Web Engineering for the Evaluation of Large Complex Web Systems: Methodologies in Web Metrics

Dalal Ibrahem Zahran

A thesis submitted in partial fulfilment of the requirements of Edinburgh Napier University, for the award of Doctor of Philosophy

August 2016

Abstract

Roaming the Internet, users sometimes encounter severe problems or feel dissatisfied using a particular site. E-government websites are the public gateways to access information and services but there is still no agreement on how to assess a government's online presence. Failure of e-government projects in achieving their goals is common and there is uncertainty about how best to evaluate an e-government website. It has been argued that existing evaluation frameworks have some methodological limitations and they mostly neglected citizens. There is a lack of an engineering approach for building web systems and the literature on measuring the quality of website is limited. There is an uncertainty in the selection of evaluation methods and some risks of standardizing inadequate evaluation practices. Managing the complexity of web applications, Web Engineering is emerging as a new discipline for the development and evaluation of web systems to promote high-quality websites. But web quality is still a debatable issue and web metrics is considered a valuable area of ongoing research. Therefore this research focuses on the methodological issues underlying web metrics and how to develop an applicable set of measurement for designing websites. The main aim is to create new metrics for web engineering and develop a generalizable measurement framework for local e-government since research in this field is limited. This study adopted a positivist quantitative research and used triangulation web evaluation methods (heuristic evaluation, user testing, automatic link checkers, and Alexa) to test multiple-case study of Saudi city websites. The proposed E-City Usability Framework is unique in integrating 3-dimension measures (website usability, e-services, and the number and type of e-services), and in using multi-orientations to cover several aspects of e-government: output (information and services), outcomes (citizen-centricity indicators), model, and model-based assessments.

Existing e-government models were criticized, and the findings employed in developing the proposed framework. The best web evaluation methods were heuristic evaluation and user testing, while link checkers and Alexa proved to be unreliable tools; nevertheless, they can be used as a useful complementary approach. Saudi city websites were ranked by website quality, e-services, and overall evaluation. Common usability problems in these websites were found to be: the sites were not citizen-centered, limited e-services and information, no e-transaction, no emergency alerts, no municipal budget, and no city council reports. They also suffered from broken links, an inactive city map, a poor eComplaint section, and a nonfunctioning search facility.

Dedication

To the memory of my father, God mercy upon him, for his belief in education, I dedicate this dissertation. He was the great person who taught me the value of study and the wealth of reading.

Acknowledgement

I would like to thank Allah Almighty for answering my prayers, for giving me the ideas, for providing the inspiration, and for helping me finish this dissertation. From the beginning of this research till the end I kept feeling that God was there in every detail and every minute.

I sincerely dedicate this study to my parents, my husband, my brothers, and my children. My mother is the angel who enlightened my road with continuous prayers and love. Also, this work could not be done without the great support of my husband. Further, my deepest gratitude goes to my sister who always helps and cares, to my brothers who are always encouraging and concerned, and to my children who surrounded me with their love.

My very special thanks go to my advisers, Dr. Malcolm Rutter, Dr. Hana Al-Nuaim, and Professor David Benyon who taught me a lot in every step of this dissertation. Their enormous efforts, valuable advice, intelligent guidance and continuous support are fully appreciated. I am deeply indebted to their suggestions and continue to be grateful.

Also, I'd like to thank the experts who participated in the test and gave me their precious time wishing them all the best especially Maysoon Abulkhair, Uhood Zamzami, Miadah Almasry and Lulwah Al-Harigy.

Finally, I'd like to present this dissertation to my beloved country, Saudi Arabia. Full respect and genuine prayers go to whoever is responsible for e-government and for the progress and comfort of the simple kind people living in this great country.

Table of Contents

List of Figures

Chapter 5

Chapter 6

List of Tables

Chapter 1

Introduction

The impressive power of Information and Communication Technology (ICT) and its technological advancements has a great impact on society and our lives. It continues to transform the way we learn, communicate, and do business with the private sector as well as governments. The tools of ICT such as computer equipment, the Internet, and mobile devices, act as different access methods to reach the world and to connect people. While it took 75 years for the telephone to reach 50 million users when it was invented, it has taken the World Wide Web only 4 years to reach the same number of users (Frey, & Osborne, 2015). This revolution of information technology continues at a rapid pace. How to deal with the positive and negative effects of these technologies are of a big concern and of a great challenge that nations must face.

The quality of websites, in reality, is often unsatisfactory and designers ignore basic web attributes like usability (Kulkarni, & Dixit, 2012). The design of everyday objects sometimes leaves the user frustrated and unable to complete a simple task. Twenty-first-century information-seekers have little patience for confusing interfaces or difficult navigation websites (Chow, 2013). Industry and business have long since understood consumers will not tolerate products that are not usable (Lee, & Kozar, 2012). If users fail to access and execute e-services due to design shortcomings, their dissatisfaction may prevent them from returning to that particular website (Huang, & Benyoucef, 2014).

Moreover, if the design is not user-centered, usability problems can have broad and severe negative impacts on websites' development and user satisfaction (Yan, & Guo, 2010). Poor human-machine interface design may lead to many problems and have a profound effect on users. Unfortunately, products and services with poor usability are still entering the market. A striking example is a nationwide public transport card in the Netherlands which caused numerous usability problems, resulting in a public outcry and resistance against the system (van Kuijk, van Driel, & van Eijk, 2015). Another example, a

third of medical device incident reports to the US Food and Drug Administration involved usage errors (Bartoo, & Bogucki, 2013). Here a usability is important because it means the difference between a widely accepted product and one that subject to recall. Also, usability problems are among the reasons for the underuse of e-government websites (Donker-Kuijer, de Jong, & Lentz, 2010).

Human Computer Interaction (HCI) is the study of how humans interact with computer systems narrowing the gap between machines and people and building bridges between hardware and humans (Yan, & Guo, 2010). The main goal of HCI is to construct systems that people find usable and useful. Since humans interact with computers through a user interface, Nielsen (2011) argues that users don't need machines to be friendly with them; they just need machines that will not stand in their way when they use them. Thus, he was the first to use the term "usability" of the interface (Buie, & Murray, 2012). The design of the user interface, or more specific, the usability of the interface is a core area in the field of HCI (Yan, & Guo, 2010; Li, Yu, & Liu, 2010). Business and industry are pioneers of the usability movement; Microsoft, IBM, and Webby awards are just a few in a long list of major companies who take usability guidelines into consideration when designing their websites (Lee, & Kozar, 2012). Also, the U.S.A. government's websites were subjected to a major overhaul through following the Research-Based Web Design and Usability Guidelines created in the year 2000 by the Department of Health and Human Services (Buie, & Murray, 2012).

In 1980 Usability Engineering has emerged and 1990 saw the rise of HCI in general (Buie, & Murray, 2012). Web Engineering, as a particular area of HCI, appeared on the scene in 1997 then became an accepted discipline from 2002 (Torrecilla-Salinas, et al., 2016). Web Engineering has adopted and improved HCI methods to be applied to web applications and other new usability evaluation methods, specifically crafted for the web, have been also developed (Fernandez, Abrahão, & Insfran, 2013). Web Engineering is defined as the set of techniques, tools, and methods that help designers develop systems on the web (Torrecilla-Salinas, et al., 2016).

Interest in the field of e-government has begun to shift slowly to more citizen-related issues such as usability and user preferences (Lofstedt, 2012). This is due to the fact that municipalities influence citizens' lives. Surveys in Europe show that 50% to 80% of the interaction between citizens and government occurs at the local levels (Montserrat, 2010).

Often e-government is evaluated by a comparing technique based on indicators that yield some sort of e-score (Zahran et al., 2015). A framework or a model defines website quality requirements by identifying measurable attributes that are further decomposed into a set of metrics, and these practices represent a method (Kulkarni, & Dixit, 2012). This will show ways for developers to engineer websites in a more user-friendly approach to improve web applications. Using web metrics assess the quality of the web engineered product and a large number of them have been proposed in the last decade. Developing models are considered as a basis to guide and measure e-government progress (Singh, Malhorta, & Gupta, 2011). However, there is a shortage of e-government models at the local and national level (Lofstedt, 2012). Nam (2014) and De Róiste (2013) have noticed that most countries launch e-government through the quick fix quick wins principle and hastily construct the e-equivalent of a bureaucratic administration while focusing on citizen-centric websites should be at the core of e-government and municipalities need to acknowledge and work to improve the online citizen-government relationship (Tsohou et al., 2013; Moraru, 2010).

1.1 ICT and Internet Usage

Internet penetration indicates the degree of progress accomplished by a country's population. How many people online reflect the level of technology awareness. Unfortunately, the world still continues to be separated by major differences in terms of the digital gap. The Internet World Statistics for the year 2016 reveal that the highest region in the world in Internet population penetration is North America at 89.0% with 320,067,193 Internet users (Table 1.1, Figures 1.1, and Figure 1.2). The Middle East is far behind with a penetration rate of 53.7% and 132,589,765 Internet users. It is ranked the 5th region amongst the world's eight regions, but the statistics also reveal that the usage growth of the Internet in the Middle East is, as much as 3,936.5% between the year 2000 and 2016, which is the second highest among the world regions (Internet World Stats, 2016).

Table 1.1: World Internet Usage Statistics for the Year 2016 (Internet World Stats, 2016)

WORLD INTERNET USAGE AND POPULATION STATISTICS JUNE 30, 2016 – Update						
World Regions	Population (2016 Est.)	Population % of World	Internet Users 30 June 2016	Penetration (%Population)	Growth 2000-2016	Users % of Table
Africa	1,185,529,578	16.2 %	**339,283,342**	**28.6 %**	7,415.6%	9.4 %
Asia	4,052,652,889	55.2 %	**1,792,163,654**	**44.2 %**	1,467.9%	49.6 %
Europe	832,073,224	11.3 %	**614,979,903**	**73.9 %**	485.2%	17.0 %
Latin America	626,054,392	8.5 %	**384,751,302**	**61.5 %**	2,029.4%	10.7 %
Middle East	246,700,900	3.4 %	**132,589,765**	**53.7 %**	3,936.5%	3.7 %
North America	359,492,293	4.9 %	**320,067,193**	**89.0 %**	196.1%	8.9 %
Oceania/ Australia	37,590,704	0.5 %	**27,540,654**	**73.3 %**	261.4%	0.8 %
WORLD TOTAL	7,340,093,980	100.0 %	3,611,375,813	49.2 %	900.4%	100.0 %

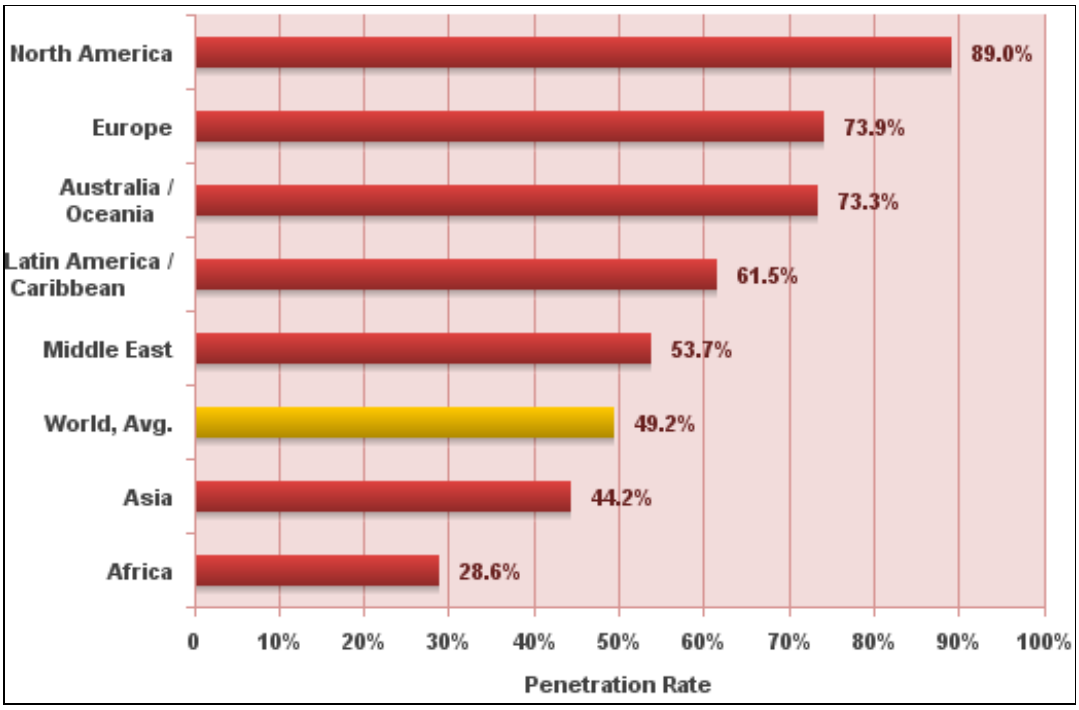

Figure 1.1: World Regions Internet Penetration Rates (Internet World Stats, 2016)

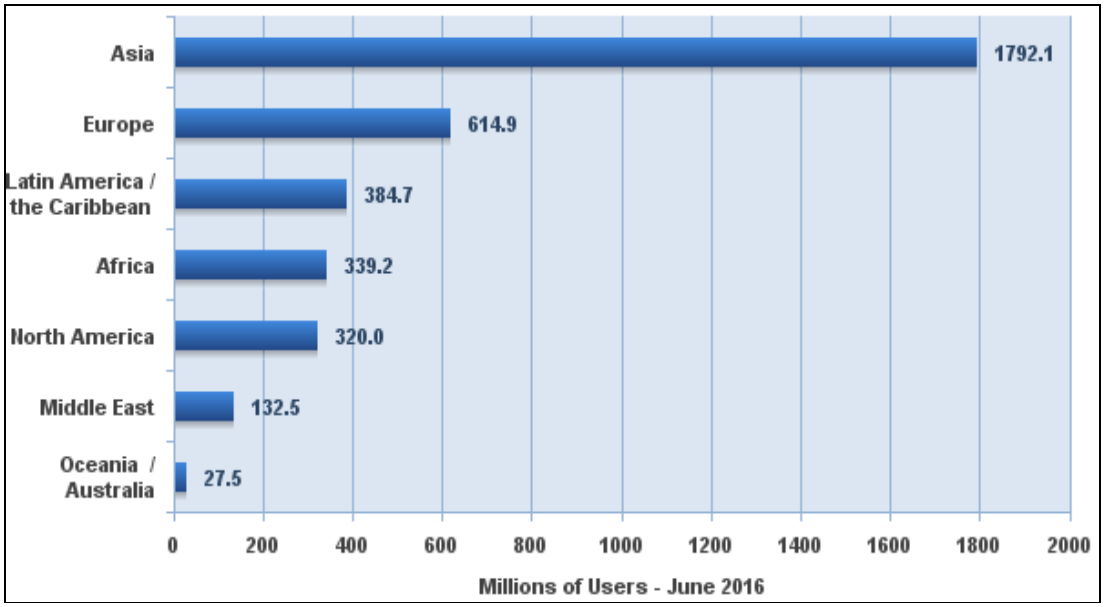

Figure 1.2: World Internet Users (Internet World Stats, 2016)

1.2 Scope and Limitation

In the more recent past, many e-government projects failed to deliver their promises in terms of obtaining important benefits and specific outcomes (Luna-Reyes & Gil-García, 2011). A targeted purpose of e-government is to increase effectiveness and efficiency of services but few studies, only recently, attempted to integrate indicators of e-government outcomes into assessment (Fraefel, Selzam, & Riedl, 2013). According to Anthopoulos et al. (2016), e-government outcomes are being questioned and several researchers debate about its potential. The authors claimed that the failure of e-government project ranges from partial failures to complete abandonments because of missing business needs and end-user satisfaction from adoption.

Reviewing scholar publications of e-government in the United States from 2007 to 2011, Snead and Wright (2014) found the mainstream is: 58% of sample studies on output, 24% on outcomes, 9% on models, 6% on processes, and 3% on theory-based assessments. Based on all the above, the research orientation for this study is intended to cover four perspectives: output (information and service on the website), outcomes (citizen-centricity indicators such as citizen satisfaction), models, and model-based assessment.

Based on the type of relation, e-government can be classified into four main categories (Figure 1.3): Government to Citizens/Customers (G2C), Government to Business (G2B), Government to Government (G2G) and Government to Employees (G2E) (Alshibly, & Chiong, 2015). This research focuses on the Government to Citizens (G2C) relationship applied at the level of local e-government only; that is e-municipalities or e-cities. In this context, the "website" is considered the main delivery channel to access e-government.

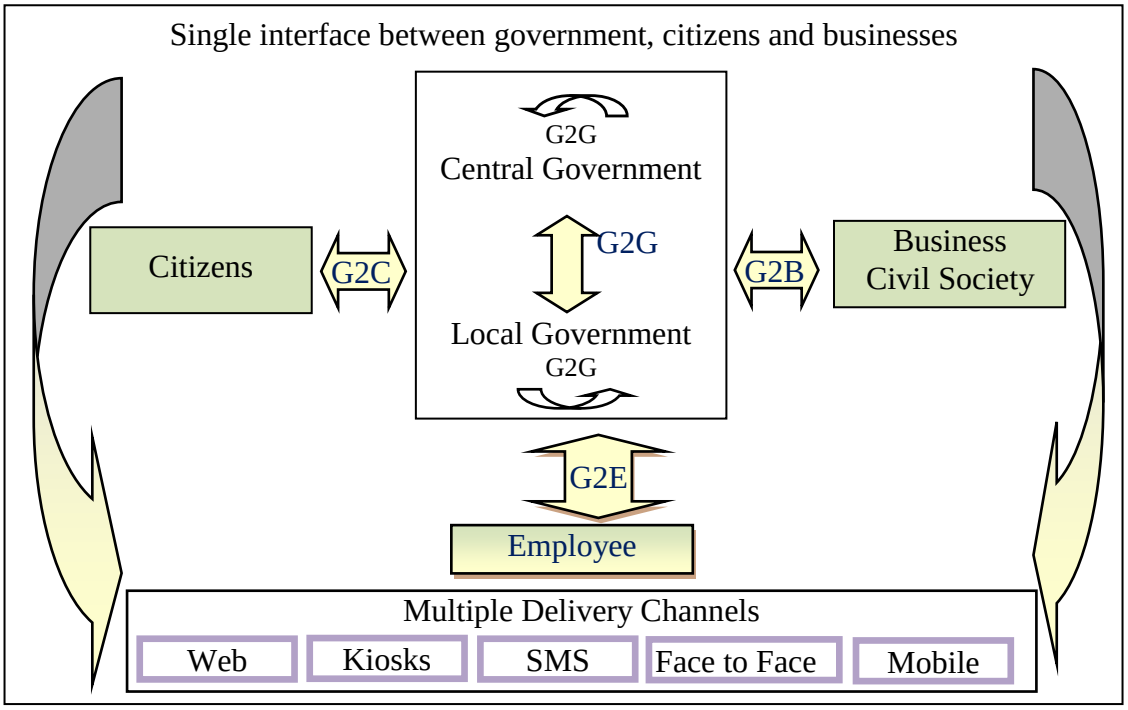

Figure 1.3: E-government Classification and Delivery Channels

People do not need a website designed to emphasize the internal bureaucratic structure, or promote the minister or department head at the expense of granting citizens fast access to services and information they need. When designing their websites, governments must not mirror their images only, but also prioritize user needs and satisfaction to pave the way toward a citizen-centered e-government approach that would attract citizens and benefit society as a whole (Ali, & Ahmad, 20015; Buie, & Murray, 2012). Therefore, the scope of this study covers the relation and interaction between the citizens and the municipal websites in the light of the main usability attributes and user-centered design models measuring the web presence of e-cities.

1.3 Statement of the Problem

Several scholars of e-government are skeptical about e-government rankings and have justifiably argued that existing e-government frameworks have some methodological limitations (Schellong, 2010; Rorissa, Demissie, & Pardo, 2011; Sandoval-Almazan & Gil-Garcia, 2008b; Debri, & Bannister, 2015). Their analysis shows a confusing picture of the measurement of e-government. A good evaluation framework for e-government at the national level is still lacking (Ataloglou & Economides, 2009; Lofstedt, 2012) as well as at the local level (Montserrat, 2010; Lofstedt, 2012). Karkin and Janssen (2014) stated that a universally accepted e-government model still needs to be developed.

In general, research assessing the limitations of e-government models is scarce (Sandoval-Almazan, & Gil-Garcia, 2008b; Lofstedt, 2012). Most of these models focused on the supply side (government) not the demand side (citizen and business) of e-government (Berntzen, & Olsen, 2009). The existing practices are pushing countries to prioritize getting good ratings for creating many services without caring whether citizens use them or not (Montserrat, 2010). Lofstedt (2012) assured that most studies focus on national e-governments at ministries' websites and rarely do they shed a light on local e-governments performance, although they are the main contact for the delivery of services, especially in developing countries. Local e-services delivery are still not paid sufficient attention even though they are very important to citizens.

Focusing on the citizens is the core of e-government and municipalities need to acknowledge the citizens' role, put it into practice and work toward improving the online citizen's government relationships. Though being a customer is only one aspect of citizenship, the culture of focusing on citizens as customers is missing in the Arab world and at the same time people's demands of e-services are escalating dramatically. Al-Nuaim (2009) stated that Arabic citizens receive their basic services from bureaucratic government agencies where employees are rarely trained in customer service or reprimanded for inefficient work and complaints from citizens are often ignored. Moreover, the Saudi IT Criminal Law does not define the privacy right nor does it mention any punishments that would be applied to companies and websites owners who do not protect their visitors'

privacy (Al-Ghaith et al., 2010). Alshehri and Drew (2010) stressed that a significant challenge is to deal with security, privacy and trust issues in Saudi governmental websites.

Although how to measure the quality of website has become a valuable area of ongoing research, the field is not yet mature (Lofstedt, 2012). Web evaluation methods are abundant in the literature yet lack studies that classify, compare, and determine the appropriate evaluation method(s) for certain purposes. Woolrych et al. (2011) warned that research that assesses usability evaluation methods has been in crisis for over a decade because of a lack of publications. They added that there are risks that inadequate evaluation practices are becoming prematurely standardized.

In addition, failure of e-government projects in achieving its goals is high and not only remains common but also continues to escalate in many developing countries (United Nations, 2014). An example of a public sector project that failed is the US Internal Revenue Service Business System Modernization (IRS BSM). It has spanned a decade and consumed more than 3 billion dollars and finally suspended (Purao, & Desouza, 2011). Also, the Saudi first national 5-year e-government plan, of providing 150 services online by the end of 2010, was not achieved as the years passed (Alfarraj, Drew, & AlGhamdi, 2011).

1.4 Motivation of the Research

Saudi Arabia is investing heavily in e-government and has a policy of transferring services online. Therefore, it is imperative to understand more on the progress of Saudi local e-government practices through evaluating several e-city websites.

In Saudi society, people often need to obtain different services and information from government entities. Citizens are frustrated with the bureaucracy and they frequently need to be excused from work and wait in long lines for hours or days to finish their tasks. E-government promises to eliminate all this waste of effort, time, and money. By just a few clicks, citizens can obtain their services whenever and wherever at their convenience 24/7. In Saudi culture, citizen centered e-government websites is a necessity, not a luxury. The need is even greater for women because she needs a legal guardian or a representative to follow up in government offices.

There are limited studies in the literature discussing the subject of Saudi local e-government and web design guidance to e-cities. In international e-government evaluation reports, such as the United Nations (2014), the Saudi national e-government scored 0.6900 out of 1 while at local level in its 2014 Digital Governance in Municipalities benchmarking only Riyadh has been evaluated and scored 35.59 out of 100 (Holzer et al., 2014).

1.5 Objectives of the Research

Most of the citizens' interaction with the government occurs at the local level but unfortunately current literature provides little web development guidance to e-cities (Lofstedt, 2012). To fill the gap of studies in local e-government and web evaluation methods and in recognition of usability as the most important metrics for implementing successful e-municipal websites, the main objective of this research is to develop an evaluation framework to assess the quality of any city website in the world. In the aim of enhancing the development of citizen-centered e-government, the developed framework is tested within the context of Saudi municipal websites. Thus, this researcher proposed how to select the appropriate evaluation method(s) through comparing and identifying the strengths and weaknesses of existing web evaluation methods.

1.6 Purpose of the Research

The main purpose of this research is:

1) To develop an evaluation framework for municipality websites. Therefore, the following research questions will be addressed:

1a. What are the major national and local e-government evaluation frameworks and what are their strengths and weaknesses?

1b. Are the metrics defined for national e-government suitable for assessing e-municipalities?

2) To test the developed framework in the context of Saudi municipality websites, a search was conducted to classify web evaluation methods. Accordingly, the following research questions will be answered:

2a. What is the ranking of Saudi municipality websites by website quality, e-services, and overall evaluation?

2b. What are the major usability problems affecting Saudi citizen use of these websites?

1.7 Methodology

This study adopted a positivist quantitative research to evaluate multiple-case study of Saudi e-city websites using a proposed e-city framework and four web evaluation methods, as follows:

1.Development of the proposed E-City Usability Framework:

a. Three-dimension assessment criteria were integrated: 1) website quality metrics; 2) e-services quality metrics and 3) the number and type of e-services since it has been realized that one or two of these dimensions are usually absent from the measurement of existing e-government frameworks.
b. The framework builds upon the strengths of ten models (three theoretical web models and seven practical e-government models) extracted from four sources in the literature: government, international organisations, academic research and educational institutions, such as the US Research-based Web Design and Usability Guidelines, the UN Digital Governance in Municipalities, the Community Benchmarks Program, Dubai E-government Excellence Model and Gartner Model. The proposed framework tries to avoid their limitations and adds appropriate metrics if absent from all models.
c. The heuristics shared in common by the majority of these models were selected if they fulfill one of the design principles of the g-quality e-government inspection method by Garcia, Maciel, and Pinto (2005). Further, selected heuristics were checked against a Folmer, Gurp, and Bosch (2003) usability framework to identify affected usability quality attributes.

2. Refinement of the developed framework:

a. Pilot test: the framework was evaluated by a usability expert to find inconsistencies or ambiguity problems.
b. A double-expert (e-government and usability expert) review: classify guidelines as objective or subjective measures, rephrase them, and approve their categorization.
c. Application of the framework on a high-ranked city website: test the New York City website heuristically using two experts. Three forms were used: the website

quality objective guidelines form, the website quality subjective guidelines form, and the e-service quality guidelines form with appropriate tasks.

3. Selecting Saudi e-city websites: the Saudi regions with the highest number of Internet users were identified, then the largest city in each region was chosen.
4. Multiple web evaluation methods (triangulation) to test Saudi e-city websites:
    a. Heuristic evaluation: two experts tested selected e-city websites, based on the website quality objective guidelines form of the proposed e-city framework (33 guidelines).
    b. User testing: five to eight users tested all selected e-city websites as follows:
        i. Users walked-through selected Saudi e-city websites to test six subjective website quality guidelines (34 to 39) of the proposed framework.
        ii. Users were given five usability tasks to evaluate the e-services in each e-city website.
        iii. Users filled in a user satisfaction questionnaire for each city website.
    c. Alexa Web analytics tool: nine metrics were calculated (domestic and global traffic ranks, page views/visitor, the speed of download, bounce rate, sites linking in, time on site, audience geography and where do visitors go on the site) as a form of validation of this tool.
    d. Automatic link checkers: seven tools were compared: Broken Link Checker, Dead Link Checker, NetMechanic, LinkTiger, Link Alarm, Web Link Validator, and Xenu, to find out which is the most reliable automatic link checker tool.
5. Outcomes from all evaluation methods (Heuristic evaluation, user testing, link checker, and Alexa) were evaluated. Results from heuristic evaluation and user testing were compiled to rank Saudi e-city websites and to identify the usability problems for each city website.

1.8 Contributions of this Research

This research aims to make several contributions on the following topics:

1. It developed a three dimension measurement framework for local e-government (E-City Usability Framework), assessing website quality, e-service quality, and the

number and type of e-services, since research in this field is mainly limited to one dimension and very little attention is given to the local level.

2. It uses multi-orientations that cover four e-government perspectives: output, outcomes (citizen-centricity indicators such as, greater information access, service quality, and citizen satisfaction), model, and model-based assessment, since most research assess narrow aspects of e-government.
3. It compares different methodologies in web evaluation, identifying their strengths and weaknesses, and proposes how to select the appropriate evaluation method(s).
4. It uses triangulation web evaluation methods to test Saudi e-city websites.
5. It checks the validity of the automatic link checker and the Alexa web analytics tool since rare studies evaluate their effectiveness in assessing websites.
6. It defines the current state of Saudi city websites and determines the potential problems encountered by users to benefit Saudi municipalities.

1.9 Structure of this Dissertation

The dissertation is organized into seven chapters including this introduction. The second chapter is the literature review which reports published research on web systems, web application engineering, web usability design principles, web quality and metrics, web evaluation methods, e-government frameworks at national and local levels, and finally Saudi e-government. The third chapter is the about the research methodology. It is followed by the development of the proposed E-City Usability Framework in chapter 4, which introduces the development process to obtain the proposed framework, the refinement steps, the inter-rater reliability, and the scoring method. Chapter 5 is about testing; it compares different web evaluation methods to determine the appropriate ones for testing e-city websites, and it describes the procedures for all usability testing methods. Chapter 6 is the result chapter and finally chapter 7 is the conclusion chapter.

Chapter 2

Literature Review

2.1 Web Systems

The evolution of the web can be traced back to 1990. Tim Berners-Lee at the European Particle Physics Laboratory wanted to exchange scientific information among researchers. He employed Internet technologies and merged them with hypertext link topologies creating a system that he called "World Wide Web" (Aghaei et al., 2012).

Since 1990 the growth of the web has increased significantly at a rapid pace. Top websites such as Google, Facebook, YouTube, Yahoo, Windows Live, Blogger, Wikipedia and Twitter are changing the world. For example, among the top USA websites in 2016, 1) Google has over 227 Million monthly unique visitors, 2) YouTube over 215 Million visitors, 3) Facebook over 129 Million visitors, 4) MSN over 127 Million visitors, 6) Amazon over 88 Million visitors, and 8) Twitter over 80 Million visitors (Quantcast US Site Rankings, 2016).

Based on their key features and technology used, web systems were classified as (Nath et al., 2014; Aghaei et al., 2012):

- Web 1.0 (a web of information) was a collection of the static read-only web in which users read information but cannot interact.
- Web 2.0 (a web of connecting people), which was defined in 2004, became a dynamic read-write web with a two-way platform. It permitted people to interact and share information online in new ways, such as blogs and wikis.
- Web 3.0 (a web of knowledge and co-operation) is about the Semantic Web and personalization. In the Semantic Web, machines can read information as humans can, applications are pieced together, the data are in the cloud and it can be run on any device (PC or mobile).

As a result, a rapid shift from simplicity to complexity has been remarkably noticed in Web evolution (Table 2.1).

Table 2.1: Web Evolution (Kumar, & Sangwan, 2011)

Simple Web	Advanced Web
Simple Web pages	Complex Web pages
Less emphasis on aesthetics/user interface	More emphasis on aesthetics/user interface
Information content static	Information is dynamic
Simple navigation	Complex navigation
Stand-alone systems	Integrated with database and other systems
High performance wasn't a major requirement	Requires high performance
Development by a single individual/small team	Requires a large development team
Used for information in none core applications	Developed in mission-critical applications

Therefore, web development becomes a complex and challenging process that must deal with a large number of heterogeneous interacting components that demand high performance systems. Information must be up to date, new functionality should be added and the whole system continues to evolve with time. Certainly, web systems should be available continuously, function well from diverse browsers and computers, and bear pressure from a large number of users (Kulkarni, & Dixit, 2012). Many factors contribute to the complexity of web systems (Figure 2.1). Therefore, developing large web systems should follow a systematic engineering approach (Kaur, & Dani, 2011).

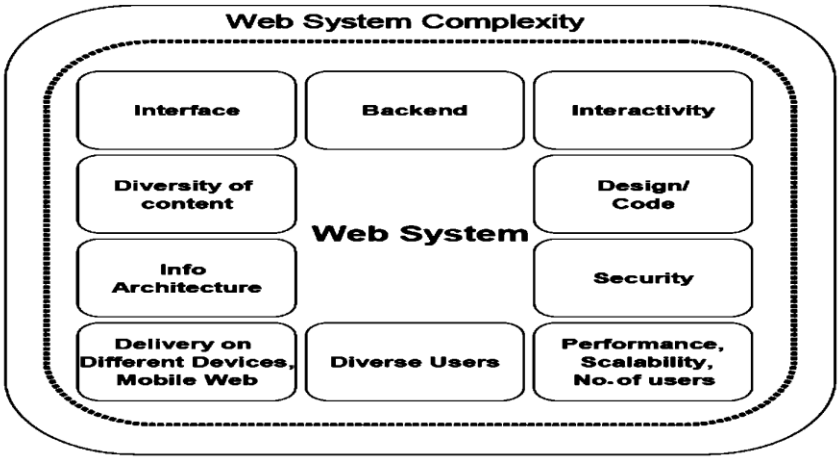

Figure 2.1: Landscape of Web Systems

2.2 Engineering of Web Applications

Although the web has become increasingly complex, the development process is still un-engineered (Karkin, & Janssen, 2014; Kumar, Dadhich, & Shastri, 2015). The development of web applications has made some improvements, but there is still a lack of an engineering approach for building web systems. In this context, "Engineering" advocates a systematic approach to develop high-quality web applications (Kumar, & Sangwan, 2011). Relative to this research, an overview of three engineering disciplines: Web Engineering, Software Engineering, and Usability Engineering, are explained next.

2.2.1 Web Engineering

Ironically, the best and also the worst thing about the Internet is that almost anyone can post a website. An ad-hoc development approach, which does not follow any method or standard, to build a complex web application system can quickly lead to poorly designed websites and may be problematic to many organizations (Kumar et al., 2015; Ali, & Ahmad, 2015). Mikkonen and Taivalsarri (2010) argue that web development is still far from maturity levels of software engineering. In turn, end users encounter some problems with an unfavorable user experience. Since 1996 in his Alertbox column, Nielsen (2016) publishes a series of top ten web design mistakes based on heuristic tests of many widely used websites. He discovered that the same web design mistakes occurred over and over again. Top 10 mistakes in web design for 2011 are (Nielsen, 2011): bad search, not changing the color of visited links, non-scannable text, fixed font size, page titles with low search engine visibility, thing that looks like advertisement, violating design conventions (consistency), opening new browser windows and not answering users' questions.

To overcome all these web shortcomings, there is a genuine need to adopt methodologies and better techniques for implementing successful websites. Progressively, Web Engineering is emerging as a new discipline addressing the unique needs and challenges of web systems. Recently, significant progress has been made in turning Web Engineering into an engineering discipline encompassing the design, development, evolution, and quality evaluation of web applications (Mikkonen, & Taivalsarri, 2010; Rio, & e Abreu, 2010). Officially, Web Engineering is defined as the application of systematic,

disciplined and engineering approaches to the successful development, deployment, and maintenance of high quality web-based systems (Ali, & Ahmad, 2015).

Furthermore, Web Engineering is bound to be a multidisciplinary field with encompassing contributions from diverse subjects (Kumar, & Sangwan, 2011) such as: human-computer interaction, user interface, systems analysis and design, software engineering, information engineering, testing, modeling and simulation, and graphic design (Figure 2.2). It seems that Web Engineering is a discipline among disciplines, cutting across Computer Science, Information Systems, Software Engineering and other non-IT specializations. Main topics of Web Engineering include, but are not limited to, the following areas: web development methodologies and models, web system testing and validation, quality assessment and assurance, web metrics and web quality attributes disciplines, performance specification and evaluation, web usability, user-centric development, and user modeling. In sum, intertwining so many disciplines introduces unique problems for organization and system development, thus the need for Web Engineering is strong (Kumar, & Sangwan, 2011).

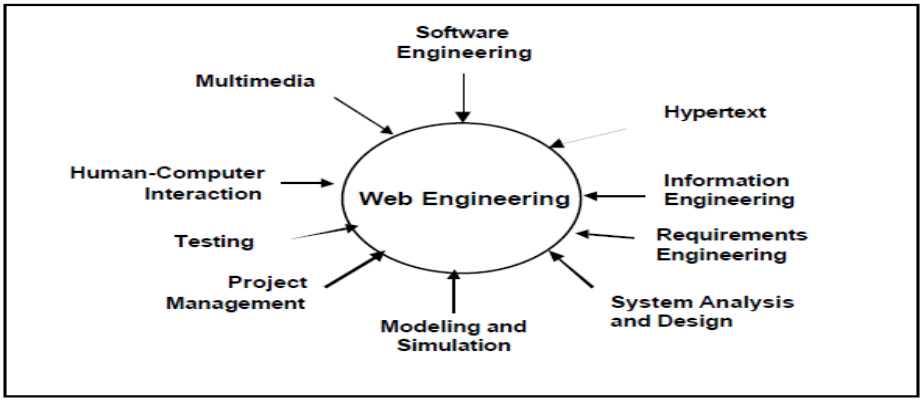

Figure 2.2: Web Engineering – A Multidisciplinary Field
(Kumar, & Sangwan, 2011)

2.2.2 Software Engineering

Software Engineering, as a discipline, emerged as a result of approaching the "software crisis" which is the apparent problem of incomplete and poorly performing software. It evolved out of the need to manage the increased size and complexity of software development. Historically, the list of software defects have plagued the software industry

and this could get worse since future systems are expected to be more complex. For example, the National Institute of Standards and Technology (NIST) estimated that software errors cost the U.S. economy $59.5 billion a year (Harter, Kemerer, & Slaughter, 2012). Further they mentioned that in the last 15 years software defects have wrecked a European satellite launch, delayed the opening of the expensive Denver airport for a year, destroyed a NASA Mars mission, induced a US Navy ship to destroy a civilian airliner, and shut down ambulance systems in London, leading to 30 deaths.

Notably and to large extent, the definition of Web Engineering is similar to the definition of Software Engineering provided by the IEEE: "Software Engineering is the application of a systematic, disciplined, quantifiable approach to the development, operation and maintenance of software" (Kumar, & Sangwan, 2011). In another word, Software Engineering is the use of techniques, methods, and tools to improve software development.

Web Engineering is similar to Software Engineering in terms of engineering discipline such as planning, modeling, construction and testing components. But while Web Engineering adopts some Software Engineering principles, it also incorporates many new approaches, methodologies, tools and guidelines to meet the unique requirements of its platforms (Kumar, & Sangwan, 2011). Hence, Software Engineering cannot be used directly for the development of web applications because its existing models do not have all the features required for the systematic development of websites. Ali and Ahmad (2015) mention that Software Engineering models, such as Waterfall model, cannot be used directly or not applicable for web systems and most researchers agree that web development is different from software systems.

Moreover, Torrecilla-Salinas et al. (2016) and Mikkonen and Taivalsaari (2010) differentiate Web Engineering as a separate field from Software Engineering and has become a discipline from 2002 to date. Similarly, Kumar and Sangwan (2011) consider Web Engineering as a newly emerging discipline in its own right. They believed Web Engineering is neither a clone nor a subset of Software Engineering, although both involve

software development. They concluded that Software Engineering and Web Engineering are different disciplines with some intersections.

2.2.3 Usability Engineering

Since its rise in the 1980s, the Human Computer Interaction (HCI) field has developed into an established branch of Computer Science, along with its fundamental concepts "Usability" and "User Centered Design" (UCD) (van Kuijk et al., 2015; Li et al., 2010). As the name implies, HCI is the study of how humans interact with computer systems, narrowing the gap between machines and people and building bridges between hardware and humans. The main goal is to construct systems that people find usable and useful (Calisir et al., 2011). According to the Association for Computing Machinery (ACM): "Usability engineering, also known as Human-Computer Interaction Engineering, is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and the study of major phenomena surrounding them" (Zahran et al., 2015).

Gradually, usability and website design have received great attention in the HCI literature (van Kuijk et al., 2015). By the 1990s, the scope of usability broadened and usability specialists study the social and organizational context in which humans learn and use computers. As a result, computer magazine software reviews include ‘usability’ as an important rating category. Many researchers as (Huang, & Benyoucef, 2014; Hasan, Morris, & Probets, 2012; Fernandez, Insfran, & Abrahão, 2011; Li et al., 2010) believe that website quality is defined principally in terms of usability. They considered also usability as a measure of success of the product, whether it is software, computer systems, or websites.

Although usability is an essential term in HCI, there is no agreement about its definition (Kulkarni, & Dixit, 2012). Different organizations and researchers have proposed different views about usability. The following definitions illustrate how usability has been perceived in 3 distinct organization standards (van Kuijk et al., 2015; Li et al., 2010):

• “The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” (ISO 9241-11)

• “The capability of the software product to be understood, learned, used, and attractive to the user when used under specified conditions.” (ISO 9126)

• “The ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component.” (IEEE 1016)

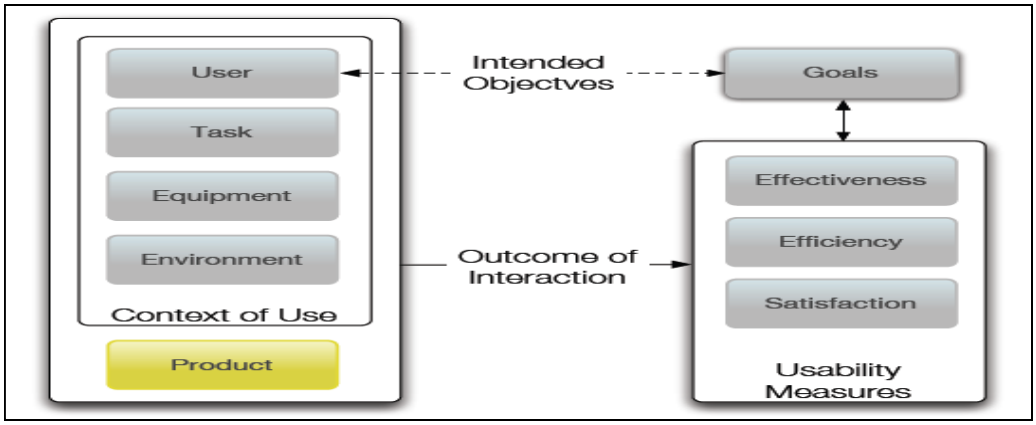

Figure 2.3: Usability According to ISO 9241-11 (Pietilä, 2010)

The ISO 9241-11 standard (Figure 2.3) takes a broader perspective on usability measurement and has been recognized as the most widely accepted definition of usability (van Kuijk et al., 2015). According to Pietilä (2010), in this standard effectiveness means the accuracy and completeness with which users achieve specified goals. Efficiency means the amount of resources, like money and people, when using the product. Satisfaction is a positive attitude towards the use of the product. The definition also takes the context of use into account which includes the users, their tasks, equipment, and the product environment. Rio and e Abreu (2010) discuss several versions of ISO till the one of 2010 (ISO25010).

The most widely accepted definition of usability used in the literature is proposed by Jacob Nielsen in 1993 (Li et al., 2010). Nielsen (1993) suggested that usability and utility together form usefulness which is the ability of a system to achieve some desired goal. The utility is the design's functionality that means it does what users need, while usability is a quality attribute that assesses how easy user interfaces are to use. Eventually usefulness, cost, compatibility and reliability lead to acceptability of a system (Figure 2.4). For Nielsen, usability is a property with five attributes (Nielsen, 1993; Calisir et al., 2011):

- Learnability: The system should be easy to learn so the user can quickly work with it.
- Efficiency: The system should be efficient to use so that the user is highly productive.
- Memorability: The system should be easy to remember, so that if the user returns to the system after some time, he doesn't have to learn from the beginning.
- Errors: The error rate of the system should be as low as possible so that users make few errors and can recover easily.
- Satisfaction: The user must like the system and feel comfortable with it.

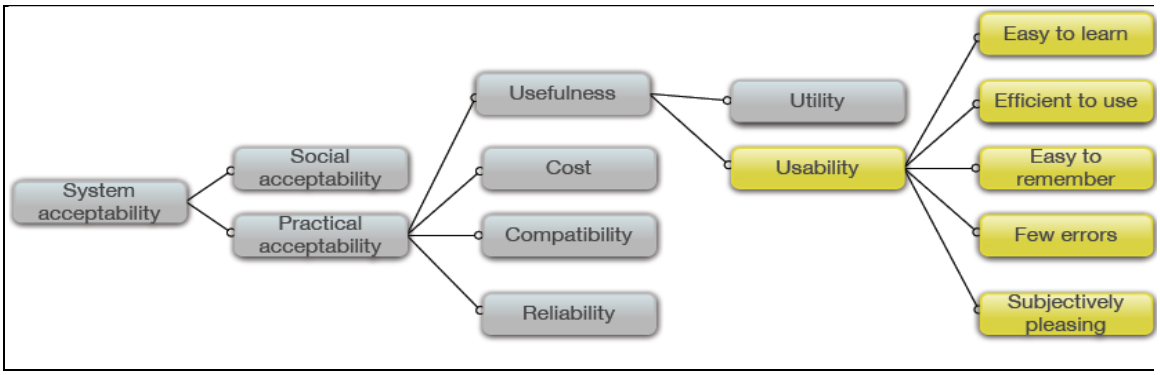

Figure 2.4: The Definition of Usability by Nielsen (Pietilä, 2010)

Li et al. (2010) stated that usability cannot be measured directly and needs to be decomposed into specific attributes, factors, and then into metrics when it is evaluated. They defined the form of measurement structure as a measure model or criteria system used to describe usability quality. The authors attributed the diversity of usability definitions to two reasons: object variety (usability study covers a broad range of systems) and divergent focus (some relate usability to business, others to social influence, or contexts, including culture and technology). The usability is a critical quality factor for interactive context and is dependent on the specific system.

On the other hand, Seffah et al. (2006) regard usability as a very confusing concept since it has not been defined in a consistent way across the standards or models and presented as high abstraction. The authors think most of these models do not include all major aspects of usability and are not well integrated into current practices. The description of the relation between metrics and high-level factors is missing plus there is a lack of aids for the

interpretation of metrics. Also, the relations between factors are not described in the models. One consequence of these weaknesses is that most web developers do not apply correctly any model in evaluating usability. This researcher doesn't agree with the argument of Seffah et al. The above explanation of Li et al. (2010) seems more accurate. Indeed, there is a growing literature for usability evaluation of website and many scholars consider usability as the most important web quality factor. Huang and Benyoucef (2014) praise usability as a well-known and well-defined in HCI research. Also, the international standards represent an agreement between global experts in measuring and defining usability. The fact that there is no global usability metric doesn't mean usability is an ill-defined concept; it only means usability metric is a valuable area of ongoing research.

2.2.3.1 User Centered Design (UCD)

UCD is a model for employing usability throughout all stages of development in order to create websites that meet users' needs and systems that place the users at the center of the design process (Chow, 2013). A UCD should consider and balance functionality and usability. The designers make sure that the user is able to use the product as intended and with a minimum effort. This approach gained popularity, prompted by a shift from a product-centric to a user-centric design (Calisir et al., 2011).

In fact, UCD can be applied to different platforms and to any system. Recently, there has been a shift toward creating a more citizen-centric e-government websites, which provides services in line with citizens' needs (Alshibly, & Chiong, 2015). With technological advances, public administrations experience a change from the bureaucratic inward-looking approach to a citizen-centric outward-looking approach that prioritizes the concerns and needs of users or customers (Nam, 2014).

If the design is not user-centered, usability problems will have broad universality, severity and a negative impact on websites' development. Poor human-machine interface design may lead to many problems. Web usability should move from a "nice to have" to a "must have" thus finding ways to reduce usability problems while designing websites (Yan, & Guo, 2010).

2.3 Web Usability Design Principles

Design principles are abstract, general and high level guides (Li et al., 2010). Design principles for usability suggest properties that have a positive effect on usability. Three well-known design guidelines, eight golden rules of interface design, Nielsen's heuristics, and the Research-Based Web Design and Usability Guidelines, are presented next just as examples.

2.3.1 Schneiderman's Eight Golden Rules of Interface Design

The user interface is the mode of interaction between the users and the system. In 1986, Schneiderman suggested eight golden rules of interface design (Shneiderman, & Plaisant, 2010):

1. Strive for consistency
2. Cater to universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Support internal focus of control
8. Reduce short-term memory load.

2.3.2 Nielsen's Heuristics

Recognized usability principles are called "heuristics". Nielsen and Molich in 1990 developed a set of 10 heuristics to evaluate screen-based products (Chow, 2013). These 10 heuristics are: visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalist design, help users diagnose and recover from errors, and help and documentation (Krenk, & McComb, 2012). However, some researchers indicated that the original heuristics were too general and vague for evaluating the web. Even though Nielsen's heuristics are general, this researcher believes that they are valuable; as an evidence, some researchers developed their own guidelines by tailoring these heuristics to their needs. For example, Garcia et al. (2005) were among the few who entail heuristic evaluation to assess the usability of e-government websites. They proposed g-quality inspection method tailored to evaluate e-government as an extension of Nielsen's heuristics evaluation method. The additional 6 heuristics were: accessibility, interoperability, security and privacy, information truth and precision, service agility, and transparency. In addition, Donker-Kuijer et al. (2010) mention that the original

Nielsen's heuristics were widely published and applied to a variety of ICT applications, including websites. Huang and Benyoucef (2014) believe that although these heuristics were developed around 20 years ago for website evaluation, they are still applicable today but need further development to address the particular needs of today's e-government.

2.3.3 Research-Based Web Design and Usability Guidelines

The Research-Based Web Design and Usability Guidelines were created by the USA Department of Health and Human Services (HHS) according to the best available up to date research to help designers build high-quality websites. The HHS guidelines have been widely used by government agencies as well as private sectors and also translated into several foreign languages. HHS has created an online version of these guidelines at www.usability.gov and explains its steps of evolution as follows (HHS, 2016):

- The project began in March 2000 with 500 guidelines. After revision, it was reduced to 398 guidelines. Each guideline shows a rating of its "Relative Importance" to the success of a website and a rating of the "Strength of Evidence" supporting the guidelines. Professional web designers, usability specialists, and academic researchers contributed to these ratings. The "Strength of Evidence" represents a consensus among researchers so the users can determine the quality of the supporting evidence.
- To determine the "Relative Importance" of each guideline, 8 website designers and 8 usability specialists evaluated each guideline and assigned a rating from 1 up to 5, the most important guidelines. To determine the "Strength of Evidence", a group of 8 usability researchers constructed judging criteria. The rating ranged from 1 up to 5, the strongest support guidelines. Consequently, the set of guidelines was reduced to 287.
- In the 2004 edition, 187 guidelines were published. Since then, it has been continually reviewed for new research information. Currently, the number of guidelines is 209.

  Some of the most important HHS guidelines (score of 5) are stated below (HHS, 2016):

  - Guideline 1: Provide useful content
  - Guideline 2: Understand and meet user's expectations
  - Guideline 3: Do not display unsolicited Windows or graphics
  - Guideline 4: Enable access to the homepage
  - Guideline 5: Avoid cluttered displays

- Guideline 6: Place important items consistently
- Guideline 7: Use clear category labels
- Guideline 8: Use meaningful link labels
- Guideline 9: Distinguish required and optional data entry fields
- Guideline 10: Facilitate scanning
- Guideline 11: Ensure that necessary information is displayed
- Guideline 12: Ensure usable search results

These guidelines are published by the HHS to be used in government websites and have been praised by many studies, such as Shneiderman (2011), Dingli and Mifsud (2011), Rinder (2012), Scowen and Regenbrecht (2009) and Buie and Murray (2012). This researcher considers the HHS guideline valuable but it is difficult to evaluate a website against so many guidelines. For example, when Nielsen succeeded in condensing usability principles to only 10 heuristics, many researchers have adopted his evaluation and built upon it. Therefore, it would be more practical if the HHS or other researchers in the field of HCI can shorten the list of Research-Based Web Design and Usability Guidelines.

2.4 Web Quality and Metrics in Web Engineering

A key element of website engineering is metrics which are used to assess and improve the quality of a web engineered product ((Singh, et al., 2011). Web metrics determine if a website performs to the expectations of the users and identify website design problems. A large number of metrics has been proposed since the 1990s (Kulkarni, & Dixit, 2012). Web metrics is concerned with quantifying different measures of websites. The first step is to define a list of factors that are important for an object and contributes to its quality and further decompose it into a set of metrics (Kaur, & Dani, 2011).

Chiou, Lin, and Perng (2010) and Bahareh (2015) noticed that different terms to evaluate a website have been used such as: criteria, factors, attributes, metrics, features, and measures, with factors being the most common term. Chiou, Lin, and Perng (2010) define factor as the set of relevant criteria. For example, “ease of use” can be a collection of criteria such as user-friendly interface, easy access to the site, or ease of navigation. They

also define a criterion to be a specific item or variable, such as loading speed, up-to-date information, or FAQ.

Singh et al. (2011) acknowledge that website engineering metrics are mainly derived from HCI, hypermedia, and software metrics. The diversity of metrics for evaluating websites indicates that there is no uniform set of metrics (Karkin, & Janssen, 2014). Therefore, web metrics is still considered a valuable area of ongoing research (Kaur, & Dani, 2011).

2.4.1 Drawbacks of Metrics

Researchers claimed that web metrics most likely define very general criteria or are not well defined (Zahran, et al., 2015; Kaur, & Dani, 2011). Also, there is a rush to develop more web metrics without empirical validations and that may lead to incorrect website evaluation (Kulkarni, & Dixit, 2012; Kaur, & Dani, 2011). Lee and Kozar (2012) criticize metrics developed in a nonscientific way, does not assess the subjective web measures, and may be company specific (e.g. Microsoft Usability Guidelines). This situation makes the use of the metrics more challenging and maybe flawed.

The lack of evaluations of existing metrics and when to reuse them leads researchers to develop more new metrics without knowing how similar these metrics are or what each metric is measuring (Vigo, & Brajnik, 2011). The authors examined automatic web metrics and deemed that the existence of several web metrics is evidence of a lack of a comparison framework that highlights how well they work and for what purposes they are appropriate.

2.4.2 Metrics and Type of Site

Web metrics are no longer one-size-fits-all and quality metrics are application domain dependent (Rio, & e Abreu, 2010; Verdegem, & Verleye, 2009). On the same point, Malak et al. (2010) mentioned that assessment models that are developed for a specific application domain, such as e-commerce or e-government, cannot be adopted to another context.

It can be argued that the design of websites is bound to follow and fulfill the firm's objectives and goal. For example, the goal of Google is to enable users to quickly identify

relevant information and leave its website quickly, thereby a clear and simple interface and instant page load increase its market and improve its brand (Chiou, et al., 2010). That means the relative importance of some features changes depending on the specific site purpose. A link rich page is a positive element for an informative site, while it probably disturbs in a service specific page where the users want to accomplish some tasks fast. Zahran et al. (2015) extend the idea further; for a website to be successful, there must be a match between the firm's objectives, the user's goals, and the website's design.

2.4.3 Research on Web Evaluation

There are many variables for measuring and defining website quality. Several studies of web quality models, such as Kumar et al. (2015) and Kulkarni and Dixit (2012), still emphasize the importance of Calero et al. (2005) research who studied published web metrics from 1992 to 2004. Using a three-dimensional web quality model (WQM), they classified 385 web metrics. The WQM defines a cube structure in which three basic aspects are considered when testing a website: web features, life-cycle processes and quality aspects (Figure 2.5).

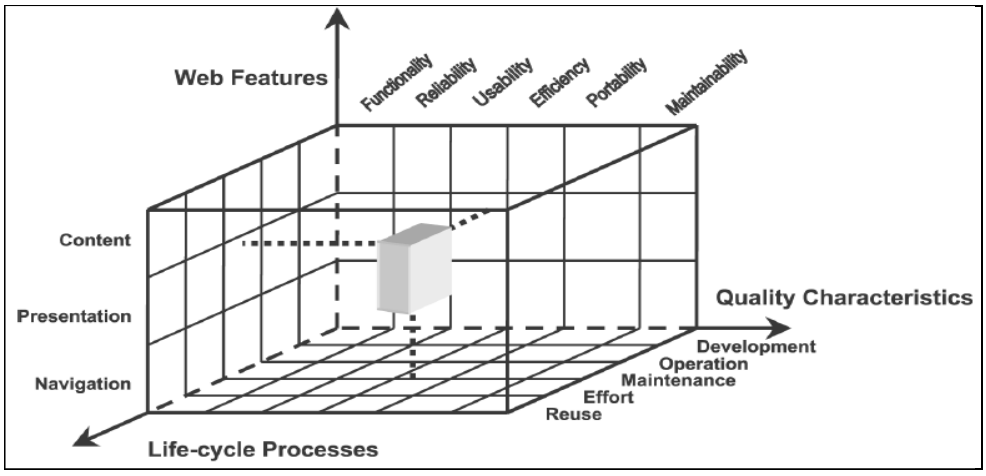

Figure 2.5: The Web Quality Model Cube (Calero et al., 2005)

The results confirm that most metrics (48% of the metrics studied) are usability metrics and 44% of them related to "presentation" (Figure 2.6). In the life cycle dimension, the majority of metrics are related to operation (43.2%) and maintenance (30%) processes. In the literature reviewed, web metric validation is not considered a major issue (theoretical

validation 3% and empirical validation 37%), a large number of metrics are automated (67%), and the triplet (usability, operation, presentation) with 149 metrics and the triplet (usability, maintenance, presentation) with 93 metrics are those with more defined metrics. In this context, validity is defined by Liu et al. (2011) as the ability of a scale instrument to measure what it is intended to measure. Validation is the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model. A validation metric provides a quantitative measure of agreement between a predictive model and physical observations. The metrics are useful either for model selection among alternative candidates or decide whether a model is acceptable before it is used for web analysis.

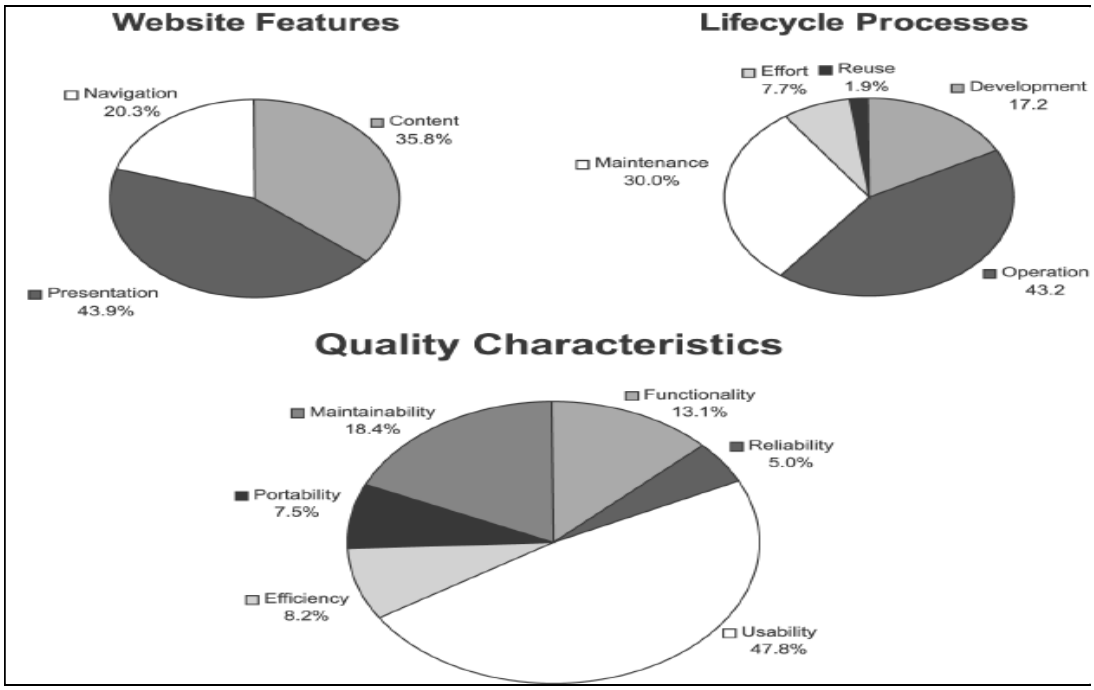

Figure 2.6: Metric Distribution across the Model Dimensions (Calero et al., 2005)

As conceived by Vigo and Brajnik (2011), traditional website quality attributes are: functionality, usability, reliability, efficiency, maintainability, and content. Whereas the quality model of another study (Lee, & Kozar, 2012) contains usability factors such as consistency, navigability, learnability, simplicity, interactivity, credibility, readability, and content relevance. Reviewing 83 publications from 1995 till 2006, Chiou et al. (2010) found ease of use and information quality as the most important factors in IS studies. Malak

et al. (2010) noticed that usability is the most studied characteristic in existing web quality modeling. From the viewpoint of Treiblmaier and Pinterits (2010), the two main questions: "what is presented?" (content) and 'how is it presented?" (design) establish the basic criteria for describing websites. After that, "ease of use" contains navigation/organization and usability. The second dimension, "usefulness", includes information or site content quality, while the third "enjoyment", is measured with constructs such as fun, entertainment, and delight (Figure 2.7).

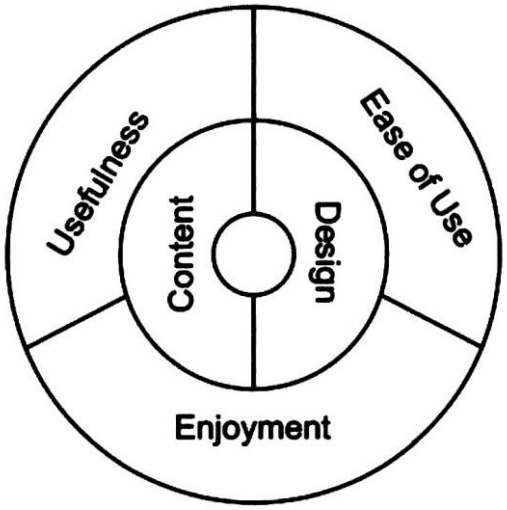

Figure 2.7: Framework for Web Metrics (Treiblmaier, & Pinterits, 2010)

Using web diagnostic tools, Dominic and Jati (2010) evaluated the quality of Malaysian University websites based on 11 quality criteria. The automatic tools are: Website optimization (online performance and speed analyzer), Check-link validator (broken links monitor), HTML validator, link popularity tool and accessibility testing software. The result confirmed that most of Malaysian University websites are neglecting the performance and quality criteria in the context of website design.

Based on the type of website and services offered, Hasan and Abuelrub (2011) summarize common quality dimensions. For example in educational websites, important attributes include currency, accuracy and comprehensibility of information, ease of use, clear layout of websites, and an attractive design. For e-government, the web quality dimensions are: quick response time, up-to-date, accurate information, effective search,

easy to understand, and secure transactions. Additionally, the authors propose an evaluation framework for any type of website with four dimensions: content, design, organization, and user friendly qualities (Figure 2.8). However, through identifying web success factors from a wide range of literature review, their proposed framework is well based on theoretical foundation but lacks empirical validation since it was not tested on any website.

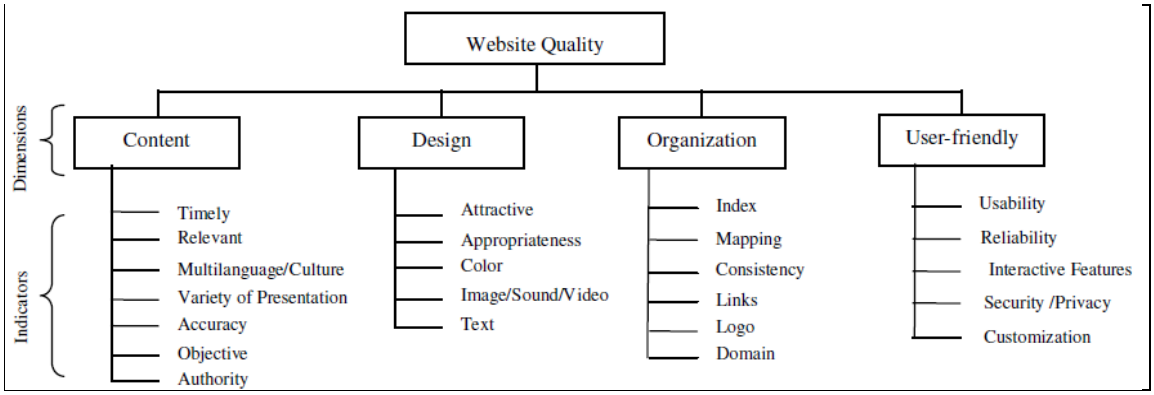

Figure 2.8: Metric Hierarchy of the proposed Framework (Hasan, & Abuelrub, 2011)

Further, Fernandez et al. (2011) performed a systematic mapping study and summarized the knowledge on web usability evaluation methods over the last 14 years, from 1996 to 2010. Figure 2.9 shows a comparison of different research fields. Since web usability evaluation method is considered a sub-topic of usability evaluation and Web Engineering that confirm the growing interest in the topic.

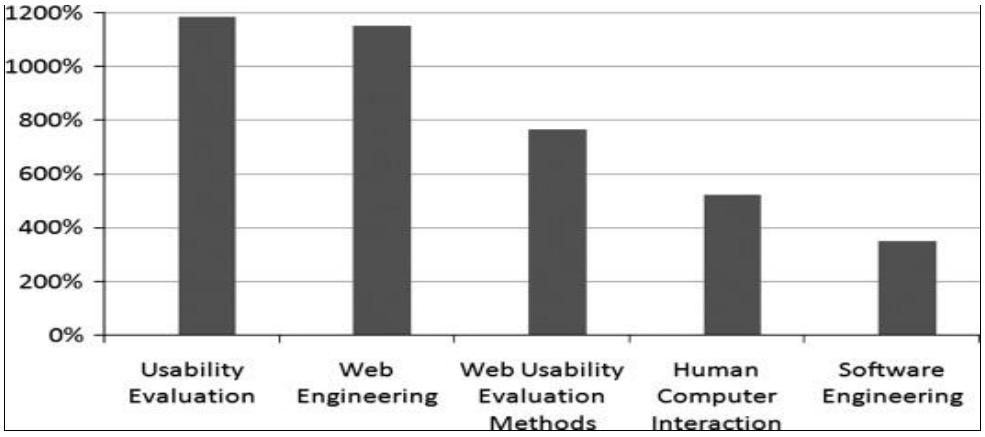

Figure 2.9: Relative Increase Means in Related Research Fields (Fernandez et al., 2011)

To conclude, Web Engineering is the implementation of engineering principles to achieve high-quality websites. There are many scopes of measuring the web since different

perceptions of quality lead to diverse criteria. But surely the lion's share of research in web development is focused around website usability. Thus, usability metrics are the most important measures for evaluating websites. Almost all studies agreed on considering usability as a critical web quality factor and have reported the benefits of a strong commitment to usability in the software development.

2.5 Web Evaluation Methods

The development of a web system is not a one-off event; it's rather a continuous process with an iterative life cycle. One of the popular models for developing a system is the ADDIE model, which stands for Analyze, Design, Develop, Implement and Evaluate (Figure 2.10) (Chow, 2013). What concerns us here is the continuous evaluation cycle. Web evaluation plays an important role in web development and should be performed repeatedly to existing websites (Kaur, & Dani, 2011). Unfortunately, evaluation of websites too often is neglected by many organizations, public or commercial. Many developers test the systems only after it fails or after serious limitations have occurred. Although testing a complex web system is difficult and may be expensive but it shouldn't be delayed until the end of the development process or performed only when problems occur.

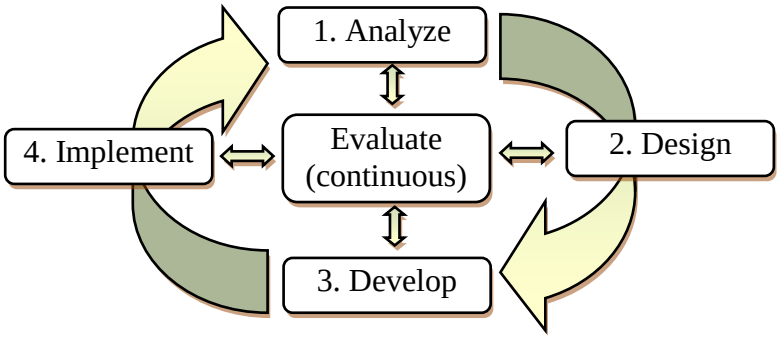

Figure 2.10: The Web Development Cycle (ADDIE Model)

In the process of analyzing websites, one can distinguish between 3 basic measurements (Stolz et al., 2005): web structure measurement; web content measurement; and web usage measurement. Structure measurement is applied to improve the website overall organization, links, and navigability. Content refers to the information, topics and services offered on the website. Since the content and structure measures alone cannot determine website success, users' satisfaction is an important dimension that deserves attention. Another view by Hasan et al. (2012) classified the evaluation methods into 3 main

categories: user-based, evaluator-based and tool-based usability evaluation methods. Tool methods are further subdivided into automatic software tools and web analytics tools. Extending Stolz et al. and Hasan et al.'s works, this research classifies web evaluation methods as follows:

1. User-based usability evaluation methods
2. Evaluator-based usability evaluation methods
3. Tool-based evaluation methods:
    a. Automatic Website evaluation tools b. Web analytics tools
4. Link analysis methods:
    a. Page Rank b. Webometrics Methods

2.5.1 User-Based Usability Evaluation Methods

Assessing the usability of an interface and recommending ways to improve it is the purview of the usability engineer. In fact, usability testing is part of the UCD approach to evaluate usability. The term "usability evaluation" is used to describe the entire test, including planning and conducting the evaluation besides presenting the results (Chow, 2013). The goal of usability evaluation is to measure the usability of the system and identify usability problems which can lead to user's confusion, error, or outright failure to complete some task. Usability problem severity is defined by Nielsen in 1993 (Joe et al., 2015):

1. Low severity problem (minor): is when the problem has little impact on few users.
2. Medium severity problem: is either a problem that has a large impact on few users or a problem that has little impact on many users
3. High severity problem (critical): is a problem that has a large impact on many users.

The user evaluation approach includes a set of methods that employs representative users to execute typical tasks on a selected system. The main aim is to record users' performance and satisfaction with the interface being tested. The most common and useful method in this category is user testing. Suggested technique during a user testing session is such as think-aloud method, field observation, questionnaires and interviews (Bahareh, 2015):

- **User Testing**

This evaluation method has different names in the literature: usability test, usability study, and user testing. User testing means observing a sample of users performing some tasks while interacting with a system, a product, or a website, in order to identify usability problems (Paz, & Pow-Sang, 2014). The purpose of a usability study is to test the system and not the users and that must be explicitly explained to tested users. As they make errors and experience some difficulties with the interface, many users feel under a lot of pressure when conducting a test; hence tests should be done with a deep respect for the users' feelings (Buie, & Murray, 2012).

When users get through a system, they work toward accomplishing specific goals in their minds. A goal is an abstract end result indicating what is to be achieved and can be reached in numerous ways. Consequently, each goal breaks down into tasks specifying what a person has to do, and then each task decomposes into individual steps that need to be undertaken. In fact, user testing must be a sampling process (Rinder, 2012). The tasks have to be real and must represent activities that people would perform daily on the application. Users should be able to do basic tasks correctly and quickly (De Róiste (2013).

To select tested tasks, the examiner begins by exploring all the tasks within the website then narrowing them down to the most important tasks to users. A good task is one that discovers a usability problem, or that is difficult to recover from. Once the list of tasks has been selected, then the next step is how to present them to the participants. One way is to use a "scenario" in which the task is embedded in a very short realistic story. A good scenario is short, in the users' words, and directly linked to tasks and concerns. Even though a scenario gives enough information for doing the task, but it does not give the steps since the point of the test is to see if a user can figure out alone the required steps. It's important to test participants individually and let them solve problems on their own. The following metrics can be collected from user testing: time for users to learn a specific function, speed of task performance, type and rate of users' errors, user retention of commands over time and user satisfaction (Rinder, 2012).

- **Think Aloud Method**

Paz and Pow-Sang (2014) and Joe et al. (2015) regard thinking aloud as an important and valuable usability engineering method for the evaluation of a user interface. Basically, it involves end users using the system while thinking out loud. By verbalizing their thoughts, the participants enable us to understand how they view or interpret the system, and what parts of the dialogue cause problems. De Róiste (2013) stated that its strength lies in the wealth of qualitative data that can be obtained from a small number of users, usually 6 to 8. Also, the users' comments can be included in the test report to make it more informative. Joe et al. (2015) make it clear that this approach is very thorough but it is time consuming and requires face-to-face and one-to-one interaction with users. In addition, analysis of data and generating final results can take a long time. However, it is believed that a well-conducted user testing can overcome some of these difficulties.

2.5.2 Evaluator-Based Usability Evaluation Methods

Evaluators inspect the interface and assess system usability. They use interface guidelines, design standards, users' tasks or their own knowledge, depending on the method, to find possible users' problems. The inspectors can be usability experts, or even novices if experts are hard to find (Huang, & Benyoucef, 2014). In this category, there are many methods such as: heuristics evaluation, cognitive walkthrough, guidelines reviews and standards inspection (Fernandez et al., 2011).

- **Heuristic evaluation**

Heuristic evaluation is the most important inspection method in which a number of evaluators assess the application and judge whether it conforms to a list of usability principles, namely 'heuristics' (Chow, 2013). There are two sets of guidelines widely used in heuristic evaluation, Nielsen's heuristic being the most common followed by Gerhardt-Powals (Lárusdóttir, 2009). Nielsen's heuristic is part of his "discount usability methods" which is easy, fast and inexpensive. During the heuristic evaluation, each evaluator goes through the system interface at least twice. The output of such evaluation is a list of usability problems with reference to the violated heuristics (Buie, & Murray, 2012; Hasan et al., 2012).

Heuristic evaluation is a very efficient usability engineering method and it is especially valuable when time and resources are scarce because experts can produce high quality results in a limited time Huang and Benyoucef (2014). A major drawback of heuristic evaluation is its high dependence on skills and the experiences of the evaluators. Another weakness of inspection methods is the great subjectivity of the evaluation (Hasan et al., 2012). Because the experts are guessing users' problems, there is a risk that they mistakenly consider some issues as problems but actually real users do not have trouble in them. Thus, these issues are often called "false problems" (Hasan et al., 2012). In principle, heuristic evaluation can be conducted by only one evaluator. Nielsen (1993) assured that single evaluator can find 35% of total usability problems, and different evaluators tend to find different problems. Huang and Benyoucef (2014) confirmed that heuristic evaluation can be conducted by a single inspector but increasing the number of evaluators is better. Bahareh (2015) believed that two evaluators are enough, while Paz and Pow-Sang (2014) estimated 3 or 5 experts for heuristic evaluation.

2.5.3 Tool-Based Evaluation Methods

Unfortunately, the complexities of the websites and technology make testing with users difficult due to time and cost constraints. Automation of testing websites is a new emerging method (Dominic & Jati, 2010).

2.5.3.1 Automatic Website Evaluation Tools

An automatic evaluation tool is a software that automates the collection of interface usage data and identifies potential web problems. There are several flavors of web testing tools: accessibility tools such as Bobby, usability testing tools such as LIFT, W3C HTML validator, and link-checker (Dominic et al., 2010). Many reasons are behind moving recently toward automated web evaluation tools (Dingli, & Mifsud, 2011): tools are fast, do not get tired, low cost, produce unbiased results, and cover the shortage of experts and inconsistent results between different experts.

According to Kulkarni and Dixit (2012), Ivory and Chevalier 2001 conducted the first study of automatic website evaluation tools. They found automatic tools identify more problems than other evaluation methods but more research is needed to validate the

guidelines embedded in the tools and to make the tools usable. A similar conclusion was reached by Dingli and Mifsud (2011), tools play an important role in the evaluation but a web professional cannot rely on them alone to improve websites.

Rodríguez et al. (2009) present a framework for assessing e-Governance maturity through analyzing municipal websites. The framework considers websites content and design and includes 152 metrics grouped into 8 features: information, functionality, truthfulness, participation, friendship, usability, accessibility and navigability. A list of penalty metrics is imposed: no sitemap, no internal search, or disabled browser back button (decrement 5 points), and missed organization name or logo (decrement 3 points). Some metrics are measured by website inspection, others by automatic tools: W3C validators, Xenu software for broken links, and a web page analyzer of source code. A survey of 16 countries was conducted to show the applicability of the framework. From each country, three municipal websites were inspected. Table 2.2 shows the percentages of fulfillment for each surveyed country. The results show that municipal websites better fulfill design metrics than content metrics with the exception of Mexico. Only 6 countries reach at least 50% of the maximum score defined for content metrics. Considerably this framework has too many metrics (152) to be tested on 48 municipal websites. Besides that, 26% of theses metrics were selected by the researchers without any justification. All that weakens theoretically the proposed framework. Another limitation of this study, there is little information on how the manual website inspection method was conducted. Also, the automatic website evaluation method, such as Xenu, was not validated in the literature yet (see section 5.4) but the authors claimed it is 100% reliable.

Table 2.2: Percentages Reached by Each Country (Rodríguez et al., 2009)

Country	e-Governance	Design	Content	Country	e-Governance	Design	Content
Spain	69.15	73.37	51.94	Chile	58.16	59.75	52.91
Australia	67.62	71.21	54.37	Nigeria	56.47	60.78	37.86
Colombia	65.17	70.90	41.26	Mexico	56.21	56.55	56.80
United States	64.92	66.56	60.19	Costa Rica	53.59	57.79	35.92
Luxemburg	63.82	69.25	41.75	Ecuador	53.51	56.76	40.29
Francia	63.74	65.94	56.31	Bolivia	49.28	52.53	35.92
Argentina	61.96	65.12	49.51	Perú	46.15	48.19	39.32
Venezuela	58.50	63.78	36.41	Puerto Rico	43.45	47.27	26.21

Comparing available studies which use automatic tools as a website evaluation method either solely or in combination with manual web assessment methods, this study found that:

- Only a limited number of studies employ automatic tools in web analysis
- Automatic tools are seldom used alone in web evaluation
- When automatic tools are used, usually it is combined with the manual assessment. Evaluations by experts or users are the mainstream approach
- Very few studies compare automatic evaluation tools and validate their effectiveness
- The most used tools are: Bobby, LIFT, W3C validators, and link checker software
- Automatic tools are not considered efficient and most of them focus on site accessibility rather than usability (Hasan et al., 2012; Scowen, & Regenbrecht, 2009)
- Information about the LIFT tool is controversial and contradictory (Zahran et al., 2014):
    - Some perceive LIFT for accessibility and few believes that it is a usability tool
    - Features measured by LIFT are inconsistent with the Research-Based Web Design and Usability Guidelines.

