following rules:

R1) Scenario of "without value suggestions": If the value of a quality attribute assigned by the Medical and IM expert (the most experienced domain expert participant), is equal to the one generated by the machine, or equal to the one assigned by the IM expert, or the Medical expert, the value is regarded as potentially accurate. In addition, if a value is not equal to the one assigned by the Medical and IM expert, but is agreed by the IM expert, the Medical expert, and the machine, the value is regarded as potentially accurate. Otherwise the value is considered problematic.

R2) Scenario of "with value suggestions": Considering the impact of machine-generated value suggestions on the novice domain expert participants (the IM expert and the Medical expert), the value of a quality attribute is regarded as potentially accurate only if it is agreed by both the machine and the Medical and IM expert. Otherwise the value is considered problematic.

According to the above rules, the potential accuracy and consistency of participant assigned quality attribute values are measured and compared in Table 7.7. These measures are generated based on the scenario test results, which are provided in **Appendix K**.

Table 7.7 Comparison of value consistency

Scenario one: without value suggestions				
No. of agreed values: **M M&IME IME ME ← assigned by** 10 5 2 3 | M = Machine 5 10 3 4 | M&IME = Medical and IM Expert 2 3 10 4 | IME = IM Expert 3 4 4 10 | ME = Medical Expert				
	M	**M&IME**	**IME**	**ME**
R1: no. of values equal to those assigned by M&IME (For M&IME: no. of values equal to those generated by the Machine, IME, or ME)	5	8	3	4
R1: no. of values equally to those agreed by the Machine, IME and ME	1	0	1	1
No. of potentially accurate values	6	8	4	5
No. of values agreed by M&IME, IME and ME	2 (no. of consistent values)			
Scenario two: with value suggestions				
No. of agreed values: **M M&IME IME ME ← assigned by** 10 6 10 9 | M = Machine 6 10 6 6 | M&IME = Medical and IM expert 10 6 10 9 | IME = IM Expert 9 6 9 10 | ME = Medical Expert				
	M	**M&IME**	**IME**	**ME**
R2: no. of values equal to those agreed by both M&IME and the Machine	6	6	6	6
No. of potentially accurate values	6	6	6	6
No. of values agreed by M&IME, IME and ME	6 (No. of consistent values)			

From the measures shown in Table 7.7, it can be seen that by referring to machine generated value suggestions, the accuracy of RQ assessment results is slightly improved for the IM expert and the Medical expert, who are relatively less experienced. The metadata suggestion feature has no positive or negative impacts on the most experienced domain expert participant (the Medical and IM expert). On the other hand, the consistency of RQ assessment results across three participants is noticeably improved. In the first scenario (without value suggestions), only 2 (20%) quality attribute values assigned by individual participants are consistent with each other. The number is tripled in the second scenario, when machine generated value suggestions were provided. It can be concluded that both the accuracy and consistency of value judgements, made by novice domain experts, are improved by providing value suggestions. This result indicates that the metadata value suggestion feature is more useful to novice domain experts, who tend to learn from value suggestions provided by the machine.

7.2.4.2 Efficiency of RQ assessment processes - Time and effort

As mentioned previously, apart from the RQ attribute values, the scenario tests also measured the time each participant spent in completing individual tasks. These data are aggregated in Table 7.8.

Table 7.8 Comparison of time spent in making value judgements

	Time spent in total (minute)		
Scenario	**Medical and IM expert**	**IM expert**	**Medical Expert**
Scenario one (without value suggestions)	3	8	15
Scenario two (with value suggestions)	2	8	8

Table 7.8 reveals that the IM expert spent the same amount of time, to complete the requested tasks in both scenarios. The participant later mentioned the decision-making process in the follow up interview. That is, a similar approach was utilised by the participant to assign quality attribute values in each scenario. The following quotation describes the IM expert's decision-making process in the scenario, when metadata value suggestions were provided:

> "It wasn't clear. So, that's why it took me more time to complete it. I did everything what I suppose to do in the previous task. So I put down values that I thought would be accurate. Then I compared [them] to the ones provided." – IM expert

By contrast, the other two participants spent relatively less time to complete the required tasks, when value suggestions were provided (in scenario two). The result implies that to some extent the metadata value suggestion feature can help domain experts to describe RQ attributes more efficiently.

However, the reduced decision-making time does not necessarily mean a saving of effort for making value judgements. As pointed out by the IM expert, some provided value suggestions required double the effort from domain expert participants. The issue was whether to trust the machine generated values. If not, the participant went to the original source, checked the resource in depth, and compared it to what was suggested by the machine. The IM expert described this concern in a situation as the following:

> "…It's a bit confusing, because in this case I have to compare. So for instance, it says 'Educational/Informative'. That's pretty much straightforward. But in order to check the publisher credentials, I still go back to the Blog (the original resource), and search for it. So, I still do the same amount of work. But it does help me, if I'm in doubt. So, if I'm doubt, if it is in 'Consumer Group' or 'Commercial body', but it suggests it is 'Consumer Group', you know I probably will choose 'Consumer Group'. But I think the amount of work is probably the same for me, anyway." – IM expert

Overall, it can be concluded that the metadata value suggestion feature reduced the time and effort those domain experts spent in evaluating and describing quality attributes. Although the feature aims to complement the medical knowledge a domain expert might lack, its effect on reducing decision-making time and effort is more obvious when the expert already has extensive medical expertise. Moreover, the results also indicate that appropriate use of this feature requires formal training.

7.2.4.3 Efficiency of RQ assessment processes - Ease of use

During semi-structured interviews, the participants were asked to complete a questionnaire to evaluate the *Ease of use*, *Usefulness* and *User satisfaction* aspects of the DED constructively. In the questionnaire, the first 8 items were designed to obtain domain expert opinions on the *Ease of use* dimension of the system for performing RQ assessment tasks (See Table 3.4). Data presentation and basic system functions, such as adding, editing, deleting, browsing, sorting, and searching data were rated against a 5-point Likert scale. In addition, the participants provided explanations about their ratings during the interview sessions. Participant ratings on *Ease of use* items are aggregated and tabulated in **Appendix L**.

The results demonstrate that all three participants highly favoured the way data was presented and organised. The Tab view for data management (Item E1 of Table 3.4), and the data entry form for adding and editing resource metadata (Item E3 of Table 3.4), both received a full score. The Tab view, as presented in Figure 6.5, was highly regarded by domain expert participants as a quick and flexible way to organise different types of portal data, e.g. resource metadata records, user accounts, and glossary terms. Adding or editing a resource metadata record via the data entry form, as presented in Figure 6.6, was also intuitive and easy to use. Furthermore, multi-dimensional data sorting (see Figure 6.7) and data searching (see Figure 6.8) both received the second highest score. Their usefulness in improving the efficiency of RQ assessment processes was recognised by the participants, as illustrated by the following comment:

> "…The representation is good. It makes things easier for me to retrieve certain resources. So it will probably make the process quicker. The fact that you can view all the resources, and sort them out, is really good…I like it's possible to sort all the columns. I think it's very useful." – IM expert

The improved data presentation and enhanced system functions are no doubt appealing features of the DED prototype system. These features together provided a satisfying user experience, appreciated by all domain expert participants. The organisation of data management views, the presentation of the data entry form, and the multi-dimensional data sorting and searching features make this system more accessible and usable than other similar systems. By this means, the efficiency of the decision-making processes of RQ assessment is improved.

7.2.4.4 Effectiveness of RQ assessment processes - Usefulness

While the system's data presentation and basic functions were evaluated for their ease of use, the system's usefulness assessment focused on three intelligent features specifically implemented for supporting RQ assessment. These intelligent features include data visualisation, URL availability checking and reporting, and the metadata value suggestion feature for describing quality attributes (see Section 6.3.3).

Ratings of the participants on corresponding questionnaire items are aggregated in **Appendix L**. In addition, participant feedback on the usefulness of these three system features was also collected from the interview discussions. The received positive and negative feedback on these system features are synthesised in **Appendix M**. Usefulness of each intelligent feature is analysed in the following discussions.

Usefulness of data visualisation

Amongst three intelligent features, data visualisation was the only one in the questionnaire (the *Usefulness* category) that received full scores. All three participants confirmed the usefulness of data visualisation, as was presented in Figure 6.9, for detecting inconsistent metadata values and categories lack of resources:

> "The graphs view is so useful, because somebody can just glance [at] it and say 'ok, maybe we need [a] few more medical or [a] few more supportive resources'." – Medical and IM expert
>
> "I think this is really helpful, as it's easier for [the] human brain to deal with diagrams than the textual [descriptions]. So it's good to have understanding of what resources are currently out there; what needs to be added. It was useful as - if I made a mistake, it will basically tell me on the diagram that I introduced some other category." – IM expert
>
> "I think this is a really good way, to show me what's happening with the system. Especially when you are able to show me the number of resources for the audience of the disease stage, which I think is very good. And also this: saying that this is the number of what has been published, [or is] incomplete and unapproved. This is a very good way for me to understand…Even as simple as what is this work expecting from me as a domain expert or a content management person. What do I have to do? How well I'm going [to do] this sort of things. So this was a very good feature." – Medical expert

It can be seen that the usefulness of viewing data in colour-coded graphs was highly regarded by all three participants, as the feature allowed them to easily monitor resource distribution and value consistency. The data visualisation feature is provided in the system's workbench view. In order to encourage domain expert users to always start from the overview of existing data, the workbench view is promoted as the default view of the system. The above feedback demonstrates the effectiveness of this design as well as the usefulness of the feature, in supporting informed decision-making on RQ.

Usefulness of URL availability checking

The feature of URL availability checking received the second best score, from the participants. Only the IM expert did not award a full score. The reason was that the IM expert expected a resolution to fix the detected broken links, which was not available at the time of evaluation. Feedback was received commenting on the feature as presented in Figure 6.10. Some of the positive feedback includes:

> "If you are trying to maintain an online system […] given the volatility of things, it's very useful to know when a link is down. So [it] can be fixed immediately...in fact it's useful to grasp the immaturity." – Medical and IM expert

> "The feature of checking URLs is fabulous for domain experts. Let's say once every month for example, you can check the broken links. It is important, because otherwise I'll have to manually do it." – IM expert

> "Yes, the URL checking is extremely helpful. It's extremely good for keeping it up to date, which I think is very important. Last thing you want is a dead link." – Medical expert

As revealed from the above comments, the URL availability checking and reporting is more relevant to the maintenance of resources included in a portal. As a URL is the only link to an external resource, having a working URL is one of the prerequisites to achieve quality information provision via a portal. Moreover, checking the availability status is just the first step. The next questions ask how to fix the detected broken links, and whether the links can be fixed. More assistance is required in performing tasks such as finding an updated URL for the same resource, or finding another similar resource as a replacement.

Usefulness of metadata value suggestion

During the semi-structured interviews, the participants were asked to reflect on their use experience of the intelligent feature, and to describe the problems they encountered and concerns they had with values suggested by the machine. The purpose was to evaluate whether the intelligent feature helped to reduce efforts for making value judgements, and to identify how this feature could be improved. By comparing the feedback of the three participants, three issues were raised that should inform future design and refinement of the intelligent feature. Table 7.9 summarises individual participants' responses to these issues.

Table 7.9 Responses of participants to machine generated value suggestions

Issue	Medical and IM expert	IM expert	Medical expert
Trust in machine generated values	No	Yes	Neutral
Extent of accuracy	It is possible for the machine to generate confirmative values which would be more useful	Machine generated suggestions would be more useful if knowing how sure they are true and accurate	Providing suggestions only is not good enough. More information is needed in terms of how sure the answer is

Resolution of conflicts	Always rely on own judgements	Prefer what the machine suggests	Choose not to make any decisions

Issue one: Trust in machine generated values

Participants from different domain expertise backgrounds have different levels of trust in machine generated metadata value suggestions. Those who are experts in the medical domain, or those who are familiar with online information diversities, are less likely to believe the capability of intelligent tools in handling complex evaluative tasks. Comments below illustrate the doubts of domain expert participants on machine-generated values.

> "Working as a so-called Expert/Expert, I wouldn't rely on the machine. Because what happens is that you go to the actual resource, and there are lots of things you need to look for. There is not enough for any computer to spill out any suggestions or whatever. You still need to deal with [it], a lot more deeply." – Medical and IM expert

> "Maybe I'm the person of the next generation; trust machines more than sometimes I trust people. Because with machines, that's algorithms. I'm sure that certain values can certainly be filled out by machine, and it will minimise the work of domain experts. It will also give me some sort of confidence." – IM expert

> "I did look at the suggestions first. I had a look at them and I thought, you know, that I just want to check it, and make sure what is being said to me is correct." – Medical expert

Establishing the trustworthiness, for domain experts of machine-generated values, is never an easy task. During the interviews, I explained to the domain expert participants how the feature worked in order to encourage trust. The participants suggested that providing explanations on how the suggestions were generated, and based on what data, would help them to understand the underlining mechanism. By this means, more trust can be built up. This can be done during a training process, when a domain expert is new to the system.

Issue two: Extent of accuracy

All participants recognised the perceived usefulness of having a machine generating some suggestions. It is recommended that the confidence level of prediction (how accurate the suggestion could be) is required together with the suggestions. The Medical and IM expert was particularly in favour of receiving some assistance from the machine, to replace certain repetitive work. In some scenarios, it is possible for a machine to generate confirmative values, the accuracy of which can be assured. For these kinds of values, the machine can pre-populate the fields in the data entry form rather than providing value suggestions, as no further effort will be required from domain experts, to check their validity. The scenario was described as follows:

> "…If you look at some records I've done, some just took me four minutes. I realise what I was doing then. I was going through all the ones from the same source, and simply just cut and paste. If I'm indexing one after another, when two articles [come] from the same (Journal) issue, again it's cut and paste. It's really quick." – Medial and IM expert

> "If I could search, I'd like to rearrange or be able to search publisher, and have the whole list of resources from the same publisher. I would be able to search on 'Journal of Clinical Oncology', and all the resources from that publisher come up to the list. So I just go down, and cut and paste, and it takes me a fraction of time." – Medial and IM expert

> "I'm sure that certain values can certainly be filled out by machine, and it will minimise the work of domain experts. It will also give me some sort of confidence. I think you can be fairly confident that there are certain values that can be proposed by the machine fairly accurately." – IM expert

While the most experienced domain expert participant (the Medical and IM expert) doubted the capacity of the machine to generate meaningful quality attribute descriptions, the other two less experienced participants focused on the reliability of value suggestions. The likely accuracy of the suggested values determined whether it would be necessary for them to double-check the value suggestions. When the likely accuracy of machine-generated values is not disclosed to domain experts, it becomes necessary to both validate the value suggestions, and to analyse the original source. This doubles the effort to make value judgements on RQ. In this sense, the accuracy or reliability, of the machine-generated values, will greatly affect domain experts' opinions on the usefulness of the metadata value suggestion feature. The following comments illustrate their concerns:

> "…Only providing the suggestion values is not good enough. I need a little more there in terms of how sure the answer is." – Medical expert

> "What I don't understand is: why do the double work? If the machine generated values are based on what experts had said, […] why would you need me to check? So, in the end, what if I'm wrong or the machine is wrong? I don't quite understand how it could help me to learn that. However, this could be automated, like certain aspects of it, like it can give certain extent of confidence." – IM expert

It is agreed by all participants that if certain resources are from the same website or the same portal, it can be assumed that they will share some common information in their descriptions. The machine can pre-populate definitive values, rather than provide suggestions for domain experts to approve. For certain quality attributes, such as *Publisher credentials* and *Review process*, it is possible to fully automate the value assignment process, when the resource is from a website that has been previously analysed by a portal. Descriptions of resources from the same website can be reused.

Issue three: Resolution of conflicts

It was mentioned by all participants that when value suggestions were provided by the machine, they chose to make their own decisions. They would visit the original web resource, and refer to the other metadata descriptions of the same resource. The Medical and IM expert tended to ignore the value suggestions, and used the usual approach to describe quality attributes. In contrast, the IM expert and the Medical expert, who were relatively less experienced, chose to check a resource on its original website. They attempted to learn how the machine came up with values as suggested. When they came up with different opinions, to what the machine suggested, they reacted in different ways. The IM expert tended to take the suggested values but with doubts, whilst the Medical expert chose to not assign any values for those uncertain fields. The following feedback describes how they resolved the conflicts, when their opinions were different to what the machine suggested.

> "In the end of the day, I still don't know with what I filled up; whether this is right or that table is right. So if it was more precisely written with regards to which elements are [accurate, and] which are not, [it would] be definitely great. "
> – IM expert

> "I wasn't sure about the one the machine says: it has 'Editorial Board'. I couldn't make out…When I looked at the website myself, I couldn't come up to that conclusion. But otherwise everything is in line with what are provided." – Medical expert

The above responses of novice domain experts imply the importance of having quality-assured value suggestions and a mechanism to make the predictive accuracy transparent. Otherwise, the usefulness of the metadata value suggestion feature will be undermined. Providing value suggestions with poor or unknown quality can affect the confidence of novice domain experts, in making value judgements, as well as their trust in machine generated values. It might result in more confusion and doubling of effort, and eventually lead to totally ignoring the value suggestions provided by the machine.

Concluding remarks on the usefulness of intelligent features

Overall, all participants agreed that the three intelligent features provided by the DED helped them to make value judgements, in a productive and consistent manner. The two novice domain experts also felt more confident about the quality of values they assigned. Moreover, it is noticed that the intelligent features are more effective in supporting RQ assessment when the domain expert already has the required medical knowledge for making value judgements. This is illustrated below.

> "…Even though I agree with certain aspects of the machine, I'm still not sure who is right. That's because, as I said, I'm not a very confident user." – IM expert

Although the usefulness of the metadata value suggestion feature was recognised by all the participants, the feedback provides areas for design refinement. For instance, as suggested by the IM expert, providing value suggestions with their predictive accuracy rates, to domain experts, can improve the usefulness of the metadata value suggestion feature.

7.2.4.5 Satisfaction with RQ assessment processes - User satisfaction

The overall user satisfaction with the system, in terms of how well the system satisfied domain expert needs, was both quantitatively measured by using the questionnaire, and qualitatively evaluated through interview discussions. The evaluation also indicates the validity and appropriateness of the system design. Participants' ratings on *User satisfaction* items of the questionnaire are aggregated in **Appendix L**. Moreover, collected from the interview discussions, participants' feedback on system utility are synthesised in **Appendix M**.

User satisfaction items, in the questionnaire, received relatively high ratings. It can be concluded that all three participants were satisfied with the utility of the DED in supporting and facilitating RQ assessment. It is noticed that the Medical expert gave the highest ratings on both overall satisfaction and system utility, but had a neutral view when asked to compare the system to other similar systems. The Medical expert participant had extensive knowledge of disease conditions of breast cancer and intimate knowledge of breast cancer patients. However, the participant had limited experience in using content management systems. Therefore, the participant had no other systems to compare with the DED.

In addition to the explicit ratings collected by the questionnaire, the following comments illustrate the perceptions of domain expert participants, on the overall system utility.

> "All those features certainly make it easier to manage the whole thing that was a pain in the past." – Medical and IM expert
>
> "I really liked it and I think it's great. I think it's quite useful and usable and presented pretty well." – IM expert
>
> "It is quite a good system, quite simple and quite easy to use. It's quite easy to understand what it is and what the purpose of the system is and what a person can do with it." – Medical expert

Although the overall system utility was satisfactory, in order to assure the effective use of the system, it is necessary to provide formal training in addition to written instructions or user manuals. One participant explicitly raised this issue during the interview:

> "I suggest that if I'm new, it would be great if I could receive some short tutorial or something. When you train a domain expert, you also need to explain where to go and what to look for [assigning] particular values." – IM expert

Furthermore, in order to evaluate how well the identified domain expert needs were addressed by the system, the interview data were analysed, reflecting on the original objectives of the system design and its functional requirements. As a result, the analysis encompassed three system utility issues, i.e. the utility of the DED in supporting RQ assessment, building medical expertise, and building IM expertise (see **Appendix M**).

The utility of the system, in terms of supporting RQ assessment and building medical expertise, received mixed feedback. However, feedback on the system utility in building IM expertise was all positive. The Medical and IM expert had used the system in a traditional fashion, for resource cataloguing and description. The utility of the system in supporting decision-making processes was not explicitly recognised by this participant. However, implicit influence of the system, on the participant's decision-making processes, was revealed from interview discussions. As illustrated by the following feedback, it was agreed by all participants that the system helped them avoid duplicate effort, minimise mistakes, and make informed decisions:

> "All I can do to minimise my mistakes would be good. It's fabulous if the system can say 'no' - that we already have the resource." – IM expert

> "If I'm thinking of a URL, I might look at the graph and say 'no we've got enough on breast reconstruction'." – Medical and IM expert

> "You are looking at so many different attributes of that particular resource. I think you've got enough information there. That explains relevance. So that's why they are important" – Medical expert

On one hand, all participants acknowledged that the system functions and features, provided for locating, managing, maintaining and analysing resources, helped them to obtain required IM skills. On the other hand, the IM expert pointed out that the system lacked an explanatory mechanism, and thus did not help much in filling the medical knowledge gap.

> "I don't think it can teach me about where and what I need to check, to fill the attributes in. Because it doesn't explain which one and why the decision was made." – IM expert

Drawing from the synthesised participant feedback, it is reasonable to conclude that the DED prototype system addresses the domain expert need for making informed decisions on RQ, and assists them in building IM expertise. However, it is relatively harder to help domain experts build medical expertise, which is recognised as a big and steep learning curve.

7.2.5 Summary of Evaluation Results

This data analysis presents encouraging evaluation results on the overall usefulness and usability of the prototype system, although different opinions were received on certain aspects of the system. As the participants had different levels of domain expertise, it was expected that they would have different interests, concerns and needs in regard to the use of the same system. Table 7.10 below provides a summary of the perceived effects of the prototype system on RQ assessment processes and outcomes, together with the associated evidence. Based on these data, it can be concluded that the prototype system provides useful and usable functions and features. These can help domain experts assess and manage RQ in a more efficient and effective manner. The accuracy and consistency of RQ assessment outcomes can also be improved through the use of this prototype system.

Table 7.10 Summary of evaluation results

Perceived effects on	Evaluation criterion	Summary of evaluation results	Data evidence
RQ assessment outcomes	Quality - *Consistency*	The consistency of decision-making outcomes was improved, and the impact was more significant on the novice domain expert participants.	Quantitative data collected from the scenario test (**Appendix K**)
	Efficiency – *Time and effort*	Time spent in making value judgements was reduced, but not the effort. When value suggestions are questionable, double effort will be required to check both the machine generated values and the original web resource.	Quantitative data collected from the scenario test (**Appendix K**); Qualitative data collected from interview discussions (**Appendix M**)
	Efficiency – *Ease of use*	The system was easy to use. However, due to the unavoidable complexity of the system, formal training is required in addition to written instructions.	Quantitative data collected from the questionnaire (**Appendix L**); Qualitative data collected from interview discussions (**Appendix M**)
RQ assessment processes	Effectiveness - *Usefulness*	The usefulness of data visualisation and URL checking and reporting features were highly regarded. The potential usefulness of the metadata suggestion feature was also recognised. However, further design improvement was requested.	Quantitative data collected from the questionnaire (**Appendix L**); Qualitative data collected from interview discussions (**Appendix M**)
	Satisfaction - *User satisfaction*	The overall utility of the system in supporting RQ assessment was satisfactory. The system supported domain expert participants making informed value judgements on RQ, but it could better facilitate RQ assessment with some design improvement. It helped in building IM expertise, but was relatively less effective in building medical expertise	Quantitative data collected from the questionnaire (**Appendix L**); Qualitative data collected from interview discussions (**Appendix M**)

Overall, both the RQ assessment processes and the outcomes were improved, with the use of the DED prototype system. The effects of the system on decision-making processes and outcomes are more obvious for novice domain experts, who require more guidance in assessing and describing RQ. The evaluation results prove the validity of the system design. Meanwhile, the study also identifies areas for design improvement. These are discussed next.

7.3 Implications for Design Improvement

The purpose of this prototype system evaluation is not only to assess the utility of the design artefact, but also to identify areas for design improvement. It was found that domain expert needs are evolving through the use of a new system, which applies new technologies to address the identified needs. The prototype system introduces new possibilities to the domain expert participants, and brings up new domain expert needs. During the interview discussions, the participants provided valuable suggestions about how the system could better assist them in understanding the tasks, and how it could be better designed to address emerging new ideas and new requirements.

Amongst all system functions and intelligent features being evaluated, the metadata value suggestion feature is a debatable one. The feature has received mixed feedback, which leads to a better understanding of domain expert needs and their interactions with the system. Major improvement can be sought on the design of this feature, e.g. adding an explanatory facility, colour coding value suggestions according to the perceived predictive accuracy, or pre-populating certain metadata fields where definitive values can be generated.

Moreover, it is critical to assure the quality of value suggestions provided by the machine. Learning from the best is important to domain experts, particularly to those less experienced. It will be problematic if the machine introduces certain values that might not be accurate. Drawing from the functional test results, it is suggested that the ML prediction models need to be built on incremental data. That is, descriptions of new resources will be used to fine-tune those models for achieving better prediction accuracy. Providing an explanation facility therefore becomes particularly essential. Domain experts need to be aware of the accuracy of the suggested values, and avoid strengthening 'wrong' decisions by repeating them. Knowing the likely accuracy of machine suggested values, can also give domain experts some measure of confidence about the decisions they make.

