The Trade-off between Usability and Security in the Context of eGovernment: A Mapping Study

Abdulla Alshamsi Keele University School of Computing and Mathematics Staffordshire ST5 5BG a.j.a.alshamsi@keele.ac.uk Nikki Williams Keele University School of Computing and Mathematics Staffordshire ST5 5BG n.k.williams@keele.ac.uk Peter Andras Keele University School of Computing and Mathematics Staffordshire ST5 5BG p.andras@keele.ac.uk

Most governments implement the latest information communication technology (ICT) to improve the online experience of their citizens and businesses. Governments put great effort into providing user-focused services that are usable, secure and accessible by portable and wireless devices (e.g. tablets, smart phones etc.). However, such devices bring with them specific problems of usability and security that affect how users interact with government digital services (GDS). This paper presents a systematic mapping study, investigating the existing problems of usability and security of GDS accessed through smart devices. It aims to uncover what evaluation methods have been used by researchers and investigate how the trade-off between usability and security is assessed in the context of GDS accessed through smart devices. The paper summarises the current knowledge available with regards to this trade-off over the last ten years. The results of the mapping study help identify several research gaps, leading to areas for new research in the domain of usability and security in the context of GDS.

Usability, Security, Trade-off, eGovernment, Digital Services, Systematic Mapping, Smart Devices

1. INTRODUCTION

The advancement of the Internet and the ubiquitous use of computing in everyday life (iPads, tablets, smartphones etc.) have influenced the growth of commercial and governmental electronic services. eGovernment is defined as the use of ICT, particularly web-based applications, to provide faster, easier and more efficient access to government digital services for the public (Huang and Benyoucef, 2014). This revolution in the ubiquitous use of computing allows citizens to interact with government digital services (GDS) at any time and in any place using their smart devices (Huang and Benyoucef, 2014). These smart devices have great mobility in delivering services to citizens and have become a main point of focus for government organisations.

At present, local and national governments in developed and developing countries aim to leverage the power of ubiquitous computing to provide fast, easy, secure and reliable GDS (González Martínez et al., 2011).

Consequently, governments aim to provide services based on users' needs that are usable and secure,

and this is critical to the successful adoption and use of GDS (Baker, 2009).

Usability and security are two related elements that have a significant influence on user communication and engagement with eGovernment, and need to be studied and understood together. Therefore, it is very important to understand usability and security in a government setting in order to provide feedback for designers so that they can develop usable and secure GDS that can be used by a wide range of citizens (Gouscos et al., 2007: Hung, Chang and Kuo, 2013).

This mapping study aims to identify aspects of usability and security in government digital services which have been researched, or where research is lacking.

Systematic mapping studies in software engineering are recommended for research areas where there is a lack of relevant, high-quality, primary study (Kitchenham and Charters, 2007). The process of the mapping study review is presented and described in detail and the findings of the review and answers to the research questions are discussed.

2. RESEARCH METHOD

A considerable amount of research has been published on various subjects within the field of usability and security of GDS. A systematic mapping study is recommended as a structured method to identify research clusters and any possible gaps in this area of research (Kitchenham and Charters, 2007). The mapping study supports the collection and categorisation of all the available studies and literature in this area, with the aim of making it simple to identify the various subtopics, and show where current research is focused. This method is selected because it provides a credible and rational evaluation of studies on the usability and security of GDS and helps identify any gaps in the current research. To achieve these goals a review protocol is developed to reduce the possibility of researcher bias and identify areas where more primary studies need to be carried out. The protocol developed is assessed and reviewed by two external experts in the field to ensure the validity of the protocol and that it meets the requirements stated by Kitchenham (Kitchenham and Charters, 2007). The first steps of the review protocol are creating research questions, identifying the search strategy and the defining the search scope. A search process, based on the research questions, is then conducted. Inclusion and exclusion criteria for studies are developed, designed in the search phase so as to assess the thoroughness of the literature search. A strategy is designed for assessing the quality of the papers collected in the search process. Next, the elements of data to be extracted from the selected literature are determined, in order to address the review questions and synthesise the data. Finally, the strategy to evaluate and analyse the data extracted from the literature is devised. In the following subsections, the detail of the procedure and the strategy followed in conducting the systematic review are described.

2.1 Research questions

Four research questions are formulated to guide the mapping study and to identify research opportunities. Based on the research objectives stated, the mapping study is driven by the following research questions:

- RQ1. What are the existing usability and security problems concerning government digital services accessed by smart devices?
- RQ2. What methods of evaluation have been used to assess the usability and security of government digital services accessed by smart devices?
- RQ3. How is the trade-off between usability and security measured and assessed in the context of government digital services?

 RQ4. What training and policies are available to the public to ensure effective usability and security of government digital services accessed by smart devices?

The findings of the proposed review questions are critically important for evidence based engineering of eGovernment services and contribute to the knowledge of usability and security within the domain of eGovernment.

2.2 Search strategy

The review includes a search strategy developed to utilise publication databases in an efficient way. This search strategy is essential for the search process to avoid including irrelevant search results. The search strategy is designed based on selecting major terms from each research question and using alternative words and synonyms in each search string. This reduces the effect of variance in the terminologies. Boolean "OR" is used to link alternate words and synonyms as well as Boolean "AND" to join major terms, if the databases allow. The search string consists of these main words: "usability" AND "security" AND "services" AND "smartphone" AND "government" AND "evaluation".

The alternate terms are connected through Boolean OR to produce a reference search string for automatic search of databases. Using the outcome from the pilot search activity, the final search strings are derived and used to retrieve the relevant papers. The created search strings are adjusted based on the search criteria of each electronic database.

The search generated 1,600 results from IEEE Xplore, 3,900 papers from ACM Digital Library and 5,700 papers from Google Scholar.