2.5.3.2 Web Analytics Tools

A relatively different method in web evaluation is the use of web analytics tools. Formerly web analytics has been defined by the Web Analytics Association as "the measurement, collection, analysis and reporting of Internet data for the purpose of understanding and optimizing web usage" (Sleeper, Consolvo, & Staddon, 2014). These tools automatically calculate statistics regarding the detailed use of a site and collect data about users' behavior. They can help, for example, in discovering navigation patterns corresponding to high web usage or to the early leaving of the website. Originally, web analytics as a business tool arose from a commercial need to aid in understanding user experience. It started with some webmasters inserting counters on their home pages to monitor web traffic. It helps online businesses improve their websites so they maximize profit (Sleeper et al., 2014). While most web analytics studies target e-commerce, the method can be applied to any website. The two data collection methods for web analytics are: 1) Server -based log files: traffic data is collected by web servers and held in log files such as access logs, agent logs, and error logs; 2) Client-based page-tagging: it requires

adding a few lines of JavaScript code to webpages to capture information about the visitors' session (for details see Clifton, 2012).

2.5.3.2.1 Google Analytics

In 2005, Google purchased a web analytics company called Urchin software. At that time, many popular websites used software solutions from Urchin to understand users' experience. Subsequently, Google released Google Analytics (GA) to the public in 2006 as a new analytics tool. The service is provided without charge for up to 5 million page views per month per account. Due to its popularity, Google placed new applicants on a waiting list until Google Analytics became generally available to the public (Zahran et al., 2014). If a Google account holder signs up for Google Analytics, Google offers code that must be inserted into each web page to be tracked. Data results are displayed in visually enhanced reports with a wealth of information on where visitors came from, what pages they visited, how long they stayed on each page, how deep into the site they navigated, where their visits ended and where they went from there (Clifton, 2012).

2.5.3.2.2 Alexa

Alexa is a leader in web traffic ranking. Alexa is a website metrics system owned by the Amazon Company which provides a downloadable toolbar for Internet Explorer users. The Alexa rating for websites is obtainable on Alexa's website http://www.alexa.com. It calculates traffic rank by analyzing the web usage of Alexa toolbar users for a three months period or more. It is a combined measure of page views and reach which is the number of users to the site. The toolbar offers search engine functionality and traffic information about the browsing sites. Users who use the Alexa toolbar contribute to the ranking of website popularity. The Alexa ranking system is a reverse order rank. Lower Alexa numerical rank is better for a website (Alexa, 2014). Although this information is of promise, Alexa ranking has some limitations. In fact, Alexa is biased towards a sample of Microsoft Windows and Internet Explorer users. Since users of other operating systems or browsers are not recorded and traffic from other Internet users is not counted, the resulting statistics are unreliable (Kaur, & Dani, 2013).

Rare studies show the relevance of Google Analytics, Alexa toolbar, or even in general automatic evaluation tools, in assessing websites (Al-Juboori et al., 2011). The authors, therefore, surveyed a combination of automatic website evaluation tools, such as WebXM, Bobby, NIST web metrics, Alexa, Link Alarm, and Web Tango. They concluded that each tool concentrates on a specific area, therefore an overall evaluation was not implemented during their research. Hasan (2009) developed a framework for evaluating three e-commerce websites in Jordan using heuristic evaluation, user testing, and Google Analytics. Jordanian companies took a long time to agree to participate in the research due to trust and security issues since to use Google Analytics, each company was asked to add script code to their server. Also, two of the selected companies did not agree to add code related to e-commerce transactions. The study pointed out that Google Analytics is useful as a preliminary step to provide a quick, easy and cheap sign of usability problems.

2.5.4 Link Analysis

Link analysis offers valuable information to study link patterns and websites' topology. According to Thelwall (2012), it is based on the notion that the quality of a web page is dependent on its links. A link from a page p to page q can be viewed as an endorsement of q by p, and as some form of positive judgment by p of q's content. There are two important methods that use link analysis: the Google PageRank and Web Impact Factor (WIF).

2.5.4.1 PageRank

A number of researchers investigated the web link structure to improve search results and proposed ranking metrics. When Page and Brin designed the Google search engine, they considered links as positive referrals and created a system called PageRank (Scowen, & Regenbrecht, 2009). Thus, PageRank, a link analysis algorithm named after Larry Page, assigns a numerical weight to each hyperlink. Each page has a calculated PageRank based on the number of links pointing to it and the quality of those links (Zahran et al., 2014). Through PageRank, the position of a page in Google search results is decided. Google takes 100 factors into account in determining the ranking of a page but PageRank is the main factor in search results' ordering (Scowen, & Regenbrecht, 2009). The PageRank scale is calculated between 0 and 10 through a complex algorithm based on the quality of external links. The PageRank metric PR(p) defines the importance of page p to be the sum of the

importance of the pages that point to p (Kaur, & Dani, 2013). If many important pages point to p, PR(p) is high.

The effectiveness of Google's search results and the adoption of PageRank by other search engines strongly indicate that PageRank is an effective ranking metric for web searches and seems to capture the importance or the quality of web pages well. In a recent survey, the majority of users are satisfied with the top-ranked results from Google and other search engines. But unfortunately, Google's PageRank is heavily biased against unpopular pages, especially those that were created recently (Zahran et al., 2014).

Essentially, PageRank is a "link-popularity" metric (Dominic, Jati, & Kannabiran, 2010). The study of Scowen and Regenbrecht (2009) proposed a correlation between usability and popularity and differentiated between popularity and success. A popular website is one that is liked by many people but a successful website is defined differently depending on the needs of the business. In other words, the study implies that success is goal-specific. For an online retailer, success may mean the percentage of visitors converted to buyers. The authors tested e-learning websites against checklist guidelines then against 5 ranking systems: Google links search, Yahoo links, Delicious links, Google PageRank and Alexa. The Google PageRank and Alexa were used to identify their correlations with usability, although neither can be relied upon as a main indicator of popularity. The study found that increased compliance with usability guidelines has a strong correlation with the increased popularity of a website. Although Alexa is not a reliable indicator, it is at least consistent with other rankings. More usable websites achieve a higher PageRank and also are more popular in Alexa. Overall, the five ranking systems showed positive correlations to each other and to the usability of the sites.

2.5.4.2 Webometrics (the WIF Method)

Studies of the web were named “Webometrics” by Almind and Ingwersen in 1997 (Thelwall, 2012). The Webometrics is “the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web, drawing on bibliometric and informetric approaches”. That means evaluation of websites can be conducted "Webometrically" (Thelwall, 2012). The main goal is to validate links as an important source

of information and to furnish its acceptance as a useful metric to measure the quality of website. It evaluates the international visibility and impact of an institution or a country on the web. Webometrics is still a young field of research that needs different theories, methods to be developed and problems to be solved (Holmberg, 2010).

The Web Impact Factor (WIF) is the most important method in Webometrics. In 1998, Peter Ingwersen proposed "WIF" by analogy with Journal Impact Factor (JIF). JIF represents the ratio of all citations to a journal to the total number of references published over a period of time. The number of citations to a journal is limited in depicting its standing (Thelwall, 2012). Outlink, link to a web page in another website, is similar to "reference" and inlink, link that a web page receives from other web pages, is similar to "citation". However, since the WIF is a snapshot of the web at a certain time and the contents of websites lack peer review and quality control, the WIF is not exactly the equivalent of the JIF but was inspired by it. From the webometrics' view, external inlinks, links received from an outsider website, are of more value and importance. The more external inlinks, the more valuable the website is. Thus, the more people link to a website the more WIF the site is getting. In turn, the higher the impact factor the higher the reputation and influence of a site (Shekofteh et al., 2010).

Sometimes, the WIF is compared to Google's PageRank. However, PageRank does not give equal weights to links; rather it varies depending on where a link is coming from. Also, a web page receiving one link from a highly linked page which has many external inlinks is ranked higher than another webpage receiving hundreds of links from lesser linked pages (Zahran et al., 2014).

Most of Webometrics studies were performed on university sites. The Cybermetrics Lab (2014) in Spain has issued the Webometrics ranking since 2004. The “Ranking Web of World Universities” is published twice a year covering 20,000 academic institutions. But very few Webometrics studies were conducted on e-government, representing a new trend. The study by Petricek et al. (2006) represents the first attempt to measure e-government Webometrically. They compared the structure of government audit office sites in Canada,

USA, UK, New Zealand and Czech Republic. The US and Canada emerge as the most internally connected and navigable sites much better than the UK.

In terms of local e-government, Holmberg (2010) examined the website of the region of Finland Proper (Varsinais Suomi) in the southwest of Finland. The aim of the study is to shed some light on the knowledge of governmental interlinking and on the real world phenomena it may indicate. The authors argue that web links information reflects the real world and relationships between organizations. Two questions are raised: does local government website interlinking in Finland Proper (Varsinais Suomi) follow geographic lines? And what are the reasons behind interlinking in local government websites there? The results indicate that interlinking between local governments in Finland follows a strong geographic, or precisely a geopolitical pattern. Primarily, links are created to reflect official cooperation and that geographic closeness is a factor in the majority of cases.

2.5.5 Comparison of Web Evaluation Methods

According to Hasan et al. (2012), evaluation by experts or users is the mainstream approach. In fact, user testing finds more major problems than other evaluation methods but is poor in uncovering minor problems. The situation is the opposite for the heuristic evaluation. Thus, they recommend conducting heuristic evaluation and user testing to take advantage of both methods. It is best to perform first heuristic evaluation to find as many "obvious" usability problems without users who are difficult to recruit in large numbers. Then, perform user testing to find remaining usability problems encountered by target users. From the point view of Joe et al. (2015), heuristic evaluation complement user testing but is not a replacement. On the contrary, Huang and Benyoucef (2014) believed that heuristic evaluation is better and detects more web design problems than user testing. Krenk and McComb (2012) arrived to another conclusion, the purpose of the website is the factor into which method is most appropriate. They compared the two methods in detail (Table 2.3) and suggested that user testing is better for evaluating dynamic websites while heuristic evaluation suited static websites.

Table 2.3: Comparing Heuristic and User Testing (Krenk, & McComb, 2012)

	Heuristic Evaluation	User Testing
Advantages	• Inexpensive and fast • A beginner can do an inspection • Identify more usability problems when compared with user testing	• Identify specific & practical problems • High confidence in the results • Users find surprising problems that would not be identified by other means • Possible to conduct even if the users do not know anything about user interface
Limitations	• Does not involve users in the evaluation process. Thus, it cannot be certain that identified problems reflect an actual user problem • The reports do not predict end-user problems as one might wish • Inspections are not as effective in determining the overall satisfaction of customers as user testing • Often need multiple evaluators to find large proportions of problems	• Expensive • Difficult and time-consuming to recruit participants and to analyze data • Should only use each user once • Location of testing may be restricted

Al-Juboori et al. (2011) state that the automation of website testing is an evolving method that cannot be considered efficient. Dingli and Mifsud (2011) consider the adoption of automatic tools is limited and demand the effectiveness of them has to be evaluated since it is difficult to have a tool that can behave like a human and exhibit common sense. They concluded that such automatic tools cannot replace an expert but it is a useful complement to standard evaluation techniques, such as user testing or heuristic evaluation. Similarly, Olsen et al. (2009) believe that automatic tools can offer a reliable first insight into the status of a website. Another limitation is the difficulty Dingli and Mifsud (2011) encountered with incorporating certain usability guidelines into the tools because of their abstract nature, therefore additional research needs to be carried out. The analyses of Cassino et al. (2015) show an improvement in cost and time when using these tools, but it doesn't measure the subjective attributes of usability such as the sense of satisfaction.

Another concern is that the market forces can cause changes that threaten automatic tools' stability and may impede its application (Zahran, et al., 2014). For example Bobby, an accessibility testing tool, was sold to Watchfire in 2004, which provided the same free service in the WebXACT tool, but Watchfire was acquired by IBM in 2007. Bobby was then discontinued as a free tool, and currently, it is included within the IBM Rational Policy

Tester Accessibility Edition (Hasan, 2009). The situation for automatic usability evaluation is even worse since very few tools exist.

On the other hand, web analytics technique solves some problems in usability evaluation since it might reduce the need for user testing. Most of the time, the data is collected automatically with high accuracy. Analytics tools offer the possibility of analyzing web traffic data for a high number of Internet users, thus increasing the reliability of the discovered errors (Hasan, 2009). However, the inaccuracy of log files as a data source is acknowledged. Another problem is the meaning of collected information and how much it describes users' behavior Rodriguez (2013). The author believes that website traffic measures are used because they are easy to capture but very often deemed to be inadequate and sometimes may generate conflicting results. Other researchers (Hasan, 2009) suggest using this method as a supplementary technique to user testing method or alone to collect the usage data of a system. Unfortunately, little research has employed web analytic tools and compared them with standard usability evaluation methods (Vaughan, & Yang, 2013).

Jalal, Biswas, and Mukhopadhyay (2010) describe the Webometric method as an imperfect tool to measure the quality of websites. According to (Zahran et al., 2014; Thelwall, 2011), question marks are raised over the entire quantitative nature of the Webometrics rankings. Search engines used in the WIF is not meant for link analysis since they are designed for content retrieval, plus they may create problems in drawing conclusions for WIF since their coverage of the web is incomplete. The lack of knowing why web links are created is a major obstacle in the Webometrics method. Thus, the motivations behind creating external links raise questions of uncertainty. Thelwall (2011) mentioned the problem of commercial search engines withdrawing their link search queries (only Yahoo remains), this is a serious threat to Webometrics and would undermine the power of link analysis.

A large number of Webometrics' studies found unexpected results and attributed that to the limitation of the WIF method. For example, a university with 99 web pages and 993 links gets an impact factor of 10 whereas another university with 87700 web pages and 12700 links obtains an impact factor below zero (Shekofteh et al., 2010). Calculating the WIF for

a website is easy but what the figures mean is arguable and debatable. Thus, Webometrics techniques are still in their experimental stage in testing (Sultana, 2015; Thanuskodi, 2011). That implies Webometrics is in the process of developing and validating its methodologies (Sultana, 2015; Thelwall, 2012).

University rankings have raised a large dispute and several studies criticize them as merely a list of criteria that mirrors the superficial characteristics of universities. For example, Jalal et al. (2010) argued that the WIF in most cases reflects unreliable results while Shekofteh et al. (2010) concluded that the WIF alone is not a good measure for ranking universities. Sultana (2015) argued that World university website ranking is not meaningful because a high link rate may not always be associated with high quality. He claimed that it is vulnerable to manipulation since the WIF outputs can be influenced by institutions who know how the Webometrics method works.

It is important to remember that Webometrics is relatively a new research field that needs further development and its results can be regarded as indicators rather than definite conclusions on the visibility and impact of a website (Thelwall, 2012). From the viewpoint of this researcher, the WIF is partially successful since it does provide some interesting and useful information such as the relationship and type of communication between universities or countries and also how a website is isolated or connected with others online. On the other hand, the method fails in the evaluation and ranking of government or universities websites since it is not a suitable tool for assessing websites' quality or content.

To conclude, there is a lack of studies that classify, compare, and determine the appropriate web evaluation methods. Krenk and McComb (2012) assured no consensus exists regarding which method is better in identifying usability problems. But web experts often recommend using several different evaluation methods since each one alone isn't free of shortcomings (Ølnes, 2013). Fernandez et al. (2011) agree and add there is no single method suitable for all circumstances. In general, the recommendation by many researchers is to conduct heuristic evaluation and user testing, while other web evaluation methods are just useful complements offering the possibility of analyzing a high number of users as an initial preview of a website.

All these different web evaluation methods are just tools to assess any kind of websites (commercial, e-government, etc.). For example, heuristic or user testing evaluate e-government websites based on certain e-government models looking for the fulfilment of specific attributes or guidelines, which is the subject of the following section.

2.6 E-Government

In the late 1990s, a new face of government known as e-government, or digital government, was introduced, following the success of the private sectors adoption of e-business and e-commerce (Coursey, & Norris, 2008). The 2013 report of the Oxford Internet Surveys (OxIS) stated that, as in many other nations, the UK take up of e-government has been slow, although incremental advances have been made over years. One reason for this, it is a difficult arena for services that are not accessed often, sometimes once a year or less. It involves tens of thousands of individuals interacting with thousands of services at all levels of government. In contrast, banking services involve millions accessing a few services, such as looking at their account balance. In 2013, 65% of Internet users in the UK said they used at least one service in the past year (Blank, 2013). All governments now are investing heavily in developing their websites and they should learn from the commercial experience and look into citizens' expectations to use e-services in earlier stages than the commerce did. By 2010, the majority of countries had embraced e-government with different level of success. Of the 192 UN Member States, 189 countries were online (United Nations, 2010). That means, 98% of countries around the world have government websites available on the Internet.

E-governments reduce travel and waiting time (moving processes from in-line to on-line), eliminate corruption, reform government, increase transparency, enhance the relationship between government and citizens, and ultimately develop democracy. E-services are cheaper, faster, and readily available 24/7 (Zhao et al., 2012; Didraga, & Brandas, 2015). Practical examples of e-government's financial benefits include the Information Network of Kansas generating a revenue of 7 million USD per year and Singapore e-Tax saving SGD 20 million per year (Mohammad, Almarabeh, & Ali, 2009).

An analysis of e-government up to the year 2004, conducted by Kunstelj and Vintar (2004), may be still valid despite the years passed (Bannister, 2007; Montserrat, 2010; Zahran et al., 2015). That study categorized e-government evaluation approaches by the dimensions (indicators) they cover:1) e-readiness includes readiness of government, citizens and businesses to e-participate, such as the ICT infrastructure; 2) back-office includes the re-engineering and digitalizing processes within the administration; 3) front-office consists of official website, e-services, and information (a. supply-side: government; b. demand-side: citizens and businesses); and 4) their impacts (improved performance indicator as a result of e-government effort, i.e., cost, trust, transparency, corruption, etc.). Luna et al. (2013) describe further the impacts as outcomes dimension in terms of efficiency, cost reduction, transparency, public participation, service quality, and customer satisfaction. Another view by Schellong (2010) perceives e-government of varying scopes: goal, input (funds, labor, or infrastructure), process, output (information, and e-services), outcome, efficiency, effectiveness, demand, and usage (adoption).

The majority of e-government studies focused on the front-office supply side of government, and less on the demand side of the citizen, while largely neglecting the back-office and the impact of e-government (Rorissa et al., 2011). The authors attributed that trend to the expensive data collection and complex processing of the back-office approach. Also, majority of studies examined central e-government while very little attention is given to the local level (Schellong, 2010; Shareef et al., 2012).

There is still no standard accepted definition of e-government but a variety of description is found in the literature. The reason of so many definitions is the multi-dimension nature of e-government. Al-Saif (2010) presents a table of seven e-government definitions from different perspectives. Bannister (2007) and Rorissa et al. (2011) believe that the definition of e-government varies according to its dimension from the very generic – the use of ICT in the formulation and execution of government and public policy; the use of ICTs in public administration to achieve innovative forms of government and governance - to the more specific - any use of ICT in public administration and services; the delivery of government services over the Internet in general and the web in particular. Also, Tsohou et al. (2013) emphasized that there was a need for impact oriented techniques

for the evaluation of e-government including citizen-centric approach. This dissertation adopts the last e-government definition in a broader approach to investigate the actual use of Saudi city websites, with an aim to contribute to impact evaluation of e-government. Particularly, it will evaluate the front office demand-side citizen-centric municipal websites and monitor some impacts of e-government, such as the e-service quality and customer satisfaction.

2.6.1 E-Government Evaluation and Frameworks at a National level

In general, an evaluation aims to determine the value and benefit derived from e-government investments and discover the current state of e-government development. It should investigate various perspectives, require the inclusion of the needs of target groups of citizens using specific e-services, and consider the social and technical context of use (Montserrat, 2010).

E-government evaluation can be classified in three ways according to their unit of analysis: websites and online services, cost-benefits, and e-government stage models (Tsohou et al., 2013). The same taxonomy with slightly different names respectively is provided by Siskos et al. (2014): websites evaluation, plans/strategy, and benchmarking e-government; and by Sandoval-Almazan and Gil-Garcia (2008a) as: citizen-centered website, managerial perspective, and e-government stage models. The first category refers to the evaluation of e-government websites according to web metrics of specific criteria (such as content, usability, functionality, security and online services) using different web evaluation methods such as user testing (Tsohou et al., 2013). In the second category, the cost-benefits, or simply benefits, are named in several other studies as: impacts, effects, outcomes, and manager objectives of e-government. Examples of these benefits are: greater information access, e-service quality, convenience, effectiveness, efficiency, transparency, democracy, cost reduction, time savings, less corruption, citizen satisfaction, etc. (Janssen et al., 2004; Kunstelj, & Vintar, 2004; Tsohou et al., 2013; Didraga, & Brandas, 2015).

The third category is about benchmarking which has long been used in business for marketing and sales purposes. In fact, countries are benchmarked on many facets, ICT, economy, education, press freedom, happiness, corruption, etc. (Rorissa et al., 2011). Often

e-government is evaluated through a benchmark which is a technique for comparing e-government performance based on indicators that yield an e-score. A model is used to derive suitable indicators for evaluating various e-government initiatives (Berntzen, & Olsen, 2009). Comparing indicators requires special care that apples are compared to apples, not oranges (Montserrat, 2010). The importance of such models lies in its offering a basis to measure and guide e-government development by drawing attention to best practices. Also benchmarking has been used by countries to monitor the efficiency and effectiveness of public spending on e-government (Rorissa et al., 2011).

Further, some of the e-government stage models are used for benchmarking purposes (Grönlund, 2011). The origin of these evolutionary models is rooted back in psychology, organization theory and IS (Information System) field (Debri & Bannister, 2015). They have been applied in various domains, including e-government by a number of authors (Layne, & Lee, 2001; Moon, 2002; Andersen, & Henriksen, 2006; Klievink, & Janssen, 2009). An e-government website's maturity model is a set of stages, from basic to advanced ones, which determines the e-government maturity. It focuses on the evolution of e-government using sequential steps, for instance from immature to mature e-government with improved quality (Fath-Allah et al., 2014). The general idea of the stage models remains the same with just different number and terms are given to stage models. Almost all of the stage models contain: web presence, interaction, transaction, integration or transformation (advanced features such as information sharing between agencies), and e-participation in some models where citizens vote online and participate in opinion surveys (Fath-Allah et al., 2014). The first IS model was developed by Nolan in 1973 using the term stages of growth, now the maturity model is widely used (Debri, & Bannister, 2015), also was referred to as stage models (Poeppelbuss et al., 2011). Poeppelbuss et al. (2011) analyze 76 maturity models in IS journals over 15 year period up to 2011, while Debri and Bannister (2015) examine 51 maturity models from 1973 up to 2015. Therefore, there are numerous stage models and benchmarks in e-government (Janssen, 2010) (see for examples Debri, & Bannister, 2015; Valdés et al., 2011; Poeppelbuss et al., 2011). The multidimensional nature of e-government is the reason for the existence of many e-government stage models.

As presented in the next section, the development of e-government models began in 2000 (Montserrat, 2010). Inspired by Schedler and Schmidt (2004), we propose to classify the e-government models into three types: organizations and consultancy firms, scholars, and official government models. Also, there are few attempts to develop e-government service quality scales.

2.6.1.1 Organizations and Consultancy Firms Evaluation Frameworks

Several organizations tried to understand the e-government phenomenon by constituting frameworks which are divided further into different numbers of stages of growth. Yildiz (2007) stated that e-government is analyzed by developing models of its stages but there is no agreement on the number of stages or requirements. The most established e-government evaluation reports published periodically and cited frequently as benchmarking are identified in Table 2.4.

Table 2.4: Popular National E-government Frameworks
(Hu et al., 2005; Moraru, 2010; United Nations, 2010)

Model	Focus	Stages				
Accenture 2000	22 Developed Countries	Publish Passive/Passive Relationship	Interact Active/Passive Interaction	Transact Active/Active Interaction		
Brown University 2001	Worldwide	Billboard "Information"	Services Delivery	Portal	Interactive Democracy	
United Nations (UN) 2002	Worldwide	Emerging	Enhanced	Interactive	Transactional	Connected
Capgemini Europe 2002	European Countries	Information	One-way Interaction	Two-way Interaction	Transaction	

1) United Nations:

The United Nations has been assessing e-government since 2002 (Berntzen & Olsen, 2009). Initially, the UN e-government model was described by Rutgers University as a three-stage model (Montserrat, 2010). Currently, it is widely used in many studies, and it has two indices: the e-government index and the e-participation as a supplementary index (United Nations, 2014). The e-government index ranks e-governments worldwide at the

national and ministry websites. Each of its three measures (online service, telecommunication infrastructure, and human capital) is a composite index that can be analyzed independently, with a value between one and zero (Table 2.5). The online service index was based on a four-stage e-government model: emerging, enhanced, transactional, and connected presence, with features for each stage (see United Nations, 2014). Further, the UN e-government model has been amended by a set of measurements, "e-participation", which is taken to be more or less directly related to democracy. The e-participation index focuses on the use of the Internet to provide information to citizens (e-information), interaction with stakeholders (e-consultation), and engagement in decision-making processes (e-decision-making) (United Nations, 2012; 2014). This index offers tools for citizens' inclusion with government such as online polls, e-voting, forums, blogs, and social networks (Facebook, Twitter, YouTube, etc.).

Table 2.5: The UN E-government Framework (United Nations, 2014)

First Class Measure	Second Class Measure	Third Class Measure
Overall Development	Online service Index	Emerging presence
		Enhanced presence
		Transaction presence
		Connected presence
	Telecommunication Infrastructure Index	Internet users / 100
		Broad banding / 100
		Tele lines / 100
		Mobile phones / 100
		PCs / 100
		Tvs / 1000
	Human capital Index	Adult literacy rate (%)
		Combined gross enrolment ratio for primary, secondary and tertiary schools (%)

2) Brown University:

Professor West and his research team at Brown University conducted an annual evaluation report of government websites since 2001 (Berntzen, & Olsen, 2009) but lately, the Brookings institution publishes the study. The report of (2008) analyzed 1,667 government websites in 198 countries for the presence of 18 features such as: phone and

address contact, publications, audio and video clips, number and type of e-services, privacy and security policies. The ranking runs along a scale from zero to 100 points (West, 2008): 4 points for the presence of 18 features totaling 72 points and one point for one service up to 28 points for 28 or more e-services. West's survey only examined the presence of services without measuring their quality. Moraru (2010) recognized West's stages as: 1) Billboard: online government information; 2) Service delivery; 3) Portal: "one stop shop" for e-government, security, and privacy; 4) Interactive democracy.

3) Accenture:

Accenture is a consulting, technology service, and outsourcing company that has been issuing annual e-government reports on developed countries since 2000 (Hu et al., 2005). The original Accenture model included two dimensions, customer relationship management (30%) and service maturity (70%) with two indices: the number of online services (service breadth) and the level of service completeness (service depth). Service maturity is decomposed into the following stages (Peters et al., 2004): (1) Publish-passive/passive relation: no communication between users and government; (2) Interact-active/passive interaction: only users can e-communicate with government; and (3) Transact-active/active interaction: two-way communication is possible.

Berntzen and Olsen (2009) record some modifications on the Accenture rankings. The 2005 Accenture index had two components, each with a weight of 50%, service maturity and customer service maturity, which were measured by four dimensions: citizen-centered, multi-channel, cross-government service, and proactive communication about the services to users. Four hundred citizens in each evaluated country were questioned about their country’s e-services, and interviews of 46 high-ranking government executives were conducted. The benchmarking performed in 2007 introduced a new indicator, citizen voice (40%), reducing the weight of service maturity to 10%.

4) Capgemini Europe:

Capgemini (2006), a company specializing in consulting, technology, outsourcing, and local professional services, focuses on evaluating the e-presence and sophistication of government websites in 32 European countries. According to Capgemini (2010), the survey

benchmarks 20 basic online services, 12 services to citizens, and 8 services to businesses with the following indicators: online sophistication, full online availability, user experience (usability, transparency, privacy, multi-channel policy, and users' feedback), and portal sophistication (most mature, user-centric, and personalized portals). Basic citizen services include income tax, job search, social security benefits, personal documents (passports, driver's license), car registration, building permission application, declaration to the police, public libraries, birth and marriage certificates, enrollment in education, announcement of moving house and health-related services. The online sophistication and availability rankings assess the 20 public services against four stages in the 2006 report then against a 5-stage maturity model from 2007 until the 2010 edition (Figure 2.11): information, one-way interaction, two-way interaction, transaction, and targetisation automation threshold (proactive, automated service delivery). This measurement evaluated the online presence and sophistication of about 10,000 websites at national, regional and local levels in the 32 European countries.

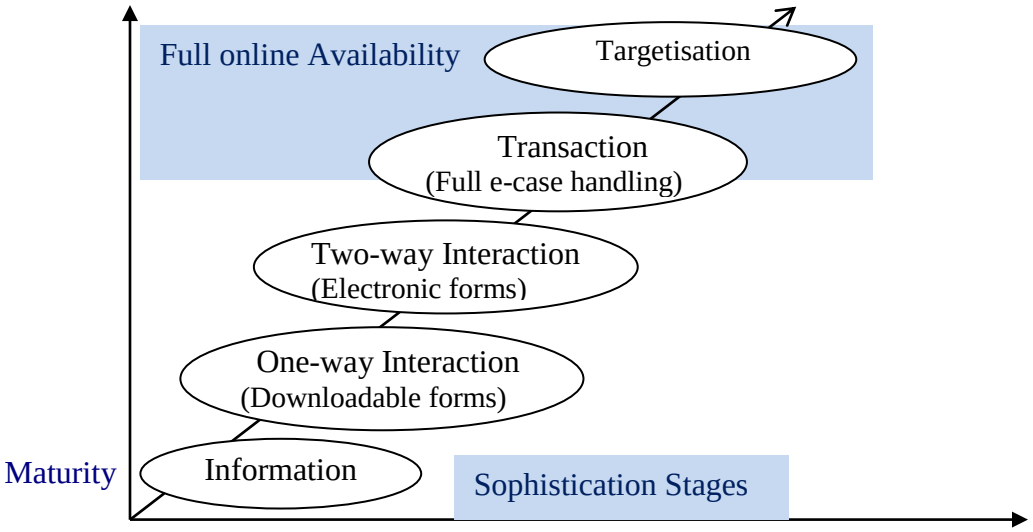

Figure 2.11: Capgemini eEurope Five Stage Maturity Model (Capgemini, 2010)

The Capgemini report (2009) claimed a paradigm shift towards customer-centric services. New patterns of relationships go from "You-Centric" model to "Me-Centric" model changing the role of the user from a passive viewer and user to an active creator of the public service delivery chain (Figure 2.12). This view of e-government is an important step toward reaching a citizen centric e-government model that benefits citizens and governments. New measurements (i.e. user-centric, transparent, citizens and business

mobility, key enablers indicators) were added in 2012 and e-survey of citizen usage and satisfaction was carried out.

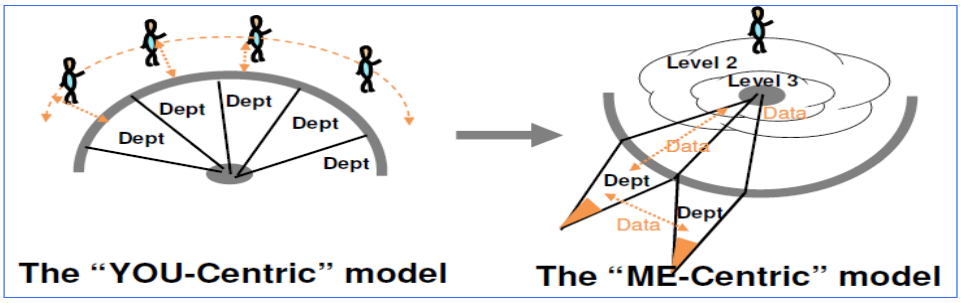

Figure 2.12: A Shift towards Customer-Centric Services (Capgemini, 2009)

2.6.1.2 Scholars Evaluation Frameworks

Sparse contributions to this vital subject are still evolving, as several scholars offer their own insights. The first e-government model was proposed by Baum and Di Maio (Gartner model) in 2000 and has four stages: Web presence, interaction stage, transaction stage, and transformation stage (a citizen-centric and responsive government) (Montserrat, 2010). Another highly cited e-government model was proposed by Layne and Lee (2001) with reference to the USA in four stages: catalogue, transaction, vertical integration (local, state and federal governments connected for similar functionalities or services of government), and horizontal integration (one-stop portal: integration across different services in which a citizen can contact one point of government and complete any level of e-transaction). The model argues that progress on e-government is a matter of technologies and organization.

Moon (2002) extended the Layne and Lee model and proposed an evaluation framework of five stages: 1.Information dissemination /catalogue; 2.Two-way communication; 3.Service and financial transactions; 4.Vertical and horizontal integration; 5.Political participation. The study surveyed 1,471 US e-municipalities with populations over 10,000 and found that larger municipalities are more advanced in e-government but they are still at an early stage of development and have not obtained many of expected outcomes. Also, Andersen and Henriksen (2006) extended the Layne and Lee model by including a customer-centric approach. They proposed a user focus "Public Sector Process Rebuilding model" with four stages: cultivation, extension, maturity, and revolution. There exist

numerous other models but these examples adequately represent the general thought of stages models. For further reading see the study of Fath-Allah et al. (2014) which compared 25 e-government stage models.

Other models such as the model of Klievink and Janssen (2009) introduces the notion of dynamic capability theory to move up from one stage to the next. Sandoval-Almazan and Gil-Garcia (2008a) evaluated Mexican portals with a mixture method of six stage e-government model (presence, information, interaction, transaction, integration, political participation) plus assessing other variables such as: usability, customization, transparency, e-services, privacy, security, broken links, design problems, and search problems. Tsohou et al. (2013) propose a reference process model for citizen-centric evaluation of e-government that identifies key performance indicators directly connect to citizen's satisfaction with e-services. Siskos et al. (2014) develop an assessment of global e-government based on eight multiple criteria of four dimensions: infrastructures, investments, e-processes, and users' attitudes.

2.6.1.3 Official Government Frameworks

Several governments developed their official frameworks to help designers build high-quality e-government sites. A good example is the USA Research-Based Web Design and Usability Guidelines which was praised by some researchers, such as Rinder (2012), and Scowen and Regenbrecht (2009). These guidelines are widely used by government agencies and private sectors, and translated into several foreign languages (HHS, 2016).

Other examples are the Australian Service Delivery Capability Model which provides a common framework to describe the capabilities required to deliver service to citizens across public agencies, and the Canadian e-Government Capacity Check which is a suite of capacity diagnosis tools to help public agencies assess their capability to deliver e-services (Valdés et al., 2011).

2.6.1.4 E-service Quality Frameworks

Most studies in e-service quality have focused on the business sector while little attention has been paid to the services of the public sector. Hence, the research in the field

of web-based e-government service quality is scarce, as follows (Jun, Liangliang, & Fubin, 2009; Connolly, Bannister, & Kearney, 2010; Stiglingh, 2014; Butt, 2014):

- Parasuraman et al. (1988) developed SERVQUAL, the first service quality measuring instrument in a traditional (offline) environment. Loiacono et al. (2000) created WebQual scale with 12 dimensions. Another scale with the same name, WebQual, by Barnes and Vidgen (2002) measures an organization's e-commerce against five attributes: usability, design, information, trust, and empathy. Yoo and Donthu (2001) developed a 9-item scale, SITEQUAL, with four dimensions: ease of use, aesthetic design, security and processing, to measure e-shopping sites. Wolfinbarger and Gilly (2003) created eTailQ, a 14-item scale with four factors: website design, reliability/fulfillment, privacy/security, and customer service.
- In 2005, Parasuraman et al. developed a multi-item scale that is divided into normal services (E-S-QUAL), and recovery services (E-Recs-QUAL). It only focused on B2C websites that sold physical products, such as Amazon.com. The E-S-QUAL has 22 items on four dimensions: efficiency, fulfilment, system availability, and privacy, while the E-Recs-QUAL has 11 items on three dimensions: responsiveness, compensation, and contact. E-S-QUAL has received the most recognition of all proposed e-service quality scales so far. Despite its power in capturing the essence of e-services, it is an imperfect tool for assessing the service quality of e-government.

The E-S-QUAL needs medications to make the scale suitable for measuring websites which are merely service based without monetary benefits. Only a limited number of studies tried this approach. For example, the study of Jun et al. (2009) proposed E-G-S-QUAL based on E-S-QUAL by taking the service characteristics of e-government websites into account. Also by adjusting the E-S-QUAL, Connolly et al. (2010) developed the E-PS-QUAL scale to evaluate the e-service quality of the Irish tax agency. However, they caution against the e-government service quality scales because further research are needed in this area and it seems that they did not capture fully all the dimensions of service quality in e-government platforms.

2.6.1.5 Limitations of National E-Government Frameworks

- E-government benchmarks:

Several scholars criticized the validity and reliability of e-government benchmarks (Codagnone et al., 2015; Siskos et al., 2014; Grönlund, 2011; Rorissa et al., 2011; Janssen, 2010; Montserrat, 2010; Salem, 2008; Bannister, 2007). Much of the criticism is on the methodologies, which sometimes are not revealed by benchmarking organizations, and tends to compare things not really comparable (Rorissa et al., 2011; Montserrat, 2010). Montserrat (2010), Salem (2008), and Bannister (2007) assure that e-government benchmarking is a booming business. Codagnone et al. (2015) perceive the public sector assessment will likely be affected by "gaming", whereby the output is adjusted or the measurements are distorted to achieve the appearance (rather than the reality) of 'good performance'. De Róiste (2013) emphasizes the importance of understanding the drivers for benchmarking e-government studies. Some studies are produced by private companies (e.g. Accenture), while others are government-sponsored reflecting specific e-government policy objectives (e.g. Capgemini). Andersen et al. (2011) agree with Bannister (2007) that e-government ranking of nations tend to be meaningless, and it is done for the beauty contest of nations rather than for the benefits of citizens. The rankings tell half the story, therefore Andersen et al. (2011) recommend to include indicators that serve end users rather than the government. Janssen (2010) has criticized benchmarks that observe only the front-end of e-government and consider the back-end as a black box. Berntzen and Olsen (2009) notice that a service may be poorly integrated with back-office but still get a high score and vice versa.

Banister (2007) provides a detailed view of major problems with benchmarking in his paper "the curse of the benchmark". He differentiates between four different sponsors: 1) Commissioned benchmarks that are paid on behalf of a government, such as Capgemini; 2) Benchmarks to sell the research findings, such as Gartner group and partially the Brown University; 3) Academic benchmarks, such as the Brown University; 4) Benchmarks for marketing to raise a firm's profile in the e-government, such as Accenture, therefore they are unlikely to publicly criticize a potential client. Moreover, few citizens are using e-services and this casts a doubt on the usefulness of benchmarks that encourage the

development of such underused services. Banister concludes that all benchmarks should, like cigarettes, carry a large health warning.

- E-government stage models:

The stage models have several problems such as: oversimplifying reality, lacking an empirical foundation, and theoretically weak. Further research is needed to establish stage models as a field of IS of theoretical value (Poeppelbuss et al., 2011; Debri & Bannister, 2015; Bannister & Connolly, 2015). It seems that stage models are unable to provide a clear vision and roadmap to organizations attempting transformation (Klievink et al., 2009). Also, there is no consideration of change mechanisms which is important to predict how e-government evolves (Debri, & Bannister, 2015). Klievink et al. (2009) suggested that stage models should be extended with a description of the transition from one stage to the next. The transition to the final stage, transformation, seems too big a gap for government organizations to achieve. Usually, countries reach the second stage easily and quickly, as it takes no great effort to supply information, forms, and emails. In contrast, a site that advances from stage 3 to stage 4 has to go through tremendous changes that require massive efforts and resources to provide transaction and a one-stop portal (Rorissa, Demissie, & Pardo, 2008).

In fact, the conceptualization into stages is doubtful. There has been some criticism aimed at stage models, focused on the evolutionary aspect and the quality assumptions of these models: 1) the assumption that evolutionary stages are independent seems not to be true empirically. An e-government website may have the characteristics of multiple stages; 2) the assumption that evolutionary stages are consecutive, linear progressing and higher stages include lower stages, seems not to be true empirically also. The models predict that the e-government evolutions occur in pre-described order; first stage 1 occurs and then stage 2 and so on, but in practice the stages occur simultaneously. It could be that some e-portals had characteristics of advanced stages but did not have features from the early stages (Sandoval-Almazan, & Gil-Garcia, 2008a; Jansen, & Ølnes, 2014).

- The UN model:

The UN model is widely used by many studies, and it is unique in including three measures (Berntzen, & Olsen, 2009). Yildiz (2007), however, has concluded that the UN

crand Layne and Lee model (2001) are oversimplifications. Siskos et al. (2014) have criticized the UN model for having too many features. The problem in ranking occurs when a website covers some but not all features in a certain stage; then, it cannot be ranked correctly as belonging to any stage, and it is difficult to distinguish between an e-government site that fulfills 100% of the stage features and one that fulfills just 20%. The authors added that the UN model assesses the quantity and not the quality of e- services.

- The Brown model:

The Brown University reports lack a detailed description of their e-government methodology (Schellong, 2010). They give more weight to the number of features and too little to services, underestimating their importance. A government website offering 28 services is presented as equal in score to another website offering hundreds of services because the maximum score for services is 28. In addition, the reports check only the presence of services without measuring their quality (Siskos et al., 2014). Rorissa et al. (2008) have examined the profiles of two government websites according to the Brown University model and concluded that the model may suggest inaccurate conclusions. A country with a single e-government website may have the same e-government index value as a country with five websites. Another criticism of the Brown University model is that it has decreased its measurement criteria over the years; in 2001, 2002, 2003, 2004, 2005, and 2006, the number of measures were 24, 25, 20, 19, 19, and 18, respectively (Holzer, & Kim, 2005). Consequently, there were inconsistencies in annual rankings from year to year; for instance, Portugal has fluctuated in ranking from position 182 to 133, 31, 86, 43, 48, 7, and then 18 in an eight-year period (Schellong, 2010).

- The Accenture model:

Regarding the Accenture model, its strength lies in the evaluation of the maturity of e-services following a hybrid methodology, quantitatively assessing the breadth and depth of e-services and qualitatively appraising the customer service delivery. Another strength is the indicator, “citizen voice”, which tried to integrate user views of e-government. On the other hand, its main weakness is continual changes in methodology and measurements, which make it impossible to compare e-government rankings over the years (Berntzen, & Olsen, 2009). Moreover, Accenture provides no details of measured services and their

maturity scores. Thus, the calculation of the indices is not reproducible. The authors added that this model lacks an evaluation of integrated services, and is limited in its application to only 22 countries. Another important issue Accenture has stopped e-government benchmarking after the 2007 report.

- The Capgemini model:

The most common critique of the Capgemini model is its focus on the government side only. Kunstelj and Vintar (2004) criticize Capgemini for its measuring the availability of 20 public services despite some of these services bringing no value to customers. They add that highlighting the technological side of e-government without considering the quality of information and usefulness of services will miss important qualitative aspects of e-government. Bannister (2007) argued that Capgemini benchmark: 1) has little credibility since it does not measure level of e-services' usage, citizen satisfaction, and how thorough the tests of e-transaction can be when one is looking at several hundred of them in a relatively short period; 2) has no measure of back office progress or service integration; 3) allows some debates around the draft evaluation results with the concerned countries before publishing, so that what emerges contains certainly some element of negotiation. In addition, this model is narrow in its scope, being concerned only with European countries.

Kromidha (2012) discusses the role of donor–benchmarker duality. He questions the purpose of some benchmarking studies and relates them to a desire to attract funds or win additional e-government business. He warns that private companies preparing the benchmarking, such as Capgemini, may be among the first to contact for premium expertise and can benefit from the benchmarking–consultancy combination. Further, Codagnone et al. (2015) believe that the method used to score services on stage 3 and 4 for EU benchmarking leaves plenty of room for gaming. Public services can be available online only for the purpose of achieving the score, without re-organisation of the services. The analysts evaluating e-government sites can only check whether a feature is present and look at the description of a service but they cannot try it online. They would have to be citizens with identity cards and own e-signatures in the countries benchmarked. The authors tried several services that were scored as fully transactional in some countries and discovered

that, in reality, at the end of the online procedure they only received a ‘pdf’ form to be delivered in person to the public office.

The recent editions of Capgemini extended the Layne and Lee model (2001) by adding a fifth stage (targetisation) to reflect proactive service delivery toward reaching a citizen-centric e-government (Kotamraju, & Der Geest, 2012). However, the new model is so late in considering “citizen-centricity” since its roots in the HCI field date back to the nineties. Jansen and Ølnes (2013) still criticize Capgemini for being too focused on the supply side of e-government and not really user oriented yet. Schellong (2010) pointed out that the model attempt to benchmark “citizen-centricity” as a constructed measure but the problem there is no clear understanding how it should be measured. He added that user-centricity indicators remain an area of testing and further improvement is needed in the future.

Grönlund (2010) questions the depth of the new EU model and considers many measures for “better government” (e.g. transparency, accountability, and participation) are shallow. He believes that the next generation of e-government research must take up the challenge and contribute to define ways of assessing them. Implementing “full case handling” is understood, while using ICT to make government better is still a great challenge. Codagnone et al. (2015) criticize the validity of the benchmarking indicators concerning the relation between the supply of e-services and their usage by citizens. Countries with sophisticated websites can have low levels of use and vice versa. An analysis of EU benchmarking data on e-government usage by citizens shows that the level of supply does not have any effect on demand. Paradoxically, the new edition has reduced the reliability of the measurement through introducing subjectivity in scoring websites. The authors concluded that there is no theoretical framework or justification of the selected indicators and the rationale for continuing this form of benchmarking is very weak.

- Layne and Lee model (2001):

Andersen and Henriksen (2006) argue that Layne and Lee model (2001) just replicated the stage models from the e-commerce area and focused on technological capabilities rather than on effectiveness in the public administration. Therefore, they proposed another evaluation approach that paves the way to customer-centric e-government. Klievink and

Janssen (2009) claim that Layne and Lee model (2001) lacks a clear theoretical foundation and the model of Andersen and Henriksen (2006) does not provide empirical evidence.

- Research-Based Web Design and Usability Guidelines:

Shneiderman (2011) and Scowen and Regenbrecht (2009) have praised the HHS guidelines of USA e-government websites for being well-designed and supported in the HCI field. Rinder (2012) and Dingli and Mifsud (2011) have confirmed that they have been validated empirically. Buie and Murray (2012) mentioned that subsets of the HHS guidelines can be tailored for particular audiences. However, one can say it may be difficult to evaluate a website against too many guidelines (about 209 guidelines); it may be better for the HHS to work on providing the web community with a shorter list of guidelines.

- Focusing on government, not citizens:

The problem with most of the national e-government models is their focusing on the supply side (government) not the demand side (citizen and business) of e-government (Berntzen, & Olsen, 2009). Two examples of the supply-side models are West and Capgemini, while the demand-side models are like Gartner and HHS guidelines (Rorissa et al., 2011; Scowen, & Regenbrecht, 2009). The imbalance of the abundance of government-side surveys compared with the scarcity of citizen-side studies has led to a misinterpretation of the final objective of e-government. The existing practices are pushing countries to prioritize getting good ratings for creating many services without caring whether citizens use them or not (Montserrat, 2010). Moreover, the majority of models, such as the UN, Capgemini, and Brown, follow a quantitative approach; only Accenture uses hybrid measures. Thus, most surveys do not evaluate qualitative issues, such as the quality of service or the citizen usage of e-government; that means higher ranking may not predict better performance (Salem, 2008).

- E-democracy:

E-government evaluation models which has a fifth stage “e-democracy” or political (citizen) participation, defined as enabling the public to participate in online public consultations, policy making and e-voting (Chatfield, & Alhujran,2009), is criticized by many studies (Debri, & Bannister, 2015; Norris, & Reddick, 2013; Klievink, & Janssen,

2009; Coursey, & Norris, 2008). For example, they agree that the early stages of Moon's model (2002) are reasonably accurate as they are taken from empirical observation, unlike the later stages which are predictive and aspirational. In fact, e-government has not reach higher mature stages as predicted by these models. It seems that e-participation reflect authors' hopes rather than be based on solid theory or sound reasoning. These studies concluded that there is no logic to believe the highest stage of such development will be e-democracy since none such models take into account politics or consider its impact.

There are other studies that criticize the UN model for the same reason. Berntzen and Olsen (2009) state that the UN model evaluates the e-government website for e-participation, but is this the right place to look for participation? Grönlund (2011) investigated the UN e-participation stage as a measure of how well governments connect to their citizens. He believed that the relation between the UN index and democracy is non-existent. Countries which are authoritarian can score high on e-participation by window-dressing their webs. Potentially the e-participation index is a misleading tool as the model is not related to the real world of government.

- Methodological limitations of e-government:

Formerly mentioned frameworks revealed that many e-government reports were based on different measurement instruments, which explains the difference in e-government rankings and the disparity of conclusions. Several scholars of e-government are skeptical about the e-government rankings and have justifiably argued that existing e-government frameworks have some methodological limitations (Schellong, 2010; Yildiz, 2007; Rorissa et al., 2011; Sandoval-Almazan, & Gil-Garcia, 2008b; Codagnone et al., 2015). Grönlund (2011) argued that the field of e-government is weak theoretically and Tsohou et al. (2013) stated that e-government evaluation is immature. Bannister and Connolly (2015) confirm that e-government has begun to develop as a field but it is still under-theorised. Grönlund (2010; 2011) and Ataloglou and Economides (2009) have concluded that a good theoretical framework for measuring e-government is still lacking. Schellong (2010) and Karkin and Janssen (2014) have said that there is no generally accepted comprehensive e-government evaluation framework and no universal standard for assessment of national e-government.

2.6.2 Local E-Government (Municipal) Models

The term local government can be considered as comprising governments that are not central, national, nor federal but includes state, provincial, regional, municipal and city governments (Lanvin, & Lewin, 2008). Similarly, Arslan (2008) refers to local governments as municipalities or e-cities and thus it is not independent of the concept of e-government. In global context, the term "municipal e-government" is used in Europe while in the US the term "local e-government" is more likely to be used (Zahran et al., 2015). This research will use both terms interchangeably. From e-cities' perspective, Kaylor, Deshazo and Van Eck (2001) derived a wider definition of e-government as the ability for anyone visiting the city website to communicate and interact with the city via the Internet in any way more sophisticated than a simple email letter to the city email address.

Increasingly, local governments impact citizens' lives and become the key players. As a result, focusing on the citizen centric websites should be at the core of e-government and through this approach municipalities need to acknowledge and work towards improving the online citizen government relationship (Moraru, 2010). Most of the frameworks evaluating municipal websites are based on Moon's framework which was discussed in section 2.6.1.2.

2.6.2.1 Digital Governance in Municipalities Worldwide

This benchmark is noteworthy for its attempt to compare e-cities globally. It has been conducted every two years since 2003 through a collaborative effort between the E-Government Institute at Rutgers, the State University of New Jersey, Global e-Policy e-Government Institute at Sungkyunkwan University and co-sponsored by the United Nations. The largest cities in the top 100 most wired countries that have official municipal websites were evaluated. Holzer, You and Manoharan (2009) justified their procedure of city sampling in accordance with Moon's study. By this manner, the largest city represents each selected country regardless of whether it is the most advanced in e-government. Montserrat (2010) criticized the sampling of the cities in this survey and considered it biased. From the positive side, it is the only one that evaluates municipal websites worldwide in term of digital governance which includes digital government (public service) and digital democracy (Holzer et al., 2009). Also, the methodology of digital governance remains constant over the years and that means all of its results are comparable.

Meanwhile, the used instrument for assessing city websites consisted of five equally weighted components (Table 2.6): security and privacy, usability, content, services, and citizen participation. To ensure reliability, each municipal website was assessed by two evaluators given clear instructions (Holzer et al., 2009). No information was given about the evaluators' background and their degree of expertise. The research applied 18-20 measures coded on a scale of (0, 1, 2, 3), 1: information about a given topic exists on the website, 2: downloadable items are available, 3: services, transactions, or interactions take place completely online. Hence, the survey instruments utilized 98 measures.

Table 2.6: E-Governance Performance Measures (Holzer et al., 2009)

E-governance Category	Key Concept	Raw Score	Weighted Score	Keywords
Security/ Privacy	18	25	20	Privacy, authentication, encryption, data management and cookies
Usability	20	32	20	User-friendly design, branding, length of homepage, targeted audiences links or channels and site search
Content	20	48	20	Access to current information, public documents, reports, publications, and multimedia materials
Services	20	59	20	Transactional services - purchase or register, interaction between citizens and government
Citizen participation	20	55	20	Online civic engagement/ policy deliberation, and citizen based performance measurement
Total	98	219	100	

2.6.2.2 MeGAP: US Municipal E-Government Assessment Project

The MeGAP (The Municipal E-Government Assessment Project) is an assessment tool for US municipal websites emphasizing online service provision. Kaylor et al. (2001) surveyed 38 American cities with a population between one and two hundred thousand people and developed a rubric for evaluating them. Functional performance dimensions were grouped into 12 categories containing 51 e-services. To rank municipalities, each service was scored on a 1–4 scale (information, contact, downloadable forms, and transaction or interaction) that yielded an e-score corresponding roughly to the stage model concepts (Flak et al., 2005). In 2005, the third version of Kaylor's survey maintained the original framework, but the catalogue expanded to 68 local services in 4 categories

(Montserrat, 2010; Flak et al., 2005): (1) Information dissemination (city codes, minutes, traffic information, municipal government directory); (2) Interactive functions (bidder applications, downloadable forms, building permit process, business license); (3) E-Commerce functions (utility payment, tax lookup and payment, code enforcement); (4) E-Democracy (e-meetings, e-forums, user customization).

2.6.2.3 Municipal Website Assessment of Community Benchmarks Program

The Maxwell School at Syracuse University established the Community Benchmarks Program (CBP) in 1999 and developed a website assessment instrument to evaluate e-municipalities in Onondaga County. Denfeld et al. (2002) reevaluated the previous study and devised the following assessment criteria: Information available: municipal meeting, minutes, budget, downloadable forms, date of website update; Contact information: phone and fax numbers, e-mail, physical address; Architecture: search, site map, link function properly, link to home page provided; Continuity of web design: consistent design of all pages; Search engines: placement of the municipality's website on Yahoo, Google and MSN for official name, popular name; General: responsiveness of town clerk, unique features of each site both well and poorly-executed, broken links. The 2002 report assigned each attribute a score of 1 if the website met the criterion or 0 if it did not. An example of a blank evaluation form for the "Information Available" criteria is in Table 2.7.

Table 2.7: An Example of a Blank Evaluation Form (Denfeld et al., 2002)

	Information Available			
	Attribute	Yes	No	Comment
1.	Date of **next** town/village board or city council meeting (acceptable in place of the date is, i.e., third Monday of the month)	1	0	
2.	Location of town/village board or city council meeting	1	0	
3.	Time of town/village board or city council meeting	1	0	
4.	Agenda of next town/village board or city council meeting (posted within 48 hours of meeting – Use last meeting agenda, if necessary)	1	0	
5.	Minutes of the last town/village board or city council meeting held within the last two months	1	0	
6.	Archive of past board or council meeting minutes	1	0	
7.	Minutes can be downloaded	1	0	
8.	Budget for the current fiscal year	1	0	
9.	Downloadable forms	1	0	
10.	A date is provided for the most recent Web update	1	0	Date _____

2.6.2.4 Key Elements for Electronic Local Authorities' Network (KEeLAN)

The Key Elements of Electronic Local Authorities' Network (KEeLAN) is a local e-Europe government framework and is also known as "Framework Programs." Started by e-Europe research, the KEeLAN model is divided into two phases measuring e-government and back-office development. The e-government stages are divided into six phases: stage 0: no Web presence; stage 1: information (about services); stage 2: interaction (downloading forms); stage 3: two-way interaction (processing of forms including authentication); stage 4: transaction (full case handling); stage 5: service integration (online service enabled by a secured network linked to various back-offices/service modules). The stages are exactly the same as in the Capgemini model, except the last one. In this context, a Web assessment tool contains questions to evaluate e-cities on 9 basic services: policy making, economic development, personal documents, credit and loans/financial support, education, building permits, environment, culture and leisure, and information dissemination. Depending on the interactivity, a score is computed to indicate the stage of the service (Arslan, 2008).

2.6.2.5 Dubai Government Websites Excellence Models

Dubai eGovernment Department developed government websites guidelines to be adopted by Dubai Government Entities. The "Government Websites Excellence Model" (GWEM) (Figure 2.13) provides the necessary control on how to ensure that government websites are designed and managed. It is based on extensive research and benchmarking to achieve maturity in line with internationally accepted website best practices and standards.

Figure 2.13: Government Websites Excellence Model (Dubai eGovernment Department, 2011)

The model is built around a “Customer-Focus” concept. A step toward a successful customer-focused website is to understand the users/customers of the website. The ability to create usable and useful website designs is highly dependent upon a clear audience definition. On the light of this concept, the model consists of 46 guidelines (Dubai eGovernment Department, 2011):

- Accessibility (6 Guidelines): provide access to the website through an easy to remember URL including an appropriate representation of the entity name under (.gov.ae) domain; provide a quick access to the website from a search engine; provide access to the website with identical and consistent results through a wide range of web browsers; provide a functional bilingual website; provide appropriate access to website files; provide access to the website for people with disabilities.
- Usability and Design (20 Guidelines) such as: provide a clear and readable entity and Dubai Government logos; provide a well-designed customer focused Homepage; provide a well-structured and effective sitemap; provide an effective and efficient Search functionality; provide a logically organized and easy to navigate website; use an appropriate design for website links; provide clear and meaningful links on the website; provide simple and easy to use forms; provide a functional print facility; provide a consistent format throughout the website.
- Content (17 Guidelines) such as: provide information about the Government Entity in "About Us"; provide Entity Contact information in "Contact Us"; provide a facility to submit feedback on the site; provide information about Government Entity e-services; provide a proper "Site Maintained By" message; provide a link to eJob, eSuggest, eComplain and Ask Dubai; provide accurate and most up to date information.
- Policies (3 Guidelines): provide information on the protection and handling of privacy, on the website terms and conditions and on the accessibility of the website.

2.6.2.6 Limitations of Local E-Government Models

Even though most of the time the interaction between citizens and government occurs at the local levels, very little research describes or analyzes existing local e-government models. There is a disproportionate number of studies focusing on national e-government

models compared with that of studies targeting local e-government models (Montserrat, 2010; Shareef et al., 2012).

Nevertheless, the UN's "Digital Governance in Municipalities Worldwide" is still the only international survey of e-cities. By supporting two different models, the UN implies that there is a difference between assessing central e-governments and assessing local ones. The methodology of digital governance has remained constant over the years, so its rankings of cities are comparable and remarkably informative. On the other side, Montserrat (2010) regards the sampling in this survey as biased. Also, the survey gives no justification for the framework measurement evaluation criteria, which constitutes a major weakness in the methodology. Each municipal website was assessed by two evaluators given clear instructions (Holzer et al., 2009). But, no information was given about the evaluators' backgrounds and their degree of expertise.

For the MeGAP of the US e-municipalities, Flak et al. (2005) believed that this model gives a more detailed analysis of the depth and breadth of municipalities than any other assessment model; but, on the other hand, the MeGAP lacks a firm theoretical foundation, doesn't assess usability, and it is a country-specific model. The major drawback of the Community Benchmarks Program is that it focuses only on the supply side of e-government. Since the two models are similar, the KEeLAN model suffers from the same problems as the Capgemini model, such as invalid benchmarking indicators and not really user oriented yet (see section 2.6.1.5).

2.6.2.7 Comparison of National and Local E-Government Models

At the national level, the existing benchmarking e-government models are very similar and are based on analogous attributes and measures; they view e-government as stages of growth and adopt four or five stages: Web presence, interaction, transaction, integration (portal), and e-participation or e-democracy (included in few models). Toonders (2010) has deemed it unclear whether the same stages of national e-government are useful for describing local e-government. Norris (2009) has cast doubt on the adequateness of stage models in municipalities. He used survey data from US municipalities over three years (2000, 2002, and 2004) and empirically examined how e-government has developed in

practice and contrasted this with the predictions of the models. The US e-municipalities did not progress through stages as anticipated. They were informational with fewer transactions and interactions and had not evolved into e-democracy yet. Norris (2009) attributed that to the e-government models having been developed in a vacuum and not being based on research or reviews of the literature. He concluded that even after 10 years of adoption, e-government has not reached higher stages of development in most countries.

Again Norris and Reddick (2013) addressed the trajectory of USA local e-government using empirical data from two nationwide surveys of American municipalities conducted in 2004 and 2011. They found American municipalities are delivering information and services online with few transactions and limited interactivity and they are mainly one way, from the government to citizens, with no evidence that it is transformative. The authors also presented more empirical studies of e-government; for example: service has been the primary focus of e-government in various locations: the United Kingdom (McLoughlin, & Cornford, 2006), Canada (Roy, 2006; 2007), Australia (Dunleavy et al., 2008), the Arab nations (Chatfield, & Alhujran, 2009), and Italy (Nasi, & Frosini, 2010). This is consistent with the conclusion of Sandoval-Almazan and Gil-Garcia (2012) who said that almost a decade after the publication of a similar study on U.S. municipalities by Moon (2002), the results of their assessing Mexican municipalities remain very similar. They believed that e-government in cities is still more rhetoric and less reality, at least in some countries.

In fact, the e-government experience differs dramatically from the national to the local level and from one country to another. Montserrat (2010) believes that the indicators and metrics defined for national e-government are not applicable at the local level. He asks, “Why are there no benchmarks at local government?” Collecting comparable data about e-municipalities is a difficult task because of differences in political and economic systems. The different role played by cities is one of the challenges that scholars must address. Montserrat (2010) confirms a clear lack of local e-government evaluation models. Shareef et al. (2012) assured that most studies focus on national e-governments, although in developing countries it is local e-governments that are the main point of contact for delivery of services. Most public services that are relevant to citizens are offered by the

local e-government, and this is a possible source of error in the assessments (Berntzen, & Olsen, 2009; Schellong, 2010).

Through its development of two models, the UN demonstrated the difference between assessing national and local e-governments. For assessing state portals, Sandoval-Almazan and Gil-Garcia (2008a) identified three approaches: 1. managerial, 2. evolutionary (e-government stages) and 3. citizen-centered perspectives. Using a mixture of the last two approaches, they assessed 32 Mexican portals against a six-stage model and also against other important variables such as usability, openness, customization, transparency, e-services, privacy, security, etc. Another contribution by Goldkuhl and Persson (2006) is a proposal to replace the one-dimensional stage models (called e-ladder) by a three-dimensional e-diamond model consisting of three polarities (informative vs performative, standardized vs individualized; separate vs coordinated). However, there are individual efforts by some authors, such as Moraru (2010) and Luna et al. (2013), who use a mixture of e-government stages and some other components they perceived important in the evaluation of municipal websites.

Upon analyzing existing normative models on municipalities (Table 2.8), it is noticeable that some of them, such as the UN Digital Governance in Municipalities and CBP, focus on general aspects of the site such as content and services. They avoid the concept of stage models and instead regard local e-government as different components or categories. Other models, such as the KEeLAN and MeGAP, follow the stage model (Arslan, 2008; Flak et al., 2005).

Table 2.8: Two Kinds of Municipalities Models

Models	Kind of Model	Descriptions				
Digital Governance in Municipalities 2003	Components	Security /Privacy	Usability	Content	Services	Citizen Participation
CBP 1999	Components	Content	Architecture	Layout		Website Design
MeGAP 2001	Stages	Information	Contact	Downloadable Forms		Transaction or Interaction
KEeLAN 2002	Stages	Information	1-way Interaction	2-way Interaction	Transaction	Service Integration

2.7 E-Government in Saudi Arabia

Saudi Arabia is moving toward e-society and e-government rapidly. Some positive encouraging signs along the road as well as some obstacles and slow growth in e-government are explored next.

2.7.1 Country Overview

The Kingdom of Saudi Arabia (KSA) is situated in Asia continent in the Middle East region. SA is a large country with an area of about 2.1 million km^2 and a population of 27.3 million (Internet World Stats, 2015). Major cities (Figure 2.14) are Riyadh (the capital) 5.451 million, Jeddah (the commercial capital and the main port on the Red Sea) 3.578 million, Mecca (the first holy city) 1.591 million, Al-Madina (the second holy city) 1.142 million and Dammam (the second port) 941,000 as estimated in 2011. The population is very young; the 2014 estimated distribution according to age is, 0-14 years: 27.6%,15-24 years: 19.3%, 25-54 years: 45.4% and 55 years and over: 7.6% (The World FactBook, 2015). This young population could be a growth driver to technology and e-government adoption as they grow up with the Internet.

Figure 2.14: Map of Saudi Arabia (The World FactBook, 2015)

Saudi Arabia is a monarchy. The cabinet of 29 ministers is appointed by the king. It is the responsibility of the Council of Ministers to formulate the High Command and oversee the implementation of internal and external policies. The country is divided to 13 provinces each with a governor. While the government is central and responsible for issuing and

adopting regulations, provincial governments can only enforce these regulations. Also, the Shura Council assists the Council of Ministers by conducting studies and raising recommendations to the Cabinet for the Prime Minister's approval and adoption. Economically, Saudi Arabia is rich in natural resources, especially oil, which is the main source of financial income. This helped rapid development in Saudi cities in all fields of infrastructure, public utilities, education, and services (Albassam, 2012; Ajaj, 2014).

The first election in the country was the election of the local municipal council in 2005 and the second one took place in 2011. These two elections were for men only and held for half of the local council seats while the government appointed the other half. Since no political parties are allowed, all candidates were independents. The municipal councils have little power and its role is a very advisory at the city mayor level. It discusses issues like budget allocation, maintenance of amenities, and street lighting (Albassam, 2012; Ajaj, 2014). In the 2015 municipal elections, women were also allowed as candidates and voters to elect two-thirds of 284 municipal councils. Only 1.48 million Saudis from a population of 20 million registered to vote in the election, including about 131,000 women and 1.35 million men (http://www.intekhab.gov.sa/).

2.7.2 Saudi E-Government Initiative

The Internet was introduced in Saudi Arabia in 1997. The IT structure began in 1998 with the first Saudi telecom company (STC). The Telecommunication Commission was established in 2001 and the MCIT (Ministry of Communication and Information Technology) in 2003 to control IT services in the country by formulating the Communication and Information Technology Authority (AlSabti, 2007). Though it was founded by a supreme royal decree in 2003, Saudi e-government program did not actually start until 2005 (Sahraoui, Gharaibeh, & Al-Jboori, 2006). That means the country started its e-government project later than many other Arabic neighboring countries (Al-Saif, 2010); for example Dubai (Geray, & Al Bastaki, 2005), Qatar (Al-Shafi, & Weerakkody, 2010) and Jordan (Mofleh, Wanous, & Strachan, 2008) initiated their e-government journey in 2000. The e-government program dubbed "Yesser", an Arabic word which means "simplify" or "make easy", plays the role of the enabler or facilitator of e-government in the public sector by building the national infrastructure and defining standards. Yesser was

initiated in cooperation with IBM to ensure an appropriate level of collaboration between different government bodies (Buragga, 2010). Nevertheless, each government entity in SA is in charge of its own digital transformation (Sahraoui et al., 2006).

In 2006, the Saudi government assigned a big budget to its e-government project 3 billion Saudi Riyals (SR) (Yesser, 2015), US $800 million, for a 5-year plan (Sahraoui et al., 2006). Then the budget increased to 4 billion riyals, US $1.2 billion, for the year 2010. Therefore, Saudi public sectors have enough government financial support to publish their own services online (Alshehri, & Drew, 2010). The vision for Saudi e-government in the first national 5-year plan (2006–2010) is: By the end of 2010, everyone in the Kingdom will be able to enjoy world-class government services offered in a seamless, user friendly and secure way by utilizing a variety of electronic means. The objective is to provide 150 e-services available to everybody anytime with 75% adoption rate and 80% user satisfaction (Yesser, 2015). Figure 2.15 depicted the time table for the e-services and the beta version of the national portal in 2007.

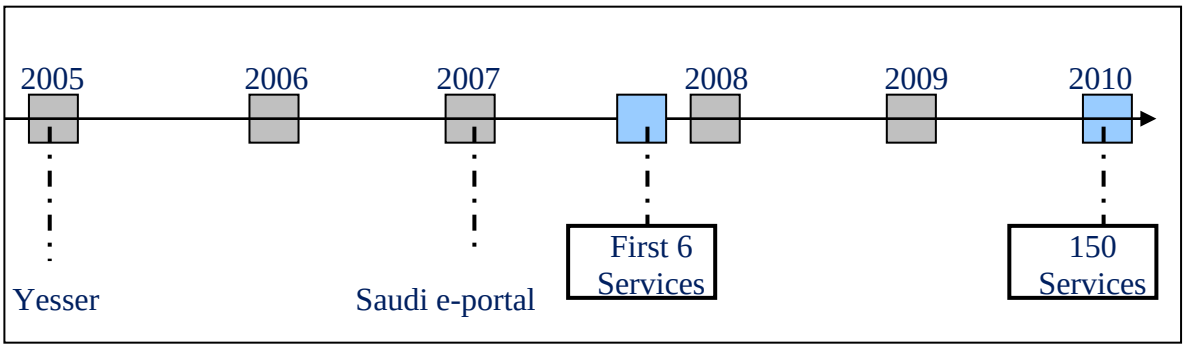

Figure 2.15: Time Table for the Initial 150 Saudi E-services

Based on a thorough review of ministries' websites, Sahraoui et al. (2006) concluded that Saudi e-government is rather far from world standards and consequently the Saudi vision of high e-services adoption and user satisfaction rates might not be reachable within the specified time horizon. Al-Shehry et al. (2006) agreed with this expectation and added that Saudi e-government suffers from design-reality gaps defined as the oversize gaps between project design and on-the-ground reality. Anthopoulos et al. (2016) considered design-reality gaps is still a problem that faces e-government projects in developing

countries and might lead to e-government failure totally or partially. Al-Shehry (2008) commented that the first Saudi national 5-year plan is an over-ambitious vision because several critical issues need to be addressed before such lofty goals can be achieved. A manager in one ministry told Al-Shehry (2008): "we need to sort out managerial problems and the re-engineering process before going online. Otherwise, we end up transferring this problem and making it more complex with computers". A study by Alfarraj et al. (2011) is not satisfied with the rank of Saudi e-government in the UN 2010 report, 58 worldwide and 4th among the Gulf countries since it is far from the expectation for 2010 as the Saudi e-government had predetermined. As the year 2011 passed, the Saudi e-government program timetable was not achieved as expected in light of what has been done so far and as indicated in the literature.

Moreover, there is an enduring debate among Saudis about e-government. The organization structure of the e-government program is composed of five parts: the supreme supervisory committee, the steering committee, the advisory group, the e-government committee in each government organization and Yesser administration. The supervisory committee consists of the Minister of Finance, the MCIT, and the Governor of the Communications and IT Commission (Yesser, 2015). According to Sahraoui et al. (2006), a major obstacle that slowed down the development of Saudi e-government is the absence of a central authority to oversee the implementation of the program. No identifiable entity is responsible for the digital migration of the entire government. Ministries and other government agencies are to separately implement their slice of the e-government plan as they wish. While the Yesser website provides some details on plans, it lacks important information on timelines, objectives and especially what has been accomplished up to date. There is no comprehensive e-government project to oversee scheduled execution of a clearly defined plan to install e-government in governmental institutions. An independent expert interviewed by Al-Shehry (2008) commented: "We have political support but there is no follow-up for e-government implementation. In other words, the e-government team cannot force ministries to change towards e-government at a specific time". This holds true even today.

Some experts interviewed by Al-Shehry (2008) believed that the MCIT should be responsible for the technical parts of ICT but the whole e-government project should be coordinated by a joint group linked directly to the Saudi Cabinet. Others argued that there should be e-government committees in every organization supported by top executives to supervise the implementation of the e-government plan in their respective organizations. Implying his dissatisfaction with the performance of MCIT, a top e-government manager suggested to Al-Shehry (2008) that e-government should be given to a ministry that has direct interaction with citizens.

The second national action plan (2012 to 2016) has the following vision (Yesser, 2015): "Enabling everyone to use effective government services, in a secure integrated and easy way, through multiple electronic channels". The human resources, communication, and change management is the most important work stream of the second action plan. Notice that this plan doesn't specify the number of e-services to be provided, adoption rate, and neither user satisfaction rate.

2.7.3 E-Readiness in Saudi Arabia

The first Arab country to link to the Internet was Tunisia 1991, then Kuwait 1992, Egypt and the UAE 1993 and Jordan 1994, while Saudi Arabia and Syria were the slowest countries in the region to allow the Internet (Wheeler, 2007). In particular, the Internet was first launched in Saudi Arabia in April 1997 (Al-Shehry, 2008). According to the Internet World Stats (2015), only 200,000 Saudis were using the Internet in the year 2000 (Table 2.9). But by 2014, the Internet users rose up very quickly and became 18,300,000 users out of 27,345,986, which is 66.9% of the estimated population. That represents 16.4% of Internet users in the Middle East. Thus, the usage growth is very large and that gives Saudis an optimistic future in the diffusion of technology and a solid ground for an e-government. However, the digital divide is still substantial and Internet penetration is relatively low. Among the 15 Middle Eastern countries, Saudi Arabia is in the 9th position in Internet penetration. More Saudi efforts are needed to catch up with other leading countries in e-government.

Table 2.9: Saudi Arabia Internet Usage Statistics (Internet World Stats, 2015)

Middle East Internet Users, Population and Facebook Statistics						
MIDDLE EAST	**Population (2014 Est.)**	**Users, in Dec/2000**	**Internet Usage 30 Jun-2014**	**% Population (Penetration)**	**Internet % users**	**Facebook 31 Dec-2012**
Bahrain	1,314,089	40,000	**1,297,500**	98.7 %	1.2 %	413,200
Iran	80,840,713	250,000	**45,000,000**	55.7 %	40.2 %	n/a
Iraq	32,585,692	12,500	**2,997,884**	9.2 %	2.7 %	2,555,140
Israel	7,821,850	1,270,000	**5,928,772**	75.8 %	5.3 %	3,792,820
Jordan	6,528,061	127,300	**5,700,000**	87.3 %	5.1 %	2,558,140
Kuwait	3,268,431	150,000	**3,022,010**	92.5 %	2.7 %	890,780
Lebanon	4,136,895	300,000	**3,336,517**	80.7 %	3.0 %	1,587,060
Oman	3,219,775	90,000	**2,584,316**	80.3 %	2.3 %	584,900
Palestine (West Bk.)	2,731,052	35,000	**1,687,739**	61.8 %	1.5 %	966,960
Qatar	2,123,160	30,000	**2,016,400**	95.0 %	1.8 %	671,720
Saudi Arabia	27,345,986	200,000	**18,300,000**	66.9 %	16.4 %	5,852,520
Syria	22,597,531	30,000	**5,920,553**	26.2 %	5.3 %	n/a
United Arab Emirates	9,206,000	735,000	**8,807,226**	95.7 %	7.9 %	3,442,940
Yemen	26,052,966	15,000	**5,210,593**	20.0 %	4.7 %	495,440
Gaza Strip	1,816,379	n/a	**n/a**	n/a	n/a	n/a
TOTAL Middle East	231,588,580	3,284,800	**111,809,510**	48.3 %	100.0 %	23,811,620

Based on Saudi Communications and Information Technology Commission (CITC, 2015) website, the number of Saudi Internet users doubled in a six years period, from 9.3 million in 2008 to 18.3 million in 2014 (Figure 2.16). A dramatic shift occurred with the Internet penetration in the country increasing to 60.1% of the population by 2014.

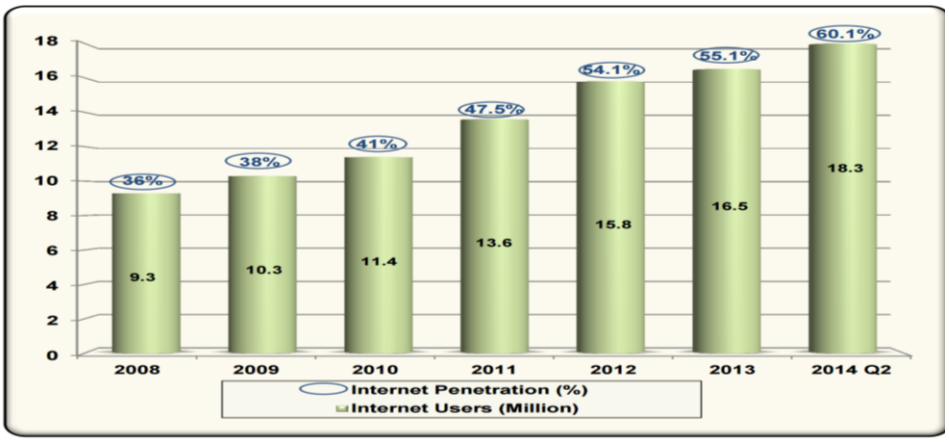

Figure 2.16: Internet Market Evolution 2008 - 2014 (CITC, 2015)

Table 2.10 provides a closer look at the Saudi Telecommunication Infrastructure Index (TII) and some selected countries from the UN Report 2014. The number of mobile subscribers grew faster than other countries and the demand for wireless broadband services has increased significantly. When compared with other developing countries, the trend toward using and owning a technology by the Saudi is good but unfortunately, there is no Saudi local production of the software or hardware materials (Al-Ghaith, Sanzogni, & Sandhu, 2010). The increased demand for ICTs is met by purchasing overseas technologies. AlZahrani (2011) warned that poor Internet service at a high cost due to a lack of competition can be a key barrier to the adoption of Saudi e-government.

