University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

12-2010

Information Quality on the World Wide Web: A User Perspective

Jaikrit Kandari
University of Nebraska-Lincoln, jaikritkandari@huskers.unl.edu

Follow this and additional works at: https://digitalcommons.unl.edu/imsediss

Part of the Library and Information Science Commons, Management Information Systems Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Other Engineering Commons

INFORMATION QUALITY ON THE WORLD WIDE WEB: A USER PERSPECTIVE

by

Jaikrit Singh Kandari

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Industrial and Management Systems Engineering

Under the Supervision of Professor Ramaratnam Ram Bishu

Lincoln, Nebraska

December, 2010

INFORMATION QUALITY ON THE WORLD WIDE WEB: A USER PERSPECTIVE

Jaikrit Singh Kandari, M.S.

University of Nebraska, 2010

Advisor: Ramaratnam Ram Bishu

Information Quality has emerged as an important measure for the success of Information Systems. At the same time the World Wide Web has established itself as the key infrastructure for information administration, exchange, and publication. Users are getting information from the web at the click of a button however they must filter sub-standard information before they can use quality content. Researchers have aimed to address this problem by suggesting various information quality frameworks. Research till date has seen twenty important Information Quality (IQ) frameworks emerge. These models though varied in their approach and application share a number of characteristics regarding their classifications of the attributes of quality. A new framework for the measurement of information quality is developed and twenty two information quality dimensions are identified for measuring information quality in context of the web from a user perspective. An online survey instrument is used for data collection. The research argues that WWW is not a homogeneous entity and should be understood from individual aspects of three independent variables of web domain, type of website and end user nationality to arrive at its conclusion. Results highlight nine IQ dimensions which are important across the whole web environment, while thirteen dimensions have contextual importance and vary across web domain and national culture.

Keywords: Information Quality (IQ), IQ Frameworks, survey instrument, reliability

To Ma and Papa

ACKNOWEDGEMENTS

Table of Contents

Abstract

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

Information Systems (IS) field has seen exponential growth both in terms of scope and complexity. Much of this can be attributed to the numerous disruptive technology innovations that have skewed IS's progress graph unlike any known field. One invention in information systems that has contributed hugely to its tremendous growth is the World Wide Web. The internet provides users easy means of interactive communication which was not possible before. Anyone can publish information on the web by simply acquiring space on a website and creating an electronic document (Herrera-Viedma et al. 2006). The scale and reach of published information on the Web dwarfs that of the print world. In the process the internet has become the largest available repository of data with the largest number of visitors searching for information (Herrera-Viedma et al. 2006).

People who need information to accomplish their tasks are finally being provided with easy online access to relevant information (Strong et al. 1997b). However, these information consumers must take into account the fact that this information is not governed by any set of standards and may not have passed the eyes of any editor. There are no rules on the type and quality of information which can be published on the internet. Hence information or data consumers have to make their own decisions regarding the quality of information before using it for their own needs. This research is aimed to understand the consumer's perspective in setting and measuring those quality standards while handling information or data on the World Wide Web.

"Data" usually refers to information at its early stages of processing and "information", the product at a later stage (Strong et al. 1997b). In the context of this

research the term "information" refers to both data and information and has been used interchangeably. Following general quality literature, Wang and Strong (1996) described information quality (IQ) as data that is 'fit-for use' by data consumers. They also described 'Data quality dimension' as a set of data quality attributes that represent a single aspect or construct of data quality. The work done by Wang and Strong in IQ is very exhaustive and extensive (Matheus 2004). Subsequent research in the field has seen some researchers adapt and expand on their work while others have taken a different approach to show their perspective of IQ. This research contends that there is a thread of commonality amongst all these existing frameworks.

This research article attempts to arrive at a convergence of ideas by identifying the common data quality dimensions which are prevalent in existing literature. It then aims to understand the significance of IQ dimensions specifically in context of the World Wide Web. The research seeks to understand the behavior of these IQ dimensions across different web domains and nationalities. The authors attempt to expand the boundaries of existing literature by trying to identify new and unknown IQ dimension(s) that could emerge as specific measures of IQ in the web environment.

This thesis is the culmination of an extensive research which was undertaken to understand IQ on the World Wide Web from a user perspective. The complete research was documented in two journal articles. The first paper is listed as chapter 2 in the thesis. It details the development of the framework which is used by the authors for their research. Within the first paper "Information Systems" gives an overview of information systems and traces its evolution and growing complexity. DeLone and McLean (2003) model was explained to highlight the importance of IQ in IS and its relevance in the web

environment. "Information quality" section is presented in the 'fit-for-use' context and literature review of existing IQ frameworks is done. "Development of the Framework" section details the five steps from identifying the threads of commonality in existing frameworks and converging on 22 IQ dimensions to the development of survey instrument and running a pilot to check its internal consistency and reliability. The first paper and chapter 2 conclude with the development of framework and instrument phase of the research.

The second paper which is prepared for submission is listed as chapter 3 in the thesis. It takes off from where the first paper culminated. "Theoretical Background and Research Objectives" section briefly explains the framework and survey development in phase one of the research. It also reiterated the research objectives. "Research Methodology" section explains the experimental design, selection of the levels for each factor in the research and data collection. "Results" section details the statistical analysis done on SAS. Each result is described and also depicted using graphs. The section also explains how the results have been divided into various quadrants for easier grouping and interpretation. The last section is the "Discussion" which talks results and interpret the result. It also aims to understand the bigger canvas as to how the research is significant to the field of IQ and its contribution. At the same time, it discusses the future questions for additional research.

Chapter 4- "Discussion" - concludes the thesis. This chapter reviews the complete research in its entirety. Some parts of the research which were not included in the two papers due to space and relevance constraints have been discussed in this chapter.

CHAPTER 2

JOURNAL PAPER 1

Information Quality on the World Wide Web: Development of a Framework

Jaikrit Singh Kandari

Erick C. Jones

Fiona F. Nah

Ram R. Bishu

University of Nebraska- Lincoln

Abstract

Data consumers are provided with easy online access to information on the World Wide Web. However, consumers face information quality problems in their quest for information. This paper focuses on the development of an instrument to measure IQ on the World Wide Web from a user's perspective. Based on a comprehensive review of the literature, twenty important Information Quality (IQ) Frameworks were identified. These models, though varied in their approach and application, share a number of characteristics regarding their classifications of the attributes of quality. The paper identifies common dimensions that exist across the existing IQ frameworks in the literature and develops a unified comprehensive framework for the measurement of IQ based on the identified thread of commonality and the intuitive approach. A survey instrument was developed and fine-tuned using iterative cognitive interview process. The proposed survey instrument comprises of 73 questions to measure 22 IQ dimensions.

Keywords: Information Quality (IQ), IQ Frameworks, survey instrument, reliability

Information Quality on the World Wide Web: Development of a Framework

INTRODUCTION

Information Systems (IS) field has seen exponential growth both in terms of scope and complexity. Much of this can be attributed to the numerous disruptive technology innovations that have skewed IS's progress graph unlike any known field. In 1943, IBM chairman Thomas Watson predicted a world market of 5 computers. Bill Gates in 1981 thought, "640 K should be enough for anybody". Today the field has gone far ahead of its data processing days to a world where the computers are 'personal', 'wi-fi', 'wireless' and 'networked'.

One invention in information systems that has contributed hugely to its explosive growth is the World Wide Web. The internet provides users easy means of interactive communication which was not possible before. Anyone can publish information on the web by simply acquiring space on a website and creating an electronic document (Herrera-Viedma et al. 2006). The scale and reach of published information on the Web dwarfs that of the print world. In the process the internet has become the largest available repository of data with the largest number of visitors searching for information (Herrera-Viedma et al. 2006).

People who need information to accomplish their tasks are finally being provided with easy online access to relevant information (Strong et al. 1997b). However there are neither rules nor standards governing the type and quality of information that a writer can put on the Web (Diligenti, Gori, & Maggine, 2004). Information consumers have to make their own decision about the quality of information before using it for their needs.

Wang and Strong (1996), in their seminal work, attempted to address these concerns about Information Quality (IQ). They applied general quality literature and described IQ as data that is 'fit-for use' by data consumers. They also described 'Data quality dimension' as a set of data quality attributes that represent a single aspect or construct of data quality. "Data" usually refers to information at its early stages of processing and "information", the product at a later stage (Strong et al. 1997b). In the context of this research the term "information" refers to both data and information and has been used interchangeably.

The work done by Wang and Strong in IQ is very exhaustive and extensive (Matheus 2004). Subsequent research in the field has seen some researchers adapt and expand on their work while others have taken a different approach to show their perspective of IQ. *This research aims to identify IQ dimensions and study their significance levels across different types of websites in individual web domains and across different nationalities.* This article details the development of a framework to measure the quality of information on the Web and subsequent development of a survey instrument for data collection

INFORMATION SYSTEMS

OVERVIEW

Information systems (IS) as defined by Davis (1999) is an umbrella term which encompasses information technology (IT) systems and applications for transactions and operations, support of administrative and management functions, organizational communication and coordination, and for adding value to products and services. Watson (2007) defines it as a socio-technical system comprised of two sub-systems: a technical

sub-system and a social sub-system. This definition is demonstrated in Figure 1. The technical sub-system is comprised of:

'*Information technology*' – it includes hardware, software and telecommunication equipment that is used to capture, process, store and distribute information. '*Process*' - maps the set of actions that an individual, a group or an organization must employ to carry out a specific business or organizational activity.

The social sub-system encompasses:

'*People*'- includes all individuals directly involved with the system. They include managers, who define the goals of the system and the end users. '*Organizational Structure*' - refers to the relationship among individuals in the people component and encompasses the hierarchical, reporting and rewards systems

Social System	Technical System
Structure	Technology
People	Process
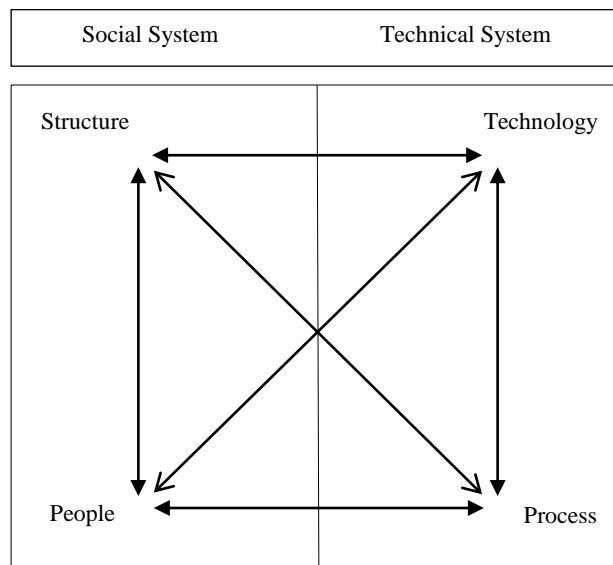
Figure 1: The Socio-Technical System (Watson, 2007)

HISTORY

Davis (1999) notes that organizations first started using computer in the mid-1950s, primarily for electronic data processing (EDP), which was simple recording, classifying, manipulating and summarizing of transaction records. By the mid- 1960s the term, management information system (MIS) had been coined and was used to define the comprehensive information processing that computer and IT systems could do for organizations. It enlarged the scope of data processing to add systems for supporting management and administrative activities including planning, scheduling, analysis, and decision making.

Around 1970's the predefined management reports proved insufficient to meet many of the decision-making needs of the management and thus decision support system was born. It provided interactive ad hoc support for the decision-making processes of managers and other business professionals. Introduction of microcomputers in to the work place by 1980's ushered IS in to an era of expert systems and knowledge management systems. The systems were capable of supporting the creation, organization and dissemination of business knowledge within the enterprise.

Mid to late 1990s saw an emergence of enterprise resource planning (ERP) which extended the use of IT beyond internal networks to integrate all facets of an organization to include its planning, manufacturing, sales, resource management, customer relations, inventory control, order tracking, financial management, human resource and marketing. Davis (1999) says, "Innovative applications based on IT created value by providing services any time, at any location, and with extensive customization. Web-based communication and transaction applications became common".

THE COMPLEXITY

With growth comes complexity and by the 1990's IT-based systems were employed to change organizational structures and processes. However, frequently they were being used without the full understanding of its applicability, effectiveness or efficacy (Myers et al. 1997). The production of IS function was proven difficult to define and measure (Scudder & Kucis, 1991) and assessing the value of IT infrastructure became the biggest single problem for the 90s.

Researchers started discussing the need to assess the contribution of IS function in late 1970s (King & Rodriguez 1978; Matlin 1977). Early focus was on the economic aspects and centered on measures of systems availability and performance (Borovits & Neumann 1979, Zmud 1979, Ives & Olson 1984). It was however McLean (1973) who called for a shift from a measurement focus on efficiency to effectiveness. It required computer professionals to measure and pursue organizational objectives, in addition to pursuing their internal departmental goals. He differentiated between efficiency and effectiveness thus: "Efficiency is concerned with doing things right; effectiveness is concerned with doing the right things" (McLean 1973).

In 1992 DeLone & McLean suggested that researchers should "systematically combine individual measures from the IS success categories to create comprehensive measurement instrument". The six dimensions in their model share a dependent relationship as well and temporal and causal relationships. The authors contend that 'system quality' and 'information quality' singularly and jointly affect both 'use' and 'user satisfaction'". Also both are direct antecedents of 'individual impact' which

ultimately has some 'organizational impact' (DeLone & McLean, 1992). The model is shown in Figure 2.
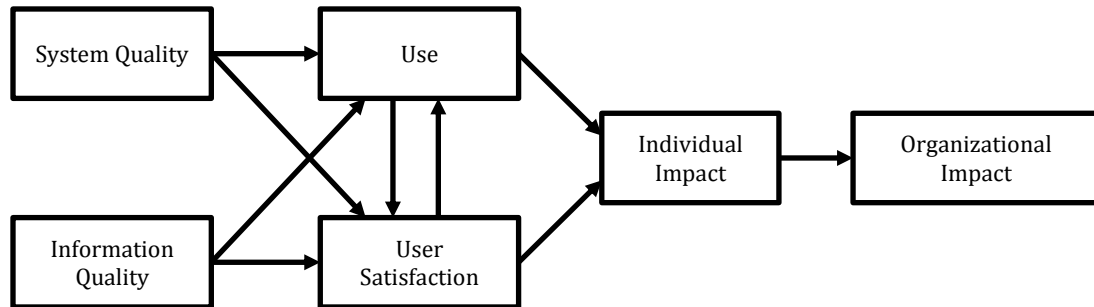
Figure 2: DeLone and McLean IS Success Model (1992)

In the DeLone and McLean model system quality refers to 'technical level measures' such as reliability of the computer system, online response time, ease of use, response time and system accuracy. Information quality refers to 'meaning level' of IS output in terms of accuracy, relevance, timeliness, adaptability and accessibility. Use is measured as reported by the users or the actual use as reported by the system in terms of queries by time, connect time or number of computer functions utilized. User satisfaction refers to measures of how the information affects the user. Individual impact deals with how the information system modifies the user's experience with the system while organizational impact contains measures about how the systems and the information provided influence the organization.

DeLone & McLean (1992) model, though not without criticism, has seen the highest acceptance by researchers. It has been cited by Wang & Strong (1996) in their literature search to justify using 'information quality' and 'user satisfaction' as the foundation of their research. This research borrows extensively from both these works in Information

Systems and Information Quality to try and understand *Information Quality from a user perspective on the World Wide Web.*

INFORMATION QUALITY

Information quality (IQ) is commonly described in the literature as a multi-dimensional concept (Ballou et al. 1998; Klein, 2001; Aladwani et al. 2002; Gendrone et al., 2004). Data Quality (DQ) is another term which is often used synonymously with IQ and is described as data that is 'fit-for-use' (Wang & Strong, 1996). Tayi & Ballou (1998) too reasoned that since IQ is relative, information considered useful for one person may not be 'fit' for another person's use.

