UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Electronics and Computer Science

Trustworthiness of Web Information Evaluation Framework

by

Jarutas Pattanaphanchai

Thesis for the degree of Doctor of Philosophy

July 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Electronics and Computer Science

Doctor of Philosophy

TRUSTWORTHINESS OF WEB INFORMATION EVALUATION FRAMEWORK

by Jarutas Pattanaphanchai

Assessing the quality of information on the Web is a challenging issue for at least two reasons. Firstly, there is little control over publishing quality. Secondly, when assessing the trustworthiness of Web pages, users tend to base their judgements upon subjective criteria such as the visual presentation of the website, rather than rigorous criteria such as the author's qualifications or the source's review process. As a result, Web users tend to make incorrect assessments of the trustworthiness of the Web information they are consuming. Also, they are uncertain of their ability to make a decision whether to trust information they are not familiar with. This research addresses this problem by collecting and presenting metadata based on useful practice trustworthiness criteria, in order to support the users' evaluation process for assessing the trustworthiness of Web information during their information seeking processes.

In this thesis, we propose the Trustworthiness of Web Information Evaluation (TWINE) application framework, and present a prototype tool that employs this framework for a case study of academic publications. The framework gathers and provides useful information that can support users' judgments of the trustworthiness of Web information. The framework consists of two layers: the presentation layer and the logic layer. The presentation layer is composed of input and output modules, which are the modules that interface with the users. The logic layer consists of the trustworthiness criteria and metadata creation modules. The trustworthiness criteria module is composed of four basic criteria, namely: authority, accuracy, recency and relevance. Each criterion consists of the items, called indicators, in order to indicate the trustworthiness of Web information based on their criteria. The metadata creation module gathers and integrates metadata based on the proposed criteria that will then be used in the output module in order to generate the supportive information for users. The framework was evaluated based on the tool, using an empirical study. The study set a scenario that new postgraduate students search for publications to use in their report using the developed tool. The

students were then asked to complete a questionnaire, which was then analysed using quantitative and qualitative methods.

The results from the questionnaire show that the confidence level of users when evaluating the trustworthiness of Web information does increase if they obtain useful supportive information about that Web information. The mean of the confidence level of their judgments increases by 12.51 percentage points. Additionally, the number of selected pieces of Web information used in their work does increase when supportive information is provided. The number of pieces of Web information selected by the users increases on average less than one percentage points. Participating users were satisfied with the supportive information, insofar as it helps them to evaluate the trustworthiness of Web information, with the mean satisfaction level of 3.69 of 5 points. Overall the supportive information provided, based on and provided by the framework, can help users to adequately evaluate the trustworthiness of Web information.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Jarutas Pattanaphanchai , declare that the thesis entitled *Trustworthiness of Web Information Evaluation Framework* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:
  - Pattanaphanchai, Jarutas. "Doctoral Consortium proposal: evaluating trustworthiness of web content using semantic web technologies." In *The 10th International Semantic Web Conference (ISWC 2011)*, pp. 325-332. Springer Berlin Heidelberg, 2011.
  - Pattanaphanchai, Jarutas, Kieron O'Hara, and Wendy Hall. "HETWIN: helping evaluate the trustworthiness of web information for web users framework using semantic web technologies." Poster presented at *The 8th International Conference on Semantic Systems (I-SEMANTICS 2012)*, Graz, Austria, 2012.
  - Pattanaphanchai, Jarutas, Kieron O'Hara, and Wendy Hall. "Trustworthiness criteria for supporting users to assess the credibility of web information." In *Proceedings of the 22nd international conference on World Wide Web companion (WWW '13 Companion)*, pp. 1123-1130. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2013.

Signed:.....

Date:.....

Acknowledgements

I would never have been able to finish my thesis without the guidance of my supervisors, help from friends, and support from my family.

First and foremost, I would like to express my deepest gratitude to my supervisors, Professor Dame Wendy Hall for her excellent guidance and patience throughout my PhD. I would also like to thank Dr. Kieron O'Hara, who gave me support and encouragement for all the works I have done, and spent mans hours patiently discussing my research. Their support was of immeasurable value to me and for that I am truly thankful. In addition, I would like to thank Dr. Gary Wills and Lester Gilbert for their comments and suggestions about the statistical issues and experimental design.

Moreover, I would like to thank the Royal Thai Government Scholarship from whom I received the funding for my PhD.

Besides my supervisors, I would like to thank my fellow labmates in the Web and Internet Science and Agents, Interaction and Complexity groups at the University of Southampton who were of great support throughout my PhD. Special thanks goes to Dr.Colin Williams, Dr.Harry Rose, and Dr.Thomas Andritsch who were willing to proofread my publications and my reports. Many thanks to Dr. Oleksandr Pryymak, Dr.Betty Purwandari, Dr.Moody Alam, and Iman Naja for great company and their support.

I would also like to thank my parents, who have always impressed upon me the importance of education and have always supported me both emotionally and financially. Many thanks to my sisters and brothers who have never spoiled me, instead they have trained me to be strong and thoughtful. In addition, I would like to thank Assistant Professor Nittida Elz, who provided me with valuable skill of doing a research independently. I have used those experiences for helping me throughout my PhD.

Finally, I would like to say thank you to all of the friends I have made at Southampton, during my PhD.

Nomenclature

H_0	Null hypothesis
H_a	Alternative hypothesis
df	Degrees of freedom
α	The probability of detecting an effect in the population when that effect does not in fact exist (Type I error)
β	The probability of failing to detect an effect in the population when it actually does exists (Type II error)
P	Statistical power $(1 - \beta)$
SE	Standard Error Mean
ES	The population effect size
t	A statistic t-test
N_{lb}	The lower bound of the expected solution
N_{ub}	The upper bound the expected solution
T_i	The suggested trustworthiness score of the i^{th} result
$S_{A,i}$	The score representing the combined *authority* of the authors of the paper, i
U_d	The usefulness score of an indicator
W_d	The relative weight of an indicator
$P_{i,d}$	A boolean variable which returns value 1 if indicator d is satisfied for the result i
A_i	The set of authors of result i
h	h-index
U_k	The usefulness score of an indicator in authority criterion
W_k	The relative weight of an indicator in authority criterion
$P_{k,a}$	A boolean variable which returns value 1 if indicator in authority criterion a is satisfied for the author k in the author list

Chapter 1

Introduction

Trust is an important factor to be considered when users consume data. This is especially true on the Web, which has an abundance of information but a lack of quality control that allows incorrect or low quality information to be published. In addition, ordinary Web users tend to base their decisions on whether to trust Web information on heuristic factors that are mainly based on surface level characteristics of the Web page (i.e. user interface design) (Fogg et al., 2003). Such characteristics are easily disguised, and Web users can arrive at the wrong conclusions about the trustworthiness of the information they consume. Therefore, it is necessary for Web users to be able to critically assess the trustworthiness of Web information. This is a non-trivial task because it is subjective, in that it depends on a person and the context in which the information is being considered. A piece of information is trustworthy in one context but it might not be trustworthy in another context; for example, a mechanic will be trusted to fix a car but not to perform brain surgery.

A number of studies have suggested that providing supportive information, such as the identity of the author (e.g. name, position, title), the expertise of the author, or the date of publication could potentially increase the Web users' confidence and help them to determine whether the information they have found is trustworthy (Rieh and Belkin, 1998; Wathen and Burkell, 2002). However, in today's Web, the relevant supportive data might not available to gather in order to assess the trustworthiness of the information.

The Semantic Web has been developed in which these issues can be addressed. The Semantic Web was first proposed by Berners-Lee et al. (2001) and is currently the focus of much work in academia. It is a technology that has been designed to make content machine-readable so computers are able to process information more effectively. The Semantic Web describes facts about things and their relationships using the Resource Description Framework (RDF) in the form of subject-predicate-object expressions. Conceptually, RDF will declare data as a graph, which has nodes (subject, object) and edge

(predicate) linking between nodes. Therefore, RDF allows data on the Web to be combined, exposed and shared across different sources by creating the link between new nodes. This ability to easily express and link data facilitates the creation of metadata, which describes information and its relationships to other information. As a result, the Semantic Web provides a solution to address the above problems by providing metadata that describes information. This allows users to use this metadata to make more accurate judgments of whether particular Web information can be trusted.

As far as using metadata to create supportive information is concerned, the criteria used to build supportive information are also important, as they will create the useful supportive information in order to help a Web user to assess the trustworthiness of information. The field of information quality research provides tools and methods that can be applied to analysing the quality of Web data and its data sources. In particular, it describes a number of quality criteria to help in assessing the quality of information (Taylor, 1986; Rieh and Belkin, 2000; Naumann, 2002; Tate, 2010).

Another relevant area of research is that of Web credibility, which is the study of factors that lead people to believe or not to believe the information they find online. Several works have studied and proposed criteria for use in assessing the credibility of Web information (Wathen and Burkell, 2002; Fogg et al., 2003; Persuasive Technology Lab, 2007). From these previous studies on information quality and Web credibility, we categorised the analytic approach for deriving criteria into two main approaches: "normative" and "descriptive" analyses. The normative analysis is an approach to derive the criteria which can advise the users when they are evaluating Web information in order to get best results. Alternatively, the descriptive analysis is an approach that obtains the criteria from the actual behaviour of the users when they are interacting with information.

Even though the criteria from the descriptive analysis are a reflection of the actual behaviours of users, the fact that they do not necessarily reflect the true trustworthiness of the information means that the decision a user makes regarding whether or not to trust the information may be no better than arbitrary. For example, the criteria from descriptive analysis include surface characteristic criteria (Fogg et al., 2003; Persuasive Technology Lab, 2007), which are easy to disguise using professionally designed templates such as those from content management systems. Therefore, the criteria from the descriptive analysis are subjective and are not rigorous enough to support the user in making a critical judgment of the trustworthiness of Web information. In contrast, the criteria from the normative analysis are objective factors that consider the trustworthiness of Web information based on strong evidence. Such objective criteria include: the authority criterion, relevance criterion, currency criterion and accuracy criterion (Rieh and Belkin, 1998; Tate, 2010), which we will discuss in more detail in Chapter 4.

Therefore, our aim was to use objective criteria that are rational in order to develop a normative model, which we combine with Semantic Web technologies in order to

present a framework called TWINE (Trustworthiness of Web Information Evaluation) that helps users to assess the trustworthiness of Web information. We describe our research hypothesis in section 1.1, then, we present our research contributions in section 1.2. Finally in section 1.3, we state our thesis structure.

1.1 Research Hypothesis

A numerous amount of information presenting on the Web makes it difficult for Web users to make reliable quality assessments whether a piece of information published on the Web was valid, legitimate, or even just interesting. Therefore, having a tool that helps them to evaluate the trustworthiness of the information can improve their ability for making judgments. We present the trustworthiness of Web information evaluation (TWINE) framework, which consists of modules for gathering and integrating useful metadata based on objective criteria using Semantic Web technologies. In addition, TWINE presents the supportive information from integrated useful metadata to the users in order to support their evaluation of the trustworthiness of Web information.

Accordingly, we set the overall hypothesis for our research as "A framework (such as TWINE) with properties of gathering, integrating and presenting supportive information using Semantic Web technologies helps users to more effectively evaluate the trustworthiness of Web information". More specifically, we divide our hypothesis into sub-hypotheses as follows:

- Using our framework, users increase their confidence in their judgement of the trustworthiness of the Web information that they find.
- Using our framework, users increase the number of pieces of trustworthy information that they select to use.
- Using our framework, users are satisfied with the supportive information insofar as it helps them to evaluate the trustworthiness of Web information.

Against this, in the next section we discuss the contributions of this thesis.

1.2 Research Contributions

The aim of our research is to propose a framework for helping Web users to critically evaluate the trustworthiness of Web information. In addition, our framework will help the users to increase their confidence to evaluate the trustworthiness of Web information and their satisfaction in using the data they have found. Therefore, our main contributions are as follows:

- The evaluation and selection of the criteria used for supporting the evaluation of the trustworthiness of Web information.
- The calculation of the criteria's weights for use in calculating the total suggested trustworthiness score of the information. The calculation of such a score will allow search results to be ranked based on their relative trustworthiness.
- The integration of metadata gathered using Semantic Web technologies based on our trustworthiness criteria to build supportive information that can be used to evaluate the trustworthiness of the information on the Web.
- The interpretation of the data model created during the above process to produce an explanation of the trustworthiness in a human-readable form to users who will, in turn, use this data to support their decisions.
- A prototype tool, which is implemented based on the proposed framework.

Our research work has been presented in the following papers:

- Pattanaphanchai, Jarutas. "Doctoral Consortium proposal: evaluating trustworthiness of web content using semantic web technologies." In *The 10th International Semantic Web Conference (ISWC 2011)*, pp. 325-332. Springer Berlin Heidelberg, 2011.
- Pattanaphanchai, Jarutas, Kieron O'Hara, and Wendy Hall. "HETWIN: helping evaluate the trustworthiness of web information for web users framework using semantic web technologies." Poster presented at *The 8th International Conference on Semantic Systems (I-SEMANTICS 2012)*, Graz, Austria, 2012.
- Pattanaphanchai, Jarutas, Kieron O'Hara, and Wendy Hall. "Trustworthiness criteria for supporting users to assess the credibility of web information." In *Proceedings of the 22nd international conference on World Wide Web companion (WWW '13 Companion)*, pp. 1123-1130. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2013.

With this in mind, the next section describes the structure of this thesis.

1.3 Thesis Structure

This thesis is structured as follows:

Chapter 2
A review of the literature that is relevant to this thesis is presented. In more detail, this

chapter presents background knowledge on the Web and the Semantic Web, which are the main technologies used in this thesis. Moreover, we discuss trust and research areas about trust on the Web; namely, information quality and Web credibility.

Chapter 3
This chapter presents the development process of the Trustworthiness of Web information (TWINE) framework. We then present the process of constructing the trustworthiness criteria module. We describe the process of composing the trustworthiness criteria module into the framework together with the metadata collection and presentation modules.

Chapter 4
This chapter presents the development process of deriving the trustworthiness criteria. We discuss the process of analysing the trustworthiness criteria from the information quality and Web credibility research areas, and selecting those to be used in our framework.

Chapter 5
This chapter describes the validation process of the proposed criteria. We demonstrate the development of an instrument to validate the indicators, which are representative of the proposed criteria to be used in the TWINE framework. We compile the questionnaire asking the expert to validate the indicators to ensure that the proposed criteria are helpful.

Chapter 6
This chapter discusses the development process of the TWINE prototype, and presents the process of implementing a prototype based on the framework proposed. In addition, we conduct usability tests for evaluating the prototype in order to ensure that the users will be able to understand what the prototype is used for, and to use it easily.

Chapter 7
This chapter presents the evaluation of the framework. We discuss the evaluation process used to assess our proposed framework through the proposed prototype. We describe the design of a study that evaluates the framework by assessing it based on the prototype. We then analyse the results of this study.

Chapter 8
Finally, this chapter provides a summary of this thesis. In addition, we discuss the conclusions that can be drawn from the results in the evaluation of the framework process. Furthermore, we discuss future extensions to this work.

Chapter 2

Literature Review

This chapter provides a discussion of the background literature that inspired the work in this thesis. It also sets out an overview of the current status with regards to the assessment of trust in the Web. In more detail, Section 2.1 provides an overview of the basic concepts of the Web that are required in order to understand trust in this area. Next, Section 2.2 describes the Semantic Web and also explains its architecture, and its technology. Afterwards, Section 2.3 is a discussion of the concept of trust. Section 2.4 is a discussion of the relevant topics of trust on the Web. Finally, in section 2.5, we draw the chapter to a close by summarising our discussions and drawing conclusions.

2.1 The World Wide Web

The World Wide Web (sometimes referred to as "WWW" or simply, "The Web") is a service on the Internet that links documents together using hypertext technology. This technology allows users to access general information on the Internet without having to consider the geographical location of the material or the operating system of either their own computer or the host computer. Documents can refer to each other through links, as displayed in Figure 2.1. Web browsers then display the linked documents by interpreting the hypertext information (HTML) from which the document is composed (Berners-Lee et al., 1994). Therefore, the Web is a network of linked document resources. More importantly, because of the universal accessibility of the Web, the Web makes the documents published on it available to a *global* audience. Not only that, but the universal accessibility of the web also enables *anybody* to publish information about whichever topic they so choose, regardless of the quality of that information or the author's knowledge of that topic. In order to facilitate and manage this, the Web consists of three components.

- Uniform Resource Identifiers (URIs) for locating resources on the Web (Berners-Lee et al., 1998).
- Protocols such as HyperText Transfer Protocol (HTTP) for accessing, distributing, potentially collaborative hypermedia information resources on the Web. HTTP allows an open-ended set of connection approaches and headers, which identify the details of a request (Berners-Lee et al., 1996).
- HyperText Markup Language (HTML) for describing the layout of documents, describing navigation among linked document resources, designing forms for implementing transactions, and including multimedia files such as images, audio or videos into documents (Berners-Lee and Connolly, 1995).