The evaluation results, particularly those measures drawn from the ML experiments on new resources, and the feedback on the metadata value suggestion feature, further confirm the appropriateness of the proposed semi-automated RQ assessment approach as opposed to automated approaches (Eysenbach and Diepgen 1998; Griffiths and Christensen 2005; Wang and Liu 2007). The concept of semi-automation is to enable part of the work to be done by the machine; another part of the work done by domain experts. Appropriate system design needs to facilitate the collaboration between machine and domain experts. The IM expert's comment "minimise the effort" summarises the ultimate goal of this type of system. That is, not only eliminating mistakes and improving decision-making consistency and accuracy, but also reducing the required human effort, and increasing productivity and decision-making confidence. In order to avoid doubling effort, the system needs to be designed in a way to help domain experts use time smartly. For instance, when assessing RQ, effort of domain experts should be required only when the machine indicates relatively low confidence about the values it generates.

In the future, the same intelligent techniques, which have been used for describing the quality aspect, can be applied to describe other aspects of resources. It is envisaged that when assessing and describing a web-based resource, the machine could provide value suggestions, for all those fields that it can possibly predict. In order to minimise the effort required from domain experts, the traffic light idea can be utilised to signify the fields that require further checking or approval. It is envisaged that similar approaches could be employed to perform ongoing evaluation on the refined quality tool.

7.4 Chapter Summary

This chapter presented the evaluation results on the usefulness and usability of the DED prototype system in supporting RQ assessment. Data collected from the functional test and the usefulness and usability study were analysed and reported.

As introduced in Chapter 3, a multi-criteria evaluation framework was applied to assess the effects of the DED on both the decision-making processes and the outcomes of RQ assessment. The criterion addressed by the functional test results was the *Predictive accuracy* of RQ assessment outcomes. Criteria addressed by the usefulness and usability study included *Consistency*, which evaluated the quality of RQ assessment outcomes, and *Time and effort*, *Ease of use*, *Usefulness*, and *User satisfaction*, which evaluated the Efficiency, Effectiveness, and Satisfaction with RQ assessment processes.

The internal functional test carried out a series of Machine Learning experiments. The usability and usefulness study employed scenario tests and semi-structured interviews. Three domain expert participants empirically evaluated the usefulness and usability of the DED in supporting and facilitating RQ assessment. Based on their use experience, areas for design improvement were identified and discussed. The evaluation results demonstrated that both the decision-making processes and the outcomes of RQ assessment were improved through the use of the DED prototype system.

Chapter 8

8 Conclusions

This chapter presents how user-sensitive resource quality assessment was conceptualised and rigorously evaluated employing a socio-technical design science research approach. It begins with a thesis summary, reviewing the user-sensitive systems development research process discussed in relevant chapters. It then summarises how research questions were addressed, presenting a set of key findings and outcomes for the healthcare domain. Reflecting on the socio-technical design science research experience, a method transferable for other domains was proposed. This method can produce domain-specific and user-sensitive resource quality assessment metrics. The later part of this chapter highlights research contributions, and summarises the adaptation of Hevner et al.'s (2004) design science research guidelines to evaluate this research within a socio-technical context. Finally, fields and directions for future research are suggested before the chapter summary.

8.1 Thesis Summary

This thesis presents multi-disciplinary and socio-technical design science research, concerned with conceptualising and supporting user-sensitive resource quality (RQ) assessment in a healthcare domain. The research is committed to finding new solutions that employ intelligent technologies to improve the scalability and sustainability of RQ assessment for metadata-driven health information portals. As a result, a semi-automated RQ assessment approach was proposed. It encompassed a user-sensitive RQ assessment framework and an intelligent quality tool, to standardise and facilitate the decision-making processes on RQ. This research applied and adapted the design science research framework and guidelines (Hevner and Chatterjee 2010; Hevner et al. 2004; March and Smith 1995). The research was informed by the theorising approach for design science research (Venable 2006), the user-sensitive design methodology (McKemmish et al. 2009), and the systems development research methodology (Burstein 2002; Nunamaker et al. 1991).

A user-sensitive RQ assessment approach was conceptualised, prototyped, and evaluated via three inter-connected research phases, namely concept building, system building, and system evaluation. In the concept building research phase, the construct of RQ and the design requirements of new RQ assessment solutions were defined. It comprised a comprehensive analysis of multi-disciplinary research literature and a case study of domain expert RQ assessment practice in the context of two metadata-driven and user-sensitive health information portals.

The development of the user-sensitive RQ assessment approach needs to be in alignment with the information quality (IQ) research literature. In Chapter 2, IQ research in organisational and web information systems was investigated. IQ and its assessment is a well-established research field in the Information Systems (IS) discipline (Klein 2001; Lee et al. 2006; Madnick et al. 2009; Naumann 2002; Price and Shanks 2005b; Stvilia et al. 2007; Wang et al. 1995b). The field provides useful theories for constructing and assessing IQ in a systematic manner. In the literature, IQ is defined and assessed for serving a specific purpose or use (Eppler 2006; Redman 1992; Wang and Strong 1996). The review of general and specific IQ solutions revealed the diversity of IQ constructs, frameworks, and assessment approaches. Due to the subjective and contextual nature of IQ and the dynamics of user needs and values, operational frameworks are imperative to enable domain experts make contextual value judgements on RQ in a timely manner. However, existing IQ research does not provide operational solutions to assess RQ from a user-sensitive viewpoint. Although existing IQ assessment approaches do not directly address the decision support needs of domain experts for RQ assessment, they form the basis for developing new solutions that follow user-sensitive design principles (McKemmish et al. 2009). As a result, this research articulated the RQ assessment processes and domain expert needs for intelligent support. The concepts of RQ and user-sensitive RQ assessment were developed for health information portals.

Drawing on the outcomes of the exploratory case study, this research concluded that consistent and efficient RQ assessment required collaborations between domain experts and intelligent tools. Given the abundance of online health information and its varying quality, manual approaches do not scale the processes. On the other hand, fully automated approaches cannot realistically replicate the role that domain experts play in RQ assessment. Therefore, as described in Chapter 4, partially automating the processes may be an optimal solution.

In addition, it was determined that quality assurance of portal content was achieved via a set of decision-making processes undertaken by domain experts through the use of the content management system (CMS) of a portal. RQ assessment is the interplay of three factors: domain experts, RQ assessment processes, and portal CMS (as depicted in Figure 1.1). Intelligent tools can be provided to assist domain experts in selecting and describing online information resources for their inclusion in a portal. Moreover, it was identified that domain expertise was multi-dimensional, including medical expertise, lay patient/user experience, and IM skills. In order to help domain experts build this multi-dimensional expertise and to support the making of informed value judgements, an intelligent quality tool was proposed for integration into the portal CMS. These findings implied the requirement of developing socio-technical solutions rather than conventional technical solutions, to address user-sensitive RQ assessment issues.

As a result, in Chapter 5, a semi-automated RQ assessment approach was conceptualised, which encompassed a user-sensitive RQ assessment framework and an intelligent quality tool. The quality framework defined the construct of RQ (as a composition of *Reliability* and *Relevancy*) and its assessment in the context of user information needs and quality perceptions. Facilitated by an attribute-based approach, the feasibility of the proposed framework was empirically tested in a healthcare domain. Based on the analysis of quality dimensions, quality indicators and quality attributes, assessment metrics were developed to qualitatively measure the *Reliability* dimension of RQ for the healthcare domain. The quality framework further informed the conceptual design of an intelligent quality tool to support RQ assessment in that context.

In Chapter 6, the system building research phase was described. The feasibility of the proposed RQ assessment approach was demonstrated through system prototyping. The Domain Expert Dashboard (DED) prototype system was developed in an operational health information portal. The prototype system was designed to provide greater functionality to meet identified domain expert needs for making informed decisions, collaboration, and learning. In the DED, machine learning (ML) techniques were applied to generate metadata value suggestions, which assisted domain experts in describing RQ attributes.

Chapter 7 reported the system evaluation research phase. Effects of the DED on RQ assessment processes and outcomes were assessed, using a multi-criteria evaluation framework. This included an internal functional test, and a usefulness and usability study with domain expert users. According to the evaluation framework, the quality of decision-making outcomes was measured in terms of predictive accuracy and consistency. The efficiency, effectiveness, and user satisfaction with the decision-making processes were measured in four areas, including time and effort spent in making value judgements, ease of use of enhanced system functions, usefulness of intelligent features, and overall user satisfaction. The evaluation of the prototype system demonstrated that both the processes and the outcomes of RQ assessment were improved through the use of DED.

8.2 Addressing Research Questions

The research findings and outcomes together addressed the central research question of this thesis, which was:

How to conceptualise and support user-sensitive resource quality assessment for metadata-driven health information portals?

This question was divided into six sub-questions, which were explored at both conceptual and practical levels. This section justifies how the research findings and outcomes address the associated sub-questions in a structure as described below.

First, the design requirements of RQ assessment solutions were articulated, resulting from the exploratory case study on available RQ assessment practice. The following three sub-questions were addressed by investigating two metadata-driven and user-sensitive health information portals:

1. What kinds of domain expertise are required in performing user-sensitive resource quality assessment in the context of metadata-driven health information portals?
2. What tasks and activities are involved in user-sensitive resource quality assessment processes?
3. What are the needs of domain experts with regard to intelligent support?

Research findings are summarised in Section 8.2.1.

Second, a user-sensitive RQ assessment framework was proposed to address the following two sub-questions at the conceptual level:

4. How to define resource quality in the context of health information portals?
5. How can existing information quality assessment theories, principles and approaches be extended and adapted to conceptualise resource quality assessment from a user-sensitive viewpoint?

The proposed user-sensitive RQ assessment framework is summarised in Section 8.2.2.

Finally, the below sub-question was addressed by the conceptualisation and instantiation of an intelligent quality tool.

6. How can intelligent technologies be applied to support user-sensitive resource quality assessment in metadata-driven health information portals?

The proposed user-sensitive RQ assessment framework and the intelligent quality tool constituted a user-sensitive RQ assessment approach. The approach was prototyped via the design and development of the DED as part of a portal CMS. The proposed architecture of the intelligent quality tool and the DED are discussed in Section 8.2.3.

The research findings and outcomes include the articulation of RQ assessment processes, the needs of domain experts for intelligent support, and the design requirements of RQ assessment solutions to support the identified needs. Based on these findings, three design artefacts were constructed, including the user-sensitive RQ assessment framework, the conceptual model of an intelligent quality tool, and the DED prototype system. In addition, the multi-dimensional domain expertise was identified, which was an outcome shared with the Smart Information portals (SIP) project.

8.2.1 Articulation of the Design Requirements

The joint domain expertise study identified the multi-faceted nature of domain expertise. It concluded that the role of domain expert requires (Evans et al. 2009):

- a good understanding of a specific disease domain (both medical and lay experience);
- intimate knowledge of portal users and their information needs;
- awareness of the kinds and styles of information most valued by portal users;
- knowledge of the structure and nature of health information systems; and
- knowledge of information management principles.

That is, domain experts evaluate the perceived usefulness of online information resources to prospective portal users, based on their medical domain knowledge, information management expertise, and lay experience. However, these expertise and skills may not be obtained by individuals, and the level of domain expertise may also vary. In order to support the multi-dimensional domain expertise, a number of domain expert tools were proposed by the joint study. This research drew an independent work on proposing and conceptualising a quality framework together with a quality tool, to work with the other domain expert tools.

In the exploratory case study detailed in Chapter 4, it was identified that the RQ assessment processes involved two connected and entwined processes, of resource selection and resource description. Decisions, such as whether to include a new resource or to publish an included resource, were made based on the combined knowledge, skills and experience of individual domain experts. Although assisted by the portal's resource selection criteria and resource description scheme, domain experts intended to use heuristic techniques and tacit knowledge when making value judgements. Different domain experts had their own decision patterns, search strategies and quality appraisal techniques, which could not be replaced by automation tools. On the other hand, existing content management systems of the two portals examined in the exploratory case study were designed mainly for organising and facilitating data creation and publishing. They did not provide extra functionality for supporting RQ assessment processes. From the domain expert activities and needs analysis detailed in Section 4.3, a semi-automated approach was proposed in order to assess RQ in a systematic, scalable and user-sensitive manner.

8.2.2 The User-Sensitive Resource Quality Assessment Framework

A user-sensitive RQ assessment framework is one of the key design artefacts of this research. As discussed in Chapter 5, it integrated and extended the context-based IQ assessment theory (Strong et al. 1997; Wang and Strong 1996). The framework suggests a systematic and operational approach to assess RQ in the user context. It defines the construct of RQ and how it can be measured by mapping resource attributes to characterised user information needs and quality perceptions (see Figure 5.2). Drawing from an examination of existing metadata models for user-sensitive resource description, a number of resource attributes were proposed for facilitating the RQ measurement.

The proposed user-sensitive RQ assessment framework is intended to provide a robust mechanism for achieving quality-assured portal content management in a systematic way. This quality framework can be used to guide resource evaluation and selection processes for new resources. It can also be used for analysing existing portal data for maintenance purposes. More importantly, such a framework respects expert knowledge of portal users, and articulates the connections between resource attributes and user characteristics in relation to RQ assessment. The framework was used to guide the design and development of a quality tool to work with domain expertise. The value of this framework is that it helps domain experts decide what to look for in an information resource, in order to make value judgements on its overall quality. It also helps domain experts to focus on certain data aspects that indicate one quality aspect. The framework reflects the quality management functionality by identifying and gathering quality indicators, and defining the level of objectivity in RQ assessment.

In this framework, RQ was denoted as a union of two composites: *Reliability* and *Relevancy*. *Reliability* was regarded as the term that best described the quality concerns of health information consumers. An initial set of sub-dimensions was selected to denote the construct of *Reliability* including *Accuracy*, *Credibility* and *Currency*. *Relevancy* was denoted as a composition of *User profile*, *Information reference* and *Subject interest*. The construct of RQ was denoted in a formula provided in Section 5.2.3. This research only defined the measure of *Reliability*.

The measure of the *Reliability* dimension of RQ was defined using an attribute-based data model, which was adapted from Wang et al.'s (1995a) attribute-based data quality management approach. A RQ attribute was defined as a collective term that referred to both subjective quality dimensions and objective quality indicators. Informed by an analysis of quality indicators and their relationships to existing metadata models, the attribute-based data model (see Figure 5.5) was instantiated for the healthcare domain, to facilitate the implementation of the proposed quality framework. The resulting assessment metrics informed the design of a quality tool to support qualitative RQ assessment as described in Chapter 5. In summary, while the subjective quality dimensions are defined by a quality framework, quality indicators of a resource can be collected or measured from the resource's original site, or can be harvested from existing metadata models (where quality indicators have been transformed and preserved). Values of collective quality attributes can then be derived from these quality indicators, enabled by intelligent technologies.

8.2.3 The Intelligent Quality Tool and the Prototype System

As described in Chapter 5, based on the proposed RQ assessment framework and the instantiated assessment metrics, an intelligent quality tool was conceptualised (see Section 5.4). The tool was devised to automatically detect, collect, compute available quality indicators, as well as to suggest values for collective quality attributes based on the transformed and preserved quality indicators. This intelligent quality tool consisted of two sub-systems. One sub-system generated quality indicators, and could be implemented as a web browser extension. The other sub-system aggregated quality indicator values, for describing quality attributes and generating data analysis reports. This sub-system was implementable as a domain expert interface of a portal CMS.

As detailed in Chapter 6, the conceptualised quality tool was prototyped in a real portal context. According to the conceptual model of the quality tool, decision-making processes of RQ assessment can be supported in two ways: to enhance the existing portal content management functionality; and to introduce intelligent features to the domain expert interface of a portal CMS. As a result, the DED prototype system was developed. It implemented the domain expert interface sub-system of the quality tool. The prototype system was designed and developed in the context of the Breast Cancer Knowledge Online (BCKOnline) portal. The portal developed a comprehensive metadata schema for user-centred resource description. The BCKOnline metadata schema included a multi-attributed quality element, which captured domain experts' collective value judgements on RQ. Based on the rich resource descriptions that encoded previous value judgements, ML techniques were applied to predict quality attribute values for new resources. A summary of the enhanced system functions and features for supporting and facilitating RQ assessment was provided in Table 6.6.

8.3 A Method for Developing Domain-Specific and User-Sensitive Resource Quality Assessment Metrics

Reflecting on the learned experience in the healthcare domain, a more generic method is proposed in order to transfer the knowledge to other contexts. Figure 8.1 illustrates how user-sensitive RQ assessment metrics could be developed for other domains.

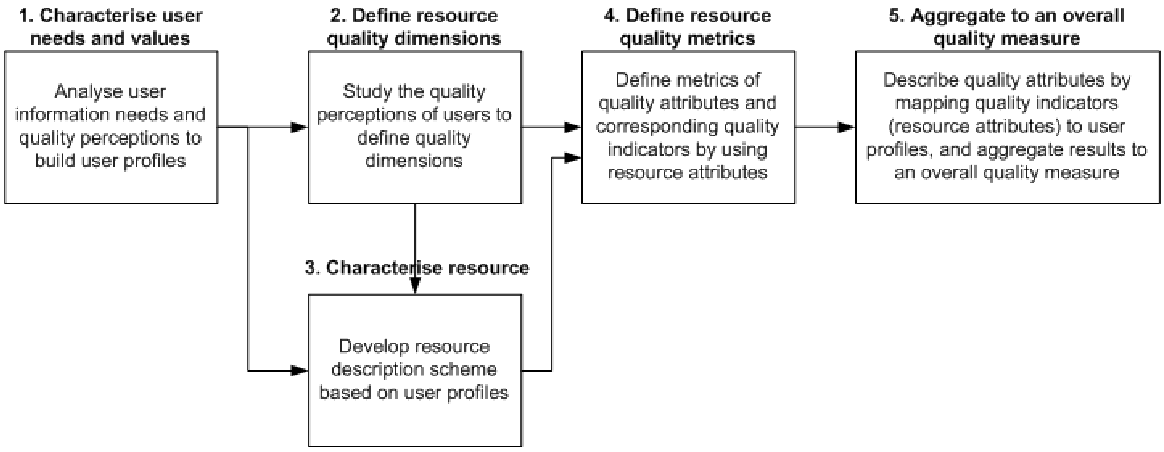


Figure 8.1 A method for developing domain-specific, user-sensitive RQ assessment metrics

The transferable method includes the following procedures for developing domain-specific and user-sensitive RQ assessment metrics:

1. The process first requires the characterisation of the user, particularly their information needs and quality perceptions. That includes modelling the information needs of a targeted user community, and building profiles for different groups of users. The traditional way to collect this type of information is to conduct a user needs study via interviews or focus groups, e.g. studies conducted by Williamson (2005) and Zaphiris et al. (2006). The emergence of web 2.0 technologies, and recent advances in usage analysis, now enable vast user data collections. Therefore, user profiles can be built for individuals via an adaptive and automated approach.
2. The next step is to define which quality dimensions to measure, in response to the identified user perceptions on information quality. User quality perceptions can be analysed via the literature analysis of works, done in a specific domain or via survey studies with the targeted user community.

3. Having user characteristics modelled and quality dimensions defined, the following step is to characterise a resource. A resource description scheme needs to be developed to reflect identified user characteristics. Particularly, attributes that describe the quality aspect of a resource need to be explicitly defined. Resource attributes can be modelled and encoded in a metadata schema for resource description purposes.
4. After resource characteristics are explicitly modelled, quality indicators can then be defined. Quality indicators can include resource attributes defined in Step 3 and some technical or indirect indicators, such as web popularity and Google PageRank (Brin and Page 1998). Quality attributes, defined by the resource description scheme and the selected quality indicators together, constitute the RQ assessment metrics. Based on these metrics, applicable intelligent techniques can be applied in a number of areas, e.g. to facilitate the quality indicator generation, resource characterisation, and quality attributes description.
5. The overall quality measure finally is drawn from a qualitative aggregation of evaluation results, by mapping resource attributes to user characteristics.

8.4 Research Contributions

This study makes theoretical, methodological, and practical contributions through the conceptualisation, construction, and evaluation of a user-sensitive RQ assessment approach for metadata-driven health information portals. The central contribution of this research is to define the concept of user-sensitive RQ assessment and provide a well-structured, theory-based, and empirically tested approach to standardise and support user-sensitive RQ assessment for health information portals.

The adoption of user-sensitive design to develop quality-assured portal content, requires domain experts to evaluate, select, and describe resources in line with user needs and values (McKemmish et al. 2009). However, existing IQ constructs and assessment frameworks are insufficient to address the user sensitivity. Moreover, very little attention was drawn to the challenges domain experts encounter, and the intelligent support they desire, in user-oriented RQ assessment processes.

This research is the first major attempt at conceptualising user-sensitive RQ assessment, in the fields of IQ research and consumer health informatics. It provides theoretical justification and empirical evidence on how the quality of information resources can be systematically assessed as a relationship between a resource and a user. The research develops an operational solution to assess RQ in a user-sensitive manner. Adapted from the context-based IQ assessment theory (Strong et al. 1997; Wang and Strong 1996), the proposed RQ framework integrates intrinsic and contextual IQ dimensions to construct the content quality of online health information resources. More importantly, the framework encodes the consensus understanding of the information needs and quality perceptions of health information consumers. Perceived *Reliability* and *Relevancy* of online resources are qualitatively assessed and aggregated by mapping resource attributes to profiled user information needs and values. By this means, RQ is assessed by domain experts as a relative concept dependent on the circumstances of individual users. Users are respected as the final judge of quality.

According to Purao (2002, p. 19), there are two types of outputs considered as significant contributions for design science research:

- an artefact that represents a symbolic, manipulable representation of concepts and abstractions in the form of operational principles; and
- an artefact that supports or controls the phenomenon of interest.

Figure 8.2 illustrates the research outcomes and contributions to theory, methodology, and practice through the lens of mixed methodologies, which integrate traditional systems development with user-sensitive paradigm in the socio-technical context.

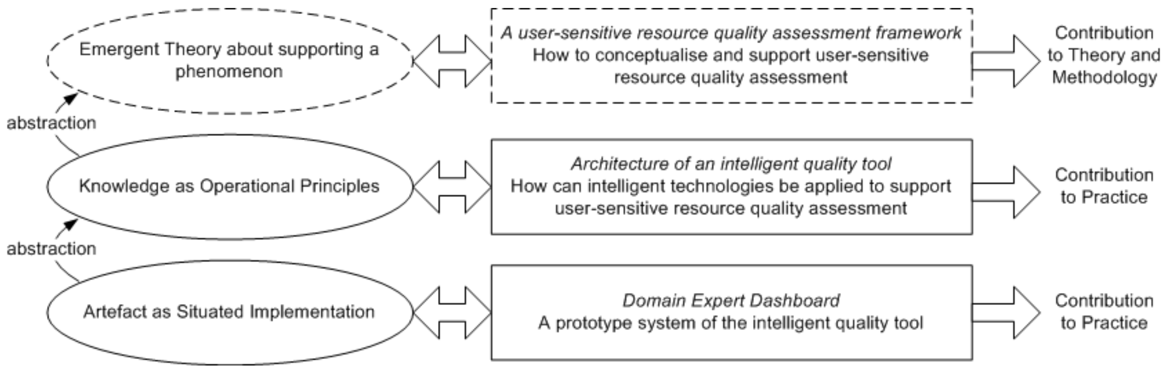


Figure 8.2 Research outcomes and contributions

In Figure 8.2, ovals on the left are adapted from the model suggested by Venable (2006, p. 17). In this model, the dashed oval at the top indicates the emergent theory. The oval in the middle indicates the knowledge generated as operational principles (reproducible knowledge) that the artefact illustrates. The bottom oval indicates the artefact as an instantiation of emergent theory. Based on the abstraction level represented by this model, outcomes of this research, and the contributions they bring to theory, methodology, and practice, are classified accordingly (shown in boxes on right in Figure 8.2). These contributions are further elaborated as follows.

8.4.1 Contributions to Theory and Methodology

This research contributes to the context-based IQ assessment theory (Strong et al. 1997; Wang and Strong 1996) by developing a user-sensitive RQ assessment framework. This study articulates what user-sensitive design means to RQ assessment in the context of metadata-driven health information portals. The design of the quality framework adapted, integrated, and extended existing IQ constructs and definitions, and incorporated the information needs and quality perceptions of health information consumers defined by the consumer health informatics literature.

In the proposed framework, RQ was defined as a composition of *Reliability* and *Relevancy* for the healthcare domain. The framework also provided an explicit construct of the user context. It defined the relationships between resources attributes and characterised user context, in order to standardise the user-sensitive RQ assessment. Perceived RQ was measured as a relative and aggregated construct, by mapping resource attributes to characterised user information needs and quality perceptions. The needs and values of health portal users were characterised based on user needs studies conducted in the consumer health informatics literature. More importantly, the framework captured the expert knowledge of RQ evaluation and description, and defined the level of objectivity in RQ assessment. It also provided an operational guideline for the application of intelligent technologies to semi-automate RQ assessment.

Furthermore, due to the socio-technical nature of the research problem, this study combined the design science research with an interpretivist, user-sensitive design methodology. The combined user-sensitive design science research focused on human (users) rather than machine (technologies) contexts. The major difference was that the adapted approach conceptualised not only the design artefacts, but also the research problem that contributed to related theories. For instance, this research encompassed two kinds of conceptualisation. One was the theory-building relating to the concept of user-sensitive RQ assessment. The other one was the concept-building relating to the intelligent quality tool and its prototype using socio-technical design and evaluation approaches.