The 'fit-for-use' model is widely adopted in quality literature and emphasizes the importance of taking a consumer's viewpoint of quality because ultimately it is the consumer who will make a judgment about the product's "fitness-for-use" (Deming 1986, Juran 1989, Juran & Gryna 1980). The model has been well received by researchers working in the field of IQ. Strong et al. (1997a) contend that this definition gives IQ a context. Shankar & Watts (2003) point out that the reason for a contextual approach is both simple and logical, because it recognizes that the attributes and dimensions used to assess IQ can vary depending on the context in which the data is to be used.

INFORMATION QUALITY FRAMEWORKS

The view that, "data cannot be assessed independent of the people who use the data" (Strong et al. 1997a), is currently the most widely accepted in quality literature. Shankar & Watts (2003), too caution against defining quality using frameworks without a context.

This article reviews major IQ frameworks collated from 1996 to 2006. While varied in their approach and application, these frameworks share distinct characteristics regarding ultimate classifications of the IQ dimensions (Knight 2008). The review is done in a chronological manner, with few exceptions when extended or derivative models have been discussed together. The chronological list of frameworks by year, author and IQ model are shown in Table 1.

#	Year	Author(s)	IQ Framework
1	1996	Wang & Strong	A Conceptual Framework for Data Quality
2		Zeist & Hendricks	Extended ISO Model
3	1997	Beck	Evaluation Criteria for web information sources
4		Harris	User-focused checklist (CARS) to help researchers look for clues regarding website IQ
5	1999	Alexander & Tate	Applying a Quality Framework to Web Environment
6		Katerattanakul & Siau	IQ of Individual Web Site
7		Shanks & Corbitt	Semiotic-based Framework for Data Quality
8	2000	Dedeke	Conceptual Framework for measuring IS Quality
9		Naumann & Rolker	Classification of IQ Metadata Criteria
10		Zhu & Gauch	Quality metrics for information retrieval on the WWW
11	2001	Leung	Adapted Extended ISO Model for Intranets
12	2002	Kahn, Strong & Wang	Mapping IQ dimension into the PSP/IQ Model
13		Liu &Chi	Evolutional Data Quality
14		Eppler & Muenzenmayer	Conceptual Framework for IQ in the Website
15		Klein	5 IQ Dimensions
16	2003	Shankar & Watts	Theoretical Model for Data Quality Assessment
17		Sturges & Griffin	Tool for Archaeological website quality evaluation
18	2004	Tombros et al.	5 dimensions for judging quality in web pages
19	2005	Stvilia, et al.	Application of 7 known IQ metrics to automated system tool to measure IQ of Wikipedia content
20	2006	Song & Zahedi	IQ dimensions that influence users judgements of Web-Based Health informediaries

Table 1: Chronological List of IQ Frameworks - Adapted from Knight (2008)

This research credits Wang & Strong (1996) for being the pioneers in the field of IQ. However it acknowledges that many more researchers like Zeist & Hendricks (1996),

Beck (1997), Harris (1997), Alexander & Tate (1999) and Shanks & Corbitt (1999) made significant first contributions to the field. Katerattanakul & Siau (1999), Leung (2001) and Klein (2002) were among the early adopters. Others like Dedeke (2000), Naumann & Rolker (2000), Zhu & Gauch (2000) and Liu & Chi (2002) gave new perspectives while Kahn, Strong & Wang (2002) and Shankar & Watts (2003) extended earlier works. Recent works like Sturges & Griffin (2003), Stvilia, et al. (2005) and Song & Zahedi (2006) have focused more closely on specific domains.

Wang and Strong (1996) in their seminal work postulated the contextual IQ paradigm. They stated four IQ areas. The first- 'intrinsic data quality'- indicates that information has quality in its own right. It includes: accuracy, objectivity, believability and reputation. The second- 'contextual data quality'- requires that information should be provided on time and in appropriate amounts. It includes: relevancy, value-added, timeliness, completeness and appropriate amount of data. The third- 'representational data quality' comprises aspects related to the format of the information and its meaning. It includes: interpretability, ease of understanding, representational consistency and concise representation. Finally the fourth - 'accessibility data quality' emphasizes that information on the web must be easily accessible but secure. It includes: accessibility and access security.

Around the same time Zeist & Hendricks (1996) presented the 'Extended ISO Model' which identified six IQ characteristics and their respective sub-characteristics. The IQ characteristic 'Functionality' includes sub-characteristics of suitability, accuracy, interoperability, compliance, security and traceability of information. Similarly 'Reliability', includes maturity, recoverability, availability, degradability and fault

tolerance of the content. 'Efficiency' of the webpage content investigates the time and resource behavior. 'Usability' includes the understandability, learnability, operability, luxury, clarity, helpfulness, explicitness, customizability and user-friendliness characteristics of information. 'Maintainability' pertains to the analyzability, changeability, stability, testability, manageability and the reusability of content while lastly 'Portability' is the adaptability, conformance, replaceability and installability of information. Leung (2001) adapted the work of Zeist & Hendricks (1996) and introduced the Adapted Extended ISO Model for Intranets. He defined IQ dimensions using the same set of characteristics and sub-characteristics but in the context of intranet environment.

The period of late 1990s saw application of IQ guidelines to build user-resources and 'how to' frameworks for the searchers of information (Knight, 2008). This was specifically directed to users of the World Wide Web. Notable frameworks were "CARS Checklist for Information Quality" (Harris, 1997), Web Evaluation Criteria (Beck 1997) and Web Wisdom (Alexander and Tate 1999). Some criteria which kept showing up and re-enforcing their importance were accuracy, objectivity and currency.

Shanks & Corbitt (1999) conceptualized a semiotic-based IQ framework. They looked at the quality of data from a cultural aspect by defining their quality dimensions in terms of socially understood constructs. Their proposed four semiotic levels are 1, 'syntactic'- when web pages should be consistent, 2, 'semantic'- ensures that information on the web pages is complete and accurate, 3, 'pragmatic'- warrants that the content on the website must be usable and useful and 4, 'social'- ensures shared understanding of meaning and an awareness of biasness on the webpage. However in subsequent semiotic approaches (Price & Shanks, 2004, 2005) the 'social' construct was removed.

Katerattanakul & Siau (1999) described four IQ categories of individual websites adapted from the dimensions by authors Wang & Strong (1996). The 'intrinsic' category ensures the accuracy and free-of-error webpage content. It includes accurate, workable and relevant hyperlinks on the webpage. 'Contextual' category warrants provision of the author's information. 'Representational' information quality refers to the organization, visual settings, typographical features, consistency, vividness and attractiveness of the webpage. 'Accessibility' ensures the navigational tools used to access and move around on the website.

Dedeke (2000), identified quality characteristics in an electronic systems environment. His data quality framework included five categories namely; ergonomic, accessible, transactional, contextual and representational where ergonomic category deals with the ease of navigation on the webpage. Accessibility quality ensures information accessibility, sharing and technical access. Transactional category is the responsiveness of a webpage, its error tolerance, efficiency and adaptability of the content. Contextual category ensures relevancy, completeness, appropriateness and timeliness of webpage content while representational quality is the consistency, conciseness, structure, interpretability, readability and contrast of the information on the webpage.

IQ criteria as defined by Naumann & Rolker (2000) included subject, object and process criteria. Subject criteria include believability, concise representation, understandability, value addition, interpretability and relevancy of information on the website. Objective criteria aim to ensure that the webpage is complete, secure, objective, timely and verifiable. Process criteria include dimensions like accuracy, response time and consistent representation. The same year Zhu & Gauch (2000) proposed a quality

metrics for information retrieval on the World Wide Web. The suggested metrics included the availability metric, the authority metric, the currency metric, information to noise ratio, cohesiveness and the popularity metrics. All the above metrics were specific to measure data quality on the webpages.

Kahn et al. (2002) introduced the mapping of IQ dimensions into the PSP/IQ Model. IQ was categorized in the context of the web by the authors. Two basic quality types were defined; the 'product quality', which was classified further as sound information and useful information and 'service quality' classified as dependable information and useable information. Sound Information includes free-of-error, concise, representation, completeness and consistent representation of information on the webpage. Relevant Information includes appropriate amount of information, relevancy, understandability, and interpretability and accuracy. Dependable Information includes timeliness, security while useable Information includes believability, accessibility, reputation, value-addition and ease of manipulation.

Liu & Chi (2002) proposed the "Evolutional Data Quality" framework which was primarily built on the foundation of Wang & Strong's (1996) four category IQ model. The model conceptualizes the process of user/information interaction into a cycle that separates IQ into two contexts, one, information production and two, information use. The four quality types are data collection, data organization, data presentation and data application. Collection includes IQ dimensions like accuracy, objectivity, trustworthiness, completeness and clarity. Organization includes reliability, consistency, storage efficiency, retrieval efficiency and navigability. Presentation includes IQ dimensions like semantic stability, faithfulness, neutrality and interpretability while

Application includes IQ dimensions like ease of manipulation, timeliness, privacy, relevancy and appropriate amount of data.

Eppler & Muenzenmayer (2002) subdivided their suggested framework into content and media quality. The content quality is concerned about the quality of the information presented on the web; it advices that the webpage content should include comprehensive, accurate, clear and applicable information. For sound information web authors must ensure that the information on the website is concise, consistent, correct and current. Media quality on the other hand is concerned about the quality of the medium used to deliver the web content. It includes convenience, timeliness, traceability and interaction of the webpage. Other quality criteria are accessibility, security, retrieval speed of the webpage and maintainability.

Klein (2002) adopted the user-driven, consumption model of Wang & Strong (1996). Her research focused on how often users encountered the IQ problems and how encountering the problem impacted their perception of the source's IQ. She identified five key IQ dimensions in the context of the web. They were namely accuracy, amount of data, completeness, relevance and timeliness.

In the last few years more models have been suggested. Shankar & Watts (2003), suggested that accuracy, completeness, timeliness, believability and relevance are the core factors while discussing a theoretical model for data quality assessment. Tombros et al. (2004) suggested five dimensions for judging quality in web pages and included IQ as one of the aspects of their identified web features. Sturges & Griffin (2003), Stvilia et al. (2005) and Song & Zahedi (2006) have contributed contextual models in the fields of archeological website quality, Wikipedia content and web-based health infomediaries

respectively. Liu & Huang (2005) in their work mention key dimensions like source, content, format and presentation, currency, accuracy and speed.

In the above review it is evident that a thread of commonality has existed in the kind of dimensions being used to measure the IQ. Some of the dimensions keep re-enforcing their importance irrespective of the context, while few are highly relevant to a particular field. This research looks at all the available framework, investigates the commonality running between them to arrive at a set of dimensions and a unique framework to measure the IQ on the World Wide Web from a user perspective.

RESEARCH OBJECTIVE

This research aims to understand IQ from a context of the World Wide Web. While some researchers have looked at this before there is no consensus over the IQ dimensions which are important in context of the web. At the same time there is some commonality in previous findings. The authors also argue that the web environment is not a homogeneous entity and each sub-group should be considered in its individual context. The study looks at three factors to study IQ in the web context from a user perspective. The three factors are: one, the web domain, two, individual websites within a web domain and three, end-user nationality. This research seeks to answer the following research questions:

1. Which of the IQ dimensions are relevant in the context of World Wide Web from the user perspective?

2. Do IQ dimensions behave differently across individual web domains of the WWW?

3. Do IQ dimensions behave differently across different websites within individual web domains?

4. Can different national cultures lead to varying IQ dimensions?

The scope of this article is limited to the development of the framework and survey instrument for data collection. The research design and findings are discussed in future works.

DEVELOPMENT OF THE FRAMEWORK

Review of IQ frameworks shows a definite commonality amongst the 20 major IQ models. This research investigates the degree of overlap in the various models to propose a new framework to measure IQ on the World Wide Web. The steps involved in the development of the framework are listed below and discussed in detail subsequently:

STEP 1: Finding common dimensions.

STEP 2: Finalize and define dimensions in context of the World Wide Web.

STEP 3: Development of the survey instrument.

STEP 1: FINDING COMMON DIMENSIONS

Wang & Strong (1996) in their seminal work postulated the contextual IQ paradigm. Their research was aimed to determine quality characteristics of data, from a data consumer's perspective. Their first survey generated an extensive list of 179 potential quality attributes which were evaluated using importance ratings, exploratory factor analysis and sorting study. The final framework proposed 15 IQ dimensions to measure data quality.

Subsequent researchers in the field of IQ have either based their work on the Wang and Strong model or introduced a new paradigm in measuring IQ. This study after extensive literature review identified twenty major frameworks in the field of information quality between 1996 and 2006. Table 2 provides a summary of the most common dimensions and the frequency with which they have appeared in the twenty IQ frameworks. It is interesting to note that all 15 dimensions proposed by Wang and Strong make the list. Timeliness appeared in 18 out of 20 frameworks. Accuracy, Accessibility, Amount of Data, Believability, Consistent Representation, Completeness, Objectivity and Relevancy showed their presence in 10 out of 20 frameworks. Usability and Usefulness with their presence 3 and 2 times in the list were at the bottom in the frequency table.

DIMENSIONS	Wang and Strong	Zeist & Hendricks	Beck	Harris	Alexander &Tate	Katerattanakul & Siau	Shanks & Corbitt	Dedeke	Naumann & Rolker	Zhu & Gauch	Leung	Kahn et al	Liu & Chi	Eppler & Muenzenmayer	Klein	Shankar &Watts	Sturges & Griffin	Tombros et al.	Stvilia et el	Song & Zahedi	TOTAL
Accuracy	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X			X	17
Accessibility	X	X			X	X	X	X	X	X	X	X		X						X	12
Amount of Data	X			X		X		X	X			X	X	X	X		X	X			11
Availability		X						X	X	X				X					X	X	7
Believability	X			X		X	X		X	X		X	X	X		X			X	X	12
Consistent Representation	X					X	X	X	X		X	X	X	X	X		X		X		12
Completeness	X		X	X		X	X	X	X			X	X	X	X	X	X	X	X	X	16
Concise Representation	X					X		X	X		X	X		X	X		X				9
Efficiency		X						X		X	X	X	X	X			X			X	9
Navigation					X	X		X					X				X				5
Objectivity	X		X	X	X		X		X	X		X	X	X	X					X	12
Reputation	X			X		X	X		X			X							X		7
Relevancy	X					X		X	X	X	X	X	X	X	X	X				X	12
Reliability		X			X			X		X		X	X						X	X	8
Security	X	X					X	X		X	X	X	X						X	X	10
Timeliness	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		18
Understandability	X	X					X	X		X	X		X						X	X	9
Value added	X					X	X		X	X						X	X	X	X		9
Usability		X					X				X										3
Useful							X												X		2
Interpretability	X						X	X		X	X	X									6
Ease of Operation		X					X			X	X	X				X			X		7
Authority			X	X	X				X	X	X			X		X	X	X			10

Table 2: Tracing IQ Dimensional commonality in existing frameworks.

STEP 2: FINALIZE AND DEFINE DIMENSIONS IN CONTEXT OF WWW

Once the most frequently occurring 23 IQ dimensions had been identified, it was important to look at them in context of the World Wide Web. With this view two new dimensions 'layout' and 'advertisement' were proposed as additions to the list, taking the total number of dimensions to 25.

A focus group of five graduate students was used to understand their perspective of the dimensions. The dimensions were defined in line with definitions understood in available literature and up to four questions were framed around each dimension. Phase one of the experiments required the students to browse a website and complete the survey questionnaire while thinking out loud. Once the individual surveys were completed, the scores for each dimension were analyzed. In phase two the researchers discussed the definitions of the constructs with the focus group and gathered feedback on the relevance of the dimensions in context of the web environment.