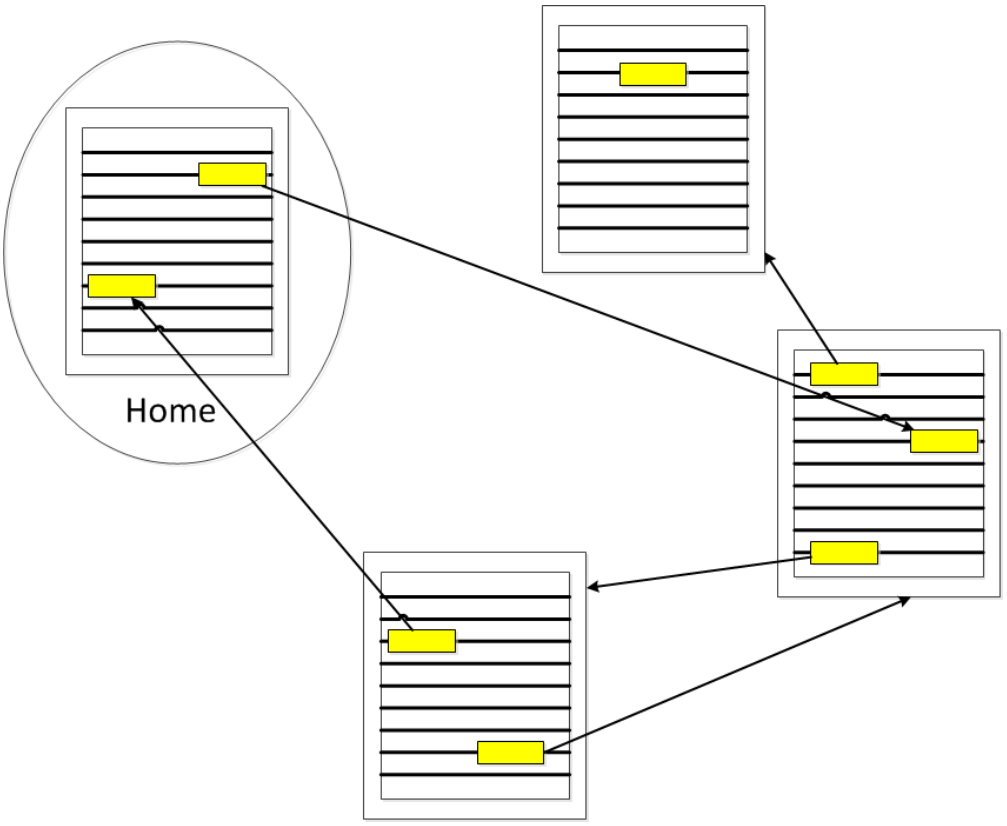


Figure 2.1: The basic Hypertext model

Although the Web itself does not define "classes" of sites (that is, the Web simply defines how documents should be defined and not what they should contain), certain common classes or types of websites have emerged through popular use of the Web. These are discussed in the next section.

2.1.1 Types of Web Sites

There are several types of web sites present on the Web such as personal website, commercial website, or news website. The type of content on a site is dependent upon that

site's purpose. For example, the purpose of a commercial website is try to sell or provide the products; thus the content on the website will present the information about the products, the price of the products and how the users can buy these products. Therefore, each type of website will have specific features which will affect the way that users evaluate the information contained in it. Take, for example, a website that presents a personal idea or concept. Here, the identity of the author will be an important feature for users to consider because it can indicate the author's level of expertise, which will in turn affect the credibility of the information the author provided. However, the likelihood that the required information can be used to evaluate a particular web site is not as clear-cut as it appears, as there are cases in which the properties of one website's typology can overlap with the other types of website. Nevertheless, the website still display the same primary content that it is intended to present. For example, for a personal website, main content is mainly about personal details such as the author's workplaces, hobbies or interests, or opinions. However, it may also present information designed to sell products which are produced by the owner of that personal website. When it comes to designing the layout and presentation of a website, a web designer will often take the *purpose* of the website into account. Therefore, a website will be presented in the form of personal design (e.g. less formal layout or the use of more informal language). As far as an overlap between the type of website is concerned, the classification of the site will help the web designer or developer to clearly understand the intended purpose of the website. In turn, they can design and present the content with a focus on the site's primary purpose (Shelly et al., 2008; Sklar, 2008).

Several studies define types of web sites based on what form of information the site intends to provide to the users. For example,

Crowder and Crowder (2008) categorised five basic types of sites; these are personal, informational, organisational, political, and commercial sites.

- The main purpose of a *personal* website is to introduce an individual's interests, ideas or biography to the public. The information provided is limited in scope to close friends, family, and the person involved.
- An *informational* site provides information on a particular topic or offers a limited amount of information without any charge.
- An *organisational* site presents organisation-specific information. This information specifically relates to the organisation in question.
- A *political* site publishes information about a particular political candidate. It aims to provide information about elections and also the candidate's social agenda in order to raise the candidate's popularity among voters.

- A *commercial* site's primary purpose is to sell products. The site presents information about products or services to encourage people to buy those products or services.

Smith (2008) identified five types of Web pages as personal, picture, topical, commercial, and entertainment types.

- A *personal* site provides a person's information to share with his/her friends, colleagues, family, and others.
- A *picture* site is a site to which users can upload their pictures to show or share online.
- A *topical* site focuses on a specific topic of concern or interest, a cause, or the passion of the creator or volunteer group.
- A *business* site basically tries to promote products or services for sale. It covers a variety of styles of presentation, depending on the goals of the site.
- An *entertainment* site is mainly for entertaining users. It might provide humorous stories or games for the users.

Shelly et al. (2008) categorised the types of web sites as personal, organisational/topical, and commercial.

- A *personal* site presents an individual's passion for something. It provides personal details or information.
- An *organisational/topical* site provides information to promote or support an organisation. For example, the World Health Organization's web site provides information concerning health-related issues, whereas, the American Kennel Club provides information focused on dog breeds, pedigrees, and shows.
- A *commercial* site aims to present information in order to promote and sell products or services.

Tate (2010) defined six types of Web pages based on the primary purpose of presenting content to the audience.

- *Advocacy Web pages* publish content or idea to influence public opinion, or to encourage activism such as to increase voter turnout, to increase membership, or to promote a cause.
- *Business* Web pages primarily promote and sell products or services.

- *Informational* Web pages have a primary purpose of providing factual information. For instance, they may provide government research reports, census data, or statistical results of research.
- *News* Web pages focus on providing current information on local, regional, national, or international events.
- *Personal* Web pages basically present a person's interests, expressions, or opinions of something. It may state the name of the author but it has no affiliated organisational support.
- *Entertainment* Web pages provide enjoyable content for the users such as humorous stories, music, and games.

MacDonald (2011) suggested six different types of website:

- *Personal* sites mainly show information on a person.
- *Resumé* sites are another type of personal site but they specifically present the details of a person's work history or portfolio.
- *Topical* sites focus on a particular area of interest.
- *Event* sites provide information on a specific event for a certain period of time. An example of an event site is a wedding website, which is created by the host in order to provide information about a wedding (directions, background information, gift registries, etc.) and which will be removed when the wedding is over.
- *Promotion* sites have a primary purpose to show off personally produced products.
- *Small business* (or e-commerce) sites are basically selling anything that can be sold online such as amazon.com.

Different types of websites will publish different kinds of information based on that site's purpose. However, some types of websites in the aforementioned studies can be considered as being the same, even if each study allocated different titles to the sites. Therefore, all these websites can be combined into one single type that encompasses all those originally cited in the literature, and all types will impact on the level of trust placed in them by the user. Consequently, we analyse the types of websites discussed above in order to create Web information domains that are used in our research; this is discussed in greater detail in the next section.

2.1.2 Web Information Domains

From the categorisation of each website in section 2.1.1, we synthesise and categorise the information domains that are used in this research. These categories are then set as the framework within which to assess user's level of trust in a domain.

We note that some of the website types have been referred to in numerous studies and furthermore, some types, although distinct in name, actually refer to the same broad type of website. For example, take the "*personal site*" type, which is classified as a site that provides information about a person (interests, ideas, expressions, and opinions). We note that the "*resumé site*" type is similar to the "*personal site*" type because its purpose is also to present personal information, but it specifically provides the details of a person's work history. Consequently, for the purpose of this investigation, we merge "*personal*" and "*resumé*" into a single domain. Moreover, the "*informational site*" type is also defined as a domain named "*informational*", whose main purpose is to provide academic publications or other factual information on a particular topic. Likewise, the "*entertainment site*" type is basically providing enjoyable content. In this way, we merge types of web site which have similar purposes for presenting information but might be defined with different names. Then, we assign a name to the merged type that is representative of the overall concept that the merged type encompasses. For example, the business, promotion and commercial site types present information of products or service for sale. We group them as one type named, "*commercial site*." Similarly with, the organisation, topical, advocacy, event, picture, and political site types, they provide information on a particular area of interest, organisation, or event. Therefore, we merge them as a group named the "*topical*" domain. Finally, the news site type is a specific type that presents current information on local, regional, or international events. Consequently, we categorise Web information domain into six domains:

- **Topical domain**: Sites within this domain will provide information related to a specific organisation. It provides information to advocate an individual's or a group's opinion.
- **Commercial domain**: The main content provided within this domain is information for promoting or selling products or services.
- **Informational domain**: The information presented in this domain is intended to provide factual information, statistical data, or academic publications (results or reports of research, and articles) which can educate users.
- **News domain**: The information in this domain specifically discusses current information on local, national, or international events.
- **Personal domain**: This domain provides information on an individual such as personal details, interests, passions, and opinions.

- **Entertainment domain**: The main purpose of the information in this domain is to entertain the users.

These six domains of Web information are used as the basic domains of our evaluation of the Web information process. Even though, the content on the websites might overlap with other domains characteristics, the evaluation process is based on the primary purpose of that web site. For instance, if the web site is a personal site which also provides information about the site owner's publication details, the evaluation process focuses its evaluation based on the fact that this is a personal website.

2.1.3 Web Information Challenge

The linked nature of documents on the Web influences use and distribution of information online because it allows users to produce information and present it widely, regardless of the physical location of the publisher. In addition, it is easy to link from one document to other documents. Although this feature has increased the popularity of using the Web, it also raises concerns about the trustworthiness of the information published on it. Unlike traditional publishing platforms such as books and newspapers which undergo an editorial process to check and assess the quality of the information before it is published, the Web allows everyone to publish their information without any checks. Therefore, assessing the trustworthiness of Web information is challenging; as a result, Web users require critical criteria and tools to do this.

One promising approach is to make available additional information concerning the provenance of the Web information the user is browsing. Without useful supportive information about the content they are consuming, users tend to base their judgement on arbitrary factors such as web site design, or their familiarity of the topics (Fogg et al., 2000, 2002, 2003). This can lead to the wrong decision being taken to use that information. Such supportive information should include important features regarding more robust normative criteria of trustworthiness. Examples may include the identity of the author (e.g. name, affiliation), the publication date of the information, and the number of times that the information has been referenced.

Moreover, the additional provided information will help to increase an individual's confidence pertaining to the trustworthiness of information, and will help to improve the accuracy of their assessments. However, the interesting issues are how we get that supportive information and how we integrate it.

A technology, called the Semantic Web (SW), has been proposed. It has been designed to make data on the web more meaningful by linking in semantic data (Berners-Lee et al., 2001; Pollock, 2009). This enables computers to work with the data more intelligently by allowing the computer to interpret the data and to know the meaning of the information.

Moreover, it allows data to be integrated easily whereas in the Web of documents, information is not machine readable and has to be gathered and assessed by human readers. This requires a very resource-heavy process in order to manage information.

In the next section, we provide the background on the Semantic Web in order to give a reference point for understanding the technology that we use when gathering useful supportive information and integrating it which is one of the main approaches in this research.

2.2 The Semantic Web

The Semantic Web can be defined in many different ways. Berners-Lee et al. (2001) stated that, "*The Semantic Web is an extension of the current Web in which information is given a well-defined meaning, better enabling computers and people to work in cooperation.*" Similarly, Harth et al. (2009) pointed out that "*The Semantic Web is the extension of the World Wide Web that enables people to share content beyond the boundaries of applications and websites.*" In addition, Uschold (2003) defined the Semantic Web as "*machine-usable content.*" Put another way, Herman et al. (2008) explained that, "*The Semantic Web is a vision: the idea of having data on the Web defined and linked in a way that it can be used by machine not just display purposes, but for automation, integration and reuse of data across various applications.*"

As a result, in our work we define the Semantic Web as a Web of linked data which are considered to be machine understandable, reusable, and interpretable. In section 2.2.1, we describe the Semantic Web architecture. Then, in section 2.2.2, we discuss the importance of the Resource Description Framework (RDF) which is used to describe the objects on the Web. We discussed named graph which is an extended RDF graph with its provenance information. In section 2.2.4, we explain the "Simple Protocol and RDF Query Language," which is used to retrieve data from the RDF. Finally, in section 2.2.5 we discuss a method that describes how to publish and interconnect the structured data (i.e. RDF).

2.2.1 The Semantic Web Architecture

Berners-Lee et al. (2001) stated that the Semantic Web is an extension of the current Web. Therefore, the Semantic Web consists of the established standards of the current Web technologies that allow information to be defined in a well-defined manner, then to be shared and integrated across resource boundaries. These features enable computers and people to work in collaboration. The Semantic Web architecture was firstly designed by Berners-Lee (2000) as a common framework that consists of layers of Web technologies and standards that function collaboratively. The latest version of the Semantic Web

architecture was presented by Berners-Lee (2006a) at the AAAI2006 conference as shown in Figure 2.2.

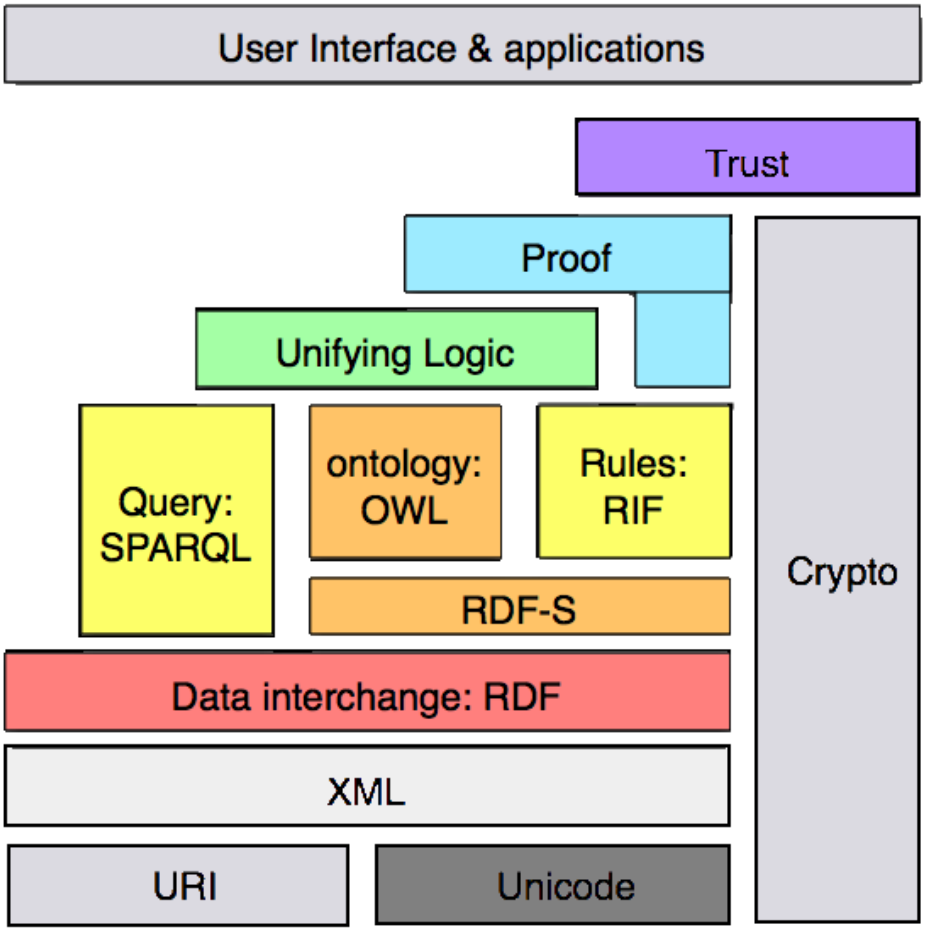


Figure 2.2: The Semantic Web Architecture

The bottom layer consists of URI and Unicode as two separate blocks. Unicode, a universal character code, is a basic character encoding for any language. Hence, it helps to process and communicate data that travel between regions of the world (Bettels and Bishop, 1993).