2.3 Search process

The scope of the search focuses on the publication period and source. The search for publications is limited to the period from January 2005 to July 2015 due to the revolution in the introduction of smart devices during this period. Three electronic data sources, or search engines, are used, ACM Digital Library, IEEE Xplore Digital Library and Google Scholar. For each data source, the return results per search are documented and the papers retrieved are manually imported into Mendeley software. At this stage, some irrelevant papers are excluded based on their titles and abstracts, before saving them to the reference manager software. The total number of papers retrieved from all sources is 11,200. The number of papers retrieved is very high for this review, so the most relevant conferences and journals, which have the most published papers in them and are linked to the field of usability and security in the context of GDS, are selected. Two of

the authors have reviewed the validity of the selection of the conferences and journals, and 15 conferences and 13 journals are selected as being most relevant to the review. The selection is based on the Institute for Scientific Information's (ISI) impact factor rating, the number of articles published and whether the authors are cited in other reputable journals and conferences (See Table 1). Through consideration of the selected journals and conferences, the number of papers retrieved is narrowed down to 690. Then, the conferences and journals with fewer papers are excluded from the review, leaving three conferences and three journals with the highest volume of papers included in the final selection. The total number of papers in the three journals and the three conferences is 129 (See Table 2).

The 129 full text papers are read in accordance with the inclusion and exclusion criteria, and the decisions are shown in the next section. After the process of inclusion and exclusion decision-making is complete, the final set of primary studies is reviewed. The total number of primary papers selected is 74. These papers are considered the most relevant to this review. Mendeley software (www.mendeley.com) is used as a reference manager tool for managing and storing the papers retrieved from the search engines, and to classify the retrieved papers.

2.4 Inclusion and exclusion criteria

It is imperative that any mapping study contains comprehensive inclusion and exclusion criteria, in order to highlight only those primary studies that provide evidence related to the research questions. The inclusion and exclusion criteria of this review study are developed based on the research questions and are used to ensure that the results obtained are reliable and categorise studies correctly according to the guidelines set by Kitchenham (Kitchenham and Charters, 2007).

No	Conferences		
1	Computer Human Interaction CHI (ACM)		
2	International conference on Digital Government Research (ACM)		
3	European conference on Information systems ECIS		
4	Australian Conference on Information systems		
5	International Conference on Availability Reliability and security ARES (IEEE)		
6	Symposium on Usable Privacy and Security SOUPS (ACM)		
7	American Conference on Information Systems		
8	International Conference on eDemocracy & eGovernment (IEEE)		
9	Symposium on Security and Privacy (IEEE)		

10	International Conference on Information Assurance and Security (IAS)
11	Symposium on Computer and Communications Security (CCS) (ACM)
12	International Conference on Security of Information and Networks
13	ACM Conference on Data and Application Security and Privacy (CODASPY)
14	ACM International Joint Conference on Pervasive and Ubiquitous Computing
15	Mobile HCI conference (MobileHCI)
No	Journals
1	European Journal of Information Systems
2	International Journal of Electronic Governance
3	Journal of Information and Software Technology
4	Journal of Systems and Software
5	International Journal of Human-Computer Interaction
6	Personal and Ubiquitous Computing International Journal
7	Security & Privacy Journal
8	Journal of Transactions on Information Forensics and Security
9	Transactions on Consumer Electronics
10	Journal of Engineering and Technology
11	International Journal of Mobile Human Computer Interaction (IJMHCI)
12	Journal of Usability Study (ACM)
13	Journal of Government Information Quarterly

Table 1: Initial conferences and journals selection

Selected Conferences	No of Papers
Computer Human Interaction CHI (ACM)	27
International Conference on Digital Government Research (ACM)	5
Mobile HCI Conference (MobileHCI) (ACM)	10
Selected Journals	
International Journal of Human-Computer Interaction	9
Security & Privacy Journal (IEEE)	13
Journal of Government Information Quarterly	10

Table 2: Final selection of conferences and journals

The following inclusion criteria are applied to all 129 papers obtained from the search process. The reviewed papers have to meet at least one of the

following inclusion criteria to be included in the review:

- I1: Papers discuss and describe any usability and security problems in the context of GDS.
 - I2: Papers report any usability and security problems of smart devices (smart phones).
 - I3: Papers report usability and security evaluation methods and guidelines in the context of GDS.
 - I4: Papers describe the methods used to evaluate usability and security of smart devices.
 - I5: Papers discuss the trade-off between usability and security and how it is measured.
 - I6: Papers discuss training guidelines, policies and security awareness in the context of GDS.

Papers that meet any of the exclusion criteria shown below are discounted from the review:

- E1: Papers that do not have a specific problem to investigate, search process, or data analysis process.
- E2: Papers focusing on the technical side (cryptography, coding etc.).
- E3: Non peer-reviewed literature.
- E4: Tutorial summary, panel discussion, technical report, book chapter or PowerPoint slides.
- E5: Papers not written in English.

After the process of inclusion and exclusion, the validity of the inclusion and exclusion process is checked. Each author randomly selects 10 papers and applies the inclusion and exclusion process, to verify whether the included and excluded papers are properly reviewed and classified. Then a meeting between the authors leads to a consensus on any disagreements about the included or excluded papers.

2.5 Quality assessment

To evaluate the quality of the papers obtained from the inclusion and exclusion phase, 11 quality assessment criteria are applied based on the recommendations of Dyba (Dings and Dyba, 2008). These 11 criteria questions are not listed in this paper and can be found in (Dings and Dyba, 2008). The criteria are used to rate the quality of reviewed papers and ensure the selected papers contribute effectively to the review study. The possible outcomes of applying the criteria are "Agree", "Partially Agree" or "Not Agree". The criterion in Q1

is used as a basis for accepting or rejecting a study. In a case where either Q1, or both Q2 and Q3, are scored Not Agree, no further quality assessment is done by the reviewer. The scoring procedure is based on attributing values to the scores: Agree = 1, Partially Agree = 0.5, Not Agree = 0. The validity of this quality assessment process is based on the suggestions of Kitchenham (Kitchenham and Charters, 2007). Accordingly, one researcher extracts the data and another check the extracted data. In this way, all the retrieved papers are assessed based on the quality criteria, by all the authors independently, and the quality assessments compared. This validity check helps resolve any scoring differences in the assessments.

Of the 74 papers assessed for quality, 71 had initially been included based on the screening criterion. All disagreements about the remaining three papers have been resolved by discussion between the three researchers, who finally agreed to include the three remaining papers in the review.

2.6 Data extraction

The 74 primary studies are read in detail to extract the data required in order to answer the review questions. A data extraction form is developed to give reliable and accurate extraction of the relevant data from each paper. Table 3 shows the data extracted, some specifically focusing on the research questions and other data required for later analysis, irrespective of the research questions.