Table 2.10: Telecommunication Infrastructure Index (United Nations, 2014)

Country	TII	Internet Users /100 inhabitants	Telephone lines /100	Mobile subscribers /100	Fixed broadband /100	Wireless broadband/100
UK	0.8534	87.02	52.58	130.78	34.04	72.06
France	0.8003	83.00	61.45	97.41	37.47	51.77
USA	0.7406	81.03	43.78	97.64	27.88	74.90
Bahrain	0.7055	88.00	22.01	161.17	13.14	78.42
UAE	0.5932	85.00	21.37	149.64	10.34	44.85
Saudi Arabia	0.5523	54.00	16.97	187.40	6.59	45.38

2.7.4 Global Ranking of National and Local Saudi E-Government

At the national level, the UN e-government reports for the year 2003, 2004, 2005, 2008, 2010, 2012 and 2014 have ranked Saudi Arabia 105, 90, 80, 70, 58, 41 and 36 respectively

out of 193 countries (Table 2.11) (United Nations, 2003; 2004; 2005; 2008; 2010; 2012; 2014). In a period of 11 years, Saudi e-government has jumped 69 positions and improved its e-government score by 0.352 which is a substantial gain based on the UN statistics.

Table 2.11: Saudi E-government Rankings
(United Nations, 2003; 2004; 2005; 2008; 2010; 2012; 2014)

Year	Rank	Score (out of 1.0)
2003	105	0.3380
2004	90	0.3858
2005	80	0.4105
2008	70	0.4935
2010	58	0.5142
2012	41	0.6658
2014	36	0.6900

The United Nations (2014) e-government development index (EGDI) gives more details on Saudi e-government (Table 2.12), and it seems that some Arab countries are doing well in EGDI index; Saudi Arabia is ranked 36 at score 0.6900. For the online services, Saudi e-government has a value of 0.7717 which means that a good number of e-services are offered to the public. The Saudi infrastructure growth is in a middle way through development with a score of 0.5523. The situation in the Human Capital Index for Saudi (0.6671) is much better and that means that citizen readiness is not an issue.

At the local level, Riyadh is the only city from Saudi Arabia that was assessed by Digital Governance in Municipalities Worldwide (Holzer et al., 2014). Based on the 2013-14 evaluation of 100 e-cities, Riyadh ranked 41 with a score of 35.59. Globally, Dubai was the 4th in e-services, the 5th in the privacy and security division, the 13th in usability and the 19th in e-participation section. For Riyadh, it is in position 20 in the privacy and security with a score of 9.45 out of 20, 23 in usability (scoring 14.38), 58 in content (scoring 6.35), 77 in e-service (scoring 2.17) and 50 in e-participation (scoring 2.17). All e-ranks and scores of Riyadh from 2003 till 2014 were combined in a table (Table 2.13). In all evaluated

categories, Riyadh scored relatively low except in usability where it has the best score 14.38 out of 20. In privacy and security, it even gets zero in 2003 and 2007 then gradually improved by 2014. The last three components, content, e-service and e-participation, Riyadh's score is low as 2.71.

Table 2.12: Some Countries from E-government Development Index
(United Nations, 2014)

Rank	Country	EGDI	Of which		
			Online Service Component	Telecommunication Infrastructure Component	Human Capital Component
1	Korea	0.9462	0.9764	0.9350	0.9273
2	Australia	0.9103	0.9291	0.8041	0.9978
3	Singapore	0.9076	0.9921	0.8793	0.8515
4	France	0.8938	1.000	0.8003	0.8812
7	United State	0.8748	0.9449	0.7406	0.9390
8	UK	0.8695	0.8976	0.8534	0.8574
18	Bahrain	0.8089	0.9370	0.7055	0.7840
32	UAE	0.7136	0.8819	0.5932	0.6657
36	Saudi Arabia	0.6900	0.7717	0.5523	0.6671
44	Qatar	0.6362	0.6535	0.5879	0.2932

Table 2.13: Riyadh E-City Ranks in Digital Governance 2003 – 2014
(Holzer, & Kim, 2003; 2005; 2007; Holzer et al., 2009; Holzer, & Manoharan, 2012; Holzer et al., 2014)

Riyadh, Saudi Arabia							
Year	Rank	Score (out of 100)	Privacy & Security (out of 20)	Usability (out of 20)	Content (out of 20)	Service (out of 20)	Participation (out of 20)
2003	57	18.697	0.00	10.313	3.404	4.211	0.769
2005	52	24.68	3.20	13.13	5.00	1.36	2.00
2007	73	18.15	0.00	10.63	5.60	1.01	0.91
2009	58	26.79	2.40	10.00	5.40	6.44	2.55
2012	45	30.66	7.78	14.07	5.40	1.97	1.46
2014	41	35.59	9.45	14.38	6.35	2.71	2.71

2.7.5 E-Services

Examples of Saudi e-services are (Sahraoui et al., 2006; Al-Saif, 2010):

- E-Payment Gateway "SADAD": It was established in 2004 by the Saudi Arabian Monetary Agency SAMA to facilitate bill payment between governments to business, business to business, and government to citizen. Over 85% of Saudi's eight million bank account holders use 5,000 ATMs, bank branches, telephone banking, Internet banking, and 45,000 "Point-of-Sale" POS terminals countrywide.
- Smart Cards: It issues national ID cards using smart card technology. This system has a computer chip for storing personal identification information, thumbprints, medical and driving records and digital certificates. At present, the Ministry of the Interior is replacing the personal identity cards by smart cards. In a later stage, it would integrate the driving license and the family card into the smart card. Efforts are being made to introduce e-passports also.

Most studies about Saudi e-government criticized the quality and quantity of online services. Sahraoui et al. (2006) attributed the delay of Saudi's appropriate e-services to a lack of detailed e-government master plan and one clear vision. The added decrees and new restructuring have done little to give motivation to a serious transactional e-government presence. Also, the lack of clear ownership over the umbrella e-government project in Saudi Arabia, epitomized by its decentralized development, eventually becomes an impediment when attempts will be made to enact a one stop portal unless outsourcing is considered as an alternative. The authors acknowledged that the online success belongs to non-government corporations like, for example, Saudi STC. On the other hand SADAD, the payment gateway, is an icon of Saudi e-government success relying on an outsourcing company called "Sejel Technologies".

Moreover, Alfarraj et al. (2011) believed that Saudi online services are poor and lack quality. Based on their conducted online survey, 11 out of 28 Saudi government authorities are at the interactive stage which means these websites do not yet provide online services. Also, two ministries (Ministry of Hajj and General Presidency of Youth Welfare) still have no online presence. Al-Khalifa (2010) noted that most Saudi e-government services are in

their initial stages and not working together collaboratively. As well Al-Shehry (2008) said that no useful progress has been made in e-services within ministries. Examining the Saudi portal, AlZahrani (2011) determined that the Yesser program itself is still broadly at stage one of e-government model. The low quality of Internet services must be solved, otherwise, Al-Shehry warned, people will draw back from using any e-services in the future, and it will be very difficult to regain their trust in e-government.

Many researchers (Alfarraj et al., 2011; Al-Fakhri et al., 2008; Alshehri, & Drew, 2010) argue that Saudi websites are only information providers rather than service providers. Most Saudi government websites are inefficient because they just provide general information and often the data is not updated. While some offer better functionality, as online forms, it is hard to find a government website where you apply for a job, arrange an appointment, or renew a license (Alshehri, & Drew, 2010). Saudi ministries need to offer more e-services to adequately serve the citizens. Examples of needed e-services are car renewal or registration in educational programs. As the majority of Saudi ministry websites lack such important services, the Saudi government should pay close attention to the slow development of its government websites and take measures to remedy this situation (Alfarraj et al., 2011).

Another important issue is that online consumers refrain from using e-services because of their concerns about security and privacy. According to Al-Ghaith et al. (2010), the Saudi constitution does not provide for a right to privacy but the introduction of the "IT Criminal Law" in 2007 defines IT crimes and their punishments. However, the IT Criminal Law does not define the privacy right nor does it mention any punishments that would be applied to companies and websites owners who do not protect their visitors' privacy. Alshehri and Drew (2010) stressed that governments should provide a secure access to their e-services to develop citizen trust. More than 46.6% of the respondents to their study saw security and privacy as the third-ranked barrier to Saudi e-government adoption and diffusion. Participants in the study felt that transferring personal information to public agencies online is not safe yet, fearing e-services websites are not secure enough to protect their private data from being misused or distorted. A significant challenge for Yesser is to deal with security, privacy and trust issues in governmental websites.

2.7.6 Research on Saudi E-Government

Eidaroos et al. (2009) have focused on the usability of Saudi e-government through adopting a heuristic evaluation approach. Compiling a heuristic checklist, three experts evaluated two Saudi agencies' websites, the first established prior to Yesser and the second supported by Yesser, noting that both websites had achieved Digital Excellence Award offered by Saudi MCIT in 2007 and 2008. The authors classified Saudi e-government as in the early stages of development and noted that tasks in the transaction stages, such as data entry forms, showed considerable weakness. Hence, web developers and Yesser should focus on usability in order to improve the ranking of Saudi e-government. This study lacked the evaluation of a large number of websites and refrained from naming the evaluated websites for no given reasons. Also, it didn't discuss the usability problems found in the two evaluated websites.

The other study by Buragga (2010) evaluated two Saudi e-government websites, Saudi Post and Saudi Railway. 173 participants were provided with 12 features to be evaluated manually and the results were compared with the outcomes of automated evaluation tools, HERA, CynthiaSays and HTML Validator. The study found that the Saudi Railway website had more severity problems than the Saudi Post website and both failed in providing the e-services that are needed by Saudi citizens and residents. The author mixed the concept of usability with accessibility and actually evaluated the accessibility of the selected websites claiming that this will detect usability problems.

Al-Khalifa (2010) stressed the importance of usability evaluation for government websites. By 2009, only 137 government websites were listed on the Saudi e-government portal and 400 e-services. These e-services are in their initial stages and still not working together collaboratively. She developed a heuristic evaluation checklist covering 6 components (design and consistency, navigation, forms, search functionality, content precision and information privacy, and help and feedback). She evaluated the usability of 14 Saudi government websites founded on their delivery of key services to the citizens. Two experts conducted the heuristic evaluation and found that the evaluated websites did not fully meet any of the six components. The score for design and consistency was high at

80% followed by the content and information privacy component 75%. The 14 websites suffered from: bad search facilities, no FAQ or online help, high number of dead links (average 5.5 dead links/ homepage), slow page loading time (average loading time 1.87 seconds) and inaccessible web forms. The research omitted listing the developed heuristic guidelines, the names of tested websites and the differences between their performance, and the software tools that calculated the number of broken links and page loading time.

A Significant number of studies were interested in the Saudi e-government adoption such as Al-Ghaith et al. (2010). The most significant factors affecting e-government adoption in Saudi Arabia are: complexity (or easy to use), privacy, compatibility, and Internet quality respectively. The perceived ease of use is the number one factor affecting e-service adoption, which reflects the importance of usability on using e-services among Saudi citizens.

"The Saudi Arabian e-government is striding ahead of its European counterparts." This conclusion of a CISCO co-sponsored study with Saudi government attracted and motivated Sahraoui et al. (2006) to embark on a critical analysis of Saudi e-government. The 2005 CISCO study found that the ICT has helped Saudi government boost customer satisfaction rates by 44% and citizens' adoption of e-services by 34%. Sahraoui et al. (2006) reached a totally different result in their research of 2006. First, most e-government strategies are scanty and reactive at best and there is no serious attempt for transforming government through ICT. Focusing only on technology is driven by a bureaucratic culture that perceives citizens as neither customers of government, nor participants in decision making. Second, analyzing 25 Saudi government websites, the authors found that only 13 out of 22 ministries (60%) have an online presence. Additionally, the content and depth of Saudi ministries' pseudo-portals are poor. At 2006, none of the ministries' sites had online transacting; at best, the passport department in the interior ministry offered e-service inquiries. The website of the Ministry of Hajj (pilgrimage) is not yet developed. Sejel Technology, a consortium of local companies, was originated to oversee new infrastructure operations and to be responsible for the visa to pilgrims on behalf of the Ministry of Hajj. The government online presence in Saudi Arabia is between stages 2 and 3 of the UN e-government model, hence not yet fully transactional. Many challenges are facing the Saudi

e-government such as: a deeply bureaucratic culture, the absence of citizen participation, the lack of an objective evaluation framework and a management framework. Overall, Saudi Arabia has been the least receptive to the government wave in the region with the exception of Oman, yet media reports and conference abound about Saudi e-government achievements.

The lack of current research on Saudi e-government motivated Alfarraj et al. (2011) also to conduct a regional comparison between Saudi Arabia and Bahrain. As no other research was found after the study of Sahraoui et al. in 2006, Alfarraj et al. felt that Saudi e-ministries needed to be re-evaluated to note any differences in the four years (from 2006 to 2010). They evaluated the same Saudi e-ministries using the same UN stage model at that time. Evaluation results were combined in table 2.14. Regarding Saudi Arabia, two ministries have no e-presence, two others fall into Stage 2 as mere information providers, 11 ministries at interactive stage not providing e-services yet, 12 ministries in stage 4 as e-services providers, and one only, the Ministry of Higher Education, reaches stage 5 (seamless). Thus, some Saudi ministries have made progress in developing their government websites but the development is still slow. For Bahrain, the majority of ministries are in the highest stage of the UN e-government model.

Table 2.14: Number of Examined Heuristics in each Principle (Alfarraj et al., 2011)

Stage No.	Stage Reached	Assessment Elements	# Saudi ministries	# Bahrain ministries
	No presence	No official website available	2	0
1	Emerging	e.g. agency name, agency phone number, address, operating hours, general frequently asked questions	-	-
2	Enhanced	e.g. organizational news, publication, online policy (security, privacy)	2	0
3	Interactive	e.g. officials' e-mail addresses, ability to post comments online, simple two-way communication, can download the organization's forms	11	6
4	Transactional	e.g. e-form, e-payment and query services	12	8
5	Seamless	Full integration across the organization	1	9

A cross-country comparative analysis of national e-government was conducted by Chatfield and Alhujran (2009) on 16 Arab countries. The sample e-government websites were evaluated by two experts using a four-stage e-government model. Based on their e-services, the 16 e-governments are clustered into one of three groups: Arab e-government

leaders, Arab e-government up-and-comers, or Arab e-government laggards. Table 2.15 displays the evaluation results of the six leading developed countries plus the 16 Arab countries. Accordingly, ten Arab countries are at the information stage (Lebanon, Saudi Arabia, Oman, Tunisia, Algeria, Syria, Morocco, Sudan, Yemen, Iraq), six provide 2-way interaction, only three (UAE, Bahrain, Qatar) offer online payment transaction, and seven have some sort of e-democracy. The results confirm that there is a wide digital divide between Arab countries and leading developed countries and even among the Arab countries themselves.

This study is a good effort to fill the gap of research on e-government in the Arab world. However, in sourcing the sample countries it depends on old data, the UN 2005 report. More important, the e-participation index has received criticism for its superficial and non-qualitative evaluation of e-government websites but the authors didn't mention that. According to Grönlund (2011) and Siskos et al. (2014) any country, no matter how undemocratic, can score high on e-participation. The connection of e-participation to democracy is not verified by the UN index which classifies highly undemocratic countries for just exhibiting some web features that cannot be used in practice. It is worrying to see many authoritarian countries score as well as the top full democracies countries in the EIU (Economist Intelligence Unit) ranking.

Table 2.15: E-Government Evaluation of the Samples (Chatfield, & Alhujran, 2009)

	One-Way Information Flows	Two-Way Interaction	Payment Transaction	E-Democracy
World E-Government Leaders (top six)				
United States	✓	✓	✓	✓
Denmark	✓	✓	✓	✓
Sweden	✓	✓	✓	✓
UK	✓	✓	✓	✓
South Korea	✓	✓	✓	✓
Australia	✓	✓	✓	✓
Arab E-Government Leaders				
UAE	✓	✓	✓	✓
Bahrain	✓	✓	✓	✓
Qatar	✓	✓	✓	✓
Arab E-Government Up-and-Comers				
Jordan	✓	✓		✓
Lebanon	✓			✓
Kuwait	✓	✓		✓
Egypt	✓	✓		✓
Saudi Arabia	✓			
Oman	✓			
Tunisia	✓			
Algeria	✓			
Morocco	✓			
Syria	✓			
Arab E-Government Laggards				
Sudan	✓			
Yemen	✓			
Iraq	✓			

Saudi e-services were examined by Al-Fakhri et al. (2008). First, at the national level, Saudi e-government is compared with the United Arab Emirates and the United States, then, at the local level, the Riyadh e-portal was compared to the Dubai e-portal. The data was collected from government websites' content features and questionnaires sent to Saudi government employees randomly. In 2005, the UAE had one of the most remarkable year-over-year gains in e-government worldwide. Its ranking was up from 60th place in 2004 to 42nd in 2005, while KSA had improved from 90th place in 2004 to 80th in 2005. The UAE was the first Arab government to launch an e-government portal offering many e-services and access to two portals, e-dirham for transactions and the e-forms portal. The e-dirham and the e-service are among the best world practice models in the report of United Nations (2005). On the other hand, the United States was the world leader in the UN e-government rankings from 2003 to 2005 and the second in Brown University ranking 2005. Launched in 2000, FirstGov is the official US government portal that provides information and services. FAQs contain a response timeframe for submitted questions about the USA government to be answered within two business days, which is comparable to the private sector. Also, the U.S. has a consultation portal as public comment on federal regulations and the Department of Education offers the "Teachers Ask the Secretary" section. On the contrary, Saudi e-government portal provides only general information and simple services.

Comparing Riyadh and Dubai e-government portals, Al-Fakhri et al. find the Riyadh site, established in 2002, provides only information about Saudi e-government and some services. On the other hand, Dubai established its portal in 2001 and has an outstanding progress over a short period of time to become one of the most advanced e-cities. Dubai offers e-services via one stop site through different channels. Its portal has 6 sections: citizens, residents, visitors, local business, foreign companies and investment in Dubai, and its strength lies in its ease of use. The visitors' section provides updated information on city activities, hotels, entertainment, etc. The citizens' and residents' sections allow people to pay fines, apply for a job and renew driving licenses. Dubai citizens make online payments for public services through a secure ePay gateway. Additionally, the questionnaire survey revealed that 40% of the participants believe there is a Saudi e-government portal, while 33% of them do not know whether it exists or not. The Saudis could consider several reforms such as increase the awareness of its e-government program among the public,

make Internet access more available to society, equip public facilities for Internet usage, develop a legal framework for e-transactions, and foster 2-way communication between government agencies and between the government and the public.

At the local level, Al-Nuaim (2009) evaluated the municipality websites of Arab capitals. Most Arab cities were absent from Holzer and Kim's 2005 study worldwide on e-municipalities and only 5 e-cities were ranked: Cairo-45, Dubai-50, Riyadh-52, Amman-65, and Beirut-66, out of 81 assessed websites. She selected for evaluation 6 Arab capitals' websites: Amman, Beirut, Muscat, Riyadh, Doha, and Kuwait, based on having the highest population and receiving the most e-government funding. An official government website for Cairo was not found. Thus it was excluded from the study. The website assessed by Holzer and Kim's 2005 report was actually a portal for Egypt and not Cairo. Also, she modified the Municipality Evaluation Checklist of the Maxwell School by adding 3 items: news, site map and the ability to find a site by guessing its URL. Using the checklist form, the researcher and five Internet experts assessed the availability and functionality of each item for the six Arab e-municipalities.

Results of this study show that Riyadh and Amman e-cities have an acceptable score (16.306, 14.93) reflecting they only have 74% and 67% of what should be available. Kuwait, Beirut, Doha and Muscat received a low score of less than 14 so they lack the basic requirements for a municipal website. The organization of all tested websites was ad hoc and lacks a good link structure, causing navigation difficulties. They were not even citizen centered websites, except Amman, and do not offer sections for businesses, residents, and visitors. The goal of these websites and the intended target users were not clear. Most are there just to have a web presence with general city information that does not affect the daily lives of citizens. Available links were not relevant to citizen needs or extra information (weather, currency). The tested websites were not updated regularly, have limited e-services and missed contact information.

2.7.6.1 The State of Research on Saudi E-Government

In conclusion, the state of research on Saudi e-government is as follows:

- Lack of research on Saudi e-government is apparent since few studies are found in the literature.

- The problem of co-sponsored study, with the Saudi e-government program, such as the CISCO's study that reached an unrealistic conclusion: "The Saudi Arabian e-government is striding ahead of its European counterparts." (Sahraoui et al., 2006). CISCO is unlikely to publicly criticize a client and probably was angling for a government contract. The problem of studies by private companies has been highlighted in section 2.6.1.5.
- The problem of some undeserving awards given to governments for their websites (Sahraoui et al., 2006; Eidaroos et al., 2009) who found two Saudi e-government websites that achieved Digital Excellence Award, offered by MCIT in two consecutive years, to be in an early stage of e-government development.
- Most studies are at national level evaluating Saudi ministries' website. Only a few studies assessed local Saudi e-municipality, Riyadh website
- Most studies evaluated Saudi e-government from the e-government dimension only using the UN e-government model.
- Only a handful of publications evaluated Saudi e-government in terms of usability.
- Most of the studies classified Saudi e-government as in the early stages of e-government development. Nevertheless, some ministries have made progress in developing their government websites but the development is still slow. However, every two years, Saudi e-government rank is improving dramatically in the UN benchmark. We agree with other scholars (Codagnone et al., 2015; Siskos et al., 2014; De Róiste, 2013; Grönlund, 2011; Andersen et al., 2011; Rorissa et al., 2011; Janssen, 2010; Montserrat, 2010; Bannister, 2007) who considered that e-government ranking of nations is meaningless and questioned the efficiency of e-government benchmarks.
- A significant number of studies were interested in the topic of e-government adoption in Saudi Arabia and the challenges facing its progress. Some mentioned the importance of usability in encouraging the Saudi people to adopt e-government.
- Most usability e-government studies used a heuristic approach with limited evaluation criteria that may discover only a few minor users' problems.
- To our knowledge, no study used a real user testing method and no study used automatic tools for usability.
- Some of the researchers evaluate e-government websites by themselves with no criteria given.

Chapter 3: Methodology

This chapter explains the methodology employed to achieve the objectives of this research which is to develop an evaluation framework for a city website to enhance citizen-centered e-government. To accomplish this, it also proposes how to select appropriate website evaluation methods. Therefore, this chapter includes a discussion of the methods used in this dissertation and the reasons for their inclusion, along with other approaches that were not considered and the causes for their exclusion.

3.1 Research Philosophy/ Paradigm

The first step in research "Paradigm" is defined as the theoretical framework, consisting of theories, methods and ways of defining data, which influences how knowledge is studied and interpreted (Mackenzie, & Knipe, 2006; Hasan, 2009). Bhattacherjee (2012) believes that research is shaped by our mental models to organize our reasoning and observations. Paradigms govern how we view the world and how we structure thoughts about what to see in that world.

There are varied claims about how many research philosophies or paradigms (see Figure 3.1 for overall existed research methods). However, the two main research paradigms are positivism and interpretivism (Saunders et al., 2012; Hasan, 2009). Each paradigm is suitable for a different kind of study and has different propositions and assumptions regarding the process of research. Bhattacherjee (2012) suggested the way researchers study social phenomena is formed by two philosophical assumptions: ontology is our assumptions about how we see the world, and epistemology is our assumptions about the best way to study the world and obtain knowledge, e.g., should we use an objective or subjective approach. The logic of a research describes the relationship between social research and theory, which could be deductive or inductive.

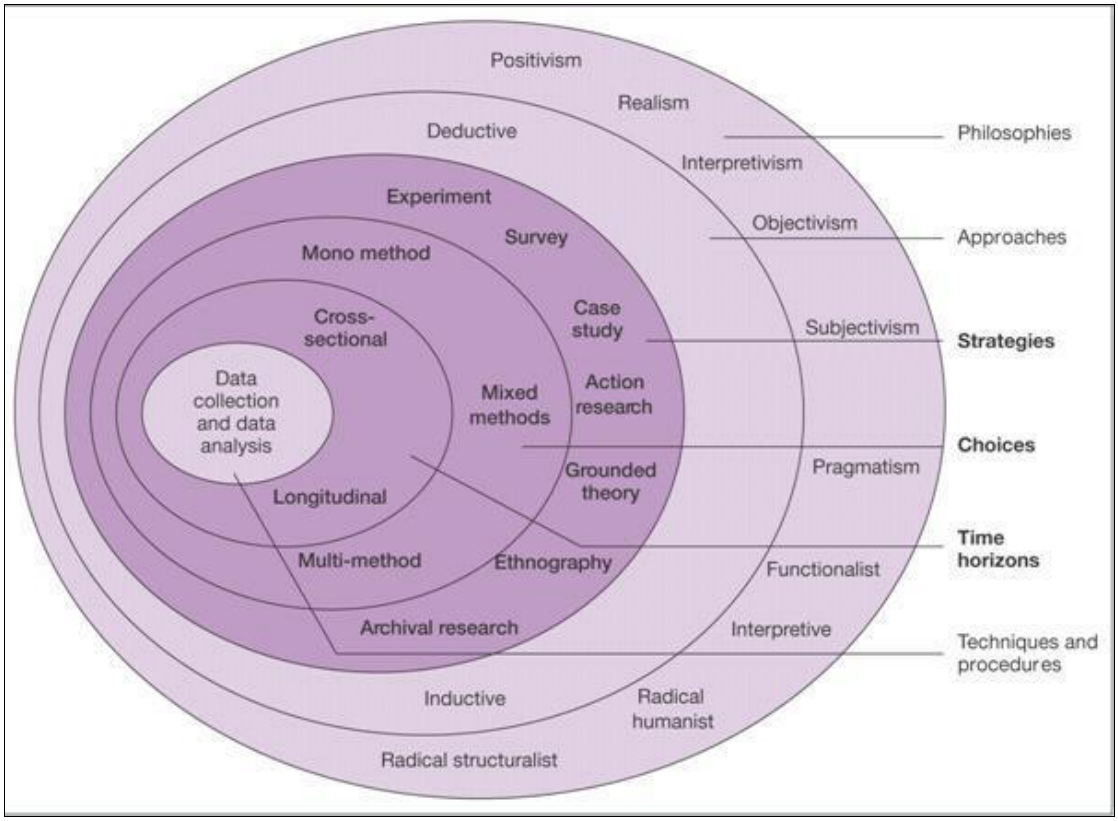

Figure 3.1: Metaphor of Research Onion (Saunders et al., 2012**)**

1. Positivist Paradigm

Positivism or a scientific method of research involves an inquiry process to understand social or human problems with an aim to test a theory or refine previous ones (Bahareh, 2015; Hasan, 2009). The variables in theory are measured by numbers and analyzed in order to decide whether or not to generalize the theory (Hasan, 2009). Thus, knowledge is acquired by observing and measuring the phenomena using the developed numeric measures. Positivism employs an objective approach to test theories by a survey or instrument and is independent of or external to the researcher (Bahareh, 2015; Hasan, 2009). Also, it employs a deductive approach to research, starting with a theory and testing the theoretical hypothesis using empirical data (Bhattacherjee, 2012).

2. Interpretivist/Constructivist Paradigm

This philosophy concentrates on the subjective meanings and interpretations of a social action, and not on the measurement of that phenomenon (Hasan, 2009). In contrast to positivism which aimed at theory testing, interpretive paradigms are typically aimed at

theory building. It employs an inductive approach that starts with observing data to derive a theory about the studied phenomenon (Bhattacherjee, 2012**).** Interpretivists use subjective data collection tools such as observation or interviews and is dependent on the researcher.

3.2 Research Approach

1. Data Type

The data type collected for a research can be quantitative, qualitative or mixed methods. Quantitative data involves numeric scores, metrics and so on, while qualitative data includes interviews, observations, etc., and is not in the form of numbers. The positivist research uses predominantly quantitative data but may utilize qualitative data, while interpretive research relies mostly on qualitative data but can benefit from quantitative data as well (Bhattacherjee, 2012).

Data type is the way one chooses to treat and analyze data according to the objective of the research. The researcher decides what collecting data method is to address their research questions. As summarized in Table 3.1, such methods may include quantitative tools, e.g. experiments or survey, or qualitative tools, such as interview or case study, or a combination of both (Bhattacherjee, 2012). Qualitative and quantitative approaches are not rigid. A study can be more qualitative than quantitative or vice versa, or mixed method in the middle of this continuum (Bahareh, 2015).

Table 3.1: Paradigms and Methods (Mackenzie, & Knipe, 2006; Bhattacherjee, 2012)

Philosophy/Paradigm	Data Type	Data Collection Methods
Positivist	Quantitative approach is dominant, numeric data, objective measure	Surveys Tests Scales Experiments
Interpretivist /Constructivist	Qualitative approach is dominant, non- numeric data, subjective measure	Interviews Observations Case study Document reviews

2. Case Study

A case study is an empirical inquiry that investigates phenomena within its real-life context, and it can be used in a positivist research for theory testing or in an interpretive

research for theory building (Bahareh, 2015). A case study has several strengths over other research methods such as experiments and survey because it can capture a richer array of data than most other research methods (Bhattacherjee, 2012**)**. The studied phenomenon can be viewed from the perspectives of several participants and may use multiple levels of analysis. There are four types of case study design: single- or multiple- case studies, and single- or multiple- unit of analysis within a case. The multiple-case design is called a comparative design, and each case is a single experiment. It involves studying two or more case studies and comparing them based on the belief that a better understanding of phenomena can be achieved by comparing them with regard to other contrasting cases (Hasan, 2009).

3.3 Selection of this Research Philosophy and Approach

The choice of design should depend on the nature of the research phenomenon being studied and the objectives of this research mentioned in chapter 1. Bhattacherjee (2012) stated that a positivist design is an appropriate paradigm if different theories existed and the researcher aims to test or integrate them and these theories consist of variables measured by numeric metrics. Hence, it is clear this study is a positivist quantitative research. Even though there are very few collected qualitative aspects, such as experts' comments and user satisfaction survey, the data type of this study is mainly quantitative. Bhattacherjee (2012) suggested that even if the researcher intended to collect quantitative data, in a questionnaire for example, he should also try to collect some qualitative data for better results. In addition to adopting a positivist quantitative research, this study will empirically evaluate a multiple-case study of one unit of analysis (five e-city websites) for comparative design. Bahareh (2015) mentioned that research designs such as multiple case studies have higher degrees of internal and external validities. Thus the multiple case study design is chosen for three reasons: it is more appropriate for theory testing, for establishing generalizability, for enhancing external validity, and for developing richer interpretations of the studied phenomenon (Bhattacherjee, 2012), all of which we aimed in this research.

3.4 Research Design

The research design is an action master plan that guides the research according to the chosen philosophy. The design of this research is divided into three main steps: framework

development, selecting case studies (Saudi e-city websites), and usability testing using web evaluation methods.

3.4.1 Development of E-City Framework

Upon reviewing the literature in e-government, it is clear that there is still no consensus on how to measure e-government website quality nor the metrics needed for such an evaluation. Although, it may be possible to measure specific features on the website, a measurement of the entire website's quality is conceptually and practically improbable. This also holds true for local e-government or city government websites. A good evaluation model for local e-government still needs to be developed. Surveys on Europe show that between 50% and 80% of the citizens' interaction with their government occurs at local e-government level (Moraru, 2010), but unfortunately, current literature provides little web development guidance to e-city websites (Lofstedt, 2012). Therefore, the main objective of this research is to propose an evaluation framework to assess the quality of city websites.

Based on the discussion of limitations of some well-known e-government models in the literature review (sections 2.6.1.5, 2.6.2.6 and 2.6.2.7), the models excluded from this research proposed e-city model are:

1. Benchmarking models: the efficiency of e-government benchmarking models are in doubt by many authors (Codagnone et al., 2015; Siskos et al., 2014; Grönlund, 2011; Rorissa et al., 2011; Janssen, 2010).
2. Stage models: several studies (Klievink et al., 2009; Poeppelbuss et al., 2011; Debri, & Bannister, 2015; Bannister, & Connolly, 2015) has criticized the "stagiest" approach as being theoretically weak and has no empirical foundations.
3. One-dimension e-government model: most of the proposed studies emphasize limited aspect of e-government; some focus on website features only, such as West's model, others highlight e-services, such as MeGAP; besides that the quality of e-services is absent from the measurement of most e-government models (Song, 2010).
4. Country specific e-government models: several studies (e.g. Flak et al., 2005; Shareef et al., 2012) has criticized country-specific models. Flak et al. (2005) have tested the USA MeGAP in Norway context and found numerous services specific to the USA but not within the responsibility of the Norwegian municipalities and vice versa. Shareef et al.

(2012) revealed that the UK e-government stage model cannot be adopted for use in the Kurdistan Region of Iraq due to various factors relating to ICT infrastructure, e-readiness, legal framework, cultural, education, and political process.

5. E-service-Quality Framework: it is immature scale, not validated yet, and not capturing fully all the dimensions of service quality in e-government platforms (more in section 2.6.1.4).

Since there is no framework that measures the quality of e-municipality websites and e-services, this research proposes the possibility of integrating 3-dimensional criteria: 1) website quality, 2) e-services quality and 3) the number and type of e-services when testing e-city websites. The 3-dimension proposed e-city framework builds upon the strengths of ten models, three theoretical and seven practical ones. The three theoretical models (Calero et al., 2005; Treiblmaier, & Pinterits, 2010; Hasan, & Abuelrub, 2011) verify that usability is the most important web quality metric. The seven models are: the US Research-based Web Design and Usability Guidelines (HHS, 2016), the UN Digital Governance in Municipalities Worldwide (Holzer et al., 2009), the Community Benchmarks Program(CBP) (Denfeld et al., 2002) and Dubai E-government Excellence Model (Dubai eGovernment Department, 2011) for website quality; the Bahrain User Interface Standards (Bahrain eGovernment Authority, 2010) and Dubai eService Excellence Model (Dubai eGovernment Department, 2009) for e-services quality; and the Gartner model (Montserrat, 2010) for classifying the type of e-services.

Karkin and Janssen (2014), Rinder (2012), and Dingli and Mifsud (2011) have praised the official American HHS guidelines for being validated empirically and supported in the HCI field. Also, HHS guidelines and Gartner are demand-side (citizen-centered) models (Rorissa et al., 2011). The UN Digital Governance in Municipalities is the single most referenced e-municipality guideline and the only international ranking of e-cities. Dubai's and Bahrain's models are built on "Customer-Focus" concept and they usually are high ranked in the UN e-city report (Dubai eGovernment Department, 2011). More details about the reasons behind choosing these 7 guidelines and some shortcomings are in next chapter.

To select the metrics for assessing website quality, web criteria analysis is conducted as follows: 1) the common e-government heuristics, agreed upon by two, three or four

guidelines, were chosen; 2) then the heuristic was selected if it fulfills one of the design principles stated on the g-quality inspection method for e-government proposed by Garcia et al. (2005). The g-quality inspection method, an extension of Nielsen's heuristic, was chosen because: a) it was developed to evaluate usability specifically in e-government websites which is necessary to fulfill these websites' intended goal and coincide with this researcher's objective, b) it was validated by some research such as Granizo et al. (2011); 3) the selected heuristics were checked against a usability framework proposed by Folmer et al. (2003) just to identify affected quality attributes; 4) the g-quality fulfilled heuristics and affected usability attributes produced the "E-City Usability Guidelines". A similar procedure is followed to obtain "E-Services Quality Guidelines" from the Dubai eService Model and Bahrain's User Interface Standards. Combining both guidelines create the first draft of "E-City Usability Guidelines" which after refinement produces the "E-City Usability Framework". The refinement steps and scoring method of the developed framework are described in the next chapter.

3.4.2 Selecting Case Studies

The selection of Saudi e-city websites was based on two criteria: the number of Internet users and the population size of the city. Regions with the highest Internet usage were selected first then the largest populated city in each region was chosen. The rationale behind the relation between high Internet usage of a large city and local e-government capacity is supported by studies such as Holzer et al. (2014).

Based on the Saudi National e-Government Portal (2013), there are 13 Saudi provinces and a total of 16 municipalities. Unfortunately, no statistics was found about Internet penetration in Saudi regions, therefore the researcher sent emails to different ministers and government centers, such as MCIT, CITC, and National Contact Center, but they did not respond. Consequently, Internet penetration in regions was based on the only available Saudi Communication and Information Technology Commission report of 2008 (CITC, 2008). Accordingly, the five selected Saudi municipal websites tested were: Jeddah, Riyadh, Al-Madinah, Eastern Region, and Qassim. For more details see section 5.2.1.

3.4.3 Testing Using Web Evaluation Methods

Reviewing literature identified another gap in knowledge related to this research's second objective, which is to test the developed framework on selected e-city websites. Unfortunately, the literature lacks research that classify, compare, and determine web evaluation methods. A detailed comparison of web evaluation methods is in section 2.5.5. Accordingly, two methods were not used in this research: 1) Link analysis methods: because it is not validated and still in the process of development, and 2) Google Analytics: since it requires inserting codes into each tracked webpage and this cannot be done in government entities.

Often web experts suggest using more than one evaluation method since each one alone isn't free of shortcomings. The recommendation by many researchers in the field is to conduct heuristic evaluation and user testing as a mainstream approach, while other web evaluation methods might be the first insight into the status of a website (Joe et al., 2015; Huang, & Benyoucef, 2014; Ølnes, 2013; Krenk, & McComb, 2012).

The use of multiple data collection methods in studying the same phenomenon for the purpose of increasing study credibility is called triangulation (Hussein, 2015). Thus, triangulation web evaluation methods were used in this research: heuristics evaluation, user testing, Alexa web analytics tool, and automatic website evaluation (the tool used was broken link checkers). The selection of the four usability evaluation methods stemmed from the aim of this research to develop an evaluation framework that identifies comprehensive usability problems from different perspectives. Heuristic evaluation and user testing are well-established website evaluation methods and they complement each other. But for the automatic link checkers and Alexa, as far as could be established, few studies use these tools and none tries to test and validate their effectiveness. Therefore, the purpose of including link checkers and Alexa is to fill the gap in research and to test how reliable are these tools in assessing e-city websites. In addition, broken links on e-government websites is a serious problem and citizens facing this error are likely to leave the website, ending up with a low citizen adoption of e-government. Therefore, it is important to conduct more research on the subject of validating automatic link checkers.

3.5 Data Collection Methods

As illustrated in Figure 3.2, heuristic evaluation and user testing checked each Saudi e-city website against the developed E-City Usability Framework that was further divided into three forms: website quality objective guidelines (Table 3.2), website quality subjective guidelines (Table 3.3), and e-service Quality subjective guidelines with its designated tasks (Table 3.4 with one example task). The objective guidelines were evaluated directly by experts while the subjective guidelines with designated tasks, whether for a website or e-service quality, were assessed by users. Using both methods allows the evaluation of user interfaces with experts and users. Also, Alexa was selected as a web analytic tool because it covers a wide area and is the best data source based on the studies of Bhat (2013) and Jowkar and Didegah (2010). Nine Alexa metrics would be collected such as traffic ranks, the speed of download, and time on site. For link checkers, seven tools were chosen, as described in the testing chapter.

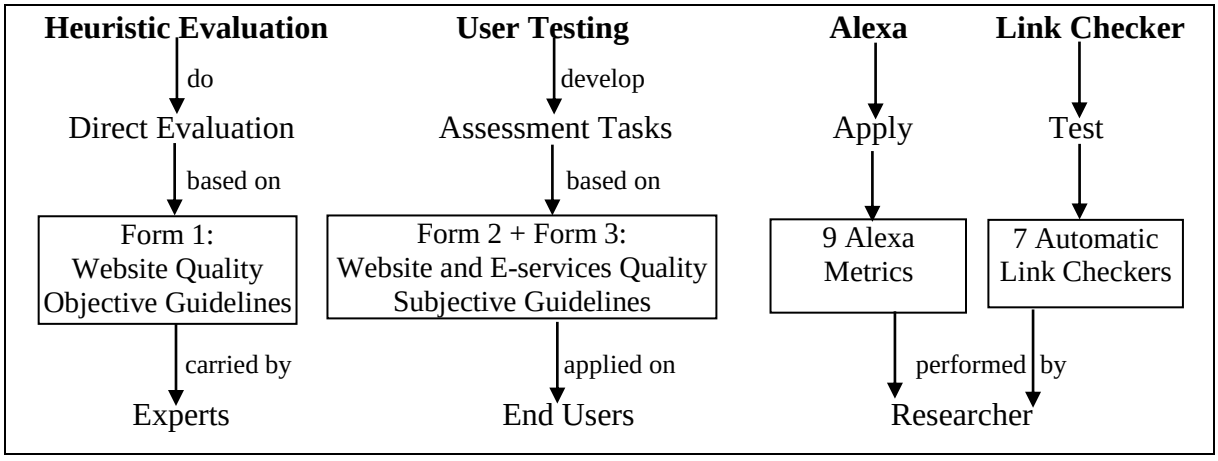

Figure 3.2: **Triangulation Web Evaluation Methods to Test E-city Websites**

Table 3.2: Website Quality Objective Guidelines Form for Heuristic Evaluation

	Guidelines	Suggested Score
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	0___.5___1
Comments:		
2	About us: mayor corner, mission, objectives of the website	0___.5___1
Comments:		

	Guidelines	Suggested Score
3	Municipal budget information	0___.5___1
Comments:		
4	City council meetings (dates, locations, agendas, minutes)	0___.5___1
Comments:		
5	FAQ with facility to ask new questions	0___.5___1
Comments:		
6	eJob: job vacancy at municipality	0___.5___1
Comments:		
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	0___.5___1
Comments:		
8	Last update date on the footer of every page	0___.5___1
Comments:		
9	Emergency alerts (road closedown, weather alerts…)	0___.5___1
Comments:		
10	Comment or eSuggest on the website	0___.5___1
Comments:		
11	Citizen satisfaction survey	0___.5___1
Comments:		
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	0___.5___1
Comments:		
13	Multilingual equivalent websites with a link on header of page	0___.5___1
Comments:		
14	The city website among top 10 hits (results) of Google and Yahoo search engines	0___.5___1
Comments:		
15	Links to related government websites open in a new window	0___.5___1
Comments:		
16	Downloadable documents/forms with appropriate access	0___.5___1
Comments:		
17	Design for common browsers access (Explorer, Chrome)	0___.5___1
Comments:		
18	Print pages properly	0___.5___1
Comments:		
19	Well-designed customer focused homepage: • Quick access to highlighted services through main menu	0___.5___1

	Guidelines	Suggested Score
	• Targeted audience group (citizens, business, tourists...)	0___.5___1
Comments:		
20	Clear entity "Logo" on every page	0___.5___1
Comments:		
21	Link to homepage from every page through "Home" or logo	0___.5___1
Comments:		
22	Sitemap	0___.5___1
Comments:		
23	Short and descriptive page titles	0___.5___1
Comments:		
24	Readable pages (font, color, background)	0___.5___1
Comments:		
25	Simple page with reasonable length of not more than 2 screens	0___.5___1
Comments:		
26	Privacy and security statement/policy	0___.5___1
Comments:		
27	All links working properly, i.e. no broken links	0___.5___1
Comments:		
28	Navigational options: • Indicator of a user is where on the site (e.g. Breadcrumbs)	0___.5___1
	• Enabled "Back button"	0___.5___1
Comments:		
29	Different colors for visited/unvisited links, underline links and no misleading cues to click	0___.5___1
Comments:		
30	Consistent design of all web pages: • Page layout (same feel and look, font, color, buttons, menus)	0___.5___1
	• Navigation and link style	0___.5___1
Comments:		

Table 3.3: Website Quality Subjective Guidelines Form

	Guidelines	Suggested Score
1	Useful and most up to date content	0___.25___.5___.75___1
Comments:		
2	eComplaint and time to resolve it	0___.25___.5___.75___1
Comments:		
3	Interactive city map (location of near-by services, transport, restaurant, hospitals, shopping)	0___.25___.5___.75___1
Comments:		
4	Meaningful images and video that don't slow downloads	0___.25___.5___.75___1
Comments:		
5	Effective search on the Header	0___.25___.5___.75___1
Comments:		
6	Logically organized short meaningful link labels	0___.25___.5___.75___1
Comments:		

Table 3.4: E-service Quality Guidelines Form with One Example Task

NA= Not Applicable

Task 1: Fill in Customer Satisfaction Survey	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1 or NA
2. Ease of navigation through the e-service process	0___.25___.5___.75___1 or NA
3. Simple forms with required fields and clear error message for invalid or incomplete data entry	0___.25___.5___.75___1 or NA
4. Online tracking for forms and e-services being processed	0___.25___.5___.75___1 or NA
5.Feedback when users waiting or submitting a request	0___.25___.5___.75___1 or NA
6. E-services completely online if possible	0___.25___.5___.75___1 or NA
7. For e-payments: • Availability of various e-payment methods (VISA, etc.)	0___.25___.5___.75___1 or NA
• Notification of e-payment via SMS or email	0___.25___.5___.75___1 or NA
Comments:	

Addressing the goal of this research to be citizen centric, the three-dimensional proposed framework evaluated the selected websites from the viewpoint of citizens who are concerned with high quality websites and a large number of high quality e-services that satisfied their needs. The e-services were evaluated on their impacts as seen from the normative view on citizens as customers. Employing heavily user testing for part of website quality and for all e-service assessment indicate that citizen's voice is the most important measure of e-government success.

The think-aloud protocol is one of the techniques that used in user testing. It refers to the user verbalizing their thoughts as they performing some tasks on a tested website. User testing sessions for this research were conducted at a room equipped with a Lenovo laptop and Internet connection. The researcher observed and took notes but, respecting users' desire, didn't record the test.

Task based approach is one of the most common methods for evaluating website usability and has been used in different studies (Rinder, 2012; Huang, & Benyoucef, 2014; Hasan et al., 2012). Mainly quantitative and some qualitative data are gathered through a task based approach. The principle behind this approach is that a variety of tasks are identified based on scenarios the users might experience when using the city website, thus a scenario task emulates real-world context. In line with Bahareh's approach (2015), the tasks in this research were extracted based on three criteria: content of the websites, similarities between websites, and most important the usability factors. Thus, a task analysis was performed and similar e-services between websites were identified, then the degree of importance of tasks was determined by connecting them to the guideline(s) of the developed framework. It worth mentioning that in formulating tasks we are not using a benchmark approach since most of the time we expect to find different e-services in Saudi e-city websites and also among different countries.

According to Bahareh (2015), a task based approach tries to use the same tasks on all websites. In case a task cannot be found across websites, an alternative task is developed for that websites. As much as possible this should result in similar tasks in all websites. Applying this to the five case studies of Saudi e-city websites, we found that:

1) For the heuristic test: two experts need to evaluate the same form (website quality objective guidelines) for all the selected websites.
2) For user testing: there are three parts:

    a) Users need to perform the same tasks on the website quality subjective form;

    b) Users need to fill in the same user satisfaction questionnaire;

    c) E-services were not the same but similar in nature and interaction for the five Saudi e-city websites and that was expected since there is no consensus for what e-services to offer online. As an example, we found only four sites offer inquiry about citizen's transaction and three sites offer inquiry about Saudis mortality in a certain period of time. However, since most of the offered services were just simple inquiry, they were, to a large extent, similar in their type of e-services and representative for that city.

In that sense: 1) we were not using a benchmarking approach which would ask the same questions in each case study because it is unrealistic, 2) we did not specify the names of e-services since we learned from the limitations of country-specific models, 3) we checked carefully what is available on the website and tested the same e-services if found, 4) if not, we look for similar type of e-services, and 5) most important we correlate scenario tasks with the proposed guidelines to ensure validity.

Finally, the methodology of this research is summarized in the next four figures covering three main subjects: 1) Development of an e-city framework, 2) Refinement of the developed framework and 3) Usability testing and web evaluation methods. Figure 3.3 (with its related diagrams a, b, and c) summarizes the entire methodology i.e. the process for constructing and testing an E-City Usability Framework. Figure 3.3a shows the method of web criteria analysis to select web metrics for the proposed framework based on web models and existing e-government models. Figure 3.3b identifies the sources of e-service quality frameworks. Figure 3.3c shows the protocol for the usability testing of selected Saudi e-city websites.

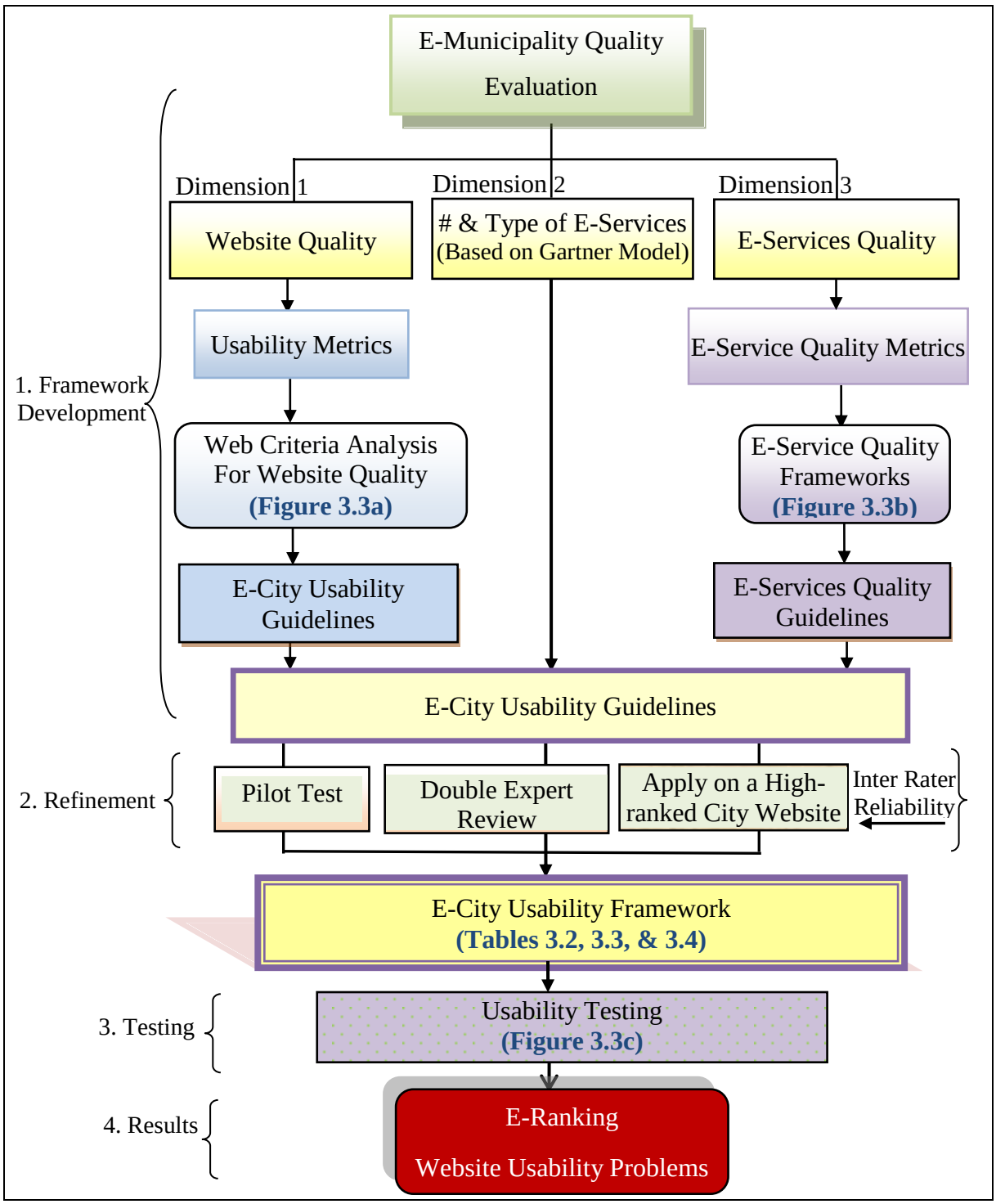

Figure 3.3: The Process for Constructing and Testing 3-Dimension City Usability Framework

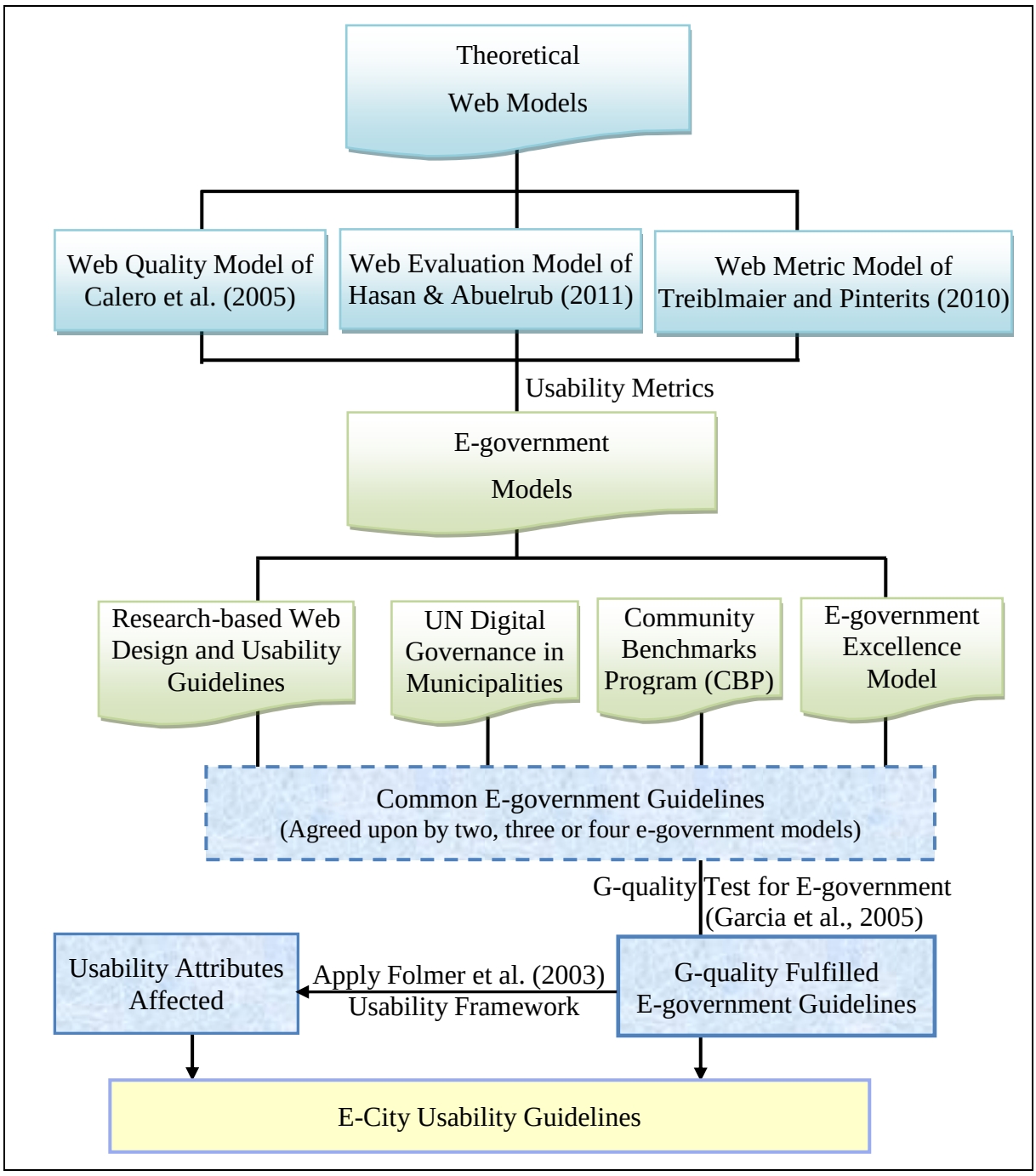

Figure 3.3a: Web Criteria Analysis for Website Quality

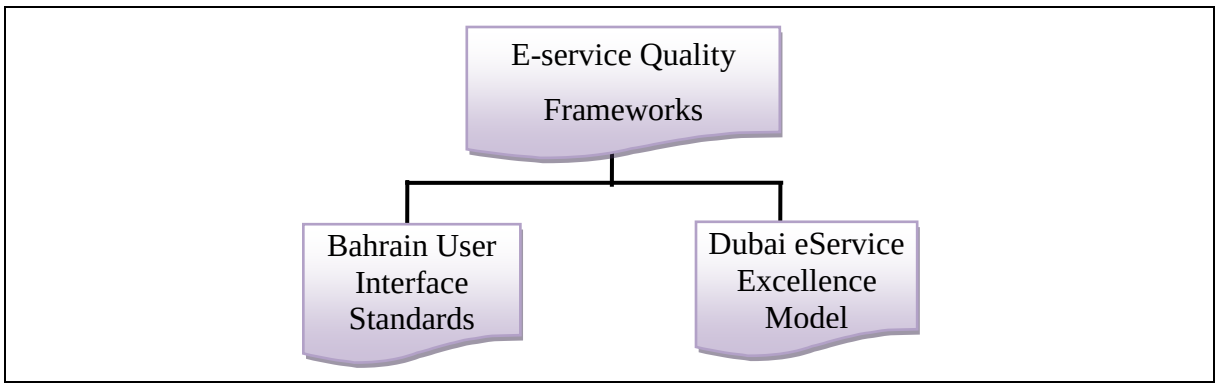

Figure 3.3b: E-Service Quality Frameworks

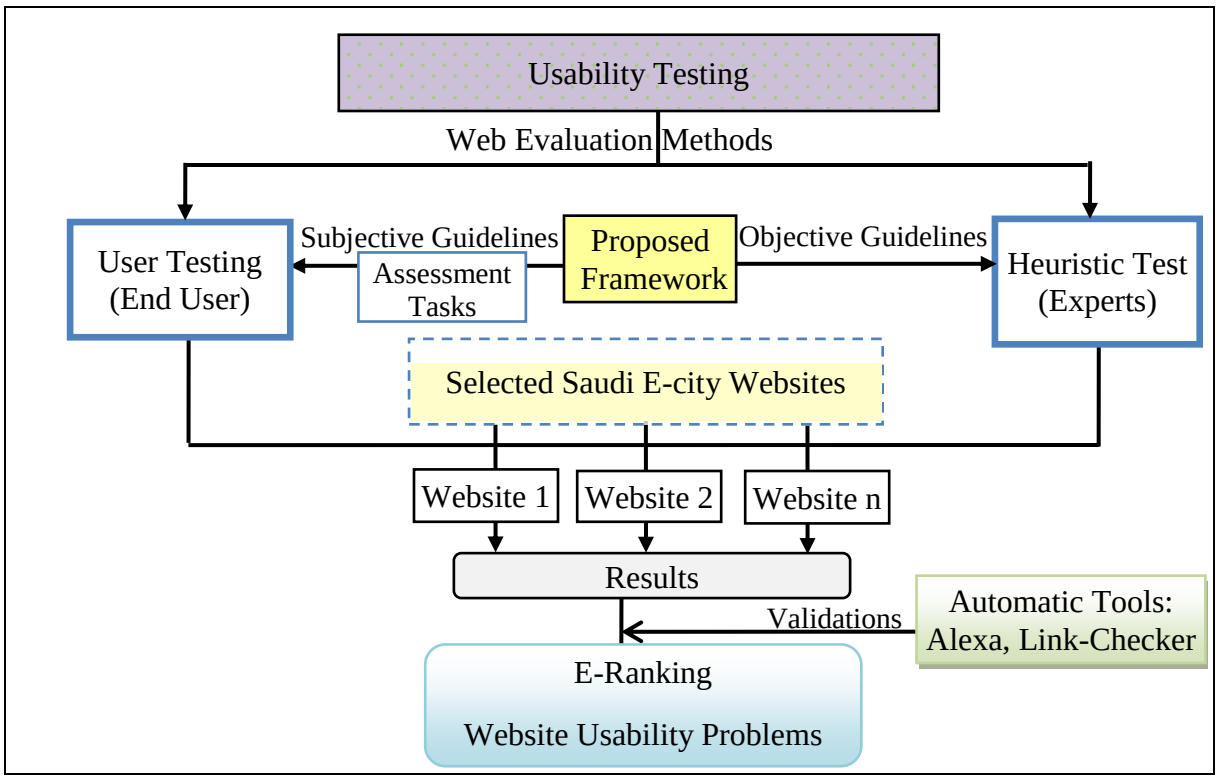

Figure 3.3c: Usability Testing

Chapter 4: Development of an E-City Usability Framework

This chapter discusses in more details how the proposed framework has been developed and what is the method for selecting the metrics for assessing website quality. Then three refinement steps to produce the final form of the E-city Usability Framework are explained. The inter-rater reliability and the scoring method are mentioned at the end of the chapter.

4.1 Development of a Measurement Framework

This research developed an E-City Usability Framework based on theoretical web metrics models and practical e-government models. The theoretical web models, discussed earlier in section 2.4.3, include three models: the web quality model by Calero et al. (2005), the web metrics model by Treiblmaier and Pinterits (2010), and the web evaluation model by Hasan and Abuelrub (2011). These theoretical models also served as a basis for determining the scope of measuring and defining web quality. The proposed framework includes metrics from practical well-known e-government models extracted from four main resource categories in the literature: 1) Government publications: guidelines for assessing e-government and e-municipalities from publications of the United States, Bahrain and Dubai; 2) International organizations: the United Nations' Digital Governance in Municipalities Worldwide survey; 3) Academic research: Gartner framework as a classification of e-government services; 4) Academic institutions: the 2002 Community Benchmarks Program (CBP) of the Maxwell School at Syracuse University on Onondaga County e-municipalities.

4.1.1 Three-Dimensional Assessment E-City Framework

According to Garcia et al. (2005), e-government websites can be divided into three types: informative, services and participative sites. That is, each e-government site presents a configuration of these three constitutive characteristics: information, services, and citizen participation. For example, an e-city website can be totally informative if it is 100% informative, 0% service and 0% participatory. The percentage of information, services, and

participatory government processes to be migrated to the web is an important factor that will indicate the government's strategy and migration maturity. This research agrees with Garcia et al. (2005) about the importance of the first two types, online information and services, and further suggests a 3-dimendional model integrating website quality metrics, the quality of e-services, and the number and type of e-services when testing e-city websites (Figure 4.1). The names of e-services are a country specific issue, therefore this research evaluates the available ones in the e-city website.

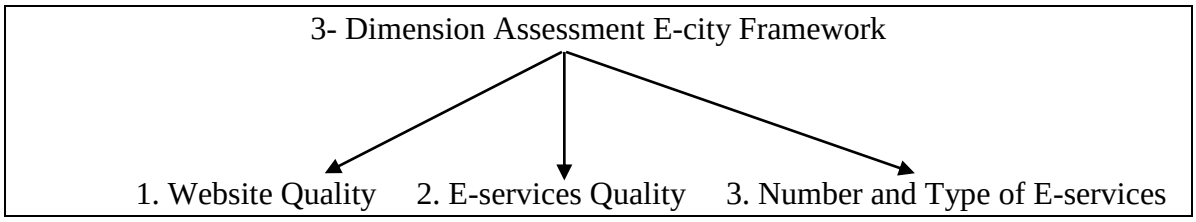

Figure 4.1: The Three-Dimensions of the Proposed E-City Usability Framework

1) Website Quality:

The theoretical web quality models incorporate usability metrics since usability is seen by many as the most important web quality measure (Huang, & Benyoucef, 2014; Hasan et al., 2012; Treiblmaier, & Pinterits, 2010). For testing the usability of municipality websites, four guidelines are chosen: 1) the USA Research-based Web Design and Usability Guidelines (HHS, 2016); 2) the UN Digital Governance in Municipalities Worldwide (Holzer et al., 2009); 3) the 2002 Community Benchmarks Program (CBP) of the Maxwell School at Syracuse University for e-municipalities (Denfeld et al., 2002) and 4) Dubai Government Websites Excellence Model (Dubai eGovernment Department, 2011) as shown on (Appendix B).

The USA is among the top 10 world leaders in the UN e-government evaluation reports and its Research-based Web Design and Usability Guidelines is created according to the best available up to date research to build high-quality websites. Each guideline shows a rating of its "Relative Importance" to the success of a website and the "Strength of Evidence" supporting that guideline. Professional web designers, usability specialists, and academic researchers contributed to these ratings. The "Strength of Evidence" represents a

consensus among researchers so the users can determine the quality of the supporting evidence (HHS, 2016). In fact, these guidelines have been validated empirically and are praised by many studies. Shneiderman (2011) wrote they are well-designed and informative guidelines. Dingli and Mifsud (2011), Rinder (2012), and Scowen and Regenbrecht (2009) praised its credibility since well-known experts in the field have reviewed its guidelines, such as Jacob Nielsen, Joseph Dumas, and Melody Ivory. Buie and Murray (2012) assure the superiority of the HHS guidelines as an authoritative government source of guidance. Moreover, Rorissa et al. (2011) mentioned that it is a demand-side model.

In addition, the Dubai model is selected because: 1) it is built on a "Customer-Focus" concept, the same scope as this research; 2) it is based on best e-practices guidelines from the UK, Canada, New Zealand and USA (Dubai eGovernment Department, 2011); 3) the 2014 UN evaluation of 100 e-city websites ranked Dubai as the first Arab e-city (Holzer et al., 2014). Globally, Dubai was ranked 9 at a score of 55.89 (out of 100), the 4th in e-services, the 5th in the privacy and security, and the 13th in usability. Also, the United Nations (2014) e-government index ranked UAE 32 at score 0.7136. For the online services, UAE e-government has a value of 0.8819 (out of 1) which means that an excellent number of e-services are offered to the public. Thus, the USA and Dubai's website guidelines may be effective in e-government context and combining their best e-practices might provide quality and validated web metrics from developed and developing countries.

Further, the assessment metrics designated specially for e-municipalities by the UN (Holzer et al., 2009) and the Maxwell School (Denfeld et al., 2002) would strengthen the proposed framework since both are devoted to assessing e-city websites. In fact, the single most referenced e-municipality guideline in the literature is the UN Digital Governance in Municipalities and besides that, it is still the only available international survey of e-cities. As well, the Community Benchmarks Program framework is praised by some studies, such as (Al-Nuaim, 2009), and its framework is based on research conducted at the Maxwell School of Syracuse University.

2) E-Services Quality:

Usefulness can be reached in e-government through ease of use and functionality that fulfills users' needs of websites. Knowing where a country stands in e-services is of great importance to both citizens and governments. Consequently, to measure e-services, this research evaluated the quality of e-services in terms of their usefulness, based on the eService Delivery Excellence Model by Dubai e-government (Dubai eGovernment Department, 2009) and User Interface Standards for e-services by Bahrain e-government (Bahrain eGovernment Authority, 2010). The main reasons for choosing Bahrain and Dubai's e-services' quality measure are: 1) they are the only specific e-services guidelines in the literature; 2) they provide an excellent example of how e-services screens should look when supporting website usability; 3) they launch many e-services, for example, Dubai e-municipality has 500 e-services out of 2000 offered by the UAE government (Dubai Smart Government, 2012); 4) they have the same tradition and culture as Saudi Arabia; 5) they are ranked high in the 2014 UN e-government report; for example Bahrain ranked 18 at score of 0.8089 which was so close to the fourth country in ranking, France (scoring 0.8938). For the online services, Bahrain e-government has an outstanding performance, a score of 9.370 (United Nations, 2014).

3) Number and Type of E-services:

Additionally, the number and type of e-services are an important measure to be considered. Some definitions show that e-government mainly means e-services (United Nations, 2008; Moraru, 2010), so merely offering information online is not e-governance; the process needs to evolve towards transactional services and this is the true start of e-government (Moraru, 2010). The classification of e-services adopted by this research is based on the Gartner model (2000) of four phases: web presence, interaction, transaction, and transformation. E-transaction is full services handling online through a two-way interaction and usually involves e-payment and submitting e-forms, while the transformation stage means integration of all e-services (back-office and front-office integration into a virtual office) and organizational changes (Montserrat, 2010). Consequently, this research takes into account the number of e-services and gives more weight to transactional and transformational online services.

However, the names of e-services are a country specific in this regard since:

1) They are different according to the nature of governance and the diverse roles by cities.

2) In the developing world, Saudi Arabia as an example, governments pass through a period of radical changes by re-engineering their public sectors and restructuring their services. Consequently, citizens are often confused and don't know which services belong to which entity of government.

3) As a concept, the usability of the website means the ability of the citizens to access the e-services and submit the request easily and efficiently online. Hence when judging the quality of the usability of the e-services, this research doesn't evaluate the physical delivery of services but only is concerned with the user interface of the website.

Consequently, this research doesn't impose or suggest certain types of e-services but evaluates the available ones on the e-city website and observes the interactions that happened online between the citizens and the e-government. Nevertheless, it might be a good idea for developed countries which have well-structured governments to explore the desired online services of their citizens and then test their e-availability and usefulness.

Thus the three-dimension proposed e-city framework builds upon the strengths of seven models (the US Research-based Web Design and Usability Guidelines (HHS, 2016), the UN Digital Governance in Municipalities (Holzer et al., 2009), the Community Benchmarks Program (CBP) for e-municipalities (Denfeld et al., 2002) and Dubai E-government Excellence Model (Dubai eGovernment Department, 2011) for website quality; the Bahrain User Interface Standards (Bahrain eGovernment Authority, 2010) and Dubai eService Excellence Model (Dubai eGovernment Department, 2009) for e-services quality; and the Gartner model (Montserrat, 2010) for classifying the type of e-services). As an attempt to produce a list of comprehensive yet manageable and practical website guidelines, this researcher tries to avoid limitations and to add other metrics if absent from all models.

4.1.2 Web Criteria Analysis

Usability metrics are the most important web quality measures according to three theoretical web models (Calero et al., 2005; Hasan, & Abuelrub, 2011; Treiblmaier, & Pinterits, 2010). For website quality, the proposed framework is based on four models (the

CBP, the US Web Design and Usability Guidelines, the UN Digital Governance in Municipalities Worldwide and Dubai's Websites Excellence Model). The number of the most important HHS usability guidelines (score of 5 or 4) is 102 out of 209, for the UN Digital Governance it's 98, for the CBP it's 49 and for Dubai Government Websites Excellence Model it's 46 guidelines. But actually, the number of guidelines is much more since some of them contain several other sub guidelines. Putting this into consideration, the Dubai model, for example, increases to about 67 guidelines. In the US guidelines, the recommendation such as "increase website credibility" contains numerous guidelines: provide a useful set of FAQ, ensure the site is up-to-date, provide links to outside sources and materials, ensure the website is arranged in a logical way and ensure the site looks professionally designed. Notice also that the last guideline is not specific enough or could mean different things to different people. Consequently, toward initially formulating the proposed e-city framework, the complete guidelines of the USA, UN, Dubai and CBP guidelines are reexamined, different metrics are compared for similarity, and therefore duplication was removed. Web assessment criteria with their reoccurrence were collected, analyzed and recorded in a table similar to Table 4.1; the check mark (✓) means the metric is mentioned in that guideline.

Table 4.1: Example of Web Criteria Analysis

	Criteria	USA Guide	UN Guide	CBP Guide	Dubai Guide
1	Graphics should not look like banners ads	✓			
2	Budget information (UN) Current municipal budget **(CBP)**		✓	✓	
3	Quick access to the site from a search engine (top 10) (D) Be easily found in the top 30 (US) Placement of the website on Yahoo, Google, MSN for official and popular names **(CBP)**	✓		✓	✓
4	Homepage link is available across all pages (D) Links to the homepage on every page (UN) Enable access to the homepage ('Home' on top) (US) Link to home page from every page visited **(CBP)**	✓	✓	✓	✓
70	Continued in Appendix C				

The procedure for web criteria analysis to select website quality metrics is as follows: 1) the three theoretical web models confirm usability as the most important web quality metrics; 2) four e-government models for website quality were chosen from the literature review; 3) the common e-government guidelines were chosen; 4) the guideline was selected if it fulfills a design principle on the g-quality inspection method for e-government proposed by Garcia et al. (2005); 5) the selected heuristics were checked against a usability framework of Folmer et al. (2003) to identify affected quality attributes; 6) the g-quality fulfilled heuristics produced the "E-City Usability Guidelines" (see Figure 3.3a in the previous chapter).

Similarly, for the e-service quality metrics, the Dubai eService Model, and Bahrain's User Interface Standards were checked and the researcher selected appropriate metrics that fulfill web design principles in the g-quality method, ending up with the "E-Services Quality Guidelines". Combining the "E-City Usability Guidelines" and "E-Services Quality Guidelines" produced the first draft of the "E-City Usability Guidelines" (Table 4.2), with their corresponding g-quality fulfilled heuristics design principles and affected usability attributes, that needs further refinements, As explained before in the methodology chapter, this research proposed framework is based on a selection from a wide range of possible options of existing frameworks in the literature for designing usability guidelines. Other e-government models were not selected because of their limitations.

Table 4.2: First Draft E-City Usability Guidelines

H = Heuristics, E = Efficiency, L = Learnability, R = Reliability, S= Satisfaction

Components	Guidelines	Fulfilled Heuristics	Affected Attributes
Content	1. Useful and most up to date content	H 14	R+S
	2. Contact information (phones, e-mails, physical address, link to customer service email, working hours)	H 15	R+S
	3. About us: vision, mission, objectives of a website	H 14	R+S
	4. Municipal Budget information	H 16	R+S
	5. Minutes: municipal meetings (date, location, agenda)	H 16	R+S
	6. FAQ with facility to ask new questions	H 10+ 14	L+R
	7. eJob	H 8	R+S
	8. News	H 14	R+S

Components	Guidelines	Fulfilled Heuristics	Affected Attributes
	9. Last update date on the footer of every page	H 14	R+S
	10. Simple forms with required fields and proper feedback	H 5 + 9	E+R
Interaction	11. Ask municipality and response time	H 15	R+S
	12. Comment or eSuggest on the website	H 15	R+S
	13. eComplaint and time to resolve it	H 15	R+S
	14. Citizen satisfaction survey	H 15	R+S
Access	15. Multilingual equivalent site with a link on header of page	H 11	L+S
	16. Meaningful images and video that don't slow downloads	H 8	E+S
	17. Quick access to the site from Google and Yahoo (top 10 search results)	H 4	E+R
	18. Links to national portal and related government websites	H 2	E+ L+R
	19. Downloadable documents/forms with appropriate access	H 9	E+R
	20. Design for common browsers access (Explorer, Chrome)	H 11	L+S
	21. Print properly pages	H 5	E+R
Structure	22. Well-designed customer focused homepage: • Quick access to highlighted services • Targeted audience group (citizens, business, tourists,..)	H 6 H 7	E+R E+S
	23. Clear entity "Logo" on every page	H 2	E+ L+R
	24. Link to homepage from every page through "Home" or logo	H 2+3	E+L+R
	25. Sitemap	H 2	E+ L+R
	26. Effective search on the Header	H 10	E+ L+R
	27. Short and descriptive page titles	H 8	E+S
	28. Readable pages (font, color, background)	H 4 + 11	L+ R+S
	29. Simple page with reasonable length (2 screens or less)	H 8	E+S
	30. Privacy and security statement/policy	H 13	R+S
Links and Navigation	31. Logically organized and short meaningful link labels	H 2	E+ L+R
	32. All links working properly (no broken links)	H 2	E+L+R
	33. Navigational options: • Indicator on user location in the site • Enabled "Back button"	H 1 H 3	E+L S
	34. Different colors for visited/unvisited links, underline links and no misleading cues to click	H 2	E+ L+R
Consistency	35. Consistent design of all pages (same feel and look, font, color, navigation bar, link style, etc.)	H 4	L+R
E-services Quality	36. Sufficient information on e-services: name, description, requirements, instructions and service centers location	H 15	E+S
	37. Ease of navigation through the e-service process	H 2	E+L+R
	38. Clear error message for invalid or incomplete data entry before form submission	H 9	E+R
	39. Online tracking for forms and e-services being processed	H 1	E+L
	40. Feedback when users waiting or submitting a request	H 1	E+L
	41. No physical visit to government offices	H 15	E+S

Components	Guidelines	Fulfilled Heuristics	Affected Attributes
	42. For e-payments: • Availability of various e-payment methods (VISA, etc.) • Notification of e-payment via SMS or email	H 7 H 13	E+S R+S

4.2 Refinement of the Developed Framework

Three refinement steps were essential to verify the preliminary developed e-city usability framework as a form of validation for the proposed guidelines (Figure 4.2): 1) a pilot test, 2) a double-expert review and 3) an application of the framework on a high-ranked city website to produce the final form of the E-City Usability Framework.

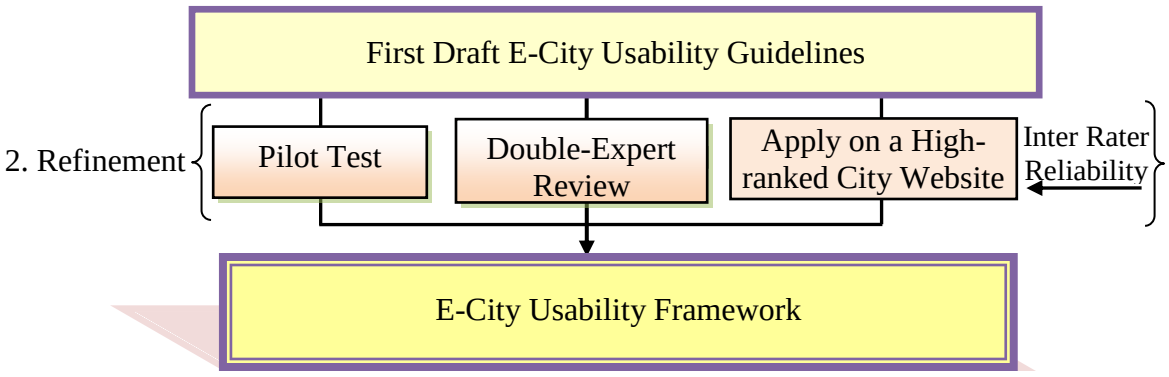

Figure 4.2: Refinement of the Proposed E-City Usability Framework

4.2.1 The Pilot Test

A copy of the developed framework was given to a usability expert, working in IT department, King Abdulaziz University in Jeddah, to check if it was clear and comprehensible. The expert was asked to conduct a pilot test to discover any problems of inconsistencies or ambiguity in the proposed guidelines checklist, taking as much time as she needed. To a large extent, most of the guidelines were found easy to understand but a few of them were not clear enough such as "eJob" and "News" guidelines, or should be rephrased such as the "Minutes" guideline and accordingly were reworded for clarification.