In addition, this applied, socio-technical design science research was informed by the user-sensitive SIP architecture. The SIP functional specification was defined at a generic and conceptual level, and was not fully defined to the extent to be prototyped. This research investigated the quality control issue of the SIP approach and prototyped it in the healthcare domain. It contributes to the further refinements of SIP concepts, architecture, and functional requirements.

These contributions advance the knowledge of information quality and its assessment from a user-sensitive viewpoint. They also enrich the understanding of the role intelligent technologies can play in RQ assessment, in terms of supporting informed decision-making, collaboration and learning. In order to transfer the knowledge consolidated from the research experience, this study proposed a more generic method for developing domain-specific and user-sensitive RQ assessment metrics for other domains.

Finally, as is the case for user supporting tools and approaches, having the system work for the purpose of targeted information consumers is a great goal. However, information quality problem solving (Lee 2003) by information consumers identifies a deeper study into unresolved conflict among different disciplines and information roles involved in the entire information value chain. Stakeholders of an information value chain include information collectors, custodians, and consumers. This thesis has provided a starting point for exploring new theoretical findings for using metadata to study information quality problem solving.

8.4.2 Contributions to Practice

The major practical contribution of this research is the description of the architecture of a quality tool. It defines components that encapsulated data management and analysis functionality, as well as the components that represented the application of intelligent techniques. Intelligent learning techniques are employed to support user-sensitive RQ assessment for metadata-driven health information portals. By this means, the study adds new knowledge of expert-oriented quality assessment support, to compliment the consumer-oriented approaches for achieving quality control of online health information.

The intelligent quality tool was designed and developed in order to solve the scalability and sustainability issues of quality-assured portal content management processes. The proposed semi-automated quality assessment solution respects the capacity of domain experts, enables them to make informed value judgements, and also facilities their collaboration and learning.

Moreover, as one of research activities undertaken in the concept building research phase, this study explored available intelligent technologies for their application in providing required functionality. Although this research selected the intelligent learning techniques for prototyping, the application of the proposed solution was not limited to the chosen techniques. The research discussed potentials of using alternative intelligent technologies to solve the problem.

Another practical contribution of this research is the DED prototype system. It provided an implementation of the intelligent quality tool, which demonstrated how intelligent learning techniques could be applied to support the decision-making processes of domain experts on RQ. It had been proved that by using the prototype system, both the RQ assessment processes and outcomes were improved. It is also believed that the system can help domain experts to develop required expertise and skills more efficiently and effectively by using enhanced system functions and intelligent features. The developed system has been used as the domain expert interface of a real health information portal (the BCKOnline portal).

Finally, the results obtained from the information users using the dashboard can facilitate the improvement of the quality of health information from the source organisations. By doing this, the entire cycle of generating, producing, storing, maintaining and using the health information can be further improved, particularly the 'reliability' aspect of the health information resource that has been explored in this study.

8.5 Adaptation of the Design Science Research within a Socio-technical context

This section presents an adaptation of Hevner et al.'s (2004, p. 83) general guidelines for conducting, evaluating, and presenting design science research. Within a socio-technical design science framework, the research employed mixed methodologies to approach research questions. How this research addresses Hevner et al.'s guidelines through the lens of socio-technical design paradigm is discussed below.

8.5.1 Guideline One: Design as an Artefact

The first guideline concerns the production of design artefacts, which can be constructs, models, methods, or instantiations. This research aimed at developing socio-technical solutions that capture and facilitate human knowledge. The primary design artefact was the conceptualisation of a semi-automated approach, using intelligent technologies to support user-sensitive RQ assessment for health information portals. The construct of RQ, and how it could be measured in the context of perceived user information needs and quality perceptions, were explicitly defined by a user-sensitive RQ assessment framework developed for the healthcare domain. As part of the solution, an intelligent quality tool was also conceptualised to implement the proposed framework. In addition to these conceptual products, an instantiation of the proposed quality tool was designed, developed, and evaluated as a significant secondary artefact of this research.

8.5.2 Guideline Two: Problem Relevance

The purpose of the design science research is to "develop technology-based solutions to important and relevant business problems" (Hevner et al. 2004, p. 83). Instead of developing purely technology-based solutions, this research provided a socio-technical solution to conceptualise and support RQ assessment. Previous research demonstrated the effects of a user-sensitive and metadata-driven approach in developing health information portals to better meet the diverse information needs of portal users (Burstein et al. 2005; Burstein et al. 2006; Fisher et al. 2002). This research was concerned with the scalability issue of laborious RQ assessment processes for this specific type of portal. A literature analysis of existing IQ assessment approaches in multiple disciplines (see Chapter 2) showed that existing IQ constructs, frameworks, assessment methods and techniques did not adequately address the diverse information needs and quality perceptions of health information consumers. Socio-technical solutions are imperative to support user-sensitive RQ assessment for metadata-driven health information portals.

The proposed user-sensitive RQ assessment solution demonstrated how intelligent technologies could be applied to improve both the processes and outcomes of RQ assessment in a portal context. The research problem owners, also the target audiences of this thesis, are mainly portal developers, domain experts, and user communities. They play different roles in the design, development, and maintenance of a portal artefact, the underlining data structure, and portal included content. The proposed user-sensitive RQ assessment solution interacts with all these aspects.

In addition, the conceptual model of the proposed intelligent quality tool presents an operational architecture. Portal designers/developers can implement and adapt the conceptual tool architecture based on their own portal settings. The developed prototype system (the DED) also provides an exemplar of quality tool implementation. It employed ML techniques to address complex technical issues, but portal designers/developers may want to use alternative intelligent technologies for problem solving. No matter what kinds of technologies are utilised to implement the defined system functionality, it is believed that the system can increase the quality appraisal capacity of domain experts. Their decision-making efficiency and effectiveness on RQ can be improved accordingly. The BCKOnline portal has integrated the DED as the domain expert interface of its CMS. It is expected that the enhanced portal CMS will significantly improve the scalability of user-sensitive RQ assessment in practice

8.5.3 Guideline Three: Design Evaluation

The third guideline concerns the rigour of artefact evaluation, via the selection and execution of evaluation methods. According to Pries-Heje et al. (2008), a design artefact can be evaluated via an 'ex ante' approach prior to its construction, or via an 'ex post' approach after the artefact is created. They propose a strategic framework to analyse the design process and design artefacts of design science research by questioning "*when* evaluation takes place, *what* is actually evaluated, and *how* it is evaluated" (p. 260). In order to validate the proposed user-sensitive RQ assessment framework and the intelligent quality tool, both the design process and the design artefacts were evaluated using various approaches.

Evaluation of design process

Guba and Lincoln (1994, p.114) propose a set of criteria for judging interpretivist research validity, the 'goodness' criteria for both qualitative and quantitative research (Lincoln and Guba 2000). These criteria were applied to evaluate the design process of this research, as discussed below.

- Credibility (paralleling internal validity): concerned with the credibility of research outcomes, i.e. the degree to which findings make sense.
- Transferability (paralleling external validity): concerned with generalizability, i.e. the applicability of the findings to other settings.
- Dependability (paralleling reliability): concerned with the adjustment of the findings to the dynamic settings of the reality.
- Confirmability (paralleling objectivity): concerned with the findings being confirmed or corroborated by others.

Methods used to ensure the quality of design process are summarised in Table 8.1.

Table 8.1 Summary of methods to ensure the quality of design process

Criterion	Methods to ensure the quality of design process
Credibility	Member checks: the research outcomes were evaluated by both the researcher and the domain expert participants (see Chapter 7)
Transferability	Description of the research setting, i.e. the context of metadata-driven health information portals (see Chapter 1) Suggestion on how the research findings and outcomes may be transferred to other domains (see Section 8.3 in this chapter)
Dependability	Comprehensive description of research methods, data collection methods and instruments, and decisions about the research to justify what was done and why (see Chapter 3)
Confirmability	Self-critical reflective analysis of the socio-technical design science research

	paradigm, the user-sensitive design methodology, and the systems development research methodology (see Chapter 3)

It is argued that the validity and rigour of the proposed user-sensitive RQ assessment approach (the design artefacts) are firstly ensured by the quality of research itself. 'Ex post' evaluation of the prototype system further demonstrates the utility, efficacy, and quality of the proposed approach.

Evaluation of design artefacts

It is widely recognised that the evaluation method must be well-selected and well-executed in order to demonstrate the utility, quality, and efficacy of the design artefact (Basili 1996; Hevner et al. 2004; Kleindorfer et al. 1998; Zelkowitz and Wallace 1998). According to the taxonomy provided by Hevner et al. (2004, p. 86), a design artefact can be evaluated via the following five types of methods:

- observational methods, e.g. case study and field study;
- analytical methods, e.g. static analysis, architecture analysis, optimization, and dynamic analysis;
- experimental methods, e.g. controlled experiment and simulation;
- testing methods, e.g. functional (black box) testing, and structural (white box) testing; and
- descriptive methods, e.g. informed argument and scenarios.

The design artefacts of this research, as was discussed previously when addressing the first guideline, encompass a user-sensitive RQ assessment framework, the conceptual architecture and instantiation of an intelligent quality tool that implements the framework. While this research delivered both the conceptualisation and instantiation design artefacts, the ultimate goal of evaluation was to assess the validity, utility and efficacy of the overall user-sensitive RQ assessment approach.

This goal was achieved by employing mixed methods to perform 'ex post' evaluation on the prototype system. In the system evaluation research phase, a functional (black box) test was carried out to assess the accuracy of the implemented intelligent feature via a series of ML experiments. Further, a usefulness and usability study (an observational evaluation method) was conducted with domain experts, to empirically evaluate the utility and efficacy of the prototype system on RQ assessment processes and outcomes.

8.5.4 Guideline Four: Research Contributions

The fourth guideline defines the research contributions of design science to provide "clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies" (Hevner et al. 2004, p. 83). Design science research needs to demonstrate one or more of the three types of research contributions based on "the novelty, generality, and significance of the design artifact" (p. 87). As this research uses a socio-technical design science approach, only transferability, or limited generalizability, is achievable in the interpretivist framework.

This research makes significant theoretical, methodological, and practical contributions via the design, construction, and evaluation of the following design artefacts:

- a user-sensitive RQ assessment framework for health information portals; and
- the conceptual model and the instantiation of an intelligent quality tool that implemented the proposed quality framework.

The novelty, significance, and transferability of these design artefacts are demonstrated in the following areas:

Novelty: The proposed concept of user-sensitive RQ assessment has not been addressed in the relevant research literature. This research adopted the user-sensitive design and context-based IQ assessment theory to define the proposed concept for the healthcare domain. Specifically, the study explored innovative ways to standardise and facilitate RQ assessment that took account of user information needs and quality perceptions.

Significance: Labour-intensity of RQ assessment is a bottleneck for scalable and sustainable content management in metadata-driven health information portals. Existing IQ assessment approaches are not sufficient to address the needs of constructing and assessing RQ in a user-sensitive manner. Thus, it is imperative and significant for this research to understand user-sensitive RQ assessment, and develop new solutions to standardise and support the human processes in practice.

Transferability: Although this research was conducted and evaluated in the specific setting of metadata-driven health information portals, the applicability, and transferability of the proposed solution was one of the key concerns of this research. Reflected on the entire socio-technical design science research process, a method was proposed for developing user-sensitive RQ assessment metrics for other domains. It needs to clarify that within the canonical design science research framework, generalizability is usually a concern.

8.5.5 Guideline Five: Research Rigour

In the fifth guideline, Hevner et al. (2004) define the rigour of the design science research to justify "the application of rigorous methods in both the construction and evaluation of the design artefact" (p. 83). Based on this guideline, Arnott and Pervan (2008) propose an operational measure of research rigour using the following two constructs:

- the rigour of the theoretical foundations by considering the effective use of appropriate reference theory; and
- the rigour of the research methodology by considering the design and the application of appropriate research methods.

The above two constructs are used to justify the rigour of this research:

Theoretical foundation: In this research, the context-based IQ assessment theory (Strong et al. 1997; Wang and Strong 1996) and user-sensitive design methodology (McKemmish et al. 2009) were justified as the foundation theories, to construct the user-sensitive RQ assessment approach for health information portals. Informed by the user-sensitive design principles, this research adapted the IQ assessment theory by building new concepts to address user-sensitivity. The formulation and construction of the proposed approach also followed Wang et al.'s (1995b) methodology for design and development of data quality management solutions.

Research methodology: This research employed mixed research methodologies to approach the central research question. The combined user-sensitive systems development research methodology governed the concept building, system building, and system evaluation research phases. The adoption and implementation of the mixed methodologies assured the rigour of the socio-technical design science research process, as well as the validity of the research outcomes. The resulting user-sensitive RQ assessment framework and the corresponding intelligent quality tool were in line with existing quality assessment approaches proposed in the relevant research fields. These artefacts followed the design requirements drawn from the exploratory case study of the available practices. Further, the usefulness and usability of the prototype system in supporting RQ assessment were tested through the use of combined evaluation methods.

8.5.6 Guideline Six: Design as a Search Process

Hevner et al.'s (2004) sixth guideline concerns the effective artefact design through an iterative search process. It is defined as "utilizing available means to reach desired ends while satisfying laws in the problem environment" (p. 83).

The major objective of this research was to investigate and define the concepts of RQ and user-sensitive RQ assessment, based on existing IQ concepts and quality assessment approaches. The research outputs, from the study's conceptualisation stage, informed later research activities undertaken in the iterative system building and system evaluation stages. These activities involved the design and development of a quality tool prototype by using selected intelligent techniques. In this case, ML techniques were used. Although the prototype system only implemented and tested one sub-system defined by the conceptual model of the quality tool, the results demonstrated adequate evidence to refine the quality tool architecture, as well as the overall RQ assessment solution. Within the limited timeframe allowed by a PhD study, an iteration of solution conceptualisation, system building, and system evaluation was executed, delivering satisfying outputs to address research questions.

8.5.7 Guideline Seven: Communication of Research

The seventh and final guideline requires the effective presentation of this socio-technical design science research to research-oriented, technology-oriented, and user-oriented audience. The targeted audiences of this research, apart from academics, are portal designers/developers, domain experts, and user communities, who are the problem owners of the user-sensitive RQ assessment issues investigated in this research. Portal designers/developers refer to those who are responsible for solving technological issues associated with the design, development, and maintenance of web portal system. Domain experts, as defined in Chapter 1, are those professionals who have medical domain expertise, lay user experience, and information management skills. They are responsible for developing and managing quality-assured portal content. User communities refer to health portal end-users, who have difficulty in finding reliable and relevant online health information. This thesis is also of interest and value to researchers in the fields of consumer health informatics and information quality.

Communications of research can happen during and after the researching process. On one hand, this research interacted actively with portal practitioners, the domain experts in particular, in different stages of the socio-technical design science research process. Both the exploratory case study and the system evaluation involved their participation. On the other hand, as the key research outcome, this research proposed a user-sensitive RQ assessment approach. The conceptualisation of this socio-technical solution was benefited on discussions with researchers, portal designers/developers, domain experts, and portal end-users, who are quality control stakeholders (Eysenbach 2005a) for online health information. Useful feedback was received when presenting this research at conferences and workshops in multiple disciplines, including information quality, information systems, and health informatics.

8.6 Future Research

Future works based on this research will include several studies that have emerged from the concept building, system building, and system evaluation research phases. In addition, this research also opens a door to several potential fields for future investigation using new technologies, such as semantic web, collective intelligence, and web 2.0. Directions for future research are outlined as follows.

8.6.1 Measurement Study

- Define the measure of the *Relevancy* dimension of RQ

This research defined the construct of RQ as a composition of *Reliability* and *Relevancy*. The overall RQ is aggregated from the qualitative measurement of *Reliability* and *Relevancy*. Both dimensions are compounds of sub-dimensions. The hierarchy of RQ dimensions was illustrated in Figure 5.1. As clarified in Section 5.3, considering the amount of research work a PhD study could carry out, only the measure of the *Reliability* dimension was defined and instantiated for the healthcare domain. Future research will investigate the measure of *Relevancy* in relation to perceived user information needs. There is a similar concept of *Relevance* defined in the field of Information Retrieval. Computational models or algorithms are proposed to measure *Relevance* based on queries of individual users (Mizzaro 1997; Saracevic 2007a). This is in contrast to the qualitative approach this research adopted in assessing *Reliability*. Therefore, how to define the qualitative measurement of *Relevancy* based on perceived user information needs is of interest for future research.

- Ontology-based RQ representation

In this research, the measure of the RQ *Reliability* dimension was represented by a relational model, which associated quality attributes to corresponding quality indicators. In the implementation of the intelligent quality tool, all the quality attributes selected for describing a resource and the available quality indicators were preserved in one flat file. However, such a flat structure may not be suitable for representing complex quality data models. It is not capable of capturing different relationships between quality attributes and their associated quality indicators. Semantic metadata, which are represented in ontologies, provide a solution to describe the complex construct of RQ (Eysenbach 2005b). Therefore, it would be interesting to explore the design requirements of a user-sensitive RQ assessment approach that uses ontology-based RQ construct and measure.

8.6.2 Collaboration for Problem-solving: Machines and Information Consumers

- Evaluate the refined prototype system with domain experts

Completing the above tasks will lead to a refined DED with enhanced intelligent features to support user-sensitive RQ assessment. In order to evaluate the utility and efficacy of the refined system and its impact on decision-making outcomes and processes, a usefulness and usability study needs to be conducted in a real portal setting. Involving a larger group of domain expert participates is considered, to evaluated the refined system.

- Sustain the quality-assured portal content management with both domain experts and end-users

Quality-assured portal content development and maintenance rely on multi-dimensional domain expertise, which may not be fully obtained by individuals. Collaborative work of multiple domain experts is required in many cases. For instance, a health information portal may use the same resource description scheme to deliver online information from multiple disease domains. Nevertheless, employing multiple domain experts to achieve quality assurance of portal content is still an expensive approach.

On the other hand, it is noticed that the required domain expertise may also be found in end-users, whose value judgements on retrieved information might also be valuable to other users. Their involvement in portal content management should be encouraged and supported by modern portal systems. Recent advances in web 2.0 technologies now enable interactive online information rating, tagging, and sharing. How these emerging technologies can be applied to enable portal end-users participating in resource recommendation and description is one interesting area for future research. Sustaining portals with both domain experts and end-users is a potentially feasible solution for achieving quality-assured online information provision.

- How information consumers collaborate with machines?

In the information quality area, the work of Lee et al. (2002) shows how collaboration between system (data quality software) and human (data quality manager) can be combined to resolve conflicting underlying principles and to solve information quality problems. This work shows how the results of data integrity test from using data quality software (integrity analyser), a kind of machine analysis finding and the principles held by accounting and marketing managers, were overruled and reconciled by a data quality manager who understands the comprehensive nature of the specific data. This reconciliation was further theorised as context-reflective problem solving (Lee 2003). The theory explains how salient but hidden metadata, the contexts of data (e.g. time, place, role, principles), and goals are explicated when facing conflicting principles in different expert domains (e.g. database management rules, marketing principles, accounting rules). Information-related decisions need to reflect on how data and metadata can provide value for information users' individual goals.

8.6.3 Enhanced Techniques for Design Improvements

- Implement the quality indicator generator

The proposed conceptual model of the intelligent quality tool encompassed two sub-systems (see Section 5.4.4). Constrained by the available time and resources, the quality tool was partially prototyped to demonstrate the concept of semi-automated RQ assessment. Based on the consideration of the extensive work done in the detection and construction of quality indicators in the relevant literature, the quality indicator generator was not prototyped in this research. However, a number of issues merit further investigation, such as how to use, preserve, and manage quality indicators in the quality indicator generator, and how the tool interacts with the metadata authoring tool.

Moreover, in this research the assessment metrics developed for measuring the *Reliability* dimension of RQ only utilised the parts of quality indicators that were captured by available metadata models. Based on the comprehensive metadata descriptions manually generated by domain experts, intelligent learning techniques were applied to provide value suggestions for describing quality attributes. It is necessary to investigate alternative approaches that can automatically describe quality attributes based on a wider range of quality indicators. It may also of interest to explore the potential of the quality indicator generator in constructing resource metadata descriptions.

- Refine the design of the quality attribute generator

As was mentioned in Section 7.1.4, removing the potentially irrelevant data attributes from the data set that was used for predicting *Publisher credentials* (one of the five selected quality attributes for prototyping) increased the predicting accuracy. Similar cases may be found with the other quality attributes. The attribute analysis tool provided by WEKA could be used to determine which data attributes (or metadata elements) are dominant factors in determining the classification model for a particular quality attribute. It is also interesting to explore the relationships amongst quality attributes. Moreover, according to the evaluation results of the implemented metadata value suggestion feature (see Section 7.2.4), it is suggested that the quality tool can pre-populate definitive values for describing certain quality attributes. For instance, the tool can pre-populate values for describing *Publisher credentials*, instead of providing suggestive guidance to domain experts. It is envisaged that traffic light symbols can be applied to indicate the predicting accuracy. Thus algorithms for generating value suggestions need to be fine-tuned in order to provide definitive values.

- Apply ML techniques in metadata authoring

It is envisaged that the ML techniques utilised for describing RQ attributes can be applied to generate value suggestions for describing other types of resource attributes. Besides, metadata values assigned by domain exerts are implicitly derived from varied kinds of quality indicator values. Therefore, intelligent techniques for detecting and processing quality indicators can also be applied to support metadata generation.

8.7 Chapter Summary

This chapter reflected on the overall socio-technical design science research process, through which a semi-automated and user-sensitive RQ assessment approach was conceptualised. The chapter summarised key research findings and outcomes, and discussed how they addressed research questions in a healthcare domain. Based on the consolidated research experience, a method transferable to other domains was proposed for developing domain-specific and user-sensitive RQ assessment metrics. The chapter highlighted the significant contributions this thesis brought to the context-based information quality assessment theory, the design science research framework, the user-sensitive SIP approach, as well as the quality control practice for online health information. The adaptation of Hevner et al.'s (2004) seven design science research guidelines to this research was discussed. The quality of this socio-technical design science research, in terms of its novelty, validity, significance, and rigour, was justified. Finally, a number of directions for future research were identified in order to further refine the proposed user-sensitive RQ assessment approach at the theoretical, technical, and practical levels.

REFERENCES

Adams S., Bont A. d. and Berg M. (2006), Looking for answers, constructing reliability: An exploration into how Dutch patients check web-based medical information, *International Journal of Medical Informatics*, vol. 75, no. 1, pp. 66-72.

Addey D., Ellis J., Suh P. and Thiemecke D. (2002), *Content Management Systems*, Glasshaus Ltd.

AGLS (2010), AGLS Metadata Standard, from http://www.agls.gov.au/.

Aguillo I. (2000), A new generation of tools for search, recovery and quality evaluation of World Wide Web medical resources, *Online Information Review*, vol. 24, no. 2, pp. 138-143.

Akbani R., Kwek S. and Japkowicz N. (2004), Applying support vector machines to imbalanced datasets, in proceedings of the 15th European Conference on Machine Learning, Pisa, Italy, 20-24 September, pp. 39-50.

Aladwani A. M. and Palvia P. C. (2002), Developing and validating an instrument for measuring user-perceived web quality, *Information and Management*, vol. 39, no. 6, pp. 467-476.

Alexander J. E. and Tate M. A. (1999), *Web Wisdom: How to Evaluate and Create Information Quality on the Web*, Lawrence Erlbaum Associates, Mahwah, NJ: Erlbaum.

Anderson J., McKemmish S. and Manaszewicz R. (2003), Quality criteria models used to evaluate health websites, in proceedings of the 10th Asia Pacific Special Health and Law Librarians Conference, Adelaide, Australia, 24-27 August, 2003, pp. 337-354.

Anderson P. F., Allee N., Chung J. and Westra B. (1999), The megasite project: A metasite comparing health information megasites & search engines, viewed 6 August, 2008, from http://www-personal.umich.edu/~pfa/megasite/.

Arnott D. and Pervan G. (2008), An assessment of DSS design science using the Hevner, March, Park, and Ram guidelines, in Hart, D. N. and Gregor, S. D. (eds.) *Information Systems Foundations: The Role of Design Science*, Australian National University, Canberra ACT, pp. 255-284.

Arthur M. (2006), *Expanding A Digital Content Management System*, Elsevier.

Asprey L. and Middleton M. (2003), *Integrative Document and Content Management: Strategies for Exploiting Enterprise Knowledge*, Idea Group Publishing.

Avison D. E., Lau F., Myers M. D. and Nielsen P. A. (1999), Action research, *Communications of the ACM*, vol. 42, no. 1, pp. 94-97.

Baeza-Yates R. and Ribeiro-Neto B. (1999), *Modern Information Retrieval*, ACM Press; Addison-Wesley, New York; Harlow, England.

Bagozzi R. P., Davis F. D. and Warshaw P. R. (1992), Development and test of a theory of technological learning and usage, *Human Relations*, vol. 45, no. 7, pp. 659-686.

Bailey B. P., Gurak L. J. and Konstan J. A. (2001), An examination of trust production in computer-mediated exchange, in proceedings of the 7th Conference on Human Factors and the Web.