Two of the authors have reviewed and checked the consistency of the data extraction process, each selecting 10 per cent of the primary studies and extracting the data for a second time, then comparing their data sheet with the primary reviewer's data sheet. Any differences found in the data extraction between the primary reviewer and the other reviewers are reconciled and resolved collaboratively (Kitchenham and Charters, 2007). The extracted data is documented and kept in a Mendeley file and Excel spreadsheet for future analysis.

Code	Field /Data	Related Research Question
D1	Extraction date	Documentation
D2	Author name	Documentation
D3	Title of publication	Documentation
D4	Publication source Documer	
D5	Year of publication	Documentation
D6	Type of publication	Documentation
D7	Aims and objectives	Documentation
D8	Research question	Documentation
D9	Security and usability problems	RQ1

D10	Security elements	RQ1,RQ3
D11	Usability attributes.	RQ1,RQ3
D12	GDS usability and security problems	RQ1
D13	Smart devices usability and security problems RQ1	
D14	Evaluation method used (name and short description)	RQ2,RQ3, RQ4
D15	Guidelines, policy, training	RQ4
D16	Statistical data used for analysis	Documentation
D17	Domain or context	RQ1,RQ2, RQ3

Table 3: Data extraction form

The papers are categorised into 3 categories and each is divided into sub-categories, in order to minimise possible misrepresentation of the data extracted. This classification of the primary studies is based on defining a set of possible answers for each research question.

2.6.1 Usability and Security Issues

Papers are categorised based on four main categories in regard to research question one (RQ1), (C1-A, C1-B, C1-C and C1-D) and each category is further classified into sub-categories. The following classifications are used to describe any paper that discusses or reports any usability or security issues in eGovernment settings:

- (I) C1-A: GDS usability problems (efficiency, satisfaction, learnability, memorability, errors, other).
- (II) C1-B: Smart device usability problems (device context, connectivity, screen size, display resolution, processing, capability or power, data entry method, other).
- (III) C1-C: Smart device security issues (authentication, access control, availability, data or message security, non-repudiation, secure storage).
- (IV) C1-D: GDS security issues (authentication, availability, confidentiality, integrity, non-repudiation, other).

2.6.2 Assessment of Usability and Security

The retrieved papers are grouped into four main categories based on research questions two and three (RQ2 & RQ3), (C2-A, C2-B, C2-C and C2-D) and each category has sub-categories:

- (I) C2-A: The focus of the assessment:
 - (a) Usability assessment.
 - (b) Security assessment.
 - (c) Trade-off between security and usability.
- (II) C2-B: Usability and the method of security evaluation:

- (a) Testing (if it involves an evaluator observing participants interacting to determine problems).
- (b) Inspection (if it involves an expert evaluator using a set of criteria to identify potential usability problems e.g., heuristic evaluation).
- (c) Inquiry (if it presents a method that collects participants' preferences or feelings from interviews or questionnaires).
- (d) Analytical modelling (if it presents an engineering method that employs various kinds of models).
- (III) C2-C: The type of study used to evaluate security and usability:
 - (a) Controlled experiment.
 - (b) Interview.
 - (c) Focus group.
 - (d) Survey.
 - (e) Case study.
- (IV) C2-D: Domain or context:
 - (a) eGovernment.
 - (b) Academic.
 - (c) Industrial.
 - (d) Medical.
 - (e) eCommerce.
 - (f) Other.

2.6.3 Training and Policies

Papers relevant to research question four (RQ4), are classified based on the following categories:

- (I) C3-A: Types of training:
 - (a) Social engineering training.
 - (b) Security awareness.
 - (c) Population awareness.
- (II) C3-B: Types of existing policies:
 - (a) Security policy.
 - (b) Acceptable use policy.
 - (c) Legalisation or regulation policy.
 - (d) Other.

2.6.4 Data Synthesis

The data extraction and data classification processes are completed according to the designed protocol and all have been assessed. Extracted data that is redundant is removed and the quality of the data rechecked. This means that all the checked data is considered suitably qualitative and valid, and therefore applicable to answer the research questions. The aim of the mapping study is to answer the research questions with effective and reliable data. The data synthesis activities are used to summarise the results of the primary studies. The extracted data is studied manually and descriptive synthesis conducted, to show the results in a tabular form. Descriptive statistics are applied to analyse and summarise the data.

3. RESULTS AND ANALYSIS

The extracted data is analysed and summarised in a structured way that assists in finding possible answers to the stated research questions. The next section provides an overview of the selected primary studies and the extracted information in regard to the research questions.

3.1 Results overview

After the filtering phases described, 74 primary studies from 3 journals and 3 conferences (across multiple years) are used for data analysis and answering the research questions. The highest volume of conference papers were published in 2009 and most of the journal papers were published in 2014, due to the release of various multi-touch interface smart devices in 2007, such as the Apple smartphone.

Of the 74 papers, 57 per cent are conference papers and 43 per cent are journal papers. This indicates that the majority of papers collected come from conference sources rather than journal publications which could be due to the longer time required by authors to publish work in journals, rather than conferences.

3.2 Research question results

Research question one (RQ1) asks, "What are the existing usability and security problems concerning government digital services (GDS) accessed by smart devices?" To find data to answer this question, the data from D9, D10, D11, D12 and D13 are analysed using the data extraction form (See Table 3). The analysis of selected papers shows that 40 papers contain relevant content and discuss usability and security problems concerning eGovernment platforms.

The answer to question one (RQ1) is divided into four sections to give a clear view of the papers reviewed.

GDS usability problems: Usability is considered an important feature that affects user interaction with government digital services. This review identifies five attributes that can affect the overall usability of GDS, based on the Jacob Nielsen usability model (De Jong and Lentz, 2006): efficiency, learnability, satisfaction, memorability and errors. Based on the review question and categorisation of the selected papers, most of the studies explain and address usability issues in an eGovernment setting based on these five attributes, as shown in Table 4. Learnability and efficiency are the most addressed attributes in the selected papers, which each present in 11 per cent of the papers reviewed. Error is the least addressed usability attribute, only present in three papers. The other two attributes are both present in 8 per cent of the papers. Among the

studies analysed, three do not clearly specify which usability attributes are addressed. These studies (Kotamraju et al., 2012; Olalere and Lazar, 2011; Kokini et al., 2012) concentrate on content analysis without stating any usability attributes. From Table 4 it can be seen that usability attributes are considered by most of the studies selected, in regard to RQ1.