4.2.2 Double-Expert Review

Based on Nielsen's view "double experts", application domain and usability experts are the best web evaluators, thus to further revise the guidelines by double experts

(e-government and usability experts). Three experts were recruited from King Abdulaziz University. The first one was chosen for her experience as a web coordinator in the Deanship of E-learning and Distance Education. The other two experts were selected from the Faculty of Computing and Information Technology as specialists in usability and e-government subjects; one being the head of IT department and the other a PhD candidate whose research is concerned with usability. All three experts are well suited for conducting usability assessment. Their different scientific backgrounds and diverse job experiences could further enrich the revision of the proposed guidelines. The design of the form given to the double-experts was based on, and similar to, the form in the study of Hassan and Li (2005). Table 4.3 presents an example of the double-expert review form and a full version is available in Appendix D. The evaluation results helped in: 1) classifying the type of guideline as an objective (guidelines with yes or no answer) or a subjective measure (guidelines that need users' opinions), 2) accepting or rephrasing the guideline, 3) approving the guideline's categorization or suggest moving the guideline to a different category and 4) proposing suggestions of improvement or additional guideline(s).

Table 4.3: Example of the Double-Expert Review Form

Objective Guidelines: Guidelines that could be answered with yes or no as available or not available.
Subjective Guidelines: Guidelines that need people's perceptions and opinions.

Guidelines Component	Guidelines	Expert's Comments			
		Type of Guideline		Acceptable phrasing/ Rephrase	Does Guideline fit into component?
		Objective	Subjective		
Content	1.Useful and most up to date content		✓	Rephrase: useful content to users	Ok
	2.Contact information (phone, e-mails, physical address, location maps, working hours)	✓		Ok	Ok
	3.About us: vision, mission, objectives of a website	✓		Ok	Move to access component

The experts' comments and suggestions were recorded and carefully considered in developing the final version of this research's E-City Usability Framework. A guideline was added to the proposed framework if: 1) it was suggested by at least two double-experts; 2) it was suggested by one expert and found in one of the guideline sources reviewed by this

study, i.e. the UN Digital Governance in Municipalities, the CBP, the US Research-based Web Design and Usability Guidelines and Dubai Government Websites Excellence Model.

For classifying the type of a guideline as objective/subjective or changing its category, the review evaluation of two out of three experts was the determining factor for any changes, while rephrasing a guideline required the researcher to check the pilot and expert review forms in addition to the four original model sources to determine the best phrasing.

As a result of the expert review, three guidelines were added to the proposed framework (emergency management, interactive city map, and social media) and two guidelines were combined under the e-service quality category (simple forms with required fields guideline and clear error message for invalid or incomplete data entry guideline). It has been advised to split some guidelines that have subsections or include several important items, causing more weight to be given to them, such as the consistency guideline. Also, some guidelines in Table 4.2 were rephrased as guidelines 5, 17, 33 and 41, or need more clarification such as the "ejob" guideline and news guideline. They also suggested other guidelines that could include more elements such as the "mayor corner" was added to the "about us" guideline and the condition "open in a new window" appended to the end of guideline 18: "links to related government websites". The proposed framework after the review by the double experts is presented in Table 4.4.

Table 4.4: The E-City Usability Framework after Expert Review

Components	Guidelines
Content	1. Useful and most up to date content
	2. Contact information (phones, e-mails, physical address, link to customer service email, working hours)
	3. About us: mayor corner, vision, mission, objectives of the website
	4. Municipal budget information
	5. City council meetings (dates, locations, agendas, minutes)
	6. FAQ with facility to ask new questions
	7. eJob: job vacancy at municipality
	8. News important to users (city statistics, projects, calendar of events, photo gallery …)
	9. Last update date on the footer of every page
	10. Emergency alerts (road closedown, weather alerts…)

Components	Guidelines
Interaction	11. Ask municipality and response time
	12. Comment or eSuggest on the website
	13. eComplaint and time to resolve it
	14. Citizen satisfaction survey
	15. Interactive city map (location of near-by services, transport, restaurant, hospitals, shopping, schools...)
	16. Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)
Access	17. Multilingual equivalent websites with a link on header of page
	18. Meaningful images and video that don't slow downloads
	19. The city website among top 10 hits (results) of Google and Yahoo search engines
	20. Links to related government websites open in a new window
	21. Downloadable documents/forms with appropriate access
	22. Design for common browsers access (Explorer, Chrome)
	23. Print pages properly
Structure	24. Well-designed customer focused homepage: • Quick access to highlighted services through the main menu • Targeted audience group (citizens, business, tourists...)
	25. Clear entity "Logo" on every page
	26. Link to homepage from every page through "Home" or logo
	27. Sitemap
	28. Effective search on the Header
	29. Short and descriptive page titles
	30. Readable pages (font, color, background)
	31. Simple page with reasonable length of not more than 2 screens
	32. Privacy and security statement/policy
Links and Navigation	33. Logically organized and short meaningful link labels
	34. All links working properly, i.e. no broken links
	35. Navigational options: • Indicator of a user is where on the site (e.g. Breadcrumbs) • Enabled "Back button"
	36. Different colors for visited/unvisited links, underline links and no misleading cues to click
Consistency	37. Consistent design of all web pages: • Page layout (same feel and look, font, color, buttons, menus) • Navigation and link style
E-services Quality	38. Sufficient information on e-services: name, description, requirements, instructions and service centers' locations
	39. Ease of navigation through the e-service process
	40. Simple forms with required fields and clear error message for invalid or incomplete data entry
	41. Online tracking for forms and e-services being processed

Components	Guidelines
	42. Feedback when users waiting or submitting a request
	43. E-services completely online if possible
	44. For e-payments: • Availability of various e-payment methods (VISA, etc.) • Notification of e-payment via SMS or email

4.2.3 Application of the Developed Guidelines on a High-ranked City Website

The third and final step to validate the proposed framework and evaluate its quality is to apply the recommended guidelines on a high-ranked city website, therefore the Seoul website was chosen because it was the first city website worldwide for five consecutive years 2003, 2005, 2007, 2009 and 2011 (Holzer, & Kim, 2003; 2005; 2007; Holzer et al., 2009, Holzer, & Manoharan, 2012).

The search for the Seoul website was conducted using two of the most popular online search engines, Google and Yahoo in English, and the keywords of the search were a combination of the name of the city, municipality, and Website. In general, the web address for governmental entities must end with the extension gov.country's name abbreviation and this regulation is important to the visitors of a governmental site since it is the only credible way to tell if a site really belongs to the government or not. Therefore, only the sites labeled .gov were considered as a true indication of the official e-government website. With this in mind, the Seoul website (http://english.seoul.go.kr/) was found from the first hit using phrases such as: city of Seoul Web site, Seoul e-city web site, Seoul Web site, Seoul Metropolitan Government, Seoul municipality website, municipality of Seoul and Seoul government web site, whereas it was found among the top 10 search results when writing other combinations like Seoul local government, and that indicates the Seoul website already complies with the proposed guideline " The city website among top 10 hits (results) of Google and Yahoo search engines ".

Furthermore, analyzing the content of the Seoul website by the researcher anticipated a positive impression and also revealed that it is a simple well-structured and easy to use website. However, closer examination disclosed two serious problems on the site: the search facility was not functioning well especially when looking for specific phrases and,

more important, there were no online services which basically hindered further evaluation of the site. Searching the literature review for answers lead to a PhD dissertation comparing the sites of Washington D.C. and Seoul city website in terms of fostering citizen participation (Chung, 2011). The author acknowledges that the content of the Seoul website in foreign languages is different from that of the website in Korean; in fact, the English website's sole purpose was to provide information for foreigners living in the city of Seoul. That means online services are mostly presented in the Korean language only. Therefore, the researcher was unable to further evaluate the Seoul website because of the language barrier.

Hence it was important to choose another city website for the pilot test of the proposed guidelines and since there is no Arabic city among the top ten high-ranked city website, therefore the New York City (NYC) website was selected for this purpose for the following reasons: 1) the website is in English and 2) the NYC website is among the top 10 city websites worldwide for five consecutive years 2003, 2005, 2007, 2009 and 2011 according to the UN Digital Governance in Municipalities (Holzer, & Kim, 2003; 2005; 2007; Holzer et al., 2009, Holzer, & Manoharan, 2012).

In examining the English NYC website, any link whose URL doesn't end with .gov will be excluded from the testing of this website. Actually, The NYC website (http://www.nyc.gov/html/index.html) was found from the first hit of Google and Yahoo using the search phrases: New York City website, New York local e-government website and New York municipality website, therefore the NYC website already fulfills the proposed guideline "The city website among top 10 hits (results) of Google and Yahoo search engines ". In addition, the NYC language gateway offers access to NYC programs, services and activities in English, Spanish, Chinese and Russian. Also, the city website provides a "NYC311" section as the main source of city government information and non-emergency services in over 50 languages while "Notify NYC" communicates localized emergency information quickly to city residents.

The procedure for testing the NYC website was to separate the proposed guidelines into objective and subjective criteria based on the review evaluation of at least two out of three

participating experts and that resulted in 37 guidelines for measuring website quality (30 objective guidelines that could be answered with yes or no and 7 subjective guidelines) plus 7 guidelines for assessing e-service quality. Then a pilot test was conducted by the researcher on the NYC website acting as a final revision of the guidelines. Consequently, regarding the three guidelines (eComplaint, comment/eSuggest and ask municipality), the NYC311 webpages were checked and a feedback link existed with three drop down menu options (suggestion, complaint, compliment), thus it was decided that eComplaint and comment/eSuggest guidelines remain the same while the "ask municipality" should be deleted since it can be performed using the comment/eSuggest function on the website.

Studies by Toonders (2010) and Norris (2009) doubted the adequateness of an e-government stage approach of evaluation and instead adopted local e-government models of different components. Therefore this research introduces a proposed "E-City Usability Framework" that is divided into seven components; the first six (content, interaction, access, structure, links and navigation, and consistency) measure the website quality (30 objective and 6 subjective guidelines totaling 36 guidelines) and the seventh one appraises the e-service quality (7 guidelines) totaling 43 guidelines; as shown in Table 4.5 with their corresponding weights.

Table 4.5: The Proposed E-City Usability Framework

Components	Guidelines	Type of Guideline	Weight
Content	1. Useful and most up to date content	Subjective	0 up to 1
	2. Contact information (phones, e-mails, physical address, link to customer service email, working hours)	Objective	0 up to 1
	3. About us: mayor corner, mission, objectives of the website	Objective	0 up to 1
	4. Municipal budget information	Objective	0 up to 1
	5. City council meetings (dates, locations, agendas, minutes)	Objective	0 up to 1
	6. FAQ with facility to ask new questions	Objective	0 up to 1
	7. eJob: job vacancy at municipality	Objective	0 up to 1
	8. News important to users (city statistics, projects, calendar of events, photo gallery…)	Objective	0 up to 1
	9. Last update date on the footer of every page	Objective	0 up to 1
	10. Emergency alerts (road closedown, weather alerts…)	Objective	0 up to 1
Interaction	11. Comment or eSuggest on the website	Objective	0 up to 1
	12. eComplaint and time to resolve it	Subjective	0 up to 1
	13. Citizen satisfaction survey	Objective	0 up to 1
	14. Interactive city map (location of near-by services, transport, restaurant, hospitals, shopping, schools...)	Subjective	0 up to 1

Components	Guidelines	Type of Guideline	Weight
	15. Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	Objective	0 up to 1
Access	16. Multilingual equivalent websites with a link on header of page	Objective	0 up to 1
	17. Meaningful images and video that don't slow downloads	Subjective	0 up to 1
	18. The city website among top 10 hits (results) of Google and Yahoo search engines	Objective	0 up to 1
	19. Links to related government websites open in a new window	Objective	0 up to 1
	20. Downloadable documents/forms with appropriate access	Objective	0 up to 1
	21. Design for common browsers access (Explorer, Chrome)	Objective	0 up to 1
	22. Print pages properly	Objective	0 up to 1
Structure	23. Well-designed customer focused homepage: • Quick access to highlighted services through the main menu • Targeted audience group (citizens, business, tourists...)	 Objective Objective	 0 up to 1 0 up to 1
	24. Clear entity "Logo" on every page	Objective	0 up to 1
	25. Link to homepage from every page through "Home" or logo	Objective	0 up to 1
	26. Sitemap	Objective	0 up to 1
	27. Effective search on the Header	Subjective	0 up to 1
	28. Short and descriptive page titles	Objective	0 up to 1
	29. Readable pages (font, color, background)	Objective	0 up to 1
	30. Simple page with reasonable length of not more than 2 screens	Objective	0 up to 1
	31. Privacy and security statement/policy	Objective	0 up to 1
Links and Navigation	32. Logically organized and short meaningful link labels	Subjective	0 up to 1
	33. All links working properly, i.e. no broken links	Objective	0 up to 1
	34. Navigational options: • Indicator of a user is where on the site (e.g. Breadcrumbs) • Enabled "Back button"	 Objective Objective	 0 up to 1 0 up to 1
	35. Different colors for visited/unvisited links, underline links and no misleading cues to click	Objective	0 up to 1
Consistency	36. Consistent design of all web pages: • Page layout (same feel & look, font, color, buttons, menus) • Navigation and link style	 Objective Objective	 0 up to 1 0 up to 1
E-services Quality	37.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	Subjective	0 up to 1
	38. Ease of navigation through the e-service process	Subjective	0 up to 1
	39. Simple forms with required fields and clear error message for invalid or incomplete data entry	Subjective	0 up to 1
	40. Online tracking for forms and e-services being processed	Subjective	0 up to 1
	41. Feedback when users waiting or submitting a request	Subjective	0 up to 1
	42. E-services completely online if possible	Subjective	0 up to 1
	43. For e-payments: •Availability of various e-payment methods (VISA, etc.) • Notification of e-payment via SMS or email	 Subjective Subjective	 0 up to 1 0 up to 1

To test the guidelines of the proposed framework on any city website, the objective guidelines can be applied directly since they could be answered with a simple yes or no. For the subjective type of guidelines, designated tasks need to be developed to assess the website under investigation and its e- services. Moreover, many educational studies require the use of judges or raters to quantify some aspect of behavior, for example, judges may be used to empirically test the viability of a new scoring rubric. Of course, the task of judging invites some degree of subjectivity in that the rating will depend upon the rater's interpretation of the concept and one strategy for reducing this subjectivity is to develop scoring rubrics or instruments (Stemler, 2004). To evaluate the NYC website, in particular, heuristic evaluation was conducted by the researcher and a colleague from the Deanship of E-learning and Distance Education in King Abdulaziz University specialized in mobile e-government and usability. The double-expert heuristic test was performed using three forms: the website quality objective guidelines form presented previously in chapter 3 (Table 3.2), the website quality subjective guidelines form (Table 4.6) and the e-service quality guidelines form (Table 4.7) with designated tasks for the NYC website. The score for an objective guideline has three values "0" (absence), "1" (complete presence) and "0.5" (incomplete presence) of the website feature, but for a subjective guideline the score has five values ranging from 0, 0.25, 0.5, 0.75 and 1 to give a more accurate rating for the subjective evaluation.

In fact, the NYC website offers many e-services but only a sample of them was tested due to: 1) time constraints, 2) a limited number of evaluators, 3) some services are impossible to examine since it requires the evaluator to enter accurate citizen data which is not attainable and 4) the objective was to validate the proposed guideline and not to rank the NYC website. However, the developed tasks are samples of the most important functions on the city website and have been created with the aid of the proposed guidelines and thought to be enough to test the applicability of the developed framework.

Table 4.6: Website Quality Subjective Guidelines Form with Connected Heuristic Tasks for New York City Website

	Guidelines	Tasks	Suggested Score
1	Useful and most up to date content	1. How do I get rid of unwanted items? 2. Read recent vital statistics about the City	0___.25___.5___.75___1
Comments:			
2	eComplaint and time to resolve it	1. Provide a feedback and complain to NYC311 that "Arabic version for most NYC311 webpages is not working"	0___.25___.5___.75___1
Comments:			
3	Interactive city map (location of near-by services, transport, restaurant, hospitals, shopping)	1.Explore the New York City online map (NYCityMap) searching for Statue of Liberty then for the Metropolitan Museum of Art 2.Find the nearest University to the Empire State Building	0___.25___.5___.75___1
Comments:			
4	Meaningful images and video that don't slow downloads	1.Watch any video on the NYC media	0___.25___.5___.75___1
Comments:			
5	Effective search on the Header	1. Search for: online birth certificate 2. Search for: citywide events calendar	0___.25___.5___.75___1
Comments:			
6	Logically organized short meaningful link labels	1.Get information about how to invite the Mayor to an event 2. What are the 10 top tours in New York City?	0___.25___.5___.75___1
Comments:			

Table 4.7: E-service Quality Guidelines Form With Connected Heuristic Tasks for New York City Website

NA= Not Applicable

Task 1: Fill in Customer Satisfaction Survey	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1 or NA
2. Ease of navigation through the e-service process	0___.25___.5___.75___1 or NA
3. Simple forms with required fields and clear error message for invalid or incomplete data entry	0___.25___.5___.75___1 or NA

4. Online tracking for forms and e-services being processed	0___.25___.5___.75___1 or NA
5.Feedback when users waiting or submitting a request	0___.25___.5___.75___1 or NA
6. E-services completely online if possible	0___.25___.5___.75___1 or NA
7. For e-payments: • Availability of various e-payment methods (VISA, etc.)	0___.25___.5___.75___1 or NA
• Notification of e-payment via SMS or email	0___.25___.5___.75___1 or NA
Comments:	

Task 2: Find your Towed Vehicle (plate #:283774747 , State: NW-New York, Type: passenger)	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1 or NA
2. Ease of navigation through the e-service process	0___.25___.5___.75___1 or NA
3. Simple forms with required fields and clear error message for invalid or incomplete data entry	0___.25___.5___.75___1 or NA
4. Online tracking for forms and e-services being processed	0___.25___.5___.75___1 or NA
5.Feedback when users waiting or submitting a request	0___.25___.5___.75___1 or NA
6. E-services completely online if possible	0___.25___.5___.75___1 or NA
7. For e-payments: • Availability of various e-payment methods (VISA, etc.)	0___.25___.5___.75___1 or NA
• Notification of e-payment via SMS or email	0___.25___.5___.75___1 or NA
Comments:	

Task 3: a) As a residence submit a Cable Television or Open Video System Complaint b) Then track the reference transaction number given in (a)	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1 or NA
2. Ease of navigation through the e-service process	0___.25___.5___.75___1 or NA
3. Simple forms with required fields and clear error message for invalid or incomplete data entry	0___.25___.5___.75___1 or NA
4. Online tracking for forms and e-services being processed	0___.25___.5___.75___1 or NA
5.Feedback when users waiting or submitting a request	0___.25___.5___.75___1 or NA
6. E-services completely online if possible	0___.25___.5___.75___1 or NA
7. For e-payments: • Availability of various e-payment methods (VISA, etc.)	0___.25___.5___.75___1 or NA
• Notification of e-payment via SMS or email	0___.25___.5___.75___1 or NA

Task 4: Check for the availability of paying a parking ticket	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1 or NA
2. Ease of navigation through the e-service process	0___.25___.5___.75___1 or NA
3. Simple forms with required fields and clear error message for invalid or incomplete data entry	0___.25___.5___.75___1 or NA
4. Online tracking for forms and e-services being processed	0___.25___.5___.75___1 or NA
5.Feedback when users waiting or submitting a request	0___.25___.5___.75___1 or NA
6. E-services completely online if possible	0___.25___.5___.75___1 or NA
7. For e-payments: • Availability of various e-payment methods (VISA, etc.)	0___.25___.5___.75___1 or NA
• Notification of e-payment via SMS or email	0___.25___.5___.75___1 or NA
Comments:	

4.3 Inter-rater Reliability

According to Stemler (2004), across all circumstances involving raters, it is important to estimate the degree of inter-rater reliability, which refers to the level of agreement between several raters on a particular instrument at a particular time, as this value has important implication for the validity of the study results. In fact, there are several statistical methods for computing inter-rater reliability. The most popular method for computing a consensus estimate of inter-rater reliability is through the use of the simple percent-agreement figure which is calculated by adding up the number of the same rating cases divided by the total number of cases rated by the two raters. The percent agreement statistic has several advantages, it has a strong intuitive appeal plus it is easy to calculate and to explain. Table 4.8 shows the assessment results of the NYC website plus the percent agreement on evaluation between the expert and this researcher whereas the data is extracted from the evaluation forms but ordered in the same manner of the proposed E-City Usability Framework (Table 4.5).

Table 4.8: Result of Two Judges Rating New York City Website

Website Quality				
Components	Guideline Number	Expert 1	Expert 2	Agreement
Content	1	1	1	√
	2	0.5	0.5	√
	3	1	1	√
	4	1	1	√
	5	0.5	1	×
	6	0.5	0.5	√
	7	1	1	√
	8	1	1	√
	9	0	0	√
	10	0.5	1	×
Interaction	11	1	1	√
	12	1	1	√
	13	1	1	√
	14	1	1	√
	15	0.5	0.5	√
Access	16	0.5	0.5	√
	17	0.75	1	×
	18	1	1	√
	19	0.5	0.5	√
	20	1	1	√
	21	1	1	√
	22	1	1	√
Structure	23.1	1	1	√
	23.2	1	1	√
	24	1	1	√
	25	1	1	√
	26	1	1	√
	27	1	1	√
	28	1	1	√
	29	1	1	√
	30	0.5	0.5	√
	31	1	1	√
Links and Navigation	32	0.5	0.5	√
	33	1	1	√
	34.1	1	1	√
	34.2	1	1	√
	35	0.5	0.5	√
Consistency	36.1	0.5	1	×
	36.2	1	1	√
Total Score on Website Quality		32.25	34 (out of 39)	
Percent Agreement on Website Quality Evaluation			≈ 90%	

E-services				
E-service #1	37	0.5	0.5	√
	38	1	1	√
	39	1	1	√
	40	NA	NA	√
	41	0.75	1	×
	42	1	1	√
	43.1 43.2	NA NA	NA NA	√ √
E-service #2	37	1	1	√
	38	0.5	0.75	×
	39	0.5	0.75	×
	40	1	1	√
	41	0.75	1	×
	42	1	1	√
	43.1 43.2	NA NA	NA NA	√ √
E-service #3	37	1	1	√
	38	0.75	0.75	√
	39	0.25	0.5	×
	40	0.25	0.5	×
	41	0.75	0.75	×
	42	1	1	√
	43.1 43.2	NA NA	NA NA	√ √
E-service #4	37	1	1	√
	38	1	1	√
	39	NA	NA	√
	40	NA	NA	√
	41	1	1	√
	42	1	1	√
	43.1 43.2	1 1	1 1	√ √
Percent Agreement on E-services Evaluation ≈ 78%				
Overall percent Agreement between the expert and the researcher ≈ 84.5%				

If two raters were in agreement on the rating to score behaviors, then these raters share a common interpretation of the guidelines. However, a typical guideline in the literature for assessing the quality of inter-rater reliability based upon consensus estimates is that they should be 70% or greater (Stemler, 2004). High inter-rater reliability was found as the percentage of agreement among raters was approximately 90% on website quality evaluation of the NYC, 78% on e-services evaluation and 84.5% for the overall agreement,

which indicates that both the expert and the researcher agree, to a large extent, on how to apply the rating scale. Moreover, the NYC received a total score of 32.25 (≈83%), out of 39 points, and 34 (≈87%) for its website quality according to the expert and the researcher respectively. On the other hand, the total score for e-services was not calculated since it was a sample test not intended to rank the NYC website but only as a form of validation of the proposed framework.

4.4 The Scoring Method

Based on the CBP model and the usability section of the UN Digital Governance model, each web metric in the developed framework, whether measuring website quality or e-services quality, was scored on a rating of "0" or "1" to assess the absence or presence of the most essential website features. Baker (2009) supports in score assignment or weighing, the norm that items receive equal scores unless definitive justifications exist for differential weighting. This could be applied as the aim of this research is to compare and diagnose usability problems of Saudi cities. A fraction of that score is given if the item had some problems and that depends on the type of guideline; the objective guidelines deserve 0, 0.5 and 1 while the subjective guidelines have five values (0, 0.25, 0.5, 0.75 and 1) to yield a more accurate rating. For example, a "0" is given if no search facility is available on the website, a "1" is given if the search is functioning well, or a fraction ("0.25", "0.5", "0.75") if the search not functioning correctly. For e-services, this research framework considers not only the quality of e-services but also their number and type corresponding roughly to the stage model concept proposed by Gartner's e-government framework. Thus, to calculate a total score for the e-services, the following steps were needed:

- Each e-transaction service (Gartner's stage 3) available on the website received a score of 3, similar to the service category in the UN model, and each e-transformation service (Gartner's stage 4) received a score of 4
- That number, transaction score of 3 or transformation score of 4, was then multiplied by the mean of applicable e-service quality guidelines (last component of the framework).

Table 4.9 explains in detail the scoring method adopted for this proposed framework. The maximum score a site can receive for the website quality is 39, inferred directly from the first six components of the framework, but the maximum score for total e-services is

unlimited depending on the type, quality and the number of e-services. For example, if website "A" got a score of 30 for website quality and has 7 transactional services; 5 of excellent quality and 2 of medium quality services, then the grand total for website "A"= total score for website quality + total score for e-services= 30 + [5*(3*1)] + [2*(3*0.5)]= 48. An overall ranking of the e-city website or a ranking by components (e.g., content component) or a ranking by website quality or e-services are possible since different components give different information about web quality.

Table 4.9: The Scoring Method

Website Quality		
Components	Max Score	
Content (10 guidelines)	10	
Interaction (5 guidelines)	5	
Access (7 guidelines)	7	
Structure (9 guidelines)	10	
Links and Navigation (4 guidelines)	5	
Consistency (1 guideline)	2	
Total score for website quality	39	
E-services		
	Score	Total
E-service # 1 (Type: transaction) E-service quality	3 (0-1)/e-service	Transaction score * (mean of the e-service quality guidelines)
E-service # N (Type: transformation) E-services quality	4 (0-1)/e-service	Transformation score * (mean of the e-service quality guidelines)
Total score for e-services		Σ Transaction + Transformation scores
Total	Σ score for website quality + Total score for e-services	

Chapter 5: Testing

This chapter identifies appropriate web evaluation methods to be applied on the selected Saudi e-city websites. The procedures for all testing and the criteria for selecting Saudi e-city websites are identified.

5.1 Web Evaluation Methods

An analysis of available web evaluation methods, discussed earlier in the literature, leads us to select the following evaluation methods:

1. User testing methods: this approach was used since it is the most important evaluation method that involves real users with real tasks and it discovers major usability problems.
2. Heuristic evaluation: this method was used throughout the refinement process of the E-City Usability Framework, such as the pilot test, the double-expert review and the testing of the guidelines on a high-ranked e-city website (NYC). Also, heuristic evaluation was conducted by two experts to test the website quality objective guidelines of the proposed framework on selected Saudi e-city websites.
3. Automatic website evaluation: this method was used on a small scale. No reliable automatic tools were identified in the literature for usability testing. Most tools go through a period of instability that impedes their application. Only a link checker was used to check specific guidelines. For example, the number of broken links provided by a link checker could be compared with the suggested guideline to ensure all links are working properly. A variety of link checker tools, such as Xenu and Web Link Validator, were available and the validated ones were used.
4. Alexa: this web analytics tool was used to measure certain web metrics on e-city websites such as website traffic rank. This reflected the number of users who visited that site (reach) and the number of web pages viewed by site visitors (page views). Also, it provided us with time on site as a possible indicator of user satisfaction as well as the speed of download of a website which could emphasise guidelines such as "ensure images do not slow download". However, doubts remain about its reliability since few studies use Alexa without questioning the tool.

5.2 Testing within the Context of Saudi E-City Websites

Multiple evaluation methods were utilized by this research to detect usability problems of selected Saudi municipal websites encountered by users.

5.2.1 Selection of Saudi E-City Websites

The selection of Saudi e-city websites was based on two criteria in the following order:

1. The number of Internet users: the rationale for selecting Saudi regions with the highest percentage of Internet users is obvious since more Internet users are more likely to use the e-government portal.
2. The city's population size: the rationale for selecting the largest city in each Saudi region originates from many e-government publications such as the Digital Governance in Municipalities Worldwide study by the UN (Holzer et al., 2009) and the study of Moon (2002); all suggest a positive relation between population and local e-government capacity. Another reason is that e-government affects a broad segment of the population in large cities.

Therefore, selecting Saudi e-city websites was based on regions with the highest Internet usage and the largest populated city in each region.

The parent organization of local governments, the Saudi Ministry of Municipal and Rural Affairs, was checked to determine the municipalities that have a web presence, but unfortunately, the ministry doesn't have an English-language website and it provides links to random e-municipalities only. Hence we turned to the Saudi e-government national portal for needed information and accordingly, the total number of main municipalities in 2013 was 16: Al Bahah, Al-Jouf, Alahsa, Assir, Hail, Holy Makkah, Jazan, Jeddah province, Al-Madinah, Municipality of Eastern Region, Najran, Northern Borders, Qassim, Riyadh, Tabuk and Taif province (Saudi National e-Government Portal, 2013), as presented in Table 5.1. However, due to the rapidly changing nature of websites, the researcher is bound to re-examine each municipality website again just before the experimental stage. Regarding how to write the name of Saudi cities in English, since there is no standard in the English translation, both the English websites of Saudi national portal, as well as the e-

municipality itself, were checked, and in case of a difference in spelling, the name of the city was written as dictated in the logo of the municipality website.

Table 5.1: Saudi Municipalities Websites

	Main Saudi Municipalities	URL
1	Jeddah Municipality	www.jeddah.gov.sa
2	Riyadh Municipality	www.alriyadh.gov.sa
3	Municipality of Eastern Region	www.eamana.gov.sa
4	Assir Municipality	http://ars.gov.sa/
5	Al-Madinah Municipality	www.amana-md.gov.sa
6	Holy Makkah Municipality	www.holymakkah.gov.sa
7	Taif Province Municipality	www.taifcity.gov.sa/
8	Al-Jouf Municipality	www.amanataljouf.gov.sa/
9	Alahsa Municipality	www.alhasa.gov.sa
10	Hail Municipality	www.amanathail.gov.sa
11	Jazan Municipality	www.jaz.gov.sa
12	Najran Municipality	www.najran.gov.sa
13	Northern Borders Municipality	www.arar-mu.gov.sa
14	Qassim Municipality	www.qassim.gov.sa/
15	Al Baha Municipality	www.mob.gov.sa/
16	Tabuk Municipality	www.tabukm.gov.sa/

Before beginning the experimental stage the researcher searched thoroughly for updated statistics about Internet usage in different Saudi regions in the following websites: Ministry of Communications and Information Technology (MCIT), Communications and Information Technology Commission (CITC), Central Department of Statistics and Information, Saudi National e-Government Portal, Saudi Telecommunication Company (STC), Discover Digital Arabia (latest trends and digital statistics in the Arab region), and also Internet search engines such as Google and Yahoo. In fact, statistics about Internet penetration in Saudi Arabia were available on an annual basis but no such statistics were found about Saudi regions except in one report in 2008 (CITC, 2008). As a final effort, the researcher sent emails to MCIT, CITC, and National Contact Center (Amer). Unfortunately, MCIT didn't reply, CITC referred to their website, and Amer sent Civil Service phone numbers which didn't answer. Consequently, based on the only available Saudi Communication and Information Technology Commission report, Internet

penetration varied across 13 Saudi provinces; Internet usage was highest in Makkah, Riyadh, Eastern, Assir and Al-Madinah provinces, etc. as ordered in Table 5.2. The regions with low Internet usage, less than 4 percent, were ignored.

Table 5.2: Share of Internet Users by Saudi Province (CITC, 2008)

	Province	2008
	Total KSA	100%
1	Makkah	27.5%
2	Riyadh	24.6%
3	Eastern	15.8%
4	Assir	7.6%
5	Al-Madinah	7.0%
6	Qassim	3.9%
7	Tabuk	3.0%
8	Hail	2.2%
9	Najran	2.0%
10	Jizan	2.0%
11	Jouf	1.7%
12	Baha	1.6%
13	Arar	1.1%

During the beginning of this phase and according to the City Population (2013) and The World FactBook (2013), the largest city in Makkah province was Jeddah with a population of 3,430,697, in Riyadh province was Riyadh with a population of 5,188,286, in Eastern province was Dammam with a population of 903,312, in Assir province was Abha with a population of 236,157 and in Al-Madinah province was Al-Madinah with a population of 1,100,093 inhabitants. Therefore, the following five municipalities were selected for this study: Jeddah, Riyadh, Dammam, Abha and Al-Madinah. Looking to Table 5.1 again we noticed: 1) all Saudi municipality web addresses end with the extension .gov.country's name abbreviation and this regulation is important to the visitors of a governmental site since it is the only credible way to tell if a site really belongs to government or not, 2) only Jeddah and Riyadh have official municipal websites with the city name, 3) whereas Abha's URL stands for its province "Assir", 4) Al-Madinah municipality utilizes the Arabic name "amana", meaning municipality, with abbreviation of the city name "md", and 5) the Eastern Region uses general meaningless name "e-amana" instead of city name.

Because of the nature of rapid updates and redesign of websites, it was essential to recheck the selection of Saudi e-city websites and the research sample before beginning the actual experimental stage of this study. Thus, there was a search for updated statistics about Saudi Internet penetration but none were found and the selected municipal websites were rechecked again. Unfortunately, continuous observation of these websites revealed that Assir website, www.ars.gov.sa , has been under construction for a long period of time, over four months. An email was sent to the Assir webmaster asking when the site will be online but there was no reply. In August 2014, a beta version of Assir site was launched but it was not functioning well (Figure 5.1). Thus, it was excluded from the test and instead "Qassim" was selected, the next website in Table 5.2.

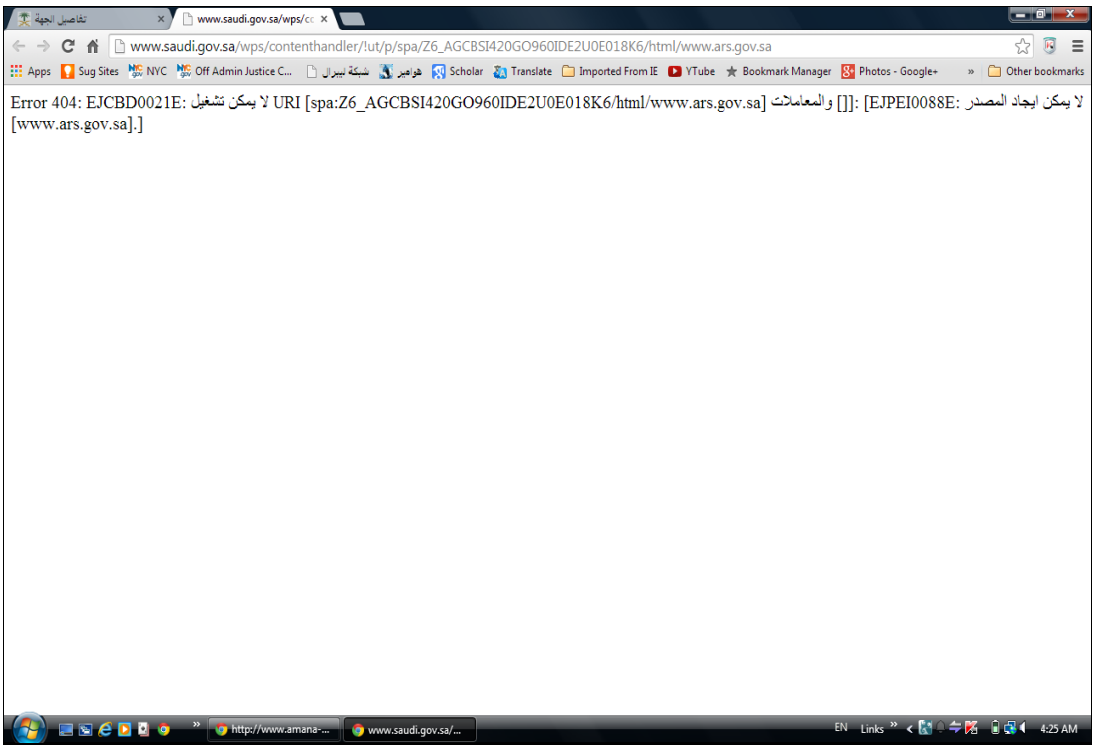

Figure 5.1: Assir Website

Hence, the five selected Saudi municipal websites to be tested are presented in Table 5.3 with their web addresses and Figures 5.2 to 5.6 show screen shots of their homepages. These local government websites offer various e-services, such as information services, online and offline forms and e-services, to the public. The primary users of these websites are all citizens and the residents of that particular region in Saudi Arabia.

Table 5.3: Selected Saudi Municipal Websites for Empirical Test

	Municipality Name	URL
1	Jeddah Municipality	www.jeddah.gov.sa
2	Riyadh Municipality	www.alriyadh.gov.sa
3	Al-Madinah Municipality	www.amana-md.gov.sa
4	Eastern Region Municipality	www.eamana.gov.sa
5	Qassim Municipality	www.qassim.gov.sa/

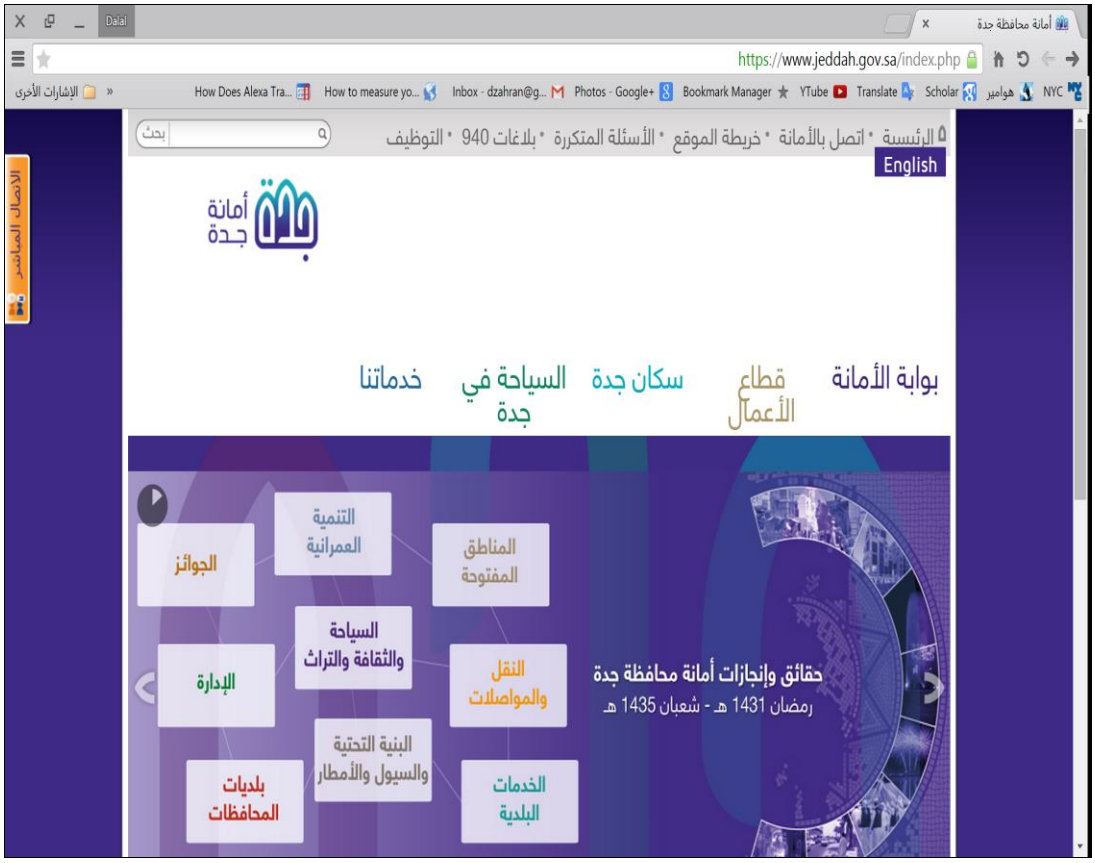

Figure 5.2: Homepage of the Jeddah E-city Website

Figure 5.3: Homepage of the Riyadh E-city Website

Figure 5.4: Homepage of the Al-Madinah E-city Website

Figure 5.5: Homepage of the Eastern Region E-city Website

Figure 5.6: Homepage of the Qassim E-city Website

5.2.2 Usability Testing

The procedure for heuristic and user testing of Saudi e-city websites, or any city website (Figure 5.7), was as follows: 1) Two experts tested the e-city website using the website quality objective guidelines form (Table 3.2); 2) Users tested the e-city website using the website quality subjective guidelines form (Table 3.3); and 3) Users tested the e-city website using the e-service quality subjective guidelines form with its designated tasks (see Appendix A). Hence, the procedure for testing Saudi e-city websites was the same as testing the New York City website except that the heuristic test was the only evaluation method used for the NYC as a form of validation of the proposed framework. But for assessing Saudi e-city websites, user testing was performed as well to gain a better insight of users' perceptions of local e-government websites. The results of user testing were combined with the heuristic tests to provide an e-score for the websites' quality. An average score for e-services was calculated from the results of testing five e-services. The average was multiplied by the number of available e-services to get a total score of e-services for the tested website.

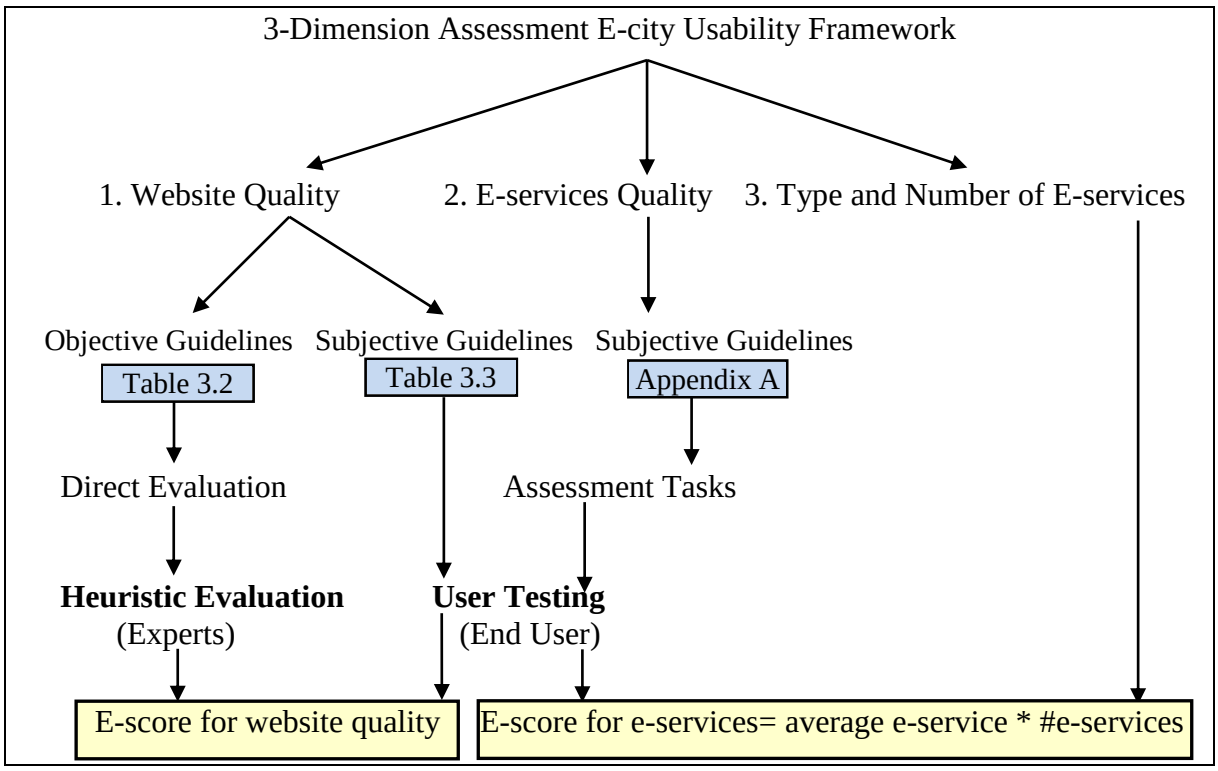

Figure 5.7: Heuristic and User-testing of Saudi City Websites using the Proposed Framework

5.2.2.1 Heuristic Evaluation

In general, heuristic evaluation was conducted by three experts in the double-expert review of the proposed framework, and by two experts in testing both NYC and selected Saudi city websites. Usually, there are two problems in recruitment: 1) recruiting experts in usability and the investigated interface is difficult and expensive, and 2) very few research discuss the "evaluator effect" with respect to detection of usability problems (Hasan et al., 2012). However, working in King Abdulaziz University makes recruiting of experts possible. To solve the second problem, we compute an inter-rater reliability score using a consensus estimate method.

The procedure of heuristic testing of Saudi e-municipalities was carried out by two experts. They evaluated the five city websites in five different sessions. At the beginning of each session, the targeted website was explored by each expert for about 15 minutes. In the second round, both experts performed a comprehensive evaluation of the selected websites guided by the website quality objective guidelines form of the proposed e-city framework (Table 3.2) to identify possible usability problems. The number of web metrics in the form is 33 guidelines. A score of "0", "0.5" and "1" was designated by the experts according to the degree of compliance with each guideline. The maximum possible score for a website was 33 while the minimum score was 0. At the end of the form, experts expressed their viewpoints and wrote down general comments on the tested website. There was no time restriction on expert in this round.

To estimate the degree of inter-rater reliability, the percentage of agreement on the evaluation between the two experts was calculated. The number of agreements on full compliance (score 1), partially compliance (score 0.5), and no compliance (score 0) guidelines were calculated also. General comments on the site by the two experts were stated and screen shots of some programming errors were recorded by the researcher. The conclusion of this section was: 1) it ranked Saudi e-city websites based on the total score given by each expert, 2) it ranked Saudi e-city websites based on the number of fully compliant guidelines, and 3) it identified the guidelines that municipality did not consider when designing their websites.

5.2.2.2 User Testing

User testing is the most fundamental usability evaluation method since it involves real users and provides direct information about how people use computers and what their exact problems are with the tested interface. Ideally, users should be able to do basic tasks on the website correctly and quickly. The main purpose of a usability test is to uncover the most serious problems disappointing users or that prevents them from completing their tasks.

A task analysis was conducted on each selected city website to identify all possible tasks. Since usability testing is a sampling process, therefore these tasks were filtered by connecting the proposed subjective guidelines with appropriate tasks in order to test the most important site features and services. Users were given short scenarios on a separate page labeled as "Task 1", "Task 2", etc. to evaluate the e-services. The method for collecting data associated with user testing was a think-aloud protocol. The researcher observed participants thinking out loud while performing the appointed tasks on selected Saudi websites.

It is an important issue to consider how many participants to include in the usability test. This research tested 5 to 8 users on each of the selected Saudi e-city website, until usability problems seem to be repeated and not much new is found, and that is based on:

1) Nielsen's discount usability engineering method that suggested a number of 5 participants could detect 85% of the potential usability problems in a product and the best results come from the first 5 users (Chow, 2013).
2) Most website usability engineering such as Stone et al. (2005) believe 3 to 5 participants were needed to see all the potential usability problems, Lewis (2006) and De Róiste (2013) considered 5 to 8 participants, were enough to discover possible usability problems in a website.
3) Zhao (2007) who first acknowledged Nielsen's method and carried out the usability test until the encountered problems started to repeat -after 5 users- and not much new was obtained.

The procedure for user testing is done as follows:

1. Users were selected for the test based on different age to cover a wide range of users of e-city websites (Table 5.4). All users have at least bachelor's degree and are regular Internet users. It is the first time for them to see the selected e-city websites. As Rinder (2012) stated, the usability of a website defines how well and easily a visitor, without formal training, can interact with the site. Joe et al. (2015) also dictated that participants were not to have used or seen the user interface before, since we are interested in first-time use learnability.

Table 5.4: Users Profile

User	Age Range	Gender	Education
1	40-49	F	Mathematics
2	20-29	M	Engineering
3	30-39	F	Hospital Administration
4	20-29	F	Economy
5	40-49	M	Communication
6	30-39	M	Administration

2. The total of G2C e-services, the total of malfunctioning e-services (such as broken links, no content, under construction web pages), and the total of transactional and transformational e-services, were calculated for e-city website.
3. G2B or G2G e-services were not considered since it is out of this study scope. In our view, citizens are the number one measure of success or failure of e-government. The proposed e-city framework employs user testing as a proxy of citizens' voice.
4. Malfunctioning e-services were excluded from usability tasks. Also, e-services that require entering private information were omitted.
5. A pilot test with one user was conducted to discover any problems before the formal user testing.
6. Ethical considerations were carefully monitored throughout this research. A consent form was signed by the participants. The test was performed in a room

equipped with a Lenovo laptop, Windows 10. It was not recorded respecting users' preference. The researcher only observed testing sessions and took notes.

7. The same users tested all selected Saudi city websites on two consecutive days to avoid users being tired. To avoid bias, the order of tested websites was different from one user to another because users usually learned from the first website and did better in the subsequent ones.
8. The session began with the researcher welcoming the users and stating the objective of the study. The steps of the test, and user's right to quit a task or even to withdraw from the test at any time, was further explained. Since most users were found reluctant, they are ensured of not recording the tests and their names would be confidential. After all, the user signed a consent form to begin the test.
9. Users have the right to "quit" a task anytime they feel frustrated, consequently, a zero was given to that task.
10. The user testing of Saudi municipal websites examined three parts: website quality, e-services quality and user satisfaction survey, as follows:
    a. Users would walk-through five Saudi e-city websites to test six subjective website quality guidelines of the proposed framework (see Table 5.5 which is an Arabic version of Table 3.3). The results were combined with the heuristic tests to provide an overview of website quality.
    b. Users were given five usability tasks to evaluate the quality of e-services in each e-city website (see Appendix A).
    c. At the end of the test, users filled in a user satisfaction questionnaire (Figure 5.8) designed to be the same for each city website.
11. Estimated session of user testing was about 2 hours for each website including a five minutes break between the three parts.
12. User testing was conducted with 6 users since it was noticed that the usability problems were repeated and not much new is found.
13. M-government facilities were not considered, such as mobile e-services, because it is outside of the scope of this study.

Table 5.5: Website Quality Subjective Tasks for Saudi City Website

Please walk-through Jeddah city website then circle a score for its performance in terms of the following criteria: فضلاً تصفح موقع أمانة جدة وضع دائرة حول العلامة المناسبة لآدائه حسب المعايير التالية:		
	Criteria	Score
1	Useful and محتوى الموقع مفيد most up to date content ومحدّث	0___.25___.5 0___.25___.5
Comments:		
2	eComplaint الشكوى الالكترونية	0___.25___.5___.75___1
Comments:		
3	Interactive city map خريطة المدينة التفاعلية	0___.25___.5___.75___1
Comments:		
4	Meaningful images الصور and videos ومقاطع الفيديو معبرة ومفيدة	0___.25___.5 0___.25___.5
Comments:		
5	Effective search البحث فعّال	0___.25___.5___.75___1
Comments:		
6	Organized short links الروابط مرتبة meaningful link labels وأسماؤها واضحة	0___.25___.5 0___.25___.5
Comments:		

1. How do you evaluate this city portal in terms of website quality?
ما هو تقييمك للبوابة من حيث جودة الموقع؟
○ Excellent (ممتاز) ○ Good (جيد) ○ Fair (مقبول) ○ Poor(سيء)

2. How do you evaluate this city portal in terms of online services quality?
ما هو تقييمك للبوابة من حيث جودة الخدمات الالكترونية؟
○ Excellent (ممتاز) ○ Good (جيد) ○ Fair (مقبول) ○ Poor(سيء)

3. How do you evaluate this city portal in terms of ease of use?
ما هو تقييمك للبوابة من حيث سهولة استخدام الموقع؟
○ Excellent (ممتاز) ○ Good (جيد) ○ Fair (مقبول) ○ Poor(سيء)

4. How do you evaluate this city portal in terms of number of e-services?
ما هو تقييمك للبوابة من حيث عدد الخدمات؟
○ Excellent (ممتاز) ○ Good (جيد) ○ Fair (مقبول) ○ Poor(سيء)

5. How many e- services you performed through any city portal this year?
كم عدد الخدمات الالكترونية التي قمت بإنجازها هذا العام من خلال موقع إلكتروني لبلدية
○ More than 10 services أكثر من 10 خدمات ○ Less than 10 services أقل من 10 خدمات ○ Noneولا خدمة

6. To what extent you depend on e-services in completion of your municipal transactions
لأي مدى تعتمد على الخدمات الالكترونية في إنجاز معاملاتك البلدية
○ Completely دائماً ○ Partially أحياناً ○ Noneلا أستخدمها

Figure 5.8: User Satisfaction Survey

5.2.2.3 Alexa Web Analytics Tool

Web traffic data contains rich and useful information about usage patterns contributing to the field of Webometrics. Generally, when studying web traffic usage data, three issues are important to consider: 1) data sources, 2) metrics definition, and 3) validity, reliability, and usefulness of these tools. The first two issues will be addressed next, while the third point will be addressed in the results chapter.

From the literature, the major data sources are as follows (Vaughan, & Yang, 2013): a) Alexa at www.alexa .com is a free service most widely used, b) Compete at www.compete.com is partially free, and c) Google Trends for websites was free when launched in 2008 at trends.google.com/websites but the site discontinued its services in September 2012. Accordingly, Alexa was chosen as a web analytics testing tool because it is: 1) the largest public free traffic data source, 2) the widest coverage area (Alexa data is collected from global users while Compete data is from American users only), and 3) the

best data source based on the studies of Vaughan (2008), Vaughan and Yang (2013), Jowkar and Didegah (2010) and Bhat (2013).

Alexa is a website metrics system analyzing the web usage of Alexa toolbar users for a three month period or more. Alexa Rank is an estimated measure of website popularity and users of the Alexa toolbar contribute to this ranking. Consequently, it is very important to know the definition of Alexa web metrics in order to measure and interpret information correctly. Among others, the following major index was defined by Alexa (2014):

1. Traffic ranks (global rank and country rank): is a combination of the number of daily visitors and the number of pages they view (pageviews) over the previous 3 months. The site with the highest visitors and page views is ranked # 1.
2. Pageviews/visitor: estimates daily unique page views per visitor to the site.
3. Speed of download: the median time it takes to load pages from the site.
4. Bounce rate: percentage of visitors leaving the website after visiting only one page.
5. Sites linking in: the number of sites that link to this site, showing its popularity.
6. Time on site (minutes): daily time on site per visitor to the site.
7. Audience geography: where are the site's visitors located (where do they come from?)
8. Where do visitors go on the site: percent of visitors to subdomain(s)
9. Reach: percentage of global Internet users who visit the site.

Among a variety of Alexa web traffic statistics, the basis used by this research for evaluating Saudi municipal sites was the following nine metrics: domestic and global traffic ranks, pageviews/visitor, speed of download, bounce rate, sites linking in, time on site, audience geography and where do visitors go on the site. Other metrics were not considered because they were not applicable indicators for the performance of e-cities, such as the global "reach" metrics since local citizens are the ones who mostly use the municipal sites.

Alexa presents its data either as an individual site overview or as site comparisons. On site overview, traffic rank data is available for sites ranked < 100,000. Thus smaller sites will not have a historical rank graph but might get other data. The site comparison option allows you to compare traffic metrics for up to 10 different websites if they have a high

volume of traffic (Alexa, 2014). Accordingly, it seems that Alexa is a measure of website popularity and not website quality or value of e-government in the cities.

5.2.2.4 Link Checker

Broken links are links that lead to a page that does not exist, such as the 404 error. Frequently Internet users facing this error were likely to leave the website, thus broken links have a negative effect on e-business or e-government entity. With the rapid growth of websites, the problem of broken links became more serious. But fortunately, there are many tools to help a website designer identify and solve this problem. To the best of our knowledge, there is no study which compares different automatic link checker tools and validates their effectiveness.

As presented in the literature review, the most used automatic web testing tools were: Bobby for accessibility, LIFT for usability and link checkers for broken links. Bobby is out of our scope since it is for site accessibility whereas testing tools for site usability were not considered efficient and reliable yet, due to the immaturity of the field. Regarding link checkers, few studies have used them without questioning their reliability. Examples of automated link checker use include the study of Mustafa and Al-Zoua'bi (2008) in which two automated tools, Web Page Analyzer and HTML Toolbox, evaluated Jordan universities' websites, and the study of Isa, Suhami, Safie, and Semsudin (2011) in which Websiteoptimization, EvalAccess, and Axandra link checker tool were used to assess Malaysia e-government websites.

To fill this gap in the research, we turned to articles published in blogs where people have posted their experiences with different link checkers and then have a discussion with other readers. Some examples of these articles were the post of Soames (2014), of Richard (2014) and of Petkova (2012). The purpose of this section is to identify and compare a sample of the most used link checker tools and also to test how reliable are these tools.

The tools most frequently cited on posts were identified and checked. The following tools were excluded from our list: paid tools e.g. WebXACT, tools only for the owner of the site e.g. Google Webmaster, slow tools e.g. A1 WS Analyzer, and the W3C Link

validation tool. The latter was one of the original link checking tools but it was slow with lengthy error messages (Soames, 2014). To find out which is the most reliable automatic link checker tool, it was decided to use seven tools: Broken Link Checker, Dead Link Checker, NetMechanic, LinkTiger, Link Alarm, Web Link Validator, and Xenu.

It was planned to test the five Saudi e-city websites with these seven link checker tools. The procedure for assessing the reliability of these tools was as follows:

1. As a preliminary trial test, these tools were applied to one Saudi e-city website, for example the Jeddah city website.
2. If there was a big difference and no agreement between these tools on the number of broken links of the Jeddah e-city website, then they were checked against a site with few bad links, i.e. a high-ranked e-city website, such as the NYC website.
3. If there was still no agreement between the seven tools on the broken links of the New York City website, then a small manually checked static website was tested by these tools plus Google Webmaster as a possible benchmark tool to compare against.

Chapter 6: Results

Multiple web evaluation methods were used to discover usability problems encountered by Saudi users of municipal websites. This chapter introduces the results of testing five Saudi municipality websites using selected web evaluation methods which are: 1) Heuristic evaluation, 2) User testing, 3) Web analytics tools such as Alexa, and 4) Automatic website testing tools such as broken link checkers. At the end, the results from web evaluation methods are combined and the usability problems of Saudi e-city websites are identified.

6.1 Heuristic Evaluation of Saudi Municipal Websites

A heuristic test of five Saudi e-city websites was carried out by two experts in e-government and usability from King Abdulaziz University during the month of June 2014. Simultaneously, both experts evaluated the selected websites using the website quality objective guidelines form of the proposed framework. The conclusion of this section ranks Saudi e-city websites, based on the total score earned, and on the number of fully compliant guidelines. Also, the guidelines that the municipality did not consider when designing their websites were identified.

6.1.1 Jeddah Municipal Website

Table 6.1 shows the assessment results of evaluating the Jeddah municipal website by two experts. A total score, a percentage agreement on the evaluation between the experts, and the number of full, partial, and non-compliance guidelines were calculated for the tested website. Experts' general written comments on the site were stated and screen shots of some website errors were captured by the researcher.

Table 6.1: Heuristic Evaluation of Jeddah Municipal Website

	Guideline Number	Expert 1	Expert 2
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	1	1
2	About us: mayor corner, mission, objectives of the website	1	1
3	Municipal budget information	0	0
4	City council meetings (dates, locations, agendas, minutes)	0	0

5	FAQ with facility to ask new questions	0.5	0.5
6	eJob: job vacancy at municipality	1	1
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	1	0.5
8	Last update date on the footer of every page	0	0
9	Emergency alerts (road closedown, weather alerts…)	0	0
10	Comment or eSuggest on the website	1	1
11	Citizen satisfaction survey	1	1
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	1	1
13	Multilingual equivalent websites with a link on header of page	1	0.5
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1
15	Links to related government websites open in a new window	1	1
16	Downloadable documents/forms with appropriate access	0	1
17	Design for common browsers access (Explorer, Chrome)	1	1
18	Print pages properly	1	1
19	Well-designed customer focused homepage: Quick access to highlighted services through the main menu	1	1
20	Targeted audience group (citizens, business, tourists...)	1	1
21	Clear entity "Logo" on every page	1	1
22	Link to homepage from every page through "Home" or logo	1	1
23	Sitemap	1	1
24	Short and descriptive page titles	1	1
25	Readable pages (font, color, background)	1	1
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5
27	Privacy and security statement/policy	1	1
28	All links working properly, i.e. no broken links	1	0.5
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	1	1
30	Enabled "Back button"	1	1
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	1	0.5
33	Navigation and link style	1	0.5
Total Score (out of 33 scores)		26 (78.8%)	24.5 (74.2%)
Percent agreement between experts on evaluation (27 out of 33 guidelines)		≈ 82%	
Agreement on "Non-compliance" (score 0)		5 guidelines (15.2%)	
Agreement on "Partial compliance" (score 0.5)		2 guidelines (6.1%)	
Agreement on "Full compliance" (score 1)		20 guidelines (60.6%)	

The total score for the Jeddah municipal website according to the first expert was 26, out of 33, (78.8%) while the second expert's score 24.5 (74.2%). An inter-rater reliability of 82% was found between both experts indicating they agreed, to a large extent, on how to apply the rating scale. That is, they agreed on the assessment of 27 guidelines: 5 guidelines of score "0", 2 guidelines of score "0.5", 20 guidelines of score "1" and they disagreed on 6 guidelines. Therefore, the Jeddah municipal website followed 20 out of the 33 guidelines reflecting 60.6% compliance with the proposed objective guidelines.

General comments by the experts on the Jeddah's municipal website were:

- The design of the site was organization centered and not user centered
- The e-services provided were limited
- Most of the offered e-services were descriptive procedures of the services rather than online transactions. Consequently, they can be categorized as "offline services", such as offline forms and information services, which is considered one-way interaction between the user and the government website, where the entire service is not occurring online
- There was no emergency alert service to the residence of Jeddah
- The English website was incomplete with several Arabic sections omitted, and some links returned the users to the Arabic site
- There was no distinction between visited and unvisited links on the site.

Screenshots of some website errors were captured by the researcher. Where possible the English webpages were presented. Otherwise, the Arabic webpages were shown instead:

- Several services were not available, such as query about a transaction (Figure 6.1), statistics page (Figure 6.2), deaths query (Figure 6.3), and query claim of Sadad payments.

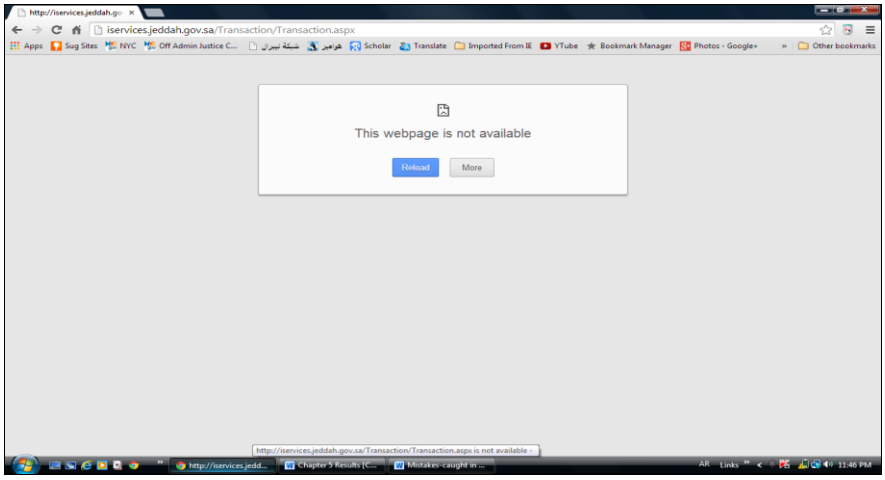

Figure 6.1: Broken Link of Transaction Query Service

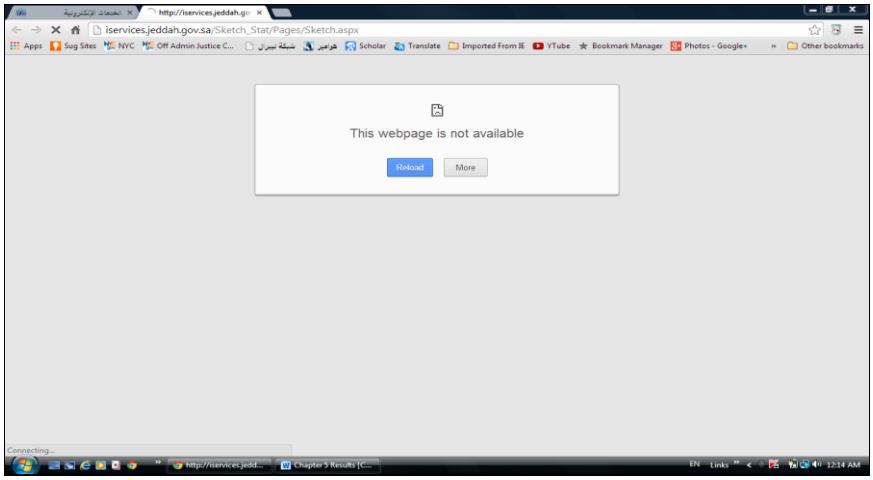

Figure 6.2: Broken Link of Statistics Service

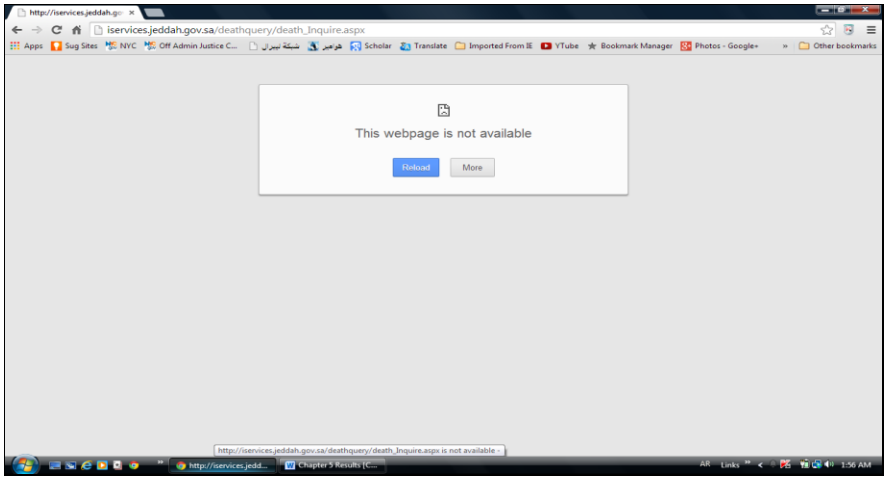

Figure 6.3: Unavailable Deaths Query Service

- Some of the e-services were not available, then after a while they returned online producing incorrect results, such as the deaths query service (Figure 6.3 and Figure 6.4) in which the number of deaths in a period of three months in Jeddah was four deaths only even though the website proclaimed mortality data is updated every 24 hours.

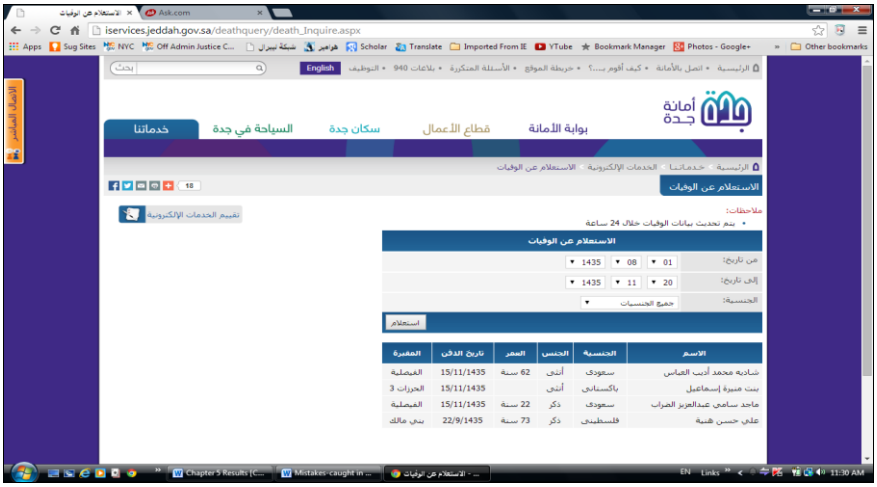

Figure 6.4: Incorrect Results of Deaths Query Service after Reconstruction

- Many of the e-services were a description of procedure services only and occasionally included a link to offline forms such as the store license renewal service (Figure 6.5).

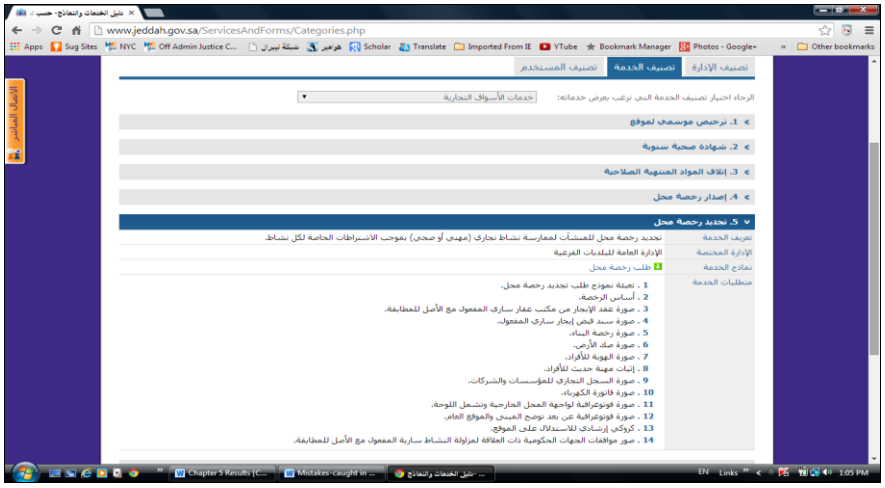

Figure 6.5: Store License Renewal Offline Service

- Jeddah latest news in the media center section was outdated back as far as 2010 (Figure 6.6)

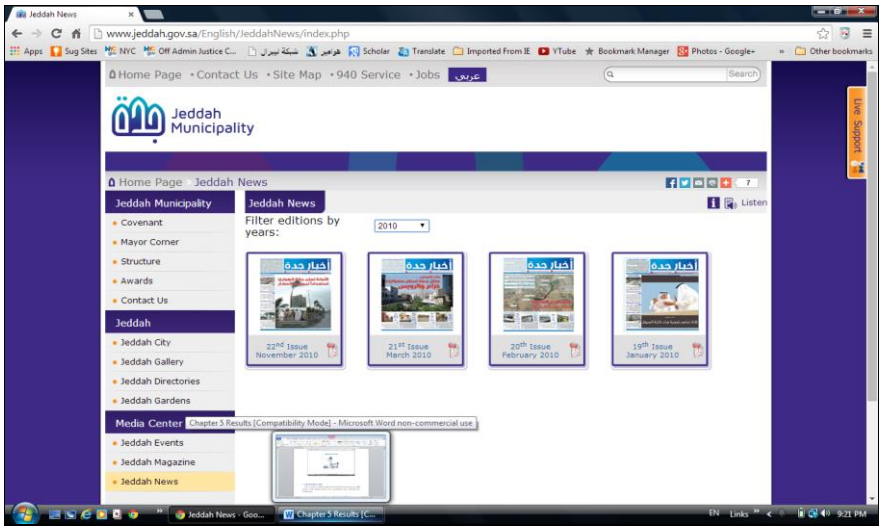

Figure 6.6: Outdated Jeddah News in the Media Center Section

6.1.2 Riyadh Municipal Website

Table 6.2 shows the assessment results of the Riyadh's municipal website by two experts. A total score, a percentage agreement on the evaluation between the experts, and the number of full, partial, and non-compliance guidelines were calculated for the tested website. Experts' general written comments on the site were stated and screen shots of some website errors were captured by the researcher.

Table 6.2: Heuristic Evaluation of Riyadh Municipal Website

	Guideline Number	Expert 1	Expert 2
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	0.5	0.5
2	About us: mayor corner, mission, objectives of the website	0	0
3	Municipal budget information	0	0
4	City council meetings (dates, locations, agendas, minutes)	0	0
5	FAQ with facility to ask new questions	0.5	0.5
6	eJob: job vacancy at municipality	0	0
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	0.5	0.5
8	Last update date on the footer of every page	0	0.5

	Guideline Number	Expert 1	Expert 2
9	Emergency alerts (road closedown, weather alerts…)	0	0
10	Comment or eSuggest on the website	0	0
11	Citizen satisfaction survey	0	0
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	0.5	0.5
13	Multilingual equivalent websites with a link on header of page	0.5	0.5
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1
15	Links to related government websites open in a new window	1	1
16	Downloadable documents/forms with appropriate access	1	1
17	Design for common browsers access (Explorer, Chrome)	1	1
18	Print pages properly	1	1
19	Well-designed customer focused homepage: Quick access to highlighted services through the main menu	0.5	0.5
20	Targeted audience group (citizens, business, tourists...)	0.5	0.5
21	Clear entity "Logo" on every page	1	1
22	Link to homepage from every page through "Home" or logo	1	1
23	Sitemap	1	1
24	Short and descriptive page titles	1	1
25	Readable pages (font, color, background)	1	1
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5
27	Privacy and security statement/policy	1	0.5
28	All links working properly, i.e. no broken links	0.5	0.5
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	1	0.5
30	Enabled "Back button"	0.5	0.5
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	0.5	0.5
33	Navigation and link style	0.5	0.5
Total Score (out of 33 scores)		18 (54.5%)	17.5 (53%)
Percent agreement between experts on evaluation (30 out of 33 guidelines)		≈ 91%	
Agreement on "Non-compliance" (score 0)		8 guidelines (24.2%)	
Agreement on "Partial compliance" (score 0.5)		12 guidelines (36.4%)	
Agreement on "Full compliance" (score 1)		10 guidelines (30.3%)	

The total score for the Riyadh municipal website according to the first expert was 18, out of 33, and that was close to the score assigned by the second expert 17.5 (53%). A high inter-rater reliability of 91% was found between the experts. They agreed on the assessment of 30 guidelines: 8 guidelines of score "0", 12 guidelines of score "0.5", 10 guidelines of score "1" and they disagreed on 3 guidelines. Therefore, the Riyadh municipal website followed 10 out of the 33 guidelines reflecting 30.3% compliance with the proposed objective guidelines. Even if we considered the partially compliant guidelines with a score of "0.5", the adherence rate is still low.

General comments by the experts on the Riyadh municipal website were:

- The e-services provided were limited
- The website was significantly slow
- Problems with the website structure: too many links and some were duplicates such as the e-services for citizen and e-services for organizations and companies. Also, the homepage was multiple pages in length and for Arabic pages, sub menus appeared on the left instead of the right
- The mayor corner was missing public speeches, emails, and contact information
- Rearranging the site through "personalization" makes some website sections disappear
- Some information was outdated such as the press file which displays four files two to four years out of date
- The site lacked important sections: eJob, emergency alerts, and citizen satisfaction survey
- Some web pages had no content or under construction, such as the weather and press release pages.

Screen shots of some website errors were captured:

- Server errors in the center of communication section (Figure 6.7) and also in Riyadh service site events (Figure 6.8).