Barnes S. J. and Vidgen R. T. (2000), WebQual: An exploration of Web-site quality, in proceedings of the 8th European Conference on Information Systems, Vienna, 3-5 July, 2000.

Barnes S. J. and Vidgen R. T. (2001), An evaluation of Cyber-Bookshops: The WebQual method, *International Journal of Electronic Commerce*, vol. 6, no. Number 1 / Fall 2001, pp. 11-30.

Barnes S. J. and Vidgen R. T. (2002), An integrative approach to the assessment of E-Commerce quality, *Journal of Electronic Commerce Research*, vol. 3, no. 3, pp. 114-127.

Barros R., Xexéo G., Pinheiro W. A. and de Souza J. (2008), A web metadata based-model for information quality prediction, in Calero, C., Moraga, M. A. and Piattini, M. (eds.) *Handbook of Research on Web Information Systems Quality*, IGI Global Information Science Publishing.

Basili V. R. (1996), The role of experimentation in software engineering: Past, current, and future, in proceedings of the 18th International Conference on Software Engineering, Berline, Germany, 25-29 March, 1996, IEEE Computer Society, pp. 442-449.

Baskerville R., Pries-Heje J. and Venable J. (2009), Soft design science methodology, in proceedings of the 4th international conference on design science research in information systems and technology, New York, USA.

Baskerville R. L. and Wood-Harper A. T. (1996), A critical perspective on action research as a method for information systems research, *Journal of Information Technology*, vol. 11, no. 3, pp. 235-246.

Batini C. (2006), *Data Quality: Concepts, Methodologies and Techniques*, Springer Berlin Heidelberg, Berlin.

Batini C., Cabitza F., Cappiello C., Francalanci C. and di Milano P. (2006), A comprehensive data quality methodology for web and structured data, in proceedings of the 1st International Conference on Digital Information Management, Bangalore, India, 6-8 December, 2006, pp. 448-456.

BCKOnline (2009), Breast Cancer Knowledge Online Project, viewed 2 September, 2010, from http://www.infotech.monash.edu.au/research/centres/cosi/projects/bcko/.

Benbasat I., Goldstein D. K. and Mead M. (1987), The case research strategy in studies of information systems, *MIS Quarterly*, vol. 11, no. 3, pp. 369-386.

Benbasat I. and Zmud R. W. (1999), Empirical research in information systems: the practice of relevance, *MIS Quarterly*, vol. 23, no. 1, pp. 3-16.

Benigeri M. and Pluye P. (2003), Shortcomings of health information on the Internet, *Health Promotion International*, vol. 18, no. 4, pp. 381-386.

Berners-Lee T., Hendler J. and Lassila O. (2001), The semantic web, *Scientific American*, vol. 284, no. 5, pp. 34-43.

Bertelsen O. W. (2000), Design artefacts: Towards a design-oriented epistemology, *Scandinavian Journal of Information Systems*, vol. 12, pp. 15-27.

Berti-Équille L. (2007), measuring and modelling data quality for quality-awareness in data mining, in Guillet, F. and Hamilton, H. J. (eds.) *Quality Measures in Data Mining*, vol. 43, Springer Berlin/Heidelberg, pp. 101-126.

Bethard S., Wetzler P., Butcher K., Martin J. H. and Sumner T. (2009), Automatically assessing resource quality for educational digital libraries, in proceedings of WWW2009 Madrid, Spain, 20-24 April 2009.

Bevan N. (1999), Quality in use: Meeting user needs for quality, *Journal of Systems and Software*, vol. 49, no. 1, pp. 89-96.

Boese E. S. and Howe A. E. (2005), Effects of Web Document Evolution on Genre Classification, in proceedings of CIKM'05, Bremen, Germany, 31 October - 5 November, 2005, pp. 632-639.

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.

Bomba D. and Land T. (2004), Constructing and validating a consumer health portal rating index, paper presented to World Medical Informatics Conference, 2004.

Bostrom R. P. and Heinen J. S. (1977a), MIS problems and failures: A socio-technical perspective. Part I: The causes, *MIS Quarterly*, vol. 1, no. 3, pp. 17-32.

Bostrom R. P. and Heinen J. S. (1977b), MIS problems and failures: A socio-technical perspective. Part II: The application of socio-technical theory, vol. 1, no. 4, pp. 11-28.

Bovee M., Srivastava R. P. and Mak B. (2003), A conceptual framework and belief function approach to assessing overall information quality, *International Journal of Intelligent Systems*, vol. 18, no. 1, pp. 51-74.

Bray T., Paoli J., Sperberg-McQueen C. M., Maler E., Yergeau F. and Cowan J. (2006), Extensible markup language (XML) 1.1, *W3C Recommendation, W3C.*

Brender J. (2006), *Handbook of Evaluation Methods for Health Informatics*, Elsevier Academic Press, Burlington, MA.

Brin S. and Page L. (1998), The anatomy of a large-scale hypertextual Web search engine, in proceedings of the 7th International World Wide Web conference, pp. 107-117.

Browning P. and Lowndes M. (2001), JISC TechWatch report: Content management systems, TSW01-02, Bristol, JISC Executive. Retrieved 6 August, 2008, from http://www.jisc.ac.uk/uploaded_documents/tsw_01-02.pdf.

Burstein F. (2002), System development in information systems research, in Williamson, K. (ed.) *Research Methods for Students, Academics and Professionals: Information Management and Systems*, 2nd, Centre for Information Studies, Charles Sturt University, Wagga Wagga, New South Wales, Australia, pp. 147-158.

Burstein F., Fisher J., McKemmish S., Manaszewicz R. and Malhotra P. (2005), User centred quality health information provision: Benefits and challenges, in proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS'05), Hawaii, USA, 3-6 January.

Burstein F. and Gregor S. (1999), The systems development or engineering approach to research in information systems: An action research perspective, in proceedings of the 10th Australasian Conference on Information Systems, Wellington, New Zealand, 1-3 December 1999.

Burstein F., Mckemmish S., Fisher J., Manaszewicz R. and Malhotra P. (2006), 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-Verlag, London, UK, pp. 359-383.

Calero C., Caro A. and Piattini M. (2008), An applicable data quality model for web portal data consumers, *World Wide Web*, vol. 11, no. 4, pp. 465-484.

Candan K. S., Liu H. and Suvarna R. (2001), Resource description framework: Metadata and its applications, *ACM SIGKDD Explorations Newsletter*, vol. 3, no. 1, pp. 6-19.

Candela L., Castelli D., Ferro N., Ioannidis Y., Koutrika G., Meghini C., Pagano P., Ross S., Soergel D., Agosti M., Dobreva M., Katifori V. and Schuldt H. (2007), The DELOS digital library reference model - Foundations for digital libraries, from http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf.

Cappiello C., Francalanci C., Missier P., Pernici B., Plebani P., Scannapieco M. and Virgillito A. (2003a), DL2: Presentation of metadata and of the quality certificate, from http://www.dis.uniroma1.it/~dq/docs/Deliverables/DL2.pdf.

Cappiello C., Francalanci C. and Pernici B. (2003b), Preserving web sites: A data quality approach, in proceedings of the 8th International Conference on Information Quality (ICIQ-03), Cambridge, Massachusetts, USA, 7-9 November, 2003, MIT, pp. 331-343.

Cappiello C., Francalanci C. and Pernici B. (2004), Data quality assessment from the user's perspective, in proceedings of the 2004 international workshop on Information quality in information systems, Paris, France, ACM Press.

Cardinaels K., Meire M. and Duval E. (2005), Automating metadata generation: The simple indexing interface, in proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, 10-14 May, 2005, ACM Press, New York, USA, pp. 548-556.

Caro A., Calero C., Caballero I. and Piattini M. (2006), Defining a data quality model for web portals, in *Web Information Systems – WISE 2006*, vol. 4255/2006, Springer Berlin / Heidelberg, pp. 363-374.

Chen G., Warren J. and Evans J. (2008), Automatically generated consumer health metadata using semantic spaces, in proceedings of the 2nd Australasian workshop on Health data and knowledge management - Volume 80, Wollongong, NSW, Australia, Australian Computer Society, Inc.

Chowdhury G. G. (1999a), The Internet and information retrieval research: A brief review, *Journal of Documentation*, vol. 55, no. 2, pp. 209-225.

Chowdhury G. G. (1999b), Template mining for information extraction from digital documents, *Library Trends*, vol. 48, no. 1, pp. 47–62.

Chowdhury G. G. (2004), *Introduction to Modern Information Retrieval*, 2nd edition, Facet Publishing, London.

CIHI (2009), The CIHI data quality framework, from http://www.cihi.ca/CIHI-ext-portal/pdf/internet/DATA_QUALITY_FRAMEWORK_2009_EN.

Civan A. and Pratt W. (2006), Supporting consumers by characterizing the quality of online health information: A multidimensional framework, in proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06), IEEE Computer Society, pp. 88a.

Cline R. J. W. and Haynes K. M. (2001), Consumer health information seeking on the Internet: the state of the art, *Health Education Research*, vol. 16, no. 6, pp. 671-692.

Cole C. (2011), A theory of information need for information retrieval that connects information to knowledge, *Journal of the American Society for Information Science and Technology*, vol. 62, no. 7, pp. 1216-1231.

Cole R., Purao S., Rossi M. and Sein M. (2005), Being proactive: Where action research meets design research, in proceedings of ICIS, pp. Paper 27.

Collins B. (1997), Infofilter project, viewed 17 June, 2008, from http://www.usc.edu/users/help/flick/Infofilter/.

Conrad J. G., Leidner J. L. and Schilder F. (2008), Professional credibility: Authority on the web, in proceedings of the 2nd ACM workshop on Information credibility on the web, Napa Valley, California, USA, ACM.

Cooke A. (2001), *A Guide to Finding Quality Information on the Internet: Selection and Evaluation Strategies*, 2nd edition, Library Association, London.

Cooke A., McNab A. and Anagnostelis B. (1996), The good, the bad and the ugly: Internet review sites, in proceedings of the 20th International Online Information Meeting, Oxford, England, United Kingdom, 3-5 December, 1996, pp. 33-40.

Corcho O. (2006), Ontology based document annotation: Trends and open research problems, *International Journal of Metadata, Semantics and Ontologies* vol. 1, no. 1, pp. 47-57.

Crouch M. and McKenzie H. (2006), The logic of small samples in interview-based qualitative research, *Social Science Information*, vol. 45, no. 4, pp. 483-499.

Cui L. (1999), Rating health web sites using the principles of citation analysis: A bibliometric approach, *Journal of Medical Internet Research*, vol. 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.

Custard M. and Sumner T. (2005), Using machine learning to support quality judgments, *D-Lib Magazine*, vol. 11, no. 10.

Davis F. D. (1989), Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS Quarterly*, vol. 13, no. 3, pp. 319-340.

Day M. (2001), Metadata in a nutshell, *Information Europe*, vol. 6, no. 2, pp. 11.

DCMI (2011), Dublin Core Metadata Initiative (DCMI), from http://dublincore.org/.

Dedeke A. (2000), A conceptual framework for developing quality measures for information systems, in proceedings of the 5th International Conference on Information Quality, pp. 126-128.

DeLone W. H. and McLean E. R. (1992), Information systems success: The quest for the dependent variable, *Information Systems Research*, vol. 3, no. 1, pp. 60-95.

Diaz J. A., Griffith R. A., Ng J. J., Reinert S. E., Friedmann P. D. and Moulton A. W. (2002), Patients' use of the Internet for medical information, *Journal of General Internal Medicine*, vol. 17, no. 3, pp. 180-185.

DIF (2010), Directory Interchange Format (DIF) Writer's Guide, from http://gcmd.gsfc.nasa.gov/User/difguide/DIF_Guide_2010.pdf.

Ding Y., Chowdhury G. G. and Foo S. (1999), Template mining for the extraction of citation from digital documents, in proceedings of the 2nd Asian Digital Library Conference, Taiwan, pp. 47–62.

DISCERN (2008), The DISCERN Instrument, viewed 7 April, 2008, from http://www.discern.org.uk/discern_instrument.php.

Duval E., Hodgins W., Sutton S. and Weibel S. L. (2002), Metadata principles and practicalities, *D-Lib Magazine*, vol. 8, no. 4.

Eboueya M. and Uden L. (2007), Benefits and limitations of portals, in Tatnall, A. (ed.) *Encyclopedia of Portal Technologies and Applications*, Information Science Reference, pp. 75-81.

English L. P. (2002), Total Quality data Management (TQdM), in Piattini, M. G., Calero, C. and Genero, M. (eds.) *Information and Database Quality*, Kluwer Academic Publishers, Boston/Dordrecht/London, pp. 85-110.

Eppler M. and Muenzenmayer P. (2002a), Measuring information quality in the web context: A survey of state-of-the-art instruments and an application methodology, in proceedings of the 7th International Conference on Information Quality, pp. 187-196.

Eppler M. and Muenzenmayer P. (2002b), Measuring information quality in the web context: a survey of state-of-the-art instruments and an application methodology, in proceedings of the 7th International Conference on Information Quality, pp. 187–196.

Eppler M. J. (2006), *Managing information quality: Increasing the value of information in knowledge-intensive products and processes*, 2nd edition, Springer, Berlin; New York.

Eppler M. J., Algesheimer R. and Dimpfel M. (2003), Quality criteria of content-driven websites and their influence on customer satisfaction and loyalty: An empirical test of an information quality framework, in proceedings of the MIT Conference on Information Quality, pp. 108–120.

Eppler M. J. and Wittig D. (2000), Conceptualizing information quality: A review of information quality frameworks from the last ten years, in proceedings of IQ 2000 Conference on Information Quality, Boston, USA, 22-23 October, 2000, pp. 83-96.

Evangelista P. F., Embrechts M. J. and Szymanski B. K. (2006), Taming the curse of dimensionality in kernels and novelty detection, in Abraham, A., de Baets, B., Köppen, M.

and Nickolay, B. (eds.) *Applied Soft Computing Technologies: The Challenge of Complexity*, Springer Berlin/Heidelberg, pp. 425-438.

Evans J., Manaszewicz R. and Xie J. (2009), The role of domain expertise in smart, user-sensitive, health information portals, in proceedings of the 42nd Hawaii International Conference on System Sciences (HICSS-42), Hawaii, USA, 5-8 January, 2009.

Evans J. R. and Lindsay W. M. (2005), *The Management and Control of Quality*, 6th edition, South-Western, Mason, Ohio.

Eysenbach G. (2000), Recent advances: Consumer health informatics, *British Medical Journal (BMJ)*, vol. 320, no. 7251, pp. 1713-1716.

Eysenbach G. (2001), An ontology of quality initiatives and a model for decentralized, collaborative quality management on the Semantic World Wide Web, *Journal of Medical Internet Research*, vol. 3, no. 4: e34.

Eysenbach G. (2002), Infodemiology: The epidemiology of (mis)information, *The American Journal of Medicine*, vol. 113, no. 9, pp. 763-765.

Eysenbach G. (2005a), Design and evaluation of consumer health information web sites, in Lewis, D. (ed.) *Consumer Health Informatics: Informing Consumers and Improving Health Care*, Springer, New York, pp. 34-60.

Eysenbach G. (2005b), MedCERTAIN/MedCIRCLE: Using semantic web technologies for quality management of health information on the web, in Lewis, D. (ed.) *Consumer Health Informatics: Informing Consumers and Improving Health Care*, Springer, New York, pp. 217-226.

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 (BMJ)*, vol. 317, no. 7171, pp. 1496-1502.

Eysenbach G. and Köhler C. (2002), How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews, *British Medical Journal (BMJ)*, vol. 324, no. 7337, pp. 573–577.

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, *Journal of American Medical Association (JAMA)*, vol. 287, no. 20, pp. 2691-2700.

Fisher J. (2009), Usability testing instruments, viewed 16 November, 2009, from http://www.sims.monash.edu.au/staff/jfisher/Usability%20instruments.htm.

Fisher J., Bentley J., Turner R. and Craig A. (2004), A usability instrument for evaluating websites - navigation elements, paper presented to OZCHI, Woollongong, University of Woollongong.

Fisher J., Burstein F., Lynch K. and Lazarenko K. (2008), "Usability + usefulness = trust": An exploratory study of Australian health web sites, *Internet Research*, vol. 18, no. 5, pp. 477-498.

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, in proceedings of the 13th Australasian Conference on Information Systems, Melbourne, Australia.

Forgionne G. A. (1999), An AHP model of DSS effectiveness, *European Journal of Information Systems*, vol. 8, no. 2, pp. 95-106.

Frické M., Fallis D., Jones M. and Luszko G. M. (2005), Consumer health information on the Internet about carpal tunnel syndrome: Indicators of accuracy, *The American Journal of Medicine*, vol. 118, no. 2, pp. 168-174.

Friedman B. (1996), Value-sensitive design, *Interactions*, vol. 3, no. 6, pp. 16-23.

Friedman B., Kahn P. H. and Borning A. (2006), Value sensitive design and information systems, in Galletta, D. and Zhang, P. (eds.) *Human-Computer Interaction in Management Information Systems: Applications*, pp. 348-372.

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 (BMJ)*, vol. 324, no. 7337, pp. 569-573.

Gaizauskas R. and Wilks Y. (1998), Information extraction: Beyond document retrieval, *Journal of Documentation*, vol. 54, no. 1, pp. 70-105.

Gertz M., O¨ zsu T., Saake G. and Sattler K.-U. (2004), Report on the Dagstuhl seminar "data quality on the web", *SIGMOD Record*, vol. 33, no. 1, pp. 127-132.

GILS (2000), Global Information Locator Service (GILS), from http://www.gils.net/.

Google (2011), Technology overview, viewed 18 June, 2011, from http://www.google.com/about/corporate/company/tech.html.

Gravano L., Chang C.-C. K., García-Molina H. and Paepcke A. (1997), STARTS: Stanford proposal for Internet meta-searching, in proceedings of the 1997 ACM International Conference on Management of Data (SIGMOD), Tucson, Arizona, USA, ACM, pp. 207-218.

Greenberg J. (2004), Metadata extraction and harvesting: A comparison of two automatic metadata generation applications, *Journal of Internet Cataloging*, vol. 6, no. 4, pp. 59-82.

Greenberg J., Spurgin K. and Crystal A. (2006), Functionalities for automatic metadata generation applications: A survey of metadata experts' opinions, *International Journal of Metadata, Semantics and Ontologies* vol. 1, no. 1 pp. 3-20.

Greenes R. A. and Shortliffe E. H. (1990), Medical informatics: An emerging academic discipline and institutional priority, *Journal of American Medical Association (JAMA)*, vol. 263, no. 8, pp. 1114-1120.

Griffiths K. M. and Christensen H. (2005), Website quality indicators for consumers, *Journal of Medical Internet Research*, vol. 7, no. 5, pp. e55.

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.

Guba E. G. and Lincoln Y. S. (1994), Competing paradigms in qualitative research, in *Handbook of qualitative research*, Sage Publications, Inc., pp. 105-117.

Hall A. and Walton G. (2004), Information overload within the health care system: A literature review, *Health Information and Libraries Journal*, vol. 21, pp. 102-108.

Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P. and Witten I. H. (2009), The WEKA data mining software: An update, *SIGKDD Explorations*, vol. 11, no. 1, pp. 10-18.

Hatala M. and Richards G. (2003), Value-added metatagging: Ontology and rule based methods for smarter metadata, in proceedings of Rules and Rule Markup Languages for the Semantic Web: Second International Workshop, Sanibel Island, FL, USA, 20 October, 2003, Springer, pp. 65-80.

Haynes D. (2004), *Metadata for Information Management and Retrieval*, Facet publishing.

Henninger M. (2008), *The hidden web: finding quality information on the net*, 2nd edition, University of New South Wales Press, Sydney, N.S.W.

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.

Herrera-Viedma E., Peis E., Olvera M., Herrera J. and Montero Y. (2003), Evaluating the informative quality of web sites by fuzzy computing with words, in Menasalvas, E., Segovia, J. and Szczepaniak, P. (eds.) *Advances in Web Intelligence*, vol. 2663, Springer Berlin/Heidelberg, pp. 954-954.

Hersh W. R. (2002), Medical informatics: Improving health care through information, *Journal of American Medical Association (JAMA)*, vol. 288, no. 16, pp. 1955-1958.

Hevner A. and Chatterjee S. (2010), *Design Research in Information Systems: Theory and Practice*, Springer, New York Dordrecht Heidelberg London.

Hevner A. R. (2007), A three cycle view of design science research, *Scandinavian Journal of Information Systems*, vol. 19, no. 2, pp. Article 4.

Hevner A. R., March S. T., Park J. and Ram S. (2004), Design science in information systems research, *MIS Quarterly*, vol. 28, no. 1, pp. 75-105.

HHOnline (2007), Evaluation of Heart Health Online: A program to assist GPs to manage depression and related illnesses in patients with coronary heart disease, viewed 2 Septermber, 2010, from http://www.beyondblue.org.au/index.aspx?link_id=6.955.

HIDDEL (2003), HIDDEL vocabulary, viewed 28 October, 2008, from http://www.medcircle.org/.

Hinkin T. R. (1995), A review of scale development practices in the study of organizations, *Journal of Management*, vol. 21, no. 5, pp. 967-988.

HONcode (2007), Health on the Net code of conduct, viewed 7 April, 2007, from https://www.hon.ch/.

Hopkins R. L. (1995), Countering information overload: The role of the librarian, *The Reference Librarian*, vol. 23, no. 49 & 50, pp. 305-333.

Hsu C.-N. and Dung M.-T. (1998), Generating finite-state transducers for semi-structured data extraction from the Web, *Information Systems*, vol. 23, no. 8, pp. 521-538.

Hudgins J., Agnew G. and Brown E. (1999), *Getting mileage out of metadata: Applications for the library*, American Library Association, Chicago.

Iivari J. and Venable J. (2009), Action research and design science research - seemingly similar but decisively dissimilar, in proceedings of the 17th European conference on information systems (ECIS), Verona, Italy, LSE Information Systems and Innovation Group, pp. 2711-2723.

Intute (2007), Intute: Health and Life Sciences - Evaluation guidelines, from http://www.intute.ac.uk/healthandlifesciences/IntuteHLS_Evaluation_Guidelines.doc.

ISO/IEC (1998), 9241-11 Ergonomic requirements for office work with visual display terminals (VDT)s - Part 11 Guidance on usability, from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=16883.

Jadad A. R. and Gagliardi A. (1998), Rating health information on the Internet: Navigating to knowledge or to Babel? , *Journal of American Medical Association (JAMA)*, vol. 279, no. 8, pp. 611-614.

Järvine P. (2007), Action research is similar to design science, *Quality and Quantity*, vol. 41, no. 1, pp. 37-54.

Jeffery K. (2000), Metadata: The future of information systems, in Brinkkemper, S., Lindencrona, E. and Sølvberg, A. (eds.) *Information Systems Engineering: State of the Art and Research Themes*, London: Springer-Verlag, paper presented at the Information Systems Engineering Symposium, Stockholm, Sweden.

John B. E. and Marks S. J. (1997), Tracking the effectiveness of usability evaluation methods, *Behaviour & Information Technology*, vol. 16, no. 4, pp. 188 - 202.

Kahn B. K., Strong D. M. and Wang R. Y. (2002), Information quality benchmarks: Product and service performance, *Communications of the ACM*, vol. 45, no. 4, pp. 184-192.

Katerattanakul P. and Siau K. (1999), Measuring information quality of web sites: Development of an instrument, in proceedings of the 20th international conference on Information Systems, Charlotte, North Carolina, United States, pp. 279-285.

Katerattanakul P. and Siau K. (2002), Information quality in Internet commerce design, in Piattini, M. G., Calero, C. and Genero, M. (eds.) *Information and Database Quality*, Kluwer Academic Publishers, Boston/Dordrecht/London, pp. 45-56.

Katz W. A. (1997), *Introduction to Reference Work*, 7th edition, New York: McGraw-Hill.

Kc M. W.-T. (2009), *Building a Prototype for Quality Information Retrieval from the World Wide Web*, PhD Thesis, University of Wollongong, Wollongong, New South Wales, Australia.

Keast G., Toms E. G. and Cherry J. (2001), Measuring the reputation of web sites: A preliminary exploration, in proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries, Roanoke, Virginia, United States, 24-28 June, 2001, ACM, pp. 77-78.

Keen P. G. W. (1987), MIS research: Current status, trends and needs, in Buckingham, R. A., Hirschheim, R. A., Land, F. F. and Tully, C. J. (eds.) *Information Systems Education: Recommendations and Implementation*, Cambridge University Press, Cambridge, England, pp. 1-13.

Kim P., Eng T. R., Deering M. J. and Maxfield A. (1999), Published criteria for evaluating health related Web sites: Review, *British Medical Journal (BMJ)*, vol. 318, no. 7184, pp. 647-649.

Klein B. D. (2001), User perceptions of data quality: Internet and traditional text sources, *Journal of Computer Information Systems*, vol. 41, no. 4, pp. 9.

Klein B. D. (2002), When do users detect information quality problems on the World Wide Web, in proceedings of American Conference on Information Systems, pp. 1101-1103.

Klein H. K. and Myers M. D. (1999), A set of principles for conducting and evaluating interpretive field studies in information systems, *MIS Quarterly*, vol. 23, no. 1, pp. 67-94.

Kleindorfer G. B., O'Neill L. and Ganeshan R. (1998), Validation in simulation: Various positions in the philosophy of science, *Management Science*, vol. 44, no. 8, pp. 1087-1099.