Smart device usability problems: Papers related to smart device usability problems are analysed and examined based on six elements, identified in Donker et al. (2010), device context, connectivity, screen size, display resolution, processing capability and data entry method. The papers retrieved in regard to RQ1 are categorised by at least one of these elements. In total 12 papers discuss the usability of smart devices. Some of the papers investigate more than one element of usability for smart devices and so are counted twice or more. The greatest volume of work focuses on device context (including the location, identities of nearby people, time, temperature, colour, weight). Seven selected papers discuss this element. The elements of connectivity and data entry methods are almost in second ranked, with three and four papers respectively. Two papers discuss display resolution, screen size or processing capability. Generally, the selected publications considered in this section, do not focus on the field of GDS.

The analysis of the papers indicates that a general usability evaluation for smartphones has been carried out, but has not focused on the problems generated by these devices in an eGovernment setting. This could be due to the unique features of smartphones, and that current smartphone platforms differ considerably in terms of functionality provided and security features.

GDS security problems: Security problems are broken down into subcategories to provide a broad view of the problems of security in a GDS setting and make evaluation easier. This review study classifies the papers in this section based on six well-known security elements: authentication, availability, confidentiality, integrity, non-repudiation and security storage. As shown in Table 4, there are only a few papers that discuss security in a GDS setting. Similar to the previous section, the total number of papers shown in Table 4 is greater than the number of papers included in the review, due to some of the studies covering more than one element of security and being counted twice or more. The most researched elements of security of GDS are authentication and integrity, with six and three papers respectively. The remaining elements are addressed by two papers each. Few papers in this review address the security of GDS, due to the complex nature of government systems and the availability of information about it for researchers.

Smart device security problems: This section discusses the number of reviewed papers

addressing security issues of smart devices. The review of this section is based on (Benantar, 2006) categorisation of the security of smart devices. Five attributes are identified that relate to the security of smart devices: authentication, access control, availability, data and message security and nonrepudiation (See Table 4 – for the sake of brevity for all tables, papers are referenced by numbers that are included in the references section). The papers relevant to this section represent 15 per cent of the total papers reviewed. Authentication is the element addressed most frequently, discussed by ten papers. Access control and non-repudiation are addressed least frequently in the reviewed papers for this section, with one paper being noted for each. There appear to be few papers addressing the security of smart devices in the setting of eGovernment.

The main findings for this question are summarised below.

Problem	No of Papers	Paper's Reference No
GDS usability problems	13	
1) Efficiency	8	8, 23, 28, 29, 36, 37, 57, 65,
2) Satisfaction	6	1, 28, 29, 36, 37, 57
3) Learnability	8	1, 8, 20, 23, 28, 29, 36, 65
4) Memorability	6	1, 8, 20, 29, 36, 65
5) Errors	3	20, 29, 36
6) Other	3	44, 45, 56
Smart devices usability problems	12	
1) Device context	7	4, 24, 42, 46, 53, 71, 76
2) Connectivity	4	49, 55, 70, 71
3) Screen size	2	62, 64
4) Display resolution	2	62, 64
5 Processing capability	2	42, 49
6) Data entry method	3	62, 64, 71
GDS security problems	7	
1) Authentication	6	2, 38, 41, 57,67, 70
2) Availability	2	2, 57
3) Confidentiality	2	18, 57
4) Integrity	3	18, 57, 70
5) Non-repudiation	2	18, 57
6) Secure storage	2	18, 57
Smart device security problems	11	
1) Authentication	10	14, 16, 21, 32, 39,54, 64, 66, 70, 74
2) Access control	1	21

3)	Availability	2	21, 32
4)	Data and message security	3	18, 54, 70
5)	Non-repudiation	1	21

Table 4: Papers addressing research question one

To conclude, the reviewed papers indicate the importance of usability and security in an eGovernment setting and the effects on users' attitudes, perceptions and interactions. The results from the selected papers related to this question indicate that the set of reviewed studies do not address the usability and security problems that may occur when services are accessed through smart devices. In addition, there is no evidence in the reviewed papers that they have considered the assessment of the problems of usability or security in an eGovernment setting, using smart devices. It can be seen from the analysis that most of these attributes have been assessed independently without considering the smart devices' requirements for GDS.

The main concerns of the reviewed studies relating to usability issues in an eGovernment context are efficiency, learnability, satisfaction and memorability. These attributes are measured and assessed in a GDS setting by most of the papers. The results show that the element of authentication in the security of GDS and smart devices is the most studied element in the papers reviewed. Finally, the analysis of the reviewed papers related to research question one shows that comprehensive assessment of security and usability is missing in the context of GDS accessed by smart devices.

RQ2 asks, "What methods of evaluation have been used to assess the usability and security of government digital services accessed by smart devices?" D14 and D17 are analysed using the data extraction sheet and summarised in Table 3. About 33 papers relate to this question out of all the papers reviewed (See Tables 5, 6 and 7). The question has two parts that need to be answered in terms of the evaluation methods used. The first part is the usability evaluation methods used to assess the usability of GDS accessed by smart devices. The second part is the evaluation methods that measure the security of GDS accessed by smart devices. The results of the two parts of the question help identify the most widely used methods in the context of GDS.

Usability evaluation methods: The analysis of the reviewed papers shows that a wide range of usability evaluation methods (UEMs) are used to improve the usability of GDS, by measuring user attitudes, perceptions and interactions with eGovernment systems. Most of these methods are used to assess the problems of GDS, and the findings vary widely from one evaluator to another for various reasons, such as the evaluator's skill or the method not being appropriately applied. The analysis of the reviewed

papers reveals that there are various classes of UEMs recognised and used in the context of eGovernment.

These methods are grouped into classes such as usability inspection methods, testing, usability inquiry analytical modelling. Testing, inspection and usability inquiry are used for formative and summative purposes in software engineering (Egelman, et al., 2008). Several of the methods tested and used come under one of the above categories, based on their attributes. The methods found in the reviewed papers that relate to this section are shown in Tables 5, 6 and 7. The analysis of the selected papers shows that the method most used to evaluate the usability of GDS is the "thinking-aloud protocol", which appears in 10 papers (See Table 5).