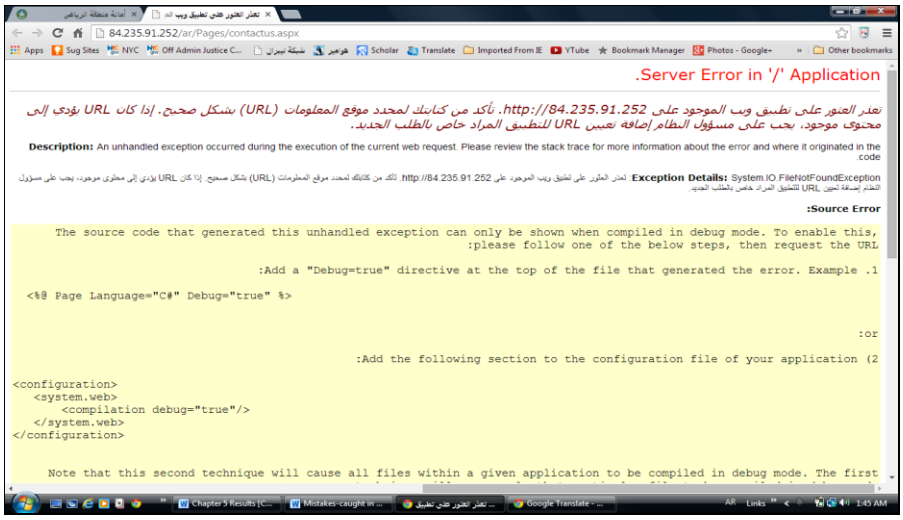

Figure 6.7: Error in the Communication Center

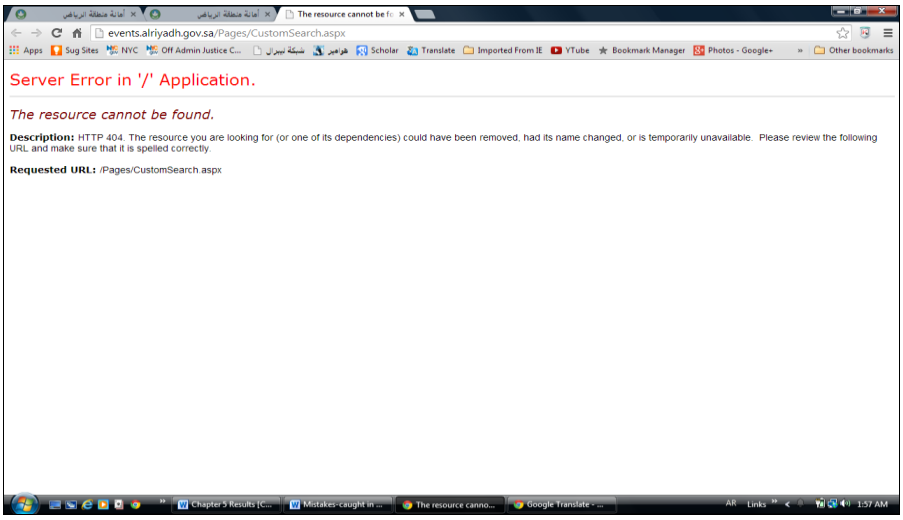

Figure 6.8: Error in Riyadh Events Link

- Link to Riyadh restaurants produced a "not found" error (Figure 6.9)

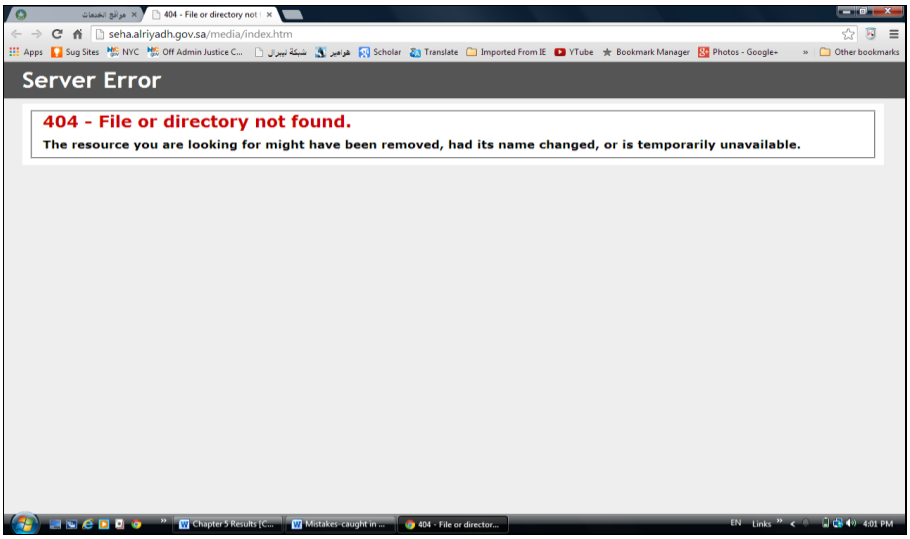

Figure 6.9: File not Found Error in Riyadh Restaurant Link

- Searching about any street in the dictionary of Riyadh's street names produced an unreadable message (Figure 6.10).

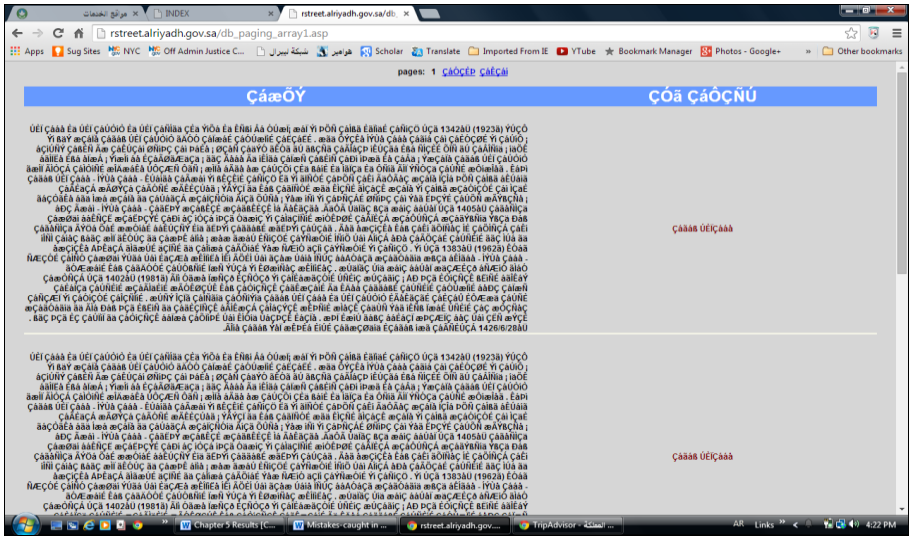

Figure 6.10: Unreadable Riyadh Street Names

- Two links for Riyadh's electronic city map (branched from e-services section) were not working (Figure 6.11) whereas a third link displayed the map.

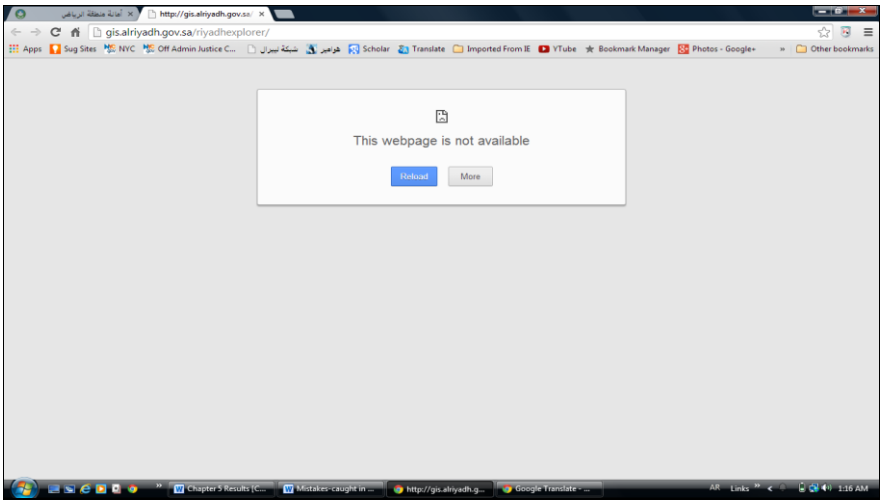

Figure 6.11: Not Found Riyadh Electronic Map

6.1.3 Al-Madinah Municipal Website

Table 6.3 shows the assessment results of the Al-Madinah municipal website by two experts. A total score, a percentage agreement on the evaluation between the experts, and the number of full, partial, and non-compliance guidelines were calculated for the tested website. Experts' general written comments on the site were stated and screen shots of some website errors were captured by the researcher.

Table 6.3: Heuristic Evaluation of Al-Madinah Municipal Website

	Guideline Number	Expert 1	Expert 2
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	0.5	0.5
2	About us: mayor corner, mission, objectives of the website	1	0.5
3	Municipal budget information	0	0
4	City council meetings (dates, locations, agendas, minutes)	0	0
5	FAQ with facility to ask new questions	0	0
6	eJob: job vacancy at municipality	0.5	0.5
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	0.5	0.5
8	Last update date on the footer of every page	0	0
9	Emergency alerts (road closedown, weather alerts…)	0	0
10	Comment or eSuggest on the website	1	1

	Guideline Number	Expert 1	Expert 2
11	Citizen satisfaction survey	1	1
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	0.5	0.5
13	Multilingual equivalent websites with a link on header of page	0.5	0.5
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1
15	Links to related government websites open in a new window	1	1
16	Downloadable documents/forms with appropriate access	1	1
17	Design for common browsers access (Explorer, Chrome)	1	1
18	Print pages properly	1	1
19	Well-designed customer focused homepage: Quick access to highlighted services through the main menu	0.5	0.5
20	Targeted audience group (citizens, business, tourists...)	0	0
21	Clear entity "Logo" on every page	1	1
22	Link to homepage from every page through "Home" or logo	1	0.5
23	Sitemap	1	1
24	Short and descriptive page titles	0.5	0.5
25	Readable pages (font, color, background)	1	1
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5
27	Privacy and security statement/policy	1	1
28	All links working properly, i.e. no broken links	0.5	0.5
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	1	1
30	Enabled "Back button"	0.5	0.5
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0.5
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	1	0.5
33	Navigation and link style	1	0.5
Total Score (out of 33 scores)		21 (63.6%)	19.5 (59.1%)
Percent agreement between experts on evaluation (28 out of 33 guidelines)		≈ 85%	
Agreement on "Non-compliance" (score 0)		6 guidelines (18.2%)	
Agreement on "Partial compliance" (score 0.5)		10 guidelines (30.3%)	
Agreement on "Full compliance" (score 1)		12 guidelines (36.4%)	

The total score for the Al-Madinah municipal website according to the first expert was 21, out of 33, (≈64%) while the second expert's score 19.5 (59%). The inter-rater reliability was 85% indicating both experts agreed, to a large extent, on how to apply the rating scale.

That is, they agreed on the assessment of 28 guidelines: 6 guidelines of score "0", 10 guidelines of "0.5", 12 guidelines of score "1"and they disagreed on 5 guidelines. It was concluded that the Al-Madinah website followed 12 guidelines out of 33, reflecting 36.4% compliance with the proposed objective guidelines.

General comments by the experts on Al-Madinah website were:

- The design of the site was organization centered and not user centered
- The e-services and information provided were limited
- Some e-services were not fully developed
- The e-services section disabled the back button and there was no link to return the user to the home page
- Some news and publications were not updated, such as the municipality magazine section which had only two magazines dating back to more than five years
- The English website was incomplete and some links returned the users to the Arabic site, for example when clicking the English e-service link
- The mayor corner and the mayor office sections were missing
- There were no emergency alerts, no FAQs, and no e-map for Al-Madinah city.

Screen shots of some website errors were captured:

- There were some web pages without content, such as general statistics and general events (Figure 6.12).

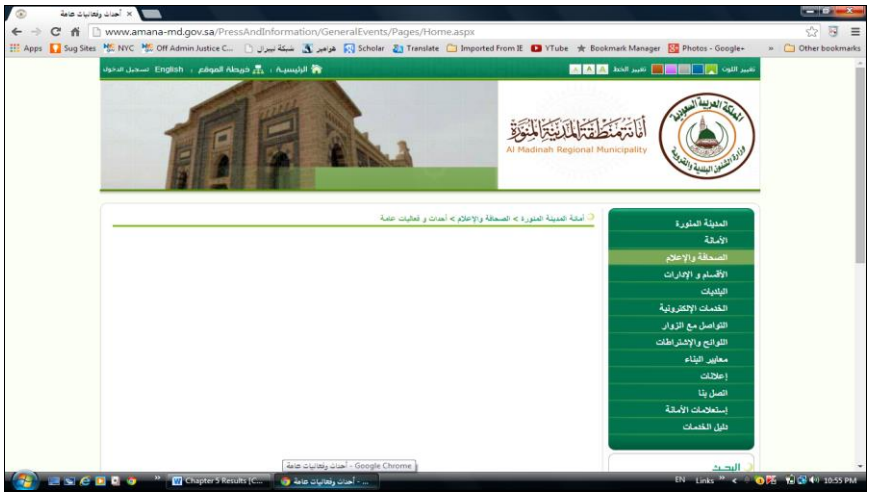

Figure 6.12: No Content Page of General Events

- There were some old news and publications such as the example in Figure 6.13

Figure 6.13: Latest Published Municipality Magazine dated five years ago

- The Mayor office section was not developed (Figure 6.14).

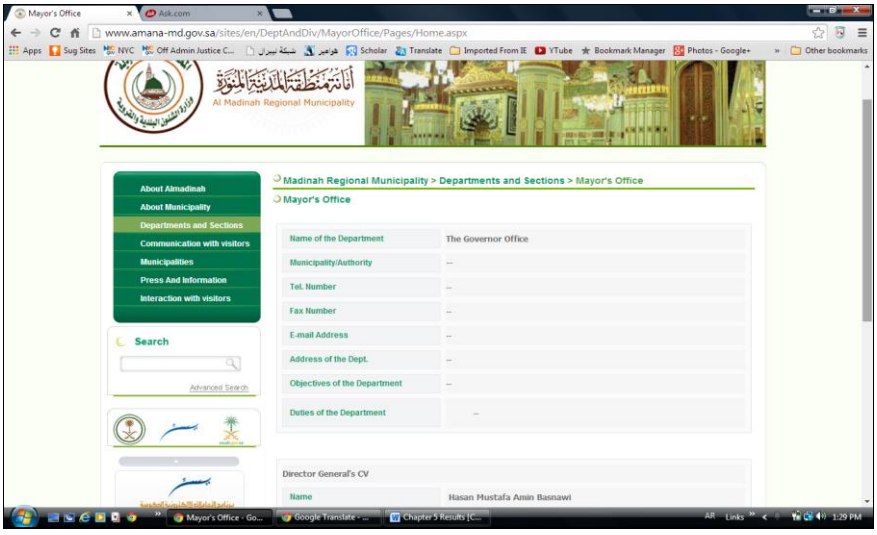

Figure 6.14: Mayor Office still under Construction

- In the English site of Al-Madinah municipality, the mayor's corner page had no content while the Arabic site contains his speech only without any contact information (Figure 6.15).

Figure 6.15: Mayor Corner with no Content Page

- The interactive map link was permanently broken (Figure 6.16)

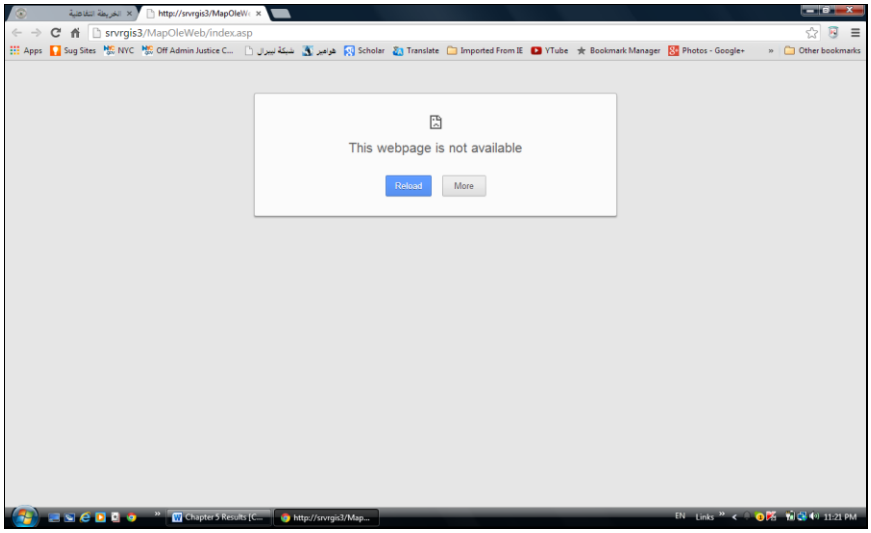

Figure 6.16: Broken Link of Interactive Map

- There was an error downloading the Al-Madinah map: "unable to resolve the server's DNS address" (Figure 6.17) on 10 June 2014.

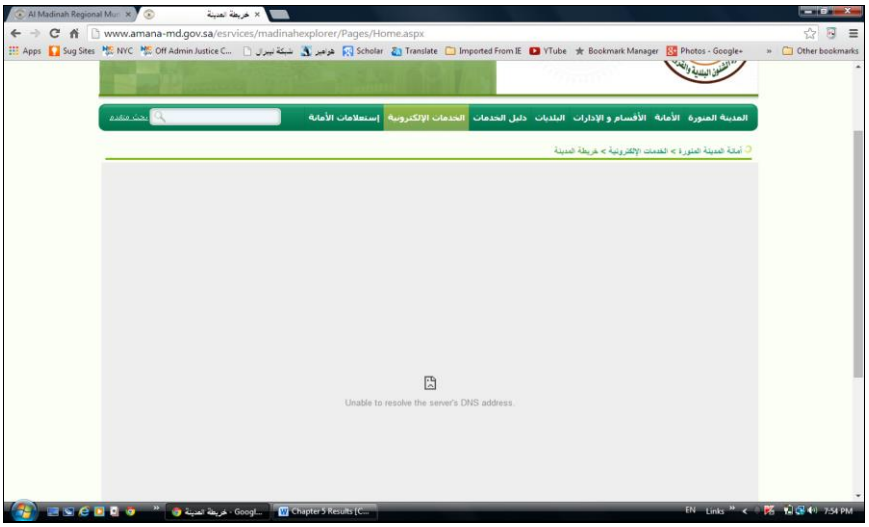

Figure 6.17: Error in Downloading Al-Madinah Map

- The death query service was under development for a long time (Figure 6.18).

Figure 6.18: Apology for unavailable Death Query Service

6.1.4 Eastern Region Municipal Website

Table 6.4 shows the assessment results of evaluating the eastern region municipal website by two experts. A total score, a percentage agreement on the evaluation between the experts, and the number of full, partial, and non-compliance guidelines were calculated. Experts' general written comments on the site were stated and screen shots of some website errors were captured.

Table 6.4: Heuristic Evaluation of Eastern Region Municipal Website

	Guideline Number	Expert 1	Expert 2
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	0.5	0.5
2	About us: mayor corner, mission, objectives of the website	1	0.5
3	Municipal budget information	0	0
4	City council meetings (dates, locations, agendas, minutes)	0	0
5	FAQ with facility to ask new questions	0.5	0.5
6	eJob: job vacancy at municipality	0	0
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	0	0
8	Last update date on the footer of every page	0	0
9	Emergency alerts (road closedown, weather alerts…)	0	0
10	Comment or eSuggest on the website	1	1
11	Citizen satisfaction survey	0	0
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	1	0.5
13	Multilingual equivalent websites with a link on header of page	0	0
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1
15	Links to related government websites open in a new window	1	1
16	Downloadable documents/forms with appropriate access	1	1
17	Design for common browsers access (Explorer, Chrome)	1	1
18	Print pages properly	0.5	0.5
19	Well-designed customer focused homepage: Quick access to highlighted services through the main menu	0.5	0.5
20	Targeted audience group (citizens, business, tourists...)	0	0
21	Clear entity "Logo" on every page	0.5	0.5
22	Link to homepage from every page through "Home" or logo	1	1
23	Sitemap	1	1
24	Short and descriptive page titles	1	1
25	Readable pages (font, color, background)	0.5	1
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5
27	Privacy and security statement/policy	1	1
28	All links working properly, i.e. no broken links	0.5	0.5

	Guideline Number	Expert 1	Expert 2
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	0	0
30	Enabled "Back button"	1	1
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	0.5	0.5
33	Navigation and link style	0.5	0.5
Total Score (out of 33 scores)		17 (51.5%)	16.5 (50%)
Percent agreement between experts on evaluation (30 out of 33 guidelines)		≈ 91%	
Agreement on "Non-compliance" (score 0)		11 guidelines (33.3%)	
Agreement on "Partial compliance" (score 0.5)		9 guidelines (27.3%)	
Agreement on "Full compliance" (score 1)		10 guidelines (30.3%)	

The total score for the eastern region municipal website according to the first expert was 17, out of 33, while the second expert's score 16.5, (50%). A high inter-rater reliability of 91% was found between the experts, indicating they agreed, to a large extent, on how to apply the rating scale. That is, they agreed on the assessment of 30 guidelines: 11 guidelines of score "0", 9 guidelines of score "0.5", 10 guidelines of score "1" and they disagreed on 3 guidelines. Therefore, the eastern region municipal website obtained low adherence rate by following 10 out of the 33 guidelines reflecting 30.3% compliance with the proposed objective guidelines. Even if we considered the partial compliance guidelines, the total adherence is still low.

General comments by the experts on the eastern region website were:

- There was no English website even though it is a must in Saudi e-government
- The e-services and information offered were limited
- The structure of the site was disorganized
- The design of the site was organization centered and not user centered

- Inconsistency in the design of some webpages; example the "cultural and tourist guide"
- Some links were either broken or inactive in the internal pages
- A malfunctioned search facility with a message: “unable to connect to the search query"
- No indicator for where the user is on the site
- There were no emergency alerts, no e-Job, no FAQs and no e-city map.

Screen shots of some website errors were captured:

- Broken links were encountered on the site; as an example "the indicative guide" link and "e-job" link which also changes the municipal web address to another unrelated URL: http://www.ekram.sa/careers as shown in Figure 6.19.

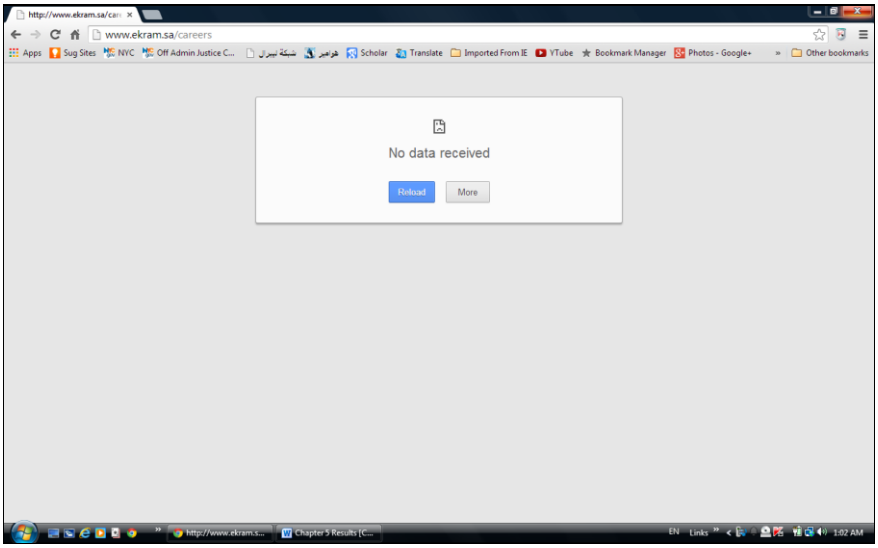

Figure 6.19: Broken Link of E-job Section

- The search function was not functioning properly displaying the message: "unable to connect to the search query" (Figure 6.20).

Figure 6.20: Malfunction of Search Services

- Some pages were under construction, either empty with titles only (Figure 6.21) or apologizing for unavailable content (Figure 6.22), such as agency services and municipal projects links.

Figure 6.21: No Content of Agency Services Page

Figure 6.22: Apology for Under Construction in the Municipal Projects Link

- Only three questions were in the Frequently Answered Questions (FAQs) (Figure 6.23).

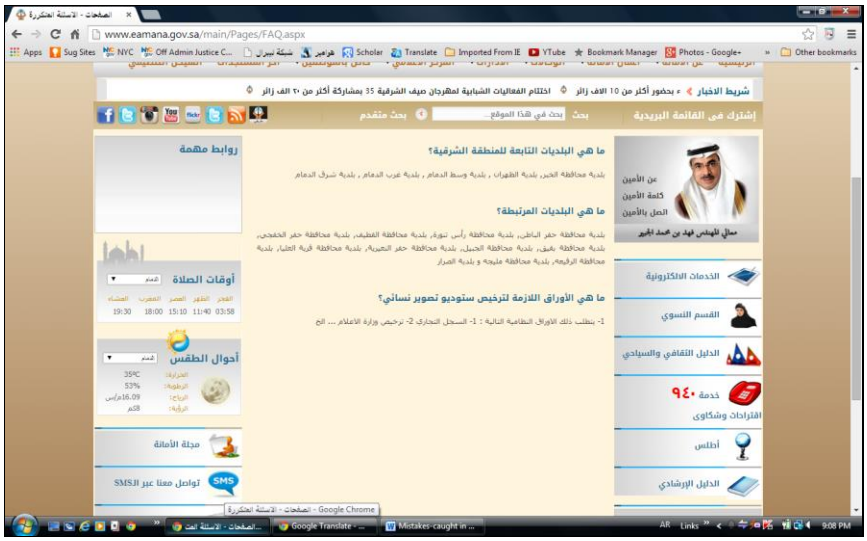

Figure 6.23: Few Questions in the FAQs Section

- Visitors poll about the site were: excellent 26.16%, very good 9.3%, bad 64.53% (Figure 6.24)

Figure 6.24: Visitors Poll about the Eastern Municipal Website

6.1.5 Qassim Municipal Website

Table 6.5 shows the assessment results of evaluating the Qassim municipal website by two experts. A total score, a percentage agreement on the evaluation between the experts, and the number of full, partial, and non-compliance guidelines were calculated. Experts' general written comments on the site were stated and screen shots of some website errors were captured.

Table 6.5: Heuristic Evaluation of Qassim Municipal Website

	Guideline Number	Expert	Researcher
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	1	0.5
2	About us: mayor corner, mission, objectives of the website	0.5	0.5
3	Municipal budget information	0	0
4	City council meetings (dates, locations, agendas, minutes)	0	0.5
5	FAQ with facility to ask new questions	0.5	0.5
6	eJob: job vacancy at municipality	0	0
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	1	0.5

8	Last update date on the footer of every page	1	1
9	Emergency alerts (road closedown, weather alerts…)	0	0
10	Comment or eSuggest on the website	1	1
11	Citizen satisfaction survey	0	1
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	0.5	0.5
13	Multilingual equivalent websites with a link on header of page	0	0
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1
15	Links to related government websites open in a new window	1	1
16	Downloadable documents/forms with appropriate access	1	1
17	Design for common browsers access (Explorer, Chrome)	1	1
18	Print pages properly	1	1
	Well-designed customer focused homepage:		
19	Quick access to highlighted services through the main menu	0.5	0.5
20	Targeted audience group (citizens, business, tourists...)	0.5	0.5
21	Clear entity "Logo" on every page	0.5	0.5
22	Link to homepage from every page through "Home" or logo	1	1
23	Sitemap	1	0.5
24	Short and descriptive page titles	1	1
25	Readable pages (font, color, background)	1	1
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5
27	Privacy and security statement/policy	1	1
28	All links working properly, i.e. no broken links	1	1
	Navigational options:		
29	Indicator of a user is where on the site (e.g. Breadcrumbs)	0.5	0.5
30	Enabled "Back button"	1	1
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0
	Consistent design of all web pages:		
32	Page layout (same feel and look, font, color, buttons, menus)	1	1
33	Navigation and link style	1	1
Total Score (out of 33 scores)		22 (66.7%)	22 (66.7%)
Percent agreement between experts on evaluation (28 out of 33 guidelines)		≈ 85%	
Agreement on "Non-compliance" (score 0)		5 guidelines (15.2%)	
Agreement on "Partial compliance" (score 0.5)		8 guidelines (24.2%)	
Agreement on "Full compliance" (score 1)		15 guidelines (45.5%)	

The total score for the Qassim municipal website according to both experts was 22, out of 33, (66.7%). An inter-rater reliability of 85% was found between the experts, indicating

they agreed, to a large extent, on how to apply the rating scale. That is, they agreed on the assessment of 28 guidelines: 5 guidelines of score "0", 8 guidelines of score "0.5", 15 guidelines of score "1" and they disagreed on 5 guidelines. Therefore, the Qassim website followed 15 out of the 33 guidelines reflecting 45.5% compliance with the proposed objective guidelines.

General comments by the experts on the Qassim municipal website were:

- There was no English website even though it is a must in Saudi e-government
- The homepage had too many links and some internal webpages were multiple pages in length
- The site suffered from link design problems: the main menu links were confused with another menu links on the right side of the page and the grouping of links was not appropriate
- Excessive advertisements were found on the site which might put its credibility as a government website in question
- The e-services offered were limited
- There was no mayor corner; just information on "contact us". The provided mayor email produced "404 not found" error
- There were no emergency alerts and no e-Job sections
- There was no distinction between visited and unvisited links on the site.
- The scroll bar was not positioned on the right side as it is supposed to be for Arabic websites.

Screen shots of some website errors were captured:

- The behavior of the website was not reliable. Sometimes the website loaded correctly while at other times the not found error was displayed (Figure 6.25)

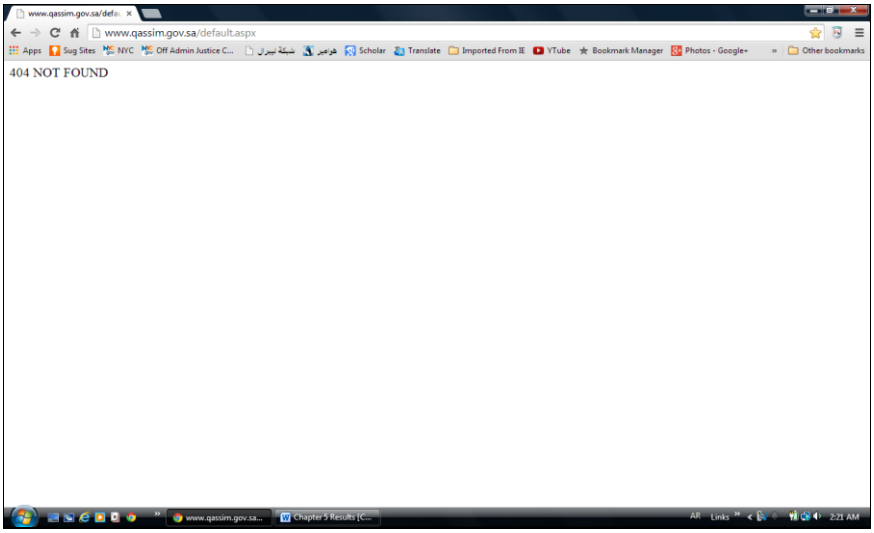

Figure 6.25: Malfunction of Qassim Website

- Many pages were under construction such as: pictures' album, all directories (emergency, government, tourist, etc.) (Figure 6.26), and "how do I do" service (Figure 6.27)

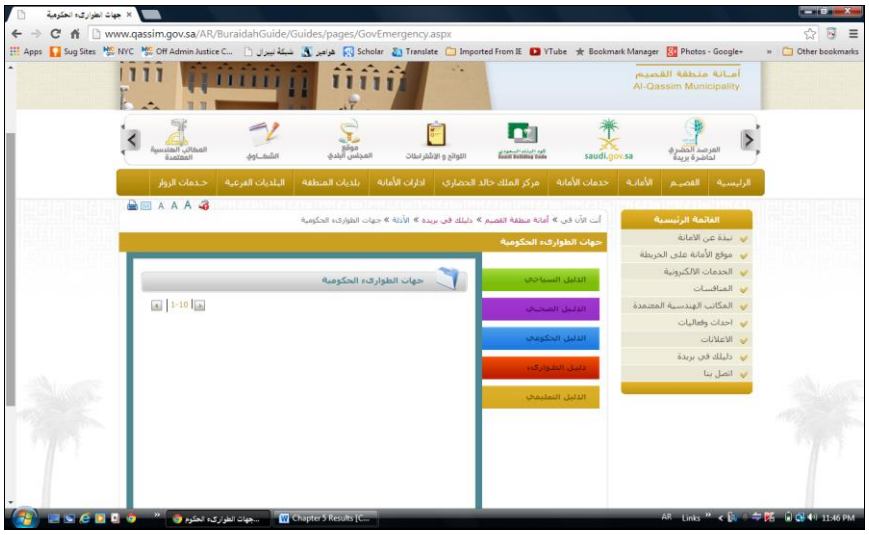

Figure 6.26: Emergency Directory under Construction

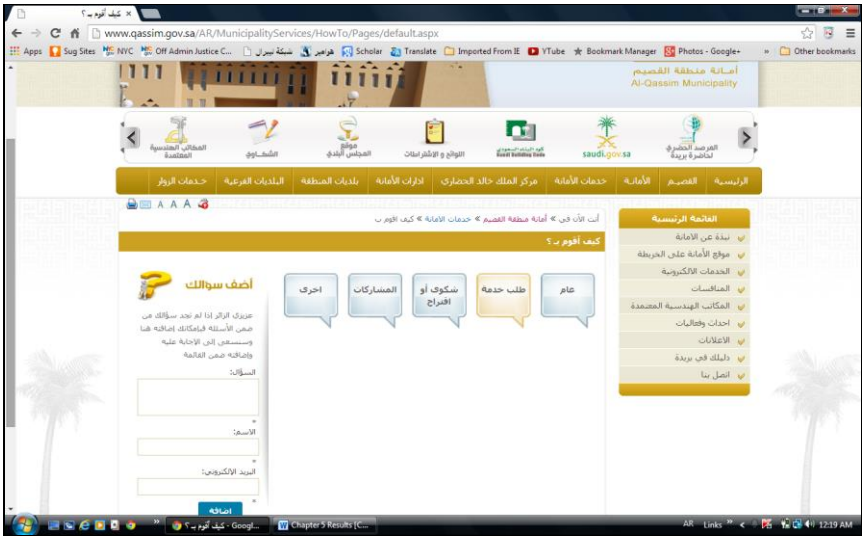

Figure 6.27: "How Do I Do" Service under Construction

- Some forms, such as an offline form "license for adding floors", were not well designed (Figure 6.28).

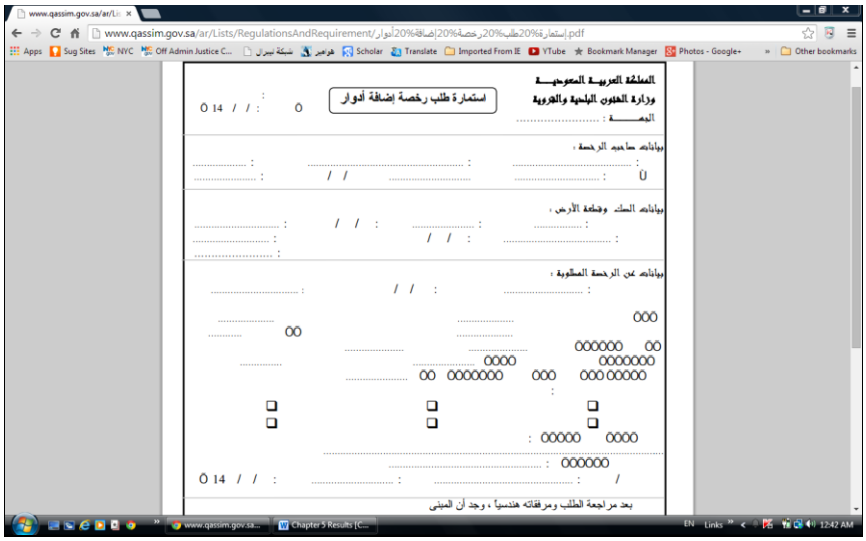

Figure 6.28: Not Well Designed Form

- The poll about visitors' satisfaction with the offered e-services via the Qassim municipal website: Totally satisfy with the e-services 29%, satisfy somewhat 6%, not satisfied 58%, and not seen 7%. (Figure 6.29)

Figure 6.29: Dissatisfaction with E-services in Visitors' Poll

6.1.6 Combined Results of Heuristic Test on Saudi Municipal Websites

This section combines previous heuristic test results in order to have a summarized view to compare the performance of the selected e-city websites. From the perspective of the two experts, Table 6.6 summarizes and ranks Saudi municipal websites based on the total score earned and also provides the total number of full compliance guidelines. Even though there were some differences in their scores, the two experts reached the same conclusion on their ranking of the e-municipalities: Jeddah, Qassim, Al-Madinah, Riyadh, and finally the Eastern Region. The top Saudi e-city "Jeddah" earned 26 points (78.8%), complied with 25 guidelines (75.8%) for the first expert, and earned 24.5 points (74.2%), complied with 21 guidelines (63.3%) for the second expert. The same score of 22 (66.7%) was given by both experts to the second e-city "Qassim" with a compliant rate of 54.5% and 48.5% respectively. The last e-municipality, Eastern Region, earned 17 points (51.5%), complied with 12 guidelines (36.4%) by the first expert, and earned 16.5 points (50%) complied with 11 guidelines (33.3%) by the second expert.

Table 6.6: Heuristic Evaluation by Two Experts Ranking E-municipalities on Total Score

Rank	Expert	City Site	Total Score	Score Percentage	Total Number of Full Compliant Guidelines	Percentage of Compliance
First site	1	Jeddah	26/33	78.8%	25/33	75.8%
	2		24.5/33	74.2%	21/33	63.3%
Second site	1	Qassim	22/33	66.7%	18/33	54.5%
	2		22/33	66.7%	16/33	48.5%
Third site	1	Al-Madinah	21/33	63.6%	16/33	48.5%
	2		19.5/33	59.1%	12/33	36.4%
Fourth Site	1	Riyadh	18/33	54.5%	12/33	36.4%
	2		17.5/33	53%	10/33	30.3%
Fifth Site	1	Eastern Region	17/33	51.5%	12/33	36.4%
	2		16.5/33	50%	11/33	33.3%

Table 6.7 presents the total number of fully or partially compliant guidelines as agreed upon by the two experts. In this regard, the top municipal website was based on the number of full compliance guidelines; if these are equal numbers then the number of partially compliant guidelines was considered. Therefore, the order of the e-municipalities ranking was: Jeddah followed 20 guidelines (60.6%), Qassim 15 guidelines (45.5%), Al-Madinah 12 guidelines (36.4%), Riyadh 10 guidelines (30.3%), the same as Eastern Region but higher in the number of partial compliant guidelines. The rank of e-cities based on the total score was the same as the order based on the total number of agreed full compliant guidelines.

Table 6.7: Heuristic Evaluation Ranking E-municipalities on Full Compliance Guidelines

	City Site	Total Number of Agreed Full Compliant Guidelines	Total Number of Agreed Partially Compliant Guidelines
1	Jeddah	20/33 (60.6%)	2/33 (6.1%)
2	Qassim	15/33 (45.5%)	8/33 (24.2%)
3	Al-Madinah	12/33 (36.4%)	10/33 (30.3%)
4	Riyadh	10/33 (30.3%)	12/33 (36.4%)
5	Eastern Region	10/33 (30.3%)	9/33 (27.3%)

None of the websites tested could be considered fully compliant with the proposed guidelines. In fact, all municipalities obtained a low adherence rate and that generally might reflect many usability problems identified during the expert testing. Table 6.8 provides important guidelines that municipalities did not consider (those scored 0 by the two experts) when designing their websites and that may indicate possible usability problems. Jeddah and Qassim did not comply with five guidelines, Al-Madinah with six, Riyadh with eight and Eastern Region with eleven guidelines. The common guidelines ignored by all five Saudi city websites were: "municipal budget information, city council meetings, emergency alerts, and different colors for visited/unvisited links", which are important to citizens and a basic component of a city website. When there is not much information on city council meetings or no emergency alerts services, citizens lose the benefits of e-government. Likewise, if there is no disclosure about municipal budget, citizens lose one of the expected outcomes from e-government, "transparency and no corruption".

Table 6.8: Heuristic Results of No Compliance Guidelines

City Site	Agreed upon Number of No Compliance Guidelines	No Compliance Guidelines
Jeddah	5/33 (15.2%)	1. Municipal budget information 2. City council meetings 3. Last update on the footer of page 4. Emergency alerts 5. Different colors for visited/unvisited links, etc.
Qassim	5/33 (15.2%)	1. Municipal budget information 2. eJob 3. City council meetings 4. Emergency alerts 5. Multilingual equivalent websites 6. Different colors for visited/unvisited links, etc.
Al-Madinah	6/33 (18.2%)	1. Municipal budget information 2. City council meetings 3. FAQs 4. Last update on the footer of page 5. Emergency alerts 6.Targeted audience group (citizen, business, …)
Riyadh	8/33 (24.2%)	1.About us: mayor corner, mission, and objectives 2. Municipal budget information 3. City council meetings 4. eJob 5. Emergency alerts 6. Comment or eSuggest 7. Citizen satisfaction survey 8. Different colors for visited/unvisited links, etc.

Eastern Region	11/33 (33.3%)	1. Municipal budget information 2. City council meetings 3. eJob 4. News important to users 5. Last update on the footer of page 6. Emergency alerts 7. Citizen satisfaction survey 8. Multilingual equivalent websites 9.Targeted audience group (citizen, business, …) 10. Indicator of a user is where on the site 11.Different colors for visited/unvisited links, etc.

Finally based on experts' comments, the most common problems encountered in Saudi municipal websites are as follows:

- The design of the sites was organization centered and not user centered
- Limited e-services and information were offered to the public
- Most of the e-services were not online transactions, but descriptive procedures of the services which can be categorized as "offline services"
- There were no emergency alerts to the residence, no e-Job, and nor any distinction between visited and unvisited links on the site
- An English website was either not provided or was an incomplete version
- Some news and publications were outdated
- Some pages had no content or under construction
- Many broken links were encountered.

6.2 User Testing of Saudi Municipal Websites

For each Saudi e-city website the total of G2C e-services, malfunctioning e-services, functioning e-transaction and e-transformation services were counted manually, as shown in Table 6.9. For example, the Jeddah e-city website had 25 transactional e-services (100% functioning), while the Riyadh e-city website had 39 e-services but eight of them not working (e.g. broken link, had no content, or under construction e-service page), therefore it actually had 31 transactional e-services (79.5% functioning).

Table 6.9: Number and Type of E-services for Five Saudi City Websites

Municipality	#G2C E-services	#Malfunctioning E-services	Type of Functioning E-service		% Functioning E-services
			#Transaction	#Transformation	
Jeddah	25	0	25	0	100%
Riyadh	39	8	31	0	79.5%
Al-Madinah	21	5	16	0	76.2%
Eastern Region	18	6	12	0	66.7%
Qassim	11	3	8	0	72.7%

Based on the discussion in section 5.2.2.2, user testing of Saudi e-city websites evaluated three parts: 1) website quality, 2) e-services quality, and 3) user satisfaction questionnaire. To provide an overview of website quality, the results of user testing will be compiled with the results of heuristic testing. For the user satisfaction questionnaire, it was found that no test user had used any of the five Saudi city websites over the previous year, nor had any of the test users depended on it for the completion of their municipal transactions. Moreover, the user testing was carried out with users until the encountered problems started to repeat. It was found that after five users no additional data was obtained, so the test was stopped after user number six. Table 6.10 shows the profile of the six users with the test date.

Table 6.10: Users Profile and Date of User Testing

User	Gender	Age	Major	Date of Test
1	F	40-49	Mathematics	11, 12 /2/2015
2	M	20-29	Engineering	13, 14 /2/2015
3	F	30-39	Hospital Administration	15, 16 /2/2015
4	F	20-29	Economy	17, 18 /2/2015
5	M	40-49	Communication	19, 20 /2/2015
6	M	30-39	Administration	22, 23 /2/2015

6.2.1 Jeddah Municipal Website

1. Website Quality

Table 6.11 shows the results of user testing six website quality guidelines (#34 to 39 of the proposed e-city framework). The average score of user testing for the website's quality

was 2.46, out of 6, (41%). Only three guidelines (1, 4, & 6) scored 50%. The lowest average score was 0.13 and 0.29, out of 1, indicating the Jeddah website had a problem in its interactive city map and eComplaint.

Table 6.11: User Testing of Website Quality for Jeddah Municipal Website

Criteria		Score						
		User 1	User 2	User 3	User 4	User 5	User 6	Average
1	Useful and most up to date content	1	0.75	0.5	0.25	0.5	0.5	**0.58**
2	eComplaint	0.75	0.5	0	0.5	0	0	**0.29**
3	Interactive city map	0.25	0	0	0.5	0	0	**0.13**
4	Meaningful images and videos	1	0.5	0.5	0.75	0.25	0.25	**0.54**
5	Effective search	0.75	0.75	0.25	0.25	0.25	0.25	**0.42**
6	Organized short meaningful link labels	0.75	0.5	0.25	0.5	0.5	0.5	**0.50**
Sum		4.5	3	1.5	2.75	1.5	1.5	**2.46**

When compiling the results from the user testing and the heuristic tests it provides the overall assessment of the websites quality. The average score of the two experts for Jeddah was 25.25 and the average of user testing was 2.46. Therefore, the quality of Jeddah municipality website was 27.71 out of 39 (71.1%) based on the heuristic and user testing (Table 6.12).

Table 6.12: Compiled Website Quality of Jeddah Municipal Website

	Guideline Number	Expert1	Expert2	Average User Testing
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	1	1	
2	About us: mayor corner, mission, objectives of the website	1	1	
3	Municipal budget information	0	0	
4	City council meetings (dates, locations, agendas, minutes)	0	0	
5	FAQ with facility to ask new questions	0.5	0.5	
6	eJob: job vacancy at municipality	1	1	
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	1	0.5	
8	Last update date on the footer of every page	0	0	
9	Emergency alerts (road closedown, weather alerts…)	0	0	
10	Comment or eSuggest on the website	1	1	
11	Citizen satisfaction survey	1	1	
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	1	1	
13	Multilingual equivalent websites with a link on header of page	1	0.5	

	Guideline Number	Expert1	Expert2	Average User Testing
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1	
15	Links to related government websites open in a new window	1	1	
16	Downloadable documents/forms with appropriate access	0	1	
17	Design for common browsers access (Explorer, Chrome)	1	1	
18	Print pages properly	1	1	
19	Well-designed customer focused homepage: Quick access to highlighted services through the main menu	1	1	
20	Targeted audience group (citizens, business, tourists...)	1	1	
21	Clear entity "Logo" on every page	1	1	
22	Link to homepage from every page through "Home" or logo	1	1	
23	Sitemap	1	1	
24	Short and descriptive page titles	1	1	
25	Readable pages (font, color, background)	1	1	
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5	
27	Privacy and security statement/policy	1	1	
28	All links working properly, i.e. no broken links	1	0.5	
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	1	1	
30	Enabled "Back button"	1	1	
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0	
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	1	0.5	
33	Navigation and link style	1	0.5	
34	Useful and most up to date content			0.58
35	eComplaint			0.29
36	Interactive city map			0.13
37	Meaningful images and videos			0.54
38	Effective search			0.42
39	Organized short meaningful link labels			0.50
Total Expert Score (out of 33)		26 (78.8%)	24.5 (74.2%)	
Average Score of the two Experts (out of 33)		25.25 (76.5%)		
Average of User Testing (out of 6)				2.46 (41%)
Total Score for Website Quality (Experts+ User Testing) (out of 39)		27.71 (71.1%)		

2. E-services Quality

The six users were given five usability tasks to evaluate the quality of the e-services on the Jeddah city website. The testing result from user #1 is presented as an example in Table

6.13. Notice that this user quit the first task, consequently, a score of zero was given to that task. The score for each e-service was calculated by multiplying the average quality of the service by 3 (transaction score), as described in section 4.4. According to user # 1, the total score for all tasks tested in the Jeddah city website was 7.07 (out of 15).

Table 6.13: Total Score of User 1 Testing E-services in Jeddah Municipal Website

User 1	
Task 1: View building regulations in Basateen district	
Guidelines	Score (out of 1)
1.Sufficient information on e-services	Quit: unable to complete the task
2. Ease of navigation through the e-service process	
3.Feedback when users waiting or submitting a request	
4. E-services completely online if possible	
Sum	
Avg quality of service	
Score for e-service # 1 =Avg*3	
Task 2: Inquire about Saudis deaths during a period of two months and also for one year	
Guidelines	Score (out of 1)
1.Sufficient information on e-services	0
2. Ease of navigation through the e-service process	0.25
3.Feedback when users waiting or submitting a request	1
4. E-services completely online if possible	1
Sum	2.25
Avg quality of service	0.56
Score for e-service # 2 =Avg*3	1.69
Task 3: Report online about a drilling in King road, Marine Science square	
Guidelines	Score (out of 1)
1.Sufficient information on e-services	0.25
2. Ease of navigation through the e-service process	0
3. Simple forms with required fields and clear error message for invalid or incomplete data entry	0
4. Online tracking for forms and e-services being processed	0
5.Feedback when users waiting or submitting a request	0
6. E-services completely online if possible	0
Sum	0.25
Avg quality of service	0.04
Score for e-service # 3 =Avg*3	0.13

Task 4: Inquire licenses shops transaction (# 320872)	
Guidelines	Score (out of 1)
1.Sufficient information on e-services	0.75
2. Ease of navigation through the e-service process	1
3.Feedback when users waiting or submitting a request	1
4. E-services completely online if possible	1
Sum	3.75
Avg quality of service	0.94
Score for e-service # 4 =Avg*3	2.81
Task 5: Inquire health centers in Basateen and Faysaleyyah districts	
Guidelines	Score (out of 1)
1.Sufficient information on e-services	1
2. Ease of navigation through the e-service process	0.25
3.Feedback when users waiting or submitting a request	1
4. E-services completely online if possible	1
Sum	3.25
Avg quality of service	0.81
Score for e-service # 5 =Avg*3	2.44
Total Score for all tasks (out of 15)	**7.07**

Table 6.14 depicted the overall testing results of six users. The total scores for all tasks evaluated by the six users were: 7.07, 8.94, 4.50, 6.50, 3.70 and 3.32 respectively. The average for all tasks performed on the Jeddah city website was 5.67 (37.8%). Notice that all users did not complete task 1 because they could not find the link to the building regulation e-service. User 6 quit task 2 also and graded service quality low similar to users 5 and 3, while users 1, 2, and 4 scores were better. This variation is expected in subjective evaluations and how patient or impatient a participant with a task depends on the type of his\her personality, which must be respected.

Table 6.14: Total Score for Six Users Testing E-services in Jeddah Municipal Website

Services	User 1	User 2	User 3	User 4	User 5	User 6
1: View building regulations in Basateen district	0	0	0	0	0	0
2: Inquire about Saudis deaths during a period of two months and also for one year	1.69	2.63	2.25	1.5	0.75	0
3: Report online about a drilling in King road, Marine Science square	0.13	0.5	0	0.88	0.13	0.5
4: Inquire licenses shops transaction (# 320872)	2.81	3	0	2.06	1.69	1.13
5: Inquire health centers in Basateen and Faysaleyyah districts	2.44	2.81	2.25	2.06	1.13	1.69
Total Score for all Tasks (out of 15)	7.07	8.94	4.50	6.50	3.70	3.32
Average for all Tasks (out of 15)	**5.67**					
Percentage of Average for all Tasks	**37.8%**					

3. User Satisfaction Survey

At the end of the test, the six users were asked to answer a survey of six questions (see Figure 5.8). As shown in Table 6.15, users did not grade Jeddah as an excellent site on any of the six evaluated features. In fact, 50% of users considered the quality of the Jeddah website as good and 5o% as fair. As for e-services quality, 83.3% of users considered it fair while 16.7% considered it poor. 50% rated the ease of using the site fair, 33.3% poor, while 16.7% good. The number of e-services was considered fair by 66.7% of users, 16.7% good, while 16.7% poor.

Table 6.15: User Satisfaction Survey of Jeddah Municipal Website

Features	Excellent	Good	Fair	Poor	None
Website quality		50%	50%		
E-services quality			83.3%	16.7%	
Ease of use		16.7%	50%	33.3%	
Number of e-services		16.7%	66.7%	16.7%	
E-services performed this year					100%
Extent of depending on portal e-services					100%

6.2.2 Riyadh Municipal Website

1. Website Quality

Table 6.16 shows the results of user testing six website quality guidelines (#34 to 39 of the proposed e-city framework). The average score of user testing for the website's quality was as low as 1.52, out of 6, (25.3%). Unfortunately, the highest score was 0.42. The lowest average scores were, 0.13 (twice), 0.21, and 0.25, out of 1, (in guidelines 2, 3, 4, & 5), indicating the Riyadh website has a problem in eComplaint, interactive city map, images and videos, and the search facility respectively.

Table 6.16: User Testing of Website Quality for Riyadh Municipal Website

Criteria		Score						
		User 1	User 2	User 3	User 4	User 5	User 6	Average
1	Useful and most up to date content	0.5	0.5	0.5	0	0.25	0.5	**0.38**
2	eComplaint	0	0	0.25	0.5	0	0	**0.13**
3	Interactive city map	0	0	0	0	0	0.75	**0.13**
4	Meaningful images and videos	0	0.5	0.25	0.25	0.25	0	**0.21**
5	Effective search	0	0.25	0.25	0.25	0.25	0.5	**0.25**
6	Organized short meaningful link labels	0.25	0.5	0.5	0.5	0.25	0.5	**0.42**
Sum		0.75	1.75	1.75	1.5	1	2.25	**1.52**

When compiling the results from the user testing and the heuristic tests it provides the overall assessment of website quality. The average score of the two experts was 17.75 and the average of user testing was 1.52. Therefore, the quality of Riyadh municipality website was 19.27 out of 39 (49.4%) based on heuristic and user testing (Table 6.17).

Table 6.17: Compiled Website Quality of Riyadh Municipal Website

	Guideline Number	Expert1	Expert2	Average User Testing
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	0.5	0.5	
2	About us: mayor corner, mission, objectives of the website	0	0	
3	Municipal budget information	0	0	
4	City council meetings (dates, locations, agendas, minutes)	0	0	
5	FAQ with facility to ask new questions	0.5	0.5	
6	eJob: job vacancy at municipality	0	0	
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	0.5	0.5	
8	Last update date on the footer of every page	0	0.5	

	Guideline Number	Expert1	Expert2	Average User Testing
9	Emergency alerts (road closedown, weather alerts…)	0	0	
10	Comment or eSuggest on the website	0	0	
11	Citizen satisfaction survey	0	0	
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	0.5	0.5	
13	Multilingual equivalent websites with a link on header of page	0.5	0.5	
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1	
15	Links to related government websites open in a new window	1	1	
16	Downloadable documents/forms with appropriate access	1	1	
17	Design for common browsers access (Explorer, Chrome)	1	1	
18	Print pages properly	1	1	
19	Well-designed customer focused homepage: Quick access to highlighted services through the main menu	0.5	0.5	
20	Targeted audience group (citizens, business, tourists...)	0.5	0.5	
21	Clear entity "Logo" on every page	1	1	
22	Link to homepage from every page through "Home" or logo	1	1	
23	Sitemap	1	1	
24	Short and descriptive page titles	1	1	
25	Readable pages (font, color, background)	1	1	
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5	
27	Privacy and security statement/policy	1	0.5	
28	All links working properly, i.e. no broken links	0.5	0.5	
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	1	0.5	
30	Enabled "Back button"	0.5	0.5	
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0	
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	0.5	0.5	
33	Navigation and link style	0.5	0.5	
34	Useful and most up to date content			0.38
35	eComplaint			0.13
36	Interactive city map			0.13
37	Meaningful images and videos			0.21
38	Effective search			0.25
39	Organized short meaningful link labels			0.42
Total Expert Score (out of 33)		18 (54.5%)	17.5 (53%)	
Average Score of the two Experts (out of 33)		17.75 (53.8%)		
Average of User Testing (out of 6)				1.52 (25.3%)
Total Score for Website Quality (Experts + User Testing) (out of 39)		19.27 (49.4%)		

2. E-services Quality

The six users were given five usability tasks to evaluate the quality of e-services in the Riyadh city website (see Appendix A). The testing result and the calculation were similar to the example given previously in Jeddah. As shown in Table 6.18, the total scores for all tasks evaluated by the six users were: 6.01, 9.19, 7.89, 8.45, 3.57 and 6.75 respectively. The average for all tasks performed on the Riyadh city website was 6.98 (46.5%).

Table 6.18: Total Score for Six Users Testing E-services in Riyadh Municipal Website

Services	User 1	User 2	User 3	User 4	User 5	User 6
1: Inquire about your transaction # 12345 dated 1435h	1.69	2.25	2.25	2.25	1.88	2.06
2: Inquire about Saudis deaths during a period of two months and also for one year	1.69	2.25	1.13	1.69	1.13	0
3: Know the engineering offices of building permits then search for "Knooz"	1.13	1.69	1.88	1.88	0	0.75
4: Inquire about healthy monitor # 3	0.75	2.25	1.88	1.88	0	2.06
5: View health certificates for your workers	0.75	0.75	0.75	0.75	0.56	1.88
Total Score for all Tasks (out of 15)	6.01	9.19	7.89	8.45	3.57	6.75
Average for all Tasks (out of 15)	**6.98**					
Percentage of Average for all Tasks	**46.5%**					

3. User Satisfaction Survey

At the end of the test, the six users were asked to answer a survey of six questions. As shown in Table 6.19, users did not grade Riyadh as an excellent site on any of the six evaluated features. Users equally graded the quality of the Riyadh website as good, fair and poor. As for e-services quality, 50% of users considered it poor, 33.3% fair, and 16.7% good. Half of tested users rated the ease of use as fair and half as good. The number of e-services was considered fair by 83.3% of users while 16.7% good.

Table 6.19: User Satisfaction Survey of Riyadh Municipal Website

Features	Excellent	Good	Fair	Poor	None
Website quality		33.3%	33.3%	33.3%	
E-services quality		16.7%	33.3%	50%	
Ease of use		50%	50%		
Number of e-services		16.7%	83.3%		
E-services performed this year					100%
Extent of depending on portal e-services					100%

6.2.3 Al-Madinah Municipal Website

1. Website Quality

Table 6.20 shows the results of user testing six website quality guidelines (#34 to 39 of the proposed e-city framework). The average score of user testing for the website's quality was 1.54, out of 6, (25.7%). Only the sixth guideline scored slightly more than 50%. The lowest average scores were, 0, 0.04, and 0.29 (twice), out of 1, (in guidelines 3, 5, 2, & 4), indicating the Al-Madinah website has a problem in its interactive city map, search facility, eComplaint, and images and videos respectively.

Table 6.20: User Testing of Website Quality for Al-Madinah Municipal Website

Criteria		Score						
		User 1	User 2	User 3	User 4	User 5	User 6	Average
1	Useful and most up to date content	0.75	0.5	0.25	0.25	0.25	0.25	**0.38**
2	eComplaint	0.75	0.5	0	0	0	0.5	**0.29**
3	Interactive city map	0	0	0	0	0	0	**0.00**
4	Meaningful images and videos	0.5	0.25	0.25	0.25	0.25	0.25	**0.29**
5	Effective search	0	0	0	0	0.25	0	**0.04**
6	Organized short meaningful link labels	0.5	1	0.25	1	0	0.5	**0.54**
Sum		2.5	2.25	0.75	1.5	0.75	1.5	**1.54**

When compiling the results from the user testing and the heuristic tests it provides the overall assessment of website quality. The average score of the two experts was 20.25 and the average of user testing was 1.54. Therefore, the quality of Al-Madinah municipality website was 21.79 out of 39 (55.9%) based on heuristic and user testing (Table 6.21).

Table 6.21: Compiled Website Quality of Al-Madinah Municipal Website

	Guideline Number	Expert1	Expert2	Average User Testing
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	0.5	0.5	
2	About us: mayor corner, mission, objectives of the website	1	0.5	
3	Municipal budget information	0	0	
4	City council meetings (dates, locations, agendas, minutes)	0	0	
5	FAQ with facility to ask new questions	0	0	
6	eJob: job vacancy at municipality	0.5	0.5	
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	0.5	0.5	
8	Last update date on the footer of every page	0	0	
9	Emergency alerts (road closedown, weather alerts…)	0	0	
10	Comment or eSuggest on the website	1	1	
11	Citizen satisfaction survey	1	1	
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	0.5	0.5	
13	Multilingual equivalent websites with a link on header of page	0.5	0.5	
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1	
15	Links to related government websites open in a new window	1	1	
16	Downloadable documents/forms with appropriate access	1	1	
17	Design for common browsers access (Explorer, Chrome)	1	1	
18	Print pages properly	1	1	
19	Well-designed customer focused homepage: Quick access to highlighted services through the main menu	0.5	0.5	
20	Targeted audience group (citizens, business, tourists...)	0	0	
21	Clear entity "Logo" on every page	1	1	
22	Link to homepage from every page through "Home" or logo	1	0.5	
23	Sitemap	1	1	
24	Short and descriptive page titles	0.5	0.5	
25	Readable pages (font, color, background)	1	1	
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5	
27	Privacy and security statement/policy	1	1	
28	All links working properly, i.e. no broken links	0.5	0.5	
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	1	1	
30	Enabled "Back button"	0.5	0.5	
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0.5	
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	1	0.5	
33	Navigation and link style	1	0.5	
34	Useful and most up to date content			0.38
35	eComplaint			0.29

	Guideline Number	Expert1	Expert2	Average User Testing
36	Interactive city map			0
37	Meaningful images and videos			0.29
38	Effective search			0.04
39	Organized short meaningful link labels			0.54
Total Expert Score (out of 33)		21 (63.6%)	19.5 (59.1%)	
Average Score of the two Experts (out of 33)		20.25 (61.4%)		
Average of User Testing (out of 6)				1.54 (25.7%)
Total Score for Website Quality (Experts + User Testing) (out of 39)		21.79 (55.9%)		

2. E-services Quality

The six users were given five usability tasks to evaluate the quality of e-services in the Al-Madinah city website (see Appendix A). The testing result and the calculation were similar to the example given previously in Jeddah. As presented in Table 6.22, the total scores for all tasks evaluated by the six users were: 7.50, 8.81, 6.94, 8.25, 6.76 and 4.69 respectively. The average for all tasks performed on Al-Madinah city website was 7.16 (47.7%) with no user being unable to complete any task.

Table 6.22: Total Score for Six Users Testing E-services in Al-Madinah Municipal Site

Services	User 1	User 2	User 3	User 4	User 5	User 6
1: Inquire about your royal grant (order # 1)	1.5	2.06	1.13	2.25	1.88	0.94
2: Inquire about Saudis deaths during a period of two months and also for one year	0.75	0.94	0.75	0.75	0.38	1.31
3: Inquire about Sultana street	0.75	1.5	0.75	0.75	0.56	0.19
4: Inquire about your transaction # 123, year 1436h	2.25	2.06	2.06	2.25	1.69	0.94
5: Inquire about investment contract # 1 in 1436h	2.25	2.25	2.25	2.25	2.25	1.31
Total Score for all Tasks (out of 15)	7.50	8.81	6.94	8.25	6.76	4.69
Average for all Tasks (out of 15)	**7.16**					
Percentage of Average for all Tasks	**47.7%**					

3. User Satisfaction Survey

At the end of the test, the six users were asked to answer a survey of six questions. As shown in Table 6.23, users did not grade Al-Madinah as an excellent site on any of the six evaluated features. 50% of users considered the quality of the Al-Madinah website poor, 33.3% good and 16.7% fair. As for e-services quality, 50% of users considered it fair and the other 50% poor. 50% rated the ease of using the site good and 50% fair. The number of e-services was considered poor by 66.7% of users, while fair by 33.3%.

Table 6.23: User Satisfaction Survey of Al-Madinah Municipal Website

Features	Excellent	Good	Fair	Poor	None
Website quality		33.3%	16.7%	50%	
E-services quality			50%	50%	
Ease of use		50%	50%		
Number of e-services			33.3%	66.7%	
E-services performed this year					100%
Extent of depending on portal e-services					100%

6.2.4 Eastern Region Municipal Website

1. Website Quality

Table 6.24 shows the results of user testing six website quality guidelines (#34 to 39 of the proposed e-city framework). The average score of user testing for the website's quality was 0.96, out of 6, (16%). Only the second guideline (eComplaint) scored 50% which was the highest score. The lowest average scores were, 0 (three guidelines 3, 4, & 5), 0.21 (guideline 1), and 0.25 (guideline 6), out of 1, indicating the Eastern website has a problem in its interactive city map, images and videos, search facilities, content, organization of links respectively.

Table 6.24: User Testing of Website Quality for Eastern Region Municipal Website

Criteria		Score						
		User 1	User 2	User 3	User 4	User 5	User 6	Average
1	Useful and most up to date content	0.5	0.25	0	0	0	0.5	**0.21**
2	eComplaint	1	0.25	0	0.75	0.5	0.5	**0.50**
3	Interactive city map	0	0	0	0	0	0	**0.00**
4	Meaningful images and videos	0	0	0	0	0	0	**0.00**
5	Effective search	0	0	0	0	0	0	**0.00**
6	Organized short meaningful link labels	0.5	0.5	0	0	0	0.5	**0.25**
Sum		2	1	0	0.75	0.5	1.5	**0.96**

Compiling the results from user testing and heuristic tests provides the overall assessment of website quality. The average score of the two experts was 16.75 and the average of user testing was 0.96. Therefore, the quality of the Eastern Region municipality website was 17.71 out of 39 (45.4%) based on heuristic and user testing (Table 6.25).

Table 6.25: Compiled Website Quality of Eastern Region Municipal Website

	Guideline Number	Expert1	Expert2	Average User Testing
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	0.5	0.5	
2	About us: mayor corner, mission, objectives of the website	1	0.5	
3	Municipal budget information	0	0	
4	City council meetings (dates, locations, agendas, minutes)	0	0	
5	FAQ with facility to ask new questions	0.5	0.5	
6	eJob: job vacancy at municipality	0	0	
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	0	0	
8	Last update date on the footer of every page	0	0	
9	Emergency alerts (road closedown, weather alerts…)	0	0	
10	Comment or eSuggest on the website	1	1	
11	Citizen satisfaction survey	0	0	
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	1	0.5	
13	Multilingual equivalent websites with a link on header of page	0	0	
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1	
15	Links to related government websites open in a new window	1	1	
16	Downloadable documents/forms with appropriate access	1	1	
17	Design for common browsers access (Explorer, Chrome)	1	1	
18	Print pages properly	0.5	0.5	
	Well-designed customer focused homepage:			

	Guideline Number	Expert1	Expert2	Average User Testing
19	Quick access to highlighted services through the main menu	0.5	0.5	
20	Targeted audience group (citizens, business, tourists...)	0	0	
21	Clear entity "Logo" on every page	0.5	0.5	
22	Link to homepage from every page through "Home" or logo	1	1	
23	Sitemap	1	1	
24	Short and descriptive page titles	1	1	
25	Readable pages (font, color, background)	0.5	1	
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5	
27	Privacy and security statement/policy	1	1	
28	All links working properly, i.e. no broken links	0.5	0.5	
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	0	0	
30	Enabled "Back button"	1	1	
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0	
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	0.5	0.5	
33	Navigation and link style	0.5	0.5	
34	Useful and most up to date content			0.21
35	eComplaint			0.5
36	Interactive city map			0
37	Meaningful images and videos			0
38	Effective search			0
39	Organized short meaningful link labels			0.25
Total Expert Score (out of 33)		17 (51.5%)	16.5 (50%)	
Average Score of the two Experts (out of 33)		16.75 (50.8%)		
Average of User Testing (out of 6)				0.96 (16%)
Total Score for Website Quality (Experts + User Testing) (out of 39)		17.71 (45.4%)		

2. E-services Quality

The six users were given five usability tasks to evaluate the quality of e-services in the Eastern Region city website (see Appendix A). The testing result and the calculation were similar to the example given previously in Jeddah. As shown in Table 6.26, the total scores for all tasks evaluated by the six users were: 1.5, 5.07, 2.44, 3.37, 2.38 and 8.25 respectively. The average for all tasks performed on the Eastern Region city website was as low as 3.84 (25.6%). The only user who could complete tasks 4 and 5 was user 6. She

succeeded in finding e-service 5 because of her knowledge of administrative terminology, since she is working as a director of administration. Also, three users were unable to complete task 1 which means 50% could not find the link to the targeted e-service.

Table 6.26: Total Score for Six Users Testing E-services in Eastern Region Site

Services	User 1	User 2	User 3	User 4	User 5	User 6
1: Report online about lights in day time (Enter required fields only)	0	1.88	0	0	2.38	2.25
2: Inquire about health certificates	0.75	0.56	0.38	0.56	0	0.75
3: Inquire about investment contracts	0.75	2.63	2.06	2.81	0	2.25
4: Inquire about building permit by identification number	0	0	0	0	0	0
5: Inquire about transaction number 12466 dated 1436	0	0	0	0	0	3
Total Score for all Tasks (out of 15)	1.50	5.07	2.44	3.37	2.38	8.25
Average for all Tasks (out of 15)	**3.84**					
Percentage of Average for all Tasks	**25.6%**					

3. User Satisfaction Survey

At the end of the test, the six users were asked to answer a survey of six questions. As shown in Table 6.27, users did not grade Eastern Region as an excellent site on any of the six evaluated features. For three features: website quality, e-services quality and number of e-services, 66.7% of users consider it poor while 33.3% fair. 66.7% rated the ease of using the site fair while 33.3% poor.

Table 6.27: User Satisfaction Survey of Eastern Region Municipal Website

Features	Excellent	Good	Fair	Poor	None
Website quality			33.3%	66.7%	
E-Services quality			33.3%	66.7%	
Ease of use			66.7%	33.3%	
Number of e-services			33.3%	66.7%	
E-services performed this year					100%
Extent of depending on portal e-services					100%

6.2.5 Qassim Municipal Website

1. Website Quality

Table 6.28 shows the results of user testing six website quality guidelines (#34 to 39 of the proposed e-city framework). The average score of user testing for the website's quality was 2.08, out of 6, (34.7%). Only two guidelines (2 & 6) scored slightly above 50%. The lowest scores were 0.04, 0.21, and 0.25, out of 1, indicating the Qassim website has a problem in its images and videos, search facility and interactive city map respectively.

Table 6.28: User Testing of Website Quality for Qassim Municipal Website

Criteria		Score						
		User 1	User 2	User 3	User 4	User 5	User 6	Average
1	Useful and most up to date content	0.5	0.5	0.5	0.25	0.5	0.5	**0.46**
2	eComplaint	1	0.75	0.25	0.75	0.75	0	**0.58**
3	Interactive city map	0.25	0.25	0.25	0.5	0	0.25	**0.25**
4	Meaningful images and videos	0	0	0.25	0	0	0	**0.04**
5	Effective search	0	0.75	0.25	0	0	0.25	**0.21**
6	Organized short meaningful link labels	0.75	0.5	0	1	0.5	0.5	**0.54**
Sum		2.5	2.75	1.5	2.5	1.75	1.5	**2.08**

Compiling the results from the user testing and the heuristic tests provides the overall assessment of website quality. Both experts arrived to the same score of 22 and the average of user testing was 2.08. Therefore, the quality of Qassim municipality website was 24.08 out of 39 (61.7%) based on heuristic and user testing (Table 6.29).

Table 6.29: Compiled Website Quality of Qassim Municipal Website

	Guideline Number	Expert1	Expert2	Average User Testing
1	Contact information (phones, e-mails, physical address, link to customer service email, working hours)	1	0.5	
2	About us: mayor corner, mission, objectives of the website	0.5	0.5	
3	Municipal budget information	0	0	
4	City council meetings (dates, locations, agendas, minutes)	0	0.5	
5	FAQ with facility to ask new questions	0.5	0.5	
6	eJob: job vacancy at municipality	0	0	
7	News important to users (city statistics, projects, calendar of events, photo gallery…)	1	0.5	

	Guideline Number	Expert1	Expert2	Average User Testing
8	Last update date on the footer of every page	1	1	
9	Emergency alerts (road closedown, weather alerts…)	0	0	
10	Comment or eSuggest on the website	1	1	
11	Citizen satisfaction survey	0	1	
12	Social media (online chat with municipality presenters, discussion forum, Facebook, Twitter, ...)	0.5	0.5	
13	Multilingual equivalent websites with a link on header of page	0	0	
14	The city website among top 10 hits (results) of Google and Yahoo search engines	1	1	
15	Links to related government websites open in a new window	1	1	
16	Downloadable documents/forms with appropriate access	1	1	
17	Design for common browsers access (Explorer, Chrome)	1	1	
18	Print pages properly	1	1	
19	Well-designed customer focused homepage: Quick access to highlighted services through the main menu	0.5	0.5	
20	Targeted audience group (citizens, business, tourists...)	0.5	0.5	
21	Clear entity "Logo" on every page	0.5	0.5	
22	Link to homepage from every page through "Home" or logo	1	1	
23	Sitemap	1	0.5	
24	Short and descriptive page titles	1	1	
25	Readable pages (font, color, background)	1	1	
26	Simple page with reasonable length of not more than 2 screens	0.5	0.5	
27	Privacy and security statement/policy	1	1	
28	All links working properly, i.e. no broken links	1	1	
29	Navigational options: Indicator of a user is where on the site (e.g. Breadcrumbs)	0.5	0.5	
30	Enabled "Back button"	1	1	
31	Different colors for visited/unvisited links, underline links and no misleading cues to click	0	0	
32	Consistent design of all web pages: Page layout (same feel and look, font, color, buttons, menus)	1	1	
33	Navigation and link style	1	1	
34	Useful and most up to date content			0.46
35	eComplaint			0.58
36	Interactive city map			0.25
37	Meaningful images and videos			0.04
38	Effective search			0.21
39	Organized short meaningful link labels			0.54
Total Expert Score (out of 33)		22 (66.7%)	22 (66.7%)	
Average of User Testing (out of 6)				2.08 (34.7%)
Total Score for Website Quality (Experts + User Testing) (out of 39)		24.08 (61.7%)		

2. E-services Quality

The six users were given five usability tasks to evaluate the quality of e-services in the Qassim city website (see Appendix A). The testing result and the calculation were similar to the example given previously in Jeddah. As shown in Table 6.30, the total scores for all tasks evaluated by the six users were: 2.19, 5.25, 5.26, 6.75, 8.07 and 3.19 respectively. The average for all tasks performed on the Qassim city website was 5.12 (34.1%). Task 3 revealed that it was difficult to find the link to the e-service. Also, user 1 gave a zero as an evaluation for both e-services 2 and 5.

Table 6.30: Total Score for Six Users Testing E-services in Qassim Municipal Website

Services	User 1	User 2	User 3	User 4	User 5	User 6
1: Inquire about your transaction # 123 for the year 1435h	2.06	2.25	1.88	2.25	2.06	1.69
2: See the official engineering offices such as Al-Rajhi office	0	0.75	0.75	1.5	2.25	0.56
3: Ask for a license to add floors to your home	0.13	0	0	0	0	0
4: Search for Job vacancies	0	2.25	2.25	2.25	1.88	0.56
5: Explore Urban observatory for Buridah	0	0	0.38	0.75	1.88	0.38
Total Score for all Tasks (out of 15)	2.19	5.25	5.26	6.75	8.07	3.19
Average for all Tasks (out of 15)	**5.12**					
Percentage of Average for all Tasks	**34.1%**					

3. User Satisfaction Survey

At the end of the test, the six users were asked to answer a survey of six questions. As shown in Table 6.31, users did not grade Qassim as an excellent site on any of the six evaluated features. 50% of users considered the quality of the Qassim website fair, 33.3% poor, and 16.7% good. As for e-services quality, 66.7% of users considered it poor and 33.3% fair. 66.7% rated the ease of using the site fair and 33.3% good. The number of e-services was considered poor by 83.3% of users while fair by 16.7%.

Table 6.31: User Satisfaction Survey of Qassim Municipal Website

Features	Excellent	Good	Fair	Poor	None
WS quality		16.7%	50%	33.3%	
E-services quality			33.3%	66.7%	
Ease of use		33.3%	66.7%		
Number of e-services			16.7%	83.3%	
E-services performed this year					100%
Extent of depending on portal e-services					100%

6.3 Alexa Web Analytics Tool

This section details the results of testing Saudi e-city websites by the Alexa web analytics tool carried out during the month of July 2014. Alexa presents its data either as an individual site overview or as site comparisons. On site overview, traffic rank data is available for sites ranked < 100,000. As described in section 5.2.2.3, nine metrics were chosen for evaluating Saudi municipal websites: domestic and global traffic ranks, page views/visitor, speed of download, bounce rate, sites linking in, time on site, audience geography, and where do visitors go on the site.

Table 6.32 shows that the highest Alexa traffic ranked city sites in Saudi Arabia were Jeddah then Riyadh with 1237 and 1417 respectively. Next Qassim was ranked 3282 followed by Al-Madinah 4249 and Eastern Region 5616, which reflected their weak performance in the tests. Globally, Jeddah and Riyadh also had the highest traffic rank of 105,231 and 161,050 respectively, whereas the Eastern Region had the lowest rank 802,029. Moreover, Al-Madinah had the highest estimated number of pages viewed (4.70) per day followed by Jeddah 3.60, Eastern Region 3.20, Qassim 2.50 and Riyadh 2.30. Thus all e-cities showed a weak performance in this parameter.

Regarding the download speed, Jeddah website was fast (1.345 seconds) contrary to the Riyadh website which was very slow (8.758 seconds). For the other three municipalities, Al-Madinah, Qassim and Eastern Region, download speed data was not provided by Alexa

and the historical traffic rank graph was unavailable which implies that they were slow and low-ranked websites. Therefore, only the Jeddah website performed well in this important feature. Further, the lower the bounce rate the better (percentage of visitors leaving the site after visiting one page). Thus the rates of Qassim (19.20%), Eastern Region (25.00%), and Al-Madinah (31.60%) were good, whereas Jeddah (41.10%) and Riyadh (54.50%) were below average. As to the number of sites linking in, Jeddah has received 284, Riyadh 247, Al- Madinah 117, Qassim 49 and Eastern Region 32 sites only, which probably indicated that Saudi city websites were not popular. The daily time on site by visitors was the highest for Al-Madinah 15 minutes, Qassim 8 minutes, Jeddah 6 minutes, and for Riyadh and Eastern Region the spent time was as low as two minutes.

Table 6.32: Web Data Statistics obtained from Alexa

City Site	Traffic Rank Saudi	Traffic Rank Global	Page views	Speed	Bounce Rate	Sites Linking In	Time on Site	Audience Geography	Where visitors go on the Site
Jeddah	1237	105,231	3.60	1.345	41.10%	284	6.00	SA 80.8%	Jeddah.gov 88.48% Iservices.Jeddah.gov 8.32%
Riyadh	1417	161,050	2.30	8.758	54.50%	247	2.00	SA 93%	alriyadh.gov 73.61% eservices.riyadh.gov 21.68%
Al-Madinah	4249	353,007	4.70	-	31.60%	117	15.00	SA 86.6%	amana-md.gov 51.91% services.amana-md.gov 49.5%
Qassim	3282	555,084	2.50	-	19.20%	49	8.00	SA 91.2% Egypt8.8%	Mail.qassem.gov 55.78% Qassim.gov 52.54%
Eastern Region	5616	802,029	3.20	-	25.00%	32	2.00	SA 100%	eamana.gov 100%

The next evaluation metric is the audience geography (where visitors come from).Table 6.32 shows that all visitors to Saudi city websites came from within the country (locally) except the Qassim website which has been seen by a small percentage of visitors from Egypt. The last index "where visitors go on the site" might indicate visitors' interest to different subdomains of the site. We found that visitors to the Al-Madinah site go frequently to the home page and also to the e-services (49.53%), while 21.68% of Riyadh visitors checked the e-services as well as the home page, and similarly only 8.3% of Jeddah

visitors used its e-services. The visitors of the other two municipalities, Qassim and Eastern Region, navigated to the home page while none go to the e-services.