Knight S.-a. (2009), *User Perceptions of Information Quality in World Wide Web Information Retrieval Behaviour*, PhD Thesis, School of Management Information Systems, Faculty of Business and Law, Edith Cowen University, Perth, Western Australia, Australia.

Knight S.-a. and Burn J. (2005), Developing a framework for assessing information quality on the World Wide Web, *Informing Science Journal*, vol. 8, pp. 159-172.

Koch T. (2000), Quality-controlled subject gateways: Definitions, typologies, empirical overview, *Online Information Review*, vol. 24, no. 1, pp. 24-34.

Kreinovich V., Nguyen H. T., Prasad N. S. and Santiprabhob P. (2004), Intelligent technologies: An introduction, *International Journal of Intelligent Systems*, vol. 19, no. 1-2, pp. 1-8.

Kubat M. and Matwin S. (1997), Addressing the curse of imbalanced training sets: one-sided selection, in proceedings of the 14th International Conference on Machine Learning, Morgan Kaufmann, pp. 179-186.

Kuechler B. and Vaishnavi V. (2008), On theory development in design science research: Anatomy of a research project, *European Journal of Information Systems*, vol. 17, pp. 489-504.

Lancaster F. W. (1979), *Information Retrieval Systems: Characteristics, Testing, and Evaluation*, 2nd edition, John Wiley & Sons, New York.

Lankes R. D. (2008), Credibility on the internet: Shifting from authority to reliability, *Journal of Documentation*, vol. 64, no. 5, pp. 667-686.

Lee M. D., Pincombe B. and Welsh M. (2005), An empirical evaluation of models of text document similarity, in proceedings of the 27th Annual Conference of the Cognitive Science Society, Citeseer, pp. 1254–1259.

Lee Y. W. (2003), Crafting rules: Context-reflective data quality problem solving, *Journal of Management Information Systems*, vol. 20, no. 3, pp. 93-119.

Lee Y. W., Pipino L., Strong D. M. and Wang R. Y. (2004), Process-embedded data integrity, *Journal of Database Management*, vol. 15, no. 1, pp. 87-103.

Lee Y. W., Pipino L. L., Funk J. D. and Wang R. Y. (2006), Journey to Data Quality, pp. 226, from.

Lee Y. W., Strong D. M., Kahn B. K. and Wang R. Y. (2002), AIMQ: A methodology for information quality assessment, *Information and Management*, vol. 40, no. 2, pp. 133-146.

Leung H. K. N. (2001), Quality metrics for intranet applications, *Information & Management*, vol. 38, no. 3, pp. 137-152.

Lincoln Y. S. and Guba E. G. (2000), Paradigmatic controversies, contradictions, and emerging confluences, in Denzin, N. K. and Lincoln, Y. S. (eds.) *Handbook of Qualitative Research*, 2nd, Sage Publications, Inc., pp. 163-188.

Liu J. (2007), *Metadata and its applications in the digital library: Approaches and practices*, Libraries Unlimited, Westport, Conn.

Liu Z. and Huang X. (2005), Evaluating the credibility of scholarly information on the web: A cross cultural study, *The International Information & Library Review*, vol. 37, no. 2, pp. 99-106.

Livari J. and Venable J. R. (2009), Action research and design science research - Seemingly similar but decisively dissimilar, in proceedings of ECIS, pp. Paper 73.

Long J. A. and Seko C. E. (2002), A new method for database data quality evaluation at the Canadian Institute for Health Information, in proceedings of the 7th International Conference on Information Quality, Boston.

Madnick S. E., Wang R. Y., Lee Y. W. and Zhu H. (2009), Overview and framework for data and information quality research, *Journal of Data and Information Quality*, vol. 1, no. 1, pp. 1-22.

Maheu M. M., Whitten P. and Allen A. (2001), *E-health, telehealth, and telemedicine: A guide to start-up and success*, Jossey-Bass Inc Pub, San Francisco, CA.

Malhotra P., Burstein F., Fisher J., McKemmish S., Anderson J. and Manaszewicz R. (2003), Breast cancer knowledge online portal: An intelligent decision support system perspective, in proceedings of 14th Australasian Conference on Information Systems, Perth, Western Australia.

Manaszewicz R., Williamson K. and McKemmish S. (2002), Breast Cancer Knowledge Online: Towards meeting the diverse information needs of the breast cancer community, in proceedings of the Electronic Networking Building Community Conference, pp. 3-7.

Mani I. (2006), Summarization of text, automatic, in Brown, K. (ed.) *Encyclopedia of Language & Linguistics*, 2nd, Elsevier, Oxford, U.K., pp. 274-291.

March S. T. and Smith G. F. (1995), Design and natural science research on information technology, *Decision Support Systems*, vol. 15, no. 4, pp. 251-266.

March S. T. and Storey V. C. (2008), Design science in the information systems discipline: An introduction to the special issue on design science research, *MIS Quarterly*, vol. 32, no. 4, pp. 725-730.

Marshall L. A. and Williams D. (2006), Health information: Does quality count for the consumer? How consumers evaluate the quality of health information materials across a variety of media, *Journal of Librarianship and Information Science*, vol. 38, no. 3, pp. 141-156.

Marshall M. N. (1996), Sampling for qualitative research, *Family Practice*, vol. 13, no. 6, pp. 522-526.

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.

Mason M. (2010), Sample size and Saturation in PhD studies using qualitative interviews, *Forum Qualitative Social Research*, vol. 11, no. 3.

McKay J. and Marshall P. (2001), The dual imperatives of action research, *Information Technology & People*, vol. 14, no. 1, pp. 46-59.

McKemmish S., Manaszewicz R., Burstein F. and Fisher J. (2009), Consumer empowerment through metadata-based quality reporting: The breast cancer knowledge online portal, *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 60, no. 9, pp. 1792-1807.

McKemmish S., Manaszewicz R. and Cheah C. (2004), BCKOnline Metadata Schema Version 1.0, from http://www.sims.monash.edu.au/research/eirg/BCKO_MetadataSchema_Version16.doc .

MedCIRCLE, MedCIRCLE Collaboration (2002), Collaboration for Internet rating, certification, labeling and evaluation of health information, viewed 12 November, 2007, from http://www.medcircle.org/metadata/index.php.

Merrill J. (1999), The Internet and Classical Civilization, *The Acquisitions Librarian*, vol. 12, no. 23, pp. 97-115.

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 (JASIST)*, vol. 58, no. 13, pp. 2078-2091.

Michnik J. and Lo M.-C. (2009), The assessment of the information quality with the aid of multiple criteria analysis, *European Journal of Operational Research*, vol. 195, no. 3, pp. 850-856.

Mizzaro S. (1997), Relevance: The whole history, *Journal of the American Society for Information Science*, vol. 48, no. 9, pp. 810-832.

Moraga Á., Caro A., Calero C. and Piattini M. (2007), Portal quality issues, in Tatnall, A. (ed.) *Encyclopedia of Portal Technologies and Applications*, Information Science Reference, pp. 747-754.

Moura A. M. d. C., Campos M. L. M. and Barreto C. M. (1998), A survey on metadata for describing and retrieving Internet resources, *World Wide Web*, vol. 1, no. 4, pp. 221-240.

Moura A. M. d. C., Pereira G. d. C. and Campos M. L. M. (2002), A metadata approach to manage and organize electronic documents and collections on the web, *Journal of the Brazilian Computer Society*, vol. 8, pp. 16-31.

Naumann F. (2002), *Quality-Driven Query Answering for Integrated Information Systems*, Springer.

Naumann F., Freytag J. C. and Spiliopoulou M. (1998), Quality-driven source selection using data envelopment analysis, in proceedings of the International Conference on Information Quality (IQ '98), MIT, Cambridge, USA.

Naumann F. and Rolker C. (1999), Do metadata model meet IQ requirements?, in proceedings of the International Conference on Information Quality (IQ '99), Cambridge, MA, USA, pp. 99-114.

Naumann F. and Rolker C. (2000), Assessment methods for information quality criteria in proceedings of the 5th International Conference on Information Quality, Cambridge, MA, USA, 20-22 October, 2000, pp. 148-162.

Norman D. A. and Draper S. W. (1986), *User centered system design: New perspectives on Human-Computer Interaction*, L. Erlbaum Associates Inc., Hillsdale, NJ, USA.

Nunamaker J. F., Chen M. and Purdin T. D. M. (1991), Systems development in information systems research, *Journal of Management Information Systems*, vol. 7, no. 3, pp. 89-106.

O'Reilly T. and Battelle J. (2009), Web squared: Web 2.0 five years on, in proceedings of the 6th Annual Web 2.0 Summit, San Francisco, CA, USA, 20-22 October, 2009.

Özel S. A., Altingövde I. S., Ulusoy Ö., Özsoyoğlu G. and Özsoyoğlu Z. M. (2004), Metadata-based modeling of information resources on the Web, *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 55, no. 2, pp. 97-110.

Passin T. B. (2004), *Explorer's Guide to the Semantic Web*, edition, Manning Publications Co.

Paynter G. W. (2005), Developing practical automatic metadata assignment and evaluation tools for internet resources, in proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries, Denver, CO, USA, ACM.

Peffers K., Tuunanen T., Rothenberger M. A. and Chatterjee S. (2007), A design science research methodology for information systems research, *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45-77.

Phillips-Wren G., Mora T M., Forgionne G. A., Garrido L. and Gupta J. N. D. (2006), A multicriteria model for the evaluation of intelligent decision-making support systems (i-DMSS), in Gupta, J. N. D., Forgionne, G. A. and Mora T, M. (eds.) *Intelligent Decision-making Support Systems*, Springer London, pp. 3-24.

PICS (2008), Platform for Internet Content Selection (PICS), viewed 25 June, 2008, from http://www.w3.org/PICS/.

Pipino L. L., Lee Y. W. and Wang R. Y. (2002), Data quality assessment, *Commun. ACM*, vol. 45, no. 4, pp. 211-218.

Porter M. E. (1991), Towards a dynamic theory of strategy, *Strategic Management Journal*, vol. 12, no. S2, pp. 95-117.

Preece J., Rogers Y. and Sharp H. (2002), *Interaction Design: Beyond Human-Computer Interaction*, 2nd edition, John Wiley & Sons, Inc, New York, USA.

Price R. and Shanks G. (2004), A semiotic information quality framework, in proceedings of the 2004 IFIP International Conference on Decision Support Systems (DSS2004), Prato, Italy, 1-3 July, pp. 658-672.

Price R. and Shanks G. (2005a), Empirical refinement of a semiotic information quality framework, in proceedings of the 38th Annual Hawaii International Conference on

System Sciences (HICSS38), Hawaii, USA, 3-6 Janurary, 2005, IEEE Computer Society, pp. 1-10.

Price R. and Shanks G. (2005b), A semiotic information quality framework: Development and comparative analysis, *Journal of Information Technology*, vol. 20, pp. 88-102.

Price S. L. and Hersh W. R. (1999), Filtering Web pages for quality indicators: An empirical approach to finding high quality consumer health information on the World Wide Web, in proceedings of AMIA 1999 Annual Symposium, Washington DC, pp. 911-915.

Pries-Heje J., Baskerville R. L. and Venable J. (2008), Strategies for design science research evaluation, in proceedings of the 16th European conference on information systems (ECIS), Galway, Ireland, pp. 255-266.

Purao S. (2002), Design research in the technology of information systems: Truth or dare, Working paper, from http://iris.nyit.edu/~kkhoo/Spring2008/Topics/DS/000DesignSc_TechISResearch-2002.pdf.

Ragetli J. (2005), Methods and tools for managing library web content, in Yu, H. (ed.) *Content and Workflow Management for Library Web Sites: Case Studies*, Information Science Publishing Hershey, USA, pp. 22-49.

Redman T. C. (1992), *Data Quality: Management and Technology*, Bantam Books, Inc., New York.

Reese G. (2000), Distributed application architecture, in Oram, A. (ed.) *Database Programming with JDBC and JAVA*, 2nd, O'Reilly & Associates, Inc.

Rieh S. Y. (2002), Judgment of information quality and cognitive authority in the web, *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 53, no. 2, pp. 145-161.

Risk A. and Dzenowagis J. (2001), Review of internet health information quality initiatives, *Journal of Medical Internet Research*, vol. 3 no. 4, pp. e28.

Rose D. E. and Levinson D. (2004), Understanding user goals in web search, paper presented to the 13th international conference on World Wide Web, New York, NY, USA.

Saracevic T. (2007a), Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance, *Journal of the American Society for Information Science and Technology*, vol. 58, no. 13, pp. 1915-1933.

Saracevic T. (2007b), Relevance: A review of the literature and a framework for thinking on the notion in information science. Part III: Behavior and effects of relevance, *Journal of the American Society for Information Science and Technology*, vol. 58, no. 13, pp. 2126-2144.

Scuse D. and Reutemann P. (2008), WEKA Experimenter Tutorial for Version 3-5-8, from http://aarnet.dl.sourceforge.net/project/weka/documentation/3.5.x/ExperimenterTutorial-3-5-8.pdf.

Seidman C. (2001), *Data mining with Microsoft SQL Server 2000 technical reference*, Microsoft Press Redmond, WA, USA.

Shankar G. and Watts S. (2003), A relevant, believable approach for data quality assessment, in proceedings of the 8th International Conference on Information Quality, pp. 178–189.

Shanks G. and Corbitt B. (1999), Understanding data quality: Social and cultural aspects, in proceedings of the 10th Australasian Conference on Information Systems, pp. 785.

Shanks G. and Price R. (2008), Data quality and decision making, in Burstein, F. and Holsapple, C. W. (eds.) *Handbook on Decision Support Systems*, Springer, Berlin, pp. 65-82.

Silberg W. M., Lundberg G. D. and Musacchio R. A. (1997), Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor - let the reader and viewer beware, *Journal of American Medical Association (JAMA)*, vol. 277, no. 15, pp. 1244-1245.

Simon H. A. (1996), *The Sciences of the Artificial*, 3rd edition, MIT Press, Cambridge, Massachusetts.

SIP (2006), Smart Information Portals: Meeting knowledge and decision support needs of health care consumers for quality online information, from http://www.sip.infotech.monash.edu.au/.

Sokvitne L. (2000), Understanding metadata, viewed 2 July, 2008, from http://service.tas.gov.au/papers/understanding_metadata.asp.

SortSite (2008), SortSite datasheet - One click site checking, compliance, and search engine optimization, Electrum smarter solutions, http://www.powermapper.com/

Srivastava J., Cooley R., Deshpande M. and Tan P.-N. (2000), Web usage mining: Discovery and applications of usage patterns from Web data, *SIGKDD Explor. Newsl.*, vol. 1, no. 2, pp. 12-23.

Strong D. M., Lee Y. W. and Wang R. Y. (1997), Data quality in context, *Communications of the ACM*, vol. 40, no. 5, pp. 103-110.

Stvilia B. and Gasser L. (2008), An activity theoretic model for information quality change, *First Monday*, vol. 13, no. 4.

Stvilia B., Gasser L., Twidale M. B. and Smith L. C. (2007), A framework for information quality assessment, *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 58, no. 12, pp. 1720-1733.

Stvilia B., Mon L. and Yi Y. J. (2009), A model for online consumer health information quality, *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 60, no. 9, pp. 1781-1791.

Sumner T. and Marlino M. (2004), Digital libraries and educational practice: A case for new models, in proceedings of the 4th ACM/IEEE Joint Conference on Digital Libraries (JCDL 2004) Tucson, Arizona, USA, ACM Press, pp. 170-178.

Tate M. and Alexander J. (1996), Teaching critical evaluation skills for World Wide Web resources, *Computers in Libraries*, vol. 16, no. 10.

Tatnall A. (2005), Portals, portals everywhere, in Tatnall, A. (ed.) *Web Portals: The New Gateways to Internet Information and Services*, Idea Group Publishing, Hershey, PA, pp. 1-14.

Tayi G. K. and Ballou D. P. (1998), Examining data quality, *Communications of the ACM*, vol. 41, no. 2, pp. 54-57.

Taylor R. S. (1986), *Value Added Processes in Information Systems*, Greenwood Publishing Group Inc.

Venable J. R. (2006), The role of theory and theorising in design science research, in proceedings of DESRIST 2006, Claremont, CA, Claremont Graduate University, pp. 1-18.

Walls J. G., Widmeyer G. R. and El Sawy O. A. (1992), Building an information system design theory for vigilant EIS, *Information Systems Research*, vol. 3, no. 1, pp. 36-59.

Wand Y. and Wang R. Y. (1996), Anchoring data quality dimensions in ontological foundations, *Communications of the ACM*, vol. 39, no. 11, pp. 86-95.

Wang R. Y., Reddy M. P. and Kon H. B. (1995a), Toward quality data: An attribute-based approach, *Decision Support Systems*, vol. 13, no. 3-4, pp. 349-372.

Wang R. Y., Storey V. C. and Firth C. P. (1995b), A framework for analysis of data quality research, *IEEE Transactions on Knowledge and Data Engineering*, vol. 7, no. 4, pp. 623-640.

Wang R. Y. and Strong D. M. (1996), Beyond accuracy: What data quality means to data consumers, *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5-33.

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.

Warner A. J. (1987), Natural language processing, *Annual review of information science and technology*, vol. 22, pp. 79-108.

Weber R. (1987), Toward a theory of artifacts: A paradigmatic base for information systems research, *Journal of Information Systems*, vol. 1, no. 2, pp. 3-17.

Wickham J. (2006), BIOME history, viewed 30 July, 2007, from http://www.intute.ac.uk/healthandlifesciences/history.html.

Williamson K. (2002), *Research methods for students, academics and professionals: Information management and systems*, 2nd edition, Centre for Information Studies, Wagga Wagga, NSW.

Williamson K. (2005), Where one size does not fit all: Understanding the needs of potential users of a portal to breast cancer knowledge online, *Journal of Health Communication*, vol. 10, no. 6, pp. 567-580.

Williamson K. and Manaszewicz R. (2003), Breast cancer information needs and seeking: Towards an intelligent, user sensitive portal to breast cancer knowledge online, *The New Review of Information Behaviour Research*, vol. 3, pp. 203-219.

Wilson P. (2002), How to find the good and avoid the bad or ugly: A short guide to tools for rating quality of health information on the internet, *British Medical Journal (BMJ)*, vol. 324, no. 7337, pp. 598-602.

Wilson T. D. (1981), On user studies and information needs, *Journal of Documentation*, vol. 37, no. 1, pp. 3-15.

Wilson T. D. (1994), Information needs and uses: Fifty years of progress, in Vickery, B. C. (ed.) *Fifty Years of Information Progress: A Journal of Documentation Review*, Aslib, pp. 15-51.

Witten I. H. and Frank E. (2005), *Data Mining: Practical Machine Learning Tools and Technologies*, 2nd edition, Elsevier.

Xie J. and Burstein F. (2011), Using machine learning to support resource quality assessment: An adaptive attribute-based approach for health information portals, in Xu, J., Yu, G., Zhou, S. and Unland, R. (eds.) *Database Systems for Advanced Applications*, vol. 6637, Springer Berlin/Heidelberg, paper presented at the 4th International Workshop on Data Quality in Integration System (DQIS'2011), Hong Kong, China, pp. 526-537.

Xie J., Burstein F. and Evans J. (2008), Intelligent support for resource quality evaluation and description in health information portals, in proceedings of the national conference for the Health Informatics Society of Australia (HIC'08), Melbourne, Australia, 31 Aug.-2 September 2008, pp. Chapter 41.

Ye S. and Chua T.-S. (2006), Learning object models from semistructured web documents, *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 3, pp. 334-349.

Ying S. and Al-Hakim L. (2010), Assuring information quality in medical platform for U-Healthcare service, in proceedings of 2010 IEEE International Conference on Wireless Communications, Networking and Information Security (WCNIS), , Beijing, China, 25-27 June, 2010, pp. 664-668.

Z39.50 (1995), Attribute Set BIB-1 (Z39.50-1995): Semantics, Z39.50 Implementers Group and Z39.50 Maintenance Agency, from http://www.loc.gov/z3950/agency/bib1.html.

Zaphiris P., Dellaporta A. and Mohamedally D. (2006), User needs analysis and evaluation of portals, in Cox, A. (ed.) *Portals: People, Processes and Technology*, Facet Publishing, London, pp. 52-62.

Zeist R. H. J. and Hendriks P. R. H. (1996), Specifying software quality with the extended ISO model, *Software Quality Journal*, vol. 5, no. 4, pp. 273-284.

Zelkowitz M. V. and Wallace D. R. (1998), Experimental models for validating technology, *IEEE Computer*, vol. 31, no. 5, pp. 23-31.

Zhu J. and Gauch S. (2000), Incorporating quality metrics in centralized/distributed information retrieval on the World Wide Web, in proceedings of the 23rd annual international ACM SIGIR conference on research and development in information retrieval, Athens, Greece, ACM Press, pp. 288-295.

Zhu Y. and Buchmann A. (2002), Evaluating and selecting web sources as external information resources of a data warehouse, in proceedings of the 3rd International Conference on Web Information Systems Engineering (WISE 2002), 12-14 December, 2002 pp. 149 - 160.

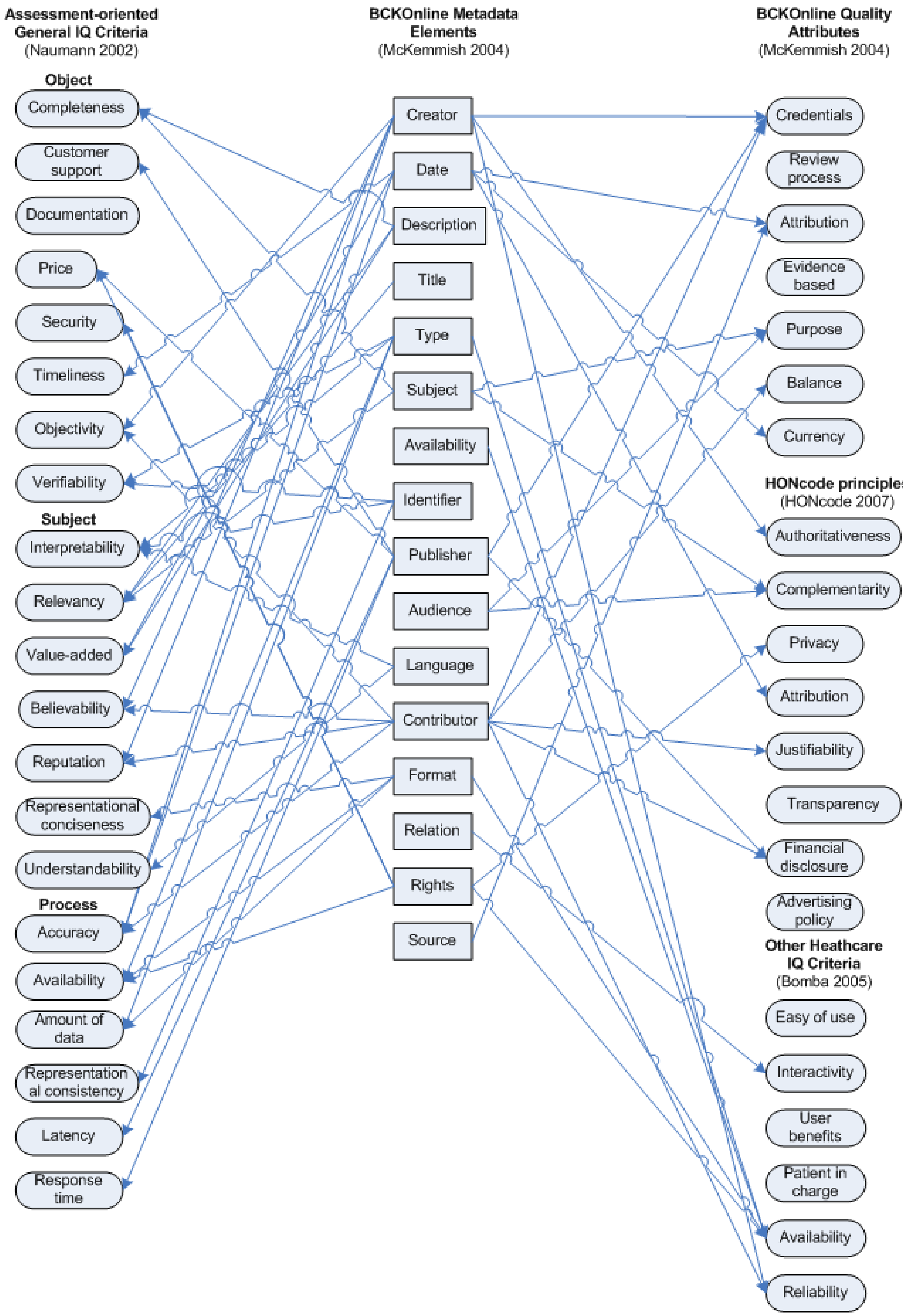
Ziebland S. (2004), The importance of being expert: The quest for cancer information on the Internet, *Social Science and Medicine*, vol. 59, no. 9, pp. 1783-1793.