This method is used at various levels of the software development life cycle and is considered to be cost-effective. An evaluator asks a participant to express his or her thoughts, feelings and opinions whilst interacting with the system. The second UEM used, which is well recognised by experts in the field, is "heuristic evaluation". This is a usability inspection method, where an expert identifies violations of the heuristic. The method is very popular in comparison to other types of usability inspection and expends fewer resources (Egelman, et al., 2008).

Usability Testing	No of Papers	Paper's Reference
Coaching method	1	28
Thinking-aloud protocol	10	1, 13, 14, 16, 33, 34, 39, 52, 63, 71,
Question-asking protocol	1	1
Teaching method	2	3, 42
Performance measurement	1	44
Log file analysis	1	9
Retrospective testing	4	24, 28, 52, 72
Remote testing	3	6, 13, 59
System usability scale (SUS)	2	62, 68
Metaphor of human	1	34
Collaboration critique method	1	7

Table 5: Usability testing methods used

Eight papers use this method and evaluate eGovernment portals (see Table 7). The remaining methods are not intensively used to examine the eGovernment setting because these techniques require special resources and the data type collected (quantitative/ qualitative) (Egelman, et al., 2008). The third class of usability evaluation method used and identified in the reviewed papers is the inquiry method. The usability inquiry method obtains information from the evaluator observing the interaction of the user with the system in real-time.

Usability Inquiry	No of Papers	Paper's Reference
Questionnaires	6	1, 26,27, 37, 42, 49
Interviews	6	16, 38, 45, 53, 54, 71
Field study	3	7, 24, 74
User feedback	5	3, 6, 16, 39, 59
Surveys	7	15, 28, 29, 54, 57, 63, 65
Focus groups	4	14, 41, 52, 53
Self-reporting logs	5	8, 48, 51, 52, 58
Case study	7	2, 17, 18, 47, 49, 53, 59
Analytical modelling	3	13, 19, 72
Simulation	5	19, 39, 46, 52, 59
Controlled experiment	12	9, 24, 25, 27, 35, 42, 44, 48, 62, 66, 70, 71
Screen snapshot	1	48

Table 6: Usability inquiry methods used

The most popular usability inquiry methods recognised in the reviewed papers are case studies, surveys, questionnaires, interviews and field studies, alongside the many other techniques shown in Table 6. The aim of these methods is to collect subjective user impressions, preferences and opinions about the characteristics of the user interface. These methods can be used by testers to gather additional data after the implementation of the system. The analysis of the related papers in this section shows that some papers describe the authors' own experiments, while some evaluate other studies. The analysis of the reviewed papers indicates no evidence that these evaluation methods have been used to assess the usability of GDS in parallel with smart device needs. There is no paper that discusses whether eGovernment services are usable when browsed and accessed by smart devices (e.g. smartphones, iPads).

Security evaluation methods: The analysis of the reviewed papers finds that papers relating to security can be divided into two categories. The first reports on technical security threats and the second report from a non-technical viewpoint. The papers reviewed focus on threat analysis of eGovernment services and risk assessment. These assessments are followed by guidelines and recommendations proposed by the authors, to mitigate the security risks identified in the analysed system. The total number of papers reviewed describing evaluation methods of security is 11.

Usability Inspection	No of Papers	Paper's Reference	
Heuristic evaluation	8	5, 23, 36, 47, 56, 64, 65, 76	

Table 7: Usability inspection methods used

There is no clearly favoured method that can be identified from the reviewed papers. Some propose

security evaluation methods following HCI guidelines and the use of security standards such as ISO/IEC 27002 and ISO/IEC 27001 for evaluating the security of systems. To conclude, it is difficult to identify any framework or security model that considers the eGovernment security requirements needed when being accessed by smart phones.

Research question three (RQ3) is. "How is the tradeoff between usability and security measured and assessed in the context of government digital services?" Analysis of D10, D11, D14 and D17 data extraction forms; reveals that 32 papers focus on usability assessment and 17 papers discuss security assessment (see Table 8). Furthermore, trade-offs in the domain of usability and security in a government setting are not addressed by the reviewed papers: only seven papers focus on settings such as eHealth and eBanking domains. The papers are too limited in scope to specifically address how the balance between usability and security is deployed in an eGovernment context. The main focus of the selected papers is on the trade-off between usability and security of passwords and logins, because passwords and logins are considered the most vulnerable aspects of a secure system.

However, in order to achieve a balance between usability and security in an eGovernment setting, it is obvious that a new framework or approach is necessary to address the specific needs of the eGovernment domain. Therefore, a newly focused assessment of trade-offs should be developed to meet the requirements of usability and security in the context of GDS. To summarise, the retrieved papers acknowledge the presence of the trade-off as evident, but actual measurement and metrics do not appear in the reviewed papers. Suggestions of how the trade-off between usability and security can be managed are provided, however there is a distinct lack of direct assessment of usability and security trade-offs in the context of eGovernment.

Paper's Focus	No of paper	Paper's Reference
Usability	32	1, 3, 5, 6, 7, 8, 16,20, 23, 24, 26, 28, 29,33, 34,40, 42, 47,50,51, 52, 56, 57, 58, 59, 62, 65, 68, 71, 72,73, 74
Security	17	2, 11, 14, 17, 18, 25, 32, 39, 40, 41, 59, 60, 63, 66, 70, 74, 76
Trade-off between usability & security	7	9, 12, 31, 61, 67, 69, 75

Table 8: Matter of assessment

RQ 4 asks, "What training and policies are available to the public to ensure effective usability and security of government digital services accessed by smart devices?" The data of D15 and D14 are analysed to identify any policies or training provided to ensure the

effectiveness of usability and security in the context of GDS. Seven reviewed papers relate to this question. The retrieved papers are classified based on the categories shown in Section 2.6.3 recommended previously in order to answer this question. The retrieved papers' main focus is on general legalisation and regulation policy of security and usability. Some are classified by security policy elements, such as making suggestions to users about the design of passwords and the design of security questions. Some papers provide recommendations about user security education and how to enhance the users' understanding of security. However, the papers related to this question are insufficient to provide any evidence about policies and training in the domain of GDS. To conclude, it is apparent from the reviewed papers that such policies and training in the context of eGovernment have not been well studied and addressed.