The URI is used for identifying resources on the WWW. The main functions of a URI are to provide a unique name to data objects across the Internet such that they can be uniquely identified, and to indicate links between those objects (Berners-Lee et al., 1998; Pollock, 2009). URIs can be used to identify things that are both network and non-network accessible such as electronic documents, services, human beings, books, or abstract concepts that do not physically exist.

In the next layer, XML (eXtensible Markup Language) is a language for describing the data, logical structure, and storage layout of a document using tags. This enables

machines to understand and extract data from documents. Moreover, it can be used for exchanging data between applications (Otto et al., 2001).

The layer above the XML layer is the data interchange layer, which uses RDF (Resource Description Framework); a language for identifying Web resources using a URI and by providing a framework explained by XML. A resource in RDF could be something which users want to describe. RDF represents Web resources and their properties as a graph of nodes and arcs. More specifically, an RDF graph is defined in the form of triples: a *subject*, the source of the relationship; a *predicate*, the labelled arc; and an *object*, the relationship's destination. The subject and the predicate are always resources, while the object can be a resource or a string (Decker et al., 2000; Manola et al., 2004; Pollock, 2009). An example of RDF is illustrated in section 2.2.2. The RDF is considered to be a medium that connects between the low level of the Semantic Web architecture and the higher layer. Also, it can be used for data interchange, which is discussed in more detail in section 2.2.2.

RDF-S is an RDF schema which is described by conceptual models (a diagram that defines theoretical entities, objects, or conditions of a system and the relationships between them). Schemas define kinds or classes of resources including their specific properties. They can be used for semantic annotations and are also used to indicate which classes and properties are expected to be used together (Nejdl et al., 2000; Manola et al., 2004).

OWL, the Web Ontology Language, extends RDF and RDF-S by adding more vocabulary terms for explaining sets of things, the resources that users want to describe, facts about those classes, relationships between classes, and characteristics of relationships. OWL is described by RDF/XML so it appears similar to RDF/XML, but it has additional reserved words and special ways to format data (McGuinness and Van Harmelen, 2004; Pollock, 2009).

The Rule Interchange Format (RIF) is a format for exchanging business rules between heterogeneous software engines. It is defined using XML; thus machines can execute these rules. The Simple Protocol and RDF Query Language (SPARQL) is a standard query language for RDF data; and is explained in more detail in section 2.2.4 (RIF Working Group, 2005; Pollock, 2009; SPARQL Working Group, 2009).

The Unifying Logic Layer is for describing a formal mathematical logic which is used to define all of the different model semantics (RDF, RDF-S, OWL, SPARQL, and RIF) into a constant model theory (RIF Working Group, 2005; Pollock, 2009; SPARQL Working Group, 2009).

The proofs are sets of rules which are created based on a set of requirements. They are used to verify identity or permission of agents or services to access or use information which leads to a conclusion of a requested service (RIF Working Group, 2005; Pollock, 2009; SPARQL Working Group, 2009).

The next layer is trust. On the Semantic Web, as was the case with the non-semantic web, there are no guarantees as to the quality of the data. Therefore, a machine needs to evaluate the data items and consider how to use data appropriately. It might need to know the origin of the data and their authorship, or the reputation of the source of information in order to assess the trustworthiness of the information. Therefore, the trust layer is concerned with verifying data or evaluating the trustworthiness of information by the machine itself. This layer tries to distinguish fake information from the genuine data. It provides an assurance of the information's quality such as its validity and a degree of confidence in the resource (Matthews, 2005; Harth et al., 2011). Trust was designed to sit at the top of the Semantic Web architecture from the Semantic Web's inception (Berners-Lee, 2006a; Golbeck, 2006). This indicates how important trust is. Therefore, when the developing on the Web we should not only be concerned with the protocol or language but we should also consider the end use of web-based information (Pollock, 2009).

In this research, we focus on an important issue of trust and using the data provided on the Web or even on the Semantic Web. However, we do not replace our approach with the trust layer. In particular, our research tries to address the challenge of assessing the trustworthiness of Web information, which is the same concept in trust layer of the Semantic Web. However, our approach focuses on the *human* role in the evaluation of the trustworthiness of information rather than the machine role which is the case for the trust layer in the Semantic Web model. The trust layer of the Semantic Web focuses on the verification of data by machine in order to model trust and to allow machines to work cooperatively. Conversely, we see the potential of developing approaches that can apply the technologies of the Semantic Web in order to help Web *users* to evaluate and verify the trustworthiness of information by providing some supportive information before they can place their trust in the information. Still, our approach could be adjusted to be implemented in the trust layer of the Semantic Web, but some functions must be edited to allow the machines to be able to use our approach to modelling trust between them.

Finally, cryptography uses encryption techniques to protect the layers below the trust layer (RIF Working Group, 2005; Pollock, 2009; SPARQL Working Group, 2009). Encryption is a method that converts secret or sensitive information from an intelligible form to an unintelligible form (without the appropriate decryption key) (Smid and Branstad, 1988; Kaliski, 1993). Therefore, it can be used to protect the data in the lower layers of the stack such as the unicode or XML blocks by converting it into another form that cannot be read by third party. By encrypting data, it can be ensured that the data cannot be tampered with by a third party.

The Semantic Web architecture was developed in order to present layers of expression and comprehension features which allow the Web to be extended in order to support the concept of an intelligent Web (Berners-Lee, 2000). Each layer in the architecture

provides functions which work together. For example, unicode and URIs are used to indicate data objects in the Internet by using XML to explain what they refer to. The RDF is used to describe the relationships between each data object. Therefore, we focus on using the technologies from the Semantic Web in the data interchange layer, which provides the features to identify objects and to describe the relationship between objects. These features can support the gathering and integration of data on the Web. We now discuss these features in more detail, starting first with a discussion on the RDF.

2.2.2 Resource Description Framework (RDF)

The RDF is the base language of the Semantic Web. It is used to describe things in a way that computers can understand and it also explains the relationship between those things. Consequently, we focus on employing this technology within our system. We now explain this technology in more detail to give some background knowledge and to express its benefits.

The RDF is a language for expressing data and metadata. It can represent metadata about Web resources such as the title, author and modification date of a Web page. RDF is not only used to represent things that can be identified on the Web but it can also refer to physical objects that cannot be directly retrieved from the Web. Moreover, RDF provides a framework for expressing and exchanging data between different applications.

The RDF represents resources as a graph which presents resources and literal values as nodes and their properties or relationships as arcs, which we call *triples* or *statements*. This representation, with nodes and arcs, allows these resources to be linked together on a global scale across the Internet. The RDF graph identifies items and relationships using URIs. To describe this expression more clearly, we use the following example: consider the statement, "There is a person identified by `http://id.ecs.soton.ac.uk/person/23796`, whose name is Jarutas Pattanaphanchai." RDF explains the various parts of the statements using a particular terminology. Firstly, the part which identifies the thing in the statement (the person in this example) is called the *subject*. Secondly, the part that identifies the property or characteristic of this subject (name in this case) is called the *predicate*, Thirdly, the part which identifies the value of that property is called the *object* (W3C, 2004). Therefore, the RDF terms of this statement are:

- The subject is the person identified by URI `http://id.ecs.soton.ac.uk/person/23796`
- The predicate is the word "name" which represents the relationship that the subject has a name
- The object is the phase "Jarutas Pattanaphanchai"

This statement could be represented as the RDF graph in Figure 2.3 in which nodes that are identified by a URI are shown as ellipses, while nodes that are literals (constant values represented by a string or number) are shown as boxes. In addition, the edge that identifies the relationships which exist between the linked nodes is also identified by a URI. In particular, RDF uses URI references (or *URIref*) which are statements that consist of a URI and an optional *fragment identifier* at the end for identifying the subjects, predicates, and objects. For example, the URI reference `http://www.example.org/index.html#section2` consists of the URI `http://www.example.org/index.html` and the fragment identifier *section2* at the end which is separated by the '#' character.

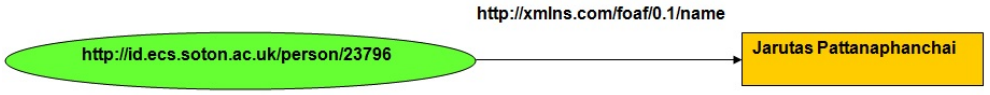


Figure 2.3: A simple RDF graph

Figure 2.3 represents an RDF statement having:

- a subject which is identified by URIref as `http://id.ecs.soton.ac.ul/person/23796`
- a predicate identified by URIref as `http://xmlns.com/foaf/0.1/name`
- an object which is described by literal as string

Since RDF uses URIrefs to name things in a statement instead of using just words, RDF refers to a set of URIrefs which is defined for some specific purpose as a *vocabulary*. For example, the set of URIrefs defined by FOAF (The Friend of a Friend) project is a vocabulary to describe people, the links between them and the things they create and do (FOAF, 2000). For instance, from Figure 2.3, the URIref `http://xmlns.com/foaf/0.1/name` means that this item explains the name of the person and that the content is, "Jarutas Pattanaphanchai". A benefit of using URIrefs to identify subjects, predicates, or objects in statements is that it can define the items more precisely(W3C, 2004).

RDF uses a specific XML syntax, called RDF/XML, for representing RDF statements in a machine processable and exchangeable term. An example of RDF/XML can be seen in Figure 2.4. This is the RDF/XML which corresponds to the graph in Figure 2.3.

The tags in RDF/XML allow programs to understand what the information means, therefore allowing programs to interpret that content properly. RDF is not limited to

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:foaf=http://xmlns.com/foaf/0.1/>
    <foaf:Person rdf:about=" http://id.ecs.soton.ac.uk/person/23796>
        <foaf:name>Jarutas Pattanaphanchai</foaf:name>
    </foaf:Person>
</rdf:RDF>
```

Figure 2.4: RDF/XML example

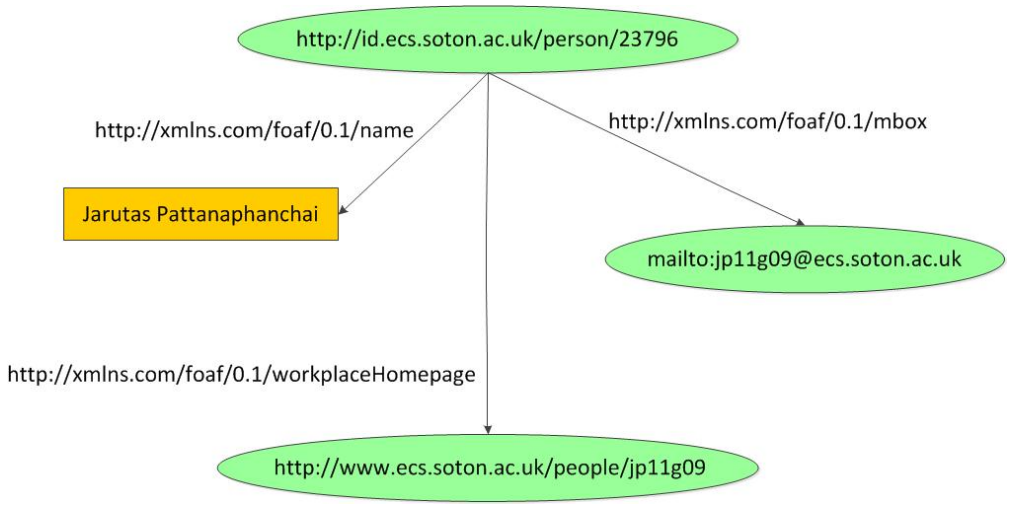


Figure 2.5: Several statements about the same resource

describing only one thing at a time. It can describe several things that explain the same resource, as shown in Figure 2.5.

Figure 2.5 shows some additional statements of Jarutas's information namely, her email address and homepage. In addition, an alternative way to write the statements is writing with triples notation. Each statement in the graph can be written as a triple consisting of a subject, a predicate, and an object in that order (Beckett, 2013). For example, the statements displayed in Figure 2.5 could be written in the triples notation as shown in Figure 2.6.

```
<http://id.ecs.soton.ac.uk/person/23796>  <http://xmlns.com/foaf/0.1/name>  ``Jarutas Pattanaphanchai'' .
<http://id.ecs.soton.ac.uk/person/23796>  <http://xmlns.com/foaf/0.1/workplaceHomepage>  <http://www.ecs.soton.ac.uk/people/jp11g09> .
<http://id.ecs.soton.ac.uk/person/23796>  <http://xmlns.com/foaf/0.1/mbox>  <mailto:jp11g09@ecs.soton.ac.uk> .
```

Figure 2.6: The triples notation example

The triples notation describes triple statement as a sequence of (subject, predicate, object) terms, separated by white-space characters and terminated by '.' after each triple. RDF structures are primarily represented using the graph model, and the triple notation is a secondary representation for convenience.

Carroll et al. (2005b) proposed a general variation on RDF, called *named graphs*. It is an RDF graph which extends the syntax and semantics of an RDF graph by adding

a name to the graph in the form of a URLRef. This feature allows RDF graphs to express metainformation about graphs in order to describe the graph itself and present the relations between graphs. We discuss named graphs and their abstract syntax and semantics in detail in the next section.

2.2.3 Named Graphs

The metainformation about an RDF graph is required for managing RDF graphs effectively such as keeping track of information process chains, restricting the usage of published information, controlling access to information, signing RDF graphs, expressing propositional attitudes, scoping assertions and logic. To allow these features, it needs to extend RDF to include a further URIref or blank node or ID to express syntactic and semantic properties, and the relationship to the RDF's triples. As a result, a named graph has been proposed to address this issue. A named graph is an RDF graph which consists of two elements, *a name (URI)* and *an rdfgraph (RDF graph)* (Carroll et al., 2005a). Named graphs can be stated in three ways: Trix, RDF/XML and TriG. Both Trix [1] and RDF/XML describe named graphs based on XML. TriG [2] states a named graph as a compact plain text format. In this research, we use Trix as a syntax to describe named graphs because it uses XML format which allows the use of XML tools such as XSLT or XQuery. The example of named graphs described the RDF in Figure 2.5 with Trix is illustrated in Figure 2.7.

```
<TriX xmlns="http://www.w3.org/2004/03/trix/trix-1/">
    <graph>
        <uri>http://id.ecs.soton.ac.uk/person/23796</uri>
        <triple>
            <uri>http://id.ecs.soton.ac.uk/person/23796</uri>
            <uri>http://xmlns.com/foaf/0.1/name</uri>
            <plainLiteral>Jarutas Pattanaphanchai</plainLiteral>
        </triple>
        <triple>
            <uri>http://id.ecs.soton.ac.uk/person/23796</uri>
            <uri>http://xmlns.com/foaf/0.1/mbox</uri>
            <uri>mailto:jp11g09@ecs.soton.ac.uk</uri>
        </triple>
        <triple>
            <uri>http://id.ecs.soton.ac.uk/person/23796</uri>
            <uri>http://xmlns.com/foaf/0.1/workplaceHomepage</uri>
            <uri>http://www.ecs.soton.ac.uk/people/jp11g09</uri>
        </triple>
    </graph>
</TriX>
```

Figure 2.7: An example of a Trix document

Named Graphs can be implemented in existing Semantic Web tools and can be of benefit to many Semantic Web application areas. For example, the TriQL.P browser (Bizer

[1] Trix is an alternative XML syntax for RDF which adds the ability to explain name and semantics on the RDF graph (Carroll and Stickler, 2007).

[2] TriG is an extension of the Turtle (Beckett and Berners-Lee, 2011) which states a group of triple statements surrounded by "{" and "}" to group triples into multiple graphs and gives a name to the graph (Bizer and Cyganiak, 2013).

et al., 2005) is a RDF browser which filters information using user-selected trust policies. The policies are dependent on the information context, the content and rating of the source, and some information about digital signatures. Furthermore, it also provides an explanation as to the reason why the browser selected the information it did to display for the users. Named Graphs are used in TRIQL.P as a data model for this application. This is an example of using RDF to implement an application, and users can obtain a benefit from the application. Consequently, RDF has the advantages of describing, linking and exchanging thing across the Internet. We note this benefit, and adopt RDF as the basic technology in our work.

2.2.4 Simple Protocol and REF Query Language (SPARQL)

The SPARQL standard is a query language and protocol for RDF. It provides an interface to interact with an RDF database. The SPARQL protocol is a method to send SPARQL queries from clients to a query processor. The protocol is described in terms of an abstract interface (independent from any specific technology or implementation) and a connection to this interface such as HTTP. Users can write queries to return information from an RDF database in the form of triple patterns. The query needs to be written in specific patterns that should be matched in a result set. This particular pattern in a query, with the target RDF model, is then be considered by a SPARQL processor. It will match the query pattern to the data in RDF and return the results. Moreover, the queries can also consist of conjunctions (logical "and") and disjunctions (logical "or") to provide more precise results from the query (W3C, 2008). An example of a SPARQL query is shown in Figure 2.8.

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?name ?mail
WHERE
{
        ?persondetail foaf:name ?name .
        ?persondetail foaf:mbox ?mail .
}
```

Figure 2.8: A basic SPARQL example

This query looks for person information (name and email address) from an RDF database. The query processor matches the pattern in the WHERE clause with all RDF instances in the graph model. The first pattern tries to match the RDF instances that have a foaf:name property. The second pattern matches all RDF instances which have a foaf:mbox property. These two patterns are inside braces, thus the query will return only results for which these two patterns are true (i.e. braces imply a logical conjunction). In addition, the symbol "?" in front of the words persondetail, name and mail indicates that they are variables (the thing that we are looking for). Each triple pattern

ends with the "." symbol. As a result, data that would be return from this query from RDF in figure 2.6 are "Jarutas Pattanaphanchai jp11g09@ecs.soton.ac.uk." In summary, SPARQL is designed to be used for querying information from an RDF data model. It can be easily used with RDF similar to the use of SQL queries with relational databases.