Feedback from the focus group suggested that 'useful' and 'usability' should not be included in the list of final dimensions as they were a subgroup of other dimensions like accuracy, value added, completeness and accessibility. Layout was understood as part of navigation and was dropped. Advertising was strongly recommended as one of the dimensions which could change the perception of IQ for the end user. The final list of 22 dimensions- used in subsequent research- with their definitions is shown in Table 3. The definitions are in line with those used in the literature by various researchers. They showcase the meaning of each dimension clearly and mark their scope in measuring the IQ in context of the World Wide Web.

#	IQ Dimension and Definition
1	Accuracy: Extent to which information is correct, reliable and certified free of error
2	Accessibility: Extent to which information is available, or easily and quickly retrievable
3	Advertising: Extent to which extra non-essential information changes perception of information
4	Amount of Data: Extent to which the quantity of volume of available information is appropriate
5	Authority: Extent to which responsibility is taken for information on the website
6	Availability: Extent to which information is physically accessible
7	Believability: Extent to which information is regarded as true and credible
8	Consistent Representation: Extent to which information is presented in the same format and compatible with previous data
9	Completeness: Extent to which information is not missing and is of sufficient breadth and depth for the task at hand
10	Concise Representation: Extent to which information is compactly represented without being overwhelming
11	Ease of Operation: Extent to which info can be manipulated for application to different tasks
12	Efficiency: Extent to which information is quick to meet the information needs for the task at hand
13	Interpretability: Extent to which information carries right symbols units etc
14	Navigation: Extent to which data are easily found and linked to
15	Objectivity: Extent to which information is unbiased, unprejudiced and impartial
16	Reputation: Extent to which information is highly regarded in terms of source or content
17	Relevancy: Extent to which information is applicable and helpful for the task at hand
18	Reliability: Extent to which information is correct and reliable
19	Security: Extent to which access to information is restricted appropriately to maintain its security
20	Timeliness: Extent to which the information is sufficiently up-to-date for the task at hand
21	Understandability: Extent to which information is clear without ambiguity and easily comprehended
22	Value-Added: Extent to which information is beneficial, provides advantages from its use

Table 3: IQ Dimensions used in the research – Definitions

STEP 3: DEVELOPMENT OF THE SURVEY INSTRUMENT

Survey questionnaire was decided as the method of data collection for the research. The initial survey was designed such that each participant was assigned to browse a website and then answer a total of 121 questions regarding the 22 dimensions.

There were up to seven questions per dimension. The questions were scoped strictly around the definitions of the dimension that they were measuring. The research used an iterative cognitive interview process and then a pilot test to arrive at the final survey instrument.

Cognitive Interviews

Cognitive interviews were used to fine tune the questions. Subjects were identified from the appropriate sub-populations for testing the survey questionnaire. The focus group consisted of 2 professors, 2 PhD students and 5 master's students. Subject recruitment was by invitation and factored respondent personalities. All subjects were outspoken and could be critical. Their browsing habits ranged from avid users of the internet to keeping it at arm's length.

The interview covered all aspects of the survey i.e. welcome page, instruction page and the survey questions. Some of the changes made based on the cognitive interviews are listed below:

- Welcome page had a picture in the initial survey instrument, which was removed on the recommendation of the participants as it was not considered salient to the survey.

- Hyperlinks from the email addresses were removed as they were found distracting.

- Welcome page was re-written to make it more appealing for participants to complete the survey.

- Questions were edited/ deleted or reclassified based on respondent feedback.

- Seven point Likert scale was changed to a five point Likert scale.

- Task Scenarios were added to the survey, which included having the participants complete a set of tasks.

Tasks scenarios were added such that users had to complete the task scenarios before proceeding to the questionnaire. It was designed such as to allow the user insight to the IQ dimensions which were being measured in the survey. Another important reason was so that users could have recent experience with web browsing. Without this step respondents were actually completing the survey based entirely from memory which could be a few hours to a few months old.

Once this issue had been addressed it led to another challenge in communication with the respondents. Sample of the initial survey in shown in Table 4 below:

"How important or unimportant is it for you that	Not important at all			Neutral			Very important
	1	2	3	4	5	6	7
information on the website be accurate							
it should be easy to retrieve information from the website							

Table 4: Initial Survey Questionnaire

Once tasks had been introduced for users to complete before the survey, some completed the survey from the perspective of the website they were browsing while others responded based on the importance of IQ dimensions in general. To resolve this confusion the authors clearly asked the users to answer each item (question) at two levels: 'Level of Importance' they associated with an IQ dimension irrespective of the website they were browsing during the survey and their 'Level of Agreement' about the presence of an IQ dimension in the website they were browsing during the survey.

Two crucial words are 'importance' and 'agreement'. In measuring Importance of an IQ dimension the authors aimed to measure the 'value/significance' respondents attach to an IQ dimension in that particular domain. For the Agreement aspect of the question the authors attempted to gauge the actual evaluation of a website being browsed by the users

based on IQ dimensions. Snapshot of the final survey questionnaire is shown in Figure 3 below:

Figure 3: Snapshot of a page from the actual survey hosted on surveymonkey.com

Feedback from the cognitive interviews also steered the authors to reduce the number of questions and the number of questions for each dimension were capped at maximum of four per IQ dimension. This reduced the number of questions to measuring IQ dimensions from 121 to 73. The estimated time for completing the survey reduced from 90 minutes to around 45 minutes. The un-randomized 73 questions are shown in Appendix I besides their respective attributes.

The final survey included 5 demographic questions. 73 questions focused on measuring the 22 IQ dimensions. One open-ended question was added for qualitative research. It was included to get feedback from the participants and seek to identify any IQ dimension(s) which might emerge as a measure of IQ on the web.

Reliability Testing

Cronbach's Alpha was used as an internal consistency technique to assess the homogeneity of the concepts in each category of the proposed research framework. Use of Cronbach's Alpha is fairly standard in most discussions of reliability. In addition, it has been used successfully in other IS instrument development (Moore & Benbasat 1991; Sethi & King 1994; Katerattanakul & Siau 1999).

The accepted level of reliability depends on the purpose of the research project. Davis (1995) suggested that the coefficient of reliability of 0.7 is sufficient for exploratory research. Some suggest that in early stages of research, reliability of 0.5 to 0.6 would be sufficient. The overall value of Cronbach's Alpha value for the instrument in the case of Importance levels is 92.2% while in case of agreement level the overall value is 96.1%.

CONCLUSION

This article has looked at the evolution of IS, its growing complexity and the important role of IQ in defining IS success. 20 Major IQ frameworks and their impact on present state of IQ have been discussed. Authors have developed a framework based on the commonality which exists amongst these frameworks and then refined it using a focus group to arrive a set of 22 IQ dimensions in context of the World Wide Web.

Survey questionnaire was decided as a means of data collection. Iterative cognitive interviews were conducted to fine tune the instrument. The final questionnaire consisted of 3 sections, 5 questions were used to measure the demographic requirements, and 73 questions were used to measures the 22 dimensions being used in this research while the

final question was an open-ended qualitative question seeking suggestions and comments from the participants. Overall Cronbach's alpha value has been reported for both the importance and the agreement aspect of the instrument. The article thus concludes phase-one of the research with the development of a survey instrument to measure IQ on the World Wide Web from a user perspective.

REFERENCES

Aladwani, A. M., & Palvia, P. C. (2002). Developing and validating an instrument for measuring user-perceived web quality. Information & Management, 39(6), 467-476.

Alexander, J. E., & Tate, M. A. (1999). Web wisdom: how to evaluate and create information quality on the web. Mahwah, NJ: Erlbaum.

Ballou, D. P., Wang, R., Pazer, H.L., & Tayi G.K. (1998). Modelling Information Manufacturing Systems to Determine Information Product Quality. Management Science, 44(4).

Beck, S. (1997). Evaluation criteria: the good, the bad & the ugly; or why it's a good idea to evaluate Web sources. http://lib.nmsu.edu/instruction/evalcrit.html

Borovits, I., & Neumann, S. (1979). Computer Systems Performance Evaluation. Lexington, MA: D.C. Health and Company.

Davis, G. B. (1999). A Research Perspective for Information Systems and Example of Emerging Area of Research. Information Systems Frontiers, 1(3), 195-203.

Dedeke, A. (2000). A Conceptual Framework for Developing Quality Measures for Information Systems. Proc. of 5th International Conference on information quality, 126-128.

DeLone, W.H., & McLean, E.R. (1992). Information Systems Success: The Quest for the Dependent Variable. Information Systems Research 3(1), 60-95.

Deming, E.W. 1986. Out of the Crisis.Cambridge, MA: Center for Advanced Engineering Study, MIT.

Diligenti, M., Gori, M., & Maggine, M. (2004). A unified probabilistic framework for

    Web page scoring systems. IEEE Transactions on Knowledge and Data Engineering,

    16(1), 4–16.

Eppler, M., & Muenzenmayer, P. (2002). Measuring information quality in the Web

    Context: A Survey of State-of-the-Art Instruments and an Application Methodology.

    Proc. of 7th International Conference on information quality, 187- 196.

Gendron, M., & D'Onofrio, M. (2002). Formulation of a Decision Support Model Using

    Quality Attributes. Proc. of the 7th International Conference on Information Quality,

    MIT IQ Conference, 2002, 305-316.

Harris, R. (1997). Evaluating Internet Research Sources. Virtual Salt.

    http://www.virtualsalt.com/evalu8it.htm.

Herrera-Viedma, E., Pasi. G., Lopez-Herrera, A., & Porcel, C. (2006). Evaluating the

    Information Quality of Web Sites: A Methodology Based on Fuzzy Computing With

    Words. Journal of the American Society for Information Science and Technology,

    57(4), 538–549.

Ives, B., & Olson, M. (1984). User Involvement and MIS Success: A Review of

    Research. Management Science, 30(5), 586-603.

Juran, J.M. 1989. Jurana on Leader for Quality: An Executive Handbook, New York: The

    Free Press

Juran, J.M., & Gryna, F.M., 1980. Quality Planning and Analysis, 2nd ed. New

    York:McGraw Hill.

Kahn, B. K., Strong, D. M., & Wang, R. Y. (2002). Information quality benchmarks:

    product and service performance. Communications of the ACM, 45(4), 184 - 192.

Katerattanakul, P., & Siau, K. (1999). Measuring information quality of web sites: Development of an Instrument. Proc. of the 20th International Conference on Information Systems (ICIS'99), Charlotte, North Carolina, United States, 279-285

King, W. R., & Rodriguez, J.I. (1978). Evaluating Management Information Systems. MIS Quaterly, 2(3), 43-51.

Klein, B. D. (2001). User perceptions of data quality: Internet and traditional text sources. Journal of Computer Information Systems, 41(4), 9-18.

Klein, B. D. (2002). Internet data quality: Perceptions of graduate and undergraduate business students. Journal of Business and Management, 8(4), 425-432.

Knight, S. (2008). User Perceptions of Information Quality in World Wide Web Information Retrieval Behavior. Dissertation: Edith Cowan University

Leung, H. K. N. (2001). Quality metrics for intranet applications. Information & Management, 38(3), 137-152.

Liu, L., & Chi, L. (2002). Evolutional Data Quality: A Theory-Specific View. Proc. Of the 7th International Conference on Information Quality, (MIT IQ Conference , 292-304.

Liu, Z., & Huang, X. (2005). Evaluating the credibility of scholarly information on the web: A cross cultural study. ScienceDirect, 37(2), 99-106.

Matheus, A. (2004). Web Design Quality Vs. Web Information Quality. Proceedings of the Ninth International Conference on Information Quality (ICIQ-04). November 5-7, 04.

Matlin, G.L. (1977). How to Survive a Management Assessment. MIS Quaterly, 1(1), 11-17.

McLean, E.R. (1973). Assessing returns from the data processing investment, In F.J.

    Gruenberger (ed.), Effective vs. Efficient Computing, pp.12-25, Englewood Cliffs,

    NJ: Prentice Hall.

Moore, G. C., & Benbasat, I. (1991). Development of an Instrument to Measure the

    Perceptions of Adopting an Information Technology Innovation. Information

    Systems Research 2(3), 192-222.

Myers, B.L., Kappelman, L.A., & Prybutok, V. R. (1997). A comprehensive Model for

    Assessing the Quality and Productivity of the Information Systems Function:

    Toward a Theory for Information Systems Assessment. Information Resources

    Management Journal, 10(1), 6-25.

Naumann, F., & Rolker, C. (2000). Assessment Methods for information quality Criteria.

    Proc. of 5th International Conference on information quality, 148-162.

Nunnally, J. C. (1967). Psychometric Theory, New York: McGraw Hill.

Parker, M.B., Harpe, D.L. & Wills, G.B. (2006) An Evaluation of Information Quality

    Frameworks for the World Wide Web, Proceedings of the 8th Annual Conference of

    WWW Applications, Bloemfontein, Free State Province, South Africa.

Price, R. and Shanks, G. (2004) A Semiotic Information Quality Framework, Proc. IFIP

    International Conference on Decision Support Systems (DSS2004): Decision

    Support in an Uncertain and Complex World, Prato, July

Price, R., & Shanks, G. (2005). A semiotic information quality framework: Development

    and comparative analysis. Journal of Information Technology, 20, 88-102.

Scudder, R.A., & Kucic, A.R. (1991). Productivity Measures for Information Systems.

    Information and Management, 20(5), 343-354.

Sethi, V., & King, W. R. (1994). Development of Measures to Assess the Extent to
Which an Information Technology Application Provides Competitive Advantage,"
Management Science 40(12), 1601-1627.

Shanks, G., & Corbitt, B. (1999). Understanding Data Quality: Social and Cultural
Aspects. Proc. 10th Australasian Conference on Information Systems, 785-797.

Shankar, G., & Watts, S. (2003). A Relevant, Believable Approach for Data Quality
Assessment. Proc. of 8th International Conference on information quality, 178-189.

Song, J., & Zahedi, F. M. (2006). Trust in health infomediaries. Decision Support
Systems, 43(2), 390-407

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997a). Data quality in context.
Communications of the ACM, 40(5), p.103 - 110.

Strong,D.M., Lee, Y.W., & Yang, R. (1997b). 10 Potholes in the Road to Information
Quality. Computer, IEEE Computer society Press, 30(8), 38-46.

Sturges, P., & Griffin, A. (2003). The Archaeologist Undeceived: Selecting Quality
Archaeological Information from the Internet. Informing Science Journal, 6, pp.221- 232.

Stvilia, B., Twidale, M. B., Smith, L. C., & Gasser, L. (2005). Assessing Information
Quality of a Community-Based Encyclopedia. 10th International Conference on
Information Quality

Tayi, G. K., & Ballou, D. P. (1998). Examining data quality. Communications of the
ACM, 41(2), 54 - 57.

Tombros, A., Ruthven, I., & Jose, J. M. (2003). Searchers' Criteria for Assessing Web
Pages. Proc. 26th annual International ACM SIGIR Conference on Research and

Development in Information Retrieval, Toronto, Canada, July 28 - Aug 01, 2003, 385-386.

Wang, R.Y., & Strong, D.M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems, 12(4), 5-34.

Watson, R.T. (2007). Information Systems: Creative Commons Attribution 3.0 License, Global Text Project. Zurich, Switzerland: Jacobs Foundation.

Zhu, X., & Gauch, S. (2000). Incorporating quality metrics in centralized/distributed information retrieval on the World Wide Web. Proc. 23rd annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, Athens, Greece, 288-295.