4. REVIEW LIMITATIONS AND THREATS TO VALIDITY

The main threat to the validity of the review is the limitation of the conference and journal selection. The 74 papers, retrieved from three conferences and three journals (editions/volumes over multiple years, 2005 – 2015), could exclude a considerable number of papers relevant to the review study. This is partly due to time concerns and partly due to the high number of papers retrieved in the automatic search. In addition, inaccuracy and bias in the retrieved papers due to the automatic search is a possible study limitation. Conducting manual searches and comparing them with the automatic searches, mitigates this bias and ensures that the search string for the automatic search retrieves all the relevant papers. Bias can come from the inclusion and exclusion process, which has an effect on the process of paper selection. Having two additional authors check the included and excluded papers mitigates this bias. Another important threat to validity is inaccuracy in the data extraction. The data extraction process is somewhat complicated, as some papers do not clearly report the methods, or what type of setting, is used. Finally, the reviewer's lack of experience in designing protocol for a mapping study is a threat to the validity of the study that should be considered. The guidelines provided by (Kitchenham and Charters, 2007), and advice from experienced practitioners of systematic reviews and mapping studies, helps reduce and avoid some of these threats.

5. CONCLUSION

This paper aims to answer four research questions with respect to usability and security in an eGovernment setting. A systematic analysis of 74

publications has been cond'ucted in order to answer the study questions.

The results highlight that very few of the papers reviewed include any kind of usability or security assessment in a GDS setting. The mapping study highlights that the majority of papers that look into the trade-off between usability and security are restricted to examination of login methods in various contexts. Therefore, the need for empirical research focusing on an eGovernment setting, in terms of evaluating methods and the trade-off between usability and security, is clearly identified.

The reviewed papers confirm that usability and security of GDS accessed with smart devices is not being addressed and there is scope for more work in this area. Future work should focus on aspects of usability and security in an eGovernment setting and aim to create an integrated framework for the assessment of these, in order to achieve an optimal and security. trade-off between usability Furthermore, this needs to be complemented by more research into awareness and education policies related to the usability and security trade-off in the context of eGovernment services accessed through smart devices.

6. REFERENCES

- Ahmad, N., Shoaib, U. and Prinetto, P. (2015) Usability of Online Assistance From Semiliterate Users' Perspective. Int. J. Hum. Comput. Interact., 31(1). 55–64. (Ref 1)
- Aichholzer, G. and Strauß, S. (2009) Understanding a complex innovation process: identity management in Austrian e-government. 10th Annu. Int. Conf. Digit. Gov. Res. (dg.o 2009). 230–239. (Ref 2)
- Akers, D., Simpson, M., Jeffries, R. and Winograd, T. (2009) Undo and erase events as indicators of usability problems. Proc. 27th Int. Conf. Hum. factors Comput. Syst. CHI '09. 659. (Ref 3)
- Alberto, P. (2010) User-centric mobile services: context provisioning and user profiling. 11th Annu. Int. Conf. Digit. Gov. Res. (dg.o 2010). 122–130. (Ref 4)
- Alonso-Ríos, D., Mosqueira-Rey, E. and Moret-Bonillo, V. (2014) A Taxonomy-Based Usability Study of an Intelligent Speed Adaptation Device. Int. J. Hum. Comput. Interact., 30(7). 585–603. (Ref 5)
- Andreasen, M. S., Nielsen, H. V., Schrøder, S. O. and Stage, J. (2007) What happened to remote usability testing? An empirical study of three methods. Proc. 25th SIGCHI Conf. Hum. Factors Comput. Syst. 1405–1414. (Ref 6)

- Babaian, T., Lucas, W. and Oja, M-K. (2012) Evaluating the collaborative critique method. Proc. SIGCHI Conf. Hum. Factors Comput. Syst. 2137– 2164. (Ref 7)
- Baker, D. L. (2009) Advancing E-Government performance in the United States through enhanced usability benchmarks. Gov. Inf. Q., 26(1). 82–88. (Ref 8)
- Ben-Asher, N., Meyer, J., Parmet, Y., Moeller, S. and Englert, R. (2009) Security and usability research using a microworld environment. Proc. 11th Int. Conf. Human-Computer Interact. with Mob. Devices Serv. Mobile HCI '09 1. (Ref 9)
- Benantar, M., 2006. Introduction to Identity-Management models. In Access control systems Security identity management and trust models. pp. 40–72 (Ref 10)
- Bhattacharya, P., Yang, L., Guo, M., Qian, K. and Yang, M. (2014) Learning mobile security . IEEE Secur, 12(1). 69–72. (Ref 11)
- Braz, C., Seffah, A. and Raihi, D. M. (2007) Designing a Trade-Off Between Usability and Security: A Metrics Based-Model. Lect. Notes Comput. Sci., 4663. 114–126. (Ref 12)
- Bruun, A., Gull, P., Hofmeister, L. and Stage, J. (2009) Let Your Users Do the Testing: A Comparison of Three Remote Asynchronous Usability Testing Methods. SIGCHI Conf. Hum. Factors Comput. Syst. 1619–1628. (Ref 13)
- Chan, Y-Y. and Wei, V. K. (2009) Teaching for Conceptual Change in Security Awareness: A Case Study in Higher Education. IEEE Secur. Priv. Mag., 7(1). 68–71. (Ref 14)
- Chilana, P. K., Ko, A. J., Wobbrock, J. O., Grossman, T. and Fitzmaurice, G. (2011) Post-Deployment Usability: A Survey of Current Practices. CHI Conf. Hum. Factors Comput. Syst. 2243–2246. (Ref 15)
- Chilana, P. K., Wobbrock, J. O. and Ko, A. J. (2010) Understanding usability practices in complex domains. SIGCHI Conf. Hum. Factors Comput. Syst. 2337–2346. (Ref 16)
- Choi, J., Chun, S. A. and Cho, J-W. (2014) Smart SecureGov. Proc. 15th Annu. Int. Conf. Digit. Gov. Res. dg.o '14. 91–99. (Ref 17)
- Choi, J., Chun, S. A., Kim, D. H. and Keromytis, A. (2013) SecureGov: Secure Data Sharing for Government Services. Proc. 14th Annu. Int. Conf. Digit. Gov. Res. 127–135. (Ref 18)
- Chowdhury, S., Poet, R. and Mackenzie, L. (2014) Passhint: memorable and secure authentication. Sigchi. 2917–2926. (Ref 19)
- de Jong, M. and Lentz, L. (2006) evaluation of municipal Web sites: Development and use of an