APPENDICES

Appendix A: Quality Initiatives for Online Health Information

Quality Initiative	Launch Year	URL
Health On the Net (HON) Code of conduct	1996	https://www.hon.ch
Guidelines for Medical and Health Information Sites on the Internet (AMA)	2000	http://www.ama-assn.org
Health Internet Ethics (Hi-Ethics)	2000	http://www.webmd.com
International eHealth code of ethics	2000	http://www.ihealthcoalition.org
URAC health Web site accreditation program	2001	http://www.urac.org
MedPICS Certification and rating of trustworthy and assessed health information on the net (MEDCERTAIN)	2000	http://www.medcertain.org
European Community (EC) quality criteria for health related websites (E-Europe Criteria)	2001	http://ec.europa.eu/information_society/eeurope/ehealth/index_en.htm
Organizing networked information (OMNI), now known as Intute: Health and Life Sciences	1996	http://www.intute.ac.uk/healthandlifesciences/
DISCERN	1999	http://www.discern.org.uk
British Healthcare Internet Association (BHIA)	1996	http://www.bhia.org
Healthinsite	2001	http://www.healthinsite.gov.au
MedlinePlus – National library of medicine	1998	http://medlineplus.gov/
Canadian Health Network (CHN)	1992	http://www.canadian-health-network.ca/
The Health summit working group: criteria for assessing the quality	1999	http://hitiweb.mitretek.org/docs/policy.html
TNQ - Quality medical information and communication(QMIC)	2002	http://www.health.tno.nl/en/news/qmic_uk.pdf

Appendix B: Matching IQ Criteria and Existing Metadata Elements



Appendix C: Analysis of Metadata Elements and Generation

C.1 Summary of BCKOnline and HHOnline Metadata Elements

Metadata Element Schema Element	**Description**	**BCKOnline** ***Qualifier*: encoding scheme**	**HHOnline** ***Qualifier*: encoding scheme**
DC Type	The type element and its Category Qualifier is used to enable the sorting and discovery of resources with reference to four major categories: Medical, Psychological, Self-help and Personal Stories.	*Category*: medical, supportive, personal	*Category*: medical, psychological, self-help, personal stories
DC Title	A name given to the resource. Choose a title that conveys the correct meaning for the resource.		
DC Identifier	An unambiguous and unique reference to the resource within a given context.	Uniform Resource Identifier (URI)	URI
DC Creator	The name of the person or organisation responsible for creating the resource content.		
DC Publisher	The entity responsible for making the resource available online.		
DC Contributor	The name of a person or organisation with an important contributory role in the creation of the resource content.		
DC Date	The date when the actual resource was created.	*Creation Date*; *Modification Date*; *Issued Date*	*Creation Date*; *Modification Date*
DC Rights	Information about rights held in and over the resource.		
DC Description	A brief textual description of the content and purpose of the resource.		
DC Subject	Glossary is disease specific and incorporates medical terminology as well as common usage terms, which are particularly appropriate for the	MESH medical subject headings; Breast cancer Victoria glossary; BCKOnline key words; BCKOnline disease	MESH medical subject headings; HHOnline glossaries

	Australian context, and current evidence based guidelines.	trajectory	
DC Language	The language of the resource.		
DC Format	Format contains the description of the physical or virtual characteristics of the medium in which the resource is stored.	Internet Media Types	Internet Media Types
AGLS Availability	The availability element has details on how the resource can be obtained or accessed, or contact information if the resource is offline.		
DC Source	Information about a current resource and from where it is derived	URI	URI
DC Relation	A reference to a related resource	URI	URI
AGLS Audience	The target audience of the resource	*Age Group*: under 40, 40-49, 50-59, over; *Disease Stages*: early breast cancer, recurrent breast cancer, advanced breast cancer; *User Type*: self, partner/spouse, friend, parent, child; *Information Preference*: plain/brief, plain/detailed, scientific/brief, scientific/detailed; *Locality of User Residence*: rural, urban & rural	*Disease Status*: at risk, one cardiac event/procedure, multiple cardiac events/procedures; *User Type*: patient, family, friend, professional; *Information Preference*: plain/brief, plain/detailed, scientific/brief, scientific/detailed
BCKOnline Quality	Information related to the quality of a resource	*Credentials Author*: cancer organisation, clinician, commercial body, consumer group, educational institution, government organization, lay author, researcher, medical organisation; *Credentials Publisher*: the same as above; *Review Process*: editorial board, peer review process, no editorial/peer review process; *Attribution of Sources*: yes, no;	*Credentials*: delete 'cancer organisations'; *Review Process*: include 'moderator/facilitator'; *Attribution*; *Evidence-based*; *Purpose*; *Balance*; *Currency*: For these qualifiers, the encoding schemes are the same as BCKOnline

	Evidence-based: meta-analysis, randomised clinical trial, case/cohort study, review, consensus opinion, personal opinion; *Purpose*: commercial, discussion forum, educational/informative, reportage of results, review; *Balance*: controversial issue – noted, controversial Issue – not noted, non-controversial Issue; *Currency*: current, non-current	

C.2 Automated Metadata Generation Methods

Element	Source	Related studies
Creator	Creator metadata; Author statement on rendered page; Domain name; Organisation from publishing context, e.g. statement on rendered page	"Creator assignment is the simplest of metadata assignment tools. Metadata is extracted from the content attribute of any Meta tag whose name is creator or dc:creator. A blacklist is then used to eliminate any unwanted terms [e.g. none, unknown]" (Paynter 2005, p. 294). "Another promising approach is to use name authority files to support Creator metadata capture, as in cataloguing and other metadata tools" (Paynter 2005, p. 295). Use of automatically or manually generated creator profiles to populate creator element as "quite often information about these people is available from different sources" (Cardinaels et al. 2005, p.549).
Publisher	Harvest from meta tags, domain name lookup	DC-DOT - assign publisher by looking up owner of domain name in DNS (Paynter 2005, p. 295).
Date	Date metadata; Date statements on rendered page; File date	Differences in creation and publication dates from manual and automated tools "normal, understandable, and/or not meaningful or not important" (Cardinaels et al. 2005, Table 2 p.554).
Description	Harvest existing abstracts, headings, etc.; Generate using text summarization algorithms	Mani's (2006) terminology; Extract versus abstract; Indicative abstract, versus informative abstract, versus critical abstract; Generic versus topic-focused; From Official Google WebMaster Central Blog; "The quality of your snippet — the short text preview we display for each web result — can have a direct impact on the chances of your site being clicked (i.e. the amount of traffic Google sends your way). We use a number of strategies for selecting snippets, and you can control one of them by writing an informative meta description for each URL." "Description field acts as summary when records are displayed and also supports user searches" (Paynter 2005, p. 296). "The iVia Description process is based on two sources: HTML Meta tags and a text summarization algorithm. The first step is to check for Metadata tags named description or dc:description, and if either are present, they are used as the description. If that fails, a text summarization program is used to extract a summary" (Paynter 2005, p. 296). The tool used is AutoAnnotator, which is based on sentence and paragraph scoring. Words are given scores based on whether "important", content or stop word, and on the kind of html markup in which they appear. Sentence, paragraph and textual division scores are then adjusted for position and the highest scoring paragraph in the highest scoring text division returned. (Paynter 2005, p. 296)
Title	Harvest from <title>, <meta> <h1> tags or text	"A list of potential titles is built up by extracting text from the following sections of the HTML Document in order: the content of any meta tag whose name is dc:title; the title tag; all h1 tags;

		the sequence of words in the first 50 letters of body text…The initial text is post-processed to remove duplicate entries, blacklist undesirable values (e.g. Homepage, Untitled Document) and remove unwanted prefixes (Welcome to, Homepage of) while preserving the order of the list. The values in the list are assumed to be in order of decreasing quality, so that when a single Title is required, the first is used" (Paynter 2005, p. 294) Other such tools include: Open Text Summariser, Extractor.
Subject	Testing of HAL decision forest algorithms to determine by analysis of content for tone (Chen et al. 2008)	INFOMINE Keyphrase assignment – "Precision is important because the Keyphrase field is relatively highly weighted in INFOMINE, and …relatively few phrases are used to represent entire documents. Recall is also important, as it approximates the degree of coverage of the document subject matter that has been attained" (Paynter 2005, p. 295). "The iVia Keyphrase assignment module combines keyphrases from two complementary sources, then ranks and post-processes the results. The first source is any HTML document meta tag named keyword or keywords. … The second source of phrase data is iVia's PhraseRate keyphrase assignment engine" (Paynter 2005, p. 295). LCSH assignment using INFOMINE collection as 'expert-assigned training data' to identify a) similar documents and b) most popular LCSH terms assigned to those documents. Also interacts with Keyphrase assignment module (Paynter 2005, p. 297-298). Limitations are noted – overlooks relationships between similar LCSH, requires training data, too simple => "We are replacing it with a system that discovers structure within the topical LCSH using expert knowledge and automatic clustering, and then induces a hierarchical classifier similar to the one proposed for LCC assignment" (Paynter 2005, p. 29). INFOMINE Category Assignment: "builds a set of category classifiers, each of which is a probabilistic binary classifier that classifies a new example as either belonging to, or not belonging to, a particular category" (Paynter 2005, p. 298).

Appendix D: Analysis of BCKOnline and HHOnline Portal Data

D.1 Top 10 Major Sources of Resources Included in the BCKOnline Portal

BCKOnline top 10 websites	No. of portal included resources from the website	Percentage in total
http://www.medscape.com/	58	6.66%
http://www.cancerbackup.org.uk/	46	5.28%
http://theoncologist.alphamedpress.org/	39	4.48%
http://www.chemocare.com/	38	4.36%
http://www.cancerlynx.com/	37	4.25%
http://www.racgp.org.au/	34	3.90%
http://www.cancersupportivecare.com/	33	3.79%
http://www.breastcancer.net/	28	3.21%
http://www.cancer.gov/	22	2.53%
http://www.supportiveoncology.net/	21	2.41%
Subtotal	**356**	**40.87%**

D.2 Top 10 Major Sources of Resources Included in the HHOnline Portal

HHOnline top 10 websites	No. of portal included resources from the website	Percentage in total
http://www.betterhealth.vic.gov.au/	88	19.95%
http://www.heartfoundation.com.au/	47	10.66%
http://www.adavic.org.au/	16	3.63%
http://www.texasheartinstitute.org/	14	3.17%
http://www.beyondblue.org.au/	12	2.72%
http://www.arcvic.com.au/	10	2.27%
http://www.blackdoginstitute.org.au/	10	2.27%
http://www.americanheart.org/	9	2.04%
http://www.heartresearchcentre.org/	9	2.04%
http://www.mja.com.au/	9	2.04%
Subtotal	**224**	**50.79%**

D.3 Health and Medical Websites Analysis

C = Copyright
D = Disclaimer
P = Privacy
AU = About Us
CU = Contact Us
LP = Linking Policy
HC = HONcode Certified
TC = TRUSTe Certified
UA = URAC Accredited
HA = HealthInsite Approved

Website	**C**	**D**	**P**	**AU**	**CU**	**LP**	**HC**	**TC**	**UA**	**HA**
http://www.medscape.com/	x	x	x	x	x	x	x	x	x	
http://www.cancerbackup.org.uk/	x	x	x	x	x		x			
http://theoncologist.alphamedpress.org/	x	x	x	x	x	x				
http://www.chemocare.com/	x	x	x	x	x					
http://www.cancerlynx.com/	x		x	x	x					
http://www.racgp.org.au/	x	x	x	x	x	x	x			x
http://www.cancersupportivecare.com/	x	x	x		x					
http://www.breastcancer.net/	x			x	x					
http://www.cancer.gov/	x	x	x	x	x	x				
http://www.supportiveoncology.net/	x			x	x					
http://www.betterhealth.vic.gov.au/	x	x	x	x	x	x	x			x
http://www.heartfoundation.com.au/		x	x	x	x					
http://www.adavic.org.au/	x	x	x	x	x					
http://www.texasheartinstitute.org/	x	x	x	x	x		x			
http://www.beyondblue.org.au/	x	x	x	x	x	x				x
http://www.arcvic.com.au/	x	x	x	x	x					x
http://www.blackdoginstitute.org.au/	x	x	x	x	x		x			x
http://www.americanheart.org/	x	x	x	x	x	x				
http://www.heartresearchcentre.org/	x		x	x	x					
http://www.mja.com.au/	x	x			x					
http://www.healthinsite.gov.au/	x	x	x	x	x	x	x			
http://www.webmd.com/	x	x	x	x	x	x	x	x	x	
http://www.healthline.com/	x	x	x	x	x					
http://www.healia.com/	x	x	x	x	x	x	x			
http://www.mayoclinic.org/	x	x		x	x					
Total	**24**	**21**	**21**	**23**	**25**	**10**	**9**	**2**	**2**	**5**

Appendix E: Analysis of BCKOnline Quality Attributes

BCKOnline quality attribute	Definition and examples	BCKOnline quality attribute encoding scheme (McKemmish et al. 2004)	Associated quality indicators	Automated tools or techniques	Related studies
Credentials	Extent to which information is regarded as credible, e.g. authors, publishers, contributors, their affiliations and relevant credentials	▪ Lay Author ▪ Clinician ▪ Researcher ▪ Consumer Group ▪ Commercial Body ▪ Educational Institution ▪ Government Organisation ▪ Medical Organisation ▪ Cancer Organisation	BCKO, DC, AGLS or HIDDEL metadata (e.g., Creator, Publisher and Contributor)	HTML parser	Haynes et al. 1994; (MedCircle Collaboration, 2003; Eysenbach 2005b)
			Terms of credentials in the Web content (e.g. "MD")	Search techniques and Web data extraction tools	(Griffiths and Christensen 2005; Griffiths et al. 2005; Naumann 2002)
			Creator Affiliation and disclosure	Search techniques and Web data extraction tools, and HTML parser	
			URL (e.g., .gov, .org or .com)	URL inspector	(Price and Hersh 1999)
			Third-party labels or seal of certification (e.g. HON)		(Price and Hersh 1999)
Evidence basis	Extend to which information is regarded as correct, reliable and certified free of error.	▪ Meta-analysis ▪ Randomised Clinical Trial ▪ Case/cohort Study ▪ Review ▪ Consensus Opinion ▪ Personal Opinion	Seal of certification (e.g. HON)		(Price and Hersh 1999)
			Web statistics (e.g. Google PageRank, Usage analysis	Web link analysis and Web usage statistics, e.g. AQA (Automated Quality Assessment) Procedure	(Griffiths and Christensen 2005; Griffiths et al. 2005; Price and Hersh 1999)
			DISCERN scores		(Griffiths and Christensen 2005; Griffiths et al. 2005)
Attribution of sources	References and sources of Web contents, and copyright information	▪ Yes/No	Provided references and copyright notice	HTML parser and Web link analysis	(Wang and Liu 2007)

Balance	Extent to which information is unbiased, unprejudiced and impartial	▪ Controversial issue – noted ▪ Controversial issue – Not noted ▪ Non-controversial issue		Search techniques and Web data extraction tools	(Griffiths and Christensen 2005; Griffiths et al. 2005)
Currency	Extent to which the information is sufficiently up-to-date for the task at hand	▪ Current ▪ Non-current	Date of creation, date of last update	HTML parser	(Wang and Liu 2007)
Review process	Extent to which information is peer reviewed, thus correct and reliable	▪ Editorial Board ▪ Peer Review ▪ Editor (proposed) ▪ Moderator (proposed) ▪ Post-publication Review Process ▪ Audience Review Process ▪ No process	Editorial review process and disclosure	HTML parser Web link analysis and Web usage statistics	(Wang and Liu 2007) (Cui 1999; Hernández-Borges et al. 1999)
Purpose	Ownership of the Website, sponsorship, advertising, underwriting, commercial funding arrangements or support, or potential conflicts of interest, and arrangements in which links to other sites are posted as a result of financial considerations.	▪ Educational or Informative ▪ Commercial (to be replaced by reference to whether resource includes Advertising) ▪ Reportage of Results ▪ Commentary (proposed) ▪ Review (proposed) ▪ Assessment or Evaluation (proposed) ▪ Discussion Forum ▪ Listserv (proposed) ▪ Bulletin Board (proposed)	Statement of purpose and disclosure	Search techniques and Web data extraction tools, and HTML parser	(Wang and Liu 2007)

Appendix F: DED Evaluation Explanatory Statement and Consent Form

F.1 Explanatory Statement

June 2010

Explanatory Statement **– Domain Experts**

Title: An Evaluation of the Domain Expert Dashboard for Sustaining Resource Quality Assessment Processes in Health Information Portals

My name is **Jue Xie** and I am conducting a research project with **Prof. Frada Burstein, Prof. Sue McKemmish and Dr. Joanne Evans** in the Faculty of Information Technology towards a Doctor of Philosophy degree at Monash University. This means that I will be writing a thesis of which is the equivalent of a 300-page book.

Why did you choose this particular person/group as participants?
All participants are or had been working for consumer-oriented health information portals as domain experts responsible for portal content management. Their expertise lies in either medical domain knowledge, or information management skills. Their working experience as a domain expert varies from 3 months to more than 5 years. We are interested in the way domain experts use the prototype system delivered by this research project, so as to evaluate the usefulness and effectiveness of the proposed quality solution underlining the system implementation.

The aim/purpose of the research
The aim of this project is to evaluate the user-sensitive quality framework and tool (UQFAT) proposed by this thesis research in terms of its usefulness and effectiveness in evaluating, describing and maintaining the quality of selected web resources for achieving quality assurance in an information portal.

Consumer Health information portals are specialised gateways to a vast number of quality controlled online resources in specific healthcare or disease domains. Quality assurance in such a portal is achieved through a set of processes undertaken by domain experts to evaluate, annotate and maintain the quality of included online resources via the domain expert interface to a portal's content management system.

Possible benefits
By agreeing to participate in this project, you help us to evaluate the usefulness and effectiveness of the prototype system in assisting domain experts to manage the quality of portal content. The information you provide will help us to understand the way the system is used so as to test the validity of the proposed quality framework, which underpins UQFAT, and refine the design of UQFAT in respect to the evaluation results. It is envisaged that the UQFAT approach can both enhance domain experts' decision-making abilities and improve the consistency of quality assessment results.

What does the research involve?
The study involves a set of semi-structured interviews that will be audio taped. Each interview involves a work through of several use cases and the submission of a survey questionnaire being completed towards the end of the session. The collected data will be analysed and aggregated according to the themes to articulate how domain experts can achieve quality assurance through the use of a quality tool in a portal's content management system. The potential fields for improvement can then be identified to refine the design of the proposed quality solution.

How much time will the research take?
Each data collection session will require 1.5 hours of your time.

Inconvenience/discomfort
There is no perceived potential inconvenience or discomfort to the participants. It is not likely that the involved system evaluation would have any foreseeable risks of harm or side effects to the participants.

Can I withdraw from the research?
Being in this study is voluntary and you are under no obligation to consent to participation. You are free to withdraw your consent and discontinue your participation in the research project any time prior to the completion of your data collection session.

Confidentiality
All information we receive from you will remain confidential. The information will be coded and identifiable by the researchers of this study only.

Storage of data
Storage of the data collected will adhere to the University regulations and kept on University premises in a locked cupboard/filing cabinet for 5 years on a password protected hard disk or in hard copies as the data were originally collected. A report of the study may be submitted for publication, but individual participants will not be identifiable in such a report.

Use of data for other purposes
The data collected in this study will be used in this thesis research, specifically for the purpose of testing the validity and feasibility of the proposed quality solution for sustaining resource quality assessment processes in consumer health information portals. Also, some of the data is likely to be published in academic journals and presented at conferences. However, due to the coding of data with the strict access of the researchers only, nobody will be named or identified in any way. Please keep in mind that it is sometimes impossible to make an absolute guarantee of confidentiality/anonymity. However, only researchers of this study will have access to the real names of the participants.

Results
If you would like to be informed of the aggregate research findings, please contact **Jue Xie on** [redacted] **or fax (**[redacted] The findings are accessible for **5 years**.

If you would like to contact the researchers about any aspect of this study, please contact the Chief Investigator:	If you have a complaint concerning the manner in which **CF10/1090 - 2010000574** is being conducted, please contact:
Prof. Frada Burstein Ph: [redacted] Fax: [redacted] Email: [redacted]	Executive Officer Monash University Human Research Ethics Committee (MUHREC) Building 3e Room 111, Research Office Monash University VIC 3800 Tel: +61 3 9905 2052 Fax: +61 3 9905 3831 Email: muhrec@adm.monash.edu.au

Thank you.

Jue Xie

F.2 Consent Form

Consent Form – **Domain Experts**

Title: An Evaluation of the Domain Expert Dashboard for Sustaining Resource Quality Assurance Processes in Health Information Portals

NOTE: This consent form will remain with the Monash University researcher for their records

I agree to take part in the Monash University research project specified above. I have had the project explained to me, and I have read the Explanatory Statement, which I keep for my records. I understand that agreeing to take part means that:

I agree to be interviewed by the researcher	☐ Yes ☐ No
I agree to be observed by the researcher	☐ Yes ☐ No
I agree to allow the interview to be audio-taped	☐ Yes ☐ No
I agree to complete questionnaires asking me to evaluate a system	☐ Yes ☐ No
I agree to make myself available for a further interview if required	☐ Yes ☐ No

and

I understand that my participation is voluntary, that I can choose not to participate in part or all of the project, and that I can withdraw from participation at any time prior to the completion of this data collection session without being penalised or disadvantaged in any way.

and

I understand that any data that the researcher extracts from the interview for use in reports or published findings will not, under any circumstances, contain names or identifying characteristics.

and

I understand that any information I provide is confidential, and that no information that could lead to the identification of any individual will be disclosed in any reports on the project, or to any other party.

and

I understand that data collected from this session will be kept in a secure storage and accessible only to the research team. I also understand that the data will be destroyed after a 5-year period unless I consent to it being used in future research.

Participant's name

Signature

Date

Appendix G: Exploratory Case Study Interview Instrument

Investigating Resource Identification, Selection and Description Processes for Smart Information Portals

First Interview Questions

1. Describe the development of the portal from your perspective? What stages were you involved in? What role(s) did you play in those stages? What expertise did you bring to the project in the domain? What expertise/experience did you gain from your role in the project?

2. Resource identification and selection processes

   a) By what means are resources identified for inclusion in the knowledge repository? For each means can you identify a specific example and explain what was involved in its assessment for inclusion? For example:

      - Prior knowledge – what percentage of resources were already known to exist prior to the construction of the knowledge repository? What kinds of resources in the user centred categories were already known to exist?
      - Alert services – what kinds of alerts services e.g. consumer health advocacy group, medical professional groups, general news services, etc.? Are different kinds of alert services useful for particular kinds of resources?
      - Web searching – what resources were identified through the use of generic or specialist search engines? Was this searching based on identified needs from user studies and/or feedback?

   b) For each of the category types used in the portal, can you identify an example and explain what was involved in its identification and assessment for inclusion?

   c) What determines the level of granularity at which a resource is assessed for inclusion?

   d) How did the Audience Qualifier scheme influence resource identification and selection activities? Where representative resources for each of the qualifiers sort? Do identification and assessment techniques vary across the qualifiers?

   e) Can you recall examples of resources that were identified but deemed unsuitable for inclusion in the repository? Did these relate to any particular type of resource and audience categories?

   f) Describe the current priority list of resources to be added to the repository? Has this list been influenced by feedback from users of the repository? If so what kinds of resources from what kinds of users?

3. Resource description processes

   a) What determines the level of granularity at which a resource is described?

   b) Describe how the metadata values for a recently added resource for each of the category types, were determined?

   c) Where any particular steps taken to assure metadata quality and consistency?

Second Interview Questions

The second interview will take place in the context of a walkthrough of resource identification, selection and description exercises with the portal's domain expert interface and metadata schema.

A set of resources will be determined (from the first interview) and the completion of the metadata schema for each will be undertaken.

Questions will be asked regarding the sourcing of metadata values and the decision making processes of the domain expert in completing those values. Opinion on the usability of the interface and the metadata schema and what might be done to improve will also be sought.

Appendix H: Scenario Test Instrument

H.1 Scenario Test: Evaluation of the Intelligent Metadata Value Suggestion Feature

November 2010

Introduction:

In this test, you will be asked to assign values for 5 attributes of the BCKOnline *Quality* metadata element for 4 INCOMPLETE resource records carefully selected from the BCKOnline portal's metadata repository. Please follow the instructions to complete two major tasks:

1. Assign values for 5 *Quality* attributes of 2 incomplete resources without value suggestions;
2. Assign values for 5 *Quality* attributes of 2 incomplete resources with value suggestions.

Please refer to Attachment A for a full description of the value encodings of 5 *Quality* attributes.