2.2.5 Linked Data

Linked data describes an approach to publish and interlink structured data on the Web. It constructs the data on the Web in machine-readable form. The meaning of information content is explicitly defined, and can be linked to- or from- other external data sets. Linked data uses RDF to define typed statements, which may refer to any objects (tangible or abstract objects) in the world (Heath and Bizer, 2011). Berners-Lee (2006b) defined the Linked Data principles as a set of rules for publishing data on the Web, in order to make all published data interconnect to each other and becomes a part of a single global data space. These principles are the following :

1. Use URIs as names for things.
2. Use HTTP URIs, so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL).
4. Include links to other URIs, so that they can discover more things.

These principles provide a basic guideline for publishing and connecting data apply to the general architecture of the Web. It relies on two technologies: URIs and HTTP. URIs is used to identify not only digital content, but also real world objects and abstract concepts such as people, books, and relationship types. The HTTP protocol provides a universal mechanism to retrieve data from the Web. The HTTP protocol enables the URIs to be looked up by dereferencing the URI in order to identify objects and abstract concepts. Publishing a data set as Linked data on the Web according to the Linked Data principles consists of three basic steps (Bizer et al., 2009):

1. Assign URIs to the entities described by the data set and these URIs are used to deference over the HTTP protocol into RDF representations.
2. Set RDF links to other data sources on the Web, so that clients can navigate the Web of Data as a whole by following RDF links.
3. Provide metadata about published data, so that clients can assess the quality of published data and choose between different means of access.

The Linked Data principles set the foundations to extend the Web into a global data space (Web of Data) rather than linked documents using the architectural principles of today's Web. Therefore, the Web of Data can be seen as an additional layer that is interwoven with the today's Web. Linked data provides a generic publishing method, which make it easier to discover and integrate data from a large scale of data sources. The Linking Open Data project (W3C SWEO Community Project, 2013) is an example of adoption and application of these Linked Data principles. This project was set up to bootstrap the Web of Data, by identifying existing data sets under open licenses, then converting these data sets to RDF according to the Linked data principles, and publishing them on the Web. Figure 2.9 demonstrates a number of data sets, recently published on the Web as Linked Data. Each node in the diagram represents a distinct data set published as Linked Data. The arcs represent the existence of links between items in the two data sets (thicker arcs indicate a greater number of links, bidirectional arcs means the outward links to the respective other exist in each data set).

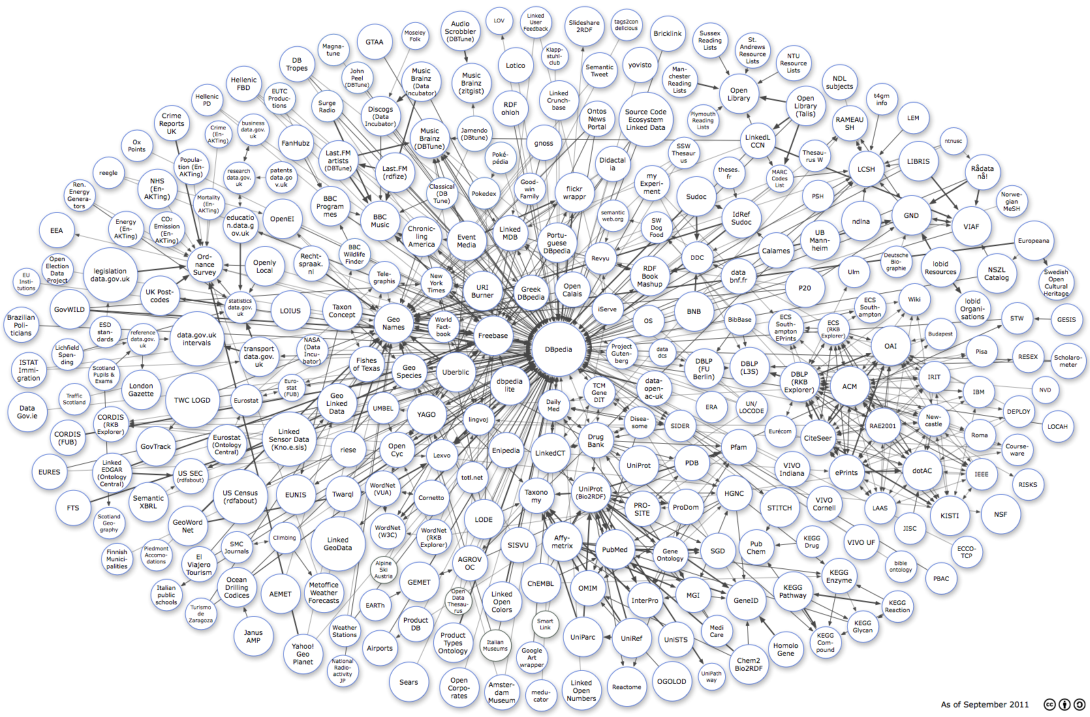


Figure 2.9: Linking Open Data cloud as of September 2011 (Cyganiak and Jentzsch, 2011).

In summary, we see the potential of using the Semantic Web technologies and semantic data, in order to create useful information to support users' judgements of the trustworthiness of Web information. In the next section, the term of trust is discussed, which is the main consideration in this work.

2.3 Trust

Trust is an important factor to be considered in the process of consuming data. This is particularly true in the Web and Semantic Web environments, which are decentralised and have little control over publishing quality. Unreliable data can lead to users making incorrect decisions. For example, a user might follow incorrect advice regarding a product or a treatment. Therefore, we should pay attention to trust and have an idea of what trust is and the benefits we can obtain from it. We state the definition of trust in section 2.3.1. Next, in section 2.3.2, we explain the levels of trust. Then, in section 2.3.3, we discuss the factors that influence trust in information. Finally, in section 2.3.4, we describe the elements that are required to establish trust.

2.3.1 Definition of Trust

The term, "trust", has been used in a wide variety of disciplines, ranging from philosophy to sociology, psychology, and computer science. Each discipline considers trust based upon differing criteria, dependent on the context. For example, sociologists tend to define trust as being structural in nature, some psychologists have considered trust as an attitude of a person towards the information, whereas economists are more likely to view trust as an optimal choice methodology (McKnight and Chervany, 1996). Accordingly, trust can be variously defined as follows:

- Trust is "assured reliance on the character, ability, strength, or truth of someone or something" (Marriam-webster, 2011).
- Trust is "confidence in or reliance on some quality or attribute of a person or thing, or the truth of a statement" (OED Online, 2011).
- Trust is "the extent to which one is willing to ascribe good intentions to and have confidence in the words and action of the others" (Cook and Wall, 1980).
- Trust is "an agent's behavioral reliance on another person under a condition of risk" (Currall and Judge, 1995).
- Trust "indicates the willingness of an agent to engage in a transaction in the absence of adequate safeguards" (Noorderhaven, 1995).
- Trust is "the expectation that arises, within a community of regular, honest, and cooperative behaviour, based on commonly shared norms, on the part of other members of that community" (Fukuyama, 1995).
- Trust is "the firm belief in the competence of an entity to act dependably, securely, and reliably within a specified context" (Grandison and Sloman, 2000).

- Trust "concerns a positive expectation regarding the behaviour of somebody or something in a situation that entails risk to the trusting party" (Marsh and Dibben, 2003).

All of these definitions define trust as a *subjective* issue that depends on the context in which the information is presented. In addition, trust is an attitude of the user towards a piece of information and their expectation of that information. This is a problem because, as a consequence of the subjective nature of trust, software may not necessarily be able to fully assess the level of trust a piece of information warrants. However, some work such as that of Golbeck et al. (2003) or Kelton et al. (2008) has been done which allows trust to be assessed based on more *objective* factors, which in turn allows us to make rational decisions regarding whether or not to trust something or someone.

Trust is an attitude towards the perceived trustworthiness of an individual. The individual's trustworthiness is a single property which defines the likelihood that someone or something will perform to expectations. Particularly, in the case of a piece of information, it can be said that information is trustworthy if it has been derived using effective means. Generally, trustworthiness is context-dependent; someone or something is trustworthy in specific respects (O'Hara, 2012). Therefore, trustworthiness can be used to support our opinion of whether or not to trust something or someone. This description is adopted for our work.

2.3.2 The Level of Trust

There has been much work on the study of trust, some of which discusses the levels of trust in a social and psychological context (Lewis and Weigert, 1985; Blomqvist, 1997; Schoorman et al., 2007). From these social and psychological perspectives, Kelton et al. (2008), defined four levels of trust described as follows:

- Individual level is "a personality trait" which can be addressed by the statement, "I trust."
- Interpersonal level is "a social tie directed from one to another." It extends the individual level statement to the statement, "I trust you."
- Relational level is "an emergent property of a mutual relationship." The statement to represent relational trust can be "You and I trust each other."
- Societal level is "a feature of a community as a whole." It can represent this trust level with the statement, "We all trust."

Therefore, the level of trust that is best suited to describe the behaviour that occurs in the information on the Web (or digital information as a whole) is interpersonal trust,

because the trust is a unidirectional relationship from the person to a piece of information or the person who created the information. For example, arguably, one can trust health information in an article published by a qualified medical doctor, but not vice versa (i.e. the doctor does not necessarily trust shomething you publish as a layperson).

In this research, we focus on trust at the interpersonal level which is a one-way attitude towards Web information. Therefore, in the remainder of this report, all trust that is discussed or described implicitly refers to *interpersonal* trust. Consequently, our framework for the trustworthiness of Web information discussed in the following sections and chapters is based on the concept of interpersonal trust.

2.3.3 The Influencing Factors on Trust in Information

Trust is a subjective issue, which is influenced by several external factors; there are (Kelton et al., 2008)

- The propensity to trust: a stable personality characteristic which influences one's willingness to extend trust to a person in a particular situation. Alternatively, it can be described as how trusting that user is.
- The context/domain: trust will be placed depending upon the context or domain in which the information is being considered. It can be expressed in the terms: "I trust you in Y context" or "I trust you to do Y." As trust is context-dependent; it brings further subjective elements, namely (O'Hara, 2012)
  - The interpretation of commitments. If one trusts another, then one must interpret the claims the 'another' make. For interpreting someone's claims, it is necessary to prove their intentions, capacities and motivations in the given context. For example, patients trust doctor A to treat them for sinusitis symptoms because they interpret the qualifications of the doctor as an otolaryngologist (as showing specialism in treating sinusitis).
  - Degree of confidence, which is the degree of belief of the trustworthiness of one in another: in other words, the confidence that one has in his/her mind about something. The degree of confidence is an important parameter in the analysis of trust because it can help someone to compare and then make judgments (e.g. a customer trusts shop A more than shop B. Therefore, the customer buys products from shop A rather than from shop B, who sells the same products). Furthermore, it helps to manage risk strategies (e.g. users will be willing to take a risk by investing their assets in a transaction which be operated by service A because the user trusts A).
  - Warrant is the positive or negative input or explanation to one's judgment. For example, customers would like to buy a product from company A, they

will tend to trust in the quality of A's product based on the commitment from company A to return money back to customers if they are not satisfied with the product.

- Social trust: trust will be affected by the reputation of the information's author or publisher.

2.3.4 Elements of Trust

In this research, we adopted the concept of composition to build trust from Kelton et al. (2008). Trust can be established based on a composition of two attitudes: confidence in that trustee will provide positive outcomes and a willingness of the trustor to react in expectation of the outcomes from the trustee. Therefore, the trustor's actions will be executed based on the elements of trust as described.

In conclusion, in this research, we define trust as an attitude towards the trustworthiness of a piece of information. Therefore, we focus on evaluating the trustworthiness of information in order to determine the elements that will be able to establish trust for the users of the information.

2.4 Trust on the Web

There has been a lot of research work undertaken about trust on the Web, focusing on issues ranging from information quality, and security, to credibility. Golbeck (2006) defined trust on the Web in three domains:

- Trust in Content: There is a massive amount of information on the Web. Everyone who connects to the Internet can publish any data on the Web. Thus, web users have to make a decision as to whether or not to trust the content when they access a page. The features which affect users' decisions of how much to trust websites are mostly based upon visual (the layout of web site, the graphic design, or the navigation on the Web) and social concepts (the reputation of site owners or recommendations from other users) (Cheskin Research and Studio Archetype/Sapient, 1999; Fogg et al., 2001b, 2003; Corritore et al., 2003).
- Trust in Services: On the Web, there are not only web sites but there are also automatic service applications, which connect to other applications or exchange information between each other. Therefore, services or agents should have a process in place to assess other services' trustworthiness in order to approve them for access and use of their information resources.

- Trust in People: One of the Web's characteristics is openness. As a result, information can come from a variety of sources. Nonetheless, the decision to trust information which comes from an unknown person is a difficult task. Social trust and reputation can help to address this problem. Users can define their trust value in other users and those scores can be accessed and accumulated to evaluate trustworthiness.

In other words, each domain considers trust in a specific factor depending on the domain. We focus on trust in content, which is a challenging issue due to the decentralised and distributed characteristics of Web data. The enormous amount of information which is published on the Web raises questions about the trustworthiness of that information. There are two areas of study that attempt to propose a solution to the assessment of the quality or credibility of the Web information, namely *information quality* and *Web credibility*. We discuss information quality in section 2.4.1 and Web credibility in section 2.4.2.

2.4.1 Information Quality

The concept of information quality (IQ) has been studied in a variety of areas such as information systems, information services, and databases. In each area of study, the researchers use the term 'information quality' in many different contexts. For example, in the information systems field, information quality may be defined as "*fitness for use*" (Tayi and Ballou, 1998), in which data that are considered as having sufficient quality for one use may not be of sufficient quality for another use; for example, a personal database which is created in different departments of a company may be correct but its content cannot be combined because they are stored in different formats. Therefore, even if the data are correct, the information can be considered of poor quality. Similarly, IQ can also be considered as "*user satisfaction*" (DeLone and McLean, 1992), which focuses on the user's attitude towards the information that is produced by the system. In addition, in the database domain, information quality is associated with accuracy of the information. Consequently, there is not a clear definition of information quality, but it can be considered that information quality is a set of criteria that can be used to assess the quality of the information that the users are consuming. Therefore, we conceive information quality as a combination of criteria or factors that can be used to measure or ensure that the information provided matches the users' expectations. We discuss the criteria that can be used to assess information quality on the Web in the next section.

2.4.1.1 Classification of Web Information Quality Criteria

Naumann (2002) classified criteria that are used to assess the quality of information into four sets:

- *Content-related*: the criteria in this set concern the basic characteristics of the actual data which are retrieved.
- *Technical*: the criteria in this classification determine the quality of data by measuring the performance of the soft- and hardware of the source, the network, and the users.
- *Intellectual*: consider the subjective aspects of the data.
- *Instantiation-related*: these criteria concern the presentation of the data.

In this research, we focus on the content-related set in which we consider the quality of the information based on the actual data we obtain. We choose this because the content of the data tends to provide solid evidence in order to make rational evaluation of the trustworthiness of the information.

2.4.1.2 Information Quality Criteria for the Web

There have been a number of studies on information quality in different contexts (e.g. data quality (Rieh and Belkin, 1998, 2000), or the value-added model (Taylor, 1986)). Each study proposed a variety of different sets of criteria that can be used to measure the quality of information.

Taylor (1986) stated the quality of information in his value-added model. His model can be considered to provide the most general framework, which can be used to develop the evaluation of information systems, information, and data. He defined five quality values for his value-added model as follows:

- Accuracy: The data and information should be error-free
- Comprehensiveness: the completeness of coverage of a particular subject or discipline
- Currency: The recency of the data
- Reliability: The consistency of quality of the system and its outputs over time
- Validity: the degree of acceptability of the data or information

From the definitions above, it is inferred that accuracy, currency, reliability, and validity are associated with data or information or outputs of a system, whereas comprehensiveness is related to the information systems.