- expert-focused evaluation tool. Gov. Inf. Q., 23. 191–206. (Ref 20)
- Dhamija, L. and Dusseault, R. (2008) The Seven Flaws of Identity Management. IEEE Security and Privacy, 6(2). 24–29. (Ref 21)
- Dings, T. & Dyba, T., 2008. Empirical studies of agile software development: A systematic review., 50, pp.833–859. (Ref 22)
- Donker-Kuijer, M. W., de Jong, M. and Lentz L. (2010) Usable guidelines for usable websites? An analysis of five e-government heuristics. Gov. Inf. Q.,27(3).254–263.(Ref 23)
- Duh, H. B-L., Tan, G. C. B. and Chen, V. H-H. (2006) Usability evaluation for mobile device: a comparison of laboratory and field tests. Proc. MobileHCI 2006. 181–186. (Ref 24)
- Egelman, S., Cranor, L. F. and Hong, J. (2008) You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. Proceeding twenty-sixth Annu. CHI Conf. Hum. factors Comput. Syst. - CHI '08. 1065. (Ref 25)
- Følstad, A., Law, E. and Hornbæk, K. (2012) Analysis in practical usability evaluation: a survey study. Proc. 2012 ACM Annu. Conf. Hum. Factors Comput. Syst. 2127–2136. (Ref 26)
- Frandsen-Thorlacius, O., Hornbæk, K., Hertzum, M. and Clemmensen, T. (2009) Non-Universal Usability? A Survey of How Usability is Understood by Chinese and Danish Users. Proc. CHI 2009. 41–50. (Ref 27)
- González Martínez, S., Luna-Reyes, L. F., Luna D. E., Gil-Garcia, J. R. and Sandoval-Almazán, R. (2011) Comparing usability of government web portals during governor change of terms. Proc. 12th Annu. Int. Digit. Gov. Res. Conf. Digit. Gov. Innov. Challenging Times dg.o '11. 327. (Ref 28)
- Gouscos, D., Kalikakis, M., Lega,I M. and Papadopoulou, S. (2007) A general model of performance and quality for one-stop e-Government service offerings. Gov. Inf. Q., 24(2007). 860–885. (Ref 29)
- Grossman, T., Fitzmaurice, G. and Attar, R. (2009) A Survey of Software Learnability: Metrics, Methodologies and Guidelines. Proc. 27th Int. Conf. Hum. factors Comput. Syst. 649–658. (Ref 30)
- Gutmann, P. and Grigg, I. (2005) Security usability. IEEE Secur. Priv., 3(4). 56–58. (Ref 31)
- Herley, C. (2014) More is Not the Answer. IEEE Security and Privacy 12(1). 14-19. (Ref 32)
- Hertzum, M. (2010) Images of Usability. Int. J., 26(6). 567–600. (Ref 33)

- Hornbæk, K. and Frokjær, E. (2005) Comparing usability problems and redesign proposals as input to practical systems development. CHI 2005 Technol. Safety, Community Conf. Proc. Conf. Hum. Factors Comput. Syst. 391–400. (Ref 34)
- Hornbæk, K. and Frøkjær, E. (2008) Making use of business goals in usability evaluation. Proceeding twenty-sixth Annu. CHI Conf. Hum. factors Comput. Syst. CHI '08. 903. (Ref 35)
- Huang, Z. and Benyoucef, M. (2014) Usability and credibility of e-government websites. Gov. Inf. Q., 31(4). 584–595. (Ref 36)
- Hung, S-Y., Chang, C-M. and Kuo, S-R. (2013) User acceptance of mobile e-government services: An empirical study. Gov. Inf. Q., 30(1). 33–44. (Ref 37)
- Inglesant, P. G. and Sasse, M. A. (2010) The true cost of unusable password policies. Proc. 28th Int. Conf. Hum. factors Comput. Syst. CHI '10. 383. (Ref 38)
- Jakobsson, M., Stolterman, E., Wetzel, S. and Yang, L. (2008) Love and authentication. Proceeding twenty-sixth Annu. CHI Conf. Hum. factors Comput. Syst. CHI '08. 197. (Ref 39)
- Johnson, E. M. and Willey, N. D. (2011) Usability Failures and Healthcare Data Hemorrhages. IEEE Secur. Priv. Mag.(March/April. 35–42. (Ref 40)
- Just, M. (2005) Designing and evaluating challengequestion systems. IEEE Secur. Priv. Mag., 2(5). 1158–1163. (Ref 41).
- Keijzers, J., Den Ouden, E. and Lu, Y. (2008) Usability Benchmark Study of Commercially Available Smart Phones: Cell Phone Type Platform, PDA Type Platform and PC Type Platform. Methods. 265–272. (Ref 42)
- Kitchenham, B. & Charters, S., 2007. Guidelines for performing Systematic Literature Reviews in Software Engineering. Engineering, 2, p.1051. (Ref 43)
- Kokini, C. M., Lee, S., Koubek, R. J. and Moon S. K. (2012) Considering Context: The Role of Mental Workload and Operator Control in Users' Perceptions of Usability. Int. J. Hum. Comput. Interact., 28(9). 543–559. (Ref 44)
- Kotamraju, N. P. and van der Geest, T. M. (2012) The tension between user-centred design and egovernment services. Behav. Inf. Technol., 31(3). 261–273. (Ref 45)
- Krauß, M. and Krannich, D. (2006) Ripcord: Rapid Interface Prototyping for Cordless Devices. Proc. MobileHCI 2006. 187–190. (Ref 46)
- Kyung, K-U., Kwon, D-S. and Yang, G-H. (2006) A novel interactive mouse system for holistic haptic