Access to the system:

The Content Management System (CMS) of the BCKOnline portal is accessible via http://bcko3.infotech.monash.edu/cms/main.do

Please use the below account to login the system in either *Google Chrome* or *Safari*:

User name: tester
Password: fit-monash

How to retrieve a resource:

Figure 1 below shows how to retrieve a resource by using its id number in the system's *Manage Resources* view. For instance, to retrieve a resource of id 1452, please follow the below steps:

1. Select 'INCOMPLETE' in the *Status* box
2. Input 'id:1452' in the *Search* box
3. Press the ENTER button

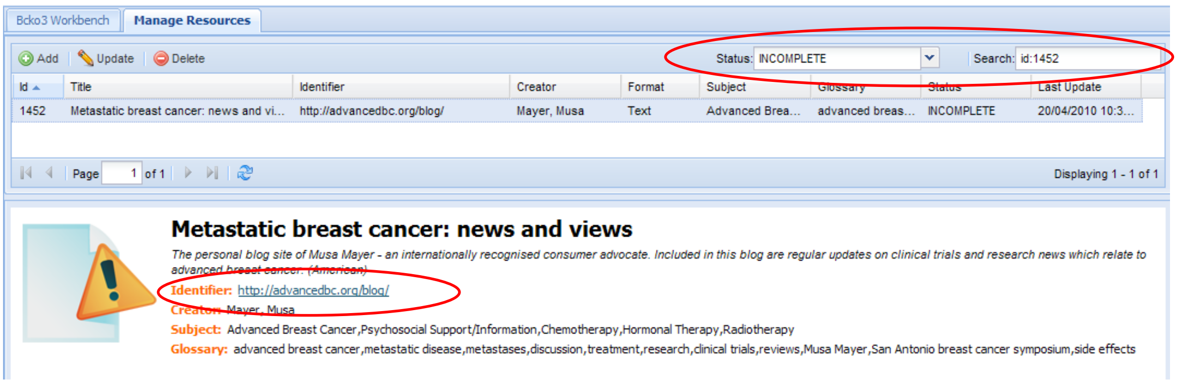


Figure 1: Retrieve the INCOMPLETE resource of id number 1452

How to open a resource via its external URL

When a resource is selected in the *Manage Resources* view, a brief description of the resource will be shown under the data list. You can follow the clickable URL (as shown in Figure 1 above) to open the resource that is external to the portal in a new Tab or a new window of the browser you use.

How to open a data entry form

To carry out the required tasks, you need to firstly retrieve a resource by using the above method, then DOUBLE CLICK the retrieved resource to open up a data entry form. In the system, those 5 *Quality* attributes can be found in the *Quality* Tab, as grouped in the RED box in Figure 2.

Figure 2: Five *Quality* attributes in the Quality Tab

Assign metadata values WITHOUT suggestions

Table 1 below lists 2 resources selected from a wide range of INCOMPLETE resources of the portal. Resource ***1452*** comes from a website that is known by the portal, which means some resources from the same site have been previously annotated and included in the portal. Resource ***1485*** comes from a website that is new to the portal.

Please **DO NOT** assign values directly in the system, instead please complete Table 2 by referring to:

1. other metadata values that are available and searchable in the system.

Please **RECORD** the time you take for assigning values for each resource.

Table 1: Two INCOMPLETE resources for value assignment without suggestions

Id	Identifier	Title
1452	http://advancedbc.org/blog/	Metastatic breast cancer: news and views
1485	http://blogs.crikey.com.au/croakey/2009/11/12/how-should-we-respond-to-the-new-breast-cancer-screening-study/	How should we respond to the new breast cancer screening study?

Table 2: Human assigned values for the above 2 INCOMPLETE resources (refer to Attachment A for description)

Id	Attribution of Sources	Balance	Purpose	Publisher Credentials	Review Process	Time Spent
1452						
1485						

Assign metadata values WITH suggestions:

Table 3 below lists the other 2 resources selected from a wide range of INCOMPLETE resources of the portal. Resource ***1507*** comes from a website that is known to the portal, whereas resource ***1543*** comes from a website that is new to the portal.

Please **DO NOT** assign values directly in the system, instead please complete Table 5 by referring to:

1. other metadata values that are available and searchable in the system, and
2. machine generated values provided in Table 4.

Please **RECORD** the time you take for assigning values for each resource.

Table 3: Two INCOMPLETE resources for value assignment with suggestions

Id	Identifier	Title
1507	http://www.breastcancer.org/symptoms/types/ask_expert/2008_07/	Triple negative breast cancer
1543	http://www.livingwithbrainmets.org/brain-metastases.html	Basics on brain mets

Table 4: Machine generated values for the above 2 INCOMPLETE resources (refer to Attachment A for description)

Id	Attribution of Sources	Balance	Purpose	Publisher Credentials	Review Process
1507	No	Controversial Issue - Noted	Educational/ Informative	Consumer Group	Editorial Board
1543	No	Controversial Issue - Noted	Educational/ Informative	Commercial Body	Editorial Board

Table 5: Human assigned values for the above 2 INCOMPLETE resources (refer to Attachment A for description)

Id	Attribution of Sources	Balance	Purpose	Publisher Credentials	Review Process	Time Spent
1507						
1543						

H.2 Attachment A: Value Encodings of Five Quality Attributes

Attribute	Value Encoding
Attribution of Sources	***Yes***: Where the content is associated with either footnotes, endnotes, a bibliography or references which clarify for the reader the source of the information and enable her to ascertain author, publication, and date. ***No***: Where the content is not associated with any of the above.
Balance	***Controversial Issue – Noted***: Where the content represents views, evidence, which are not endorsed by current best practice, or clinical practice guidelines, and where this fact has been acknowledged explicitly in the content. ***Controversial Issue – Not Noted***: Where the content represent views, evidence, which are not endorsed by current best practice, or clinical practice guidelines, and where this fact has not been acknowledged anywhere in the content. ***Non-Controversial Issue***: Where the content represents views, or evidence which are consistent with current best practice and have been endorsed via official consensus statements or clinical practice guidelines.
Purpose	***Commercial***: Where the content represents an overt attempt to 'sell' a product or service to the user. Advertising material, which informs the user as to the existence of a product or service, may not be construed as 'commercial' if the primary purpose of the webpage is deemed informative. ***Discussion Forum***: Where the content has been produced with the intention of publishing this on a discussion bulletin board/chat room, etc. The material may represent a response to a previous query, or be eliciting information from others, or simply presenting a personal point of view. ***Educational/Informative***: Material whose primary function is to inform/'educate' the reader by providing her with any or all of the following: background information; current status of knowledge on the issue; research results; potential application of such research; statistical analyses. ***Reportage of Results***: Where the content represents the results of a single or several studies. ***Review***: Where the content represents an overview of the current state of knowledge, which may be drawn from a variety of evidence levels. For example data may be derived from both randomised clinical trials and case/cohort studies.

Publisher Credentials	***Cancer Organisation***: bodies and/or groups which focus on the treatment, management and research of cancer. For example, COSA (Clinical Oncology Society of Australia), Cancer Council, etc. ***Clinician***: a suitably qualified health practitioner engaged in the clinical practice of medicine. ***Commercial Body***: any organization which is constituted under the Trade Practices Act as a 'commercial' organization. This may include both private and public bodies. ***Consumer Group***: any organization whose members are primarily lay consumers. The organization may be formally affiliated with other umbrella groups, or may be legally constituted as a charity, support group, or advocacy group. ***Educational Institution***: organizations whose primary purpose is teaching. For example, universities or teaching hospitals affiliated with academic departments or faculties. ***Government Organisation***: organisations and departments, which are funded or supervised by federal or state governments. For example: the American National Institutes of Health, the Australian Dept. of Health and Ageing. ***Lay Author***: an individual who has no officially recognized 'expertise' or qualifications in the medical arena. ***Researcher***: a suitably qualified health practitioner, but one whose primary area of expertise resides not in clinical practice, but in applied research and discovery. ***Medical Organisation***: this may include both individual bodies as well as affiliated groups such as the Royal Australian Colleges and their respective regional divisions. Hospitals and University Teaching Hospitals are also included in this category.
Review Process	***Editorial Board*** ***Peer Review Process*** ***No Editorial/Peer Review Process***

H.3 Attachment B: An Introduction of Domain Expert Dashboard Functions and Features

Domain Expert Dashboard

An Introduction of System Functions and Features for Evaluation

Purpose of this document

This document is created for providing a brief introduction of major functions and features of the Domain Expert Dashboard system. The system was developed as an improved domain expert interface of the BCKOnline portal's content management system (CMS). This document needs to be used in accompany with the scenario test instrument and the semi-structured interview instrument.

Introduction of the BCKOnline Portal

Currently health information portals do not provide customised, timely information to health consumers. Customised, in-time access to information impacts significantly on the ability of women with breast cancer to make health and lifestyle decisions. Although there is a plethora of information resources available online, the 'one size fits all' approach does not take into account the specific conditions such as disease stage, age, and treatment options. Breast Cancer Knowledge Online (BCKOnline) offers a customised search function, allowing women or their supporters to decide how they want to search for information specific to their context. Three options exist: a simple search, a search on a list of topics which cover the entire disease trajectory and treatment options or a personalised search. The portal matches the woman's profile with the information resources specifically selected and described in the portal and also retrieves only information relevant to that person's profile. All resources are external to the portal: only the URL is stored in a database with a description of each resource. The creation of the resource description and maintenance of the database is currently relied on human expertise. This project will use advances in social technologies and involve users in the creation of the resource descriptions using a user-sensitive design methodology.

The Domain Expert Dashboard Prototype System of

Access to the system:

The Content Management System (CMS) of the BCKOnline portal is accessible via http://bcko3.infotech.monash.edu/cms/main.do

Please use the below account to login the system in either *Google Chrome* or *Safari*:

User name: tester
Password: fit-monash

Introduction of Basic System Functions

Portal content here includes user account information, metadata descriptions of external online resources, and the glossary terms. Therefore, portal data in the context of this document refer to any kinds of data as mentioned above. To facilitate the data management, the following basic functions are provided:

- Views/Tabs for managing User accounts, Resources, and Glossary terms
- Flexible navigation to different views
- Browse data in a list view
- Add or Edit data in the system provided with the input validation
- Delete one or multiple data entries at one time provided with the delete confirmation dialogue

- Input validation
- Help messages

Introduction of System Features

The prototype system is featured by a number of tools and interactive interface design, listed as the following:

- Data visualisation (a)
- Sorting data (b)
- Retrieving data (b)
- URL availability checking (b)
- Grouping metadata in Tabs in the data entry form (c)
- Metadata value suggestion (c)

The above features are implemented within different data management views, which are introduced in the following sub-sections. The number in brackets next to each feature indicates the sub-section in which the feature is introduced.

a) Workbench view
For the purpose of monitoring resource distribution and value consistency, data are visualised in charts in the workbench view.

- Pie chart of portal included resources: resource status is colour coded representing PUBLISHED, UNAPPROVED or INCOMPLETE resources; the size of colour-coded area in the chart represents the proportion of resources of a certain status.
- Pie chart of URL status: URL status is colour coded representing VALID, INVALID, or UNCHECKED resource URL status; the size of colour-coded area in the chart represents the proportion of resource URLs of a certain status.
- Bar chart: colour represents the resource status; length represents the number of resources in a certain category

b) Resource management view
The resource list is characterised by the following features:

- Filter: a dropdown list is provided for filtering resources by their status
- The width of each column is adjustable
- Can directly go to any page of the list
- The list content can be refreshed anytime by just pressing the Refresh button
- Select a resource in the list to see its description underneath the list
- Sorting data in ascending or descending order
- Searching data on specific fields, e.g. Title or Id
- Checking the status of one or multiple URL(s) at one time. The link status can be either VALID, INVALID or UNCHECKED

c) Data entry form for meta-tagging resource

- The form is used to Add/Edit the metadata record of an external resource. Open up the form is easy, fast and flexible by left-clicking the Add/Edit button or by double-clicking the selected resource.
- Basic descriptive metadata are promoted in the top half of the form, while the rest part of metadata is grouped in different Tabs. Normally if one metadata element has multiple attributes, it will be placed in a separate Tab, e.g. Audience and Quality element.
- The function of providing metadata value suggestion for the Quality element is implemented and is functioning in the command line. However constrained by time, the feature has not been integrated into the user interface, e.g. the Quality Tab of the data entry form.

Appendix I: DED Evaluation Interview Instrument

Semi-structured Interview Instrument

Evaluation of the Domain Expert Dashboard System

November 2010

Purpose of this document

This document is created for guiding a one-hour semi-structured interview for the evaluation of a Domain Expert Dashboard system. The prototype system is implemented as an enhanced domain expert interface with new features to the content management system (CMS) of the BCKOnline portal. The interview participant is selected for his/her expert medical domain knowledge or information management expertise.

Interview procedures

The interview consists of two parts: a 20-minute tutorial on system functions and features, and a 40-minute evaluation of the *Usefulness*, *Satisfaction* and *Ease of Use (USE)* of the system.

Tutorial (20 minutes)

A brief introduction will be given on the following aspects of the prototype system:

1. System basic functions for evaluation:
    - Views/Tabs for managing User accounts, Resources, and Glossary terms
    - Instant navigation to different views
    - Browse data in a list view
    - Add or Edit data in the system provided with the input validation
    - Delete one or multiple data entries at one time provided with the delete confirmation dialog
    - Input validation
    - Help messages/tips
2. System features for evaluation:
    - Data visualisation
    - Sorting data
    - Retrieving data
    - URL availability checking
    - Grouping metadata in Tabs in the data entry form
    - Metadata value suggestion

System Evaluation (40 minutes)

The Domain Expert Dashboard system for evaluation is featured by some major improvement, as listed in Table 1 on the next page. In this interview, the participant is asked to evaluate these system functions and features in terms of their *Ease of Use*, *Usefulness* and the overall *Satisfaction*.

Questions about the participant:

1. Do you have experience in managing data for information portals? If yes, how long had you been in this role (please CIRCLE one answer below)? Were you dealing with online health information?

Less than 3 months	3-6 months	6-12 months	1-2 years	More than 2 years

2. How you describe your experience level of using web portals or search engines for finding health information on the Internet (please CIRCLE one answer below)?

No experience	Limited experience	Some experience	Very experienced

3. How you describe your experience or knowledge level of portal targeted users or communities when undertaking the role for portal content management? (please CIRCLE one answer below)

No experience /knowledge	Limited experience /knowledge	Some experience /knowledge	Very experienced /knowledgeable

4. Which quadrant in the below diagram best describe your domain expertise?

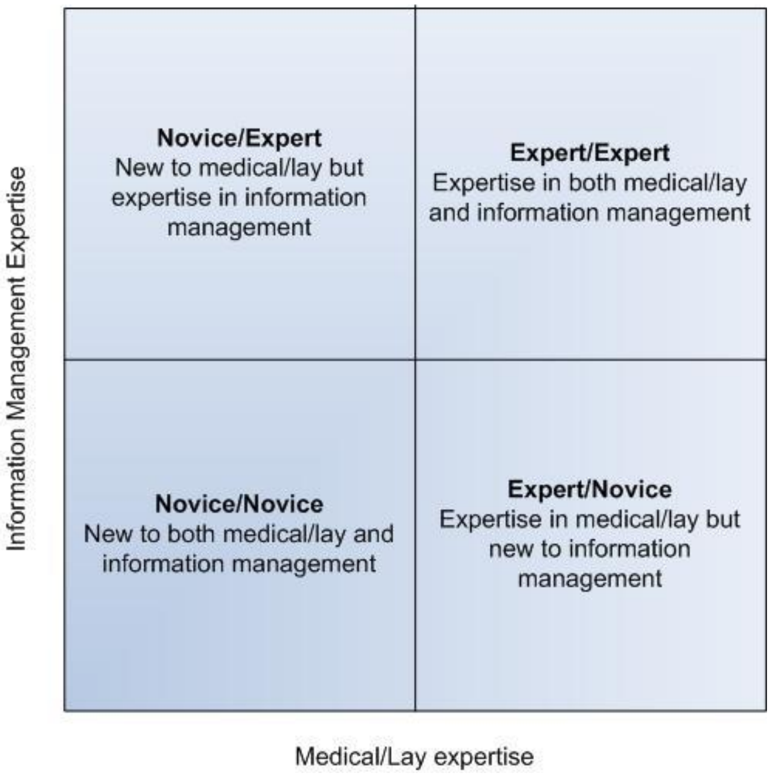


Figure 3: Domain Expertise Quadrant Diagram

USE questionnaire:

5. Based on your experience with similar systems, please rate your agreement with each statement listed in Table 1 to evaluate the *Ease of Use*, *Usefulness*, and overall *Satisfaction* of using the system for managing portal content. Please CIRCLE the level of agreement that applies (where 1 means *Strongly disagree*, 2 means *Disagree*, 3 means *Neither disagree nor agree*, 4 means *Agree*, and 5 means *Strongly agree*):

Table 1: System functions or features under USE Evaluation

Ease of Use:					
Please rate your agreement with the following statements about how you feel about the Ease of Use of the system. Please ***CIRCLE*** *the level of agreement that applies:*	**Strongly Disagree**	**Disagree**	**Neither disagree nor Agree**	**Agree**	**Strongly Agree**
E1. It is quick and flexible to open or switch between data management views	1	2	3	4	5
E2. It is easy to browse data.	1	2	3	4	5
E3. It is easy and flexible to add or edit data in data entry forms.	1	2	3	4	5
E4. It is easy, flexible and safe to delete data records.	1	2	3	4	5
E5. It is easy and flexible to sort data.	1	2	3	4	5
E6. It is easy and flexible to retrieve data.	1	2	3	4	5
E7. I can use the above features without written instructions.	1	2	3	4	5
E8. I can use it successfully and can recover from mistakes quickly and easily.	1	2	3	4	5
Usefulness:					
Please rate your agreement with the following statements about how you feel about the Usefulness of the system feature. Please ***CIRCLE*** *the level of agreement that applies:*	**Strongly Disagree**	**Disagree**	**Neither disagree nor Agree**	**Agree**	**Strongly Agree**
U1. Data visualisation in the workbench view is useful for monitoring resource distribution and value consistency.	1	2	3	4	5
U2. The URL checking and reporting feature is helpful.	1	2	3	4	5

U3. Grouping metadata in different Tabs helped me to better describe a resource in the data entry form.	1	2	3	4	5
U4. The metadata value suggestion feature is helpful.	1	2	3	4	5
U5. By using the above mentioned features, I felt more confident about the quality of metadata values I assigned.	1	2	3	4	5
U6. I found the above mentioned features would help me to be more productive and consistent.	1	2	3	4	5
Satisfaction:					
*Please rate your agreement with the following statements about how you are satisfied with the system. Please **CIRCLE** the level of agreement that applies:*	**Strongly Disagree**	**Disagree**	**Neither disagree nor Agree**	**Agree**	**Strongly Agree**
S1. I am satisfied with the system for data management.	1	2	3	4	5
S2. The system works the way I want it to work.	1	2	3	4	5
S3. I prefer to use the system comparing to the other similar systems.	1	2	3	4	5

Open-ended questions:

6. Based on your rating in Table 1, please specify which system features you like the most? Please explain why?
7. Based on your rating in Table 1, please specify which system features you dislike? Please explain why?
8. Describe any frustrations or problems you experienced while using the system.
9. Do you agree that the system as a whole helped you to assess resources for their inclusion in the portal in a systematic and consistent manner? Please elaborate.
10. Using Figure 1 to measure your medical and information management expertise, do you feel the system can help you quickly build expertise in the area that you are initially lacking in?
11. Based on your experience in using the system, how do you think such a system can be improved in terms of supporting and facilitating resource quality assessment?
12. Any final comments on the system?

Appendix J: Functional Test Results

J.1 Headers of ARFF files

The following code snippet presents an ARFF (Attribute-Relation File Format) file header, in which 35 data attributes are defined. For training and testing a classifier for each resource quality attribute, an individual ARFF file needs to be compiled. The below file header was extracted from the ARFF file that was created for building classification model for *Attribution of Sources*.

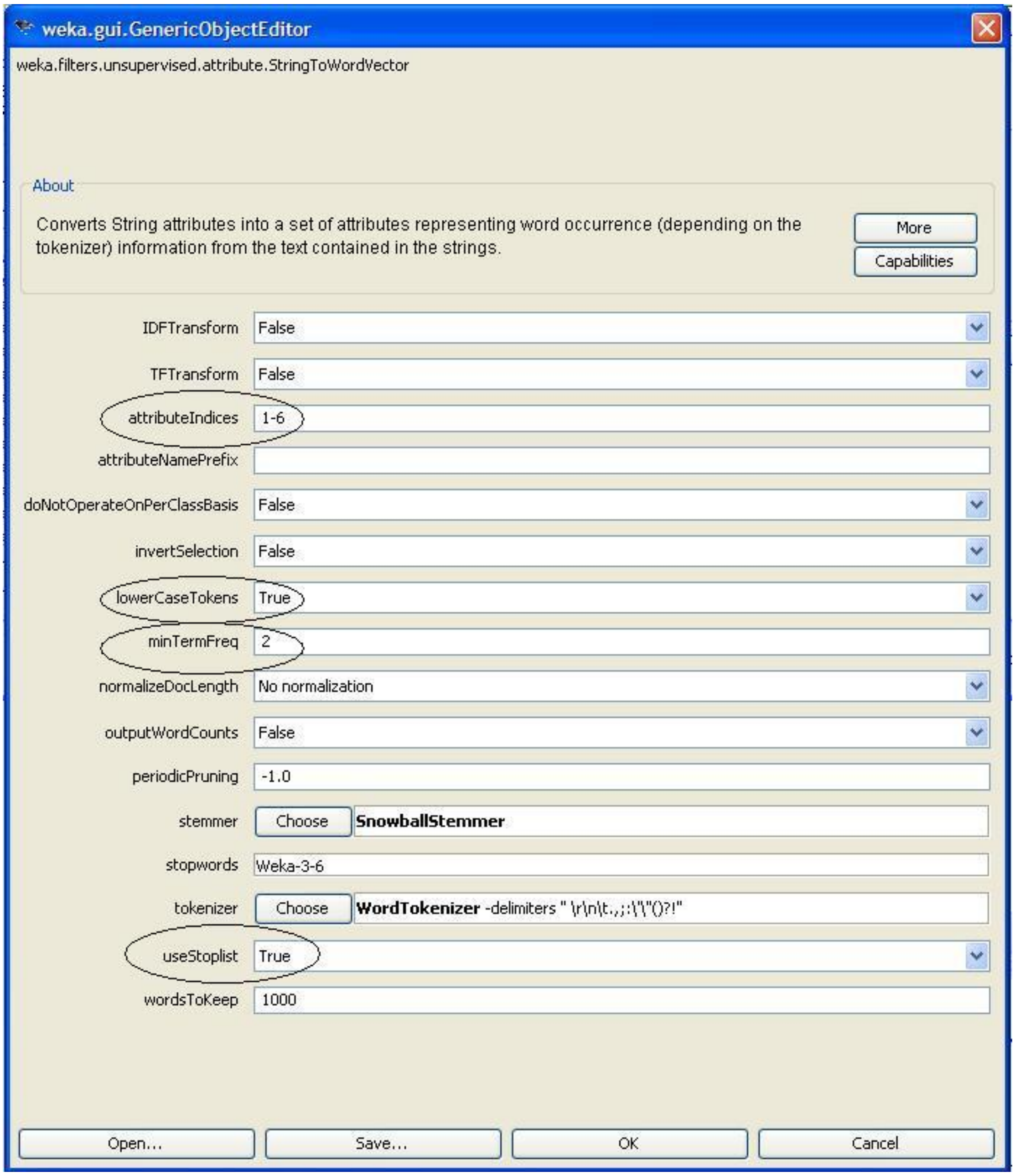
```
@relation 'Bcko Schema'

@attribute title string
@attribute creator string
@attribute description string
@attribute publisher string
@attribute rights string
@attribute glossory string
@attribute 'audience_ageGroup:Over 70' {No,Yes}
@attribute audience_ageGroup:50-69 {No,Yes}
@attribute audience_ageGroup:40-49 {No,Yes}
@attribute 'audience_ageGroup:Under 40' {No,Yes}
@attribute 'audience_diseaseStage:Early Breast Cancer' {No,Yes}
@attribute 'audience_diseaseStage:Recurrent Breast Cancer' {No,Yes}
@attribute 'audience_diseaseStage:Advanced Breast Cancer' {No,Yes}
@attribute audience_userType:Self {No,Yes}
@attribute audience_userType:Partner/spouse {No,Yes}
@attribute audience_userType:Friend {No,Yes}
@attribute audience_userType:Parent {No,Yes}
@attribute audience_userType:Child {No,Yes}
@attribute subject:Diagnosis {No,Yes}
@attribute 'subject:Type of Breast Cancer' {No,Yes}
@attribute 'subject:Early Breast Cancer' {No,Yes}
@attribute 'subject:Prevention/Risk factors' {No,Yes}
@attribute subject:Surgery {No,Yes}
@attribute subject:Radiotherapy {No,Yes}
@attribute 'subject:Hormonal Therapy' {No,Yes}
@attribute subject:Chemotherapy {No,Yes}
@attribute 'subject:Recurrent Breast Cancer' {No,Yes}
@attribute 'subject:Advanced Breast Cancer' {No,Yes}
@attribute 'subject:Palliative care' {No,Yes}
@attribute 'subject:Alternative/Complementary therapies' {No,Yes}
@attribute 'subject:Psychosocial Support/Information' {No,Yes}
@attribute type:Medical {No,Yes}
@attribute type:Supportive {No,Yes}
@attribute type:Personal {No,Yes}
@attribute quality_attributeOfSource {No,Yes}
```

In the headers of those ARFF files, which use 35 data attributes to build classifiers, only one data attribute is different to each other. That is the attribute to be classified. Their definitions in ARFF file headers are outlined in the below code snippet.

```
@attribute quality_attributeOfSource {No,Yes}

@attribute quality_balance {'Non-controversial Issue','Controversial
Issue - Noted','Controversial Issue - Not Noted'}

@attribute quality_purpose {Educational/Informative,Review,'Discussion
Forum','Reportage of Results',Commercial}

@attribute quality_credential_publisher {'Cancer
Organisation',Clinician,'Commercial Body','Consumer Group','Educational
Institution','Government Organisation','Lay Author',Researcher,'Medical
Organisation'}

@attribute quality_reviewProcess {'Editorial Board','Peer Review
Process','No Editorial/Peer Review Process'}
```

J.2 Filter Settings

The WEKA built-in filter 'Unsupervised.StringToWordVector' was selected to normalise string-typed data attributes (the first six data attributes defined in the above ARFF file header) with the following minor changes of the default parameter settings: Lower case tokens was enabled; Minimum term frequency was set to 2; Snowball stemmer was selected; and useStoplist was enabled. A screenshot of the filter parameter configuration in WEKA explorer is provided below with those non-default settings circled.