In general, most Saudi city websites did not perform well on the web, according to Alexa. Sometimes Alexa was unable to display graphics and data about some municipal websites, such as the download speed of Al-Madinah, Qassim and Eastern Region, because of their performance. All investigated Saudi city websites had global web traffic ranks far more than 100,000, reflecting their unpopularity. Domestic web traffic ranks were also poor (> 1000) with the first ranked city site, Jeddah, in the country at 1237. A weak performance in the attribute of page views was noticed with only the Al-Madinah site performing well. Further, all municipality websites were very slow except Jeddah, and also all had a very low number of sites linking in, reflecting their weak correlation with others on the Web. On the bounce rate, all performed well except Riyadh and Jeddah, in which the percentage of visitors leaving the site after one page was high. For the metric "time on site", visitors stayed on the sites for less than 8 minutes, except Al-Madinah 15 minutes. That might imply they were unsatisfied with the content of most Saudi city websites. Also, probably there was a problem on the e-services since citizen visits to e-services sections was very low on most Saudi municipal websites. Overall, none of the municipalities obtained a constant rank throughout all the metrics; for example, Jeddah was a fast website but did not perform well in other metrics.

On the other hand, the option of site comparisons on Alexa allows us to compare traffic key metrics over time for up to 10 different websites if they have a high volume of web traffic. The next two figures depict a sample screen shot of testing five Saudi municipal websites as shown on site comparisons view for July 2014. Figure 6.30 and 6.31 displays a graph of historical global traffic trends over a six month period of time, from February to July 2014. According to this data, the ranking was: Jeddah, Riyadh, Al-Madinah, Qassim, and Eastern Region.

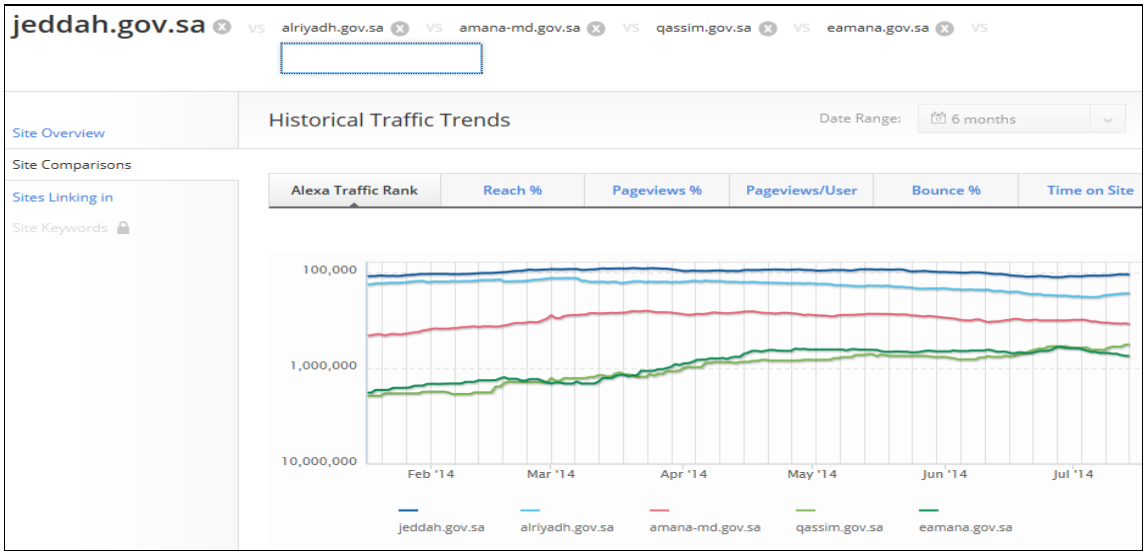

Figure 6.30: Comparison between five Saudi Municipal Websites at Alexa (page 1)

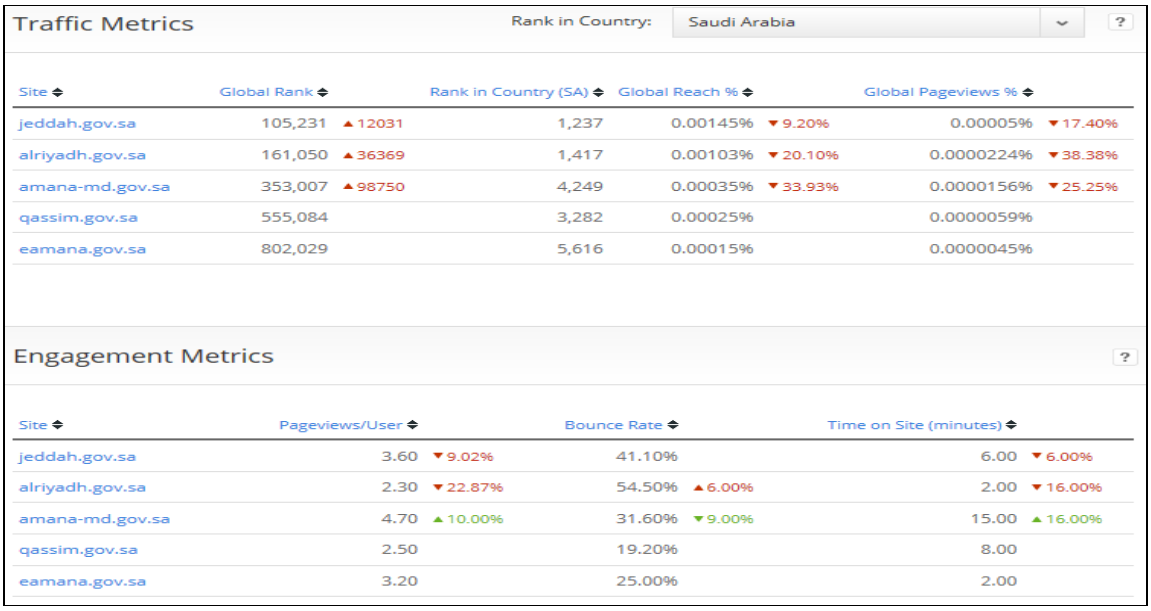

Figure 6.31: Comparison between five Saudi Municipal Websites at Alexa (page 2)

6.3.1 Reliability and Usefulness of Alexa

The purpose of this section is to find out if Alexa web traffic data provides useful and reliable information. Very few studies have been carried out using Alexa traffic data without assessing its reliability and questions remain regarding the relative advantages and disadvantages of the Alexa tool (Vaughan & Yang, 2013). In fact, web traffic data are underutilized but yet we can discover information that may otherwise be unavailable to us, such as web traffic data of multiple websites (Vaughan, 2008).

This research encountered problems with the Alexa tool as follows:

1) Alexa data are collected from users who have installed the Alexa toolbar into their browsers. Thus, Alexa ranks are based on the traffic of a limited number of users that may not be a representative sample of the Internet population (Vaughan & Yang, 2013). This potential bias of Alexa because of the way it collects data from its toolbar users compromises its reliability even though Alexa claims normalization of its data to correct this bias.

2) Traffic rank can be inaccurate since it is easy to manipulate the data. With a little effort, a user can or through some friends install the toolbar and surf a site every day, hence the site rank could jump up toward top ranked websites.

3) There is no rule for measuring metrics since Alexa doesn't specify the standard of low or high measures as the acceptable average figure for each metric and what could be considered low. For example, among the few studies in the literature, Jowkar and Didegah (2010) used Alexa data to evaluate Iranian newspapers' websites without specifying the scientific basis for metric measurement. Another study (Bhat, 2013) considered a good domestic traffic rank to be less than 1,000 and a weak one greater than 10,000 without supporting references. According to Inc. Magazine (2011), a bounce rate (percentage of visitors leaving the site after one page) higher than 80% is bad, of 50% is average and below 30% is very good; again these figures lack any foundation.

4) There are no guidelines on how to conclude or deduce valuable information from such a large amount of rich raw data. There are some examples from Rodriguez (2013) that show how Alexa data analysis is complicated. Returning to the previous example of Inc. Magazine, a high bounce rate is bad, but Rodriguez found in some cases the opposite is true. That is if the information is well targeted and the site provides all visitors' needs on a page then a high bounce rate is good. Another example, a high number of page views per visitor is good. However, it might also mean that it is difficult for the visitors to find the information they are looking for. Also, in other circumstances, a low number of pageviews is good. Consider Google, ideally, it will have 2 page views per visitor per visit since you will see (1) their homepage and (2) the search results, before clicking on a link to another

page. Therefore, Alexa data analysis is complicated, especially with the absence of guidelines. Another point: sometimes it is hard to make sense from Alexa data since the motivation behind visitors' actions cannot be predicted, such as the time on site metric can be highly but misleading if the user left his computer, for one reason or another, then returned to the site after a while.

5) The correlation between Alexa data and the performance of e-municipalities, in particular, has not been firmly established yet. According to the authors of "web traffic map of Spanish municipalities: building a ranking" (Ferras, Real, & Rosado, 2012), there was no direct correlation between population size, economics, and web traffic. Large urban municipalities do not occupy key positions in Alexa ranking. However, it is possible that comparing the local web traffic rank of municipalities of different population size is misleading. For instance, if all Al-Madinah inhabitants (about one million) and half of Riyadh inhabitants (about two and a half million) have visited their city websites, the web traffic rank of Riyadh will outperform Al-Madinah.

In conclusion, Alexa is an indicator rather than a metric tool to evaluate a website and its ranking is not accurate nor reliable. This is due to its methodology which relies on sampling through installed tool bars. However, this tool contains rich data on how websites are being used and we need to find systematic ways to extract useful information from such a large amount of raw data. Web usage data are challenging for researchers and it is a fertile field yet unexplored. A good approach is to compare Alexa's results with other outcomes from more established web evaluation methods for comparison and validation. If they correlate, this may confirm that the web traffic data contains useful information on the quality of the tested websites. For example, in our study, Alexa indicated that there is a problem on the e-services on most Saudi municipal websites. This correlates with heuristic and user testing results which showed that Saudi municipalities suffered from limited e-services on their websites. However, more research is needed to gain more knowledge into how Alexa data can complement traditional web evaluation methods such as the heuristic and user testing.

6.4 Link Checker: Automatic Website Evaluation Tools

Broken links have a negative effect on e-government websites. The purpose of this section is to identify and compare a sample of the most used link checker tools and also to test how reliable are these tools. Seven link checkers were chosen: Broken Link Checker, Dead Link Checker, NetMechanic, LinkTiger, Link Alarm, Web Link Validator, and Xenu.

Table 6.33 presents the seven tools ordered according to the price and area of coverage. All the tools are web-based services except the Web Link Validator and Xenu which are desktop applications that need installation locally. The best tools in terms of price and coverage are Dead Link Checker and Xenu since they are free and can test the entire targeted website. The third tool, Link Tiger, tested the entire website with a limited 15-day free trial which is enough time for doing our tests. Broken Link Checker is free for three thousand webpages, LinkAlarm tests 100 webpages with 15-day free trial, and Web Link Validator is free for 500 links only. According to their website, Broken Link Checker won the Web tool prize in 2012 and was used by the London Olympics 2012 site. NetMechanic, which is free for 5 webpages, is a tool that offers other features besides identifying broken links, such as assessing HTML code, browser compatibility, load time and spell checking.

Table 6.33: Sample of Broken Link Checkers for Testing Websites

	Name of Link Checker	Type	Price	Coverage
1	Dead Link Checker	Online tool www.deadlinkchecker.com/	Free	Entire website
2	Xenu	Desktop s\w	Free	Entire website
3	LinkTiger	Online tool www.linktiger.com/	Limited 15-day free trial	Entire website
4	Broken Link Checker	Online tool www.brokenlinkcheck.com/	Partially Free	3000 Webpages
5	Link Alarm	Online tool www.linkalarm.com/	Limited 15-day free trial	100 Webpages
6	Web Link Validator	Desktop s\w	Partially Free	500 links
7	NetMechanic (Broken links, browser compatibility, load time, etc.)	Online tool www.netmechanic.com	Partially Free	5 Webpages

During the month of August 2014, seven automatic link checker tools were applied to the Jeddah municipality website as a preliminary trial test. Table 6.34 shows a big

difference in the test results of the Jeddah city site, even among tools covering the entire website. Some examples of these tools are in Figures 6.32, 6.33, 6.34 and 6.35. There was no agreement at all between the seven tools on the number of broken links or even the total number of links. The number of broken links was given as 25,725 by Xenu, 11,576 by LinkTiger, 4,149 by Broken Link Checker, 4,012 by Dead Link Checker and as low as 314 by Link Alarm, 99 by Web Link validator, and 0 (zero) by NetMechanic. Besides the number of broken links, the Link Alarm gave a site score (63 out of 100) and a link failure rate of 8.1 which is, as they claimed, worse than the benchmark link failure rate of 1.3% for the e-government category. However, NetMechanic provided further information about the Jeddah site such as the load time (14.95 sec.) and the browser compatibility (18 problems). In addition, it was noticed that some link checker tools encountered some problems when evaluating a large website. For example, Xenu gave three different results (25725, 8923 and 2 broken links) for the Jeddah site even though the three tests were performed consecutively on the same day. It seems that we need more investigations on these seven tools; maybe one of them is reliable only, or all are not reliable.

Table 6.34: Comparison of Broken Link Checkers on Jeddah Municipality Website

	Name of Link Checker	Coverage	# Links	Results
1	Dead Link Checker	Whole website	18,130 URLs	Broken links: **4,012**
2	Xenu	Whole website	2,986 URLs	Broken links: **25,725**
3	LinkTiger	Whole website	34,031 links	Broken links: **11,576**
4	Broken Link Checker	3000 Webpages	-	Broken links: **4,149**
5	Link Alarm	100 Webpages	3,509 internal 312 external Total 3,821 links	Broken links: **314** Site score: **63** (of 100) Link failure: **8.1**% worse than the benchmark link failure rate of **1.3**% for the category Government
6	Web Link Validator	500 links	448 internal 52 external	Broken links: **99** (20%) 205 pages (41%)
7	NetMechanic	5 Webpages	146 URLs	Broken links: **0** Load time: 14.95 sec. Browser compatibility:18 problems

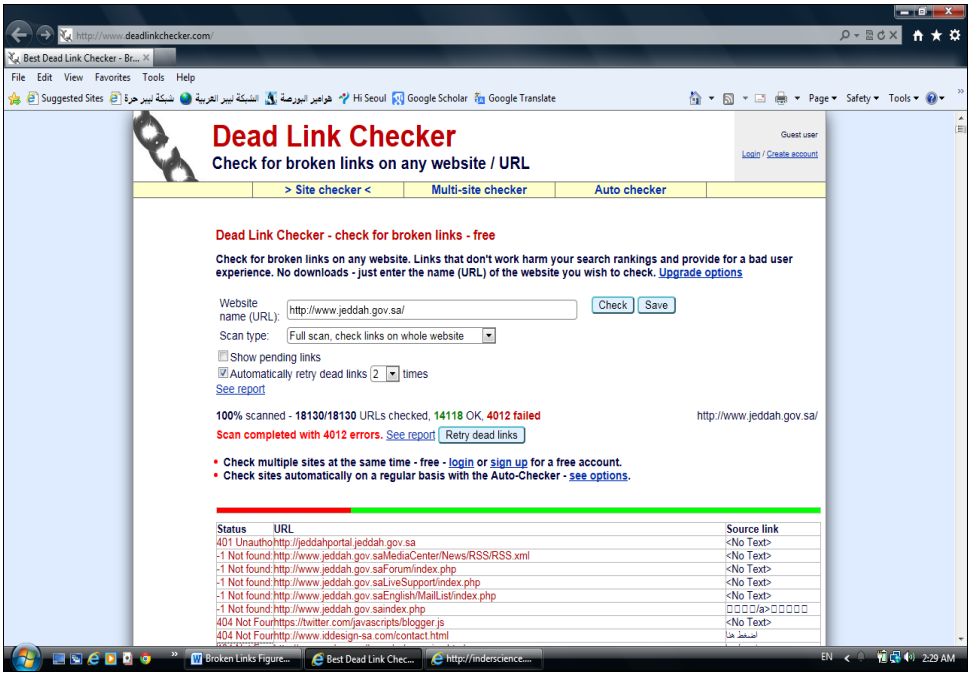

Figure 6.32: Dead Link Checker 4012 Broken Links on Jeddah Municipal Website

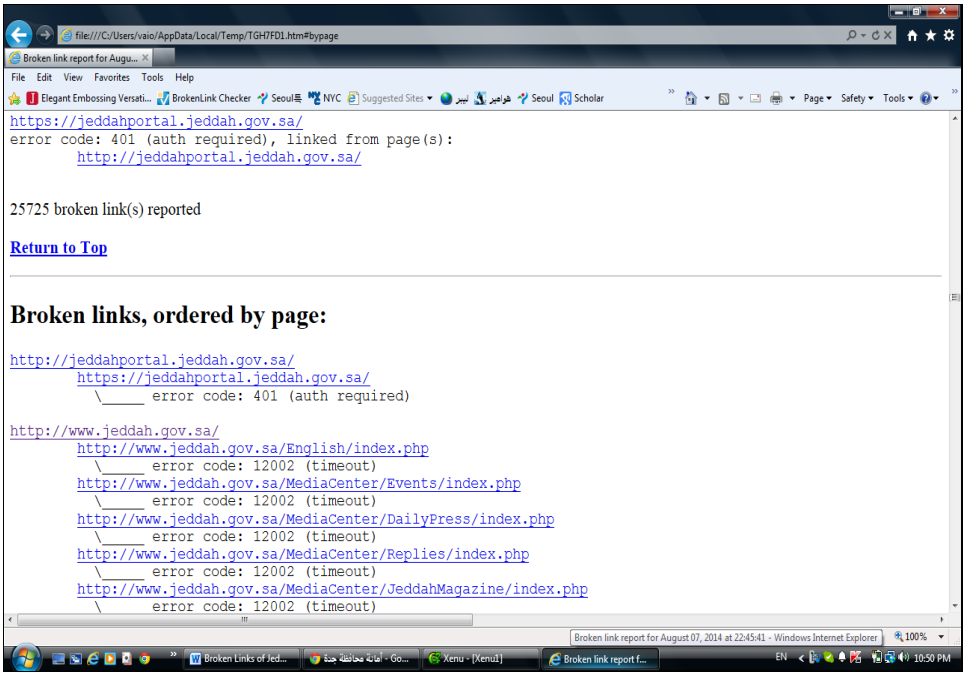

Figure 6.33: Xenu 25,725 Broken Links on Jeddah Municipal Website

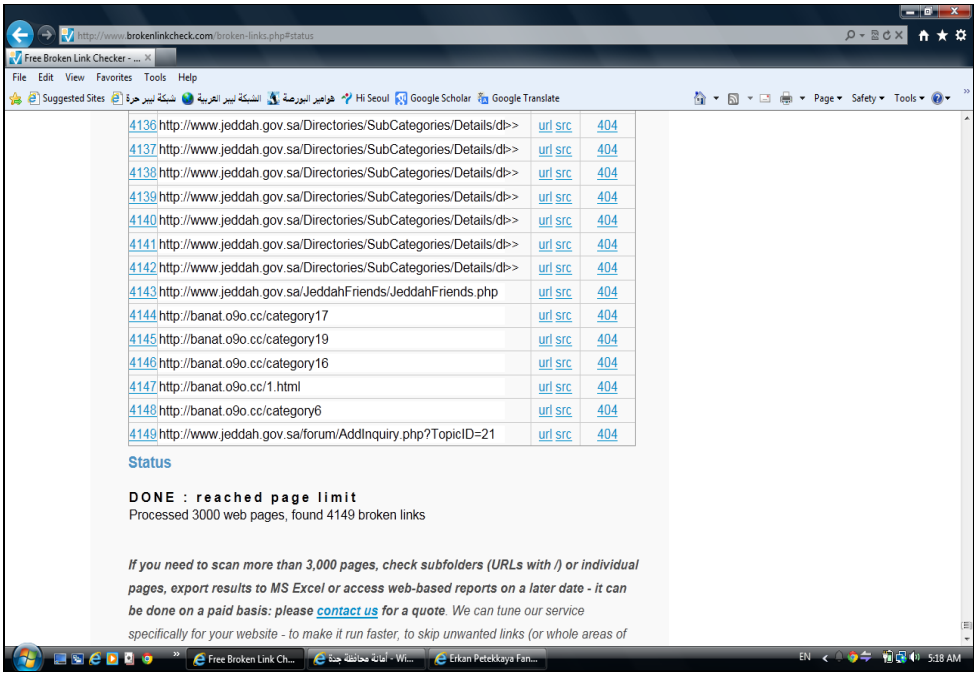

Figure 6.34: Broken Link Checker 4149 Broken Links on Jeddah Municipal Website

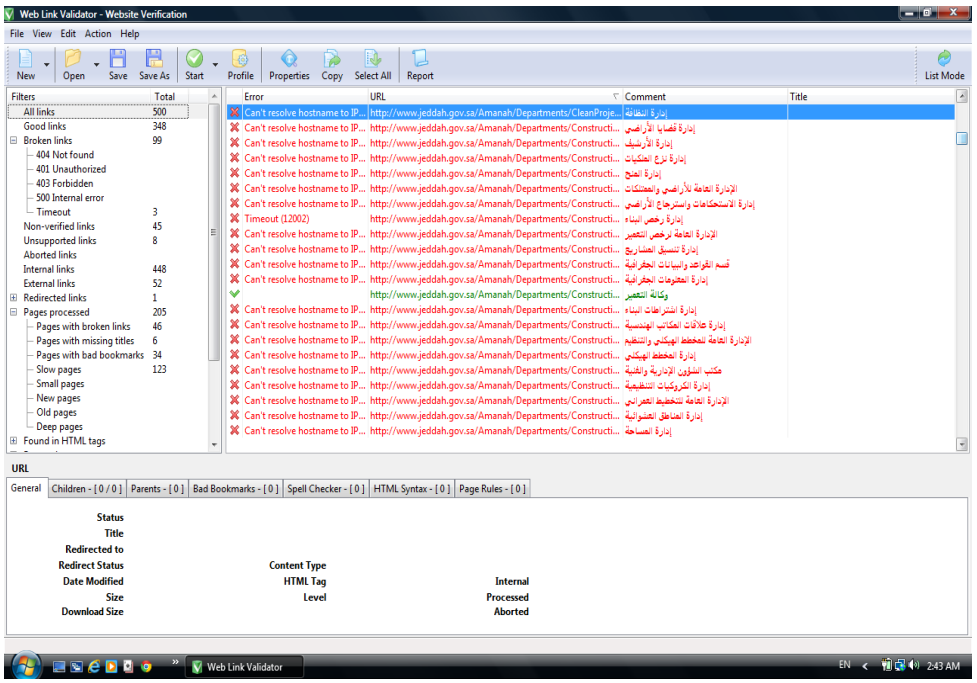

Figure 6.35: Web Link Validator 99 Broken Links on Jeddah Municipal Website

Since these results, with very high broken links, seem to be unreliable, it might be a good idea to check the seven tools against a site with few bad links, in other words against a globally high-ranked municipal site such as the New York City site, in order to know if the tools were exaggerating the test results of Jeddah. Automatic testing of the New York City site had a lower link failure for most tools, ranging from 0, 20, 22, 133, 292 and 840, except Xenu's 14,547 broken links (Table 6.35). However the same problem is still ongoing, namely the lack of agreement among the seven tools, especially those that cover the whole website such as Dead Link Checker, Xenu, and LinkTiger, on both the number of links and broken links. Further inspection of Jeddah and New York results revealed that Xenu tends to detect a high number of broken links in both cases. On the contrary, NetMechanic (Figure 6.36) is always indicating no broken links on the two sites; that might be attributed to the fact that it examined five webpages only and usually the inner webpages are the ones that suffer from broken links more than the home or main pages.

Table 6.35: Comparison of Broken Link Checkers on New York City Website

	Name of Link Checker	Coverage	# Links	Results
1	Dead Link Checker	Whole website	14,019 URLs	Broken links: **840**
2	Xenu	Whole website	5,808 URLs	Broken links: **14,547**
3	LinkTiger	Whole website	5,999 links	Broken links: **133**
4	Broken Link Checker	3000 Webpages	-	Broken links: **20**
5	Link Alarm	3000 Webpages	6177internal 3822external Total links 9999	Broken links: **292** Site score: 90 Link failure: 2.8% worse than the benchmark link failure rate of 1.3% for the category Government
6	Web Link Validator	500 links	398 internal 102 external	Broken links: **22** 127 webpages
7	NetMechanic	5 Webpages	225 URLs	Broken links: **0**

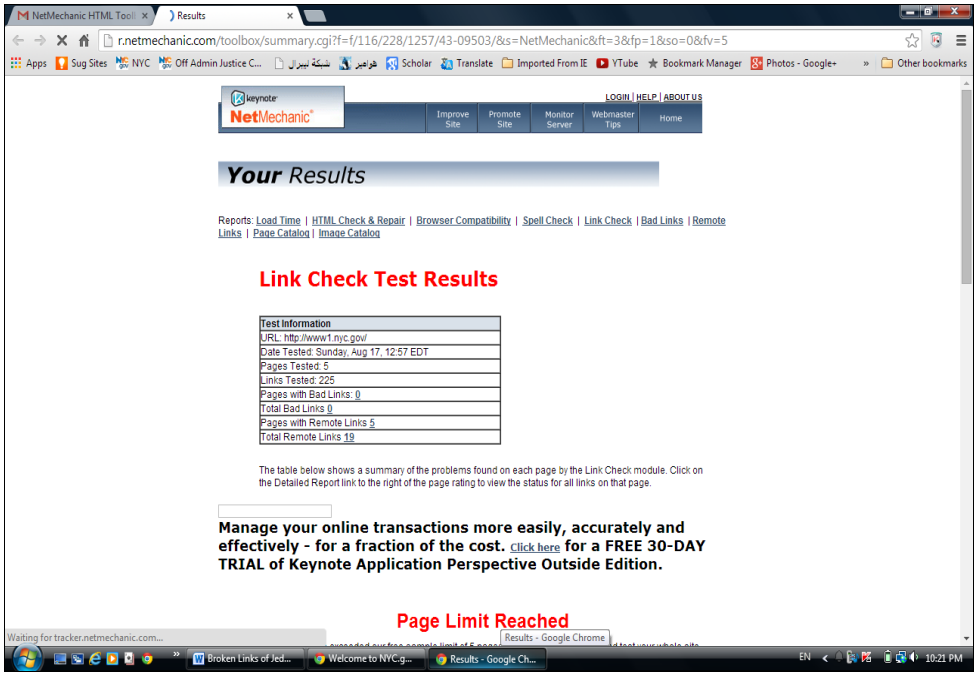

Figure 6.36: NetMechanic Reporting No Broken Links on New York City Website

Since there was no agreement among the seven tools on the broken links or even the number of total links on the two e-cities, Jeddah and New York, in addition to the fact that it is hard practically to check links of such huge websites manually, therefore to assess the reliability of these tools a website with a limited number of pages (the Edinburgh Shetland Fiddlers' Society: http://e-s-f.fsnet.co.uk/index.htm) was tested by the seven tools plus Google Webmaster. Manual link checking results revealed 4 broken links in 27 webpages. Based on Table 5.36, the obtained results confirmed that: 1) Google Webmaster couldn't find any broken links on this site (Figure 6.37); 2) all tested tools were not accurate in their results, either not identifying broken links (Google Webmaster, Broken Link Checker and NetMechanic), exaggerated them (Xenu 32 and Web Link Validator 7 broken links), underestimated their actual number (LinkAlarm 2 broken links as in Figure 6.38) or was near but not exactly (Dead Link Checker 3 and LinkTiger 5 broken links). Therefore, we come to the conclusion that auto link-checkers are not reliable tools even though some are useful and succeeded partially in identifying some broken links. Although more research is needed in the field of automatic link-checkers, webmasters can employ several tools to check their links.

Table 6.36: Comparison of Broken Link Checkers on a small owned Website

	Link Checkers	Coverage	# Links	Results
	Google Webmaster (Possible Benchmark)	Whole website	133 links	Broken links: **0**
1	Dead Link Checker	Whole website	135 URLs	Broken links: **3**
2	Xenu	Whole website	151 URLs	Broken links: **32**
3	LinkTiger	Whole website	166 links	Broken links: **5**
4	Broken Link Checker	3000 Webpages	-	Broken links: **0** 27 webpages
5	LinkAlarm	100 Webpages	412 internal 112 external Total 524 links	Broken links: **2** Site score: 97 (of 100) Link failure: 1.5% 27 webpages
6	Web Link Validator	500 links	172 links	Broken links: **7** 28 webpages
7	NetMechanic	5 Webpages	35 URLs	Broken links: **0**

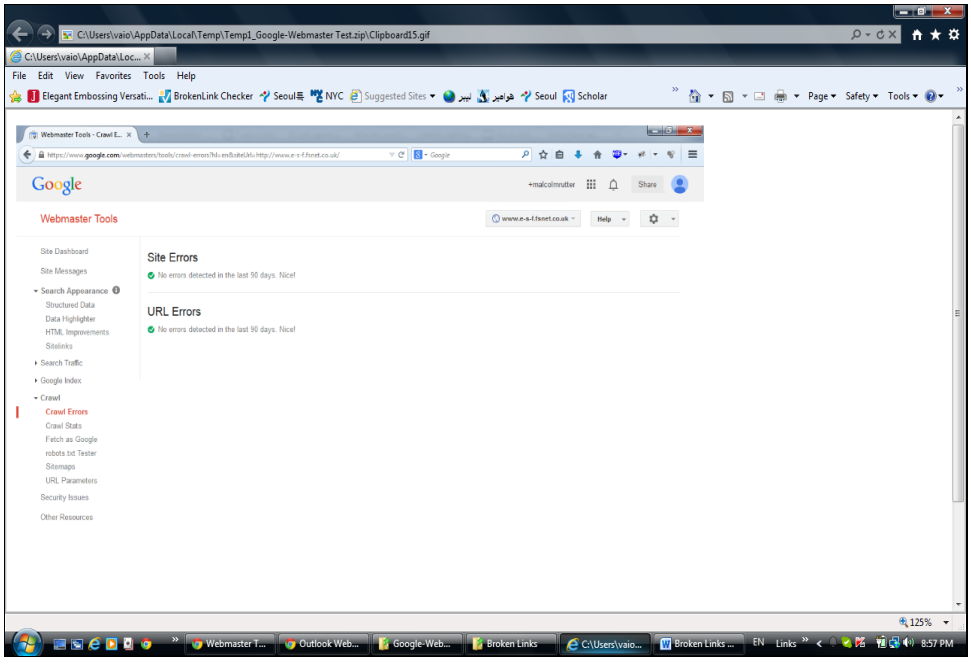

Figure 6.37: Google Webmaster Reporting

No Broken Links on a small owned Website

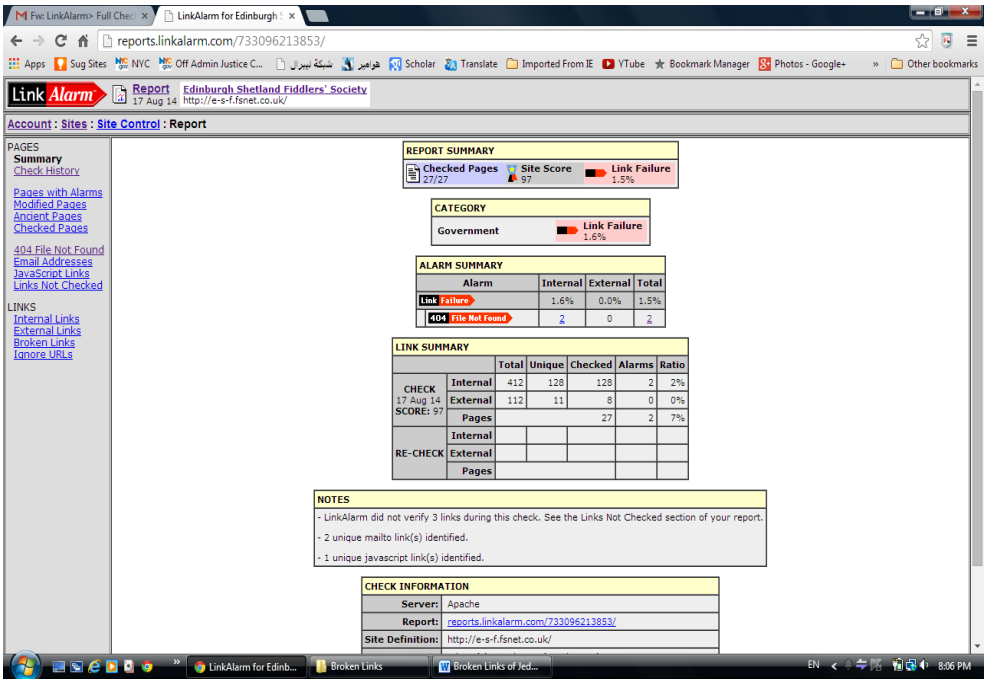

Figure 6.38: LinkAlarm Reporting 2 Broken Links on the Sample Website

6.4.1 Broken Links on Five Saudi City Websites

After testing Jeddah and even though link checkers are not reliable tools, the next step was to assess four other Saudi municipal websites since we'd like to compile different testing results. The aim is to compare and draw a conclusion about automatic link checkers, expert evaluation, and user testing, regarding broken links. Since the free trial of some tools had already ended and since the tools themselves are merely indicators of the presence of broken links in a site, four tools only were used here: Dead Link Checker, Xenu, Broken Link Checker and Web Link Validator. Table 6.37 combined the results of all tests performed on five Saudi municipal websites. During the tests, errors connecting to the site (timeout) were encountered often by Dead Link Checker and Web Link Validator in the slow website of Eastern Region. Xenu had similar problems in evaluating both the Riyadh and Al-Madinah slow sites. Again, there was no agreement between the four tools regarding the number of broken links or the number of links detected in the five Saudi municipal websites.

Table 6.37: Comparison of Broken Link Checkers on five Saudi City Websites

	Jeddah Broken Links	Riyadh Broken Links	Al-Madinah Broken Links	Qassim Broken Links	Eastern Region Broken Links
Dead Link Checker	**4,012** (18,130 links)	**3,656** (12,270 links)	**2,721** (6941 links)	**63** (6137 links)	**214** (703 links) ***error**
Xenu	**25,725** (2986 links)	**109** (446 links) ***error**	**10,688** (1108 links) ***error**	**64,714** (1822 links)	**2,636** (402 links)
Broken Link Checker	**4,149** (3000 pages)	**45** (3000 pages)	**0** (370 pages)	**904** (3000 pages)	**1,175** (730 pages)
Web Link Validator	**99** (20%) (500 links)	**49** (11%) (448 links)	**3** (1%) (500 links)	**7**(1%) (500 links)	**28** (6%) (456 links) ***error**

6.5 Compiled Testing Results

At first, we compiled the results obtained from the heuristic and user testing. An overall ranking of the e-city website is given as well as ranking by website quality or e-services since different components give different information about web quality. Finally, we looked at the findings of the Alexa web analytic tool and automatic link checkers in the context of heuristic and user testing.

- **Website Quality**

Based on experts' evaluation and user testing, the ranking of Saudi municipalities by website quality is stated in Table 6.38. Jeddah was the leading e-municipality in website quality at 71.1% of maximum points, Qassim 61.7%, Al-Madinah 55.9%, Riyadh 49.4%, and finally Eastern Region site by 45.4%.

Table 6.38: Ranking of Saudi City Websites by Website Quality

City Website	Total Score for Website Quality based on Heuristic and User Testing	Rank
Jeddah	27.71 (71.1%)	1
Qassim	24.08 (61.7%)	2
Al-Madinah	21.79 (55.9%)	3
Riyadh	19.27 (49.4%)	4
Eastern Region	17.71 (45.4%)	5

- **E-services**

The ranking of Saudi e-municipalities by e-services are presented in Table 6.39. The average for one e-service was calculated from the results of user testing of five e-services. Then that number was multiplied by the number of e-services to obtain a total score for all e-services in each city website. Accordingly, Riyadh was the leading e-municipality in e-services. Obviously, a limited number of e-services was offered through Saudi city portals. For example, the leading city in e-services, Riyadh, had only 31 e-services. Moreover, it was noticed that the average for Al-Madinah's e-services is the highest, 7.16. That can be attributed to the limited number of offered e-services on the site (16) and to the type of e-services which are a very simple inquiry. Generally, the five Saudi city websites implemented mainly simple e-services that didn't involve e-payment transaction at all.

Table 6.39: Ranking of Saudi City Websites by E-services

City Website	#Functioning E-services	Average for Five E-services	Average for One E-service	Average for all E-services	Rank
Riyadh	31	6.98	1.4	43.4	1
Jeddah	25	5.67	1.13	28.25	2
Al-Madinah	16	7.16	1.43	22.88	3
Eastern Region	12	3.84	0.77	9.24	4
Qassim	8	5.12	1.02	8.16	5

- **Overall Ranking**

The total for an e-municipality equals the total score for website quality plus the total score for e-services. In that regards, Riyadh was number one with a total of 62.67 followed by Jeddah 55.96, Al_Madinah 44.67, Qassim 32.24 and Eastern Region 26.95 (Table 6.40). It is worth noting that the total can accommodate an unlimited number of e-services. It obvious that Saudi city websites are poor and outdated. They provide little information, few simple e-services and are not citizen-centered websites.

Table 6.40: Overall Ranking of Saudi City Websites

City Website	Scores for Website Quality	Scores for E-services	Total	Rank
Riyadh	19.27	43.4	62.67	1
Jeddah	27.71	28.25	55.96	2
Al-Madinah	21.79	22.88	44.67	3
Qassim	24.08	8.16	32.24	4
Eastern Region	17.71	9.24	26.95	5

- **User Satisfaction Survey**

User satisfaction survey revealed the following results:

a. All users said that they did not believe that any city website was excellent in any of the four examined features (website quality, e-services quality, ease of using the site, number of e-services).
b. Only the Eastern Region was graded low-fair-poor on the scale of the evaluation, while the other e-municipalities ranged from good-fair-poor.
c. All users said that they had not used any of the five city portals for a year. Nor had they depended on it to achieve their municipal e-services. That means there was no adoption of e-government at the municipal level in the opinion of the six users.

- **Alexa Testing**

As previously mentioned in section 6.3.1, Alexa is an indicator rather than a metric tool to assess a website because it is clearly not reliable. However, web traffic data contains rich

information on how websites are being used that may otherwise be unavailable to us. A good solution is to compare Alexa data with other well-established web evaluation methods to find out if the data are useful. Therefore the following observations can be made:

- Download speed: Alexa showed that all Saudi municipal websites were slow except Jeddah. In user testing, users complained about the slowness of the Riyadh municipal website while fewer complaints were recorded for Al-Madinah, Qassim and Eastern Region.
- Where visitors go on the site: Alexa indicated that there was a problem with the e-services in most Saudi municipal websites. In fact, citizens' visits to e-services were very low except the Al-Madinah site (49.53%). To some extent, this correlated with the heuristic and user testing which showed that all Saudi municipalities suffered from limited e-services on their websites. In the satisfaction survey, users said they were not using online services to accomplish their municipal transactions.
- Time on site: Alexa showed that visitors stayed on Saudi municipal sites for a short period of time, except Al-Madinah 15 minutes, which may mean a low interest of visitors in these websites. This is contradicted by users finding Jeddah, for example, a more interesting website with a lot of information about the city.
- Number of sites linking in: Alexa reflected that Saudi city websites were not popular and suffered from a weak connection to others on the web, since all had a very low number of sites linking in. In general, information about popularity cannot be deduced from user testing or heuristic evaluation.
- Bounce rate: all did well except Riyadh and Jeddah, in which the percentage of visitors leaving the site after viewing one page was as high as 54.5% and 41.1% respectively. As explained before in section 4.3.1, Alexa data analysis is complicated and the motivation behind visitors' actions unknown; thus sometimes a high bounce rate is bad but in other cases it may be good. If we look into the overall ranking of Saudi city websites based on heuristic and user testing, we find Riyadh number one and Jeddah number two, much better than other Saudi e-municipalities.
- Local traffic rank: Alexa revealed that the highest traffic ranked city sites in Saudi Arabia were Jeddah (ranked at 1237), Riyadh (1417), Qassim (3282), Al-Madinah (4249) and Eastern Region (5616). Therefore local web traffic rankings were poor

for all Saudi city websites. Actually, this rank is a combination of the number of daily visitors and page views over a period of time and this kind of information simply couldn't be extracted from user or heuristic testing.

- **Broken Links**

Table 6.41 combines different results of evaluating broken links in the five Saudi city websites according to:

1. The heuristic test: two experts scored the guideline: all links were working properly, i.e. no broken links
2. The auto link checker tools: four auto tools (Dead Link Checker, Xenu, Broken Link Checker and Web Link Validator) estimated number of broken links in each city website
3. User testing: a number of malfunctioning e-services were identified so these services were excluded from the test.

Table 6.41: Broken Links Evaluation of Saudi City Websites

City Website	Expert1 Score (Out of 1)	Expert2 Score (Out of 1)	Dead Link Checker	Xenu	Broken Link Checker	Web Link Validator	#Malfunctioning E-services
Jeddah	1	0.5	4012	25725	4149	99	0
Riyadh	0.5	0.5	3656	109	45	49	8
Al-Madinah	0.5	0.5	2721	10688	0	3	5
Eastern Region	0.5	0.5	214	2636	1175	28	6
Qassim	1	1	63	64714	904	7	3

Moreover, some screen shots were presented of broken links in the heuristic evaluation section such as in Figures 6.4, 6.5, 6.10, 6.11, 6.12, 6.14, 6.19, 6.20, 6.22, 6.28. Also, most of the participants noted that Qassim was not a stable website; sometimes it loaded correctly but other times it displayed page not found error. Regarding link checkers, we can come to the conclusion that automatic link-checkers are not reliable tools even though some are useful and succeed in identifying some broken links. Other evaluation methods, such as heuristic and user testing, can only point to the existence of some broken links on all Saudi websites but cannot count the number of them manually due to a large number of links.

Moreover, the existence of a number of malfunctioning e-services is catastrophic in any e-government website. Therefore, the development of a valid automatic link checker tool is the only way to help website developers since other web evaluation methods cannot identify all possible broken links.

6.6 Usability Problems of Saudi Municipal Websites

The usability problems found by the heuristic evaluation and user testing for each Saudi city website are presented in Table 6.42. The usability problems discovered by the heuristic tests were drawn from the guidelines that municipalities violated when designing their websites (those scored 0 by the two experts) (see Table 6.8). Other usability problems were obtained from the participants in user testing and their comments. Moreover, all participants were unsatisfied in terms of offered e-services, because they were few (for example, the leading city in e-services, Riyadh, had only 31 active e-services), and in terms of the type of e-services for being simple and mostly not e-transactions. Generally, the five tested Saudi e-cities are not citizen-centric websites.

Table 6.42: Usability Problems of Saudi City Websites

City Website	Usability Problems by the two Experts	Usability Problems by User Testing
Jeddah	1. No municipal budget information 2. No city council meetings 3. No update date on the footer of web pages 4. No emergency alerts 5. No distinction between visited/unvisited links and misleading cues to click	6. Not useful content and not updated 7. Hard to find eComplaint 8. Bad interactive city map 9. Low-quality videos and slow download of some images 10. Not effective search service 11. Bad organization of some links 12. Disabling the Back button in the pages of e-services

City Website	Usability Problems by the two Experts	Usability Problems by User Testing
Riyadh	1.No municipal budget information 2.No city council meetings 3.No emergency alerts. 4.No distinction between visited/unvisited links and misleading cues to click 5.No eJob section 6. Missing about us (mayor corner, mission, and objectives) 7. No comment or eSuggest section 8. No citizen satisfaction survey	9. Not much useful and not up to date content 10.Bad eComplaint section 11.Bad interactive city map 12.Very few videos and not good images 13.Not effective search 14.Bad organization of some links 15.Long homepage and not utilizing well the white spaces 16.Annoying pop up ads 17.Slow website
Al-Madinah	1. No municipal budget information 2. No city council meetings 3. No update date on the footer of web pages 4. No emergency alerts 5. No FAQs section 6. Not targeting audience group such as citizen, business, etc.	7. Not much useful and up to date content 8.Bad eComplaint section and sometimes the page is unavailable 9.No interactive city map 10.No videos and some meaningless images 11.No search facility 12.Inappropriate names of links
Eastern Region	1. No municipal budget information 2. No city council meetings 3. No update date on the footer of pages 4. No emergency alerts 5. No distinction between visited/unvisited links and misleading cues to click 6. No eJob 7. Unimportant news to users 8. No citizen satisfaction survey 9. No English website 10. No indicator of a user is where on the site 11. Not targeting audience group such as citizen, business, etc.	12. Not useful and not up to date content 13. Not good eComplaint section 14. No interactive city map 15. No images and videos 16. No search facility 17. Serious problems in main links: sometimes un-clickable, bad names and not well organized 18. Bad website in general
Qassim	1. No municipal budget information 2. No city council meetings 3. No emergency alerts 4. No English website 5. No distinction between visited/unvisited links and misleading cues to click 6. No eJob.	7.Not useful and not up to date content 8. Not good eComplaint section 9. Bad interactive city map 10. No images and videos 11.Bad search facility 12.Bad organization and names of some links on the site

Chapter 7: Conclusion

7.1 Summary of the Study

The main objectives of Web Engineering are to promote the development of high quality and successful websites. A key enabler of website success measurement is website metrics which determine if a website performs to the expectations of the users and identify website design problems. Many researchers consider usability as the most important metrics for evaluating websites. Further, questions were raised about e-government evaluation and rankings and it was argued that existing frameworks have some methodological limitations. Citizens do not want an interface designed to reflect the internal bureaucratic structure, or promote the official chief, at the expense of granting citizens fast access to the needed services and information. A good e-government evaluation framework at local level still needed to be developed.

Therefore, the main aim of this research was to develop a comprehensive framework for evaluating any city website in the world, to address a specific gap in the literature regarding the lack of such framework. In the process of developing this framework, it was found that the literature on measuring the quality of website is limited and more research is still needed in that area. Researchers were often confused and unable to choose the appropriate method(s). To address this limitation in the web evaluation, this study also contributed to the classification of web evaluation methods and proposed the appropriate methods for testing e-government websites.

The proposed E-City Usability Framework integrated 3-dimensional assessment measures: website quality metrics, e-services quality metrics, and the number and type of e-services. To refine the framework, pilot test, double-expert review, and application on a high-ranked city website were conducted. Also inter-rater reliability, as a percent agreement on evaluation between different experts, was checked to validate the study result.

After that, web evaluation methods were classified and the appropriate methods were identified. Selected Saudi city websites were tested by four web evaluation methods:

heuristic evaluation, user testing, automatic link checkers, and the Alexa web analytics tool. First, two experts heuristically evaluated these city websites based on the website quality objective guidelines form of the proposed e-city framework (33 guidelines). Second, user testing was conducted with six users assessing: six subjective website quality guidelines (34 to 39), five e-services, and answering a user satisfaction questionnaire. Third, Alexa was utilized to calculate nine metrics and the validity of this tool was discussed. Fourth, seven link-checkers were applied on Saudi city websites and the results were compared to find out which is the most reliable tool, if found.

The result showed that the best web evaluation methods to test e-city websites are the heuristic evaluation and user testing. The other two methods, automatic link checkers and Alexa tool, are unreliable tools to assess a website; they merely are indicators rather than metric tools. Also, it was found that users of the tested Saudi city websites suffered from many usability problems and were not satisfied with the offered information and services.

7.2 Achieving the Aims

The main aim of this research, to develop an evaluation framework for city websites in an attempt to raise awareness of usability and web evaluation methods to gain the benefits of e-government, was met through developing an E-City Usability Framework. To accomplish that, the first research question (1a), what are the major national and local e-government evaluation frameworks and what are their strengths and weaknesses, was addressed through extensive literature review. It was a challenging task, especially for local e-government, since these frameworks are limited within the literature with a lack of studies that combine, classify and assess them. We proposed to classify the e-government models into three kinds: organization and consultancy firms, scholars, and official government models. Major organizations developing national e-government evaluation frameworks were Accenture, Brown University, UN, and Capgemini Europe, while two examples of scholars' models were Gartner and Layne and Lee, and a representative of official government guide was the Research-Based Web Design and Usability Guidelines.

Most of these national models were not validated empirically, focused on the government not citizens' side, and they had assessed e-government websites in terms of

evolutionary stages which was an inefficient approach. Another problem was that most studies focus on the national level and neglected local e-governments even though municipal websites are the closest to people's life. Local e-government evaluation frameworks were few: the Digital Governance in Municipalities Worldwide, the MeGAP, the Community Benchmarks Program, and the KEeLAN frameworks. The drawbacks of such frameworks were that they are country specific, lack a theoretical foundation, and focus on the government side only. The strengths and weaknesses of e-government frameworks at national and local levels were detailed in sections 2.6.1.5 and 2.6.2.6 respectively, while the comparison between them was presented in section 2.6.2.7. It was concluded that a good evaluation framework at the national level and local level was still lacking.

Moreover, this research's proposed 3-dimension framework, measuring website and e-services qualities in addition to the number and type of e-services, is a contribution to the fields of e-government and web evaluation methods. The method to create the proposed e-city framework could be considered a novel advance as a general and comprehensive approach to tackle the problem of developing an e-city framework. It tried to bypass the limitations of existing e- evaluation frameworks while building on the strength of ten government models. The heuristics shared in common by the majority of these models were selected if they fulfilled one of the design principles of the g-quality e-government inspection method by Garcia et al. (2005). Further, these common selected guidelines were checked against the Folmer et al. (2003) usability framework to identify affected usability quality attributes.

The second research question (1b), are the metrics defined for national e-government suitable for assessing e-municipalities, was examined. Very few studies in the literature discussed this issue. However, it was found that the metrics for national e-government were not suitable for assessing e-municipalities. In other words, a framework for local e-government is different than a framework for national e-government. Consequently, the researcher developed a citizen-centered city framework. The different role of cities is one of the challenges that must be addressed in the proposed E-City Usability Framework.

To test the developed framework on Saudi municipality websites, a search was conducted to identify, classify, and choose the appropriate web evaluation method(s). Only link analysis and Google Analytics were excluded from the methods of testing in this research. Link analysis was found to be unreliable and Google Analytics required inserting codes into webpages which were not allowed by governments. Therefore the appropriate web evaluation approaches to test selected Saudi city websites were confined to four methods: heuristic evaluation, user testing, automatic link checkers, and the Alexa web analytics tool.

Accordingly, the research question (2a) what is the ranking of Saudi municipality websites by website quality, e-services, and overall, can be answered based on the heuristic and user testing methods. The ranking of Saudi cities by website quality was: Jeddah leading at 71.1% of maximum points, Qassim 61.7%, Al-Madinah 55.9%, Riyadh 49.4%, and Eastern Region 45.4%. While the ranking of Saudi e-municipalities by e-services was: Riyadh with 31 e-services scored on average 43.4, Jeddah with 25 e-services scored 28.25, Al-Madinah with 16 e-services scored 22.8, Eastern Region with 12 e-services scored 9.24, and Qassim with 8 e-services scored 8.16. The overall ranking was: Riyadh with a total of 62.67 scores followed by Jeddah 55.96, Al-Madinah 44.67, Qassim 32.24 and Eastern Region 26.95.

From the experts' evaluation, the result of ranking e-cities, based on the number of fully compliant guidelines, was: Jeddah followed 20 guidelines (60.6%), Qassim 15 guidelines (45.5%), Al-Madinah 12 guidelines (36.4%), Riyadh and Eastern Region 10 guidelines (30.3%). Therefore, all tested websites obtained a low adherence rate with the proposed guidelines.

Web usage data are challenging for researchers and it is a fertile field yet unexplored. This research found that Alexa is an indicator rather than a metric tool to evaluate a website. Also, its web ranking is not accurate nor reliable due to its methodology which relies on sampling by installed tool bars, without any rules for measuring metrics, no guidelines on how to deduce valuable information from a large amount of raw data and no firmly established correlation between Alexa's data and the performance of e-municipalities. To overcome this situation, we proposed to compare Alexa's data with other web evaluation methods but found that Alexa's results either correlated, contradicted, or

could not be extracted from heuristic or user testing. As an example of its correlation, Alexa indicated there was a problem on the e-services on most Saudi municipal websites which is the same result obtained from the heuristic and user testing. As an example of contradiction, the short time on Jeddah site indicated by Alexa opposed users' findings that Jeddah was an interesting website in terms of information offered. Examples of metrics that can be extracted from Alexa but not found by the heuristic or user testing were the number of sites linking in and global web traffic rank.

Combining the results of broken links from the heuristic evaluation, user testing, and several automatic link checker tools showed that link checkers were not reliable tools even though they succeeded in identifying some broken links. On the other hand, heuristic and user testing can only point to the existence of some broken links on Saudi tested websites but cannot count the number of them manually due to a large number of links. Unfortunately, a number of broken e-services links were found in four, out of five, Saudi city websites.

The last research question (2b), what are the major usability problems affecting Saudi citizen use of these websites, was tackled. Usability problems found by the heuristic evaluation and user testing were identified for each Saudi city website (see Table 6.42). The most common usability problems in Saudi city websites were the site was not user-centered, limited e-services and information were offered, and most e-services were simple and not online transactions. The tested websites lacked: useful content, important e-services, valuable information, emergency alerts, municipal budget, city council meetings, the distinction between visited and unvisited links, and dated news. They also suffered from broken links, inactive city map, lack of an eComplaint section, and nonfunctioning search facility.

To conclude, this study contributes to the fields of usability and e-government website evaluation in the following aspects:

1) The E-City Usability Framework integrated 3-dimension measures: website quality, e-services quality, and the number and type of e-services since two of these dimensions were absent and local e-government is neglected from existing evaluation frameworks.

2) The E-City Usability Framework is unique in how it was developed: 1) it has a theoretical base since it was built upon the strengths often e-government models; 2) the method of selecting web metrics in this research was based on the g-quality inspection method developed and tested by Garcia et al. (2005), then later validated empirically by Granizo et al. (2011).
3) This research evaluated several e-government dimensions: output, outcomes of citizen-centricity, and model-based assessment, since most research assessed a narrow aspect of the e-government topic area.
4) The E-City Usability Framework contributed on how to measure the impacts or outcomes of e-government through: a) defining metrics for website quality, b) defining metrics for e-services quality, c) defining scoring method to assess the number and type of e-services, and d) conducting user testing to check user satisfaction with a city website.
5) Using triangulation web evaluation methods to test Saudi city websites (heuristic evaluation, user testing, link checkers, and Alexa) is a good opportunity to compare and check the possibility of incorporating more methods into the assessment process. Triangulation of method provides more comprehensive findings than an individual approach.
6) The reliability of automatic link checker and Alexa was addressed, since most studies use them without questioning their credibility.
7) Defining the current state and ranking of Saudi city websites and determining the potential problems encountered by users when visiting these websites, since few studies existed. The results are important and hope to benefit Saudi municipalities and their web developers in order to improve Saudi city websites for the efficiency of the Saudi government and the satisfaction of their citizens.

7.3 Discussion

7.3.1 Implications for Practice

There is a debate in the literature regarding whether user testing or heuristic evaluation is a better method in terms of detecting web design problems. Huang and Benyoucef (2014) believed that heuristic evaluation is better, while Krenk and McComb (2012) suggested that user testing is better for evaluating dynamic websites and heuristic evaluation for static websites. Other researchers, such as Joe et al. (2015), remain neutral and stated that

heuristic evaluation complements user testing but is not a replacement. This research found user testing is better in evaluating e-services, while heuristic evaluation is better in assessing website quality. This is based on the researcher's observations of user testing, on experts' comments, and on literature review. For example, an expert may discover design problems, such as the bad organization of links as a reason for not finding easily a specific information on the website, while a user may quit a task, feel frustrated and unconfident. On the other hand, Codagnone et al. (2015) give an example of trying several services that were scored by experts as transactional and they discover the reality is different. At the end of completing the online procedure, the authors received a 'pdf' form to be delivered in person to the public office. Thus, a service which should have been scored 3 (2-way interaction) was instead assigned a score of 4 (fully transaction) by experts. Hasan (2009) argued that heuristic evaluators cannot play the role of users and cannot judge the severity of usability problems in an interface for actual users. Anyway, the debate is still open and no consensus as to which web evaluation method is better in identifying usability problems (Krenk, & McComb, 2012).

In addition, it was expected that e-government would lead to a wide range of benefits. Recently, the importance of the e-government outcomes has been recognized but the research is still in its infancy. De Róiste (2013) stated that the current state of e-government evaluation ignores citizens' demand, usability measurements, and the more abstract goals of e-government such as transparency and public participation. Berger (2015) argued that if citizens' demand for e-government does not meet their expectations, governments might not achieve the expected outcomes from e-government. Alshibly and Chiong (2015) prefer to view citizens as customers and regard user satisfaction as the most important proxy of e-government success. The research described in this thesis has led to a total agreement with all three studies.

Tsohou et al. (2013) stated that there is a need for citizen-centric e-government evaluation. To fill the gap in the e-government impact studies, Andersen et al. (2011) measure the effectiveness and efficiency of e-mail response from a user perspective. Although Denmark is highly ranked in international e-government benchmarking, the authors found slow and incomplete responses by especially central government. This

researcher found that citizen-centric approach is the appropriate method which can connect the output (information and services on the website) with the expected outcomes from e-government (such as quality of services and citizen satisfaction). The outcomes or benefits from e-government are very important in determining the success of e-government. Citizen satisfaction is influenced by the output, the quantity, and quality of e-services and e-information, all of which we included in the proposed framework. That is, citizens are satisfied if they find valuable information and a variety of good e-services.

Schellong (2010) pointed out that there are no good metrics and no clear understanding of how citizen-centricity should be measured. Specially, outcomes of multiple factors, such as citizen satisfaction, are not easy to gauge. Grönlund (2010) considered many measures of "better government" (e.g. transparency and accountability) are shallow. The future e-government research must contribute to define ways of assessing better government. He assured that implementing "full case handling" is understood while using ICT to make government better is still a great challenge. The E-City Usability Framework contributed to measure the outcomes of e-government through: 1) defining metrics for website quality, 2) defining metrics for e-services quality, 3) defining a scoring method to assess the number and type of e-services, and 4) conducting user testing to hear citizens' voices and check user satisfaction with a city website and its e-services.

Moreover, we agree with other scholars (Codagnone et al., 2015; Siskos et al., 2014; De Róiste, 2013; Grönlund, 2011; Andersen et al., 2011; Rorissa et al., 2011; Janssen, 2010; Montserrat, 2010; Bannister, 2007) who said e-government ranking of nations is meaningless and questioned the validity of benchmarks. As an evidence, most studies found Saudi e-government in an early stage of development (see section 2.7.6), however, Saudi e-government is always improving dramatically in the UN ranking. That means the UN e-government ranking, as assured by Codagnone et al. (2015), Montserrat (2010), Salem (2008), and Bannister (2007), is a booming business rather than a scientific evaluation of e-government.

Further, there is a debate in the literature regarding whether the field of e-government has firm theoretical foundations. Grönlund (2011) argued that the e-government doesn't

have theoretical bases and Tsohou et al. (2013) stated that e-government evaluation is immature and ambiguous. On the contrary, Bannister and Connolly (2015) confirm that e-government has begun to develop as a field but the progress is slow. It is under-theorised but the claim that theory is absent from e-government literature is not in line with evidence. We think even though there are a lot of e-government studies around, they all seem to revolve in the same area as a field of study. E-government is at an early stage of development for a long time. After about 15 years of adoption and high government budget spending, e-government failed to deliver its promises and has not reached a higher stage of development in most countries. That doesn't only mean that progress is slow but also that there is something wrong in both the field of study and government interest in the subject.

7.3.2. Generalization of the Proposed Framework

Among the aims of this research was the possibility of the generalization of the proposed E-City Usability Framework to assess any e-city website. In general, testing a city website is a challenging and time-consuming task. If one wants to generalize the proposed framework to test a sophisticated city website that has many e-services, then the researchers may face a problem on how to test the e-services and how to select the sampling evaluation tasks. This problem faced us when assessing the NYC website which has so many e-services. Fortunately, the aim of this step was to test the validity and the applicability of the proposed framework and not to rank the NYC website, so the tested tasks have not been a problem and were carefully chosen with the aid of the proposed guidelines. In the literature, limited studies mentioned this problem of testing a large amount of information and services available on e-government websites. Bannister (2007) questioned the credibility of some e-government models and how thorough the tests of e-services can be when one is looking at hundreds of them in a short period of time. Also, Donker-Kuijer et al. (2010) stated that in light of the size of most e-government websites, usability testing of all parts does not seem feasible within time and financial constraints. They suggested that expert evaluations might be a good solution. We disagree with this solution since sometimes experts' evaluation alone of e-services could be inaccurate; they look at the description of a service but actually do not execute it online, as explained in the study of Codagnone et al. (2015).

If we look for a solution to this problem in existing well-known e-city frameworks, such as the UN Digital Governance in Municipalities, we found that the UN framework did not provide detailed evaluation methods for testing e-services. It is not clear how the UN framework selects a sample of thousands of e-services and on what basis. Also, it is not clear whether the UN framework is evaluating the quality of e-services or only counting them. It seems that this framework is employing heuristic tests but user-testing is not used at all. User testing is an important web evaluation method and a sample of e-services must be tested to discover usability problems facing real users. How to select the evaluation tasks and which e-services to be tested are a real challenge for any researcher and that has not been addressed yet.

Therefore to a large extent, the proposed e-city framework could be generalized but practically a large number of experts is needed to solve the problem of exploring and choosing the appropriate sampling e-services tasks. A possible approach is that double-experts explore and categorize the e-services (such as simple inquiry e-services, e-transaction, e-payment, and transformation services) on the site, then a sample from each type of e-services can be selected as possible targets for evaluation tasks. Finally, some of these tasks could be assessed by experts and others by participants.

7.4 Limitations

It is beyond the focus of this research to consider the back-office operation because of the complicated nature and the difficulty of access and disclosure of Saudi back-office government.

As described in section 1.2, G2G, G2B, and G2E were excluded from the evaluation of this study since it is too broad and unfeasible to evaluate the four sections of e-government. Since we regard the relation between government and citizens as the most important proxy of e-government success, we choose G2C as the scope of this study.

There is no consideration and no metrics for e-democracy or e-participation in the proposed framework since it is not currently based on theory or sound reasoning. We believe that e-government development will not reach e-democracy for political reasons.

That is in line with many studies (Debri, & Bannister, 2015; Norris & Reddick, 2013; Grönlund, 2011, Klievink & Janssen, 2009; Coursey, & Norris, 2008) which consider e-democracy not related to the real world of government and reflects only hopes and aspirations.

Another limitation of this research is that it does not cover in its website evaluation other delivery channels of e-government such as mobile government and SMS (Short Message Service), nor assesses municipalities' usage of social media such as Twitter, Facebook, Instagram, and YouTube.

7.5 Future Work

The Web is a dynamic environment by its nature and usability evaluation is a continuous process that must be conducted iteratively on a regular basis, therefore this study can be extended in the near future by reevaluating Saudi city websites to monitor how well they are progressing and whether there are any changes on their website design.

It would be interesting to conduct an analytical study of Saudi city website development feasibility by sending questionnaires to Saudi e-city developers and webmasters in order to know what they follow in terms of usability guidelines or e-government frameworks.

The proposition stated in this research, that user testing is better in evaluating e-services, while heuristic evaluation is better in assessing website quality, deserves more research. Also how to measure the outcomes of e-government through user-centricity indicators (such as efficiency, cost reduction, transparency, less corruption, service quality, and citizen satisfaction) remain an area of a great challenge that needs further research.

Moreover, very few studies have been carried out using Alexa web traffic data or assessing its reliability, and the questions remain about the relative advantages and disadvantages of this tool. The correlation between Alexa and the performance of e-municipalities has not been firmly established yet. In fact, this tool contains rich data on how websites are being used but we need to find systematic ways to extract useful information from such a large amount of raw data. Additionally, more research is needed

to gain broader knowledge into how Alexa data can complement established web evaluation methods, such as heuristic evaluation or user testing. Also, link checking remains an unexplored field since to our knowledge we could not find a publication that compares link checker tools or identifies the best.

As explained before, the generalization of the proposed E-City Usability Framework to assess any e-city website could be a valuable direction for future research. More investigation is needed on how to test a sophisticated city website that has many e-services. The procedure for selecting a representative sample from a large number of e-services is a bottleneck problem that needs to be addressed.

M-government (Mobile government) is the next generation of e-government and the future trend in contacting citizens with their governments. People are more eager to obtain government information and e-services through one device, their mobile phones. Governments can use m-government in a case of unexpected emergencies or natural disasters to quickly disseminate real-time needed information and publicize early disaster warnings to all citizens. However, m-government is still in its early stage of development and it has its own problems and challenges that must be tackled.

Saudi municipalities need to work hard on improving the online citizen-government relationships. Through all phases of website development, website usability should be considered and a designated usability expert on the web development team is necessary. A major challenge for municipality managers and webmasters remains not only to increase the overall level of e-information but go further for the more complex e-transactions. The more Saudi citizens are online than in line, the more confidence others will have in e-government and government and that definitely strengthens the relationship between the government and its citizens. Eventually, citizens can become partners in transforming Saudi society toward e-society and contributing to a better future for all.

References

Aghaei, S., Nematbakhsh, M. A., & Farsani, H. K. (2012). Evolution of the World Wide Web: From WEB 1.0 TO WEB 4.0. International Journal of Web & Semantic Technology, *3*(1), 1.

Ajaj, M. H. A. (2014). The role of local municipal branches in planning, developing and managing urban growth in Saudi Arabian cities: a case study of the city of Jeddah

Albassam, B. A. (2012). Political Reform in Saudi Arabia: Necessity or Luxury? MJSS, 1, 2.

Alexa. (2014). Retrieved 15.1.2014 from http://www.alexa.com/

Al-Fakhri, M., Cropf, R., Higgs, G., & Kelly, P. (2008). e-Government in Saudi Arabia: between Promise and Reality. International Journal of Electronic Government Research, *4* (2), 59-82.

Alfarraj, O., Drew, S., & AlGhamdi, R. (2011). EGovernment Stage Model: Evaluating the Rate of Web Development Progress of Government Websites in Saudi Arabia. (IJACSA) International Journal of Advanced Computer Science and Applications, *2* (9), 2011

Al-Ghaith, W., Sanzogni, L., & Sandhu, K. (2010). Factors Influencing the Adoption and Usage of Online Services in Saudi Arabia. The Electronic Journal of Information Systems in Developing Countries (EJISDC), 2010, *40* (1), 1-32

Ali, A. H. M., & Ahmad, N. (2015). HYBRID MEHODLODGY (HM) FOR DEVELOPING WEB-BASED SYSTEMS. International Journal of Information Research and Review, 2 (11), 1327-1357

Al-Juboori, A. F. M. A., Na, Y., & Ko, F. (2011). Web Site Evaluation: Trends and Existing Approaches. In Networked Computing (INC), The 7th International Conference, IEEE, 155-160.

Al-Khalifa, H. (2010). Heuristic Evaluation of the Usability of E-Government Websites: A Case from Saudi Arabia.ICEGOV2010, October 25-28, 2010, Beijing, China. ACM

Al-Nuaim, H. (2009). How "E" Are Arab Municipalities? An Evaluation of Arab Capital Municipal Web Sites. IGI Publishing. International Journal of Electronic Government Research, (1), 50-63.

AlSabti, K. (2007). The Saudi Government in the Information Society. National E-transactions Conference, Al-Riyadh, KSA

Al-Saif, M. (2010). The Awareness of the E-government System among Saudi Citizens. Master thesis, The University of Sheffield, UK, September 2010

Al-Shafi, S., & Weerakkody, V. (2010). Factors Affecting E-government Adoption in the State of Qatar. European and Mediterranean Conference on Information Systems 2010, April 12-13 2009, Abu Dhabi, UAE

Al-Shehry, A. (2008). Transforming Towards E-government in The Kingdom of Saudi Arabia: Technological and Organizational Perspectives. Doctoral Thesis, De Montfort University, Leicester, UK.

Al-Shehry, A., Rogerson, S., Fairweather, N., Prior, M. (2006). The Motivations for Change towards E-government Adoption: Case Studies from Saudi Arabia. eGovernment Workshop '06 (eGOV06), Brunel University, West London, UK.

Alshehri, M., & Drew, S. (2010). Challenges of e-Government Services Adoption in Saudi Arabia from an e-Ready Citizen Perspective. World Academy of Science, Engineering and Technology 66 2010

Alshibly, H., & Chiong, R. (2015). Customer empowerment: does it influence electronic government success? a citizen-centric perspective. Electronic Commerce Research and Applications, *14*(6), 393-404.

AlZahrani, A. (2011). Web-based e-Government Services Acceptance for G2C: A Structural Equation Modeling Approach. A doctoral Thesis, De Montfort University, Leicester, United Kingdom, March 2011

Andersen, K., & Henriksen, H. (2006). E-government Maturity Models: Extension of the Layne and Lee Model. Government Information Quarterly 23, 236 – 248. Available at www.sciencedirect.com

Andersen, K. N., Medaglia, R., Vatrapu, R., Henriksen, H. Z., & Gauld, R. (2011). The forgotten promise of e-government maturity: Assessing responsiveness in the digital public sector. Government Information Quarterly, *28*(4), 439-445.

Anthopoulos, L., Reddick, C. G., Giannakidou, I., & Mavridis, N. (2016). Why e-government projects fail? An analysis of the Healthcare. gov website.Government Information Quarterly, *33* (1), 161-173.

Arslan, A. (2008). Assessment of the Turkish Local e-Governments: An Empirical Study. International Journal of Human Sciences, *5*(2), 1303-5134. Available at: http://www.insanbilimleri.com

Ataloglou, M., & Economides, A. (2009). Evaluating European Ministries' Websites. International Journal of Public Information Systems, *5*(3), 147–177.

Bahareh, B. (2015). Enhancing usability of e-commerce platforms by utilizing the usability factors: An investigation into user preferences.

Bahrain eGovernment Authority (2010). User Interface Standards. Retrieved 15.2.2012 from http://www.bahrain.bh/pubportal/wps/wcm/connect/44ea880043946988846ebea42b14669a/UI_standardsV2.2%28NEAF%29.pdf?MOD=AJPERES&CACHEID=44ea880043946988846ebea42b14669a

Baker, D. L. (2009). Advancing E-Government performance in the United States through enhanced usability benchmarks. Government Information Quarterly, *26*(1), 82-88.

Bannister, F. (2007). The curse of the benchmark: an assessment of the validity and value of e-government comparisons. International Review of Administrative Sciences, *73*(2), 171-188.

Bannister, F., & Connolly, R. (2015). The great theory hunt: Does e-government really have a problem?. Government Information Quarterly, *32*(1), 1-11.

Barnes, S. J., & Vidgen, R. T. (2002). An integrative approach to the assessment of e-commerce quality. J. Electron. Commerce Research, *3*(3), 114-127.

Bartoo, G., & Bogucki, T. (2013). Essentials of usability engineering in point-of-care devices. In Point-of-Care Healthcare Technologies (PHT), 2013 IEEE,184-187. IEEE.

Berger, J. B. (2015). Formative evaluation and user engagement: A model to ensure value from e-government. In Electronic Government and Electronic Participation: Joint Proceedings of Ongoing Research, PhD Papers, Posters and Workshops of IFIP EGOV and EPart 2015, *22*, 193.

Berntzen, L., & Olsen, M. (2009). Benchmarking E-government- A Comparative Review of Three International Benchmarking Studies. 3rd International Conference on Digital Society IEEE, 77-82.

Bhat, M. H. (2013). Evaluating Indian Newspaper Websites using Alexa Internet. Library Review, 62(6/7), 5-5.

Bhattacherjee, A. (2012). Social science research: principles, methods, and practices. University of South Florida, Textbooks Collection. Book 3. http://scholarcommons.usf.edu/oa_textbooks/3

Blank, G. (2013). OxIS 2013 top-line findings: Internet use continues to grow; big increases in low-income households, the Oxford Internet Institute, University of Oxford.

Buie, E., & Murray, D. (2012). Usability in government systems: User experience design for citizens and public servants. Elsevier.

Buragga, K. (2010). Effectiveness of Accessibility and Usability of Saudi E-Government: an Objective Study. ICGST International Journal on Computer Network and Internet Research, CNIR, 2010, *10* (1), 7-19.

Butt, M. (2014). Result-oriented e-government evaluation: Citizen's perspective. Webology, *11*(2), 1.

Calero, C., Ruiz, J., & Piattini, M. (2005). Classifying Web Metrics Using the Web Quality Model. Online Information Review, Emerald Group Publishing Limited, 29 (3), 227-248

Calisir, F., Gumussoy, C. A., Bayraktaroglu, A. E., & Saygivar, E. (2011). Usability and functionality: A comparison of key project personnel's and potential users' evaluations. World Academy of Science, Engineering and Technology, 59, 204-209.

Capgemini (2006). Online Availability of Public Services: How is Europe Progressing? European Commission Directorate General for Information Society and Media

Capgemini (2009). Smarter, Faster, Better eGovernment: 8th Benchmark Measurement, November 2009. European Commission Directorate General for Information Society and Media

Capgemini (2010). Digitizing Public Services in Europe: Putting Ambition into Action. European Commission Directorate General for Information Society and Media

Cassino, R., Tucci, M., Vitiello, G., & Francese, R. (2015). Empirical validation of an automatic usability evaluation method. Journal of Visual Languages & Computing, 28, 1-22.

Chatfield, A., & Alhujran, O. (2009). A Cross-Country Comparative Analysis of E-Government Service Delivery among Arab Countries. Information Technology for Development, *15* (3), 151–170. Published online 15 June 2009 in Wiley InterScience (www.interscience.wiley.com).

Chiou, W. C., Lin, C. C., & Perng, C. (2010). A strategic framework for website evaluation based on a review of the literature from 1995–2006.Information & management, *47* (5), 282-290.

Chow, A. (2013). The Usability of Digital Information Environments: Planning, Design, and Assessment. Trends, Discovery, and People in the Digital Age, 13.

Chung, S. (2011). Fostering Citizen Participation Through Innovative Mechanisms In Governance, Policy, And Decision Making: Comparing Washington, DC, And Seoul. PhD Dissertation. ARIZONA STATE UNIVERSITY, May 2011

CITC. (2015). Communications and Information Technology Commission. Available at: http://www.citc.gov.sa/English/Pages/default.aspx

CITC (2008). Internet Usage in the Kingdom of Saudi Arabia. Retrieved 28.7.2013 from http://www.citc.gov.sa/English/Reportsandstudies/Studies/Documents/IT%20014%20E%20-%20IndividualReport2008.pdf

City Population. (2013). Retrieved 25.7.2013 from http://www.citypopulation.de/SaudiArabia.html

Clifton, B. (2012). Advanced web metrics with Google Analytics. John Wiley & Sons.

Codagnone, C., Misuraca, G., Savoldelli, A., & Lupiañez-Villanueva, F. (2015). Institutional isomorphism, policy networks, and the analytical depreciation of measurement indicators: The case of the EU e-government benchmarking. Telecommunications Policy, *39*(3), 305-319.

Connolly, R., Bannister, F., & Kearney, A. (2010). Government website service quality: a study of the Irish revenue online service. European Journal of Information Systems, *19*(6), 649-667.

Coursey, D., & Norris, D. (2008). Models of E-Government: Are They Correct? An Empirical Assessment. Public Administration Review, May/Jun 2008, *68*(3), 523-536

Cybermetrics Lab. (2014). Webometrics Ranking of World Universities. Available at: http://www.webometrics.info/

Debri, F., & Bannister, F. (2015). E-Government Stage Models: A Contextual Critique. In System Sciences (HICSS), 2015 48th Hawaii International Conference, 2222-2231. IEEE.

Denfeld , Z., Miller, A., Feinerman, F., Nine, C., Hacken, L., Reeder, K. , Hoyle P., & Tamoutselis, P. (2002). 2002 Assessment of Municipal Web Sites in Onondaga County. Community Benchmarks Program. Maxwell School of Citizenship and Public Affairs. Syracuse University.

De Róiste, M. (2013). Bringing in the users: The role for usability evaluation in eGovernment. Government Information Quarterly, *30*(4), 441-449.

Didraga, O., & Brandas, C. (2015). Comparative Study on E-Government Indicators between Romania and the European Union. Informatica Economica Journal, *19*(1), 67-76.

Dingli, A., & Mifsud, J. (2011). USEFul: A Framework to Mainstream Web Site Usability through Automated Evaluation. International Journal of Human Computer Interaction (IJHCI), *2*(1)

Dominic, P., & Jati, H. (2010). Evaluation Method of Malaysian University Website: Quality Website Using Hybrid Method. IEEE

Dominic, P., Jati, H., & Kannabiran, G. (2010). Performance Evaluation on Quality of Asian E-government Websites - an AHP Approach.

Donker-Kuijer, M. W., de Jong, M., & Lentz, L. (2010). Usable guidelines for usable websites? An analysis of five e-government heuristics. Government Information Quarterly, *27*(3), 254-263.