In a survey, Rieh and Belkin (1998) asked scholars who are experienced or expert in the evaluation of the quality of information on the Web about the factors which they

use to judge the quality of information. The researchers identified seven facets of the assessment of information quality, described as follows:

- Source: Considers the characteristics of the place or the resources that data comes from. Moreover, the source can be considered on two levels:
  - The institutional level, which involves the characteristics of the institution: for example, the URL (whether the domain is "edu", "gov", or "com"), the particular type of the institution (e.g. information from a university, college, or museum is more likely to be trustworthy or correct), and the reputation of the institution.
  - The individual level, which involves the features of the author or creator of the information: for instance, the identification of the author (their name, contact address, and institution name), or the author's affiliation (the position of author/creator, occupation).
- Content: Whether the information on the web is useful for the users
- Format: The graphic design and information structure
- Presentation: The writing style
- Currency: The recency of information
- Accuracy: The accuracy of the information presented and the reliability of links on the Web
- Speed of loading: The length of time taken to retrieve the information

In 2000, Rieh and Belkin (2000) undertook more studies on the judgment of information quality by scholars when they are interacting with information on the Web. This study collected the data based on actual searching behaviours. They concluded six major categories of criteria for evaluating the information quality and cognitive authority as follows:

- Characteristics of Information Objects: The categories that were used to identify the characteristics of information objects are described as follows:
  - Types of information objects, such as a book, dissertation, journal article, personal homepage, or others
  - Title: the title indicates an overview or review of the issue
  - Content focused on things on the page
  - Organisation or structure of the page

  - Presentation considers the way that the page is written such as no misspelled words, good syntax, no typos
  - Graphics considers the way that the page is designed in term of graphic arts
  - Functionality considers the functions in the page and whether it works as intended

- Characteristics of sources: The categories that were used to identify the characteristics of information objects are described as follows:
  - URL type was considered based on the top-level-domain such as .org, .gov, .edu.
  - The type of source was considered based on the organization from which the information was produced.
  - The reputation of the source was considered based on how well-known the source of the information is.
  - One or a collective source considered whether the information was produced from a single person's opinion or by a group of people.
  - The author's or creator's credentials were considered based on the author's or creator's affiliation.

- Knowledge is based on the user's own personal experience or familiarity with a source or the information itself. It was categorised in two facets: the type and mode of obtaining the knowledge:
  - Types of knowledge
    * Domain knowledge is associated with the knowledge of a topic area, including the source of the information in the domain.
    * System knowledge refers to the knowledge of system functionalities and system structures in a web site or other information systems available on the Web
  - Mode of obtaining knowledge
    * First-hand experience means the user knows the information based on their personal experience.
    * Second-hand knowledge means the user obtains the knowledge from others.

- The situation refers to the conditions under which the information is being considered.

- Ranking in search output

- General assumptions about the information. For example, users may not trust information which is provided by a salesperson because they assume the information he or she provides will be biased towards placing their product in a good light.

Tate (2010) proposed the information quality criteria for Web resources as follows:

- Authority is the degree to which a person or organisation is perceived as having enough valid knowledge to provide material on a given subject area. The elements that can be used to evaluate the quality of Web resources are the author's qualifications (which are assessed by his or her background, experience, and formal credentials related to the subject area), or the publisher's reputation.
- Accuracy is the degree to which the information is free from errors. The elements which can be used to indicate accuracy, such as the peer review process, or the reputation of the source.
- Objectivity is the degree to which the material conveys the neutral facts and is not influenced by personal feelings or other biases, for example, the intent of the organization or person in providing the information.
- Currency is the degree to which the information is up-to-date. The relevant elements here are: the date on which the material was first published, or the date on which the material was last revised, or the date on which the material was first placed on the web server (if applicable).
- Coverage is the scope of topics and the depth to which those topics are focused in a work. This can be considered based on the table of contents of a book, or an index or site map on the Web.

Even if these studies investigated and proposed different criteria to determine the quality of Web information, it is interesting to note that a number of criteria such as currency, accuracy, and authority appear three times across the four studies mentioned above. Therefore, it is interesting to investigate and analyse these criteria in the different domains to find out whether they have any features in common across those studies in order for us to derive the basic empirical criteria for assessing the trustworthiness of Web information for use in our framework.

2.4.1.3 Information Quality Assessment

Information quality assessment is the process of measuring the quality of a piece of information that is being consumed by the users and comparing the assessment results with the users' quality requirements (Bizer and Cyganiak, 2009). Information quality assessment consists of information quality assessment metrics (measuring an information quality based on quality indicators) (Pipino et al., 2005).

Information quality assessment metrics can be divided into three categories based on the type of quality indicator (Bizer and Cyganiak, 2009):

- *Content-based* metrics use the information content itself to be the indicators. Therefore, the metrics used to analyse the information content depend on the type of information being assessed. For example, in the case of HTML pages, information is in the form of natural language texts. Thus, it is possible to apply text analysis methods to calculate the assessment scores by matching terms or phrases against a document.

- *Context-based* metrics evaluate the quality of information using meta-information about the information content e.g. information about the creator of the information, or the date or time that the information was created. One important quality indicator for assessing the quality of information is meta-information about the identities of the information provider. Other meta-information that should be considered is the identities of the contributors and the publishers as well as the source of information.

- *Rating-based* metrics use explicit opinions of information consumers or domain experts on the information itself, information sources, or information providers. The assessment process calculates the score from the collected ratings.

In this research, we focus on the context-based metrics which base the assessment of the quality of Web information using metadata about that Web information. This is because the rating-based metric might be misled by biased opinions, and the content-based metrics are limited in assessing the information based on narrow aspects of considering using text analysis. In addition, content-based metrics requires time to train the system to assess the quality of information.

2.4.1.4 Evaluation of Information Quality on the Web

There have been several works on assessing information quality on the Web. These have proposed frameworks, models or application tools that can be used to evaluate the quality of Web information.

One of the information quality research domains has proposed an information quality assessment framework, which can be used to evaluate the quality of information on the Web in general or specific domain. Knight and Burn (2005) presented the IQIP (Identify, Quantify, Implement, and Perfect) model as an approach to handling the selection and implementation of the algorithm to evaluate the quality of retrieved information on an Internet search engine. Their model proposed a method to assess the quality of information based on criteria which were selected from the established information quality literature. In addition, it is proposed that those criteria are then quantified (given a value and ranking) within the context of three dimensions (the user, the environment, and the task). Then, Web crawler algorithms for those dimensions are created and also metadata

about the page it crawls are produced. In addition, the model considers the fit of the results from crawling to the users' needs. Therefore, it provides feedback to the crawler and improves its ability to continue crawling for relevant information. Similarly, Stvilia et al. (2007) proposed a general IQ assessment framework which considers the quality of information in different dimensions; namely, scope of IQ problems, related activities, and taxonomy of IQ dimensions organised in a systematic way based on sound theories and practice. The framework claimed to be able to be used as a valuable knowledge resource and to offer guidance for developing IQ measurement models in many different aspects. Correspondingly, Bizer and Cyganiak (2009) proposed the WIQA - Information Quality Assessment Framework - which is a set of software components that can be used by applications for processing information of an uncertain quality. The main concept of the framework is to gather information from the Web page. Particularly, users need go to a certain Web page before it can extract data from the page and store them in the data store of the system. After that, the collected data are filtered using a wide range of different quality-based policies. In addition, the framework presented a feature of generating explanations about the filtering process, which is provided to the users to help them understand why certain information is selected to present to the users.

Another research domain that gives attention to the quality of the information is that of Information Retrieval (IR), particularly in the search engine technology. The lack of enforceable standards regarding information publishing leads to an information correctness problem and lots of conflicting information, which is retrieved by the search engines. Consequently, these issues draw researchers' attention to the quality of the information. Bizer and Cyganiak (2009) not only proposed a framework to assess the quality of information; they also presented an application which employed their proposed framework to display the framework implementation in a real-world scenario. The application was implemented as a browser called, the WIQA browser. The WIQA browser extracts structure information on the web pages it visits; then stores this information together with provenance meta-information as a set of Named Graphs. Then, the browser uses filtering and an explanation engine to filter stored information and to generate explanations about filter decisions, which are then presented to the users.

Ramachandran et al. (2009) proposed a trustworthy and high-quality information retrieval system. The system provided an enhanced web search engine which provides the trustworthiness of search results. It used five factors (provenance, authority, age, popularity, and related links) to calculate the trustworthiness of Web information. In addition, Ramachandran's retrieval system used the WIQA framework (Bizer and Cyganiak, 2009) to rank the search results of their system.

Turning to another proposed application focusing on the trust domain, an evidence-based content trust model for spam detection was proposed by Wang et al. (2010). In their research, they used two types of evidence for detecting Web spam. The first is *text feature-based evidence*, which considers text in the page. For example, counting the

number of words in the page, can help to indicate that the page was spam by overflow with a number of popular words into the page (keyword stuffing). The second type of evidence is *information quality-based evidence*, which used the quality criteria (currency, availability, information-to-noise ratio, authority, popularity, and cohesiveness) to obtain data and to compute quality scores. Even though, the Web spam detector of this work can help users to filter Web spam based on rational evidence, the process of actually labelling spam was still manual (the researchers asked volunteer students to label the Web spam based on the rules of spam Web pages that the researchers provided them with). From studying Wang et al. (2010), we have derived an important requirement for our framework, which is an automated process of collecting different types of evidence from several sources. However, our work differs from that of Wang and his colleagues in that we are not manually labelling information. Instead, we automatically collect some metadata from several sources in order to create the supportive information to the users.

Based on the literature review, we have learned some important requirements for designing our framework, which are discussed in Chapter 3. We take concepts of evaluating the quality of Web information into account. However, we intend to provide a framework that can be used in practice to help users assess the trustworthiness of Web information.

Our framework does not consider the trustworthiness of Web information in only one perspective in IQ. We also consider the quality of a Web page; a topic which is discussed in more detail in the next section. In addition, we also propose a prototype which is implemented based on our proposed framework to show how well the proposed framework performs in practice. However, our work differs from the WIQA browser of Bizer and Cyganiak (2009) in that it will automatically gather information based on the search terms entered by the users. Therefore, users do not need go to a certain Web page, from which the tool can extract structure information. Moreover, it differs from Ramachandran et al. (2009) in that the results from our application provides some explanation together with the search results to make it clearer to the users as to why that information can be trusted.

Furthermore, our framework is not a detection tool that detects and discards untrustworthy information for the users (as was the case with Wang and his colleagues' work). Instead, our framework aims to train users to be more critical in evaluating the trustworthiness of Web information by providing them with basic information; and by also providing them with additional supportive information to help them assess the information. In addition, we adopt the concept of displaying an explanation from Bizer and Cyganiak (2009). The explanation gives more detail to the user and it enables users to have more confidence in using the system which is implemented based on the framework. In addition, the explanation will help them to learn about the critical factors that should be assessed when evaluating information.

2.4.2 Web Credibility

There is another concept which is similar to evaluating the trustworthiness of Web information for making a decision on whether to trust this information. This concept focuses on the credibility of information. Now, we discuss Web credibility in more detail to state the concepts relevant to trust.

2.4.2.1 The Definition of Credibility

Tseng and Fogg (1999) summarised that the word 'believability' and 'credibility' are used interchangeably in most cases. Therefore, it can be inferred that *credible* information is *believable* information. There have been several research projects about credibility which agree that the credibility of information should be evaluated in multiple dimensions for archiving obtained reasonable credibility assessment results. From this study, the majority of researchers proposed key components for evaluating credibility (namely trustworthiness and expertise). Tseng and Fogg (1999) defined "trustworthiness as a property of being well-intentioned, truthful, and unbiased." Therefore, the trustworthiness in this research focuses on the confidence in the source that produced the information. They also define "expertise as being knowledgeable, having experience, and being competent." This component focuses on the knowledge and skill of the source. Therefore, evaluating credibility should consider both trustworthiness and expertise in order to receive an overall credibility judgement.

2.4.2.2 The Types of Credibility

Tseng and Fogg (1999) proposed four types of credibility: presumed, reputed, surface, and experienced.

- Presumed credibility is based on general assumptions or stereotypes of the perceiver. An example of such a stereotype is that of a salesperson who is generally dishonest. Because of this assumption, people do not believe everything the salesperson says. This kind of credibility type judges the credibility of information based on general assumptions without any other indication.
- Reputed credibility assesses credibility based on the third party certification or reports. Example are, awards (such as the Nobel Prize), certificates (JAVA or Sun certificate), or official titles (such as Doctor and Professor). These certification from third parties increases the confidennce of the credibility of a person because they are verified by others.

- Surface credibility is judged based upon simple inspection. For example, professional dress (such as wearing a suit), and a well-designed book cover or Web page will be used as an indicator of credibility of that person or information.
- Experience credibility arises from users' direct experience with someone or something over time. Users then use this experience to evaluate the expertise or trustworthiness of subsequent statements or suggestions from those objects. For instance, customers who buy food from a shop and always get good quality food will attribute that shop with a high degree of credibility in food production.

Huynh et al. (2006) proposed four different types of trust and reputation: interaction trust, role-based trust, witness reputation and certified reputation.

- Interaction trust is based on past experience of direct interactions between the agents.
- Role-based trust is evaluated based on role-based relationships between the agents (evaluator and target). For example, the evaluator agent might trust any other agent that is owned, or certified, by the evaluator agent's owner.
- Witness reputation is built based on reports about the target agent's behaviour from other agents which interacted with the target agent. These reports will be used to derive the trustworthiness of the target agent from the views of its witnesses.
- Certified reputation is built from the certified references of the target agent itself which are provided to the interacting agent in order to gain the trust from them.

From the descriptions of Huynh et al. (2006), interaction trust was defined to be the same as the experience credibility of Tseng and Fogg (1999). Similarly, the witness reputation and certified reputation types from Huynh et al. (2006) mentioned the same idea of reputed credibility as proposed by Tseng and Fogg (1999), whereas role-based trust is a new definition. Accordingly, five types of credibility or trust can be categorised: presumed, reputed, surface, experience, and role-based.

2.4.2.3 Evaluating the Credibility of Web Sites

In this respect, the individual Web site is considered as the source of information. Therefore, the assessment of credibility in this context tries to evaluate the credentials of the Web site in question. This concept focuses on the source of the information.

There have been a number of studies on the Web site credibility issues proposed by Fogg and his colleagues in the Stanford Web Credibility Research project (Fogg et al., 2000).

They conducted an online survey to gather the comments and answers from more than 1,400 participants when they evaluated the credibility of Web sites. The participants were asked to read some randomised statements describing a Web site element. Then, they ranked on a scale of -3 to +3 how much certain elements of the page affected their belief in the site's credibility. The results from the survey concluded that five domains affect the credibility of a Web site; there are real-world presence, errors on the web, user interface/navigation, advertisement, and technical problems. In a subsequent study, they investigated specific factors which affect a Web site's credibility (Fogg et al., 2001a). They focused on banner advertisements, author photographs and names. The results show that a low-reputability banner advert reduces the credibility of Web content. Alternatively, the author's photograph had significant effects on the credibility of the article (increasing the credibility of the Web page in the case of a formal photograph). In contrast, the name of the author had limited effects. The results from this study shown that naïve users tend to employ the subjective criteria in order to make a decision. In another study, Fogg and colleagues (2003) set up a survey to gather comments from more than 2,600 participants in order to ascertain the factors employed by users in assessing the credibility of a Web site on several topics such as health, news, travel, and business. The results identified a number of features which were noticed when users evaluated the credibility of a Web site: "*design appearance, information design/structure, information focus, company name, usefulness of information, accuracy of information, name recognition and reputation, advertising, bias of information, tone of the writing, identity of site sponsor, functionality of site, customer service, past experience with site, information clarity, performance on a test, readability, and affiliations* (Fogg et al., 2003)." The feature which was mentioned most frequently was the appearance of the Web site. The second most common features were information structure and information focus. Specifically, the top ten issues which participations considered can be grouped in the following ways: five issues concerning the provided information; three that focus on the design; and two that concern the source characteristics.

Wathen and Burkell (2002) reviewed the literature related to the credibility of information and proposed a model to evaluate the credibility of on-line information. The model consists of three phases of assessment. The first stage evaluated the credibility of the medium itself based on its surface characteristics (such as the presentation, interface design, organisation of information). The next step assessed the source and information in the Web site (such as the author's credentials or expertise; currency; accuracy). The final step judged credibility based on self-knowledge of the users' own expertise, domain knowledge, and information need. The assessment process of their model is iterative in that, if the results from any evaluation are negative, the user is likely to leave the Web site and find a new one, whereas, if it passes, then the user will move on to consider the next phase. However, the evaluation process of the user has an exception to skip from the surface assessment to the content if the users have high level of need for the

information. In conclusion, the information itself has a significant influence; more so than the appearance of the Web site.