Zmud, R.W. (1979). Individual Differences and MIS Success: A Review of the Empirical Literature. Management Science, 25(10), 966-979.

Appendix I

Kindly indicate

1. Your agreement with the statement made about the specific website (Name of Website)

2. Importance YOU associate to the statement being true for this specific domain of

(Name of Domain)

#	CONSTRUCT	QUESTIONS
1	Accuracy	information on the website is accurate
		information provided on the website is credible for accuracy of content
		information on the website is free of grammatical, spelling and typographical errors
2	Accessibility	irrespective of browser and hardware types, the information on the website is easy to view in different physical settings
		it is easy to obtain needed information from the website
		it is easy to retrieve information from the website
3	Advertising	the website has zero advertising
		no pop ups are used for advertising on the website
		advertising does not conflict with information access and usage
4	Amount of Data	website has enough information to meet your task needs
		website has neither too elaborative nor too specific information
		information on the website contains adequate details
5	Authority	someone takes responsibility for the information provided on the website
		proprietary information establish proper and credible ownership
		the website clearly provides the source of information and contact info
		the website lists recommendation or ratings from outside source
6	Availability	the information provided online on the website is also available by other means
		the website lists alternatives to obtain the same service (information)
		the website provides information with a 'human touch' so you can either speak or meet with a representative
7	Believability	you believe the information on the website before using it to for any purpose
		you trust the information on the website before accepting it for use
		you become convinced about the trustworthiness of information on the website before using it
		you are convinced about information on the website to be credible
8	Consistent Representation	the information on the website has consistent presentation across various webpages and links
		the webpages are compatible with regards to fonts, layouts and presentation etc.
		the information is always presented in the same format on the website
		the information uses consistent language, symbols, units and format across all webpages
9	Completeness	the information on the website is complete

		the website provides full information without directing you to other sources
		the website does not share information in bits and pieces but in its entirety
10	Concise Representation	the information on the website is to the point
		the information on the website is exhaustive and complete yet compact
		the information on the website is not repetitive
11	Ease of Operation	the information on the website (if not restricted or paid) can be downloaded or available for saving
		the website allow you to make changes(add/edit/remove content) on information for which you take responsibility
		the website allow easy steps for accessing and editing/updating your user accounts/ids on the website
		you are able to customize the information and its presentation in your user account on the website
12	Efficiency	the information on the website helps improve your work efficiency
		the information on the website helps in saving time while trying to complete scenarios and other tasks
		the search for recent information on the website appear reverse chronologically (latest to previous)
13	Interpretability	the information on the website uses correct yet identifiable symbols eg USD for US dollars and CAD for Canadian dollars
		the information on the website uses international and local units for easy interpretation e.g. kgs and pounds or liter and ounces
		international protocols are used for information on currency, date, metrics etc. on the website
14	Navigation	the browser title clearly indicates the homepage of a website
		the website provides easy navigation to needed information
		the homepage/main page of the website contains an index or site map for easy navigation to needed information
		information flow and site navigation on the website are clear and not confusing
15	Objectivity	the information on the website is based on facts
		the information on the website is objective
		the website clearly demarcates individual/group opinion and factual information
		the information on the website is impartial
16	Reputation	information on the website be used only on basis of past reputation and recognition
		the website already has a fine reputation before you look at the information being provided on it
		the website is regarded or known for being a credible source of information
17	Relevancy	the website only provides information relevant to the task at hand
		the website only provides related information helpful to solve the task at hand
		the information on the website does not lead to different directions than needed for the task
18	Reliability	information input by user (if allowed) self corrects or triggers exceptions e.g. Client D.O.B of 2/31/2009 should not be allowed
		information on website does not get corrupt over a period of time
		personal and confidential information provided by customer is not sold to

		third party or used for financial gains
		information provided is backed by facts and does not change over time
19	Security	the website has proper safeguards against unauthorized use of available or stored information
		the website highlights credible security measures while handling secure information like credit card info
		the website is a reviewed site and effectively counters viruses, malware and hackers
20	Timeliness	the website clearly mentions when it was last updated
		any time sensitive information on the website clearly mentions date of last update
		the website provides timestamp for all information posted
21	Understandability	the information provided on the website is easily understood
		the information on the website is clear and unambiguous
		the website uses easy to understand language for better comprehension and understanding
22	Value added	the information on the website offers you an advantage of letting to know more than you already do
		the information on the website adds value to your knowledge
		the website provides beneficial information which helps the task at hand and also adds to your pool of knowledge

CHAPTER 3

JOURNAL PAPER 2

Information Quality on the World Wide Web

A Framework for Measurement and Its Validation

Jaikrit Singh Kandari

Erick C. Jones

Fiona F. Nah

Ram R. Bishu

University of Nebraska- Lincoln

Abstract

The scale and reach of published information on the World Wide Web dwarfs the printed paper world. Users are getting information from the web at the click of a button however they must filter sub-standard information before they can use quality content. Researchers have aimed to address this problem by suggesting various information quality frameworks. This article contends that these models though varied in their approach and application, share a greater commonality. It seeks to identify the common attributes that exist across these frameworks. A new framework for the measurement of information quality is developed and twenty two information quality dimensions are identified for measuring information quality in context of the web from a user perspective. An online survey instrument is used for data collection. The research argues that WWW is not a homogeneous entity and should be understood from individual aspects of three independent variables of web domain, type of website and end user nationality to arrive at its conclusion. Results highlight nine IQ dimensions which are important across the whole web environment, while thirteen dimensions have contextual importance and vary across web domain and national culture.

Keywords: Information Quality (IQ), IQ Frameworks, Hofstede's Cultural Dimensions

Information Quality on the World Wide Web

A Framework for Measurement and Its Validation

INTRODUCTION

The World Wide Web is arguably the largest available repository of data with the largest number of visitors searching for information (Herrera-Viedma et al. 2006). The scale and reach of published information on the web dwarfs the printed paper world. In many cases it happens without efficient information quality control (Herrera-Viedma et al. 2006). There are neither rules nor standards governing the type and quality of information that a writer can put on the web (Diligenti, Gori, & Maggine, 2004). One consequence of this oversight presents itself in the form of bad information.

The problem of information quality (IQ) has not escaped researchers' attention. Following general quality literature, Wang and Strong (1996) described information quality (IQ) as data that is 'fit-for use' by data consumers. They propose that assessing information quality (IQ) involves understanding it from the user's point of view. This research adopts their point of view and contends that data cannot be assessed independent of the people who use it.

Kandari et al. (2010) have reviewed twenty major IQ frameworks which have been proposed in literature since 1996. They identified the common dimensions that exist across the existing IQ frameworks in the literature and developed a unified comprehensive framework for the measurement of IQ based on the identified thread of commonality and the intuitive approach. A survey instrument with 22 IQ dimensions was then designed and validated for reliability. This research moves forward from the "Development of a Framework" phase to the measurement, analysis and validation phase

of the research. It reports the ANOVA results of data collection for three independent variables namely, web domain, type of website within a domain and national culture. The implications of the results for information quality in context of the World Wide Web and from a user perspective are discussed.

THEORETICAL BACKGROUND AND RESEARCH OBJECTIVES

INFORMATION QUALITY

Information quality (IQ) is described as data that is 'fit-for-use' (Wang & Strong, 1996). The 'fit-for-use' model is widely adopted in quality literature and emphasizes the importance of taking a consumer's viewpoint of quality because ultimately it is the consumer who will make a judgment about the product's "fitness-for-use" (Deming 1986, Juran 1989, Juran & Gryna 1980). The model has been well received by researchers working in the field of IQ. Wang and Strong (1996) described 'Data quality dimension' as a set of data quality attributes that represent a single aspect or construct of data quality. "Data" usually refers to information at its early stages of processing and "information", is the product at a later stage (Strong et al. 1997). In the context of this article the term "information" refers to both data and information and has been used interchangeably.

INFORMATION QUALITY FRAMEWORKS

Kandari et al. (2010) reviewed twenty major IQ frameworks in IQ literature that have been proposed by researchers in the field of IQ, since its inception in 1996. The authors in agreement with Knight (2008) contend that despite the varied research contexts of IQ frameworks there exists a remarkable commonality amongst the eventual elements identified by various researchers as being important 'dimensions' of IQ. Kandari et al.

(2010) identified twenty three most frequently occurring dimensions in IQ literature. Table 1 provides a summary of the most common dimensions and the frequency with which they have appeared in the identified twenty IQ frameworks.

DIMENSIONS	Wang and Strong	Zeist & Hendricks	Beck	Harris	Alexander &Tate	Katerattanakul & Siau	Shanks & Corbitt	Dedeke	Naumann & Rolker	Zhu & Gauch	Leung	Kahn et al	Liu & Chi	Eppler & Muenzenmayer	Klein	Shankar &Watts	Sturges & Griffin	Tombros et al.	Stvilia et al	Song & Zahedi	TOTAL
Timeliness	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		**18**
Accuracy	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X			X	**17**
Completeness	X		X	X		X	X	X	X			X	X	X	X	X	X	X	X	X	**16**
Accessibility	X	X			X	X	X	X	X	X	X	X		X						X	**12**
Believability	X			X		X	X		X	X		X	X	X		X			X	X	**12**
Consistent Representation	X					X	X	X	X		X	X	X	X	X		X		X		**12**
Objectivity	X		X	X	X		X		X	X		X	X	X	X					X	**12**
Relevancy	X				X		X	X	X	X	X	X	X	X	X	X				X	**12**
Amount of Data	X			X		X		X	X			X	X	X	X		X	X			**11**
Security	X	X						X	X		X	X	X	X					X	X	**10**
Authority			X	X	X					X	X	X			X		X	X	X		**10**
Concise Representation	X					X		X	X		X	X		X	X		X				**9**
Efficiency		X						X		X	X	X	X	X			X			X	**9**
Understandability	X	X						X	X		X	X		X					X	X	**9**
Value added	X							X	X		X	X					X	X	X	X	**9**
Reliability		X			X				X		X		X	X					X	X	**8**
Availability		X						X	X	X				X					X	X	**7**
Reputation	X			X		X	X		X			X							X		**7**
Ease of Operation		X						X			X	X	X				X			X	**7**
Interpretability	X							X	X		X	X	X								**6**
Navigation				X	X		X						X				X				**5**
Usability		X					X				X										**3**
Useful							X													X	**2**

Table 1: Tracing IQ Dimensional commonality in existing frameworks.

DEVELOPMENT OF FRAMEWORK

Twenty three IQ dimensions identified in Table 1 were spread across various IQ perspectives. It was important to look at them in context of the World Wide Web. With this in mind, two new dimensions, 'layout' and 'advertisement' were added to take the IQ list to 25 dimensions. The main reason for adding advertising was that it was strongly felt that inclusion of unsought information could lead to poor perception of information being sought, more so if the advertisement was not relevant to the end user. Each construct (dimension) was defined in line with definitions understood in IQ literature. Kandari et al. (2010) gives complete details on the testing of these 25 dimensions with a focus group of five respondents. The results and feedback from this activity were used to arrive at a 'final set of 22 IQ dimensions' which have been used in subsequent research. The final IQ dimensions with their definitions are shown in Table 2.

#	IQ Dimension and Definition
1	Accuracy: Extent to which information is correct, reliable and certified free of error (Wang & strong 1996)
2	Accessibility: Extent to which information is available, or easily and quickly retrievable (Wang & strong 1996)
3	Advertising: Extent to which extra non-essential information changes perception of information (Kandari et al. 2010)
4	Amount of Data: Extent to which the quantity of volume of available information is appropriate (Wang & strong 1996)
5	Authority: Extent to which responsibility is taken for information on the website
6	Availability: Extent to which information is physically accessible
7	Believability: Extent to which information is regarded as true and credible (Wang & strong 1996)
8	Consistent Representation: Extent to which information is presented in the same format and compatible with previous data (Wang & strong 1996)
9	Completeness: Extent to which information is not missing and is of sufficient breadth and depth for the task at hand (adapted from Wang & strong 1996)
10	Concise Representation: Extent to which information is compactly represented without being overwhelming (Wang & strong 1996)
11	Ease of Operation: Extent to which info can be manipulated for application to different tasks (adapted from Wang & strong 1996)
12	Efficiency: Extent to which information is quick to meet the information needs for the task at hand (Knight 2008)
13	Interpretability: Extent to which information carries right symbols units etc (adapted from Wang & strong 1996)
14	Navigation: Extent to which data are easily found and linked to

15	Objectivity: Extent to which information is unbiased, unprejudiced and impartial (Wang & strong 1996)
16	Reputation: Extent to which information is highly regarded in terms of source or content (Wang & strong 1996)
17	Relevancy: Extent to which information is applicable and helpful for the task at hand (Wang & strong 1996)
18	Reliability: Extent to which information is correct and reliable
19	Security: Extent to which access to information is restricted appropriately to maintain its security (adapted from Wang & strong 1996)
20	Timeliness: Extent to which the information is sufficiently up-to-date for the task at hand (adapted from Wang & strong 1996)
21	Understandability: Extent to which information is clear without ambiguity and easily comprehended (Wang & strong 1996)
22	Value-Added: Extent to which information is beneficial, provides advantages from its use (Wang & strong 1996)

Table 2: IQ Dimensions used in the research with Definitions

RESEARCH OBJECTIVES

WORLD WIDE CONTEXT

Some researchers have looked at IQ in the context of the World Wide Web and even though there is some overlap in the proposed frameworks suggested to measure IQ there is however no consensus over the IQ dimensions which are important for the web. This leads us to main objective of the research:

To identify IQ/DQ dimensions that are relevant in the context of World Wide Web from a user perspective?

WEB DOMAIN AND WEBSITE TYPE CONTEXT

In this light a review of Table 1 shows that some IQ dimensions keep re-enforcing their importance across different IQ contexts unlike others which do not occur as frequently. The authors contend that this is because the web environment is not a homogeneous entity. It can be classified in to different sub-groups and each sub-group should be considered in its individual context. The study looks at three factors to study IQ

in the web context from a user perspective. The three factors are: one, the web domain, two, individual websites within a web domain and three, end-user nationality. Each of this sub-group should be considered in its individual context and thus following hypotheses are proposed:

H1. The significance of individual IQ dimensions varies across different web domains?

H2. The significance of individual IQ dimensions varies across individual website types within individual web domains?

NATIONAL CULTURE PERSPECTIVE

This research argues that any website can potentially be visited by people from many different countries. These users may view and use a website differently depending on their cultural backgrounds (Faiola, 2005). The term "culturability" emphasizes the importance of the relationship between culture and usability in WWW design (Dong &Lee, 2008). A number of cross-cultural web design studies, grounded in Hall (1959, 1976) and Hofstede's (1980, 1980, 1983, 1984, 1985), are available in literature. Their focus however lay in deriving characteristics of webpage design for different cultural contexts (Marcus, 2000 & Yuan et al. 2005). None of the existing frameworks have looked at variations in IQ perception with changes in the national culture. This research believes it is an important area which cannot be overlooked. The following hypothesis is proposed:

H3: The significance of individual IQ dimensions varies across different national cultures in the WWW?

RESEARCH METHODOLOGY

EXPERIMENTAL DESIGN

The research uses a 2^3 or a 2*2*2 complete factorial between-subject research design. The model is shown in Table 3.

Factors	Description	Level 1	Level 2
Domain	Web Domain	News	e-commerce
Type(Domain)	Website within each Domain	High Ranked (HR)	Low Ranked (LR)
Nation	End User Nationality	USA	INDIA

Table 3: The between-subject research model

- *Web Domain*: was tested at two levels for domains: NEWS and e-commerce. The selection was made based on the huge impact they have on the World Wide Web.

- *Nationality*: was tested at two levels, USA and INDIA for national culture.