- display in a human-computer interface. Int. J. Hum. Comput. Interact., 20(3). 247–270. (Ref 47)
- Leon, P. G., Ur, B., Balebako, R., Cranor, L. F., Shay, R. and Wang, Y. (2011) Why Johnny Can't Opt Out: A Usability Evaluation of Tools to Limit Online Behavioral Advertising. Proc. SIGCHI Conf. Hum. Factors Comput. Syst. 589–598. (Ref 48)
- Lesemann, E., Woletz, N. and Koerber, S. (2007) Combining methods to evaluate mobile usability. Proc. 9th Int. Conf. Hum. Comput. Interact. with Mob. devices Serv. - MobileHCI '07. 444–447. (Ref 49)
- Lewis, J. R. (2014) Usability: Lessons Learned... and Yet to Be Learned. Int. J. Hum. Comput. Interact., 30(9). 663–684. (Ref 50)
- Lindgaard, G. (2007) Usability Testing: What Have We Overlooked ?. Proc. 2007 ACM Annu. Conf. Hum. Factors Comput. Syst. CHI '07. 1415–1424. (Ref 51)
- Mankoff, J., Fait, H. and Tran, T. (2005) Is Your Web Page Accessible ? A Comparative Study of Methods for Assessing Web Page Accessibility for the Blind. Chi 2005. 41–50. (Ref 52)
- Markova M., Aula A., Vainio T., Wigelius H. and Kulju M. (2007) MoBiS-Q: a tool for evaluating the success of mobile business services. Proc. 9th Int. Conf. Hum. Comput. Interact. with Mob. devices Serv. 238–245. (Ref 53)
- Muslukhov, I., Boshmaf, Y., Kuo, C., Lester, J. and Beznosov, K. (2013) Know Your Enemy: The Risk of Unauthorized Access in Smartphones by Insiders. Proc. 15th Int. Conf. Human-computer Interact. with Mob. devices Serv. (MobileHCI '13). 271–280. (Ref 54)
- Nylander, S., Lundquist, T. and Brännström, A. (2009) At home and with computer access: why and where people use cell phones to access the internet. Proc. SIGCHI Conf. Hum. Factors Comput. Syst. 1639–1642. (Ref 55)
- Olalere, A. and Lazar, J. (2011) Accessibility of U.S. federal government home pages: Section 508 compliance and site accessibility statements. Gov. Inf. Q., 28(3). 303–309. (Ref 56)
- Papadomichelaki, X. and Mentzas, G. (2012) e-GovQual: A multiple-item scale for assessing e-government service quality. Gov. Inf. Q., 29(1). 98–109. (Ref 57)
- Petrie, H. and Power, C. (2012) What do users really care about?. Proc. 2012 ACM Annu. Conf. Hum. Factors Comput. Syst. CHI '12. 2107. (Ref 58)
- Petrie, H., Hamilton, F., King, N. and Pavan, P. (2006) Remote usability evaluations With disabled people. Proc. SIGCHI Conf. Hum. Factors Comput. Syst. CHI '06. 1133. (Ref 59)

- Pfleeger, S. L. and Cunningham, R. K. (2010) Why measuring security is hard. IEEE Secur. Priv., 8(4). 46–54. (Ref 60)
- Putnam, R. D. (2000) Better Together: Report of the Saguaro Seminar on Civic Engagement in America. 89–93. (Ref 61)
- Raptis, D., Tselios, N., Kjeldskov, J. and Skov, M. B. (2013) Does Size Matter? Investigating the Impact of Mobile Phone Screen Size on Users' Perceived Usability, Effectiveness and Efficiency. Proc. Int. Conf. Human-computer Interact. with Mob. devices Serv. (MobileHCI '13). 127–136. (Ref 62)
- Renaud, K. (2012) Blaming noncompliance is too convenient: What really causes information breaches?. IEEE Secur. Priv., 10(3). 57–63. (Ref 63)
- Reynaga, G., Chiasson, S. and Van Oorschot, P. C. (2015) Heuristics for the Evaluation of Captchas on Smartphones. Proc. Int. Conf. Human-Computer Interact. with Mob. Devices Serv. Mobile HCI '15. 126–135. (Ref 64)
- Ruba, A., Hartmut, H. and Viswanath, V. (2014) A Usability Evaluation of the Obamacare Website. Gov. Inf. Q., 31(4). 669–680. (Ref 65)
- Sasamoto, H., Christin, N. and Hayashi, E. (2008) Undercover: authentication usable in front of prying eyes. Proceeding twenty-sixth Annu. SIGCHI Conf. Hum. factors Comput. Syst. 183– 192. (Ref 66)
- Sasse, A., (2015) Scaring and Bullying People into Security Won't Work. IEEE Secur. Priv., 13(3). 80–83. (Ref 67)
- Sauro, J. and Kindlund, E. (2005) A method to standardize usability metrics into a single score. Proc. SIGCHI Conf. Hum. factors Comput. Syst. (CHI '05). 401–409. (Ref 68)
- Shay, R., Cranor, L. F., Komanduri, S., Durity, A. L., Huh, P., Mazurek, M. L., Segreti, S. M., Ur, B., Bauer, L. and Christin, N. (2014) Can long passwords be secure and usable? Proc. 32nd Annu. ACM Conf. Hum. factors Comput. Syst. CHI '14. 2927–2936. (Ref 69)
- Sieger, H. and Möller, S. (2012) Gender differences in the perception of security of mobile phones. Proc. 14th Int. Conf. Human-computer Interact. with Mob. devices Serv. companion. 107–112. (Ref 70)
- Taylor, P., Zhang, D. and Adipat, B. (2009) Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. Int. J. Hum. Comput. Interact., 18(3), November 2014. 37–41. (Ref 71)

- Tohidi, M., Buxton, W., Baecker, R. and Sellen, A. (2006) Getting the right design and the design right. Proc. ACM CHI 2006 Conf. Hum. Factors Comput. Syst., 1. 1243–1252. (Ref 72)
- Verdegem P. and Verleye G. (2009) User-centered E-Government in practice: A comprehensive model for measuring user satisfaction. Gov. Inf. Q., 26(3). 487–497. (Ref 73)
- Von Zezschwitz, E., Dunphy, P. and De Luca, A. (2013) Patterns in the wild: a field study of the usability of pattern and pin-based authentication on mobile devices. Proc. Mob. HCI 2013 Secur. Priv. 261–270. (Ref 74)
- Whalen, T. (2011) Security as if People Mattered. Secur. Privacy, IEEE, 9(4). 64–67. (Ref 75)
- Yeratziotis, A., Pottas, D. and Van Greunen, D. (2012) A Usable Security Heuristic Evaluation for the Online Health Social Networking Paradigm. Int. J. Hum. Comput. Interact. 28(10). 678–694. (Ref 76)