Princeton Survey Research Associates studied the factors which influence users when evaluating the credibility of a Web site and choosing to visit it. From their study, it was seen that users have different credibility standards for different types of site. However, there were factors that made users think a website is credible, dependent on the type of website (Associates Princeton Survey Research, 2002). For example,

- E-commerce site
  - A statement of all fees that users will be charged for using the site (shipping cost, transaction fees and handling fees)
  - A statement of how the site will use users' personal details such as their name, address, credit card number, etc.
  - An explanation of the expected delivery date or confirmation of the users' reservations
  - A statement of the site's policies for returning unwanted items or cancelling a reservation
  - The contact address of the site's staff in case the users have any problems
  - The site's privacy policies
- News, information web site
  - The site's privacy policy
  - Advertising is clearly labelled and distinguished from news or information on the site
  - The contact address of a person who is responsible for the content on the site
  - A prominently displayed page for corrections and clarifications
  - The name of people who are responsible for the content on the site
  - The financial relationships between the site and other sites

Nine key factors in deciding to visit a Website and assessing its credibility are (Associates Princeton Survey Research, 2002)

- The site is easy to navigate and to find what you want
- Being able to trust the information on a website
- Being able to easily identify the sources of information on a website
- Knowing the website is updated frequently with new information

- Being able to find out the important facts about a website
- Knowing who owns the website
- What business and organisations financially support the site
- The site displays seals of approval from other groups
- The site displays awards and certificates from other groups

Alternatively, Miyamori et al. (2008) proposed an evaluation approach which assesses the information credibility of Web information based on four criteria (content, sender, appearance, and social valuation). In addition, an evaluation approach comprises manually-annotated data based on evaluation criteria. Moreover, the researchers also proposed the prototype system, WISDOM, which is an information credibility analysis system based upon natural language processing. The WISDOM prototype provides credible information from different perspectives based on assessing and judging information from their data evaluation process. The WISDOM system needs to collect web pages using a spider which it then stores in its local storage. Then, the stored web pages are analysed and classified based on their criteria as mentioned. The results of analysis are stored with tags in the form of XML data. The users can locate credible information on their specific topics of interests by inputting a topic keyword with the browser. The results will show a list of related Web pages which were retrieved from their analysis result storage, classified by content, sender, concept, or opinion. In contrast our framework aims to support users to evaluate the trustworthiness of Web information without the software making a decision on behalf of the users, but to instead provide additional information that can help to support the user's decision. Moreover, our work does not require a spider function to gather Web pages. In addition, we provide an explanation to the users in order to give them more confidence as to why the information should be trusted.

2.5 Conclusion

In this chapter, we described the definitions of the Web and the Semantic Web. Moreover, we explained the architecture of the Semantic Web and the base technologies that we use in our work in order to provide background knowledge. Subsequently, we discussed the definition of trust which has been adopted in a wider domain.

In our work, we define trust as an opinion held by the user as to the likelihood that a Web resource is trustworthy. We also explained the challenge of applying trust on the Web. In addition, we discussed information quality and Web credibility, which is a concept related to trust. We also discussed the definition of information quality and Web

credibility, and the evaluation methods for evaluating websites and Web information in each study.

We discussed that while prior work has discussed the assessment of the quality and credibility of web information, the work thus far has been lacking insofar as, for some criteria, it is hard to collect (these data might not be provided) or it is subjective information that cannot be gathered from the website, such as the content's bias or intention. In addition, some approaches are time-consuming for gathering useful metadata (need to open a certain page before being able to extract metadata) or labelling the data manually in order to use these labels for filtering the trustworthy information from untrustworthy information and present this information to the users.

From the review, we employ the concepts from two research areas of trust on the Web information quality (detailed in section 2.4.1) and Web credibility (detailed in section 2.4.2) to design our framework. We introduce a framework which collects the metadata of information based on the criteria that can be used to support users' assessments of the trustworthiness of Web information including the explanation for why a piece of information should be trusted. We use RDF and SPARQL as base technologies to store, integrate, and query metadata. The use of RDF provides other benefits, as well as describing features in the Web. Moreover, it also provides features to increase the efficiency of implementing an application using the Web. Therefore, it is a good opportunity to use RDF and SPARQL, which are basic technologies in the Semantic Web, to address the issue of trust. We see the opportunity to fill the gap in the literature by evaluating the trustworthiness of Web information using the Semantic Web technologies. In the next chapter, we discuss the details of the construction of our framework.

Chapter 3

Development of the Trustworthiness of Web Information Evaluation Framework (TWINE)

In this chapter, we describe the derivation of the architecture for the Trustworthiness of Web Information Evaluation framework (TWINE). This framework is a conceptual application framework that aims to help Web users evaluate the trustworthiness of the Web information they use, and acts as a supporter that gathers and provides useful information that can support users' judgments of the trustworthiness of Web information.

We describe the functional architecture of the TWINE framework in section 3.1. In section 3.2, we explain the construction of the presentation layer. Then in section 3.3, we discuss the construction of the logic layer of the TWINE framework; specifically, we present the trustworthiness criteria that we use in our framework. Each criterion consists of the items, called indicators, in order to indicate the trustworthiness of Web information based on their criteria. These indicators can be modified in practical detail in order to provide flexible features to adopt for use in different areas. In addition, they can be changed and further investigated in future studies as the Semantic Web provides more information or as new techniques emerge. Finally, we draw the chapter to a close by summarising the architecture of the TWINE in section 3.4.

3.1 The Functional Architecture of the TWINE Framework

The purpose of this research is to provide a framework that can be used to implement practical tools that can help web users to evaluate the trustworthiness of Web information in a rigorous and easy way. Our framework consists of two layers, which are the presentation layer and the logic layer as shown in Figure 3.1:

- The presentation layer is the layer that interacts with end users. It provides an interface for accepting the search term from the users and displaying the search results with supportive information to the users.
- The logic layer is the layer that gathers metadata based on the trustworthiness criteria, and integrates the collected metadata to build integrated metadata graphs.

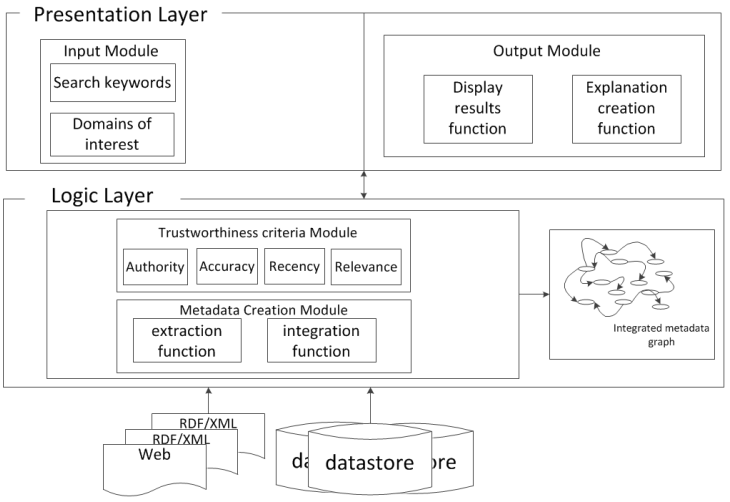


Figure 3.1: The Trustworthiness of Web Information Evaluation framework(TWINE)

More details of each layer are discussed in the following sections.

3.2 The Construction of the Presentation Layer

In order to allow the framework to collaborate with the users and the logic layer, we need a layer that works as a medium between end users and the application. This is the presentation layer, which consists of input and output modules.

- Input module: we designed our framework to integrate the features of a search engine. This allows Web users to look for information based upon key terms. The

search results also provides supportive information along with the search results to support their judgment. Consequently, a method to automatically find the information from specific keywords is added into the framework under the "input module". Furthermore, a domain of interest for the information is also important because it can help to refine the returned information such that it is relevant to the user's interests, and to determine the indicators in criteria that are used to gather information for display to the users.

- Output module: the result display is also an important part of the framework because it is the part that communicates to the users. Specifically, it needs to present the gathered supportive information in a meaningful and understandable way to users in order for them to be able to use this information when assessing the trustworthiness of Web information. In addition, the explanation as to why a piece of information should be trusted is also helpful for supporting web users' evaluations of the information. We therefore add to our framework the "output module" whose job it is to display search results with supportive information including the explanations.

In the next section, we discuss the logic layer, which works as a background process to provide results and useful supportive information.

3.3 The Construction of the Logic Layer

The logic layer consists of two modules; namely, the trustworthiness criteria and metadata creation modules:

3.3.1 Trustworthiness Criteria Module

This module provides the criteria that are used for gathering data in order to create the metadata for supporting the user's evaluation. Our framework suggests that the trustworthiness of Web information can be assessed by four criteria:

- *Authority* indicates the reputation of the source that produced the content. It can be considered on two levels: the institutional and the individual level.
- *Accuracy* is based on how accurate is the expressed information, or on information regarding the editorial process through which the information must pass before it is published; for example, whether or not the information has been peer reviewed.
- *Recency* focuses on how recently the information was created or modified.

- *Relevance* indicates whether the content meets the user's needs, (i.e. whether or not it is useful for them). This can be assessed by looking at the title of the information or the references of the information.

Each criterion consists of several indicators that are used to assess the trustworthiness of Web information. In addition, these indicators can be changed if the Semantic Web provides more information or new techniques are proposed.

The indicators of each criterion describe the terms that can be evaluated in the practical works. Therefore, the combination of criteria and their indicators in the framework provides a conceptual method to evaluating the trustworthiness of Web information. An overview of the trustworthiness criteria module in the TWINE framework is shown in Figure 3.2.

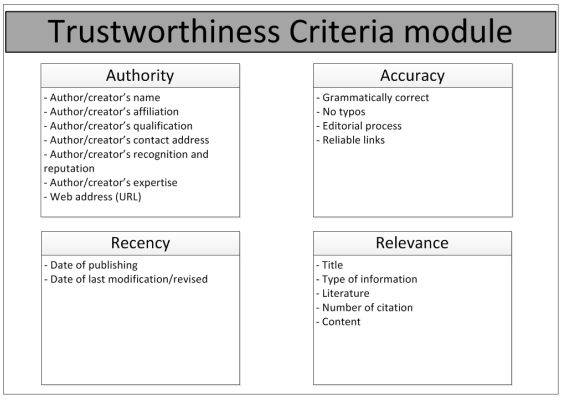


Figure 3.2: This diagram presents the trustworthiness criteria module of the TWINE framework. Each criterion is shown at the top of each box. Indicators related to each criterion are listed in the box.

3.3.2 Metadata Creation Module

A main purpose of this research is to develop a framework that will help Web users to evaluate the trustworthiness of Web information based on the proposed criteria, it is necessary to have a module that can gather the supportive information based upon the proposed criteria automatically. Accordingly, this module gathers metadata that are useful either from the Web itself or other sources that provide useful data by extracting or querying from those sources. The process of extracting and querying data from the aforementioned sources must take into account the provenance of that information as well as the information's content. The provenance information indicates the quality of the metadata by stating by whom the gathered metadata is asserted or quoted. After collecting metadata, this gathered metadata are integrated in order to create a metadata graph in the form of RDF graphs. The results from this module provides an integrated metadata graph which combines the gathered supportive information.

3.3.3 An Integrated Metadata Graph

An integrated metadata graph is built from the trustworthiness criteria and metadata creation modules. We use a Named Graph technique in order to create our metadata graph. In addition, each metadata graph also incorporates the provenance information of itself and are attached to the metadata when it is built as a metadata graph. This information can help to secure the accuracy of the metadata.

3.4 Conclusion

In this chapter, we illustrated the architecture of the Trustworthiness of Web Information Evaluation framework (TWINE). We described the modules of the framework and their respective functions. TWINE was developed to support web users in order to evaluate the trustworthiness of web information by providing the supportive data which is gathered and integrated based on trustworthiness criteria. The key purpose of the framework tries to fill the gap between conceptual guidelines to help users assess the trustworthiness of Web information and the practical action from users in the real world. It can help web users to evaluate Web information in a more rigorous way and also to educate them to realise the importance of considering the trustworthiness of the Web information they use.

The literature review from Chapter 2 provided the foundation ideas of criteria for TWINE through the information quality and Web credibility studies. We synthesised the criteria and indicators from these two researches (the details of the derivation of our criteria will be discussed in Chapter 4). In our research, the trustworthiness of Web information can be determined by four criteria: authority, accuracy, recency, and relevance. Each criterion in the framework consists of individual indicators, which can be edited and evaluated in empirical studies.

The framework consists of two layers: the presentation layer and the logic layer. The presentation layer is composed of input and output modules which are the modules that interface with the users. The logic layer consists of the trustworthiness criteria and metadata creation module that gathers and integrates metadata to support users' evaluation based on the trustworthiness criteria. In the next chapter, we discuss in more detail the process of developing our trustworthiness criteria.

Chapter 4

Development of the Trustworthiness Criteria

In Chapter 2, we discussed studies from two domains about trust in the Web information environment. One is Information Quality which concerns the quality of the information that is provided on the Web. The other is Web credibility, which focuses on the reliability of the information provider and the information itself. Both studies present several criteria that can be used as guidance for assessing the trustworthiness of a website. However, the criteria from each piece of research have their limitations; for some criteria it is hard to collect metadata directly from the website (this data might not be provided) or it is subjective information that cannot be gathered from the website, such as subjective bias or intention in the content. This chapter presents the details of the process of analysing the trustworthiness criteria from the two aforementioned domains and the selection of criteria to be used in our framework in the following sections.

4.1 Summary of the Criteria from the Information Quality and Web Credibility Research Areas

In this section, we investigate and analyse the criteria from the two areas of study discussed in sections 2.4.1 and 2.4.2, which emerged from different topics and participants, but which have much common ground to support their integration. We select the common criteria between information quality and Web credibility studies. The fact that these criteria are common to both domains indicates that they can be used within different settings in order to evaluate the trustworthiness of the information on the Web. Consequently, we obtain two lists of criteria and indicators from each research area, which are displayed in Table 4.1 and Table 4.2.

Table 4.1: A synthesised list of criteria and indicators derived from the "information quality" research area

Criteria	Indicators	Source
Accuracy	- Grammatically correct - No misspelled words - No typos - Comprehensiveness (table of content/ site map) - System reliability (functionality) - Validity - Page structure - Graphic design - Presentation - Speed of loading - Reliability of links	Taylor (1986) Rieh and Belkin (1998) Rieh and Belkin (2000) Tate (2010)
Authority	- Web address (URL) - Author's name - Author's credentials - Author's affiliation (position/occupation) - Author's qualification (title) - Author's contact address - Type of sources - The reputation of the source or publisher	Rieh and Belkin (1998) Rieh and Belkin (2000) Tate (2010)
Currency	- First published - Last modification/revised - Up-to-date (recent)	Taylor (1986) Rieh and Belkin (1998) Tate (2010)
Characteristics of the information	- Content - Type of information objects (book, journal, article, etc.) - Title - Ranking in search output (predictive) - Tone - Writing style - Intention of providing information - Information structure - Literature	Rieh and Belkin (1998) Rieh and Belkin (2000) Tate (2010)
Preference	- Users' own personal experience of	Rieh and Belkin (2000)

Table 4.1: A synthesised list of criteria and indicators derived from the "information quality" research area

Criteria	Indicators	Source
	a topic area or system - General Assumption (e.g., users may not trust information which is provided by a salesperson because they assume the information they provide will be biased towards placing their product in a good light.) - The conditions for considering the information	Tate (2010)

Table 4.1 summarises the synthesised list of criteria and indicators from the research in section 2.4.1 on information quality. There are a total of five criteria, the first of which considers the *accuracy* of the presented information, which includes aspects such as the spelling of words, grammatical accuracy, editorial process (i.e. peer-review), information layout, and system functionality (e.g. how well the website on which the information is hosted works). The second criterion is *authority*, which focuses on the source of information. For example, the author's or creator's name, affiliation, and contact details; the type of source (i.e., book, journal, webpage); and the information's Web address (URL). The third criterion, *currency*, is concerned with how recent the information is. Therefore, it considers the date and time at which the information was published or the last time the publication was modified. The fourth criterion focuses on *characteristics* of the information. These characteristics are used to evaluate the quality of the information based on the information itself, and include the type, title, writing style, tone, search engine ranking, structure, and literature (both cited and citing) of the information. The last criterion, the *preference* criterion, judges the information based on the Web user's own experience on the topic and system, and the user's general assumptions about, and requirements from, the information.