- *Type of website within a domain*: Two website ranking portals were used to select two websites within each domain. www.compete.com can compare two websites for up to two years based on unique visitors, page views, average stay etc. while www.alexa.com can compare websites for past one year based on traffic rank, reach, page views, time on site, search percentage etc. A comparison snapshot of www.amazon.com vs. www.planetonline.com on the two portals is shown in Figure 1.

Figure 1: Snapshot of www.amazon.com vs. www.planetonline.com

This study contends that within an individual web domain there can be a varied range of individual websites ranging from badly designed, low ranked and poor quality websites to well designed, high ranked and good quality websites. Hence 2 levels for "type of website" were selected for each web domain in the experiment. The websites selection was based on historical data and the cumulative rank on the two ranking portals. By design, one high ranked website (represented with HR) and another low ranked (LR) website were selected to represent two extreme ends of the web domain spectrum. NEWS websites of a foreign country were selected by design to minimize the learning effect and bias of respondents which could arise if NEWS website of a host country was assigned. The websites selected for each domain are shown in Table 4.

DOMAIN	TYPE	
	High Ranked Website (HR)	Low Ranked Website (LR)
NEWS	www.bbc.uk	www.star.co.uk
e-commerce	www.amazon.com	www.planetonline.com

Table 4: Website selection within each Domain

MODEL

ANOVA was used to analyze the data and the following mathematical model was used:

$Y = \mu + \alpha_i + \beta_j + \gamma_{k(i)} + (\alpha\beta)_{ij} + (\beta\gamma)_{jk(i)} + \varepsilon_{ijkl}$

where i, j, k are at two levels and μ is the overall mean of the scores

Main Effect Model Components:

α_j The effect due to i^{th} level of Factor 'Domain'

β_j The effect due to j^{th} level of Factor 'Nationality'

$\gamma_{k(i)}$ The effect due to k^{th} level within i^{th} level Factor 'Type (Domain)'

Two-way Interaction Model Components:

$(\alpha\beta)_{ij}$ The effect of being in level i of Factor 'Domain' and level j of Factor 'Nationality'

$(\beta\gamma)_{jk(i)}$ The effect of being in level j of Factor 'Nationality' and level k within level i of

Factor 'Type(Domain)'

Error Components:

ε_{ijkl} The unexplained part of the score

This leads to the following (generalized) null hypotheses:

1:      H_0: There is no difference in the means of independent factors

H_a: The means are not equal for independent factors

2:      H_0: There is no interaction between independent factors

H_a: There is interaction between independent factors

DATA COLLECTION

INSTRUMENT DEVELOPMENT

Survey questionnaire was used for data collection. Kandari et al. (2010) details the development of the survey instrument. The final questionnaire had a total of 79 questions, 5 measured the demographics of the sample, 73 quantified 22 IQ dimensions while one was an open ended to get a qualitative feedback from the users.

SAMPLE

The study was conducted in an academic setting (undergraduates, graduates, faculty/staff) at two large universities, one in USA and other in INDIA. An e-mail pre-notification invited a random sample of students and non-students. Subjects for the study were also recruited using fliers around the university campus requesting participation.

Sample of Indian participants in United States was screened for permanent residents, citizens or first generation-Indians.

ONLINE SURVEY HOSTING

The final survey was hosted at www.surveymonkey.com. It provided the sophistication needed to host a results database that would first, automatically update and summaries of results when new data was entered into the system, second, generate reports in the desired format and third, provides an easy solution to the security requirements of an SSL connection.

RESULTS

DEMOGRAPHICS

A total of 184 participants responded to the survey questionnaire with 23 subjects in each cell. This is shown in Table 5 below.

	News	News	e-commerce	e-commerce	Total
	www.bbc.uk	www.star.co.uk	www.amazon.com	www.planetonline.com	
USA	23	23	23	23	**92**
INDIA	23	23	23	23	**92**
Total	**46**	**46**	**46**	**46**	**184**

Table 5: Data Distribution -Balanced Cell

Demographic division is shown in Table 6. The number of female respondents were 73 (40%) compared to 111 male participants (60%). Graduate students with the highest percentage of respondents were at 46% while undergraduates and faculty/staff/others had around 26% each. 79% of the respondent population was between 19-30 years of age. Internet usage between 2-5 hours was a day was the most common for 88 (47%) users while 44 (24%) browsed for less than 2 hours daily.

	Female	Male		
Gender	73	111		
	Undergraduate	**Graduate**	**Faculty/Staff/Other**	
Academic Status	48	84	52	
	19-30 yrs.	**31-45 yrs.**	**46-60 yrs.**	**>60 yrs.**
Age	147	32	4	1
	<2 hr./day	**2-5 hr./day**	**5-10 hr./day**	**>10 hr./day**
Usage	44	88	42	10

Table 6: Demographic Data

ANALYSIS

DESCRIPTION OF SURVEY QUESTIONS

The survey questionnaire consisted of 73 questions (items) which measured 22 IQ dimensions (constructs). Snapshot of one of the actual survey pages is show in Figure 2.

Figure 2: Snapshot of a page from the actual survey hosted on surveymonkey.com

The respondents had to reply to each question at two levels: Their 'Level of Agreement' about the presence of an IQ dimension in the website they were browsing during the survey and the 'Level of Importance' they associated with an IQ dimension irrespective of the website they were browsing during the survey.

Two crucial words are 'importance' and 'agreement'. In measuring Importance of an IQ dimension the authors aimed to measure the 'value/significance' respondents attach to an IQ dimension in that particular domain. For the Agreement aspect of the question the authors attempted to gauge the actual evaluation of a website being browsed by the users based on IQ dimensions.

RESULTS FOR IMPORTANCE ASPECT

Table 7 summarizes main effects and interaction effects of three independent variables: 'Domain', 'Nationality' and 'Type of website nested under Domain' for the "Importance Aspect" of the research. Each of the 22 dimensions is a dependent variable.

Summary Results						
Dimension	**Mean**	**Importance Significance**				
		D	**N**	**T(D)**	**D*N**	**N*T(D)**
Accuracy	**4.17**	NS	NS	NS	NS	NS
Advertising	**3.85**	NS	NS	NS	NS	NS
Amount of Data	**3.9**	NS	NS	NS	NS	NS
Believability	**4.13**	NS	NS	NS	NS	NS
Consistent Representation	**3.74**	NS	NS	NS	NS	NS
Completeness	**3.87**	NS	NS	NS	NS	NS
Concise Representation	**3.79**	NS	NS	NS	NS	NS
Navigation	**4.05**	NS	NS	NS	NS	NS
Understandability	**4.19**	NS	NS	NS	NS	NS
Accessibility	**4.14**	0.0057	NS	NS	NS	NS
Reputation	**3.86**	0.0009	NS	NS	NS	NS
Relevancy	**3.7**	<.0001	NS	NS	NS	NS
Authority	**3.74**	NS	0.0002	NS	NS	NS
Timeliness	**3.76**	NS	0.0042	NS	NS	NS
Availability	3.41	0.0145	<.0001	NS	NS	NS
Ease of Operation	3.43	0.002	0.0037	NS	NS	NS
Value Added	**3.97**	0.0083	0.0181	NS	NS	NS
Interpretability	**3.6**	NS	0.0107	0.0022	NS	NS
Security	**4.2**	0.0033	NS	NS	0.028	NS
Efficiency	**3.59**	NS	NS	NS	0.0106	NS
Reliability	**3.89**	0.0474	NS	NS	NS	0.02
Objectivity	**3.91**	NS	NS	NS	NS	0.0417

Table 7 ANOVA Summary for Importance Levels

Notations:
D: Domain; N: Nationality; T (D): Type of website nested within Domain; NS: Not-Significant

Summary results in Table 7 show clearly that nine out of the twenty two dimensions are not significantly impacted by any of the independent factors. These dimensions include *Accuracy*, *Advertising*, *Amount of Data*, *Believability*, *Consistent Representation*, *Completeness*, *Concise Representation*, *Navigation* and *Understandability*.

Domain has a significant effect on dependent measures *Accessibility, Reputation and Relevancy*. The mean values of these dimensions across two levels of web domains e-commerce and NEWS are plotted in Figure 3, 4 and 5 respectively.
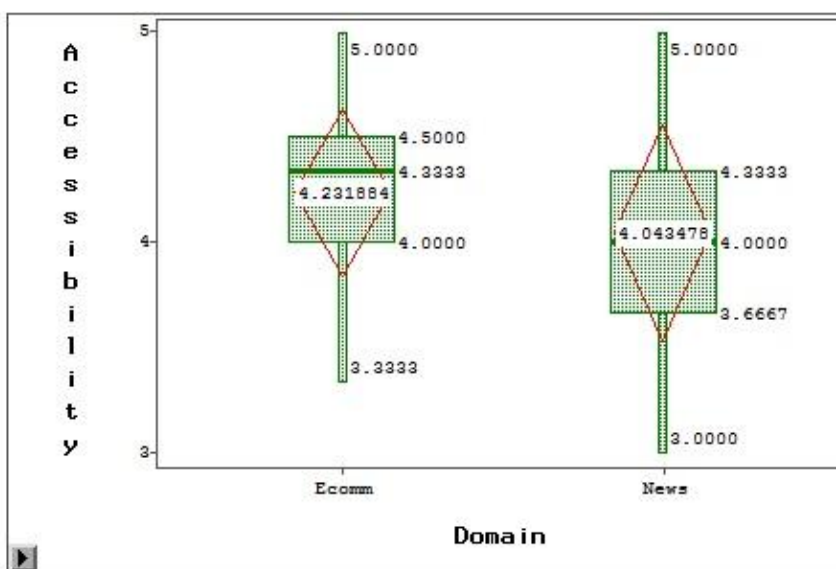
Figure 3: Accessibility mean vs. Web Domain
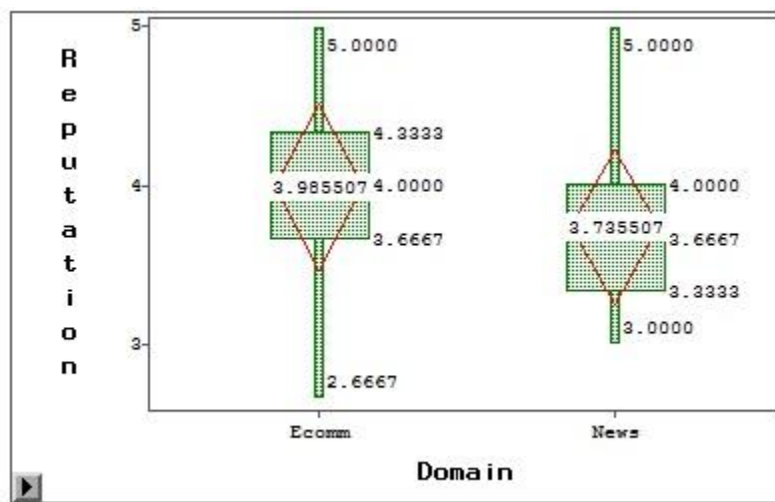

Figure 4: Reputation mean vs. Web Domain
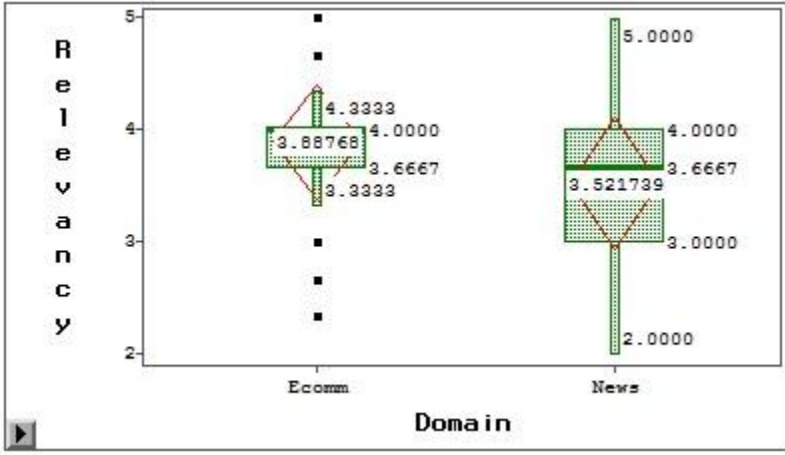
Figure 5: Relevancy mean vs. Web Domain

Main effect Nation has significant impact on two dependent IQ dimensions of *Authority* and *Timeliness*. The mean values of these dimensions are plotted against two levels of nationality, namely, INDIA and USA. The graphs are shown in Figure 6 and 7 respectively.
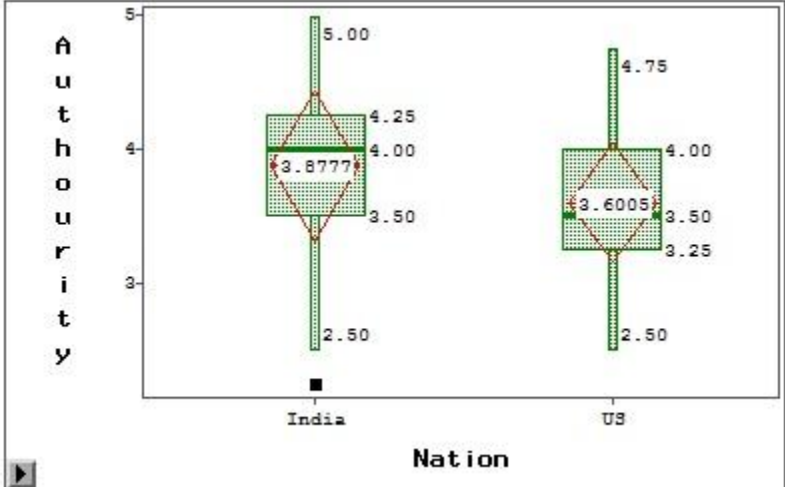
Figure 6: Authority mean vs. Nationality
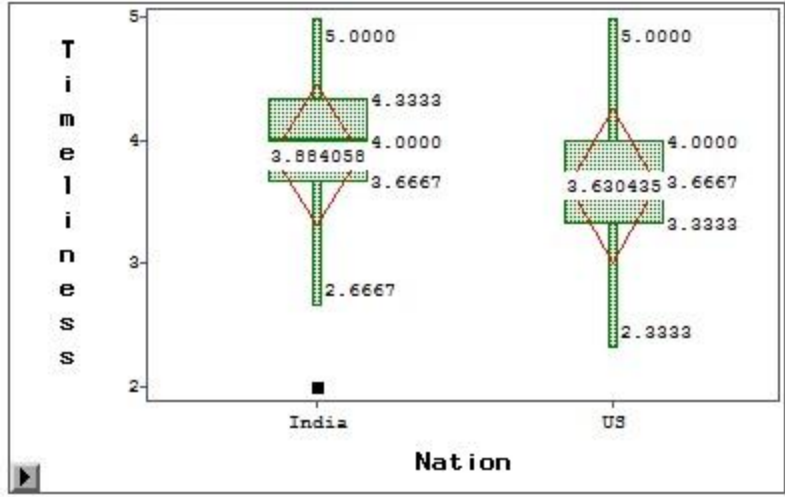
Figure 7: Timeliness mean vs. Nationality