Dubai eGovernment Department (2011). Government Websites Excellence Model, Dubai Government Websites Guidelines, Version 3.0. Retrieved 15.2.2012 from http://www.deg.gov.ae/SiteCollectionImages/Content/DeG%20Documents/Dubai_Government_Website_Guidelines_English.pdf

Dubai eGovernment Department (2009). eServices Delivery Excellence Model, eServices Standards & Implementation Guidelines, Version 3.1. Retrieved 15.2.2012 from http://www.deg.gov.ae/SiteCollectionDocuments/Content/English/eServices%20Delivery%20Excellence%20Model.pdf

Dubai Smart Government (2012). e4all, the official magazine of Dubai Smart Government. Issue No. 99 - January 2012, p.10. Retrieved 7.8.2013 from http://www.deg.gov.ae/SiteCollectionImages/Content/DeG%20Documents/January2012_En.pdf

Eidaroos, A., Probets, S., & Dearnley, G. (2009). Heuristic Evaluation for e-Government Websites in Saudi Arabia. IN: Aloqbi, A., Alsini, I. and Alzimani, K.(eds.). SIC: The Third Saudi International Conference, 5th- 6th June 2009: Conference Proceedings.

Fath-Allah, A., Cheikhi, L., Al-Qutaish, R. E., & Idri, A. (2014). E-Government Maturity Models: A Comparative Study. International Journal of Software Engineering & Applications, *5*(3), 71.

Fernandez, A., Abrahão, S., & Insfran, E. (2013). Empirical validation of a usability inspection method for model-driven Web development. Journal of Systems and Software, *86*(1), 161-186.

Fernandez, A., Insfran, E., & Abrahão, S. (2011). Usability Evaluation Methods for the Web: A Systematic Mapping Study. Information and Software Technology, *53*(8), 789-817.

Ferras, C., Real, D., & Rosado, C. (2012). Web Traffic Map of Spanish Municipalities: Building a Ranking. In Information Systems and Technologies (CISTI), 2012 7th Iberian Conference on (pp. 1-6). IEEE.

Flak, L. Olsen, D., & Wolcott, P. (2005). Local E-Government in Norway: Current Status and Emerging Issues. Scandinavian Journal of Information Systems, *17*(2),41–84

Folmer, E., Gurp, J., & Bosch, J. (2003). A Framework for Capturing the Relationship between Usability and Software Architecture. Available at: http://publications.jillesvangurp.com/spisf.pdf

Fraefel, M., Selzam, T., & Riedl, R. (2013). Organizational requirements for building up national e-Government infrastructures in federal settings. In System Sciences (HICSS), 2013 46th Hawaii International Conference, 1642-1651. IEEE.

Frey, C. B., & Osborne, M. (2015). Technology at Work. The Future of Innovation and Employment, Citi GPS: Global Perspectives & Solutions

Garcia, A.C.B., Maciel, C. & Pinto, F.B. (2005). A Quality Inspection Method to Evaluate EGovernment Sites, Electronic Government: 4th International Conference, EGOV, August 22-26 2005, Springer Berlin / Heidelberg, 198-209

Geray, O., & Al Bastaki, M. (2005). Dubai eGovernment Initiative: Concept, Achievements and the Future Pillars of Success. Dubai eGovernment, UAE, UNDP/OECD Conference

Goldkuhl, G., & Persson, A. (2006). From e-ladder to e-diamond–re--conceptualising models for public e-services.

Granizo, C. J., Yanez, P. L., Ramirez, D. P., & Machado, P. C. (2011). Usability in E-government Sites. In Information Technology: New Generations (ITNG), 2011 Eighth International Conference, 453-458. IEEE

Grönlund, Å. (2010). Ten years of e-government: The 'end of history'and new beginning. In International Conference on Electronic Government,13-24. Springer Berlin Heidelberg.

Grönlund, Å. (2011). Connecting eGovernment to real government-the failure of the UN eParticipation index. In Electronic Government, 26-37. Springer Berlin Heidelberg.

Harter, D. E., Kemerer, C. F., & Slaughter, S. A. (2012). Does software process improvement reduce the severity of defects? A longitudinal field study. Software Engineering, IEEE Transactions, *38* (4), 810-827.

Hasan, L. (2009). Usability Evaluation Framework for E-commerce Websites in Developing Countries. A Doctoral Thesis. Loughborough University

Hasan, L., & Abuelrub, E. (2011). Assessing the quality of websites. Applied Computing and Informatics, *9*(1), 11-29.

Hasan, L., Morris, A., & Probets, S. (2012). A comparison of usability evaluation methods for evaluating e-commerce websites. Behaviour & Information Technology, *31*(7), 707-737.

Hassan, S., & Li, F. (2005). Evaluating the Usability and Content Usefulness of Web Sites: A Benchmarking Approach. Journal of Electronic Commerce in Organizations, 2005, *3*(2), 46-67

HHS. (2016). Research-Based Web Design and Usability Guidelines. Retrieved Jan 2016 from http://www.usability.gov

Holmberg, K. (2010). Co-inlinking to a municipal Web space: a webometric and content analysis. Scientometrics, *83*(3), 851-862.

Holzer, M., & Kim, S. (2003). Digital Governance in Municipalities Worldwide (2003): An Assessment of Municipal Websites throughout the World. E-Governance Institute, National Center for Public Productivity. Rutgers, the State University of New Jersey. Newark. Available at: http://unpan1.un.org/intradoc/groups/public/documents/aspa/unpan012905.pdf

Holzer, M., & Kim, S. (2005). Digital Governance in Municipalities Worldwide (2005): A Longitudinal Assessment of Municipal Websites throughout the World. E-Governance Institute,

National Center for Public Productivity. Rutgers, the State University of New Jersey. Newark. Available at: http://unpan1.un.org/intradoc/groups/public/documents/aspa/unpan022839.pdf

Holzer, M., & Kim, S. (2007). Digital Governance in Municipalities Worldwide (2007): A Longitudinal Assessment of Municipal Websites throughout the World. E-Governance Institute, National Center for Public Productivity. Rutgers, the State University of New Jersey. Newark. Available at: http://unpan1.un.org/intradoc/groups/public/documents/aspa/unpan031374.pdf

Holzer, M., You, M., Manoharan, A. (2009). Digital Governance in Municipalities Worldwide (2009): A Longitudinal Assessment of Municipal Websites throughout the World. E-Governance Institute, National Center for Public Performance, Rutgers University, Newark. Available at: http://www.scribd.com/doc/36936196/Digital-Governance-Municipalities-Worldwide-2009

Holzer, M., Manoharan, A. (2012). Digital Governance in Municipalities Worldwide ((2011-12): A Longitudinal Assessment of Municipal Websites throughout the World. E-Governance Institute, National Center for Public Performance, Rutgers University, Newark. Available at: https://spaa.newark.rutgers.edu/sites/default/files/files/EGov/Publications/Digital-Governance-Municipalities-Worldwide-2012.pdf

Holzer, M., Zheng, Y., Manoharan, A., Shark, A. (2014). Digital Governance in Municipalities Worldwide ((2013-14): A Longitudinal Assessment of Municipal Websites throughout the World. E-Governance Institute, National Center for Public Performance, Rutgers University, Newark. Available at: https://spaa.newark.rutgers.edu/sites/default/files/files/EGov/Publications/Digital-Governance-Municipalities-Worldwide-2014.pdf

Hu, Y., Xiao, J., Pang, j., & Xie, K. (2005). A Research on the Appraisal Framework of E-Government Project Success. Yat-sen University, China. ACM International Conference. Available at http://portal.acm.org/citation.cfm?doid=1089551.1089647

Huang, Z., & Benyoucef, M. (2014). Usability and credibility of e-government websites. Government Information Quarterly, *31*(4), 584-595.

Hussein, A. (2015). The use of triangulation in social sciences research: Can qualitative and quantitative methods be combined? Journal of Comparative Social Work, *4*(1).

Inc. Magazine. (2011). How to Reduce Your Website's Bounce Rate. Available at: http://www.inc.com/guides/2011/01/how-to-redce-your-website-bounce-rate.html

Internet World Stats. (2015). Available at: http://www.internetworldstats.com/
Internet World Stats. (2016). Available at: http://www.internetworldstats.com/stats.htm

Isa, W. A. M., Suhami, M. R., Safie, N. I., & Semsudin, S. S. (2011). Assessing the usability and accessibility of Malaysia e-government website. American Journal of Economics and Business Administration, *3*(1), 40.

Jalal, S. K., Biswas, S. C., & Mukhopadhyay, P. (2010). Web impact factor and link analysis of selected Indian universities. Annals of Library and Information Studies, *57*(2), 109-121.

Jansen, A., & Ølnes, S. (2013). Benchmarking E-Government Quality- Whose Quality Are We Measuring. In EGOV/ePart Ongoing Research, 43-53.

Jansen, A., & Ølnes, S. (2014). UNDERSTANDING PUBLIC E-SERVICES- Deconstructing e-Services to Identify Quality Aspects. In Norsk konferanse for organisasjoners bruk av IT, *22* (1).

Janssen, D., Rotthier, S., & Snijkers, K. (2004). If you measure it they will score: An assessment of international e-government benchmarking. Information Polity, *9*(3-4), 121-130.

Janssen, M. (2010). Measuring and benchmarking the back-end of e-Government: a participative self-assessment approach. In Electronic Government, 156-167. Springer Berlin Heidelberg.

Joe, J., Chaudhuri, S., Le, T., Thompson, H., & Demiris, G. (2015). The use of think-aloud and instant data analysis in evaluation research: Exemplar and lessons learned. Journal of biomedical informatics, *56*, 284-291.

Jowkar, A., & Didegah, F. (2010). Evaluating Iranian newspapers' web sites using correspondence analysis. Library Hi Tech, *28*(1), 119-130.

Jun, Z., Liangliang, C., & Fubin, L. (2009). ES-QUAL: Its applicability in evaluating E-government web sites service quality. In Information Engineering and Electronic Commerce, 2009. IEEC'09. International Symposium, 515-518. IEEE.

Karkin, N., & Janssen, M. (2014). Evaluating websites from a public value perspective: A review of Turkish local government websites. International Journal of Information Management, *34*(3), 351-363.

Kaur, A., & Dani, D. (2013). The Web Navigability Structure of E-Banking in India. International Journal of Information Technology and Computer Science (IJITCS), *5*(5), 29. Retrieved from http://www.mecs-press.org/ijitcs/ijitcs-v5-n5/IJITCS-V5-N5-4.pdf

Kaur, A., & Dani, D. (2011). The Systematic Review of Web Measurement. International Journal of Computer Science & Engineering Survey (IJCSES), *2*(3), 59-73.

Kaylor, C., Deshazo, R., & Van Eck, D. (2001). Gauging E-government: A Report on Implementing Services among American Cities. Government Information Quarterly *18*(2001), 293–307

Klievink, B., & Janssen, M. (2009). Realizing joined-up government—Dynamic capabilities and stage models for transformation. Government Information Quarterly, *26*(2), 275-284.

Kotamraju, N., & Der Geest, T. (2012). Behaviour & Information Technology, 2012, *31*(3), 261-273

Krenk, K., & McComb, S. (2012). Comparing Different Methods for Evaluating Usability of Website Interfaces. In IIE Annual Conference. Proceedings (1). Institute of Industrial Engineers-Publisher.

Kromidha, E. (2012). Strategic e-government development and the role of benchmarking. Government Information Quarterly, *29*(4), 573-581.

Kulkarni, R. B., & Dixit, S. K. (2012). Empirical and Automated Analysis of Web Applications. International Journal of Computer Applications, *38*(9), 1-8.

Kumar, N., Dadhich, R., & Shastri, A. (2015). Quality Models for Web-based Application: A Comparative Study. International Journal of Computer Applications, *125*(2).

Kumar, S. Sangwan, S. (2011). Adapting the Software Engineering Process to Web Engineering Process. International Journal of Computing and Business Research (IJCBR), *2*(1)

Kunstelj, M., & Vintar, M. (2004). Evaluating the Progress of E-government Development: A Critical Analysis. Information Polity, IOS Press, *9*(2004), 131–148

Lanvin, B., & Lewin, A. (2008). The Next Frontier of E-Government: Local Governments May Hold the Keys to Global Competition. The Next Frontier of E-Government, World Bank, 51-64

Lárusdóttir, M. (2009). Listen to Your Users: The Effect of Usability Evaluation on Software Development Practice. A Doctoral Dissertation, Department of Information Technology, UPPSALA University, Sweden

Layne, K., & Lee, J. (2001). Developing fully functional E-government: A four stage model. Government information quarterly, *18*(2), 122-136

Lee, Y., & Kozar, K. A. (2012). Understanding of website usability: Specifying and measuring constructs and their relationships.Decision support systems, *52*(2), 450-463.

Lewis, J. R. (2006). Sample sizes for usability tests: mostly math, not magic. Interactions, *13*(6), 29–33.

Li, X., Yu, Z., & Liu, C. (2010). Developing Usability Measure Structure: Process And Principals. Xi'an Research Inst. Of Hi-Tech, China. 2010 2nd International Conference on Computer Engineering and Technology, 4, 181-183

Liu, Y., Chen, W., Arendt, P., & Huang, H. Z. (2011). Toward a better understanding of model validation metrics. Journal of Mechanical Design, *133*(7), 071005.

Lofstedt, U. (2012). E-Government- Assessment of Current Research and Some Proposals for Future Directions. International journal of public information systems, 1(1).

Loiacono, E., Watson, R. T., & Goodhue, D. L. (2000). WebQual: A Web Site Quality. Instrument Working paper Worcester Polytechnic Institute.

Luna, D. E., Gil-Garcia, J. R., Luna-Reyes, L. F., Sandoval-Almazan, R., & Duarte-Valle, A. (2013). Improving the performance assessment of government web portals: A proposal using data envelopment analysis (DEA). Information Polity, *18*(2), 169-187.

Luna-Reyes, L. F., & Gil-García, J. R. (2011). Using institutional theory and dynamic simulation to understand complex e-Government phenomena. Government Information Quarterly, *28* (3), 329-345.

Mackenzie, N., & Knipe, S. (2006). Research dilemmas: Paradigms, methods and methodology. Issues in educational research, *16* (2), 193-205.

Malak, G., Sahraoui, H., Badri, L., & Badri, M. (2010). Modeling web quality using a probabilistic approach: An empirical validation. ACM Transactions on the Web (TWEB), *4* (3), 9.

Mikkonen, T., & Taivalsaari, A. (2010). The mashware challenge: bridging the gap between web development and software engineering. InProceedings of the FSE/SDP workshop on Future of software engineering research, 245-250, ACM.

Mofleh, S., Wanous, M., & Strachan, P. (2008). Developing Countries and ICT Initiatives: Lessons Learnt from Jordan's Experience. The Electronic Journal on Information Systems in Developing Countries (EJISDC), 2008, *34* (5), 1-17

Mohammad, H., Almarabeh, T., & Ali, A. (2009). E-government in Jordan. European Journal of Scientific Research, *35* (2), 188-197. Available at: http://www.eurojournals.com/ejsr.htm

Montserrat, J. (2010). Local e-Government Bench-learning: Towards a new methodological framework to benchmark electronic services provision and adoption in local public administration

Moon, M. (2002). The Evolution of E-Government among Municipalities: Rhetoric or Reality? Public administration Review, *62*(4), 424-434

Moraru, G. (2010). Anatomy of E-government: Assessment of Municipal E-government Services in Romania. CEU eTD Collection.

Mustafa, S. H., & Al-Zoua'bi, L. (2008). Usability of the Academic Websites of Jordan's Universities An Evaluation Study. Yarmouk University, Faculty of Information Technology2008.

Nam, T. (2014). Determining the type of e-government use. Government Information Quarterly, *31*(2), 211-220.

Nath, K., Dhar, S., & Basishtha, S. (2014). Web 1.0 to Web 3.0-Evolution of the Web and its various challenges. In Optimization, Reliabilty, and Information Technology (ICROIT), 2014 International Conference, 86-89. IEEE.

Nielsen, J. (1993). Usability engineering. San Francisco, CA: Morgan Kaufmann

Nielsen, J. (2011). Alertbox Column: Top 10 Mistakes in Web Design. Available at: http://www.useit.com/alertbox/9605.html

Nielsen, J. (2016). Nielsen's Articles. Available at: https://www.nngroup.com/articles/author/jakob-nielsen/

Norris, D. (2009). E-government among American Local Governments: Adoption, Impacts, Barriers and Lessons Learned. The International Conference for Administrative Development: Towards Excellence in Public Sector Performance, Riyadh, Saudi Arabia, November 1-4, 2009.

Norris, D. F., & Reddick, C. G. (2013). Local E-Government in the United States: Transformation or Incremental Change? Public Administration Review,*73*(1), 165-175.

Ølnes, S. (2013). Small is Beautiful or Bigger is Better? Size of Municipalities and Quality of Websites. In ECEG2013-13th European Conference on eGovernment: ECEG 2013 (389). Academic Conferences Limited.

Olsen, M., Nietzio, A., Snaprud, M., & Fardal, F. (2009). Benchmarking and Improving the Quality of Norwegian Municipality Web Sites. 5th Int'l Workshop on Automated Specification and Verification of Web Systems

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. Journal of Retailing, *64*(1), 12-40.

Paz, F., & Pow-Sang, J. A. (2014). Current Trends in Usability Evaluation Methods: A Systematic Review. In Advanced Software Engineering and Its Applications (ASEA), 2014 7th International Conference, 11-15, IEEE.

Peters, R., janssen, M. and Engers, T. (2004). Measuring E-government Impact: Existing Practices and Shortcomings. New York, USA: ACM, Available at http://www.leibnizcenter.org/docs/peters/peters_janssen_vanengers_egov_measurement_2004.pdf

Petkova, L. (August, 2012). Broken Links Software - XENU, Broken Links Checker or Link Tiger? Retrieved October 2014 from: http://blog.monitor.us/2012/08/broken-links-software-xenu-broken-links-checker-or-link-tiger/

Petricek, V., Escher, T., Cox, I., & Margetts, H. (2006). The Web Structure of E-Government - Developing a Methodology for Quantitative Evaluation. IW3C2, WWW 2006, Edinburgh, UK.

Pietilä, M. (2010). Toward Strategic Usability – User Knowledge as a Basis for New Service Development. Master Thesis, Helsinki University of Technology

Poeppelbuss, J., Niehaves, B., Simons, A., & Becker, J. (2011). Maturity models in information systems research: literature search and analysis.Communications of the Association for Information Systems, *29*(27), 505-532.

Purao, S., & Desouza, K. (2011). Looking for clues to failures in large-scale, public sector projects: A case study. In System Sciences (HICSS), 2011 44th Hawaii International Conference, IEEE, 1-10

Quantcast US Site Rankings (2016). https://www.quantcast.com/top-sites-1 (accessed 2 Aril 2016)

Richard. (February, 2014). 10 Essential Broken Link Checker Tools. Retrieved October 2014 from: http://codecall.net/2014/02/12/best-broken-link-checker-tools/

Rinder, J. M. (2012). The Importance of Website Usability Testing (Doctoral dissertation, University of Oregon).

Rio, A., & e Abreu, F. B. (2010). Websites Quality: Does It Depend on the Application Domain?. In Quality of Information and Communications Technology (QUATIC), 2010 Seventh International Conference, 493-498. IEEE.

Rodriguez, A. (2013). How to measure your tourism website's effectiveness with Alexa. The Business of Travel. Available online at: http://www.freenomads.com/blog/?p=1060#sthash.rCJEat2x.uDO8EivZ.dpbs

Rodríguez, R., Estevez, E., Giulianelli, D, & Vera, P. (2009). Assessing e-Governance Maturity through Municipal Websites Measurement Framework and Survey Results. 6th Workshop on

Software Engineering, Argentinean Computer Science Conference, San Salvador de Jujuy, Argentina, Oct 2009.

Rorissa, A. , Demissie, D. , & Pardo, T. (2008). Toward a Common Framework for Computing e-Government Index. ICEGOV2008, December 1-4, 2008, Cairo, Egypt.

Rorissa, A., Demissie, D., & Pardo, T. (2011). Benchmarking e-government: A comparison of frameworks for computing e-government index and ranking. Government Information Quarterly, 28(3), 354-362.

Sahraoui, S., Gharaibeh, G., & Al-Jboori, A. (2006). E-Government in Saudi Arabia: Can it Overcome Its Challenges? eGovernment Workshop '06 (eGOV06), September 11, 2006, Brunel University, West London, UK

Salem, F. (2008). Benchmarking the e-Government Bulldozer: Beyond Measuring the Tread Marks. Journal of Measuring Business Excellence, *11*(4), 9-22.

Sandoval-Almazan, R., & Gil-Garcia, J. R. (2012). Are government internet portals evolving towards more interaction, participation, and collaboration? Revisiting the rhetoric of e-government among municipalities. Government Information Quarterly, 29, S72-S81.

Sandoval-Almazan, R. & Gil-García, J. R. (2008a). E-Government portals in Mexico. Electronic Government: Concepts, Methologies, Tools and Aplications, *1*(1), 1726–1734.

Sandoval-Almazan, R., & Gil-Garcia, J. R. (2008b). Limitations of Evolutionary Approaches to E-government. Handbook of Research on Public Information Technology, Hershey, IGI Global, 22-30.

Saudi National e-Government Portal. (2013). Retrieved 25.7.2013 from www.saudi.gov.sa/

Saunders, M., Lewis, p., & Thornhill, A. (2012). Research methods for business students, 6th edition. Pearson Education.

Schedler, K., & Schmidt, B. (2004). Managing the e-government organization. International Public Management Review, *5*(1).

Schellong, A. R. (2010). Benchmarking EU e-government at the crossroads: a framework for e-government benchmark design and improvement. Transforming Government: People, Process and Policy, *4*(4), 365-385.

Scowen, G., & Regenbrecht, H. (2009). Increased popularity through compliance with usability guidelines in e-learning Web sites. International Journal of Information Technology and Web Engineering (IJITWE), *4*(3), 38-57.

Seffah, A., Donyaee, M., Kline, R., & Padda, H. (2006). Usability measurement and metrics: A consolidated model. Software Quality Journal, *14*, 159–178

Shareef, S. M., Jahankhani, H., & Dastbaz, M. (2012). E-government stage model: Based on citizen-centric approach in regional government in developing countries. International Journal of Electronic Commerce Studies, *3*(1), 145-164.

Shekofteh, M., Shahbodaghi,A., Sajjadi, S., & Jambarsang, S. (2010). Investigating Web Impact Factors of Type 1, Type 2 and Type 3 Medical Universities in Iran. Journal of Paramedical Sciences (JPS), *1*(3).

Shneiderman, B. (2011). Tragic errors: usability and electronic health records. interactions, *18*(6), 60-63.

Shneiderman B, Plaisant C. (2010). Designing the User Interface. 5th ed. Boston, MA: Pearson Addison-Wesley.

Singh, Y., Malhotra, R., & Gupta, P. (2011). Empirical Validation of Web Metrics for Improving the Quality of Web Page. Int. Jal. Advanced Computer Science and Applications, *2*(5).

Siskos, E., Askounis, D., & Psarras, J. (2014). Multicriteria decision support for global e-government evaluation. Omega, 46, 51-63.

Sleeper, M., Consolvo, S., & Staddon, J. (2014). Exploring the benefits and uses of web analytics tools for non-transactional websites. In Proceedings of the 2014 conference on Designing interactive systems, 681-684. ACM

Snead, J. T., & Wright, E. (2014). E-government research in the United States. Government Information Quarterly, *31* (1), 129-136.

Soames, C. (September, 2014). 5 essential link checking tools for SEOs, bloggers and content editors. Retrieved October 2014 from: http://www.smartinsights.com/search-engine-optimisation-seo/link-building/site-link-checking-tools/

Song, C. (2010). Validating IS Success Factors: An Empirical Study on Webbased State or Local E-government Systems. AMCIS 2010 Proceedings, 191

Stemler, S. (2004). A Comparison of Consensus, Consistency, and Measurement Approaches to Estimating Interrater Reliability. Practical Assessment, Research & Evaluation, 9(4), Yale University PACE Center. Retrieved March 26, 2013 from: http://PAREonline.net/getvn.asp?v=9&n=4

Stiglingh, M. (2014). A measuring instrument to evaluate e-service quality in a revenue authority setting. Public Relations Review, *40*(2), 216-225.

Stolz, C., Viermetz, M., Skubacz, M. (2005). Guidance Performance Indicator- Web Metrics for Information Driven Web Sites. IEEE/WIC/ACM International Conference on Web Intelligence 2005, Proceedings, 186–192

Stone, D., Jarrett, C., Wodroffe, M., & Minocha, S. (2005). User Interface Design and Evaluation. San Francisco, CA: Morgan Kaufmann Publisher, Elsevier

Sultana, S. (2015). A webometric analysis of private university websites in Bangladesh (Doctoral dissertation, University of Dhaka).

Thanuskodi, S. (2011). Webometric analysis of private engineering college websites in Tamil Nadu. Journal of communication, *2*(2), 73-81.

Thelwall, M. (2011). A comparison of link and URL citation counting. InAslib proceedings, *63*(4), 419-425. Emerald Group Publishing Limited.

Thelwall, M. (2012). A history of webometrics. Bulletin of the American Society for Information Science and Technology, *38*(6), 18-23.

The World FactBook. (2015). Available at: https://www.cia.gov/library/publications/the-world-factbook/index.html

Toonders, R. (2010). Dutch Municipality E-government: Explaining E-government Adoption and Development. Master Thesis. Eindhoven University of Technology

Torrecilla-Salinas, C. J., Sedeño, J., Escalona, M. J., & Mejías, M. (2016). Agile, Web Engineering and Capability Maturity Model Integration: A systematic literature review. Information and Software Technology, *71*, 92-107.

Treiblmaier, H., & Pinterits, A. (2010). Developing Metrics for Web Sites. The Journal of Computer Information Systems, Spring 2010, *50*(3).

Tsohou, A., Lee, H., Irani, Z., Weerakkody, V., Osman, I. H., Anouze, A. L., & Medeni, T. (2013). Proposing a reference process model for the citizen-centric evaluation of e-government services. Transforming Government: People, Process and Policy, *7* (2), 240-255.

United Nations. (2003). UN Global E-government Survey 2003. United Nations, New York, USA. Available at: http://unpan1.un.org/intradoc/groups/public/documents/un/unpan016066.pdf

United Nations. (2004). Global E-government Readiness Report 2004: Towards Access for Opportunity. United Nations, New York, USA. Available at: http://unpan1.un.org/intradoc/groups/public/documents/un/unpan019207.pdf

United Nations. (2005). Global E-government Readiness Report 2005: From E-Government to E-Inclusion. United Nations, New York, USA. Available at: http://unpan1.un.org/intradoc/groups/public/documents/un/unpan021888.pdf

United Nations. (2008). UN E-Government Survey 2008:From E-Government to Connected Governance. United Nations, New York, USA. Available at: http://unpan1.un.org/intradoc/groups/public/documents/un/unpan028607.pdf

United Nations. (2010). E-Government Survey 2010- Leveraging E-Government at a Time of Financial and Economic Crisis. United Nations, New York, USA. Available at: http://unpan1.un.org/intradoc/groups/public/documents/un/unpan038851.pdf

United Nations. (2012). E-Government Survey 2012: E-Government for the People. United Nations, New York, USA. Available at: http://unpan3.un.org/egovkb/Portals/egovkb/Documents/un/2012-Survey/unpan048065.pdf

United Nations. (2014). E-Government Survey 2014: E-Government for the Future we want. United Nations, New York, USA. Available at: http://unpan3.un.org/egovkb/Portals/egovkb/Documents/un/2014-Survey/E-Gov_Complete_Survey-2014.pdf

Valdés, G., Solar, M., Astudillo, H., Iribarren, M., Concha, G., & Visconti, M. (2011). Conception, development and implementation of an e-Government maturity model in public agencies. Government Information Quarterly, *28*(2), 176-187.

van Kuijk, J., van Driel, L., & van Eijk, D. (2015). Usability in product development practice; an exploratory case study comparing four markets. Applied ergonomics, *47*, 308-323.

Vaughan, L. (2008). A new frontier of informetric and webometric research: mining Web usage data. *Collnet Journal of Scientometrics and Information Management*, *2*(2), 29-35.

Vaughan, L., & Yang, R. (2013). Web traffic and organization performance measures: Relationships and data sources examined. Journal of Informetrics, *7*(3), 699-711.

Verdegem, P., & Verleye, G. (2009). User-centered E-Government in practice: A comprehensive model for measuring user satisfaction.Government Information Quarterly, *26*(3), 487-497.

Vigo, M., & Brajnik, G. (2011). Automatic web accessibility metrics: Where we are and where we can go. Interacting with Computers, *23*(2), 137-155. Available at: http://dx.doi.org/10.1016/j.intcom.2011.01.001

West, D. (2008). Improving Technology Utilization in Electronic Government around the World, Brookings, 2008. Available at http://www.brookings.edu/reports/2008/0817_egovernment_west.aspx

Wheeler, D. (2007). Digital Governance and Democratization in the Arab World. Idea Group Inc (IGI), 2007, Section: State E-Government and Country Profiles/ Category: Africa and Arab Countries, 327-335. Available at http://www.igi-global.com/viewtitlesample.aspx?id=11524

Wolfinbarger, M., & Gilly, M. C. (2003). eTailQ: dimensionalizing, measuring and predicting etail quality. Journal of retailing, *79*(3), 183-198.

Woolrych, A., Hornbæk, K., Frøkjær, E., & Cockton, G. (2011). Ingredients and Meals rather than Recipes: A Proposal for Research that does not treat Usability Evaluation Methods as Indivisible Wholes. International Journal of Human-Computer Interaction, *27*(10), 940-970

Yan, P., & Guo, J. (2010). The Research of Web Usability Design. Chongqing Three Gorges University. IEEE 2010.

Yesser (2015). Saudi E-government Program. Available at http://www.yesser.gov.sa/

Yildiz, M. (2007). E-government Research: Reviewing the Literature, Limitations, and Ways Forward. Government Information Quarterly *24*(2007), 646-665

Yoo, B., & Donthu, N. (2001). Developing a scale to measure the perceived quality of an Internet shopping site (SITEQUAL). Quarterly journal of electronic commerce, *2*(1), 31-45.

Zahran, D. I., Al-Nuaim, H. A., Rutter, M. J., & Benyon, D. (2015). A Critical Analysis of E-Government Evaluation Models at National and Local Municipal Levels. The Electronic Journal of E-Government, *13* (1), 28-42

Zahran, D. I., Al-Nuaim, H. A., Rutter, M. J., & Benyon, D. (2014). A comparative approach to web evaluation and website evaluation methods. International Journal of Public Information Systems, *10*(1).

Zhao, J. (2007). An Empirical Usability Evaluation of a Web-based Public Participation Geographic Information System and Discussion Forum. M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering Technical Report No. 250, University of New Brunswick, Fredericton, New Brunswick, Canada.

Zhao, F., José Scavarda, A., & Waxin, M. F. (2012). Key issues and challenges in e-government development: An integrative case study of the number one eCity in the Arab world. Information Technology & People, *25*(4), 395-422.

Appendices

Appendix A: Usability Tasks

Table A.1: Usability Tasks to Evaluate E-service Quality for Jeddah City Website

• Please perform the following tasks:	
Task 1: View building regulations in Basateen district (Abhur municipality) أعرض دليل اشتراطات البناء بحي البساتين (بلدية أبحر)	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1
2. Ease of navigation through the e-service process	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request	0___.25___.5___.75___1
4. E-services completely online if possible	0___.25___.5___.75___1
Comments:	
Task 2: Inquire about Saudis deaths during a period of two months and also for one year استفسر عن وفيات السعوديين خلال فترة شهرين وأيضاً لمدة سنة	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1
2. Ease of navigation through the e-service process	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request	0___.25___.5___.75___1
4. E-services completely online if possible	0___.25___.5___.75___1
Comments:	
Task 3: Report online about a drilling in King road, Marine Science square بلغ إلكترونياً عن حفر بطريق الملك ميدان علوم البحار	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1
2. Ease of navigation through the e-service process	0___.25___.5___.75___1
3. Simple forms with required fields and clear error message for invalid or incomplete data entry	0___.25___.5___.75___1
4. Online tracking for forms and e-services being processed	0___.25___.5___.75___1
5.Feedback when users waiting or submitting a request	0___.25___.5___.75___1
6. E-services completely online if possible	0___.25___.5___.75___1
Comments:	

Task 4: Inquire licenses shops transaction (# 320872) استعلم عن معاملة رخصة محلك (رقم المعاملة: 320872)	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1
2. Ease of navigation through the e-service process	0___.25___.5___.75___1
3. Online tracking for forms and e-services being processed	0___.25___.5___.75___1
4. E-services completely online if possible	0___.25___.5___.75___1
Comments:	
Task 5: Inquire health centers in Basateen and Faysaleyyah districts استعلم عن المراكز الصحية بحي البساتين وحي الفيصلية	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements, instructions and service centers' locations	0___.25___.5___.75___1
2. Ease of navigation through the e-service process	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request	0___.25___.5___.75___1
4. E-services completely online if possible	0___.25___.5___.75___1
Comments:	

Table A.2: Usability Tasks to Evaluate E-service Quality for Riyadh City Website

• Please perform the following tasks: فضلاً أنجز الخدمات التالية:	
Task 1: Inquire about your transaction # 12345 dated 1435h أستعلم عن معاملتك رقم 12345 عام 1435هـ	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 2: Inquire about Saudis deaths during a period of two months and also for one year استفسر عن وفيات السعوديين خلال فترة شهرين وأيضاً لمدة سنة	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 3: Know the engineering offices of building permits then search for “Knooz” office أستعرض المكاتب الهندسية لرخص البناء ثم أبحث عن مكتب "كنوز" أو "السويلم"	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1

4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 4: Inquire about healthy monitor # 3 استعلم عن المراقب الصحي رقم 3	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 5: View health certificates for your workers أستعرض الشهادات الصحية لعمالك	
Guidelines	Suggested Score
1.Sufficient information on e-services: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	

Table A.3: Usability Tasks to Evaluate E-service Quality for Al-Madinah City Website

• Please perform the following tasks:	فضلاً أنجز الخدمات التالية:
Task 1: Inquire about your royal grant (order # 1)	استعلم عن منحة برقم الطلب 1
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 2: Inquire about Saudis deaths during a period of two months and also for one year	استفسر عن وفيات السعوديين خلال فترة شهرين وأيضاً لمدة سنة
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 3: Inquire about Sultana street	استعلم عن شارع "سلطانة"
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1

4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 4: Inquire about your transaction # 123, year 1436h استعلم عن معاملتك برقم 123 لعام 1436	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 5: Inquire about investment contract # 1 in 1436h استعلم عن استثمار برقم العقد 1 عام 1436هـ	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	

Table A.4: Usability Tasks to Evaluate E-service Quality for Eastern Region Website

• Please perform the following tasks: فضلاً أنجز الخدمات التالية:	
Task 1: Report online about lights in day time (Enter required fields only) بلغ إلكترونياً عن أعمدة إنارة مضاءة في النهار بالدمام حي فيصلية (أدخل الحقول المطلوبة فقط)	
Guidelines	Suggested Score
1.Sufficient information on e-servicesas: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3. Simple forms with required fields and clear error message for invalid or incomplete data entry النموذج بسيط يحدد الحقول المطلوبة ورسالة الخطأ واضحة عند إدخال بيانات غير صحيحة أو ناقصة	0___.25___.5___.75___1
4. Online tracking for forms and e-services being processed إمكانية تتبع النموذج والخدمة عند التنفيذ	0___.25___.5___.75___1
5.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
6. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 2: Inquire about health certificates استعلم عن الشهادات الصحية	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 3: Inquire about investment contracts استعلم عن عقود الاستثمار برقم الهوية	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name,	0___.25___.5___.75___1

description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	

Task 4: Inquire about building permit by identification number استعلم عن رخصة البناء برقم الهوية	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	

Task 5: Inquire about transaction number 12466 dated 1436 استعلم عن معاملة رقم 12466 تاريخ 1436	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	

Table A.5: Usability Tasks to Evaluate E-service Quality for Qassim City Website

• Please perform the following tasks: فضلاً أنجز الخدمات التالية:	
Task 1: Inquire about your transaction # 123 for the year 1435h استعلم عن معاملتك رقم 123 لعام 1435هـ	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 2: See the official engineering offices such as Al-Rajhi office أعرف المكاتب الهندسية المعتمدة مثل مكتب الراجحي	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 3: Ask for a license to add floors to your home أطلب رخصة إضافة أدوار لمنزلك	
Guidelines	Suggested Score
1.Sufficient information on e-servicesas: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3. Simple forms with required fields and clear error message for invalid or incomplete data entry	0___.25___.5___.75___1

النموذج بسيط يحدد الحقول المطلوبة ورسالة الخطأ واضحة عند إدخال بيانات غير صحيحة أو ناقصة	
4. Online tracking for forms and e-services being processed إمكانية تتبع النموذج والخدمة عند التنفيذ	0___.25___.5___.75___1
5.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
6. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 4: Search for Job vacancies ابحث عن الوظائف الشاغرة بالأمانة	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	
Task 5: Explore Urban observatory for Buridah استعرض المرصد الحضري لحاضرة بريدة	
Guidelines	Suggested Score
1.Sufficient information on e-services such as: name, description, requirements and instructions معلومات كافية عن الخدمة مثل: وصف الخدمة، متطلباتها، وتعليمات استخدامها	0___.25___.5___.75___1
2. Ease of navigation through the e-service process سهولة الوصول للخدمة والتنقل خلالها	0___.25___.5___.75___1
3.Feedback when users waiting or submitting a request التجاوب والتفاعل مع المستخدم عند إنتظاره لتنفيذ الخدمة أو تقديم طلب	0___.25___.5___.75___1
4. E-services completely online if possible الخدمة كاملة عبر الانترنت	0___.25___.5___.75___1
Comments:	

Appendix B: Dubai Government Websites Excellence Models

- Accessibility (6 Guidelines):
    - Provide access to the website through an easy to remember URL including an appropriate representation of the entity name under (.gov.ae) domain.
    - Provide a quick access to the website from a search engine
    - Provide access to the website with identical and consistent results through a wide range of web browsers
    - Provide a functional bilingual website
    - Provide appropriate access to website files
    - Provide access to the website for people with disabilities
- Usability and Design (20 Guidelines):
    - Provide a clearly defined website header and footer
    - Provide a clear and readable entity & Dubai Government logos
    - Provide a functional link to the official portal of Dubai Government
    - Provide a well-designed customer focused Homepage
    - Provide a functional Homepage link available across all the website pages
    - Provide a well-structured and effective Sitemap
    - Provide an effective and efficient Search functionality
    - Provide a logically organized and easy to navigate website
    - Provide a proper and easy to use navigation facility
    - Use an appropriate design for website links
    - Provide clear and meaningful links on the website
    - Provide active internal and external links
    - Direct old website links to new and updated ones
    - Provide simple and easy to use forms
    - Provide proper and easy to understand guidelines for completing the online forms
    - Provide functioning and properly working forms
    - Provide a functional print facility on the website pages
    - Provide a consistent font style across the website pages

    - Provide a consistent format throughout the website
    - Provide well designed website page titles
- Content (17 Guidelines):
    - Provide information about the Government Entity in "About Us" Section
    - Provide Entity Contact information in "Contact Us" page
    - Provide a facility to submit feedback on the website
    - Provide effective and efficient Frequently Asked Questions (FAQ) on the website
    - Provide sufficient information about Government Entity services & eServices
    - Provide accurate website copyright information
    - Provide a proper "Site Maintained By" message
    - Provide a functional link to eJob
    - Provide a functional link to eSuggest
    - Provide a functional link to eComplain
    - Explain the complaint handling procedures on the website
    - Provide a functional link to Ask Dubai
    - Define/Use proper and meaningful metadata on almost every page on the website
    - Provide accurate dates on the website pages
    - Provide accurate and most up to date information on the website
    - Provide correct grammar and spellings content
    - Provide appropriate and well-designed online advertisements
- Policies (3 Guidelines):
    - Provide information regarding the protection and handling of privacy in the website
    - Provide information regarding the website terms and conditions
    - Provide information on the accessibility of the website.

Appendix C: Web Criteria Analysis

	Criteria	USA Guide	UN Guide	CBP Guide	Dubai Guide
1	Graphics should not look like banners ads	✓			
2	Budget information **(UN)** Current municipal budget **(CBP)**		✓	✓	
3	Quick access to the site from a search engine (in the top 10) (D) Be easily found in the top 30 (US) Placement of the site on Yahoo, Google, MSN for official and popular names **(CBP)**	✓		✓	✓
4	Homepage link is available across all pages (D) Links to the homepage on every page **(UN)** Enable access to the homepage ('Home' on top) (US) Link to home page from every page visited **(CBP)**	✓	✓	✓	✓
5	Provide useful content (US) Provide accurate and most up to date info (D)	✓			✓
6	Establish user requirements	✓			
7	Understand and meet user's expectations	✓			
8	Involve users in establishing user requirements	✓			
9	Set and state goals	✓			
10	Focus on performance before preference	✓			
11	Consider many user interface issues	✓			
12	Access the website through an easy to remember URL with appropriate representation of the entity name under (gov.ae) (D)				✓
13	Do not display unsolicited windows or graphics (US) Ensure that images do not slow downloads (US) Use video, animation and audio meaningfully (US) Provide customer focused homepage (limit heavy images, flash and video for quick download). (D) Examine graphics **(UN)**	✓	✓		✓
14	Access site through a wide range of browser (D) Design for common browsers (US)	✓			✓
15	Provide a functional bilingual website (D) Access in more than one language **(UN)**		✓		✓
16	Provide appropriate access to files (file name, date, description, size, format, link to free download needed program) (D) Downloadable documents **(UN)** Downloadable forms **(CBP)**		✓	✓	✓
17	Provide a clear header (logo, homepage link, search, About us, bilingual link) and footer (contact us, polices, copyright, last update, site maintained by)				✓
18	Provide a clear entity and Dubai government logo (Clickable logo directs to the homepage) (D) Include logos on every page (US)	✓			✓
19	Provide a link to the national e-government portal in the page header consistently (D) Listing of external links **(UN)** Links to related content: (government entities) (US) Increase website credibility: (logical organization, site looks professionally designed, archive past content, up-to-date, links to outside sources and link to other credible sites) (US)	✓	✓		✓

	Criteria	USA Guide	UN Guide	CBP Guide	Dubai Guide
20	Minimize the number of clicks or pages: the most common tasks completed on the fewest number of clicks (US)	✓			
21	Provide customer focused homepage (first good impression for users, About us, address user needs, quick access to highlighted e-services, grouped by **customer segments**, by service category, usage frequency, by need, include useful information to users, reasonable size of page for fast loading, limit scrolling, heavy images, flash, video for quick download). (D) Page Length (too long: 2 or more screen) **(UN)** Minimize page download time (US) Avoid cluttered displays (US) Targeted audience (citizen, business) **(UN)** Eliminate horizontal scrolling (US)	✓	✓		✓
22	Provide an effective sitemap.(D) Sitemap **(UN)** Ensure the homepage looks like a homepage (important links, sitemap, search) (US) Sitemap **(CBP)**	✓	✓	✓	✓
23	Provide an effective search working properly (D) Search tool **(UN)** Ensure the homepage looks like a homepage (important links, sitemap, search) (US) Ensure usable search results (US) Design search engines to search the entire site (US) Provide a search option on each page (US) Search capability to help the user access info more easily **(CBP)**	✓	✓	✓	✓
24	Provide logically organized and easy to navigate site (should be different from the rest of content, short and descriptive navigation menu title) (D) Easy to navigate site **(UN)** Differentiate and group navigation elements (US)	✓	✓		✓
25	Provide easy to use navigation facility: indicate where the user is in the site (Breadcrumbs), have a link to homepage) (D) Provide navigational options: don't disable back button because this confuse users (US) Provide feedback on users' location: (Breadcrumbs, change color of visited links) (US)	✓			✓
26	Use an appropriate design for links (underline, different colors for visited and non-visited links, avoid link style on non clickable content) (D) Blue links, purple visited links and underline to indicate link **(UN)** Provide feedback on users' location: (change color of visited links) (US) Avoid misleading cues to click (US)	✓	✓		✓
27	Provide clear and meaningful links on the site (D) Use meaningful link labels) (US)	✓			✓
28	Provide active internal & external links (external links open in a new page & in related language) (D) Check links **(UN)**		✓	✓	✓

	Criteria	USA Guide	UN Guide	CBP Guide	Dubai Guide
	Each link functioning properly **(CBP)**				
29	Direct old website links to new and updated ones				✓
30	Provide simple and easy to use forms (mark required field). (D) Forms (required field, field labels aligned with field, field accessible by tabs, confirmation page, if errors did users have to reenter info) **(UN)** Distinguish required & optional data entry field (US)	✓	✓		✓
31	Provide proper and easy to understand guidelines for completing the online forms (online instruction)				✓
32	Provide functioning and properly working forms (confirmation screen upon submitting, reference # to **follow up**, date of request completion, print) (D) Online tracking system for forms, petition, etc. (UN)		✓		✓
33	Provide a print facility on the website pages (D) Develop pages that will print properly (US)	✓			✓
34	Provide well web page titles (short, clear) (D) Provide descriptive page titles (US)	✓			✓
35	Government information in "About Us": vision, **mission**, objectives, contact info (D) Provide customer focused homepage (first good impression for users, **purpose** of the site) (D) **Mission** statements (UN)		✓		✓
36	Provide contact information in "Contact Us": (physical address, location maps, entity & branches service centers phone number, fax, e-mail, customer service e-mail for user inquires, opening hours). (D) Contact info + info about location of offices. (UN) Contact info: phone of mayor or town supervisor, municipal clerk, council members and various departments, fax, e-mails, physical address **(CBP)**		✓	✓	✓
37	Provide a simple feedback form for interaction (D) Provide a functional link to eSuggest (D) Allow comments or feedback (UN)		✓		✓
38	Effective FAQ with facility to ask questions (D) FAQ (U) Increase website credibility: (FAQ) (US)	✓	✓		✓
39	Provide sufficient information about e-services: (service name, description, procedures, forms, time, service centers location)				✓
40	Provide accurate copyright information				✓
41	Provide a proper "Site Maintained By" message				✓
42	Provide a functional link to eJob (D) Posting job vacancies **(UN)**		✓		✓
43	Provide a functional link to eComplain (D) Complaints **(UN)** Explain complain procedures & the time to resolve them (D)		✓		✓
44	Use meaningful metadata on every web page				✓
45	Provide last updated date on the homepage (D) Update of website **(UN)**		✓	✓	✓

	Criteria	USA Guide	UN Guide	CBP Guide	Dubai Guide
	Date of most recent website update **(CBP)**				
46	Provide correct grammar and spellings content (same translation for bilingual websites)				✓
47	Provide appropriate online advertisements				✓
48	Provide information on the protection and handling of privacy (D) A privacy or security statement/policy **(UN)**		✓		✓
49	Provide information on the site terms and conditions include usage of content and registration (D)				✓
50	Provide a consistent format (same look & feel, font, color, navigation menu, link style, reasonable image size, contrast sharply with a plain background, no flash introductory page). (D) Consistent navigation bar, color, font, etc. **(UN)** Font color **(UN)** Place important items consistently (US). Unsure visual consistency (US) Use black text on plain, high-contrast background (US) Consistent design of all pages (header, footers, navigation bar, body of each page, font, color and background design) **(CBP)**	✓	✓	✓	✓
51	customer service e-mail for user inquires (D) Provide a functional link to Ask Dubai (D) Request information **(UN)** Responsiveness of town clerk: email city clerk to ask "What are the hours that the municipal offices are open?" number of days/hours to receive a response. **(CBP)**		✓	✓	✓
52	Perform online satisfaction survey (D) Citizen satisfaction survey **(UN)**		✓		✓
53	Minutes of public **(UN)** Minutes: date, time, location and agenda of next municipal meeting, last meeting, archived and downloadable minutes **(CBP)**		✓	✓	
54	City code and regulations		✓		
55	City charter and policy priority		✓		
56	Documents, reports or books		✓		
57	GIS capabilities		✓		
58	Emergency management or alert mechanism		✓		
59	Human resources information		✓		
60	Calendar of events		✓		
61	Pay utilities, taxes, fines		✓		
62	Apply for permits		✓		
63	Apply for licenses		✓		
64	E-procurement		✓		
65	Property assessment		✓		
66	Customize the city page		✓		
67	Webmaster response		✓		
68	Report violations of administrative laws and regulations		✓		
69	Newsletter **(UN)** Create a positive first impression of your site (up-to-date news, present key topics area) (US)	✓	✓		
70	Online discussion forums		✓		

Appendix D: Double-Expert Review Form of E-City Usability Framework

Component	Guidelines	Expert's Comments			
		Type of Guideline		Acceptable phrasing/ Rephrase	Does Guideline fit into component?
		Objective	Subjective		
Content	1. Useful and most up to date content				
	2. Contact information (phone, fax, e-mails, physical address, location maps, link to customer service email, working hours)				
	3. About us: vision, mission, objectives of a website				
	4. Budget information				
	5. Minutes: municipal meetings (date, location, agenda)				
	6. FAQ with facility to ask new questions				
	7. eJob				
	8. News				
	9. Last update date on the footer of every page				
	10. Simple forms with required fields and proper feedback				
Interaction	11. Ask municipality and response time				
	12. Feedback, comment or eSuggest on the website				
	13. eComplain and time to resolve it				
	14. Citizen satisfaction survey				
Access	15. Bilingual equivalent websites with a link on header of page				
	16. Meaningful images and video that don't slow downloads				
	17. Quick access to the site from Google and Yahoo(top 10 search results)				
	18. Links to national portal and related government websites				
	19. Downloadable documents/ forms with appropriate access				
	20. Design for common browsers				
	21. Print properly pages				
Structure	22. Customer focus homepage: • Quick access to highlighted services • Targeted audience group (citizens, business)				
	23. Clear entity "Logo" on every page				
	24. Link to homepage from				

Component	Guidelines	Expert's Comments			
		Type of Guideline		Acceptable phrasing/ Rephrase	Does Guideline fit into component?
		Objective	Subjective		
	every page through "Home"/logo				
	25. Effective sitemap				
	26. Effective search on the Header				
	27.Short and descriptive page titles				
	28. readable pages (font, color, background)				
	29. Simple page with reasonable length (2 or less screen)				
	30. Privacy and security statement/policy				
Links	31. Logically organized and short meaningful link labels				
	32. All links working properly (no broken links)				
	33. Navigational options: • Indicator on user location in the site • enabled "Back button"				
	34. Different colors for visited/ unvisited links, underline links and no misleading cues to click				
Consistency	35. Consistent design of all pages (same feel and look, font, color, navigation bar, link style)				
E-services Quality	36. Sufficient information on e-services: name, description, procedures, instructions and services centers location				
	37. Ease of navigation through the e-service process				
	38. Clear error message for invalid or incomplete data entry before form submission				
	39. Online tracking for forms and e-services being processed				
	40. Feedback when users waiting or submitting a request				
	41. No physical visit to government offices				
	42. For e-payments: • Availability of various e-payment methods (VISA) • Notification of e-payment via SMS or email				

*Objective Guidelines: Guidelines that could be answered with yes or no as available or not available
*Subjective Guidelines: Guidelines that need people's perceptions, opinions and judgments

Expert's suggestions:

Appendix E: Publications

E 1: Published Paper 1

International Journal of Public Information Systems, Vol 10, No 1, 2014

URL: http://www.ijpis.net/ojs/index.php/IJPIS/article/view/126

A COMPARATIVE APPROACH TO WEB EVALUATION AND WEBSITE EVALUATION METHODS

Dalal Ibrahem Zahran
Dept. of Computer Science
King Abdulaziz University
Jeddah, Saudi Arabia
dzahran@kau.edu.sa

Hana Abdullah Al-Nuaim
Dept. of Computer Science
King Abdulaziz University
Jeddah, Saudi Arabia
hnuaim@kau.edu.sa

Malcolm John Rutter
School of Computing
Edinburgh Napier University
Scotland, UK
M.Rutter@napier.ac.uk

David Benyon
School of Computing
Edinburgh Napier University
Scotland, UK
D.Benyon@napier.ac.uk

Abstract

There is still a lack of an engineering approach for building Web systems, and the field of measuring the Web is not yet mature. In particular, there is an uncertainty in the selection of evaluation methods, and there are risks of standardizing inadequate evaluation practices. It is important to know whether we are evaluating the Web or specific website(s). We need a new categorization system, a different focus on evaluation methods, and an in-depth analysis that reveals the strengths and weaknesses of each method. As a contribution to the field of Web evaluation, this study proposes a novel approach to view and select evaluation methods based on the purpose and platforms of the evaluation. It has been shown that the choice of the appropriate evaluation method(s) depends greatly on the purpose of the evaluation.

Keywords: Web Evaluation Methods; Website Evaluation Methods; Web Engineering; Usability Evaluation Methods.

1. Introduction

Web development is a complex and challenging process that must deal with a large number of heterogeneous interacting components (Murugesan, 2008). Although the construction of Web applications has evolved some discipline, there is still a lack of an engineering approach for building Web systems, and the entire development process is still un-engineered (Ahmad et al., 2005).

An ad-hoc development approach to building complex Web systems quickly leads to poorly designed websites that may cause disasters to many organizations (Ahmad et al., 2005). Nielsen (2011) discovered that the same Web design mistakes occurred over and over again, leading him to publish a series of top-ten Web design mistakes based on testing widely used websites. Progressively, "Web Engineering" is emerging as a new discipline addressing the unique needs and challenges of Web systems and is officially defined as: "The application of systematic, disciplined and quantifiable approaches to development, operation, and maintenance of Web-based Information Systems" (Deshpande et al., 2002). The main topics of Web engineering include, but are not limited to, the following areas: Web development methodologies and models, Web system testing and validation, quality assessment, Web metrics and Web quality attributes disciplines, performance specification and evaluation, Web usability, and user-centric development (Kumar and Sangwan, 2011; Murugesan, 2008).

Unfortunately, evaluation of websites is too often neglected by many organizations, public or commercial, and many developers test systems only after they fail or after serious complications have occurred. Although testing a complex Web system is difficult and may be expensive, it shouldn't be delayed until the end of the development process or performed only after users report problems. The development of a Web system is not a one-off event; it's rather a user-centered continuous process with an iterative life cycle of analysis, design, implementation, and testing (Murugesan, 2008). In this context, testing plays an important

role in Web development, and therefore several methods have been proposed by scholars for evaluating websites. Yet, research that assesses evaluation methods has been in crisis for over a decade, with few publications and risks that inadequate evaluation practices are becoming standardized (Woolrych et al., 2011). In fact, the notion of website evaluation is often confused with Web evaluation in the literature. It is important to know the scope and purpose of evaluation: Are we evaluating the Web or specific website(s)? Also, is the goal to redesign the website, for example, or to obtain Web-ranking and traffic statistics? We need a different focus on evaluation methods and a new categorization system according to the purpose and platforms of evaluation.

Therefore, and to fill a gap in the literature of Web evaluation methods, the following are the objectives of this paper: (1) to distinguish between Web and website evaluation methods; (2) to identify the strengths and weaknesses of the respective approaches; and (3) to recommend the appropriate evaluation method(s) for assessing the Web/website based on the purpose of the evaluation.

2. Related Work

2.1. Web Metrics

Palmer (2002) focused on the need of metrics and emphasized that metrics help organizations generate more effective and successful websites. A survey by Hong (2007) on Korean organizations found that a key enabler of website success measurement is website metrics. These metrics play two important roles: They determine if a website performs to the expectations of the users and the business running the site, and they identify website design problems.

An earlier attempt to measure the Web was introduced in 1996 by Bray, who tried to answer questions such as the size of the Web, its connectivity, and the visibility of sites (Dhyani et al., 2002). Stolz et al. (2005) introduced a new metric assessing the success of information-driven websites that merged user behavior, site content, and structure while utilizing user feedback.

Calero et al. (2005) studied published Web metrics from 1992 to 2004. Using a three-dimensional Web quality model (WQM), they classified 385 Web metrics. The WQM defines a cube structure in which three aspects are considered when testing a website: Web features, life-cycle processes, and quality aspects. The results confirm that most metrics (48% of the metrics studied) are usability metrics, and 44% of them related to "presentation". In this respect, usability is a quality attribute that assesses how easy user interfaces are to use and also refers to methods for improving ease-of-use during the design process (Nielsen, 2012b). In the life cycle dimension, the majority of metrics are related to operation (43.2%) and maintenance processes (30%) (Figure 1). In addition, a large number of metrics are automated (67%).

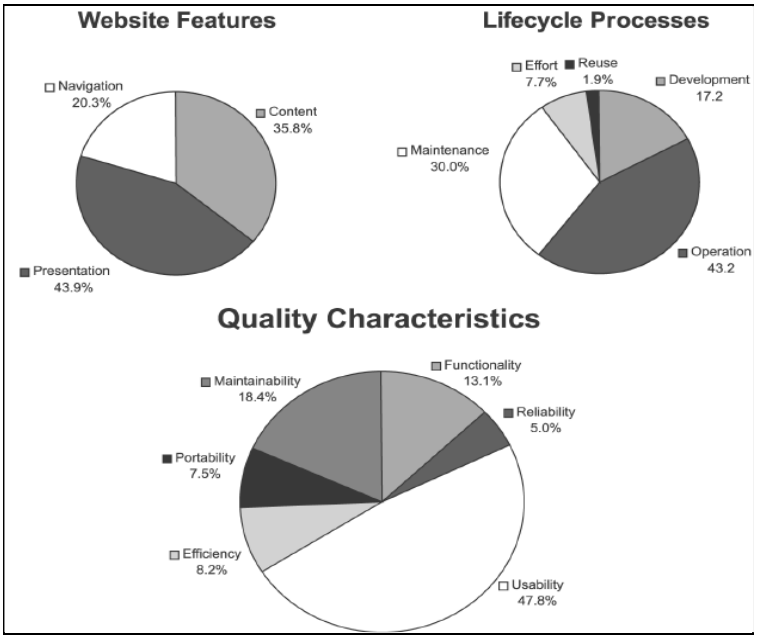

Figure 1. Metric Distribution across the Model Dimensions (Calero et al., 2005)

Dominic and Jati (2010) evaluated the quality of Malaysian University websites based on 11 quality criteria, such as load time, frequency of update, accessibility errors, and broken links, using the following Web diagnostic tools: Websiteoptimization (online performance and speed analyzer), Checklink validator, HTML validator, link popularity tool, and accessibility testing software. From the viewpoint of Treiblmaier and Pinterits (2010), there are two basic criteria for describing websites: "What is presented?" (Content) and "How is it presented?" (Design). The dimension "Ease of Use" contains navigation/organization and usability, the "Usefulness" dimension includes information or site content quality, while the third dimension is "Enjoyment" (Figure 2).

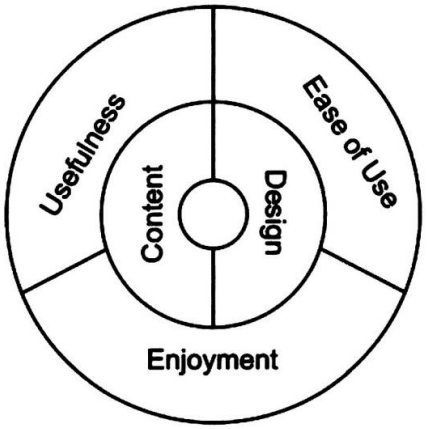

Figure 2. Framework for Web Metrics (Treiblmaier and Pinterits, 2010)

2.2. Trends and Existing Evaluation Approaches

Reviewing previous studies on existing evaluation methods reveals the following problems:

a) Researchers in the field use the terms "Web evaluation methods" (WEMs) and "website evaluation methods" (WSEMs) interchangeably. That is, they do not differentiate between diverse platforms of assessment methods; neither do they consider the purpose of the evaluation. For example, some studies evaluate the Web as a whole phenomenon for the purpose of site ranking or the connectivity and visibility of sites, such as Dhyani et al. (2002) and Stolz et al. (2005). Others assess specific websites against certain attributes aiming to discover the usability problems of the site, such as the studies of Calero et al. (2005), Dominic and Jati (2010) and Treiblmaier and Pinterits (2010).

b) Researchers in the field seldom classify evaluation methods. Nielsen and Mack (1994) classified usability evaluation methods (UEMs) into four categories: automatic (software evaluation), empirical (user testing), formal (evaluation models), and informal (expert evaluation), and later Ivory and Hearst (2001) categorized them into five categories: testing, inspection, inquiry, analytical modeling, and simulation. Recent attempts by Fernandez et al. (2011) adopted the same taxonomy as Ivory and Hearst. Unfortunately, those classifications of evaluation methods are few, old, and missing newer approaches, as neither of these taxonomies reflects, for example, Web analytics or link analysis aspects of UEMs.

c) Researchers in the field often applied the method(s) on different websites but seldom analyzed them or identified their strengths and weaknesses. For instance, link analysis methods have been used widely, but very few authors, such as Jalal et al. (2010), Noruzi (2006), and Shekofteh et al. (2010), evaluate them. Also, Fernandez et al. (2011) and Hasan (2009) indicated that there is little detail about the benefits and drawbacks of each method. Woolrych et al. (2011) warned that research that assesses UEMs has been in crisis for over a decade because of fewer publications. There are also risks that inadequate evaluation practices are becoming prematurely standardized.

d) Few compare evaluation methods or look at a combination of them. Summarizing the knowledge on UEMs over the last 14 years (1996 till 2009), Fernandez, et al. (2011) confirmed that studies often compare a limited number of evaluation methods. Also, Woolrych et al. (2011) argue that very few comparative studies investigate evaluation methods. Reviewing studies from 1995 till 2006, Chiou et al. (2010) stated that there was very limited research exploring the strategies of website evaluation.

A sample of studies using or comparing evaluation methods (explained in the next section) is presented in Table 1. Most of the research uses one or a few techniques only, and the literature is lacking the identification and classification of WEMs. It is worth noting that user testing and heuristics evaluation are traditional methods defined earlier by Nielsen (1993), whereas webometrics is a relatively new and evolving approach.

Table 1. Web Evaluation Methods

Authors	User Testing	Heuristics Evaluation	Automatic Tools	Analytics Tools	Google Analytics	Alexa	PageRank	Webometrics
Brajnik(2004a; 2004b; 2008); Ivory & Chevalier (2002); Dingli & Mifsud (2011); Dominic et al. (2010); Berntzen & Olsen (2009); Olsen et al. (2009); Ataloglou & Economides (2009)			√					
Palmer (2002)				√				
Hasan et al. (2009)					√			
Cho & Adams (2005)							√	
Noruzi (2005; 2006); Björneborn (2004); Jeyshankar & Babu (2009); Holmberg & Thelwall (2009); Li (2003); Thelwall & Zuccala (2008); Boell et al. (2008); Petricek et al. (2006); Shekofteh et al. (2010); Aminpour et al. (2009)								√
Nielsen (1993); Stone et al. (2005); Folmer & Bosch (2004); Lárusdóttir (2009)	√	√						
Prom (2007)				√	√			
Fang (2007)				√	√		√	
Scowen (2007)			√			√	√	
Matera et al. (2006)	√	√	√	√				
Hasan (2009)	√	√	√	√	√			

3. Classification of Evaluation Methods

The development of a Web system is a continuous process with an iterative life cycle of analysis, design, implementation, and testing (Murugesan, 2008). In the process of analyzing websites, Stolz et al. (2005) distinguished between three basic measurements: Web structure measurement (organization and navigability/links), Web content measurement, and Web usage measurement (as page view, sessions, frequency, unique users, and duration). Another view by Hasan (2009) categorized the assessment pattern into user, evaluator, and tool-based UEMs. But what we need really is a different focus on evaluation methods and a new categorization system according to the purpose and platforms of evaluation. Therefore, we propose a distinction between Web and website evaluation methods. We also stress the need for a more systematic identification of those methods.

Based on the previous discussion of classifying the assessment approaches to Web or website evaluation methods and extending Stolz et al. and Hasan's work, the following taxonomy of evaluation method is proposed:

1. Website evaluation methods (WSEMs):
    A. User-based usability evaluation methods
    B. Evaluator-based usability evaluation methods

C. Automatic website evaluation tools (Bobby, LIFT, etc.)

2. Web evaluation methods (WEMs):
   A. Web analytics tools: (Google analytics, Alexa)
   B. Link analysis methods:
      i. PageRank
      ii. Webometrics methods.

3.1. Website Evaluation Methods (WSEMs)

The WSEMs measure a limited number of websites, manually or automatically, based on assigned criteria to achieve a high-quality website. Manual evaluation includes experts or real user testing, while automatic assessments employ different software-testing tools. The output of such an evaluation is a list of usability problems and recommendations to improve the tested website.

3.1.1. User-based Usability Evaluation Methods

The whole process of design for usability, user testing, and redesign is called User-centered Design (Folmer and Bosch, 2004; Nielsen, 1993). The term "usability evaluation" is used to describe the entire test, including planning and conducting the evaluation and presenting the results. The goal of a usability evaluation is to measure the usability of the system and identify usability problems that can lead to user confusion, errors, or dissatisfaction (Lárusdóttir, 2009). The user evaluation approach includes a set of methods that employs representative users to execute some tasks on a selected system. The users' performance and satisfaction with the interface are then recorded. The most common, valuable, and useful method in this category is user testing. Suggested techniques during a user-testing session include the think-aloud method, field observation, questionnaires, and interviews (Hasan, 2009):

User Testing

According to Stone et al. (2005), when users use a system, they work towards accomplishing specific goals in their minds. A goal is an abstract end result indicating what is to be achieved, and it can be attained in numerous ways. Consequently, each goal breaks down into tasks specifying what a person has to do, and then each task decomposes into an individual step that needs to be undertaken. In fact, user testing must be a sampling process, and users should be able to do basic tasks correctly and quickly. To select tested tasks, the examiner begins by exploring all the tasks within the website then narrowing them down to those that are the most important to users. A good task is one that discovers a usability problem or one that reveals an error that is difficult to recover from. The next step is how to present selected tasks to the participants, and one way to do this is to use a “scenario” in which the task is embedded in a realistic story. A good scenario is short, in the users' words, and directly linked to the user's everyday tasks and concerns. It does not give the steps for doing the task, since the point of the test is to see if a user can figure out the required steps alone.

It is important to test users individually and let them solve problems on their own. Actually, the purpose of a usability study is to test the system and not the users, and this aspect must be explicitly explained to tested users (Nielsen, 1993; Stone et al., 2005). The

following metrics can be collected from user testing: time for users to learn a specific function, speed of task performance, type and rate of users' errors, user retention of commands over time, and user satisfaction (Abras et al., 2004). Moreover, how many participants to include in a user testing is a major issue in the usability field. Usually, three to five participants are needed to see all the potential usability problems (Nielsen, 1993; Stone et al., 2005). Nielsen confirmed that the best results come from the first five users and that roughly 85% of the usability problems in a product are detected with five participants.

The Think-aloud Method

Lárusdóttir (2009) and Nielsen (1993) regard thinking aloud as the single most valuable usability evaluation method, and Nielsen (2012a) still holds the same opinion, as he titled his article, "Thinking Aloud: The #1 Usability Tool." Basically, this method involves an end user using the system while thinking out loud. By verbalizing their thoughts, the test users enable us to understand how they view or interpret the system and what parts of the dialogue cause problems. Its strength lies in the wealth of collected qualitative data that can be obtained from a small number of users. The users' comments can be included in the test report to make it more informative. However, to some extent, thinking aloud seems an unnatural setting for users, and sometimes it may give a false impression of the actual cause of usability problems if too much weight is given to the users' justifications (Nielsen, 1993).

3.1.2. Evaluator-based Usability Evaluation Methods

Evaluators or experts inspect the interface and assess system usability using interface guidelines, design standards, users' tasks, or their own knowledge, depending on the method, to find possible user problems (Lárusdóttir, 2009). The inspectors can be usability specialists or designers and engineers with special expertise (Matera et al., 2006). In this category, there are many inspection methods, such as cognitive walkthrough, guideline reviews, standard inspection, and heuristic evaluation (Hasan, 2009).

Heuristic Evaluation

Heuristic evaluation is a very efficient usability engineering method, and it is especially valuable when time and resources are scarce. A number of evaluators assess the application and judge whether it conforms to a list of usability principles, namely "heuristics" (Hasan, 2009). There are two sets of guidelines that are widely used in heuristic evaluation, Nielsen's (1993) heuristics being the most common, followed by Gerhardt-Powals' (1996) (Lárusdóttir, 2009). Nielsen's heuristics are part of the so-called "discount usability methods" which are easy, fast, and inexpensive. During the heuristic evaluation, each evaluator goes individually through the system interface at least twice, and the output of such evaluation is a list of usability problems with reference to the violated heuristics (Matera et al., 2006). In principle, heuristic evaluation can be conducted by only one evaluator, who can find 35% of total usability problems (Nielsen, 1993), but another view by Matera et al. (2006) believes that better results are obtained by having five evaluators and certainly not fewer than three for reasonable results.

3.1.3. Automatic Website Evaluation Tools

Automatic evaluation tools are software that automates the collection of interface usage data and identify potential Web problems. The first study of automatic tools was conducted by Ivory and Chevalier (2002), who concluded that more research was needed to validate the embedded guidelines and to make the tools usable. Thus Web professionals cannot rely on them alone to improve websites. Brajnik (2004b) mentioned several kinds of Web-testing tools: accessibility tools such as Bobby, usability tools such as LIFT, performance tools such as TOPAZ, security tools such as WebCPO, and classifying website tools such as WebTango. He stated that the adoption of tools is still limited due to the absence of established methods for comparing them and also suggested that the effectiveness of automatic tools has to be itself evaluated (2004a). In fact there are many automated tools available as either Web-based services or desktop applications. A recent popular free Web-based accessibility tool is Cynthia Says (http://www.cynthiasays.com/) which is a product from HiSoftware that allows you to enter the URL to be analyzed in to the sight and get a report on how it complies with Section 508 standards and/or the Web Content Accessibility Guidelines (WCAG). Table 2 shows some studies that use different kinds of automatic website evaluation tools.

Table 2. Examples of Automated Web Site Evaluation Studies

Name of the Study	Author / Year	Automatic tools
Assessing e-governance Maturity through Municipal Websites: Measurement Framework and Survey	(Rodríguez et al., 2009)	1-W3C validators 2-Xenu s\w (broken links) 3-Weight & image resolution 4-Source code analyzer
Quantitative Assessment of European Municipal Web Sites Development and Use of an Evaluation Toll	(Miranda, Sanguino, & Banegil 2009)	1-Google search engine 2-Link popularity check 3-Chronmeter (access speed)
Local E-government: Reconstructing Limassol's Municipality (Cyprus) Web Site to Provide Functional and Effective E-services	(Zevedeos, 2006)	1-WebXact (Bobby) 2-Lynx (accessibility) 3-Vischeck (color) 3-W3C Markup Validator 4-W3C CSS validator 5-W3C Link Checker
Performance Evaluation on Quality of Asian E-government Websites – an AHP Approach	(Dominic et al., 2010)	1-Website optimization (website performance and speed analyzer) 2-W3C checklink 3-Link popularity 4-Accessibility s\w Tawdis tester 5-Color-blind webpage filter
Evaluating Global E-government Sites: A View Using Web Diagnostic Tools	(Choudrie, Ghinea, & Weerakkody, 2004)	1-WebXact (accessibility, quality & privacy) 2-Netmechanic (Links) 3-W3C HTML validator 4-Vizcheck (color)

3.2. Web Evaluation Methods (WEMs)

The WEMs study the Web as a whole by calculating statistics about the detailed use of a site and providing Web-traffic data, visibility, connectivity, ranking, and the overall impact of a site on the Web.

3.2.1. Web Analytics Tools

Web analytics have been defined by the Web Analytics Association as "the measurement, collection, analysis and reporting of Internet data for the purpose of understanding and optimizing Web usage" (Fang, 2007). These tools automatically calculate statistics about the detailed use of a site helping, for example, in discovering navigation patterns corresponding to high Web usage or to the early leaving of a website (Matera et al., 2006). Originally, Web analytics is a business tool that started with some webmasters inserting counters on their home pages to monitor Web traffic. While most Web analytics studies target e-commerce, the method can be applied to any website (Prom, 2007). The two data collection methods for Web analytics are server-based log files (traffic data is collected in log files by Web servers) and client-based page-tagging (requiring the addition of JavaScript codes to webpages to capture information about visitors' sessions) (Hasan, 2009). The two well-known Web analytics tools are Google Analytics and Alexa.

Google Analytics

Google purchased a Web analytics company called Urchin software in 2005 and subsequently released Google Analytics to the public in 2006 (Fang, 2007; Hasan et al., 2009). The service is free for up to five million page views per month per account. Once signed up for Google Analytics, Google offers users code that must be inserted into each Web page to be tracked. Visual data results are displayed with a wealth of information as to where visitors came from, what pages they visited, how long they stayed on each page, how deep into the site they navigated, etc. (Fang, 2007).

Alexa

Alexa is a website metrics system owned by the Amazon Company that provides a downloadable toolbar for Internet Explorer users. It calculates traffic rank by analyzing the Web usage of Alexa toolbar users for three months or more as a combined measure of page views and reach (the number of visitors to the site). Although this information is useful, Alexa ranking is biased towards MS Windows and Internet Explorer users (Scowen, 2007).

3.2.2. Link Analysis Methods

Link analysis studies websites' topology, assuming that the quality of a Web page is dependent on its links. There are two important methods that use link analysis: PageRank and webometrics.

PageRank

A number of researchers investigated the Web link structure to improve search results and proposed ranking metrics. When Page and Brin designed the Google search engine, they considered links as positive referrals and created a system called PageRank. Google PageRank is a link analysis algorithm named after Larry Page that assigns a

numerical weight to each hyperlink, and each page has a calculated PageRank based on the number and quality of links pointing to it (Scowen, 2007). Google takes 100 factors into consideration when determining the ranking of a page, but PageRank is the main factor in search-result ordering. The PageRank metric PR(p) defines the importance of page p to be the sum of the importance of the pages that point to p, and the PR(p) is high if many important pages point to p. The effectiveness of Google's search results and the adoption of PageRank by other search engines strongly indicate that it is an effective ranking metric for Web searches, but unfortunately it is heavily negatively biased against unpopular pages, especially those created recently (Cho and Adams, 2005).

Scowen (2007) tested e-learning websites against checklist guidelines then against five ranking systems: Google links search, Yahoo links, Delicious links, Google PageRank, and Alexa. The Google PageRank and Alexa were used to know their correlations with the usability of the website, although neither can be relied upon as a main indicator of popularity. He found that increased compliance with usability guidelines has a strong correlation with increased popularity of a website. Although Alexa is not a reliable indicator, it is at least consistent with other rankings. Thus, more usable websites achieve a higher PageRank and are also more popular in Alexa. Overall, the five ranking systems showed positive correlations to each other and to the usability of the sites.

Webometrics and the WIF Method

Björneborn (2004) has proposed webometrics as "the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web, drawing on bibliometric and infometric approaches." This means evaluation of websites can be conducted "webometrically" with the goal to validate links and furnish its acceptance as a useful metric to measure the Web. Webometrics assess the international visibility and impact of an institution or a country on the Web (Jeyshankar and Babu, 2009), but it is still a nascent field of research (Björneborn, 2004; Holmberg and Thelwall, 2009).

The Web Impact Factor (WIF) is the most important method in webometrics. In 1998, Peter Ingwersen proposed WIF through an analogy with the Journal Impact Factor (JIF) (Noruzi, 2005; Li, 2003) that represents the ratio of all citations to a journal to the total references published over a period of time (Dhyani et al., 2002). Since it is a snapshot of the Web and lacks peer review and quality control, the WIF is not exactly the equivalent of the JIF, but it was inspired by it (Thelwall and Zuccala, 2008). In this method, external inlinks are of more value and importance (Aminpour et al., 2009); the more people link to a website, the more WIF the site is getting and, in turn, the higher the impact factor, the higher the reputation and influence of a site (Jeyshankar and Babu, 2009; Shekofteh et al., 2010). Sometimes the WIF is wrongly compared to PageRank method. PageRank does not afford equal weight to links, and weightings vary depending on from where a link is coming (Boell et al., 2008).

Most of webometrics studies were performed on university sites such as the Cybermetrics Lab (2010), which has issued the “Ranking Web of World Universities” since 2004. A study by Thelwall and Zuccala (2008) measured the international interlinking to and from different European universities. Figure 3 shows European links from university

networks with the width of arrows proportional to the number of pages between universities. Results show the dominance of the large, richer western European nations, especially the UK and Germany (de) as central actors on the Web and also strongly connected with each other. The importance of Switzerland (ch) is apparent, since it is connected strongly to the UK and Germany, weakly to seven countries, and medium to one country, France (fr). In turn, France is connected strongly to Germany, weakly to nine countries, and medium to four countries: Italy (it), Belgium (be), Switzerland (ch), and the Netherlands (nl). Poland (pl) is also well-connected and has a significant presence as a newcomer.

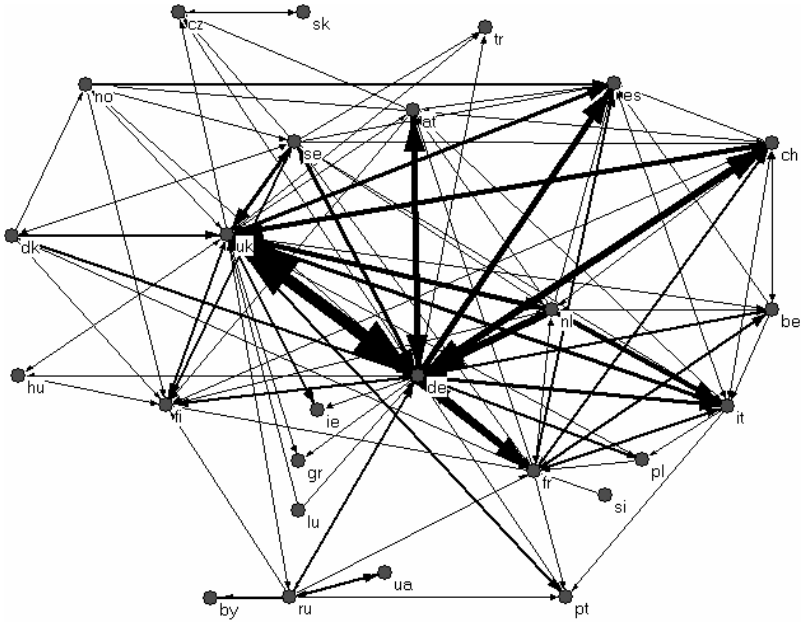

Figure 3. European Link Network. (Thelwall and Zuccala, 2008)

A few webometrics studies have been conducted on e-government, representing a new application of the WIF method. The first attempt to measure e-government webometrically was the study by Petricek et al. (2006), which compared the audit office sites in five countries and showed that the US and Canada emerge as the most connected sites, more than the UK, New Zealand, and Czech Republic.