Table 4.2: A synthesised list of criteria and indicators derived from the "Web credibility" research area

Criteria	Indicators	Source
Surface	- Presentation (colour, design, font) - Organisation of information - Tailoring - Layout - Images	Wathen and Burkell (2002) Fogg et al. (2003) Associates Princeton Survey Research (2002)

Table 4.2: A synthesised list of criteria and indicators derived from the "Web credibility" research area

Criteria	Indicators	Source
	- Readability - Display page	
Source	- Expertise - URL suffix - Name recognition and reputation - Identity of site operator - Author's or creator's affiliation - The author's contact address - The person's or organisation's name who is responsible for creating, or maintaining the information	Wathen and Burkell (2002) Fogg et al. (2003) Associates Princeton Survey Research (2002)
Content	- Matching a user's previous knowledge - The level of requirement - The ease of applying the information to a user's situation - Information bias - Writing tone	Wathen and Burkell (2002) Fogg et al. (2003)
Commercial	- Advertising - Customer service - Statement of fee - Privacy policy - Service policy - Contact address - Financial relationships	Fogg et al. (2003) Associates Princeton Survey Research (2002)
Usability	- Navigation - Download speed - Information structure - Site functionality (speed of processing/ loading) - Past experience of a user - A user's own test set	Wathen and Burkell (2002)) Fogg et al. (2003)
Currency	- Up-to-date	Wathen and Burkell (2002)
Accuracy	- No error/typos	Wathen and Burkell (2002)

Table 4.2: A synthesised list of criteria and indicators derived from the "Web credibility" research area

Criteria	Indicators	Source
	- Information accuracy - Information clarity	Fogg et al. (2003)
Relevance	- Matches the user's needs - Information focus - Company or author motives - Information usefulness	Fogg et al. (2003)

Table 4.2 summarises the synthesised list of criteria and indicators from Web credibility research, which can be categorised into eight criteria. First, the *surface* criterion focuses on the layout and presentation of the information. Second, the *source* criterion evaluates the credibility of the information based upon attributes of the source of the information such as its URL, the identity of the site operator and the person's name whose responsibility it is to maintain the information. Third, the *content* criterion considers the trustworthiness of information based upon the information itself, the preferences of the user and the user's past experience. Fourth, the *commercial* criterion focuses mainly on advertisements within the site, customer service and payment processes. Fifth, the *usability* criterion considers the performance of the website in terms of its ability to serve the information and its functionality such as ease of navigation, download speed and information structure. Sixth, the *currency* criterion is mainly focused on how up-to-date the information is using the proxy of the date of publication of the information. Seventh, the *accuracy* criterion is focused on then clarity and accuracy of the information, and finally, the *relevance* criterion considers how well the topic of the information matches the user's need, and the overall usefulness of this information.

Nonetheless, each unique set of criteria presented in the different pieces of research as mentioned has its limitations (e.g., it is hard to collect the information based on that criterion directly from the Web; or it only slightly reflects the credibility of the information content itself). Therefore, we have to select the criteria that can be used in practice and that have a significant impact on the evaluation of the trustworthiness of Web information. We discuss the details in the following section.

4.2 Synthesis of Criteria from the Summarising Lists

We analysed and synthesised the criteria from the studies into information quality and the credibility of Web resources that were discussed in section 2.4. We produced a set of two lists containing the summary of criteria and indicators in each study as displayed in Table 4.1 and Table 4.2 from section 4.1.

From Table 4.1 and Table 4.2, it can be seen that some of these criteria can be adopted for implementation in practice, such as the currency criterion, that assesses the information based upon the information's last modification timestamp and the authority of the information, which can be evaluated based on the author or creator's details. However, some of the proposed criteria require data that are difficult to gather. Furthermore, some criteria do not have a significant impact on the trustworthiness of information.

Therefore, an initial assessment of which criteria and indicators should be included or excluded from the framework must analyse how practical they are to implement and how significant their effects are in helping Web users evaluate the trustworthiness of Web information. As an example, consider the indicators of the preference criterion in Table 4.1, these indicators are based on subjective factors such as the user's experiences, and assumptions, and the circumstances under which the user is considering the information. In addition, the indicators of the surface criterion in Table 4.2 are characteristics which can be disguised easily by using professionally-designed templates offered with content management platforms. Consequently, it is easy to make any information look good, to give the (possibly false) impression of trustworthiness. In addition, some indicators do not have significant impact on the trustworthiness of a piece of information. For example, a component of the usability criterion in Table 4.2 is the download speed of the document, which is the time the document takes to load. This may indicate the performance of the system and may influence a user's perceived trust of that information but it does not reflect on the information itself. Similarly, the commercial criterion in Table 4.2 is focused on the advertising and financial characteristics of the information. Hence, they are more likely to negatively affect the trustworthiness of information (Fogg et al., 2002) because mostly advertisements try to selectively describe positive aspects of the product while actively hiding negative points. The articles that are designed to advertise a product may be deemed less trustworthy by users.

As a result, we exclude some criteria and those kinds of indicators from the synthesised list. In summary, we exclude one criterion from Table 4.1 (the preference criterion), and four criteria from Table 4.2; surface, content, commercial, and usability criteria. We then generate a new list of indicators as shown in Table 4.3 and Table 4.4.

Table 4.3: A new synthesised list of criteria and indicators in information quality research which excludes the preference criterion

Criteria	Indicators	Source
Accuracy	- Grammatically correct - No misspelled words - No typos - Editorial process - Reliability of links	Taylor (1986) Rieh and Belkin (1998) Rieh and Belkin (2000) Tate (2010)
Authority	- Web address (URL)	Rieh and Belkin (1998)

Table 4.3: A new synthesised list of criteria and indicators in information quality research which excludes the preference criterion

Criteria	Indicators	Source
	- Author's name - Author's credentials - Author's affiliation (position/occupation) - Author's qualification (title) - Author's contact address - Type of sources - The reputation of the source's or publisher's	Rieh and Belkin (2000) Tate (2010)
Currency	- First published - Last modification/revised - Up-to-date (recent)	Taylor (1986) Rieh and Belkin (1998) Tate (2010)
Relevance	- Content - Type of information objects (book, journal, article, etc.) - Title - Ranking in search output (predictive) - Literature	Rieh and Belkin (1998) Rieh and Belkin (2000) Tate (2010)

Table 4.4: A new synthesised list of criteria and indicators of Web credibility research which excludes the surface, content, commercial, and usability criteria

Criteria	Indicators	Source
Source	- Expertise - URL suffix - Name recognition and reputation - Identity of site operator - Author's or creator's affiliation - The author's contact address - The person's or organisation's name who is responsible for creating, or maintaining the information	Wathen and Burkell (2002) Fogg et al. (2003) Associates Princeton Survey Research (2002)
Currency	- Up-to-date	Wathen and Burkell (2002)

Table 4.4: A new synthesised list of criteria and indicators of Web credibility research which excludes the surface, content, commercial, and usability criteria

Criteria	Indicators	Source
Accuracy	- No error/typos	Wathen and Burkell (2002)
	- Information accuracy	Fogg et al. (2003)
Relevance	- Matches the user's needs - Information focus - Information usefulness	Fogg et al. (2003)

These criteria and indicators from Table 4.3 and Table 4.4 are used in the next section to generate the criteria used in our proposed framework.

4.3 Generating the Criteria

In this section, the criteria and indicators for use in our framework are generated, using the two lists of criteria and indicators from section 4.2. We arranged the criteria and indicators from Table 4.3 and Table 4.4 into groups, shown in Table 4.5 for comparing and analysing in the next step. The table comprises two columns: column A (indicators from information quality) and column B (indicators from Web credibility).

We analysed these eight criteria to discover patterns and similarities between them. From Table 4.5, it can be seen that some similarities exist in the *meanings* of the criteria, even if they are referred to by different names. For instance, the *authority criterion* in column A (information quality) and *source criterion* in column B (Web credibility) represent the details or identification of the source of the information. Similarly, the *characteristics of the information criterion* in the information quality column and the *relevance criterion* in the Web credibility column both express the relevance of the information to the user's requirements. Therefore, we reclassify these two sets of criteria into new criteria, namely the *authority* and the *relevance* criteria respectively. The remaining criteria *accuracy* and *currency*, have no cross-over in meaning between information quality column and Web credibility column. However, we named the *currency* criterion with the new name, *recency* criterion, in order to best represented the criterion. We use the indicators in these two criteria as they stand.

The main concept of our framework is to suggest rational criteria that can help users to evaluate the trustworthiness of Web information. Consequently, given the discussion above and the results from Table 4.5, we select four criteria for use in the TWINE framework:

- An authority criterion relates to the author's identification and credentials.

Table 4.5: A comparison of criteria and indicators from information quality and Web credibility research areas

Column A Indicators from Information Quality research		Column B Indicators from Web Credibility research	
Accuracy	- Grammatically correct - No misspelled words - No typos - Editorial process - Reliability of links	Accuracy	- No error/typos - Information accuracy
Authority	- Web address (URL) -Author's name - Author's credentials - Author's qualification (title) - Author's affiliation (position/occupation) - Author's contact address - The source's or publisher reputation - Type of sources	Source	- Expertise - URL suffix - Affiliation - Identity of site operator - Name recognition and reputation - The author's contact address - The responsible person's name
Currency	- First published - Last modification/revised - Up-to-date (recent)	Currency	- Up-to-date
Characteristics of Information	- Type of information objects (book, journal, article, etc.) - Content -Title - Literature - Ranking in search output (predictive)	Relevance	- The familiar of a user with the topic - Information focus - Information usefulness

- An accuracy criterion relates to the accurate expression of the information.
- A recency criterion relates to how up-to-date the web information is.
- A relevance criterion relates to the matching between content and user's needs.

We then examined and synthesised indicators in each criterion of these two research areas. An initial analysis identified that some indicators have a direct mapping between the two domain areas. Some indicators are unique insofar as they only appear in one column. However, some indicators can be merged to best represent the meaning of the indicators.

From the list of indicators in Table 4.5, we synthesised indicators in each criterion based on the categories as described above. Firstly, for the accuracy criterion we selected two indicators from the information quality column and added all unique indicators in this criterion into a new column, called column C, the synthesised indicators as shown in Table 4.6.

Table 4.6: A synthesised list of unique indicators from accuracy criterion

Column A Indicators from Information Quality research		Column B Indicators from Web Credibility research		Column C Synthesis of indicators	
Accuracy	- Editorial process - Reliability links	Accuracy		Accuracy	- Editorial process - Reliability links

Now, we consider the indicators which might have a direct mapping between these two research areas. We found that both the information quality column and Web credibility column have the indicator, "no typos". In addition, in the information quality column, it has the indicator "no misspelled words" which has the same meaning as no typos. Thus, we grouped them together and added the resulting indicator ("no typos") into the synthesised indicators in column C as shown in Table 4.7.

Table 4.7: A synthesised list of direct mapping indicators from accuracy criterion

Column A Indicators from Information Quality research		Column B Indicators from Web Credibility research		Column C Synthesis of indicators	
Accuracy	- No misspelled words - No typos	Accuracy	- No error/typos	Accuracy	- No typos

The remaining indicators were combined to create a new component depending on similarity in meanings or those that best represented the indicators. An indicator, called grammatically correct, in Table 4.8 column C is the combination of the indicators "grammatically correct" from column A and "information accuracy" from column B. The term "grammatically correct" was selected to represent a factor that reflects the accuracy of information based on how accurately it is expressed.

Table 4.8: A synthesised list of combination indicators from accuracy criterion

Column A Indicators from Information Quality research		Column B Indicators from Web Credibility research		Column C Synthesis of indicators	
Accuracy	- Grammatically correct	Accuracy	- Information accuracy	Accuracy	- Grammatically correct

For further examination, we used the same procedure for creating the indicators in remaining criteria. As a result, we produced a synthesised list of indicators for each criterion as shown in Table 4.9.

In summary, the alignment process has produced four criteria and 18 indicators as potential criteria and indicators for inclusion in our proposed framework as shown in Table 4.10.

Table 4.9: Alignment of the synthesised criteria and indicators from two research areas

Column A Indicators from Information Quality research		Column B Indicators from Web Credibility research		Column C Synthesis of indicators	
Accuracy	- Grammatically correct - Editorial process - Reliability of links - No typos - No misspelled words	Accuracy	- Information accuracy - No error/typos	Accuracy	- Grammatically correct - Editorial process - Reliability of links - No typos
Authority	- Web address (URL) - Author's name - Author's affiliation (position/occupation) - Author's credentials - Author's contact address - Author's qualification (title) - The reputation of the source or publisher - Type of sources	Source	- URL suffix - Expertise - Name recognition and reputation - Affiliation - The author's contact address - Identity of site operator - The responsible person's name	Authority	- Web address (URL) - Expertise - Author's recognition and reputation - Author's/creator's name Author's contact address - Author's qualification - Author's affiliation
Currency	- First published - Up-to-date (recent) - Last modification/revised	Currency	- Up-to-date	Recency	- Date of publication - Date of last modification/revised
Characteristics of Information	- Type of information objects (book, journal,article, etc.) - Ranking in search output (predictive) - Content - Title - Literature	Relevance	- The familiar of a user with the topic - Information usefulness -Information focus	Relevance	- Type of information - Number of citations - Content - Title - Literature

Table 4.10: A potential criteria and indicators list for inclusion in the proposed framework

Criteria	Indicators
Authority	- Author's/creator's qualification - Author's/creator's expertise - Author's/creator's contact address - Author's/creator's name - Author's/creator's affiliation - Web address (URL) - Author's/creator's recognition and reputation
Accuracy	- Grammatically correct - No typo - Editorial process - Reliability links
Recency	- Date of publishing - Date of last modification/revised
Relevance	- Title - Type of information - Number of citations - Content - Literature

After completing this phase, we obtain a list of potential criteria and indicators for use in the construction of the trustworthiness criteria of our framework. The next chapter demonstrates an expert validation study to investigate the proposed trustworthiness criteria of the TWINE framework.

Chapter 5

Validation of the Criteria used within the TWINE Framework

In this chapter, we demonstrate the development of an instrument to validate the criteria and their indicators which are used in the TWINE framework (see Chapter 3). In section 5.1, we present an approach for developing a questionnaire to elicit experts' opinions. Then, we describe and analyse the results of the questionnaire in section 5.2. Finally, in section 5.3, we summarise the chapter.

5.1 Validating the Trustworthiness Criteria in TWINE

Validation is the process of evaluating how well an instrument works or fulfills its function (Anastasi and Urbina, 1997; Oluwatayo, 2012). In order to validate our trustworthiness criteria used in the TWINE framework as discussed in Chapter 4, we seek the opinions of experts. This approach provides useful feedback on the quality of our presented framework. Moreover, the results from the validation study are used to revise the indicators used in the criteria. An overview of the expert evaluation process for the trustworthiness criteria in the TWINE is shown in Figure 5.1.

The expert evaluation process consists of five steps; these are defining indicators for each criterion, designing the questionnaire, identifying potential experts, recruiting expert participants, and conducting the validity study. We explain each step in detail in the following sections.

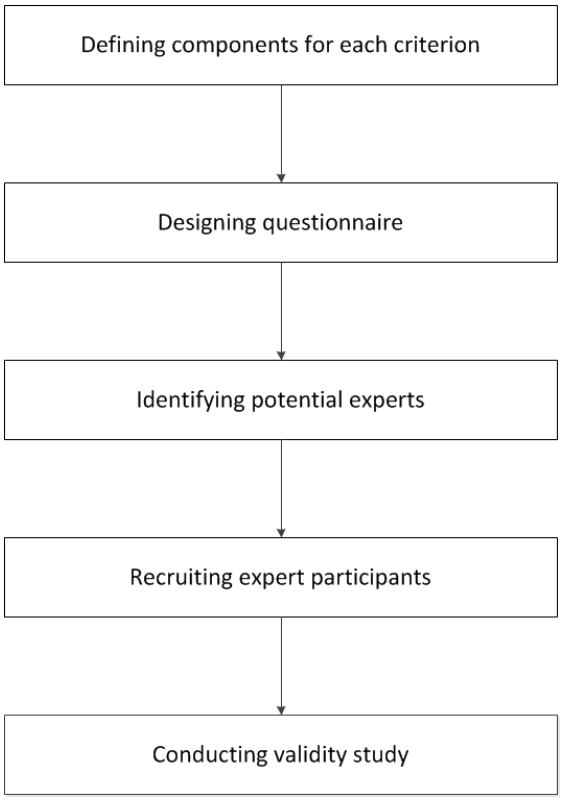


Figure 5.1: The validation processes of the criteria for the framework

5.1.1 Defining Indicators of Criteria for the Development of an Expert Evaluation Instrument

This step describes a process to generate indicators which are representative of each criterion. We used items generated from the synthesis process in Chapter 4 as shown in Table 4.10 in order to create indicators. These items were developed according to the usefulness of each item for evaluating the trustworthiness of information provided on the Web.

In this research, we set the scenario for our case study in academic publications because it is the scenario for which users who are not familiar with the area or who have less experience about the topic need a tool that can help them to evaluate the trustworthiness of Web information they are looking for. In addition, the trustworthiness of these pieces of information has an impact on the results of the work the users are researching. It is important that users are able to obtain trustworthy publications to reference in their work. Moreover, a new source of published academic journals (e.g., open-access journals) has been introduced recently. However, some open-access journals have been criticised for their poor quality control (Bohannon, 2013). As a result, academic publications are not an easy area in which to assess the trustworthiness of information they provide.

In a questionnaire, we set the scenario that the experts were supervisors of new undergraduate students, who were looking for information to use in their work. It was

intended to ask each expert for their opinions on how useful each of the criteria is in order to support the evaluation of the trustworthiness of Web information. Therefore, we constructed a list of indicators based on sets of corresponding criteria; these indicators would also be generally available for academic publications. However, some of the indicators in the criteria were edited to make them suit the scenario better.