Summary Table 7 shows that both factors domain and nation show significant main effects for IQ dimensions *Value Added*, *Availability* and *Ease of Operation*. Graphs for the variations in IQ dimensions versus 2 levels of domain (e-commerce and NEWS) and 2 levels of Nationality (INDIA and USA) are plotted in Figures 8(a): (b), 9(a): (b) and 10(a): (b) respectively.
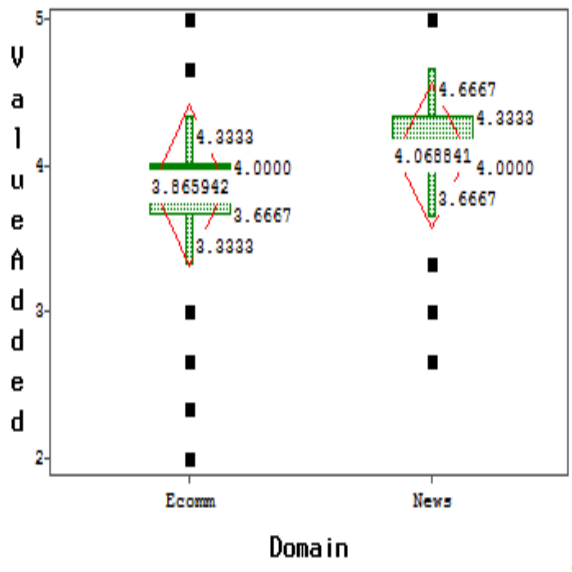
Figure 8a

Value added mean vs. Web Domain
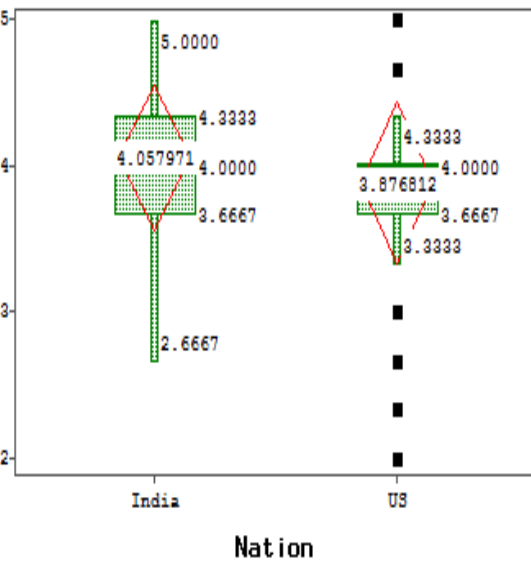
Figure 8b

Value added mean vs. Nationality

Figure 9a

Availability mean vs. Web Domain
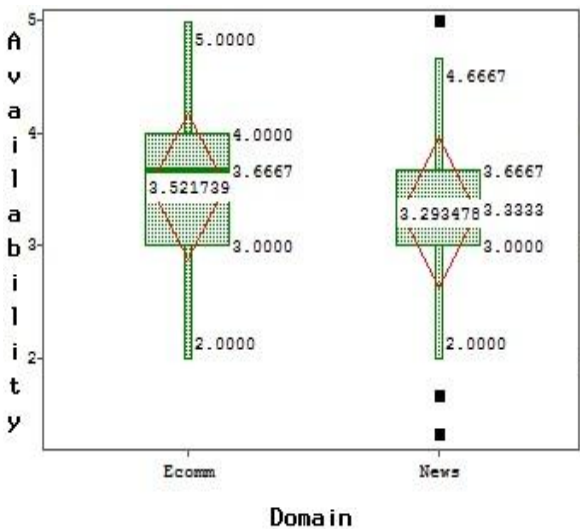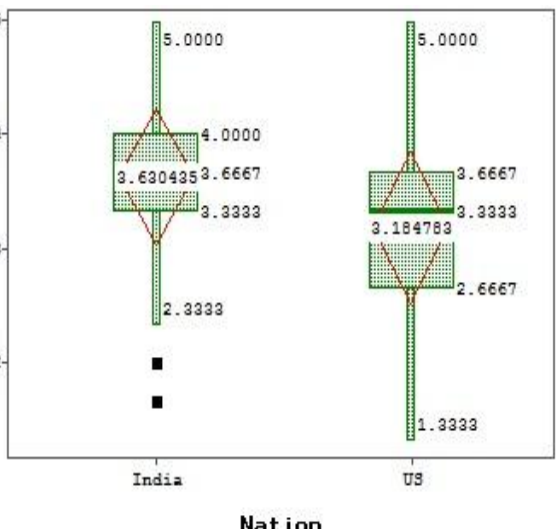
Figure 9b

Availability mean vs. Nationality

Figure 10a

Ease of Operation mean vs. Web Domain
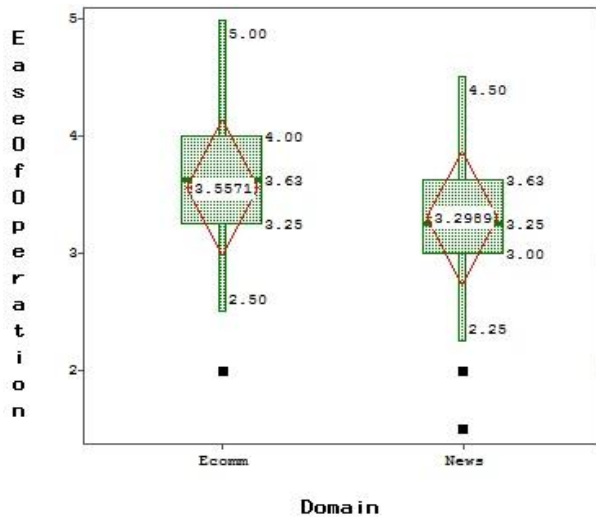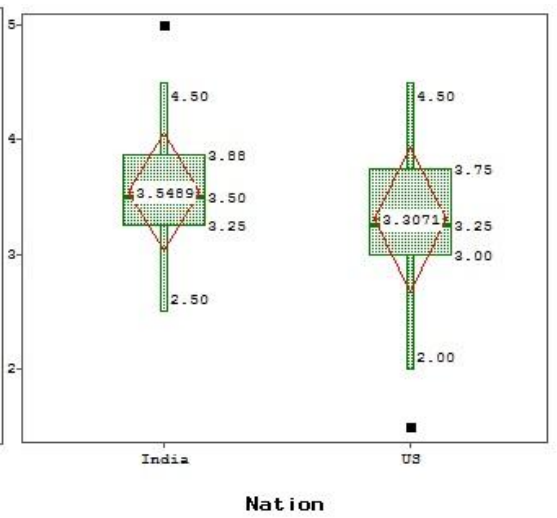
Figure 10b

Ease of Operation mean vs. Nationality

Security shows significant variation with domain and domain*nation interaction. Domain*nation interaction also significantly effects *Efficiency*. These interactions for Security vs. Domain * Nation and Efficiency vs. Domain*Nation are plotted in Figure 11 and 12 respectively.
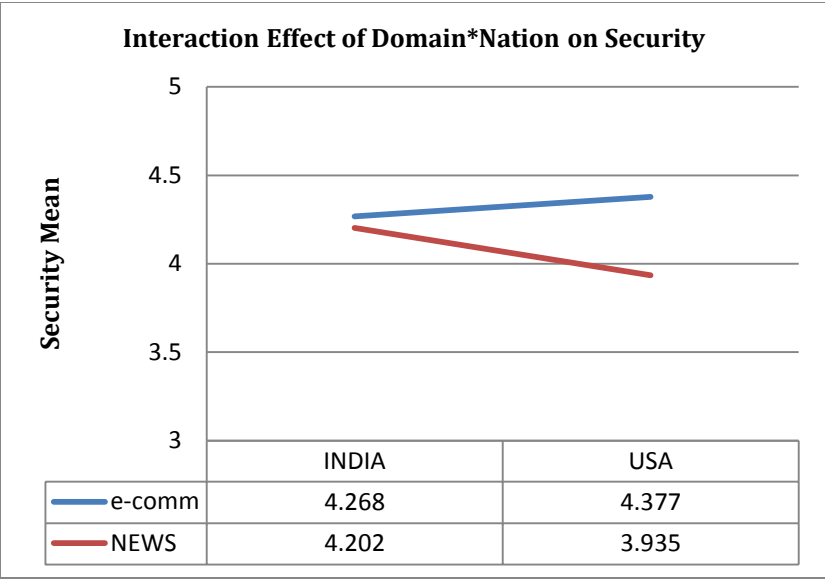
Figure 11: Interaction Effect of Domain*Nation vs. Security
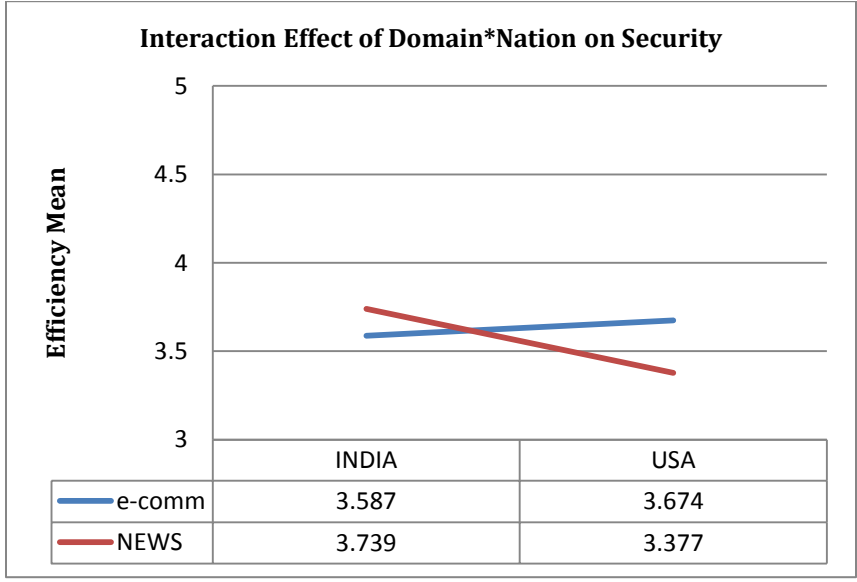
Figure 12: Interaction Effect of Domain*Nation vs. Efficiency

Factors Nation and Type (Domain) have significant effect on *Interpretability*. Graph of Interpretability mean values against two levels of nationality is shown in Figure 13 while type nested under domain is not plotted as it is not of interest to the authors.
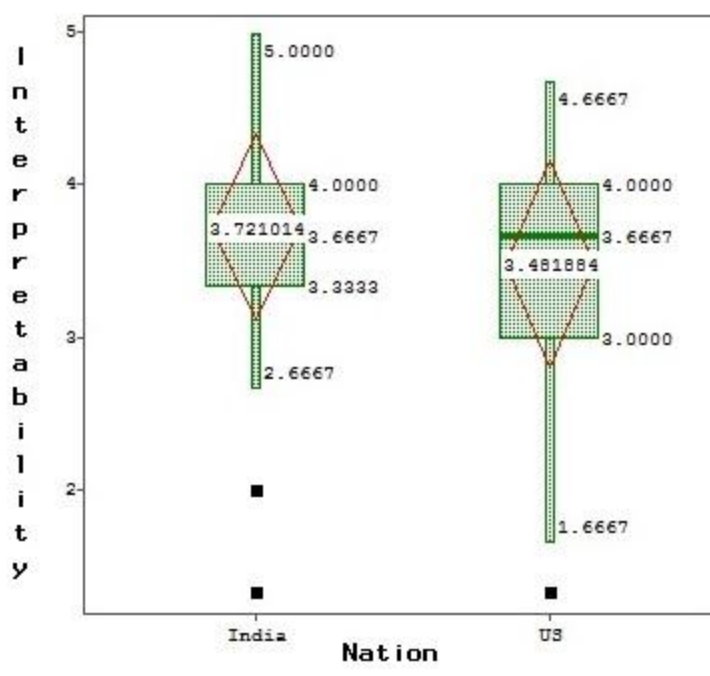
Figure 13: Interpretability mean vs. Nationality

Remaining dimensions of *Reliability* and *Objectivity* are significantly affected by two-way interactions which involve Type (Domain). This is not of interest to the authors and has not been pursued.

INTERPRETATION OF IMPORTANCE RESULTS:

The results in summary table 7 can be classified into four categories of a 2*2 results quadrant, as shown in Figure 14.
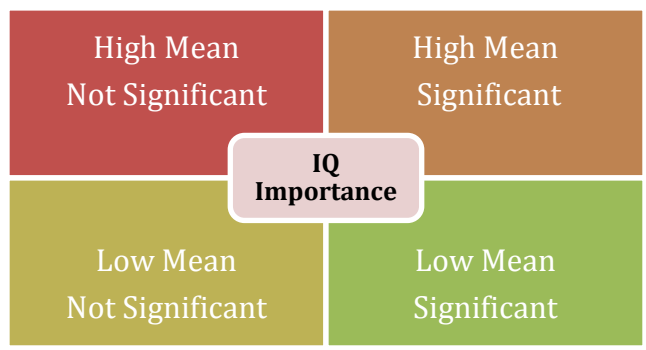
Figure 14: 2X2 results quadrant

These quadrants can be interpreted as discussed below:

- High value of mean and no significant independent factors will suggest that the IQ dimension is important across all nationalities, domain and web site types.

- High value of importance mean and significant main effect and/or interaction effect would mean that IQ dimension though important depends on one or more of the independent factors.

- Low value of mean and none of the independent factors as significant, it can be reasonably deduced that the dimensions are not important to IQ on the World Wide Web from a user perspective.

- Low mean and significant independent factors would mean that even though the dimension is not important from a user perspective in information quality it is still significantly impacted by the independent factors considered in the study.

Mean values of the IQ dimensions in Table 7 indicate that 20 out of 22 IQ dimensions have mean values above 3.5. Six of these twenty values are above the 4.0 mark. The values are plotted in Figure 15. These values have been assigned to the results quadrant.
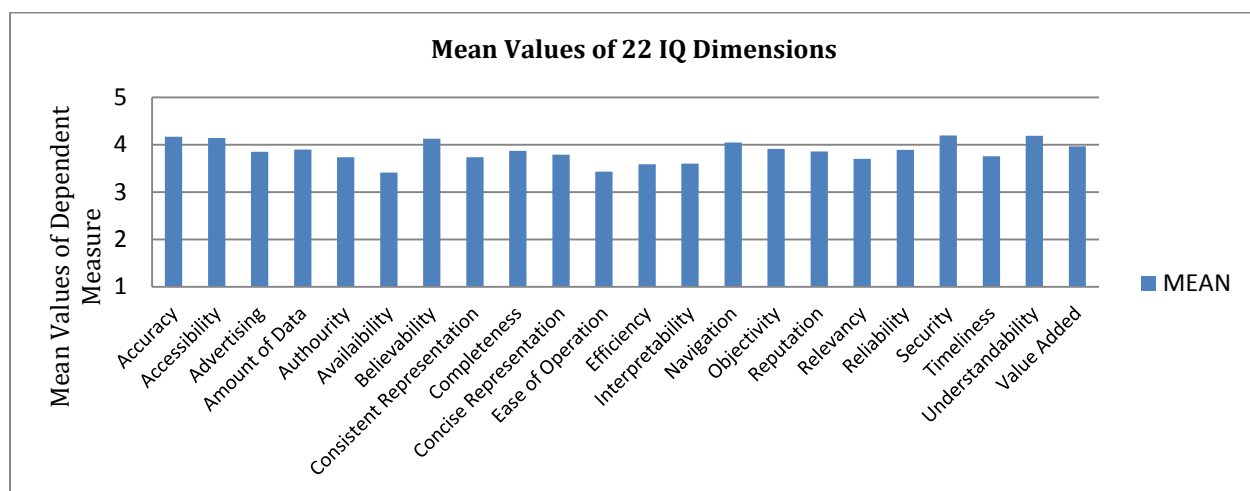
Figure 15: Mean values of twenty two IQ dimensions

Means above 3.5 are considered higher and below 3.0 are considered low. Eleven dimensions fall in Ist quadrant, nine in the IInd quadrant, while none fall in quadrant IIIrd and IVth. This division is shown in Figure 16. Two dimensions, *Availability* and *Ease of Operation* with mean values between 3.0 and 3.5 have not been assigned to any quadrant.