J.3 Test Results

Test results presented below are experiment outputs generated by WEKA explorer using the 10-fold stratified cross-validation method.

Attribution of Sources (35 attributes)

=== Summary ===

Correctly Classified Instances	658	84.359 %
Incorrectly Classified Instances	122	15.641 %
Kappa statistic	0.6644	
Mean absolute error	0.1564	
Root mean squared error	0.3955	
Relative absolute error	33.1194 %	
Root relative squared error	81.3935 %	
Total Number of Instances	780	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.894	0.238	0.859	0.894	0.876	0.828	No
	0.762	0.106	0.817	0.762	0.788	0.828	Yes
Weighted Avg.	0.844	0.188	0.843	0.844	0.842	0.828	

=== Confusion Matrix ===

```
  a   b   <-- classified as
 431  51 |   a = No
  71 227 |   b = Yes
```

Balance (35 attributes)

=== Summary ===

Correctly Classified Instances	611	78.3333 %
Incorrectly Classified Instances	169	21.6667 %
Kappa statistic	0.5116	
Mean absolute error	0.2704	
Root mean squared error	0.3496	
Relative absolute error	89.4134 %	
Root relative squared error	90.0008 %	
Total Number of Instances	780	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.857	0.356	0.82	0.857	0.838	0.751	Non-controversial Issue
	0.644	0.143	0.704	0.644	0.673	0.751	Controversial Issue - Noted
	0	0	0	0	0	?	Controversial Issue - Not Noted
Weighted Avg.	0.783	0.282	0.78	0.783	0.781	0.751	

=== Confusion Matrix ===

a	b	c	<-- classified as
437	73	0	a = Non-controversial Issue
96	174	0	b = Controversial Issue - Noted
0	0	0	c = Controversial Issue - Not Noted

Purpose (35 attributes)

=== Summary ===

Correctly Classified Instances	701	89.8718 %
Incorrectly Classified Instances	79	10.1282 %
Kappa statistic	0.579	
Mean absolute error	0.2457	
Root mean squared error	0.325	
Relative absolute error	217.7083 %	
Root relative squared error	137.7028 %	
Total Number of Instances	780	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.973	0.447	0.921	0.973	0.946	0.774	Educational/Informative
	0.654	0.026	0.746	0.654	0.697	0.855	Review
	0.25	0.001	0.8	0.25	0.381	0.848	Discussion Forum
	0.217	0.007	0.5	0.217	0.303	0.811	Reportage of Results
	0	0	0	0	0	0.48	Commercial
Weighted Avg.	0.899	0.38	0.884	0.899	0.886	0.783	

=== Confusion Matrix ===

a	b	c	d	e	<-- classified as
639	14	1	3	0	a = Educational/Informative
26	53	0	2	0	b = Review
12	0	4	0	0	c = Discussion Forum
14	4	0	5	0	d = Reportage of Results
3	0	0	0	0	e = Commercial

Publisher Credentials (35 attributes)

=== Summary ===

Correctly Classified Instances	568	72.8205 %
Incorrectly Classified Instances	212	27.1795 %
Kappa statistic	0.6561	
Mean absolute error	0.1761	
Root mean squared error	0.287	
Relative absolute error	99.7929 %	
Root relative squared error	96.7155 %	
Total Number of Instances	780	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.783	0.066	0.672	0.783	0.723	0.912	Cancer Organisation
	0	0.003	0	0	0	0.597	Clinician
	0.719	0.092	0.723	0.719	0.721	0.849	Commercial Body
	0.742	0.069	0.728	0.742	0.735	0.907	Consumer Group
	0.571	0.009	0.696	0.571	0.627	0.892	Educational Institution
	0.629	0.021	0.722	0.629	0.672	0.906	Government Organisation
	0.167	0.001	0.5	0.167	0.25	0.859	Lay Author
	0	0	0	0	0	?	Researcher
	0.776	0.081	0.783	0.776	0.779	0.888	Medical Organisation
Weighted Avg.	0.728	0.071	0.726	0.728	0.726	0.885	

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	i	<-- classified as
90	0	3	7	1	1	0	0	13	a = Cancer Organisation
0	0	2	1	0	1	0	0	0	b = Clinician
7	1	141	22	2	3	1	0	19	c = Commercial Body
9	0	22	115	1	2	0	0	6	d = Consumer Group
0	0	4	3	16	1	0	0	4	e = Educational Institution
8	1	6	4	0	39	0	0	4	f = Government Organisation
0	0	4	1	0	0	1	0	0	g = Lay Author
0	0	0	0	0	0	0	0	0	h = Researcher
20	0	13	5	3	7	0	0	166	i = Medical Organisation

Publisher Credentials (4 attributes)

=== Summary ===

Correctly Classified Instances	640	82.0513 %
Incorrectly Classified Instances	140	17.9487 %
Kappa statistic	0.7731	
Mean absolute error	0.1751	
Root mean squared error	0.2854	
Relative absolute error	99.2412 %	
Root relative squared error	96.1554 %	
Total Number of Instances	780	

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.852	0.033	0.817	0.852	0.834	0.952	Cancer Organisation
0	0	0	0	0	0.314	Clinician
0.842	0.074	0.793	0.842	0.817	0.915	Commercial Body
0.826	0.062	0.766	0.826	0.795	0.923	Consumer Group
0.821	0.009	0.767	0.821	0.793	0.932	Educational Institution
0.823	0.007	0.911	0.823	0.864	0.955	Government Organisation
0	0.001	0	0	0	0.51	Lay Author

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0	0	0	0	0	?	Researcher
	0.818	0.041	0.884	0.818	0.85	0.919	Medical Organisation
Weighted Avg.	0.821	0.048	0.814	0.821	0.816	0.921	

=== Confusion Matrix ===

a	b	c	d	e	f	g	h	i	<-- classified as
98	0	2	4	1	2	0	0	8	a = Cancer Organisation
0	0	3	0	0	1	0	0	0	b = Clinician
2	0	165	22	1	1	0	0	5	c = Commercial Body
6	0	17	128	0	0	1	0	3	d = Consumer Group
0	0	0	1	23	0	0	0	4	e = Educational Institution
2	0	6	0	0	51	0	0	3	f = Government Organisation
0	0	4	2	0	0	0	0	0	g = Lay Author
0	0	0	0	0	0	0	0	0	h = Researcher
12	0	11	10	5	1	0	0	175	i = Medical Organisation

Review Process (35 attributes)

=== Summary ===

Correctly Classified Instances	660	84.6154 %
Incorrectly Classified Instances	120	15.3846 %
Kappa statistic	0.383	
Mean absolute error	0.2601	
Root mean squared error	0.3342	
Relative absolute error	132.2259 %	
Root relative squared error	106.8427 %	
Total Number of Instances	780	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.945	0.626	0.882	0.945	0.912	0.657	Editorial Board
	0.423	0.028	0.6	0.423	0.496	0.861	Peer Review Process
	0.283	0.025	0.486	0.283	0.358	0.728	No Editorial/Peer Review Process
Weighted Avg.	0.846	0.525	0.826	0.846	0.832	0.681	

=== Confusion Matrix ===

a	b	c	<-- classified as
613	20	16	a = Editorial Board
39	30	2	b = Peer Review Process
43	0	17	c = No Editorial/Peer Review Process

Appendix K: Scenario Test Results

Note: Potentially problematic values are marked in grey colour

S1: without suggestions	**Machine**	**Medical and IM expert**	**IM expert**	**Medical expert**
Id: 1452				
Attribution of Sources	**No**	Yes	Yes	Yes
Balance	**Non-Controversial Issue**	**Controversial Issue - Noted**	**Controversial Issue - Not Noted**	**Non-Controversial Issue**
Purpose	Educational/Informative	Educational/Informative	**Discussion Forum**	Educational/Informative
Publisher Credentials	Consumer Group	Consumer Group	**Lay Author**	**Researcher**
Review Process	**Editorial Board**	No Editorial/Peer Review Process	No Editorial/Peer Review Process	No Editorial/Peer Review Process
Id: 1485				
Attribution of Sources	Yes	Yes	**No**	**No**
Balance	Controversial Issue - Noted	Controversial Issue - Noted	Controversial Issue - Noted	**Controversial Issue - Not Noted**
Purpose	**Educational/Informative**	Discussion Forum	**Educational/Informative**	Discussion Forum
Publisher Credentials	Commercial Body	Commercial Body	**Lay Author**	**Researcher**
Review Process	Editorial Board	**No Editorial/Peer Review Process**	Editorial Board	Editorial Board
No. of problematic values	**4**	**2**	**6**	**5**

S2: with suggestions	**Machine**	**Medical and IM expert**	**IM expert**	**Medical expert**
Id: 1507				
Attribution of Sources	**No**	**Yes**	**No**	**No**
Balance	Controversial Issue - Noted	Controversial Issue - Noted	Controversial Issue - Noted	Controversial Issue - Noted
Purpose	Educational/Informative	Educational/Informative	Educational/Informative	Educational/Informative
Publisher Credentials	Consumer Group	Consumer Group	Consumer Group	Consumer Group
Review Process	**Editorial Board**	**No Editorial/Peer Review Process**	**Editorial Board**	**Editorial Board**
Id: 1543				
Attribution of Sources	No	No	No	No
Balance	**Controversial Issue - Noted**	**Non-Controversial Issue**	**Controversial Issue - Noted**	**Controversial Issue - Noted**
Purpose	Educational/Informative	Educational/Informative	Educational/Informative	Educational/Informative
Publisher Credentials	Commercial Body	Commercial Body	Commercial Body	Commercial Body
Review Process	**Editorial Board**	**No Editorial/Peer Review Process**	**Editorial Board**	**Not Sure**
No. of problematic values	4	**4**	**4**	**4**

Appendix L: Tabulated Questionnaire Results

		Point assigned by (5-point Likert scale, where 1 means *Strongly disagree*, 2 means *Disagree*, 3 means *Neither disagree nor Agree*, 4 means *Agree*, and 5 means *Strongly agree*)			
Item code	**Aspect of the DED system for evaluation**	**Medical and IM expert**	**IM expert**	**Medical expert**	**Average**
E1	Data management views	5	5	5	5.0
E2	Browse data	4	4	5	4.3
E3	Add/Edit data entry (resource)	5	5	5	5.0
E4	Delete data	5	4	5	4.7
E5	Sort data	5	5	4	4.7
E6	Retrieve data	5	4	5	4.7
E7	Use without written instructions	2	5	2	3.0
E8	Recover from mistakes	4	5	5	4.7
U1	Usefulness of data visualisation	5	5	5	5.0
U2	URL checking and reporting	5	4	5	4.7
U3	Grouping metadata	4	3	5	4.0
U4	Metadata value suggestion	4	3	4	3.7
U5	Confidence in metadata quality	5	3	4	4.0
U6	Productivity and consistency	5	4	5	4.7
S1	Overall satisfaction	4	4	5	4.3
S2	Overall utility of the system	4	4	5	4.3
S3	Preference of use	4	5	3	4.0

Appendix M: Synthesis of DED Evaluation Interview Data

Feedbacks on the usefulness of DED system features

System feature	Medical and IM expert	IM expert	Medical expert
Data visualisation	*Positive feedback*: "The graphs view is so useful, because somebody can just glance [at] it and say 'ok, maybe we need [a] few more medical or [a] few more supportive resources'."	*Positive feedback*: "I think this is really helpful, as it's easier for [the] human brain to deal with diagrams than the textual [descriptions]. So it's good to have understanding of what resources are currently out there; what needs to be added. It was useful as - if I made a mistake, it will basically tell me on the diagram that I introduced some other category. Yeah, that's really good. Because it can pick up mistakes."	*Positive feedback*: "I was particularly impressed by the data visualisation. I think this is a really good way, to show me what's happening with the system. Especially when you are able to show me the number of resources for the audience of the disease stage, which I think is very good. And also this: saying that this is the number of what has been published, [or is] incomplete and unapproved. This is a very good way for me to understand…Even as simple as what is this work expecting from me as a domain expert or a content management person. What do I have to do? How well I'm going [to do] this sort of things. So this was a very good feature."
URL availability checking	*Positive feedback*: "If you are trying to maintain an online system […] given the volatility of things, it's very useful to know when a link is down. So [it] can be fixed immediately...in fact it's useful to grasp the immaturity."	*Positive feedback*: "The feature of checking URLs is fabulous for domain experts. Let's say once every month for example, you can check the broken links. It is important. Because otherwise I'll have to manually do it. I think it's great that you've got that one. I really like that."	*Positive feedback*: "Yes the URL checking is extremely helpful. It's extremely good for keeping it up to date, which I think is very important. Last thing you want is a dead link."

Metadata value suggestion	*Positive feedback*: "It is always a lot quicker and a lot easier if I can be reminded." *Negative feedback*: "At this pointing time, for me, I need to work from the source rather than computer generated suggestions. Because each resource is individual, each site is individual and even within those things, different things can happen depends on the nature of the resource that you are looking at."	*Positive feedback*: Yeah it's good if I know that I can fill it up and then I can compare to what's already out there, and if it can explain to me whether I'm wrong and why. It does make me go back and check so makes me analyse the resources more carefully probably. So, I do familiarise the resources and I feel better and critically evaluate them. *Negative feedback*: "I don't know... But if my opinion agrees with the opinion of the machine, well that's fabulous, but what if it's not. I'm in doubt then. I'm not sure like if I'm right or wrong. So, instead of throwing things to them but just explain to them what the system provides and how it can improve the experience, not necessarily straightforward."	*Positive feedback*: "With the suggestions, which definitely helped. And [with] most of them I was fine and most of them I agree."

Feedbacks on the utility of the DED system

Utility of the system	Medical and IM expert	IM expert	Medical expert
Supporting resource quality assessment	*Positive feedback*: "The only time I would do that is: if I'm thinking of a URL, I might look at the graph and say 'no we've got enough on breast reconstruction'." *Negative feedback*: "I haven't used the system for that purpose. The way I decide to put in a resource comes from all the documents that I'm searching from. It's independent of the system.	*Positive feedback*: "Yeah, all I can do to minimise my mistake would be good. And in particular, I think it's really good if there are few domain experts that are working on the same portal, instead of introducing the same resource in the system 2,3,4,5 times, it's fabulous if the system can say 'no' - that we already have the resource. So, you know things like that. So yeah, I think it's good. Then, I would agree."	*Positive feedback*: "The ways things have been categorised definitely helped. You can see why it's been included. You are looking at so many different attributes of that particular resource. I think you've got enough information there. That explains relevance. So that's why they are important. Yes I do agree the system does the job in a systematic manner."
Building medical expertise	*Positive feedback*: "In terms of the information, it certainly can help grows the medical expertise, because you learn the language. It's a learning curve and through repetition and familiarity. Hopefully, the knowledge base expends, increases, and improves."	*Negative feedback*: "Don't know, because it doesn't explain to me where I need to search for. You know there are certain things that can help me with like minimise my mistakes, but I don't think it can teach me about where and what I need to check, to fill the attributes in. Because it doesn't explain which one and why the decision was made"	*Positive feedback*: "I use to always stay away from the thinking that I could never deal with this sort of thing, because it is too complicated. But it has taught me quite a bit. And it has taught me that it is quite simple and it can be quite simple. And it's easy to understand what I can do and what I can't do."
Building IM expertise	*Positive feedback*: "IM expertise, yes probably because you know what you look for in the document or the resource."	*Positive feedback*: "Definitely the way information is presented and all the features are already there can help as long as you explain to the novice domain expert what stages he/she needs to go through in order to become more experienced. So take a look of what has been done by others and look at the publishers."	*Positive feedback*: "I think it does. I use to always think that things to do with metadata and content management are too complicated for me. If the system is to tell me to view the resources and add metadata or information about a resource, I can do it. It makes me capable of doing something."

GLOSSARY

Consumer Health Informatics

Also known as e-Health; the branch of health or medical informatics that analyses consumers' needs for information; studies and implements methods of making information accessible to consumers; and models and integrates consumers' preferences into medical information systems.

Source: Eysenbach, G. (2000), Recent advances: Consumer health informatics, British Medical Journal, vol. 320, no. 7251, pp. 1713-1716.

Content Management System (CMS)

A Content Management System (CMS) is the system that allows the institution to manage its local resources. Resources stored in the CMS may be presented through the institutional portal. An institution may operate more than one CMS in order to manage different kinds of content (Web pages, learning resources, e-prints, etc.).

Source: JISC (Joint Information Systems Committee) Information Environment Architecture Portal FAQ, http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/faq/portal/

Domain Experts

People with first hand experience and extensive knowledge of the medical, supportive and psychosocial information needs of the community; trained domain experts work for information portals also play the role as subject and metadata experts with librarians' skills in resource identification, selection and description.

Source: SIP (2006), Smart Information Portals: Meeting knowledge and decision support needs of health care consumers for quality online information, from http://www.sip.infotech.monash.edu.au/

Health Informatics

Health informatics (also called healthcare informatics, medical informatics, nursing informatics, or biomedical informatics) is a discipline at the intersection of computer science, information science, and healthcare. It deals with the resources, devices, and methods required to optimise the acquisition, storage, retrieval, and use of information in health and biomedicine. Health information tools include not only computers but also clinical guidelines, formal medical terminologies, and information and communication systems.

Source: Wikipedia, http://en.wikipedia.org/wiki/Health_informatics

Information Quality (IQ)

The ability of a piece of information to meet user requirements.

Source: Tayi, G. K. and Ballou, D. P. (1998), Examining data quality, *Communications of the ACM*, vol. 41, no. 2, pp. 54-57.

IQ of Online Health Information

The degree to which web-based health information positively affects a user's health outcomes, quality of life, or disease-specific clinical end points.

Source: Risk, A. and Petersen, C. (2002), 'Health Information on the Internet: Quality Issues and International Initiatives', *JAMA*, vol. 287, no. 20, pp. 2713-2715.

Intelligent Technologies

In the context of this research, technologies that can help domain experts in solving complex decision-making problems are regarded as intelligent technologies. Varying types of intelligent technologies have been used to capture different aspects of expert decision-making, such as probabilistic techniques, fuzzy logics, neural networks, classifiers and statistical learning.

Source: Kreinovich V., Nguyen H. T., Prasad N. S. and Santiprabhob P. (2004), Intelligent technologies: An introduction, *International Journal of Intelligent Systems*, vol. 19, no. 1-2, pp. 1-8.

Metadata

Data about data; data definitions describing aspects of actual data items, or the context or semantics of using the data. In the context of web search, metadata plays an essential role in finding information and understanding its structure, provenance, relevance, and quality.

Source: Henninger, M. (2008), The hidden web: finding quality information on the net, 2nd, University of New South Wales Press, Sydney, N.S

Jeffery K. (2000), Metadata: The future of information systems, in Brinkkemper, S., Lindencrona, E. and Sølvberg, A. (eds.) *Information Systems Engineering: State of the Art and Research Themes*, London: Springer-Verlag, paper presented at the Information Systems Engineering Symposium, Stockholm, Sweden.

Ontology

Ontology is an old world from philosophy. An ontology is a theory about the nature of existence, of what types of things exist; Metadata. In computer science, ontology is a structured data model of concepts and relationships between those concepts. It may include a set of terms and relationships among those terms that used to describe and represent the area of knowledge. The most widely cited definition of ontology is "a specification of a conceptualization.

Source: Berners-Lee, T., Hendler, J. and Lassila, O. (2001), The semantic web, *Scientific American*, vol. 284, no. 5, pp. 34.

Gruber, T. R. (1993), A translation approach to portable ontology specifications, *Knowledge Acquisition*, vol. 5, no. 2, pp. 199-220.

Quality Control

Quality control of online health information is emerging from consumer health informatics as a research field to study the determinants and distribution of health information and misinformation on the web, in order to protect health information consumers from fraudulent or harmful online information.

Source: Eysenbach G. (2002), Infodemiology: The epidemiology of (mis)information, *The American Journal of Medicine*, vol. 113, no. 9, pp. 763-765.

Eysenbach G. (2005), Design and evaluation of consumer health information web sites, in Lewis, D. (ed.) *Consumer Health Informatics: Informing Consumers and Improving Health Care*, Springer, New York, pp. 34-60.

Quality Tool

In this thesis, a quality tool refers to a computer system that supports and facilitates decision-making processes on the quality of online information resources. As a quality tool is devised mainly for supporting resource quality assessment, it can also be termed as a quality assessment tool.

Resource Quality (RQ)

In this thesis, RQ is defined as the extent to which information contained in a web-based resource meets the information needs and quality perceptions of individual users.

Semantic Web

An evolving extension of World Wide Web, which makes web resources not only readable by human-being, but also meaningful to computers, software agents in particular, to find, share, exchange and integrate data automatically, efficiently and effectively. The Semantic Web is about two things. It is about common formats for integration and combination of data drawn from diverse sources. It is also about language for recording how the data relates to real world objects.

Source: W3C, http://www.w3.org/2001/sw/

User-Centred Design

User centred design is a multi-disciplinary activity, which incorporates human factors and ergonomics knowledge and techniques with the objective of enhancing effectiveness and productivity, improving human working conditions, and counteracting the possible adverse effects of use on human health, safety and performance.

Source: Bevan N. (1999), Quality in use: Meeting user needs for quality, *Journal of Systems and Software*, vol. 49, no. 1, pp. 89-96.

User-Sensitive Design

User-sensitive design is both a philosophy and a methodology that integrates user information needs analysis (Wilson 1981; Wilson 1994), user-centred design (Norman and Draper 1986; Preece et al. 2002) and value-sensitive (Friedman et al. 2006) approaches, and systems development research techniques within an interpretivist framework. The philosophical view of user-sensitive places user's needs and values at the heart of system design and development.

Source: McKemmish S., Manaszewicz R., Burstein F. and Fisher J. (2009), Consumer empowerment through metadata-based quality reporting: The breast cancer knowledge online portal, *Journal of the American Society for Information Science and Technology (JASIST)*, vol. 60, no. 9, pp. 1792-1807.

User Satisfaction

Satisfaction is described as a combination of comfort and acceptability of use: Comfort refers to overall physiological or emotional response to user of the system (whether the user feels good, warm, and pleased, or tense and uncomfortable). Acceptability of use may measure overall attitude towards the system, or the user's perception of specific aspect, such as whether the user feels that the system supports the way they carry out their tasks, do they feel in command of the system, is the system helpful and easy to learn.

Source: Bevan N. (1999), Quality in use: Meeting user needs for quality, *Journal of Systems and Software*, vol. 49, no. 1, pp. 89-96.

Web

A shortened form of World Wide Web, abbreviated as WWW, is a system of interlinked hypertext documents accessed via the Internet. With a web browser, one can view web pages that may contain text, images, videos, and other multimedia and navigate between them via hyperlinks.

Source: Wikipedia, http://en.wikipedia.org/wiki/World_Wide_Web

Web 2.0

Active client-side applications that use all kinds of data on the web and provide intelligent data sharing features.

Source: O'Reilly T. (2005), What is web 2.0: Design patterns and business models for the next generation of software, http://oreilly.com/web2/archive/what-is-web-20.html.

Web Information Portal

A gateway to a virtual, distributed knowledge repository of information resources, which have been selected, evaluated and described by trained domain experts with the explicit aim of providing customised information of high quality and relevance to a community of users.

Portals or gateways are web applications that give users immediate access to descriptions of web resources, and control over when and how they share information.

Source: Cooke, A. (2001), *A guide to finding quality information on the Internet: selection and evaluation strategies*, 2nd, Library Association, London.

Eboueya M. and Uden L. (2007), Benefits and limitations of portals, in Tatnall A. (ed.) Encyclopedia of portal technologies and applications, Information Science Reference, pp. 75-81.

SIP (2006), Research Proposal of Smart Information Portals: Meeting knowledge and decision support needs of healthcare consumers for quality online information.

Webpage

A document or information resource that is suitable for the World Wide Web and can be accessed through a web browser and displayed on a monitor or mobile device. This information is usually in HTML or XHTML format, and may provide navigation to other web pages via hypertext links. Web pages frequently subsume other resources such as style sheets, scripts and images into their final presentation.

Source: Wikipedia, http://en.wikipedia.org/wiki/Webpage

Website

A collection of related web pages containing images, videos or other digital assets. A web site is hosted on at least one web server, accessible via a network such as the Internet or a private local area network through an Internet address also called URL.

Source: Wikipedia, http://en.wikipedia.org/wiki/Website

Web Link Analysis

Web search ranking algorithms that use eigenvector-based centrality metrics, including Google's PageRank, Kleinberg's HITS algorithm, and the TrustRank algorithm. Link analysis is also conducted in information science and communication science in order to understand and extract information from the structure of collections of web pages. For example the analysis might be of the interlinking between politicians' web sites or Blogs.

Source: Wikipedia, http://en.wikipedia.org/wiki/Network_theory#Web_link_analysis