4. Analysis of Evaluation Methods

This section examines existing evaluation methods individually, regardless of any proposed categorization in order to identify the strengths and weaknesses of each method.

Automatic website evaluation tools attract attention because they are fast, consistent, produce unbiased results, and obviate the shortage of experts and inconsistent results between them (Ataloglou and Economides, 2009; Dingli and Mifsud, 2011; Dominic et al., 2010). Also, these tools can offer an initial overview of the status of a website (Olsen et al., 2009). However, automation of website testing is an evolving method that cannot be considered efficient (Al-Juboori et al., 2011). Berntzen and Olsen (2009),

Brajnik (2008), and Dingli and Mifsud (2011) concluded that automatic tools cannot replace human evaluators but should assist them. Ivory and Chevalier (2002) predicted that automation is a useful complement to standard evaluation techniques. Manual evaluations provide more details than automatic tests, which cannot capture the whole picture. Anything requiring assessment is likely to be poorly machine testable (Brajnik, 2004b).

Another concern is that the market forces can cause changes that threaten automatic tools' stability. For example, Bobby, an accessibility testing tool, was sold in 2004 to Watchfire, which provided the same free service in the WebXACT tool, but Watchfire was acquired by IBM in 2007. Bobby was then discontinued as a free tool, and currently it is included within the IBM Rational Policy Tester Accessibility Edition (Hasan, 2009). In fact, automatic tools are seldom used alone in website evaluation; also, very few studies compare the tools and validate their effectiveness (Al-Juboori et al., 2011; Brajnik, 2004a, 2004b). The most-used tools are Bobby, LIFT, W3C validators, and link-checker software. Most automatic tools focus on site accessibility rather than usability, and they are not considered efficient (Hasan, 2009; Scowen, 2007). Even the very few tools for usability often neglect structural and navigational problems (Matera et al., 2006). Further, information about LIFT is contradictory; some conceive LIFT as a test for accessibility and some as a usability tool. Also, features measured by LIFT are inconsistent with the USA Research Web Design and Usability Guidelines (Scowen, 2007).

On the other hand, Web analytics tools solve some problems in Web evaluation, since they might reduce the need for user testing, and often the data is collected automatically with high accuracy. They offer the possibility of analyzing a high number of visitors, thus increasing the reliability of the discovered errors; however, the inaccuracy of log files as a data source is acknowledged (Hasan, 2009). Another serious problem is the meaning of the collected information and how much it describes users' behavior (Matera et al., 2006). Palmer (2002) believes website traffic measures are used because they are easy to capture but are very often deemed to be inadequate and sometimes may generate conflicting results.

A Web analytics tool such as Alexa has some limitations; it is biased towards a sample of MS Windows and Internet Explorer users. The resulting statistics are unreliable since users of other operating systems or browsers are not recorded, and traffic from other Internet users is not counted (Scowen, 2007). Unfortunately, there are only a few studies that show the value of Google Analytics in assessing websites; Hasan (2009) developed a framework for evaluating three e-commerce sites in the kingdom of Jordan using heuristic evaluation, user testing, and Google Analytics. Jordanian companies took a long time to agree to participate in the research due to trust and security issues, since they were asked to add script code to their servers.

Noruzi (2006) considers the webometric method as an imperfect tool to measure the quality of websites. Questions are raised over the entire quantitative nature of the webometrics rankings (Björneborn, 2004). The tool used in the WIF analysis is not meant for the task, and search engines are designed for content retrieval, not link analysis; plus, they may create problems in drawing conclusions for the WIF since their coverage of the Web is incomplete. The lack of knowing why Web links are created is a major obstacle in

the webometrics method; thus the motivations behind creating links raise questions of uncertainty (Noruzi, 2006). Also, some webometrics' studies found unexpected results and attributed them to the limitations of the WIF method. For example, a university with 993 links and 99 Web pages, by division, gets an impact factor of 10, whereas another one with 12,700 links and 87,700 Web pages obtains an impact factor below zero (Shekofteh et al., 2010).

Based on webometrics evaluation, university rankings have raised a large dispute, and several studies criticize them as merely a list of criteria that mirrors the superficial characteristics of universities. Noruzi (2006) argued that world university website ranking is dangerous and not meaningful because a high link rate may not always be associated with high quality. It is vulnerable to manipulation, since the WIF can be influenced by institutions that know how this method works. Shekofteh et al. (2010) concluded that the WIF alone is not a good measure for ranking universities, and Noruzi (2006) stated that with about 10 years of criticism, it seems that there is no obvious alternative yet. Webometrics is relatively a young field of research that needs different theories to be built, methods to be developed, and problems to be solved (Björneborn, 2004; Holmberg and Thelwall, 2009). Calculating the WIF for a website is easy, but what the figures mean is arguable. Thus, the researches on webometrics are in the process of developing and validating its methodologies.

Matera et al. (2006) supported Nielsen (1993) in considering heuristic evaluation as a very efficient method when time and resources are scarce because experts can produce high-quality results in a limited time. But a negative aspect is its high dependence on skills and the experiences of the evaluators. They concluded that novice evaluators with no usability expertise are poor evaluators, usability experts are 1.8 times as good, while application domain and usability experts (double experts) are 2.7 as good. Another weakness of this method is the great subjectivity of the evaluation; there is a risk that the experts mistakenly consider some issues as problems but actually real users do not have trouble with them; this is often referred to as "false problems" (Lárusdóttir, 2009).

According to Nielsen (1993), user testing with the think-aloud technique finds more major Web problems than other evaluation methods but is poor in uncovering minor ones, and the situation is the opposite for the heuristic evaluation. Since they complement each other, he recommends first conducting a heuristic evaluation to find as many "obvious" usability problems then performing user testing to find the remaining problems. Likewise, Hasan (2009) reached the same conclusion of Nielsen and added that Google Analytics is a useful quick preliminary step to discover general usability problems. She found that user testing is good for identifying major usability problems in four areas: navigation, design, the purchasing process, and accessibility and customer service, while the heuristic evaluation identifies minor usability problems in eight areas: navigation, internal search, the site architecture, the content, the design, accessibility and customer service, inconsistency and missing capabilities, plus addressing security and privacy issues. Other Web experts recommend using several different evaluation techniques, since each one alone is not free of shortcomings (Ivory and Chevalier, 2002).

The overall recommendation by many researchers is to conduct heuristic evaluation and user testing to find most usability problems. Other evaluation methods are just useful complements offering the possibility of analyzing a high number of users as an initial preview of a website. Consequently, evaluations by experts or users are the mainstream approach, and probably the future trend is a mixture of automatic and manual website evaluations.

5. Selection of Appropriate Evaluation Method(s)

Kaur and Dani (2013) evaluated the state of navigability of Indian banking websites and found that Alexa and Google PageRank do not have significant correlations with navigability metrics, indicating that popularity and importance are not good indicators of website navigability; therefore, the traffic data and the back-links of the websites are not meaningful measures of site navigation assessment. Cho and Adams (2005) added that PageRank is not a metric of page quality. Further, Hong (2007) stated that most organizations use Web metrics to determine site traffic or popular content but seldom used them to improve navigation. Jalal et al. (2010) and Noruzi (2006) concluded that the webometric method is an imperfect tool to measure the quality of websites and that it reflects unreliable results in most cases.

The findings of these five studies support the argument that WEMs, such as the Web analytics tools and the link analysis methods, do not discover navigation problems accurately nor do they measure website quality. Further, it seems that WEMs are complementary approaches since they do not definitely discover usability problems of a site, rather they indicate their probability.

On the other hand, even though usability testing demonstrates how real users interact with a website and the exact problems they face, it cannot measure the success of a site or describe the interactions of large numbers of users with it (Hasan, 2009). This highlights the weakness that WSEMs, such as user, evaluator, or automatic evaluation methods, cannot provide traffic data, Web ranking of a site, or its online visibility among others.

Therefore, the choice of the appropriate evaluation method depends greatly on the purpose of the evaluation. If it is intended to redesign the website and wanted to discover most of its potential usability problems, then the best evaluation methods are user testing and expert evaluation, while an automatic tool or Google analytics is a useful complement in this situation. If the goal of the evaluation is to redesign a website then WSEM is the best approach, while WEMs are not useful enough in this circumstance. Similarly, if the goal is to clarify the extent of online correlation with other institutions/countries or to know the ranking of a website and how much traffic it attracted, then the best way is to use WEMs, link analysis methods, and Web analytics tools, respectively. Figure 4 shows how the purpose of Web evaluation determines the type of method; the dotted arrow is toward a complementary method.

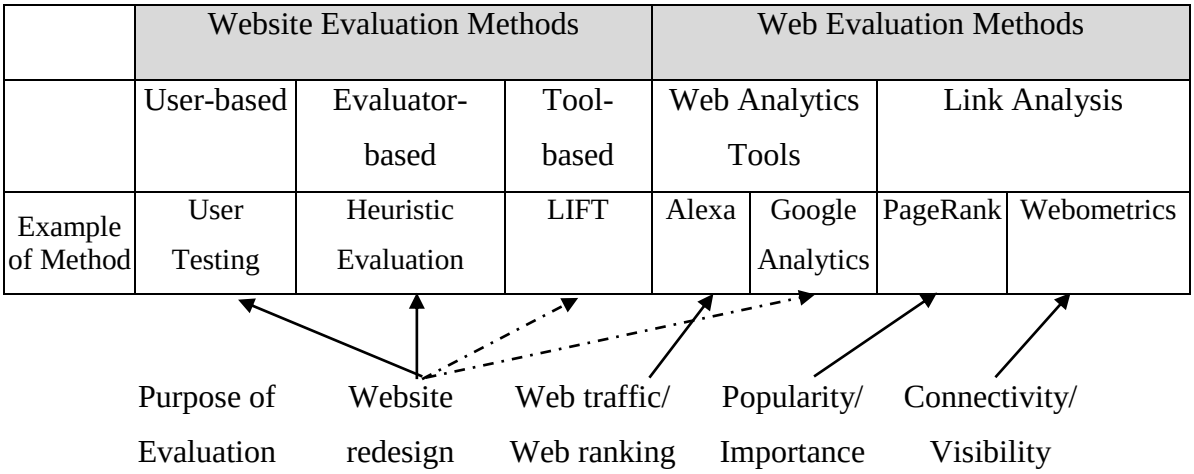

Figure 4. Purpose of Web Evaluation Determines the Appropriate Method Type

6. Conclusion

To address the challenge of developing complex Web systems, "Web Engineering" is an emerging discipline for the implementation of engineering principles to promote high quality websites that attract visitors. How to measure the Web has become a valuable area of ongoing research, but unfortunately the field is not yet mature; Web evaluation methods are scattered over the literature with a lack of studies that classify, compare, and determine the appropriate evaluation method(s).

Previous studies confused the term “Web evaluation methods” with “website evaluation methods,” since they did not distinguish between diverse platforms of assessment methods and also did not address the purposes behind such evaluation. For example, some studies evaluated the Web in terms of ranking and connectivity of sites, while others assessed specific websites to discover their usability problems.

A novel approach to view evaluation methods is proposed, and a new categorization system has been suggested based on the purpose and platforms of evaluation. As a contribution to the field of Web evaluation, we have identified existing evaluation methods and accordingly classified them into two types: (1) website evaluation methods including user-based UEMs such as user testing and think aloud, evaluator-based UEMs such as heuristics evaluation, and automatic website evaluation tools and (2) Web evaluation methods including Web analytics tools (Google analytics, Alexa) and link analysis consisting of PageRank and webometrics methods.

Analyzing existing evaluation methods resulted in the following conclusions: First, standard evaluation techniques are user testing and heuristic evaluation. Second, tool-based evaluation methods offer a first insight into the status of a website. Automatic testing is a useful complementary tool but it is an evolving method with little evidence of its efficacy. Similarly, Web analytics tools provide some useful website traffic measures. However, the resulting statistics of Alexa, for example, are unreliable since it covers a limited number of Internet users. Also, Google Analytics is a quick preliminary step to discover usability

problems, but its uses are limited due to trust and security issues. Third, link analysis methods try to validate links as a useful metric to measure the Web, but actually PageRank and webometrics methods can be regarded as indicators rather than definite conclusions on the visibility and impact of a website. For example, the WIF is partially successful; it does provide some useful information such as the relationship and type of communication between universities/countries and also how a website is isolated or connected with others online. On the other hand, the method is not appropriate for the ranking of websites since it is not a suitable tool for assessing a website's quality.

The purpose of Web evaluation determines the appropriate method(s) to be used. If the purpose is to redesign the website, then the scope of evaluation is WSEM, and therefore, as stated by the literature, the best evaluation methods are user testing and expert evaluation, while automatic and Web analytics tools (complementary) could provide a first insight into the status of the website. Similarly, if Web ranking and traffic statistics are of interest, then the scope of evaluation is WEMs; thus the best way is to use a Web analytics tool such as Alexa.

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-Centered Design. Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, *37*(4), 445-56.

Ahmad, R., Li, Z., & Azam, F. (2005). Web Engineering: A New Emerging Discipline. IEEE 2005 International Conference on Emerging Technologies. September 17-18, Islamabad.

Al-Juboori, A. F. M. A., Na, Y., & Ko, F. (2011). Web Site Evaluation: Trends and Existing Approaches. In Networked Computing (INC), 2011 The 7th International Conference, IEEE, 155-160.

Aminpour, F., Kabiri, P., Otroj, Z., & Keshtkar, A. (2009). Webometric Analysis of Iranian Universities of Medical Sciences. Scientometrics, *80* (1), 253–264.

Ataloglou, M., & Economides, A. (2009). Evaluating European Ministries' Websites. International Journal of Public Information Systems, *5*(3), 147–177.

Berntzen, L., & Olsen, M. (2009). Benchmarking E-government- A Comparative Review of Three International Benchmarking Studies. In Digital Society, 2009. ICDS'09. Third International Conference on IEEE, 77-82.

Björneborn, L. (2004). Small-World Link Structures across an Academic Web Space: A Library and Information Science Approach. Royal School of Library and Information Science, Denmark. Retrieved from http://vip.db.dk/lb/phd/phd-thesis.pdf

Boell, S., Wilson, C., & Cole, F. (2008). A Webometric Analysis of Australian Universities Using Staff and Size Dependent Web Impact Factors (WIF). University of New South Wales, School of Information Systems, Technology and Management (SISTM), Sydney, Australia

Brajnik, G. (2004a). Comparing Accessibility Evaluation Tools: A Method for Tool Effectiveness. Universal Access in the Information Society, *3*(3-4), Springer Verlag, 252-263

Brajnik, G. (2004b). Using Automatic Tools in Accessibility and Usability Assurance. 8th International ERCIM UI4ALL Workshop, June 2004, Vienna, Austria. Springer, Berlin, 219-234

Brajnik, G. (2008). Beyond Conformance: The Role of Accessibility Evaluation Methods. Keynote paper, 2nd International Workshop on Web Usability and Accessibility IWWUA08, September, 2008, Auckland, New Zealand. Retrieved from http://www.cast.org/learningtools/Bobby/index.html

Calero, C., Ruiz, J., & Piattini, M. (2005). Classifying Web Metrics Using the Web Quality Model. Online Information Review, Emerald Group, *29* (3), 227-248

Chiou, W. C., Lin, C. C., & Perng, C. (2010). A Strategic Framework for Website Evaluation based on a Review of the Literature from 1995–2006. Information & Management, 47(5), 282-290.

Cho, J., & Adams, R. (2005). Page Quality: In Search of an Unbiased Web Ranking. In Proceedings of the 2005 ACM SIGMOD conference, Baltimore, Maryland, 551-562

Choudrie, J., Ghinea, G., & Weerakkody, V. (2004). Evaluating Global e-Government Sites: A View using Web Diagnostic Tools. Brunel University, Uxbridge, UK. Electronic Journal of e-Government, 2 (2), 105-114

Cybermetrics Lab. (2010). Webometrics Ranking of World Universities. Retrieved from http://www.webometrics.info/

Deshpande, Y., Murugesan, S., Ginige A., Hansen, S., Schwbe, D., Gaedke, M., & White, B. (2002). Web Engineering. Journal of Web Engineering, Rinton Press, *1*(1), 003-017.

Dhyani, D., Ng, W., & Bhowmick, S. (2002). A Survey of Web Metrics. Nanyang Technological University. ACM Computing Surveys, *34*(4), 469–503.

Dingli, A., & Mifsud, J. (2011). USEFul: A Framework to Mainstream Web Site Usability through Automated Evaluation. International Journal of Human Computer Interaction (IJHCI), *2*(1), 10

Dominic, P., & Jati, H. (2010). Evaluation Method of Malaysian University Website: Quality Website Using Hybrid Method. In Information Technology (ITSim), 2010 International Symposium, IEEE, *1*, 1-6

Dominic, P., Jati, H., & Kannabiran, G. (2010). Performance Evaluation on Quality of Asian E-government Websites - an AHP Approach. International Journal of Business Information Systems, *6*(2), 219-239.

Fang, W. (2007). Using Google Analytics for Improving Library Website Content and Design: A Case Study. Library Philosophy and Practice 2007, 1-17

Fernandez, A., Insfran, E., & Abrahão, S. (2011). Usability Evaluation Methods for the Web: A Systematic Mapping Study. Information and Software Technology, *53*(8), 789-817.

Folmer, E., Bosch, J. (2004). Architecting for Usability: A Survey. Journal of Systems and Software, 2004, *70*(1), 61–78. Retrieved from http://dissertations.ub.rug.nl/FILES/faculties/science/2005/e.folmer/c2.pdf

Hasan, L. (2009). Usability Evaluation Framework for E-commerce Websites in Developing Countries. A Doctoral dissertation. Loughborough University

Hasan, L., Morris, A., & Probets, S. (2009). Using Google Analytics to Evaluate the Usability of E-commerce Sites. In Human Centered Design, Springer Berlin Heidelberg, 697–706

Holmberg, K., & Thelwall, M. (2009). Local Government Web Sites in Finland: A Geographic and Webometric Analysis. Scientometrics, *79*(1), 157–169.

Hong, I. (2007). A Survey of Web Site Success Metrics Used by Internet-dependent Organizations in Korea. Internet Research, Emerald Group, *17*(3), 272-290

Ivory, M. & Chevalier, A. (2002). A Study of Automated Web Site Evaluation Tools. University of Washington, Department of Computer Science2002

Ivory, M. Y., & Hearst, M. A. (2001). The State of the Art in Automating Usability Evaluation of User Interfaces. ACM Computing Surveys (CSUR), *33*(4), 470-516.

Jalal, S., Biswas, S. & Mukhopadhyay, P. (2010). Web Impact Factor and Link Analysis of Selected Indian Universities. Annals of Library and Information Studies, *57*, 109 – 121. Retrieved from http://eprints.rclis.org/16164/1/Annals-57-2.pdf

Jeyshankar, R., & Babu, R. (2009). Websites of Universities in Tamil Nadu: A Webometric Study. Annals of Library and Information Studies, *56*(2), 69-79. Retrieved from http://nopr.niscair.res.in/bitstream/123456789/5939/1/ALIS%2056(2)%2069-79.pdf

Kaur, A., & Dani, D. (2013). The Web Navigability Structure of E-Banking in India. International Journal of Information Technology and Computer Science *(IJITCS)*, 5(5), 29. Retrieved from http://www.mecs-press.org/ijitcs/ijitcs-v5-n5/IJITCS-V5-N5-4.pdf

Kumar, S. Sangwan, S. (2011). Adapting the Software Engineering Process to Web Engineering Process. International Journal of Computing and Business Research, *2*(1)

Lárusdóttir, M. (2009). Listen to Your Users: The Effect of Usability Evaluation on Software Development Practice. A Doctoral Dissertation, Department of Information Technology, UPPSALA University, Sweden

Li, X. (2003). A Review of the Development and Application of the Web Impact Factor. Online Information Review, Emerald Group, *27*(6), 407-417. Retrieved from http://www.emeraldinsight.com/researchregister

Matera, M., Rizzo, F., Carughi, G. (2006). Web Usability: Principles and Evaluation Methods. Department of Electronic Information, Milano, Italy

Miranda, F., Sanguino, R., & Banegil ,T. (2009). Quantitative Assessment of European Municipal Web Sites - Development and Use of an Evaluation Tool. Internet Research, 19(4), 425-441, Emerald Group Publishing Limited

Murugesan, S. (2008). Web Application Development: Challenges and the Role of Web Engineering. University of Western Sydney, Australia.

Nielsen, J. (1993). Usability Engineering. San Francisco, CA: Morgan Kaufmann

Nielsen, J. (2011). Top 10 Mistakes in Web Design. Retrieved from http://www.useit.com/alertbox/9605.html

Nielsen, J. (2012a). Thinking Aloud: The #1 Usability Tool. Retrieved from http://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/

Nielsen, J. (2012b). Usability 101: Introduction to Usability. Retrieved from http://www.nngroup.com/articles/usability-101-introduction-to-usability/

Nielsen, J. & Mack, R. L. (Eds.) (1994). Usability Inspection Methods. John Wiley & Sons, New York.

Noruzi, A. (2005). Web Impact Factors for Iranian Universities. Webology, *2*(1), Retrieved from http://webology.ir/2005/v2n1/a11.html

Noruzi, A. (2006). The Web Impact Factor: A Critical Review. The Electronic Library, *24*(4), 490-500.

Olsen, M., Nietzio, A., Snaprud, M., & Fardal, F. (2009). Benchmarking and Improving the Quality of Norwegian Municipality Web Sites. Automated Specification and Verification of Web Systems, 115

Palmer, J. (2002). Web Site Usability, Design, and Performance Metrics. Information Systems Research, *13*(2), 151-167

Petricek, V., Escher, T., Cox, I., & Margetts, H. (2006). The Web Structure of E-Government - Developing a Methodology for Quantitative Evaluation. In Proceedings of the 15th international conference on World Wide Web, ACM, 669-678

Prom, C. (2007). Understanding On-line Archival Use through Web Analytics. ICA-SUV Seminar, Dundee, Scotland. Retrieved from http://www.library.uiuc.edu/archives/workpap/PromSUV2007.pdf

Rodríguez, R., Estevez, E., Giulianelli, D, & Vera, P. (2009). Assessing e-Governance Maturity through Municipal Websites- Measurement Framework and Survey Results. 6th Workshop on Software Engineering, Argentinean Computer Science Conference, San Salvador de Jujuy, Argentina, Oct 2009.

Scowen,G. (2007). Increased Website Popularity through Compliance with Usability Guidelines. A Doctoral Dissertation. Information Science, University of Otago, New Zealand

Shekofteh, M., Shahbodaghi,A., Sajjadi, S., & Jambarsang, S. (2010). Investigating Web Impact Factors of Type 1, Type 2 and Type 3 Medical Universities in Iran. Journal of Paramedical Sciences (JPS), *1*(3).

Stolz, C., Viermetz, M., Skubacz, M., & Neuneier, R. (2005). Guidance Performance Indicator- Web Metrics for Information Driven Web Sites. IEEE/WIC/ACM International Conference on Web Intelligence 2005, 186–192

Stone, D., Jarrett, C., Wodroffe, M., & Minocha, S. (2005). User Interface Design and Evaluation. San Francisco, CA: Morgan Kaufmann

Thelwall, M., & Zuccala, A. (2008). A University-Centred European Union Link Analysis. Scientometrics, *75*(3), 407–420.

Treiblmaier, H., & Pinterits, A. (2010). Developing Metrics for Web Sites. Journal of Computer Information Systems, *50*(3).

Woolrych, A., Hornbæk, K., Frøkjær, E., & Cockton, G. (2011). Ingredients and Meals rather than Recipes: A Proposal for Research that does not treat Usability Evaluation Methods as Indivisible Wholes. International Journal of Human-Computer Interaction, *27*(10), 940-970.

Zevedeos, R. (2006). Local e-Government: Reconstructing Limassol's municipality (Cyprus) Web Site to Provide Functional and Effective E-services. Master Dissertation. University of Sheffield.

E 2: Published Paper 2

A Critical Analysis of E-government Evaluation Models at National and Local Municipal Levels

Dalal Ibrahem Zahran

Dept. of Computer Science
King Abdulaziz University
Jeddah, Saudi Arabia

dzahran@kau.edu.sa

Hana Abdullah Al-Nuaim

Dept. of Computer Science
King Abdulaziz University
Jeddah, Saudi Arabia

hnuaim@kau.edu.sa

Malcolm John Rutter

School of Computing
Edinburgh Napier University
Scotland, UK

M.Rutter@napier.ac.uk

David Benyon

School of Cmputing

Edinburgh Napier University

Scotland, UK

D.Benyon@napier.ac.uk

Abstract

The importance of e-government models lies in their offering a basis to measure and guide e-government. There is still no agreement on how to assess a government online. Most of the e-government models are not based on research, nor are they validated. In most countries, e-government has not reached higher stages of growth. Several scholars have shown a confusing picture of e-government. What is lacking is an in-depth analysis of e-government models. Responding to the need for such an analysis, this study identifies the strengths and weaknesses of major national and local e-government evaluation models. The common limitations of most models are focusing on the government and not the citizen, missing qualitative measures, constructing the e-equivalent of a bureaucratic administration, and defining general criteria without sufficient validations. In addition, this study has found that the metrics defined for national e-government are not suitable for municipalities, and most of the existing studies have focused on national e-governments even though local ones are closer to citizens. There is a need for developing a good theoretical model for both national and local municipal e-government.

Keywords: E-government, Municipality, E-government Evaluation Models, Web Evaluation, Usability, Citizen-centric Websites.

1. Introduction

Following the success of e-commerce in the late 1990s, a new face of government known as e-government was introduced (Coursey & Norris, 2008). The European Union (EU) defined e-government as the use of information and communication technologies (ICT) in public administrations to improve public services and democratic processes (Moraru, 2010). In development, e-government has lagged behind e-commerce; a survey conducted in the UK said that, while 85% of Internet users have searched for or bought goods and services online, and 50% of users do e-shopping at least once a month, only 39% had interacted with e-government in the last year (Petricek, Escher, Cox & Margetts, 2006). According to the United Nations (2012) e-government report, the level of e-government usage is low worldwide; in EU27 countries, the average usage rate is 32%, and in the Organization for Economic Co-operation and Development (OECD) countries it is about

40%. The United Nations report attributes limited adoption of e-government mainly to privacy and security concerns plus a lack of usability (whether the site is designed for easy use by citizens). The more citizen-centric e-government services areand the stronger the user focus, the more likely their adoption is, indicating a shift from what services governments can provide to what citizens really need.

To frame local entities in the e-government context, the term "local government" defines governments that are not central or national but are state, provincial, regional, municipal, or city governments (Lanvin & Lewin, 2006). Arslan (2008) refers to local governments as municipalities or e-cities, whereas Zevedeos (2006) distinguishes the term *municipal e-government*, used in Europe, from *local e-government*, more likely used in the United States. From the e-cities' perspective, Kaylor, Deshazo, and Van Eck (2001) derived a wider definition of e-government: the ability for anyone visiting the city website to communicate and interact with the city via the Internet in any way more sophisticated than a simple email letter to the city email address. Recent surveys in Europe show that 50% to 80% of the interaction between citizens and government occurs at the local levels (Moraru, 2010). Thus municipalities are key influences in citizens' lives. Focusing on citizen-centric websites, says Moraru , should be at the core of e-government, and municipalities need to acknowledge and work to improve the online citizen-government relationship.

Precisely how well are e-governments progressing around the world, and how can one measure website quality? Often e-government is evaluated by a benchmark, which is a technique for comparing e-government based on indicators that yield some sort of score (Flak, Olsen, & Wolcott, 2005). A framework or a model is a set of concepts, values, metrics, and practices that represent a method of viewing reality. The importance of developing models lies in their supplying a basis to measure and guide e-government (Berntzen & Olsen, 2009). It is argued, however, that e-government research is hampered by a want of comparison or comprehensive analysis of e-government models. What is lacking is more in-depth analysis and a deeper recognition of e-government models at national and local levels. Sandoval-Almazan and Gil-Garcia (2008b) stated that research assessing the limitations of e-government models is scarce. In order to fill a gap in the literature on e-government evaluation models, this paper has the following objectives: (1) to identify major e-government evaluation frameworks at national and local levels, (2) to

determine the strengths and weaknesses of their methodologies, and (3) to compare the evaluation models and note whether they differ between national and local e-governments.

This paper is organized into seven sections including this introduction. The second section briefly introduces the concept of metrics, models, and web quality in terms of usability. The third section identifies the methodology followed. The first part of the fourth section considers e-government models at national level while the second part is about e-government models at local level, with the limitations of each. The fifth section compares national e-government models to local ones. Following that is the discussion section, and the seventh section concludes the paper.

2. Web Quality, Metrics and Models

Web quality is still a debatable issue, and there are many parameters for measuring the Web, as different perceptions of quality lead to diverse criteria. Consequently, Web metrics are considered a valuable area of ongoing research (Calero, Ruiz, & Piattini, 2005). Gibson (2006) says that the lion's share of research in Web development is focused on website usability and metrics. Other researchers such as Signore (2005) and Calero et al. (2005) have also noted that website quality is defined in terms of usability. Aikio (2006) has described usability as a measure of the success of a product, whether it is software, computer systems, or any other product. To define an appropriate set of metrics, one needs to determine a list of quality factors that are important for an object (Freire, Fortes, Turine & Paiva, 2008). So Web metrics cannot be regarded as one-size-fits-all and existing research indicates that they differ, to some extent, by website categories, such as government and commerce (Hong, 2007). For a website to be successful there must be a match among the organization's objectives, the user's goals, and the website's design (Bélanger et al., 2006).

According to the Association for Computing Machinery (ACM): "Usability engineering, also known as human-computer interaction engineering, is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and the study of major phenomena surrounding them" (Folmer & Bosch, 2004). User-centered design is a broad philosophy, and there is a variety of methods for designing usable systems that place the users at the center of the design (Hasan, 2009). Thus, Web usability has moved from being a "nice-to-have" to being a "must-have" (Yan & Guo, 2010). Usability

cannot be measured directly, however, it needs to be decomposed into specific attributes and then into metrics. The form of measurement structure is a model or criteria system used to describe usability quality (Li, Yu, & Liu, 2010).

3. Methodology

A critical analysis of e-government evaluation models at national and local levels was undertaken. The methodology of this research follows systematic online searches in order to find major e-government models developed since the year 2000, which represented the onset of e-government models (Hu, Xiao, Pang & Xie, 2005; Montserrat, 2010). A cross-search among several computer and technology databases was employed to retrieve related articles. The literature review spanned the broad spectrum of journals specifically focused on e-government benchmarking and evaluation. Hence, a large number of models assessing national e-government were found in the literature, while less research has addressed the area of local e-government models. To the best of our knowledge, we did not find research that covered e-government models at national and local levels under one umbrella.

Hence, the data collected includes about 60 scientific articles examining different aspects of national and local e-governments in addition to a large number of well-established e-government reports that have been published periodically by international organizations\companies such as the United Nations (UN), Accenture, and Capegemini. Table 1 presents a sample of e-government studies and the models they addressed.

Table 1: Sample of E-government Evaluation Models

References	E-government Models
Rorissa, Demissie & Pardo (2011)	UN, West, Layne & Lee
Sandoval-Almazan & Gil-García (2008a)	UN, Layne & Lee, Moon
Andersen & Henriksen (2006)	UN, Layne & Lee, World Bank, Moon
Arslan (2008)	UN, West, Layne & Lee, Moon, KEeLAN, UN Digital Governance in Municipalities
Berntzen & Olsen (2009)	UN, West, Accenture, Layne & Lee
Bevan (2005); Scowen (2007); Ivory & Megraw (2005)	HHS guidelines
Coursey & Norris (2008)	UN, Layne & Lee, Gartner

References	E-government Models
Denfeld et al. (2002)	Community Benchmarks Program (CBP)
Flak et al., 2005	UN, West, Accenture, Capgemini, MeGAP
Heeks (2006)	UN, West, Accenture, Capgemini
Hu et al. (2005)	UN, West, Accenture
Jansen (2005)	UN, West, Accenture
Kaylor et al. (2001)	MeGAP
Kunstelj & Vintar (2004)	UN, West, Accenture, Capgemini, KEeLAN
Montserrat (2010)	UN, Capgemini, MeGAP, Gartner, Layne & Lee ,UN Digital Governance in Municipalities
Moraru (2010)	West, MeGAP, Layne & Lee, World Bank, UN Digital Governance in Municipalities
Salem (2008)	UN, West, Accenture, Capgemini

Further, an analysis of e-government up to the year 2004, conducted by Kunstelj and Vintar (2004) categorized existing e-government approaches by the aspects (indicators) they cover: 1) e-readiness includes readiness of government, citizens and businesses to e-participate; 2) the back-office includes the reengineering and digitalizing processes within the administration; 3) the front-office consists of a number of online services and information (a. supply-side; b. demand-side); and 4) their effects and impacts. Also the authors have showed that the majority of e-government studies focused on the front-office supply side, "the government", and less on the demand side, "the citizens and businesses", while largely neglecting the back-office and the impact of e-government. Rorissa et al. (2011) attributed that trend to the expensive data collection and complex processing of the back-office approach. Therefore, this research will not address other e-evaluation tracks such as the driving success factors behind e-government, including social utilization and e-readiness, e.g. the availability of ICT infrastructure and online penetration. Also, it is beyond the scope of this research to consider the effects and impacts of e-government. Other critical variables, such as usability, are discussed in this paper for the sake of understanding some e-government models that included usability as a measurement attribute.

The main purpose of this research is to identify major e-government models, whether national or local ones, and to pinpoint their weaknesses and problems as well as their strengths. For our analysis, we selected the well-known and frequently cited e-government models in the academic community and the practice field. At the national level, we adopted and extended the classification of e-government models proposed by Schedler and Schmidt (2004). Since fewer basic models were proposed on local e-government, we intend to describe all that we found. Therefore, for this study, the sample of national e-government models includes the UN, West (or Brown University), Accenture, Capgemini, Gartner, Layne and Lee and HHS guidelines. For municipal evaluation models the sample includes the UN Digital Governance in Municipalities Worldwide, MeGAP (Municipal E-Government Assessment Project), CBP (Community Benchmarks Program), and KEeLAN (Key Elements of Electronic Local Authorities' Network).

4. E-Government

By 2012, most countries had embraced e-government with varying levels of success; of the 193 United Nations Member States, only three countries were not online, so about 98% of the world's countries have government websites available on the Internet (United Nations, 2012). In spite of a variety of descriptions, there is still no standard accepted definition of e-government. The World Bank conceives e-government as the use of ICT, such as the Internet and mobile devices, to transform relations with citizens and businesses, and between branches of government (Lanvin & Lewin, 2006). According to the United Nations (2008), e-government is the use of ICT to improve the activities of public-sector organizations and deliver services to citizens. A common intersection between different e-government definitions is the digitization of governmental operations and processes.

E-governments reduce travel and waiting time (moving processes from in-line to on-line), eliminate corruption, reform government, increase transparency, enhance the relationship between government and citizens, and ultimately develop democracy (Al-adawi, Yousafzai, & Pallister, 2005). E-services are cheaper, faster, and readily available 24/7. Practical examples of e-government's financial benefits include the Information Network of Kansas generating a revenue of 7 million USD per year and Singapore e-Tax saving SGD 20 million per year (Mohammad, Almarabeh, & Ali, 2009).

4.1 E-Government Evaluation Models at National Level

A model is used to derive suitable indicators for evaluating various e-government initiatives (Berntzen & Olsen, 2009). The importance of creating such models lies in its offering a basis to measure and guide e-government development by drawing attention to best practices. Actually, the construction of e-government models began in 2000 (Montserrat, 2010). Inspired by Schedler and Schmidt (2004), we propose to classify the e-government models into three kinds: organizations and consultancy firms, scholars, and official government models.

1. Organizations and Consultants E-government Evaluation Models

Several organizations tried to understand the e-government phenomenon by constituting models which are divided further into different numbers of stages of growth with specified features that must be fulfilled in each stage. Heeks (2006) has confirmed that stage models have their origins in private-sector e-commerce models, and Yildiz (2007) has said that e-government is studied by developing models of its stages. Unfortunately, there is no agreement among organizations on the number of stages and requirements. The most established e-government evaluation reports, published periodically and cited frequently, are identified in Table 2.

Table 2: E-government Evaluation Models developed by Organizations

Model	Focus	Stages			
UN 2002	Worldwide	Emerging	Enhanced	Transactional	Connected
Brown University 2001	Worldwide	Billboard "Information"	Services Delivery	Portal	Interactive Democracy
Accenture 2000	22 Developed Countries	Publish Passive/Passive Relationship	Interact Active/Passive Interaction	Transact Active/Active Interaction	
Capgemini Europe 2002	European Countries	Information	One-way Interaction	Two-way Interaction	Transaction

- **The UN Model:**

The United Nations has been assessing e-government since 2002 (Berntzen & Olsen, 2009). Initially, the UN e-government model was described by Rutgers University as a three-stage model (Montserrat, 2010). Currently, it is well established and widely used in many studies, and it has two indices: the e-government index (Table 3) and the e-participation index. The e-government index ranks e-governments worldwide at the national and ministry websites. Each of its three measures (online service, telecommunication infrastructure, and human capital) is a composite index that can be analyzed independently, with a value between one and zero. The recent online-service index was based on a four-stage e-government model: emerging, enhanced, transactional, and connected presence, with specified features for each stage (see United Nations, 2012).

Table 3: The UN E-government Model (United Nations, 2012)

First Class Index	Second Class Index	Third Class Index
Overall Development	Online-service Measure	Emerging presence
		Enhanced presence
		Transactional presence
		Connected presence
	Telecommunication Infrastructure Measure	PCs / 100
		Internet users / 100
		Broadbanding / 100
		Telelines / 100
		Mobile phones / 100
		TVs / 1000
	Human-capital Measure	Adult literacy rate (%)
		Combined gross enrolment ratio for primary, secondary, and tertiary schools (%)

- **The Brown University (West) Model:**

Professor West and his research team at Brown University have conducted an annual evaluation report of government websites since 2001 (Berntzen & Olsen, 2009). The report analyzes government websites worldwide for the presence of 18 features, such as phone and address contact, publications, audio and video clips, number and type of e-services, privacy, and security policies. The ranking runs along a scale from zero to 100 points (West, 2008): 4 points for the presence of 18 features totaling 72 points, and one point for

one service up to 28 points for 28 or more e-services. Moraru (2010) recognized West's stages as: (1) Billboard: online government information; (2) Service delivery; (3) Portal: "one-stop shop" concept, security, and privacy; (4) Interactive democracy.

- **The Accenture Model:**

Accenture is a consulting, technology services, and outsourcing company that has been issuing annual e-government reports on developed countries since 2000 (Hu et al., 2005). The original Accenture model included two dimensions -customer relationship management (30%) and service maturity (70%) - with two indices: the number of online services (service breadth) and the level of service completeness (service depth). Service maturity is decomposed into the following stages (Peters, janssen & Engers, 2004): (1) Publish—passive/passive relation: no communication between users and government; (2) Interact—active/passive interaction: only users can e-communicate with government; and (3) Transact—active/active interaction: two-way communication is possible.

Berntzen and Olsen (2009) record some modifications on the Accenture rankings. The 2005 Accenture index had two components, each with a weight of 50%, service maturity and customer service maturity, which were measured by four dimensions: citizen-centered, multi-channel, cross-government service, and proactive communication about the services to users. Four hundred citizens in each evaluated country were questioned about their country's e-services, and interviews of 46 high-ranking government executives were conducted. The 2007 report introduced a new indicator, citizen voice (40%), reducing the weight of service maturity to 10%.

- **The Capgemini Europe Model:**

Capgemini (2006), a company specializing in consulting, technology, outsourcing, and local professional services, focuses on evaluating the e-presence and sophistication of government websites in 32 European countries. According to Capgemini (2010), the survey benchmarks 20 basic online services, 12 services to citizens, and 8 services to businesses with the following indicators: online sophistication, full online availability, user experience (usability, transparency, privacy, multi-channel policy, and users' feedback), and portal sophistication (most mature, user-centric, and personalized portals). Basic citizen services include income tax, job search, social-security benefits, personal documents (passports,

driver's license), car registration, building permission application, declaration to the police, public libraries, birth and marriage certificates, enrollment in education, announcement of moving house and health-related services. The online sophistication and availability rankings assess the 20 public services against four stages in the 2006 report then against a 5-stage maturity model in the 2010 report: information, one-way interaction, two-way interaction, transaction, and automation threshold (proactive, automated service delivery).

2. Scholars E-Government Models

Sparse contributions to this vital subject are still evolving, as several scholars offer their own insights. The first e-government model was proposed by Baum and Di Maio (Gartner model) in 2000 and has four stages: Web presence, interaction stage, transaction stage, and transformation stage (a citizen-centric and responsive government) (Montserrat, 2010). Another highly cited e-government model was proposed by Layne and Lee (2001) with reference to the USA in four stages: catalogue, transaction, vertical integration (connecting government agencies), and horizontal integration (one-stop portal) (Andersen & Henriksen, 2006).

3. Official Government Frameworks

Several governments developed their own official frameworks to help designers build high-quality e-government websites. A good example is the USA Research-Based Web Design and Usability Guidelines created by the Department of Health and Human Services (HHS) according to the best available up-to-date research. The HHS guidelines were praised by some researchers such as Scowen (2007), Bevan (2005), and Ivory and Megraw (2005). These guidelines are widely used by government agencies and private sectors, and also translated into several foreign languages. The project began in 2000 with 500 guidelines, but shortly was reduced to 398 and now 209 guidelines. Each guideline has a rating for its "Relative Importance" to the success of a website and a rating of the "Strength of Evidence" supporting the guideline. To determine the "Relative Importance," eight website designers and eight usability specialists assigned each guideline a rating from 1, for the least important guidelines, to 5, for the most important. The "Strength of Evidence" represents a consensus among a group of 8 usability researchers so that the users can verify the quality of the supporting evidence (HHS, 2013).

4.1.1 Limitations of National E-Government Evaluation Models

The majority of models, such as the UN, Capgemini, and Brown, are based on objective measures and follow a quantitative approach; only Accenture uses hybrid measures (Salem, 2008). The quantitative method may lead to a dilemma if not designed properly. Curtin (2006) has said that higher ranking may not predict better performance, since most surveys do not evaluate qualitative issues such as the quality of service or the citizen usage of e-government. Furthermore, Jansen (2005) has showed that a number of experts have interpreted the framework differently.

Another problem with most of these models is their focusing on the supply side (government) not the demand side (citizen and business) of e-government (Berntzen & Olsen, 2009). Two examples of the supply-side models are West and Capgemini, while the demand-side models are like Gartner and HHS guidelines (Rorissa et al., 2011; Flak et al., 2005; Scowen, 2007). The imbalance of the abundance of government-side surveys compared with the scarcity of citizen-side studies has led to a misinterpretation of the final objective of e-government. The existing practices are pushing countries to prioritize getting good ratings for creating many services without caring whether citizens use them or not (Montserrat, 2010). Accenture tries to overcome this shortcoming and uses interviews to determine the citizen's point of view.

Unfortunately, most countries launch e-government through the "quick fix, quick wins" principle and hastily construct the e-equivalent of a bureaucratic administration (Kunstelj, & Vintar, 2004). Usually, countries reach the second stage easily and quickly, as it takes no great effort to supply information, forms, and emails. In contrast, a website that advances from stage 3 to stage 4 has to go through tremendous changes that require massive efforts and resources to provide transaction and a one-stop portal (Rorissa, Demissie, & Pardo, 2008).

In fact, the conceptualization into stages is doubtful. There has been some criticism aimed at stage models, focused on the evolutionary aspect and the quality assumptions of these models: 1) The assumption that evolutionary stages are independent seems not to be true empirically. An e-government website may have characteristics of multiple stages; **2)** The assumption that evolutionary stages are consecutive, linear progressing and higher

stages include lower stages, seems not to be true empirically also. The models predict that the e-government evolutions occur in pre-described order; first stage 1 occurs and then stage 2 and so on, but in practice the stages occur simultaneously. It could be that some e-portals had characteristics of advanced stages but did not have features from the early stages (Sandoval-Almazan & Gil-Garcia, 2008a, 2008b; Goldkuhl & Persson, 2006).Other researchers have criticized Web metrics proposed for Web systems because they tend to be simplistic and define very general criteria (Signore, 2005) or are not well defined at all (Calero et al., 2005). There is a rush to develop more Web metrics without any kind of validations, which may make the use of them dangerous and difficult. Calero et al. (2005) have found 3% of metrics in the literature validated theoretically and 37 % validated empirically; there are also hundreds of Web metrics available, but no guidelines for their use. The lack of evaluations of existing metrics leads researchers to develop more new metrics without knowing how similar these metrics are or what each metric is measuring (Vigo & Brajnik, 2011).

The UN model is widely used by many studies, and it is unique in including three measures (Berntzen, & Olsen, 2009). Yildiz (2007), however, has concluded that the UN and Layne and Lee models are oversimplifications. Abanumy, Mayhew, and Al-Badi (2003) have criticized the UN model for being too general and having too many features. The problem in ranking occurs when a website covers some but not all features in a certain stage; then, it cannot be ranked correctly as belonging to any stage, and it is difficult to distinguish between a ministry that fulfills 100% of the stage features and one that fulfills just 20%. The authors have tried to solve this problem by splitting each UN stage into three layers.

The Brown University reports lack a detailed description of their e-government methodology (Schellong, 2009). They give more weight to the number of features and too little to services, underestimating their importance. A government website offering 28 services is presented as equal in score to another website offering hundreds of services, because the maximum score for services is 28. In addition, the reports check only the presence of services without measuring their quality. Rorissa et al. (2008) have examined the profiles of two government websites according to the Brown University model and concluded that the model may suggest inaccurate conclusions. A country with a single e-government website may have the same e-government index value as a country with five

websites. Another criticism for the Brown University model is that it has decreased its measurement criteria over the years; in 2001, 2002, 2003, 2004, 2005, and 2006, the number of measures were 24, 25, 20, 19, 19, and 18, respectively (Holzer & Kim, 2005). Consequently, there were inconsistencies in annual rankings from year to year; for instance, Portugal has fluctuated in ranking from position 182 to 133, 31, 86, 43, 48, 7, and then 18 in an eight-year period (Schellong, 2009).

Essentially, the strength of the Accenture model lies in the evaluation of the maturity of e-services following a hybrid methodology, quantitatively assessing the breadth and depth of e-services and qualitatively appraising the customer service delivery. Another strength is the new indicator introduced in 2007, "citizen voice," which integrates user views of e-government. On the other hand, this model's main weakness is its continual changes in methodology and measurements, which make it impossible to compare e-government rankings over the years (Berntzen & Olsen, 2009). Furthermore, Accenture provides no details of measured services and their maturity scores. Thus, the calculation of the indices is not reproducible. Kunstelj and Vintar (2004) have implied that Accenture lacks an evaluation of integrated services, and since it focuses on only 22 countries, this model is limited in its application.

The most common critique of the Capgemini model is its focus on the government side only. Also, Kunstelj and Vintar (2004) criticize the Capgemini for its measuring the availability of 20 public services despite some of these services bringing no value to customers. They add that highlighting the technological side of e-government without considering the quality of information and usefulness of services will miss important qualitative aspects of e-government. A further problem is that higher stages of the model do not necessarily imply the existence of lower stages: for example, a service can reach stage 3 or stage 4 without offering downloadable forms. In addition, this model is narrow in its scope, being concerned only with European countries. However, the Capgemini (2009) report has claimed a paradigm shift toward customer-centric services. New patterns of relations go from the "you-centric" model to the "me-centric" model, changing the role of the user from that of a passive viewer and user to that of an active creator of the public-service delivery chain.

Scowen (2007) has praised the official American HHS guidelines for being supported in the Human Computer Interaction (HCI) field, and Ivory and Megraw (2005) have said that

they are clear and have been validated empirically. Bevan (2005) has compared them with the ISO 9241-151 standard and the JISC (Joint Information Systems Committee) guidelines for the UK academic websites. The JISC carried out an extensive search to adapt the best guidelines to its services, and thus confirmed the superiority of the HHS guidelines even though they were not exhaustive and omit some material specific to e-commerce. The study also shows how subsets of the HHS guidelines can be tailored for particular audiences. Nevertheless, very few studies assess these guidelines. One can say it may be difficult to evaluate a website against too many guidelines (currently 209 guidelines). When Nielsen (1993) succeeded in condensing usability principles to only 10 heuristics, many researchers adopted his evaluation and built upon it. Likewise, it may be better for the HHS to work on providing the Web community with a shorter list of guidelines.

Formerly mentioned frameworks revealed that many e-government reports were based on different measurement instruments, which explains the difference in e-government rankings and the disparity of conclusions. Furthermore, several scholars of e-government are skeptical about the e-government rankings and have justifiably argued that existing e-government frameworks have some methodological limitations (Schellong, 2009; Yildiz, 2007; Rorissa et al., 2011; Sandoval-Almazan & Gil-Garcia, 2008b). Their analysis shows a messy picture of the measurement of e-government. Yildiz (2007) has indicated that the "stagi-est" approach to e-government is unsatisfactory and that the development of stages does not necessarily follow neatly in a linear order. He adds that such models are not applicable to e-government, especially in developing countries. Ataloglou and Economides (2009) and Peters et al. (2004) have concluded that a good theoretical framework for measuring the impact of e-government is still lacking. Schellong (2009) has said that a relevant and universally accepted e-government model still needs to be developed.

From the preceding discussion, it is evident that there is no generally accepted comprehensive e-government evaluation framework and no universal standard for assessment at the national e-government level (Ataloglou & Economides, 2009; Jansen, 2005).

4.2 Municipal E-Government Evaluation Models

Most of the frameworks evaluating municipal websites are based on Moon's work. Moon (2002) has proposed a framework of five stages: information dissemination/catalogue, two-way communication, service and financial transactions, vertical and horizontal integration, and political participation. Having surveyed 1,471 US e-municipalities with populations over 10,000, he has found that larger municipalities are more advanced in e-government; nevertheless, they are still at an early stage of development and have not reached many of their expected outcomes. The four most popular local e-government models within the literature are the following.

1. Digital Governance in Municipalities Worldwide

Co-sponsored by the UN, this benchmark compares the largest e-cities globally every two years since 2003. It is still the only framework that evaluates municipal websites worldwide in terms of digital governance, which includes digital government (delivery of public service) and digital democracy (Holzer, You, & Manoharan, 2009). Moon (2002) selected the largest city in a country to represent that country regardless of its advances in e-government; the UN study sampled cities by the same principle. Montserrat (2010) considers the sampling in this survey to have been biased, but its methodology remains constant over the years. The instrument for assessing city websites consisted of five components: security and privacy, usability, content, services, and citizen participation. The research applied 18–20 measures coded on either a scale of 1: information exists on the website; 2: downloadable items are available; and 3: services, transactions, or interactions are completely online, or a dichotomy of two points, (0, 3) in the "service" and "citizen participation" and (0, 1) in the "privacy" and "usability" categories. Hence, the survey instruments used 98 measures (see Holzer et al., 2009). To ensure reliability, each municipal website was assessed by two evaluators who were given clear instructions.

2. MeGAP

The MeGAP (The Municipal E-Government Assessment Project) is an assessment tool for US municipal websites emphasizing online service provision. Kaylor et al. (2001) surveyed 38 American cities with a population between one and two hundred thousand

people and developed a rubric for evaluating them. Functional performance dimensions were grouped into 12 categories containing 51 e-services. To rank municipalities, each service was scored on a 1–4 scale (information, contact, downloadable forms and transaction or interaction) that yielded an e-score corresponding roughly to the stage model concepts (Flak et al., 2005). In 2005, a third version of Kaylor's survey maintained the original framework, but the catalogue expanded to 68 local services in 4 categories (Montserrat, 2010; Flak et al., 2005): (1) Information dissemination (city codes, minutes, traffic information, municipal government directory); (2) Interactive functions (bidder applications, downloadable forms, building permit process, business license); (3) E-Commerce functions (utility payment, tax look-up and payment, code enforcement); (4) E-Democracy (e-meetings, e-forums, user customization).

3. Municipal Website Assessment of Community Benchmarks Program

The Maxwell School at Syracuse University established the Community Benchmarks Program (CBP) in 1999 and developed a website assessment instrument to evaluate e-municipalities in Onondaga County. Denfeld et al. (2002) re-evaluated the previous study and devised the following assessment criteria:

- Information available: municipal meeting, minutes, budget, downloadable forms, date of website update;
- Contact information: phone and fax numbers, e-mail, physical address;
- Architecture: search, site map, link function properly, link to home page provided;
- Continuity of Web design: consistent design of all pages;
- Search engines: placement of the municipality's website on Yahoo, Google, and MSN for official name, popular name;
- General: responsiveness of town clerk, unique features of each site (both well and poorly executed), broken links.

The 2002 report assigned each attribute a score of 1, if the website met the criterion, or 0, if it did not. An example of a blank evaluation form for the "Information Available" criteria is in Table 4.

Table 4: An Example of a blank evaluation form (Denfeld et al., 2002)

	Information Available			
	Attribute	Yes	No	Comment
1	Date of **next** town/village board or city council meeting (acceptable in place of the date is, i.e. third Monday of the month)	1	0	
2	Location of town/village board or city council meeting	1	0	
3	Time of town/village board or city council meeting	1	0	
4	Agenda of next town/village board or city council meeting (posted within 48 hours of meeting—use last meeting agenda if necessary)	1	0	
5	Minutes of the last town/village board or city council meeting held within the last two months	1	0	
6	Archive of past board or council meeting minutes	1	0	
7	Minutes can be downloaded	1	0	
8	Budget for the current fiscal year	1	0	
9	Downloadable forms	1	0	
10	A date is provided for the most recent Web update	1	0	

4. Key Elements for Electronic Local Authorities' Network (KEeLAN)

The Key Elements of Electronic Local Authorities' Network (KEeLAN) is a local e-Europe government framework and is also known as "Framework Programs." Started by e-Europe research, the KEeLAN framework is divided into two phases measuring e-government and back-office development. The e-government stages are divided into six phases: stage 0: no Web presence; stage 1: information (about services); stage 2: interaction (downloading forms); stage 3: two-way interaction (processing of forms including authentication); stage 4: transaction (full case handling); stage 5: service integration (online service enabled by a secured network linked to various back-offices/service modules). The stages are exactly the same as in the Capgemini model, except the last one. In this context, a Web assessment tool contains questions to evaluate e-cities on 9 basic services: policy making, economic development, personal documents, credit and loans/financial support, education, building permits, environment, culture and leisure, and information dissemination. Depending on the interactivity, a score is computed to indicate the stage of the service (Arslan, 2008).

4.2.1 Limitations of Municipal E-government Evaluation Models

Even though most of the time the interaction between citizens and government occurs at the local levels, one can find very little research that describes or analyzes existing local e-government models. In fact, there is a disproportionate number of studies focusing on

national e-government models compared with that of studies targeting local e-government models (Heeks, 2006).

Nevertheless, the UN's "Digital Governance in Municipalities Worldwide" is still the only international survey of e-cities. By supporting two different models, the UN implies that there is a difference between assessing central e-governments and assessing local ones. The methodology of digital governance has remained constant over the years, so its rankings of cities are comparable and remarkably informative. On the other side, Montserrat (2010) regards the sampling in this survey as biased. Also, the survey gives no justification for the framework measurement evaluation criteria, which constitutes a major weakness in the methodology. Each municipal website was assessed by two evaluators given clear instructions (Holzer et al., 2009). But, no information was given about the evaluators' backgrounds and their degree of expertise.

For the MeGAP of the US e-municipalities, Flak et al. (2005) believed that this model gives a more detailed analysis of the depth and breadth of municipalities than any other assessment model; but, on the other hand, the MeGAP lacks a firm theoretical foundation, doesn't assess usability, and it is a country-specific model. The major drawback of the Community Benchmarks Program is that it focuses only on the supply side of e-government. Since the two models are similar, the KEeLAN model suffers from the same problems as Capgemini (quantitative approach, focus on government only).

5. Comparison of National and Municipal E-Government Models

There is still no agreement on how to measure e-government and devise metrics for the Web. At the national level, the existing e-government models are very similar and are based on analogous attributes and measures; they view e-government as stages of growth and adopt four or five stages: Web presence, interaction, transaction, integration (portal), and e-participation or e-democracy (included in few models).

On the other hand, Yildiz (2007) has criticized the "stagi-est" approach to assessing national e-government and complained that there is no agreement on the number of stages and requirements. Toonders (2010) has deemed it unclear whether the same stages of national e-government are useful for describing local e-government. Norris (2009) has cast doubt on the adequateness of stage models in municipalities. He used survey data from US

municipalities over three years (2000, 2002, and 2004) and empirically examined how e-government has developed in practice and contrasted this with the predictions of the models. The US e-municipalities did not progress through stages as anticipated. They were informational with fewer transactions and interactions and had not evolved into e-democracy yet. Norris attributed that to the e-government models having been developed in a vacuum and not being based on research or even reviews of literature, so that, even after 10 years of adoption, e-government has not reached higher stages of development in most countries.

Again Norris and Reddick (2013) addressed the trajectory of US local e-government using empirical data from two nationwide surveys of American local governments conducted in 2004 and 2011. They found American local e-governments are delivering information and services online with few transactions and limited interactivity and they are mainly one way, from government to citizens, with no evidence that it is transformative. The authors also presented more empirical studies of e-government; for example: service has been the primary focus of e-government in various locations such as the United Kingdom (McLoughlin and Cornford 2006), Canada (Roy 2006, 2007), Australia (Dunleavy et al. 2008), the Arab nations (Chatfi eld and Alhujran 2009), and Italy (Nasi and Frosini 2010). This is consistent with the conclusion of Sandoval-Almazan and Gil-Garcia (2012) who said that almost a decade after the publication of a similar study on U.S. municipalities by Moon (2002), the results of their assessing Mexican local e-government remain very similar. They believed that e-government in municipalities is still more rhetoric and less reality, at least in some countries. In fact, the e-government experience differs dramatically from the national to the local level and from one country to another. Montserrat (2010) believes that the indicators and metrics defined for national e-government are not applicable at the local level He asks, "Why are there no benchmarks at local government?" Collecting comparable data about e-municipalities is a difficult task because of differences in political and economic systems. The different role played by cities is one of the challenges that scholars must address. Montserrat also confirms a clear lack of local e-government evaluation models. Heeks (2006) says that most studies focus on national e-government, although in developing countries it is local governments that are the main point of contact for delivery of services. Most public services that are relevant to

citizens are offered by the local e-government, and this is a possible source of error in the assessments (Berntzen & Olsen, 2009; Schellong, 2009).

Through its development of two models, the UN demonstrated the difference between assessing national and local e-governments. For assessing state portals, Sandoval-Almazan and Gil-Garcia (2008a) identified three approaches: 1. managerial, 2. evolutionary (e-government stages) and 3. citizen-centered perspectives. Using a mixture of the last two approaches, they assessed 32 Mexican portals against a six-stage model and also against other important variables such as usability, openness, customization, transparency, e-services, privacy, security, etc. Another contribution by Goldkuhl and Persson (2006) is a proposal to replace the one-dimension stage models (called e-ladder) by a three-dimension e-diamond model consisting of three polarities (informative vs performative, standardized vs individualized; separate vs coordinated).

Upon analyzing existing normative models on municipalities (Table 5), it is noticeable that some of them, such as the UN Digital Governance in Municipalities and CBP, focus on general aspects of the site such as content and services. They avoid the concept of stage models and instead regard local e-government as different components or categories. Other models such as, the KEeLAN and MeGAP, follow the stage model (Arslan, 2008; Flak et al., 2005). Also worth mentioning are the individual efforts by some authors, such as Sandoval-Almazan and Gil-Garcia (2008a), Moraru (2010) and Luna, Gil-Garcia, Luna-Reyes, Sandoval-Almazan & Duarte-Valle (2013), who use a mixture of e-government stages and some other components they perceived important in the evaluation of municipal websites.

Table 5: Two Kinds of Municipalities Models

Models	Kind of Model	Descriptions				
Digital Governance in Municipalities Worldwide 2003	Components	Security/ Privacy	Usability	Content	Services	Citizen Participation
Community Benchmarks Program 1999	Components	Content	Architecture		Layout	Website Design
MeGAP 2001	Stages	Information	Contact		Downloadable Forms	Transaction or Interaction
KEeLAN 2002	Stages	Information	1-way Interaction	2-way Interaction	Transaction	Service Integration

6. Discussion

There is still no agreement on how to measure governments online, and this has become a valuable area of ongoing research. The situation remains arbitrary since there is a rush to develop more e-government models without any validations, and most of these models are not based on solid research. Several scholars, such as Schellong (2009), Yildiz (2007), and Rorissa et al. (2011), are skeptical about the e-government rankings, and their analyses show a confusing picture of e-government.

We have classified three kinds of national e-government models : organizations' models (UN, Brown University, Accenture, and Capgemini), scholars' models (such as the Gartner model and the Layne and Lee model), and official government models (e.g. USA Research-Based Web Design and Usability Guidelines). For local e-government, the most popular models are the UN Digital Governance in Municipalities, the U.S. MeGAP, the Community Benchmarks Program, and the KEeLAN Europe model.

The common limitations of most e-government models include focusing on the government rather than the citizen side, using quantitative measures, and not considering qualitative issues such as the quality of services, constructing the e-equivalent of a bureaucratic administration, or defining very general criteria without sufficient validations.

The UN national model has been widely used by many studies, but it has been criticized for being too general and having so many features. The Brown University model assigns more weight to the number of features and too little to services. The Brown and the Accenture models have changed their measurement criteria over the years, so they are inconsistent in their annual rankings; the Accenture model, moreover, lacks an evaluation of integrated services and has been applied to only 22 countries. The Capgemini model, limited to European countries, focuses on the government side only and checks the availability of e-services without measuring their quality. Some studies have praised the official American HHS guidelines for being clear and validated empirically, but it is difficult to evaluate a website against 209 guidelines.

It is hard to find research that discusses the limitations of e-government model at the local level. Nevertheless, the UN model of Digital Governance in Municipalities is still the only international survey of e-cities, and its methodology has remained constant over the

years. On the other hand, no justification has been given for its evaluation criteria, nor any information about the evaluators' backgrounds. Although it gives a detailed analysis of municipalities, the MeGAP is a country-specific model particular to US municipalities and also lacks an assessment of website quality. The major drawbacks of Community Benchmarks Program and the KEeLAN models are their focus on the government side and consideration of only quantitative measures.

Comparing national with local e-government models has revealed interesting findings. The existing e-government models are very similar in viewing e-government in terms of stages of growth. Many, however, have criticized th "stagi-est" approach; the stages are not independent or consecutive, and there are no agreements on the number of stages and requirements. Furthermore, the adequacy of stage models for assessing municipalities is suspected. Some of the existing local government models avoid the stage approach and instead adopt the concept of viewing local e-government as a different component.

The model requirements for e-government vary from those for e-commerce, the e-experience differs from national to local governments, and there is a disparity between cities in politics, economics, and type of public services. Thus, the metrics defined for national e-government are not applicable to municipalities, and the different roles played by cities make the development of a city model far more challenging for scholars.

Most studies have focused on national e-governments despite local governments being the main point of contact with citizens, and this may lead to misreading the aims of e-government. Thus, there is a clear lack of local e-government evaluation models. Most e-government reports, however, have focused on the government, thus enhancing the image of the government and not prioritizing citizens' needs or facilitating their lives. Yet, a complete view of e-government in cities is not possible (Montserrat, 2010). Also e-government has had too little user testing to convey the voices of citizens. A well-developed citizen-centric website could greatly benefit the outcomes expected from e-government.

E-government models use good practices to assess development, but they are still an inaccurate reflection of the real situation. It is evident that there is no comprehensive e-government evaluation model and no standard for assessment; therefore, there is a genuine need for developing a good theoretical model for national and local e-governments that are clearly distinct.

7. Conclusion and Future Research

This research examined major e-government evaluation models at national and local levels. The stage model approach seems to be the prevailing trend in the evaluation of e-government and has been taken for granted even though many studies have showed its limitations, as described earlier in this paper. We believe, like Goldkuhl and Persson (2006), that the use of e-government stage models seems to be misleading and erroneous and should be abandoned. Another solution that exploits the strengths and reduces the weaknesses of this method could be to think about it as components rather than stages (Sandoval-Almazan & Gil-Garcia, 2008b).Then the right combination of components should be the focus of future research. A good proposal here is to include, in such a way, a combination of website quality and e-services quality.

Also we fully support a more comprehensive evaluation, such as the study of Luna et al. (2013), which considers the front-office factors (information, interaction, transaction, integration and participation) and other factors such as technology (number of internet, computers, mobile users), organization\institution (government efficiency index), and context (global competitiveness and infrastructure indexes).

The UN model seems to have more strength at the national level because it is comprehensive and has three indexes of measurement: online service, telecommunication infrastructure, and human capital (see Table 3). But this applies only under one condition: to think of the stages as components. At the local level, the UN Digital Governance in Municipalities seems to be the most solid because it is comprehensive and assesses five important components (security and privacy, usability, content, services, and citizen participation). It could stand one improvement, however: the evaluation criteria under each component should be amended based on validated metrics, such as the HHS guidelines.

The field of local e-government needs further research. The general country structure is important in the development of an evaluation model, as online services differ from country to country due to differences in political and economic systems (Moraru, 2010; Montserrat, 2010; Flak et al., 2005). Thus, each country can determine its e-services by reviewing its political system and conducting polls to determine citizens' needs.

E-government is not delivered through websites only and not restricted to a specific technology. As technology evolves, e-government is extending to different delivery channels, such as mobile devices (m-government) and new platforms like social media (Twitter, Facebook, etc.). Montserrat (2010) stated that local administrations are introducing web 2.0 technologies into e-services, and yet there are no e-government stage models that take them into account. Sandoval-Almazan and Gil-Garcia (2012) reckoned that without a plan and set of rules, social media could become disorganized and provide poor results. Therefore, we intend to address this subject in a future paper.

References

Abanumy, A., Mayhew, P., & Al-Badi, A. (2003). An Exploratory Study of E-government in two GCC Countries. The 2003 International Business Information Management Conference, December 16 – 18th, Cairo, Egypt.

Aikio, K. (2006). Integrating Usability Engineering with Software Engineering: A Preliminary View on Aspects Surrounding the Topic of Usability Integration. University of Oulu, Finland.

Al-adawi, Z., Yousafzai, S. & Pallister, J. (2005). Conceptual Model of Citizen Adoption of E-government. The Second International Conference on Innovations in Information Technology (IIT'05). Available at http://www.it-innovations.ae/iit005/proceedings/articles/G_6_IIT05-Al-Adawi.pdf

Andersen, K., & Henriksen, H. (2006). E-government Maturity Models: Extension of the Layne and Lee Model. Government Information Quarterly 23, 236 – 248. Available at www.sciencedirect.com

Arslan, A. (2008). Assessment of the Turkish Local e-Governments: An Empirical Study. International Journal of Human Sciences, *5*(2). Available at http://www.insanbilimleri.com

Ataloglou, M., & Economides, A. (2009). Evaluating European Ministries' Websites. International Journal of Public Information Systems, *5*(3), 147–177.

Bélanger, F., Fan, W., Schaupp, L., Krishen, A., Everhart, J., Poteet, D., Nakamoto, K. (2006). Web Site Success Metrics: Addressing the Duality of Goals. Communications of the ACM, *49*(12), 114-116.

Berntzen, L., & Olsen, M. (2009). Benchmarking E-government- A Comparative Review of Three International Benchmarking Studies. 3rd International Conference on Digital Society IEEE, 77-82.

Bevan, N. (2005). Guidelines and Standards for Web Usability. Proceedings of HCI International 2005, Lawrence Erlbaum

Calero, C., Ruiz, J., & Piattini, M. (2005). Classifying Web Metrics Using the Web Quality Model. Online Information Review, Emerald Group Publishing Limited, *29* (3), 227-248

Capgemini (2006). Online Availability of Public Services: How is Europe Progressing? European Commission Directorate General for Information Society and Media

Capgemini (2009). Smarter, Faster, Better eGovernment: 8th Benchmark Measurement, November 2009. European Commission Directorate General for Information Society and Media

Capgemini (2010). Digitizing Public Services in Europe: Putting Ambition into Action. European Commission Directorate General for Information Society and Media

Coursey, D., & Norris, D. (2008). Models of E-Government: Are They Correct? An Empirical Assessment. Public Administration Review, May/Jun 2008, *68*(3), 523-536

Curtin, G. (2006). Issues and Challenges: Global E-Government/E-Participation Models, Measurement and Methodology. Workshop on E-Participation and E-Government: Understanding the Present and Creating the Future, United Nations. Budapest, Hungary

Denfeld , Z., Miller, A., Feinerman, F., Nine, C., Hacken, L., Reeder, K. , Hoyle P., & Tamoutselis, P. (2002). 2002 Assessment of Municipal Web Sites in Onondaga County. Community Benchmarks Program. Maxwell School of Citizenship and Public Affairs. Syracuse University.

Flak, L. Olsen, D., & Wolcott, P. (2005). Local E-Government in Norway: Current Status and Emerging Issues. Scandinavian Journal of Information Systems, *17*(2), 41–84

Folmer, E., Bosch, J. (2004). Architecting for usability: a survey. The Journal of Systems and Software, 2004, *70*(1), 61–78. Available at http://dissertations.ub.rug.nl/FILES/faculties/science/2005/e.folmer/c2.pdf

Freire, A., Fortes, R., Turine, M., & Paiva, D. (2008). An Evaluation of Web Accessibility Metrics based on their Attributes. SIGDOC'08, 2008, Lisbon, Portugal.

Gibson, C. (2006). Website Engineering in the Real World. ACM SE'06, March, 2006, Melbourne, Florida, USA

Goldkuhl, G., & Persson, A. (2006). From e-ladder to e-diamond–re--conceptualising models for public e-services.

Hasan, L. (2009). Usability Evaluation Framework for E-commerce Websites in Developing Countries. A Doctoral Thesis. Loughborough University

Heeks, R. (2006). Understanding and Measuring eGovernment: International Benchmarking Studies. Paper prepared for UNDESA workshop, E-Participation and E-Government: Understanding the Present and Creating the Future, Budapest, Hungary, 2006.

HHS. (2013). Research-Based Web Design and Usability Guidelines. Available at http://www.usability.gov

Holzer, M., & Kim, S. (2005). Digital Governance in Municipalities Worldwide (2005): A Longitudinal Assessment of Municipal Websites throughout the World. E-Governance Institute, National Center for Public Productivity. Rutgers, the State University of New Jersey. Newark. Available at http://unpan1.un.org/intradoc/groups/public/documents/aspa/unpan022839.pdf

Holzer, M., You, M., Manoharan, A. (2009). Digital Governance in Municipalities Worldwide (2009): A Longitudinal Assessment of Municipal Websites throughout the World. E-Governance Institute, National Center for Public Performance, Rutgers University, Newark. Available at http://www.scribd.com/doc/36936196/Digital-Governance-Municipalities-Worldwide-2009

Hong, I. (2007). A Survey of Web Site Success Metrics Used by Internet-dependent Organizations in Korea. Internet Research, Emerald Group Publishing Limited, *17*(3), 272-290

Hu, Y., Xiao, J., Pang, j., & Xie, K. (2005). A Research on the Appraisal Framework of E-Government Project Success. Yat-sen University, China. ACM International Conference. Available at http://portal.acm.org/citation.cfm?doid=1089551.1089647

Ivory, M., & Megraw, R. (2005). Evolution of web site design patterns. ACM Transactions on Information Systems, 23(4), 463-497

Jansen, A. (2005). Assessing E-government Progress – Why and What. NOKOBIT, 1504-1697. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.107.9423&rep=rep1&type=pdf

Kaylor, C., Deshazo, R., & Van Eck, D. (2001). Gauging E-government: A Report On Implementing Services Among American Cities. Government Information Quarterly *18*(2001), 293–307

Kunstelj, M., & Vintar, M. (2004). Evaluating the Progress of E-government Development: A Critical Analysis. Information Polity, IOS Press, *9*(2004), 131–148

Lanvin, B., & Lewin, A. (2006). The Next Frontier of E-Government: Local Governments May Hold the Keys to Global Competition. World Bank, Global Information Technology Report, 2007, 51-63
.
Li, X., Yu, Z., & Liu, C. (2010). Developing Usability Measure Structure: Process And Principals. Xi'an Research Inst. Of Hi-Tech, China. 2010 2nd International Conference on Computer Engineering and Technology, 4, 181-183

Luna, D. E., Gil-Garcia, J. R., Luna-Reyes, L. F., Sandoval-Almazan, R., & Duarte-Valle, A. (2013). Improving the performance assessment of government web portals: A proposal using data envelopment analysis (DEA). Information Polity, 18(2), 169-187.

Mohammad, H., Almarabeh, T., & Ali, A. (2009). E-government in Jordan. European Journal of Scientific Research, 35 (2), 188-197.

Montserrat, J. (2010). Local e-Government Bench-learning: Towards a new methodological framework to benchmark electronic services provision and adoption in local public administration

Moon, M. (2002). The Evolution of E-Government among Municipalities: Rhetoric or Reality? Public administration Review, *62*(4), 424-434

Moraru, G. (2010). Anatomy of E-government: Assessment of Municipal E-government Services in Romania. CEU eTD Collection.

Nielsen, J. (1993). Usability engineering. San Francisco, CA: Morgan Kaufmann

Norris, D. (2009). E-government among American Local Governments: Adoption, Impacts, Barriers and Lessons Learned. The International Conference for Administrative Development: Towards Excellence in Public Sector Performance, Riyadh, Saudi Arabia, November 1-4, 2009.

Norris, D. F., & Reddick, C. G. (2013). Local E-Government in the United States: Transformation or Incremental Change? *Public Administration Review*,*73*(1), 165-175.

Peters, R., janssen, M. & Engers, T. (2004). Measuring E-government Impact: Existing Practices and Shortcomings. In Proceedings of the 6th international conference on Electronic commerce, 480-489. Available at http://wing.comp.nus.edu.sg/downloads/keyphraseCorpus/136/136.pdf

Petricek, V., Escher, T., Cox, I., & Margetts, H. (2006). The Web Structure of E-Government - Developing a Methodology for Quantitative Evaluation. IW3C2, WWW 2006, Edinburgh, UK.

Rorissa, A. , Demissie, D. , & Pardo, T. (2008). Toward a Common Framework for Computing e-Government Index. ICEGOV2008, December 1-4, 2008, Cairo, Egypt.

Rorissa, A., Demissie, D., & Pardo, T. (2011). Benchmarking e-government: A comparison of frameworks for computing e-government index and ranking.Government Information Quarterly, 28(3), 354-362.

Salem, F. (2008). Benchmarking the e-Government Bulldozer: Beyond Measuring the Tread Marks. Journal of Measuring Business Excellence, *11*(4), 9-22.

Sandoval-Almazan, R., & Gil-Garcia, J. R. (2012). Are government internet portals evolving towards more interaction, participation, and collaboration? Revisiting the rhetoric of e-government among municipalities. *Government Information Quarterly*, *29*, S72-S81.

Sandoval-Almazan, R. & Gil-García, J. R. (2008a). E-Government portals in Mexico. Electronic Government: Concepts, Methologies, Tools and Aplications, 1(1), 1726–1734.

Sandoval-Almazan, R., & Gil-Garcia, J. R. (2008b). Limitations of Evolutionary Approaches to E-government. Handbook of Research on Public Information Technology, Hershey, IGI Global, 22-30.

Schedler, K. , & Schmidt, B. (2004). Managing the e-government organization. International Public Management Review, *5*(1).

Schellong, A. (2009). EU eGovernment Benchmarking 2010+: General Remarks on the Future of Benchmarking- Digital Government in the EU.

Scowen,G. (2007). Increased Website Popularity through Compliance with Usability Guidelines. Dissertation. Information Science, University of Otago, New Zealand

Signore, O. (2005). A Comprehensive Model for Web Sites Quality. In Proceedings of the Seventh IEEE International Symposium on Web Site Evaluation, Budapest, Hungary. 30-38

Toonders, R. (2010). Dutch Municipality E-government: Explaining E-government Adoption and Development. Master Thesis. Eindhoven University of Technology

United Nations. (2008). UN E-Government Survey 2008: from E-government to Connected Governance. United Nations, New York, USA. Available at: http://unpan1.un.org/intradoc/groups/public/documents/un/unpan028607.pdf

United Nations. (2012). E-Government Survey 2012- E-Government for the People. United Nations, New York, USA. p.39. Available at http://unpan1.un.org/intradoc/groups/public/documents/un/unpan048065.pdf

Vigo, M., & Brajnik, G. (2011). Automatic web accessibility metrics: Where we are and where we can go. Interacting with Computers, *23*(2), 137-155. Available at http://dx.doi.org/10.1016/j.intcom.2011.01.001

West, D. (2008). Improving Technology Utilization in Electronic Government around the World, Brookings, 2008. Available at http://www.brookings.edu/reports/2008/0817_egovernment_west.aspx

Yan, P., & Guo, J. (2010). The Research of Web Usability Design. Chongqing Three Gorges University. IEEE 2010.

Yildiz, M. (2007). E-government Research: Reviewing the Literature, Limitations, and Ways Forward. Government Information Quarterly *24*(2007), 646-665

Zevedeos, R. (2006). Local e-Government: Reconstructing Limassol's municipality (Cyprus) Web Site to Provide Functional and Effective E-services. Master Dissertation. University of Sheffield.