High Mean Not Significant		High Mean Significant		
Understandability	4.19	Accessibility	4.14	Domain
Accuracy	4.17	Reputation	3.86	Domain
Believability	4.13	Relevance	3.7	Domain
Navigation	4.05	Authority	3.79	Nation
Amount of Data	3.90	Timeliness	3.76	Nation
Completeness	3.87	Value Added	3.97	Domain, Nation
Advertising	3.85	Interpretability	3.6	Nation, Type (Domain)
Concise Representation	3.79	Objectivity	3.91	Nation*Type (Domain)
Consistent Representation	3.74	Efficiency	3.59	Domain* Nation
		Security	4.2	Domain, Domain* nation
		Reliability	3.89	Domain, Nation* Type (Domain)
NA		NA		
Low Mean Not Significant		Low Mean Significant		

Figure 16: Summary Results Divided into Results Quadrant

Accessibility, Reputation and *Relevance* are 'Domain' dependent which means that though they are considered important measures of IQ, their importance will vary by domain. In Figure 3 mean scores of Importance levels for IQ dimension 'Accessibility' were plotted for NEWS and e-commerce domains. The box plot shows that maximum value for both domains is 5 which suggest that respondents consider the IQ dimensions

"very important" on the Likert scale. However for NEWS the minimum value is 3.0 suggesting that at least one respondent was "Neutral" and did not give any importance rating to 'Accessibility'. The domain e-commerce had a minimum score of 3.33 suggesting that 'Accessibility' was regarded favorably as an important IQ dimension. Also e-commerce had a higher mean at 4.23 compared to NEWS's 4.04. 50% respondents scored importance of accessibility in e-commerce between 4.0 and 4.5 while the same percentage scored accessibility between 3.667 and 4.33 for the NEWS domain. Thus we see that though 'Accessibility' is an important IQ dimension it will rank higher in the e-commerce domain than the NEWS domain.

By the same interpretation of the box plot it can be argued that, *Reputation* and *Relevancy* are relatively more important in the e-commerce domain compared to the NEWS domain. *Authority* and *Timeliness* which show significant effect of main factor Nationality are higher ranked IQ dimensions from the perspective of an Indian user than an American respondent. Relative rankings of the IQ dimensions are shown in Table 8.

Comparision of Relative Significance of a Dimensions Between 2 Factor Levels				
Dimensions	**Relative Importance**	**Factor Level 1**	**Factor Level 2**	**Relative Importance**
Accessibility	⬆	e-commerce	NEWS	⬇
Reputation	⬆	e-commerce	NEWS	⬇
Relevancy	⬆	e-commerce	NEWS	⬇
Authority	⬆	INDIA	USA	⬇
Timeliness	⬆	INDIA	USA	⬇

Table 8: Relative rankings of IQ dimensions within 2 levels of same factor

Value Added is impacted by both the Domain and Nation. Using box plot in Figure 8(a) and 8(b) it can be interpreted that *Value Added* has more relevance to NEWS in the web domain and INDIA in factor nationality. *Security* is significantly affected by both the factors and their main effects are more pronounced than the interaction while for *Efficiency*, the interaction effects of Domain and Nation mask the main effects and are more pronounced.

Accuracy, *Advertising*, *Amount of Data*, *Believability*, *Consistent Representation*, *Completeness*, *Concise Representation*, *Navigation* and *Understandability* fall in quadrant II and not affected by any factors. These IQ dimensions are thus important for all the web site types for both web domains and across nationalities.

ONE-WAY ANOVA

The research efforts have so far concentrated on understanding the impact of Nationality, Domain and their interaction on the "Importance Aspect" of the IQ dimensions. The role of the nested factor "Type of Website within a domain" is unclear.A one-way ANOVA was performed on the data that was collected to better understand how the "Type of website" played a part in IQ measurement from the user's perspective. Respondents rated the high ranked (HR) or low ranked (LR) website- they had been assigned (by design)- on the 22 IQ dimensions. The rating scale ranged from Strongly Disagree to Strongly Agree. The data for "Type of website" was divided based on the two domain levels of e-commerce and NEWS. ONE -Way ANOVA on "Type" was used for analysis. Summary results for e-commerce and NEWS are shown in Table 9 and Table 10 respectively.

Summary Results for e-commerce					
Agreement Aspect				**Importance Aspect**	
Dimension	Pr>F	MEAN (HR)	MEAN (LR)	Dimension	Pr>F
Accuracy	0.0002	3.91	3.44	Accuracy	NS
Accessibility	<.0001	3.97	3.22	Accessibility	NS
Advertising	0.0006	3.34	2.72	Advertising	NS
Amount of Data	<.0001	3.72	3.20	Amount of Data	NS
Authority	0.0036	3.63	3.21	Authority	NS
Availability	NS	NS	NS	Availability	NS
Believability	<.0001	3.88	2.88	Believability	NS
Consistent Representation	0.0022	3.91	3.54	Consistent Representation	NS
Completeness	<.0001	3.68	2.93	Completeness	NS
Concise Representation	0.0266	3.42	3.10	Concise Representation	NS
Ease of Operation	0.0226	3.41	3.14	Ease of Operation	NS
Efficiency	<.0001	3.49	2.94	Efficiency	NS
Interpretability	NS	NS	NS	Interpretability	NS
Navigation	<.0001	3.93	3.48	Navigation	NS
Objectivity	0.0296	3.63	3.36	Objectivity	NS
Reputation	<.0001	3.91	2.90	Reputation	NS
Relevancy	NS	NS	NS	Relevancy	NS
Reliability	0.001	3.57	3.15	Reliability	NS
Security	0.0001	3.74	3.09	Security	NS
Timeliness	NS	NS	NS	Timeliness	NS
Understandability	0.0005	3.95	3.48	Understandability	NS
Value Added	0.0002	3.70	3.11	Value Added	NS

Table 9: One –Way ANOVA Summary for e-commerce

Notations: NS: Not Significant

Table 9 shows that for the importance aspect "Type of website" did not impact any of the IQ dimensions. The participant could be browsing a low ranked website or a high ranked website as assigned in the survey and yet it had no bearing on their ratings for the "Importance of the IQ dimension". However in an extreme turnaround for agreement

aspect, 18 out of 22 dimensions were significantly impacted by the type of website the user was browsing. User rating for the mean values of these dimensions clearly indicates that the higher ranked websites had done a better job in addressing the IQ dimension than their lower ranked counterparts.

Summary Results for NEWS					
Agreement Aspect				Importance Aspect	
Dimension	Pr>F	MEAN (HR)	MEAN (LR)	Dimension	Pr>F
Accuracy	0.0002	4.04	3.54	AccuTracy	NS
Accessibility	<.0001	4.08	3.28	Accessibility	NS
Advertising	<.0001	3.36	2.79	Advertising	NS
Amount of Data	<.0001	3.86	3.31	Amount of Data	NS
Authority	NS	3.45	3.36	Authority	NS
Availability	NS	3.07	3.00	Availability	NS
Believability	<.0001	3.94	3.15	Believability	0.0449
Consistent Representation	NS	3.73	3.47	Consistent Representation	NS
Completeness	0.0172	3.64	3.25	Completeness	NS
Concise Representation	<.0001	3.73	3.15	Concise Representation	NS
Ease of Operation	0.0202	3.39	3.11	Ease of Operation	NS
Efficiency	<.0001	3.57	2.96	Efficiency	NS
Interpretability	<.0001	3.91	3.35	Interpretability	0.0056
Navigation	0.0006	4.12	3.64	Navigation	NS
Objectivity	NS	3.61	3.38	Objectivity	NS
Reputation	<.0001	3.76	3.15	Reputation	NS
Relevancy	0.0487	3.30	3.02	Relevancy	NS
Reliability	0.0349	3.47	3.26	Reliability	NS
Security	0.0053	3.54	3.14	Security	NS
Timeliness	NS	3.93	3.67	Timeliness	NS
Understandability	<.0001	4.09	3.59	Understandability	NS
Value Added	<.0001	4.15	3.50	Value Added	NS

Table 10: One –Way ANOVA Summary for e-commerce

Notations: NS: Not Significant

Table 10 shows that for the importance aspect "Type of website" did not impact 20 out the 22 IQ dimensions. For agreement aspect, 18 out of 22 dimensions were

significantly impacted by the type of website the user was browsing. Mean values of IQ dimensions were greater for higher ranked websites than lower ranked websites.

In summary analysis of One- Way ANOVA shows that the participant rating for the importance aspect were purely based on the "value" they attached to an IQ dimension within a domain, while for the agreement aspect the website was rated poorly if the IQ dimension was missing and vice versa.

DISCUSSION

This research identified 20 major IQ frameworks which exist in IQ literature. Up on thorough review it acknowledges that there is an inherent commonality amongst different models. A framework was developed based on the idea of commonality and then refined in context of the World Wide Web.

Data collected was analyzed using ANOVA model in SAS. In context of this research the authors suggest a ranking order for IQ dimension as shown in Table 11. For practitioners wanting to develop websites with high IQ, the first nine dimensions in relative rankings are *Understandability, Accuracy, Believability, Navigation, Amount of Data, Completeness, Advertising, Concise Representation* and *Consistent Representation.* Results show that these dimensions are not significantly impacted by any of the factors and more importantly these dimensions maintain high mean values for their importance ratings across all factor levels. The authors contend that the nine dimensions form the core group which cannot be neglected while developing a high IQ website. In other words it can be interpreted to mean that these dimensions will hold their relative

importance across the World Wide Web. *Generalizing these results means that the 'set of nine' is the core set valid for any website, for any domain and for all nationalities in the World Wide Web.*

			Summary Results				
RANK	Dimension	Mean	Importance Significance				
			D	N	T(D)	D*N	N*T(D)
1	Understandability	**4.19**	NS	NS	NS	NS	NS
2	Accuracy	**4.17**	NS	NS	NS	NS	NS
3	Believability	**4.13**	NS	NS	NS	NS	NS
4	Navigation	**4.05**	NS	NS	NS	NS	NS
5	Amount of Data	**3.9**	NS	NS	NS	NS	NS
6	Completeness	**3.87**	NS	NS	NS	NS	NS
7	Advertising	**3.85**	NS	NS	NS	NS	NS
8	Concise Representation	**3.79**	NS	NS	NS	NS	NS
9	Consistent Representation	**3.74**	NS	NS	NS	NS	NS
10	Security	**4.2**	**0.0033**	NS	NS	**0.028**	NS
11	Value Added	**3.97**	**0.0083**	**0.0181**	NS	NS	NS
12	Objectivity	**3.91**	NS	NS	NS	NS	**0.0417**
13	Reliability	**3.89**	**0.0474**	NS	NS	NS	**0.02**
14	Accessibility	**4.14**	**0.0057**	NS	NS	NS	NS
15	Reputation	**3.86**	**0.0009**	NS	NS	NS	NS
16	Relevancy	**3.7**	**<.0001**	NS	NS	NS	NS
17	Timeliness	**3.76**	NS	**0.0042**	NS	NS	NS
18	Authority	**3.74**	NS	**0.0002**	NS	NS	NS
19	Interpretability	**3.6**	NS	**0.0107**	**0.0022**	NS	NS
20	Efficiency	**3.59**	NS	NS	NS	**0.0106**	NS
21	Availability	3.41	**0.0145**	**<.0001**	NS	NS	NS
22	Ease of Operation	3.43	**0.002**	**0.0037**	NS	NS	NS

Table 11: Relative Rankings of IQ Dimensions

Dimensions 10 through 20, showed high mean values. *Security, Value Added, Objectivity and Reliability* have high mean scores from 4.2 to 3.89. They were significantly impacted by the main effect or interaction effects. In essence it suggests that though their relative rankings will vary across factors of Domain, Nation and Type (Domain), these dimensions should be given enough attention in all spheres since any one

factor or their interaction could play a role in altering user perception of IQ on the website.

Accessibility, Reputation and Relevancy are ranked in context of this research in Table 11. Domain showed a main effect. Up on further analysis it was seen that all the three dimensions had higher importance relevance for e-commerce than compared to NEWS. Hence web developers should take notice of the web domain before ensuring the presence of these IQ dimensions.

Timeliness followed by *Authority, Interpretability* and *Efficiency* were ranked from 17 to 20. All the four IQ dimensions showed Nation as a contributing factor in their ratings. Authors argue that web developers only make an educated guess about the nationality of the user who could potentially browse the website. Hence even though *Timeliness* and *Authority* have higher mean scores than *Relevancy* they are lower in the rankings.

Means values of 3.41 and 3.4 for *Availability* and *Ease of Operation* are less than 3.5 which are needed to qualify in the high mean quadrant. Neither do the numbers qualify for the low mean quadrant (3.0 and below). Both IQ dimensions are impacted significantly by Domain and Nation. Thus while in this research they lie in the zone of indecision, it is possible that for other domains they might qualify for quadrant I or slip down to quadrant IV. Both the dimensions have been ranked the lowest in the IQ dimensions.

One Way ANOVA was used to understand the effect of type of website on the IQ dimensions. It was also used as a method to validate the types of websites which were selected within each domain. The ANOVA result showed that for e-commerce none of

the mean values with importance ratings for the IQ dimensions were different across the two levels of the domain. Hence it was validated that type of website did not play a role in user attaching "value" to a particular IQ dimension. While in agreement aspect, mean values for 20 out of 22 dimensions were significantly higher for high ranked website (www.amazon.com) when compared to those of lower ranked website (www.planetonline.com). The same pattern was seen for the NEWS domain. Thus it can be argued that any website which has a better presence of IQ dimensions will surely have better standards of Information Quality.

This study has brought out some interesting findings. The authors' argument that World Wide Web is not a homogeneous entity but a sum of parts has not been rejected by the results of this study. As a part of future research if the set of nine dimensions is incorporated as a part of the search engine algorithm then hopefully the search results for a query will direct the user to better websites with higher information quality.

One of the limitations of the study was domain has not been used in its traditionally understood meaning. The meaning as used in this study is actually a big section within a .com domain and the use of actual domains was beyond the scope of this work. It would make an interesting future research to study the IQ dimensions in the traditional definition of web domains i.e. a '.edu' vs. a '.com' vs. a '.gov' vs. a '.org' domain.

REFERENCES

Alexander, J. E., & Tate, M. A. (1999). Web wisdom: how to evaluate and create information quality on the web. Mahwah, NJ: Erlbaum.

Beck, S. (1997). Evaluation criteria: the good, the bad & the ugly; or why it's a good idea to evaluate Web sources. http://lib.nmsu.edu/instruction/evalcrit.html

Dedeke, A. (2000). A Conceptual Framework for Developing Quality Measures for Information Systems. Proc. of 5th International Conference on information quality, 126-128.

Deming, E.W. 1986. Out of the Crisis.Cambridge, MA: Center for Advanced Engineering Study, MIT.

Diligenti, M., Gori, M., & Maggine, M. (2004). A unified probabilistic framework for Web page scoring systems. IEEE Transactions on Knowledge and Data Engineering, 16(1), 4–16.

Eppler, M., & Muenzenmayer, P. (2002). Measuring information quality in the Web Context: A Survey of State-of-the-Art Instruments and an Application Methodology. Proc. of 7th International Conference on information quality, 187- 196.

Faiola, A. (2005). Cross-culture cognition and online information design: Identifying cognitive styles among web designers of diverse national origin. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.

Hall, E. T. (1959). The silent language. New York: Doubleday.

Hall, E. T. (1976). Beyond culture. New York: Anchor Press.

Harris, R. (1997). Evaluating Internet Research Sources. Virtual Salt http://www.virtualsalt.com/evalu8it.htm

Herrera-Viedma, E., Pasi. G., Lopez-Herrera, A., & Porcel, C. (2006). Evaluating the

Information Quality of Web Sites: A Methodology Based on Fuzzy Computing With

Words. Journal of the American Society for Information Science and Technology,

57(4), 538–549.

Hofstede, G. (1980). Culture's Consequences: International Differences in Work-Related

Values. Beverly Hills, CA: Sage Publications.

Hofstede, G. (1983). The cultural relativity of organizational practices and theories.

Journal of International Business Studies, 76-88.

Hofstede, G. (1984). Cultural dimensions in management and planning. Asia Pacific

Journal of Management, 1(2), 81-99.

Hofstede, G. (1985). The interaction between national and organizational value system.

Journal of Management Studies, 22(4), 347-357.

Juran, J.M. 1989. Jurana on Leader for Quality: An Executive Handbook, New York: The

Free Press

Juran, J.M., & Gryna, F.M., 1980. Quality Planning and Analysis, 2nd ed. New

York:McGraw Hill.

Kahn, B. K., Strong, D. M., & Wang, R. Y. (2002). Information quality benchmarks:

product and service performance. Communications of the ACM, 45(4), 184 - 192.

Kandari, J., Jones, E.C., Nah, F., & Bishu, R.(2010). Information Quality on the World

Wide Web: A User Perspective. UNL: MS Thesis

Katerattanakul, P., & Siau, K. (1999). Measuring information quality of web sites:

Development of an Instrument. Proc. of the 20th International Conference on

Information Systems (ICIS'99), Charlotte, North Carolina, United States, 279-285.

Klein, B. D. (2002). Internet data quality: Perceptions of graduate and undergraduate business students. Journal of Business and Management, 8(4).

Knight, S. (2008). User Perceptions of Information Quality in World Wide Web Information Retrieval Behavior. Dissertation: Edith Cowan University

Leung, H. K. N. (2001). Quality metrics for intranet applications. Information & Management, 38(3), 137-152.

Liu, L., & Chi, L. (2002). Evolutional Data Quality: A Theory-Specific View. Proc. Of the 7th International Conference on Information Quality, (MIT IQ Conference , 292-304.

Marcus, A., & Gould, E. W. (2000). Crosscurrents, cultural dimensions and global Web user-interface design. Interactions, 7(4), 32-46.

Naumann, F., & Rolker, C. (2000). Assessment Methods for information quality Criteria. Proc. of 5th International Conference on information quality, 148-162.

Shanks, G., & Corbitt, B. (1999). Understanding Data Quality: Social and Cultural Aspects. Proc. 10th Australasian Conference on Information Systems, 785-797.

Shankar, G., & Watts, S. (2003). A Relevant, Believable Approach for Data Quality Assessment. Proc. of 8th International Conference on information quality, 178-189.

Song, J., & Zahedi, F. M. (2006). Trust in health infomediaries. Decision Support Systems, 43(2), 390-407

Strong,D.M., Lee, Y.W., & Yang, R. (1997). 10 Potholes in the Road to Information Quality. Computer, IEEE Computer society Press, 30(8), 38-46.

Sturges, P., & Griffin, A. (2003). The Archaeologist Undeceived: Selecting Quality

Archaeological Information from the Internet. Informing Science Journal, 6, pp.221-232.

Stvilia, B., Twidale, M. B., Smith, L. C., & Gasser, L. (2005). Assessing Information Quality of a Community-Based Encyclopedia. 10th International Conference on Information Quality

Tombros, A., Ruthven, I., & Jose, J. M. (2004). Searchers' Criteria for Assessing Web Pages. Proc. 26th annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Toronto, Canada, July 28 - Aug 01, 2003, 385-386.

Wang, R.Y., & Strong, D.M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems, 12(4), 5-34.

Yuan, X., Liu, H., Xu, S., & Wang, Y. (2005). The impact of different cultures on e-business Web design–Comparison research of Chinese and Americans. In *Proceedings of the 11th International Conference on Human-Computer Interaction* [CD-ROM]. Las Vegas, NV: Mira Digital Publishing.

Zeist, R. H. J., & Hendriks, P. R. H. (1996). Specifying software quality with the extended ISO model. Software Quality Management, 5(4), 273-284.

Zhu, X., & Gauch, S. (2000). Incorporating quality metrics in centralized/distributed information retrieval on the World Wide Web. Proc. 23rd annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, Athens, Greece, 288-295.

CHAPTER 4

DISCUSSION

The study is based on the argument that there is a dimensional commonality in existing IQ frameworks. It also contends that the World Wide Web is not a homogenous entity and is a sum of its many individual entities and thereby each should be considered in their own context. The research has been able to show that at least 20 of the 22 dimensions are important IQ dimensions in the World Wide Web. Two IQ dimensions which lie in the zone of uncertainty for the result quadrant are domain dependent and it will be interesting to note how they behave in other domain types.

DEVELOPMENT OF FRAMEWORK

Based on the thread of commonality amongst IQ frameworks between 1996-2006, twenty three most frequently occurring IQ dimensions were identified. These were refined for the overall context of the World Wide Web by using a focus group. One new IQ dimension- Advertising was added while two dimensions were dropped from the list resulting in the final 22 dimensions. It is interesting to note that Advertising ranks at number 7 in the overall ranking of those 22 dimensions based on the ANOVA analysis. It forms the core group of nine IQ dimensions which have been identified as the most important across all web domains, nationality and web site types and thus by extension across the World Wide Web.

THE RESULTS QUADRANT

The authors felt it was best to divide the results from the ANOVA analysis in a results quadrant. The expectation was that all the results will fit into the four quadrants and will thereby give a clear picture about the importance of the IQ dimensions in context of the World Wide Web. The finding of the research was surprising since 20 of the 22 IQ dimensions fit into quadrant I and II, while none qualified to be in the the IIIrd and IVth quadrant.

ANOVA model was analyzed in SAS. *Understandability, Accuracy, Believability, Navigation, Amount of Data, Completeness, Advertising, Concise Representation* and *Consistent Representation* are not significantly impacted by any of the factors and more importantly these dimensions maintain high mean values for their importance ratings across all factor levels. The authors contend that the nine dimensions form the core group which cannot be neglected while developing a high IQ website.

Eleven IQ dimensions showed high mean values. IQ dimensions of *Security, Value-Added, Objectivity, Reliability, .Accessibility, Reputation, Relevancy, Timeliness, Authority, Interpretability* and *Efficiency* fall into this group. They were significantly impacted by the main effect or interaction effects. Their relative rankings will vary across factors of Domain, Nation and Type (Domain), these dimensions should be given enough attention in all spheres since any one factor or their interaction could play a role in altering user perception of IQ on the website.

Mean values of 3.41 and 3.4 for *Availability* and *Ease of Operation* are less than 3.5 which are needed to qualify in the high mean quadrant. Neither do the numbers qualify for the low mean quadrant (3.0 and below). Both IQ dimensions are impacted

significantly by Domain and Nation. Thus while in this research they lie in the zone of indecision, it is possible that for other domains they might qualify for quadrant I or slip down to quadrant IV. Both the dimensions have been ranked the lowest in the IQ dimensions. Also as future work it will be interesting to do a factor analysis and see if these 22 dimensions can be categorized as four to five major categories.

One Way ANOVA was used to understand the effect of type of website on the IQ dimensions. It was also used as a method to validate the levels of websites which were selected within each domain. The ANOVA result showed that for e-commerce none of the mean values with importance ratings for the IQ dimensions were different across the two levels of the domain. Hence it was validated that type of website did not play a role in user attaching "value" to a particular IQ dimension. While in agreement aspect, mean values for 20 out of 22 dimensions were significantly higher for higher ranked website (www.amazon.com) when compared to those of lower ranked website (www.planetonline.com). The same pattern was seen for the NEWS domain. Thus it can be argued that any website which has a better presence of IQ dimensions will surely have better standards of Information Quality.

This study has brought out some interesting findings. The authors' argument that World Wide Web is not a homogeneous entity but a sum of parts has not been rejected by the results of this study. However one of the limitations of the study was domain has not been used in its traditionally understood meaning. The meaning as used in this study is actually a big section within a .com domain and the use of actual domains was beyond the scope of this work. It would make an interesting future research to study the IQ

dimensions in the traditional definition of web domains i.e. a '.edu' vs. a '.com' vs. a '.gov' vs. a '.org' domain.

As a part of future research if the set of nine dimensions is incorporated as a part of the search engine algorithm then hopefully the search results for a query will direct the user to better websites with higher information quality.

APPENDIX A

E-mail Pre-notice

I am writing to ask you to help us better understand user perspective on the quality of content presented on the internet. I am a graduate student in the industrial Engineering department at UNL. I am working on my thesis titled "Information Quality on the World Wide Web – A user perspective".

You are one of a sample taken from the UNL faculty and students randomly selected for this study. If you agree to participate in this online survey, you will receive an e-mail detailing the steps. The survey should take about 30-45 minutes to complete and is approved by the Institutional Review Board. The survey is confidential and your participation is voluntary.

In case you prefer to receive the survey at a different e-mail address or if you do not wish to be contacted further regarding this research please drop us a brief e-mail at jaikritkandari@huskers.unl.edu or rbishu@unl.edu . You are welcome to contact us at 402-613-6650 or 40-472-2393 in case you have any questions or need any further clarifications. If you have enquiries about your rights as a research participant please contact the Institutional Review Board at 402-472-6965.

I hope that you would agree to participate in this important project to help understand web human interaction better.

Sincerely,
Jaikrit Kandari
Graduate Student
Industrial and Management Systems Engineering, 175 NH
University of Nebraska Lincoln

APPENDIX B

E-mail Invitation – www. amazon.com

Dear

Thank you for volunteering to participate in the research survey. The survey should take about 30-45 minutes to complete. **You will be testing the website: www.amazon.com.** Please complete the survey in ONE session for smoother data collection.

The link to the URL for the SURVEY is:
http://www.surveymonkey.com/s/3egi

Use one of the methods to open the link
Click on the link to open the website in a new browser window.
Open a new browser window. Copy the link above and paste it in the address bar.
Open a new browser window and type the URL in the address bar.

The survey is confidential and your participation is voluntary. If you have any questions about being part of the study you may contact us via email or phone at 402-613-6650, 402-472-2393, jaikritkandari@yahoo.com or rbishu@unl.edu . If you have any questions about your rights as a research participant please contact the Institutional Review Board at 402-472-6965.

I appreciate your participation in this important project to help us information quality from a user perspective.

Sincerely,
Jaikrit Kandari
Graduate Student
Industrial and Management Systems Engineering, 175 NH
University of Nebraska- Lincoln

APPENDIX C

E-mail Invitation – www.planetonline.com

Dear

Thank you for volunteering to participate in the research survey. The survey should take about 30-45 minutes to complete. **You will be testing the website: www.planetonline.com** Please complete the survey in ONE session for smoother data collection.

The link to the URL for the SURVEY is:
http://www.surveymonkey.com/s/1ebi

Use one of the methods to open the link
Click on the link to open the website in a new browser window.
Open a new browser window. Copy the link above and paste it in the address bar.
Open a new browser window and type the URL in the address bar.

The survey is confidential and your participation is voluntary. If you have any questions about being part of the study you may contact us via email or phone at 402-613-6650, 402-472-2393, jaikritkandari@yahoo.com or rbishu@unl.edu . If you have any questions about your rights as a research participant please contact the Institutional Review Board at 402-472-6965.

I appreciate your participation in this important project to help us information quality from a user perspective.

Sincerely,
Jaikrit Kandari
Graduate Student
Industrial and Management Systems Engineering, 175 NH
University of Nebraska- Lincoln

APPENDIX D

E-mail Invitation – www.bbc.co.uk

Dear

Thank you for volunteering to participate in the research survey. The survey should take about 30-45 minutes to complete. **You will be testing the website: www.bbc.co.uk** Please complete the survey in ONE session for smoother data collection.

The link to the URL for the SURVEY is:
http://www.surveymonkey.com/s/7ngi

Use one of the methods to open the link
Click on the link to open the website in a new browser window.
Open a new browser window. Copy the link above and paste it in the address bar.
Open a new browser window and type the URL in the address bar.

The survey is confidential and your participation is voluntary. If you have any questions about being part of the study you may contact us via email or phone at 402-613-6650, 402-472-2393, jaikritkandari@yahoo.com or rbishu@unl.edu . If you have any questions about your rights as a research participant please contact the Institutional Review Board at 402-472-6965.

I appreciate your participation in this important project to help us information quality from a user perspective.

Sincerely,
Jaikrit Kandari
Graduate Student
Industrial and Management Systems Engineering, 175 NH
University of Nebraska- Lincoln

APPENDIX E

E-mail Invitation – www.thestar.co.uk

Dear

Thank you for volunteering to participate in the research survey. The survey should take about 30-45 minutes to complete. **You will be testing the website: www.thestar.co.uk** Please complete the survey in ONE session for smoother data collection.

The link to the URL for the SURVEY is:
http://www.surveymonkey.com/s/5nbi

Use one of the methods to open the link
Click on the link to open the website in a new browser window.
Open a new browser window. Copy the link above and paste it in the address bar.
Open a new browser window and type the URL in the address bar.

The survey is confidential and your participation is voluntary. If you have any questions about being part of the study you may contact us via email or phone at 402-613-6650, 402-472-2393, jaikritkandari@yahoo.com or rbishu@unl.edu . If you have any questions about your rights as a research participant please contact the Institutional Review Board at 402-472-6965.

I appreciate your participation in this important project to help us information quality from a user perspective.

Sincerely,
Jaikrit Kandari
Graduate Student
Industrial and Management Systems Engineering, 175 NH
University of Nebraska- Lincoln

APPENDIX F

Task Scenarios – www.amazon.com and www.planetonline.com

TASK SCENARIOS:

The task scenarios and questions that follow in the survey are related to the website: www.amazon.com

Before we begin the questions, please complete the following three task scenarios to get some idea about the website. The tasks are designed to help you give a better feedback to the survey questions. You do not need to answer any questions but kindly make a mental note of your experience and your impression of the website as you complete these tasks.

Task 1

Open website www.amazon.com using any browser (Internet Explorer, Mozilla Firefox, Google Chrome etc)

Understand the basic layout of the homepage.

Check out the main tabs/links, the font type, size and color across the different web pages and come back to the homepage from any link page that you opened in this process.

Check for any copyright information on the homepage? Also look for any contact information

Watch out for advertisements if any?

Task 2

Search for the book named "Good to Great by Jim Collins (Hard Cover)" or check out some electronic item of your choice. Check for information provided about the item, pricing and purchasing options.

Check for security features provided for use of credit card or personal information.

Task 3

Check if options for registration, login and sign off are available.

Check if option is available to contact customer care or leave comments, queries etc.

APPENDIX G

Task Scenarios – www.bbc.co.uk and www.thestar.co.uk

TASK SCENARIOS:

The task scenarios and questions that follow in the survey are related to the website: www.bbc.co.uk

Before we begin the questions, please complete the following three task scenarios to get some idea about the website. The tasks are designed to help you give a better feedback to the survey questions. You do not need to answer any questions but kindly make a mental note of your experience and your impression of the website as you complete these tasks.

Task 1

Open website www.bbc.co.uk using any browser (Internet Explorer, Mozilla Firefox, Google Chrome etc)
Understand the basic layout of the homepage.
Check out the main tabs/links, the font type, size and color across the different web pages and come back to the homepage from any link page that you opened in this process.
Check for any copyright information on the homepage? Also look for any contact information
Watch out for advertisements if any?

Task 2

Check out a few top stories in sports or politics and their presentation.
Check the sites disclosure on privacy and security, if any

Task 3

Check if option is available to contact customer care or leave comments, complaints, queries etc.
Check if relevant credit is given to information source