For example, considering the accuracy criterion, according to the scenario, academic publications were normally expected to be grammatically correct and contain no typos. Academic content needed to pass the process of peer review. Therefore, we could indicate the accuracy of information based on the editorial process, which would cover all of the aspects of being grammatically correct, containing no typos and having reliable links. Similarly for a component, "content", in the relevance criterion, we used the information's abstract to represent the content, because an abstract in an academic publication summarised the concepts and major points of the work from which we were informed about the content. As a result, we obtained a total of 13 indicators for the framework as shown in Table 5.1, which also shows the relationships between the criteria and the indicators in the framework.

Table 5.1: A list of indicators in experts' validation

Criteria	List of indicator elements
Authority	Element1: The name of the content creator (e.g. author's name or a name of organization)
	Element 2: The creator's or author's affiliation
	Element 3: The creator's or author's position
	Element 4: The creator's or author's title (e.g. Dr or Professor)
	Element 5: The physical address of the organisation
	Element 6: Brief details about the content creator's experience
Accuracy	Element 7: Information of the editorial process (e.g. has the content passed peer-review or has it been reviewed by others?
Recency	Element 8: The publication date of the content
	Element 9: The last modification date of the con-tent
Relevance	Element 10: Number of times that the information has been referenced in other documents
	Element 11: Publication medium (e.g. book, journal, article, blog, etc.)
	Element 12: An overview of the content (e.g. title, abstract, etc.)

Table 5.1: A list of indicators in experts' validation

Criteria	List of indicator elements
	Element 13: A list of references
4 criteria	13 Elements

5.1.2 Designing the Questionnaire

Our framework is designed to help the novice Web user to assess the trustworthiness of information found on the Web, and our focus is the academic domain. We used the generated elements from section 5.1.1 to create a questionnaire. The purpose of this questionnaire was to allow an expert to rate the effect of the elements on the evaluation of the trustworthiness of Web information. It aimed to provide a better understanding of the factors that influence the assessment of the trustworthiness of Web information. The results from this were used to refine our designed framework. There were four sections in this questionnaire.

- Section 1: The effect of the presence of each element on the person's confidence in their ability to evaluate the trustworthiness of Web information.
- Section 2: The effect of the absence of each element in the person's confidence in the trustworthiness of Web information.
- Section 3: The importance of the elements in assessing the trustworthiness of Web information.
- Section 4: Additional elements which should be considered.

Experts were asked to rate elements corresponding to the purpose of each section.

In section 1, we asked experts to rate how useful each item is in order to evaluate the trustworthiness of Web information. The response options used a four-point scale format as shown in Table 5.2.

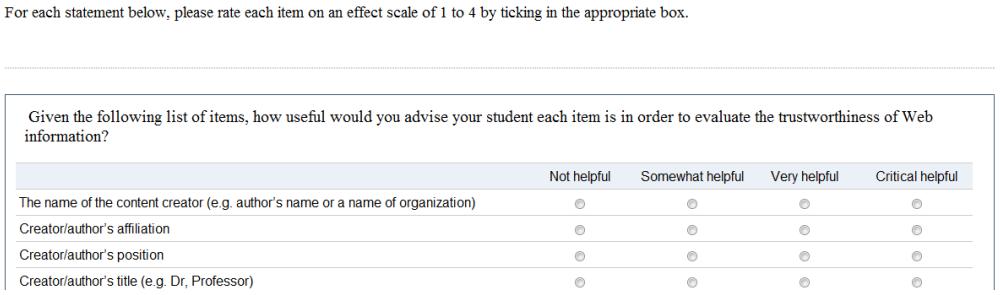
Table 5.2: Rating criteria for the expert evaluation in Section 1

Rating	Definition
1 = "Not helpful"	The presence of this element does not affect one's evaluation of the trustworthiness of Web information.
2 = "Somewhat helpful"	While the presence of this element helps to build one's confidence in the evaluation of trustworthiness of Web information, its absence does not seriously detract from one's confidence.

Table 5.2: Rating criteria for the expert evaluation in Section 1

Rating	Definition
3 = "Very helpful"	This element is needed to be truly confident of one's evaluation of the trustworthiness of the Web information. However, without this element, one can still have some confidence in one's evaluation.
4 = "Critically helpful"	This element is essential in order to evaluate the trustworthiness of Web information. Without it one cannot have any confidence in one's evaluation of the trustworthiness of said information.

An example question from Section 1 is shown in Figure 5.2. We set the scenario of the study by asking experts to imagine themselves in a situation in which they were an academic advisor to new undergraduate students who were starting their studies at the university. The students came to the expert to ask for advice about the indicators that they should look for on the Web that indicated that they could trust the information on the Web. Then, we gave a list of elements and asked experts to rate the usefulness of each element in order to evaluate the trustworthiness of Web information.


For each statement below, please rate each item on an effect scale of 1 to 4 by ticking in the appropriate box.

Given the following list of items, how useful would you advise your student each item is in order to evaluate the trustworthiness of Web information?

	Not helpful	Somewhat helpful	Very helpful	Critical helpful
The name of the content creator (e.g. author's name or a name of organization)				
Creator/author's affiliation				
Creator/author's position				
Creator/author's title (e.g. Dr, Professor)				

Figure 5.2: An example of a question for rating the usefulness of the presence of elements for evaluating the trustworthiness of Web information.

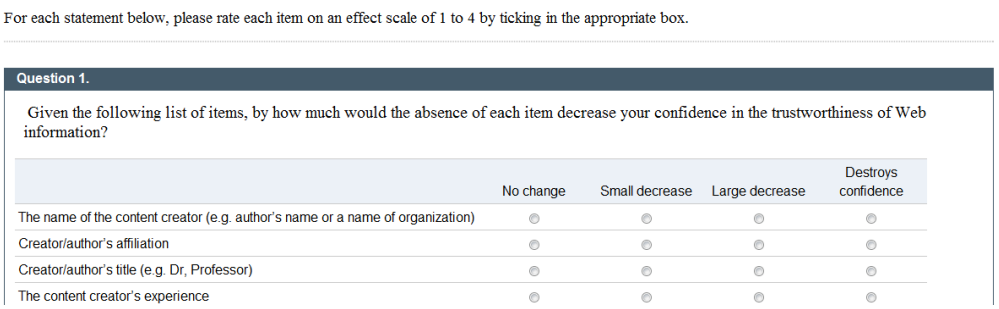
In Section 2, we asked experts to rate how much the absence of each element would decrease their confidence in the trustworthiness of Web information. The response options used a four-point scale as shown in Table 5.3. We gave the same scenario to the experts as we did in Section 1 of the questionnaire. An example question is shown in Figure 5.3.

Table 5.3: Rating criteria for expert evaluation in Section 2

Rating	Definition
1 = "No change"	The absence of the element does not decrease my confidence in the trustworthiness of the Web information.
2 = "Small decrease"	The absence of this element will decrease the trust I place in the Web information, but the Web information can still be trustworthy without it.

Table 5.3: Rating criteria for expert evaluation in Section 2

Rating	Definition
3 = "Large decrease"	The absence of this information is damaging to the Web information's trustworthiness. However, other features of the Web information may redeem ***some*** trust.
4 = "Destroys confidence"	If this information is not present, one cannot place any trust in the Web information.


For each statement below, please rate each item on an effect scale of 1 to 4 by ticking in the appropriate box.

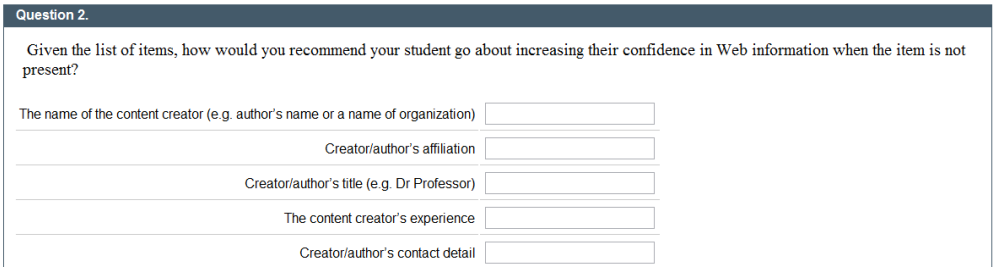
Question 1.

Given the following list of items, by how much would the absence of each item decrease your confidence in the trustworthiness of Web information?

	No change	Small decrease	Large decrease	Destroys confidence
The name of the content creator (e.g. author's name or a name of organization)	○	○	○	○
Creator/author's affiliation	○	○	○	○
Creator/author's title (e.g. Dr, Professor)	○	○	○	○
The content creator's experience	○	○	○	○

Figure 5.3: An example question for rating the effect of the absence of elements on the perceived trustworthiness of Web information.

In addition, we asked the experts to give suggestions on how to increase the confidence in Web information in the case that certain elements were not present on the Web. An example question is shown in Figure 5.4.


Question 2.

Given the list of items, how would you recommend your student go about increasing their confidence in Web information when the item is not present?

The name of the content creator (e.g. author's name or a name of organization)	
Creator/author's affiliation	
Creator/author's title (e.g. Dr Professor)	
The content creator's experience	
Creator/author's contact detail	

Figure 5.4: An example question for recommendations for increasing the confidence in Web information in the case that elements are not provided.

In Section 3, we asked experts to rank the three most important elements from the given list for supporting the assessment of the trustworthiness of Web information as shown in Figure 5.5.

Finally, in Section 4, we provided open-ended questions to ask for the experts' suggestions on additional items which should be considered. Some example questions are shown in Figure 5.6.

The expert evaluation questionnaire used in this study can be found in Appendix A.

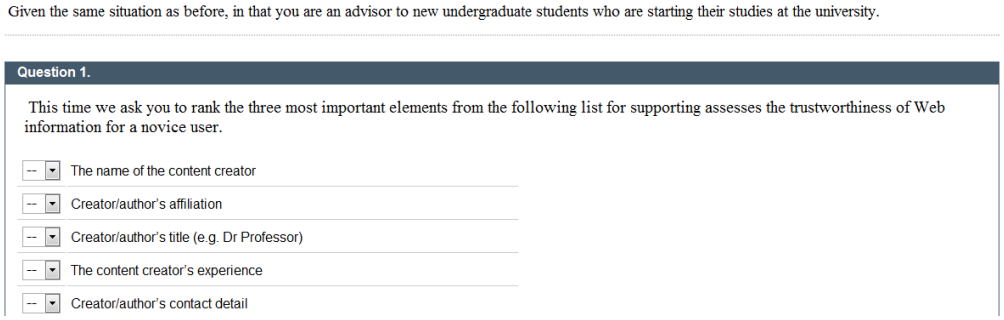


Figure 5.5: An example question for ranking the three most important elements for supporting assessment of the trustworthiness of Web information.

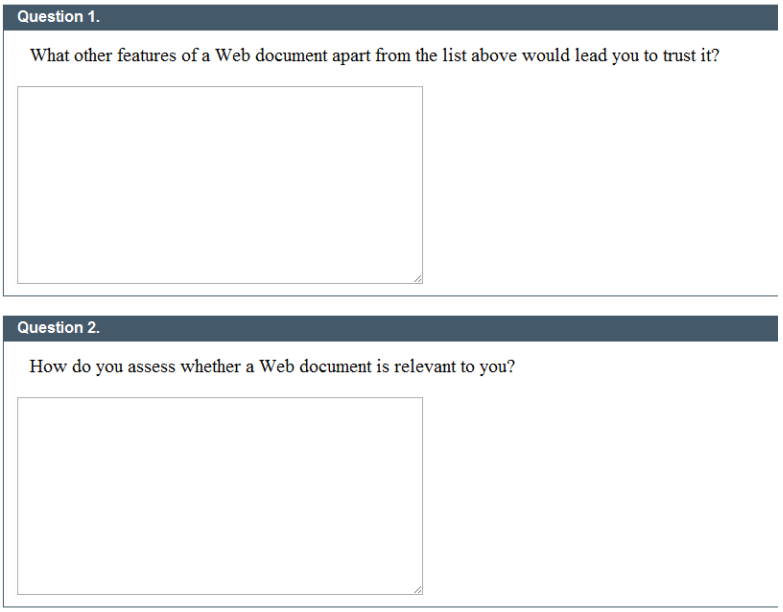


Figure 5.6: Examples of open-ended questions asking the experts to suggest additional elements which should be considered

5.1.3 Identifying Potential Expert Participants

After we finished the questionnaire, the next step was to identify potential participants. We discuss the process we used to define the number of polled experts in section 5.1.3.1 and then we discuss the selection of potential experts to participate in the study in section 5.1.3.2.

5.1.3.1 Defining the Sample Size

We need to estimate the minimum number of participants who need to attend, in order to ensure the quality of the study. We used a *priori* power analysis to estimate sufficient elaborate sample sizes in our study (Cohen, 1992b; Prajapati et al., 2010).

We designed a questionnaire to ask the experts' opinions on the usefulness of elements for evaluating the trustworthiness of Web information. In addition, we were looking for items that affect the experts' confidence in their ability to evaluate the trustworthiness of Web information. The effect on the experts' confidence could be in one of two directions (i.e. increase or decrease). As a result, the most appropriate statistical test for this study was the two-tailed probability test.

In our study, we chose α-level[1] as 0.05 based on Fisher's suggestion in his study for the probability of detecting an effect in the population, when that effect does not in fact exist (Fisher and Bennett, 1973). In addition, we set β[2] as 0.2 which is the suggested maximum acceptable probability of a Type II error based on Cohen's suggestion (Cohen, 1992a). Moreover, we set a desired statistical power[3](P) as 0.8 (1-β) because we wanted at least an 80% chance of detecting a statistically significant effect from the study findings.

We set the effect size[4] (d) as 1.0 (large effect), according to the effect size conventions of Cohen's suggestion (Cohen, 1992b), as we wished to be able to detect whether the experts believe that the presence of the elements was really useful for assessing the trustworthiness of Web information.

The minimum sample size is a proportion of a statistical power to an effect size of the study with α-level. We used the G*Power[5] program for calculating the number of participants we need to recruit, as shown in Figure 5.7.

t tests – Means: Difference from constant (one sample case)			
Analysis:	A priori: Compute required sample size		
Input:	Tail(s)	=	Two
	Effect size d	=	1.0
	α err prob	=	0.05
	Power (1–β err prob)	=	0.8
Output:	Noncentrality parameter δ	=	3.1622777
	Critical t	=	2.2621572
	Df	=	9
	Total sample size	=	10
	Actual power	=	0.8030969

Figure 5.7: A *priori* power analysis function in G*Power.

G*Power does this by calculating the power level for each number of agents:

[1] α is the probability of detecting an effect in the population when that effect does not in fact exist (Type I error). In other words, it is a false-positive (Field, 2009).

[2] β is the probability of failing to detect an effect in the population when it actually does exist (Type II error). Rather, it is a false-negative (Field, 2009).

[3] P is the ability of a statistical test to detect an effect if it exists in the sample size.

[4] Effect size is the magnitude of the difference between groups of studies. Large differences (large effect size) are easier to detect (Cohen, 1992b).

[5] G*Power is a general stand-alone power analysis program for a variety of frequently used statistical tests (i.e. t tests, F tests, z tests, etc.) (Erdfelder et al., 1996).

$$N \in [N_{lb}, N_{ub}]$$

where N_{lb} is the lower bound of the expected solution, and N_{ub} is an arbitrary upper bound.

G*Power then selects the smallest N for which the calculated power level is greater than the specified power level (Erdfelder et al., 1996). Given the α-level, the effect size (d) and desired statistical power (P) as mentioned above, we obtained the total sample size as being 10 participants, with α-level of 0.05 and the power of a statistical test to detect an effect of the sample size to the real-world scenario as 0.8.

5.1.3.2 Defining the Panel of Potential Participants

The purpose of our study is to evaluate the criteria which are used in our framework that help users to assess the trustworthiness of Web information. Therefore, participants should be familiar with the process of the evaluation of information, in particular, within the Web environment. Moreover, they should have experience in assessing the quality or credibility of Web information. As we discussed in section 5.1.3.1, the number of experts we used in our validity study is 10.

In this study, we had five experts who were librarians and five experts who were academic researchers. Those experts were recruited based on their experience in assessing the credibility or quality of information on the Web. The profile of each category of expert is described as follows:

- Academic researcher: Experts in this category were research fellows who have had experience in searching for information for their research and selecting the publications to reference in their work.
- Liaison staff in the library: We compiled a list of liaison staff from the suggestions of Academic Liaison of Faculty of Physical & Applied Sciences, Southampton University. All five experts have had experience in evaluating the credibility of information on the Web. Moreover, they were involved in selection or de-selection of journals in various areas. In addition, one of them has been specifically trained in evaluating the credibility of information. Furthermore, two experts created a tutorial session for evaluating the credibility of information on the Web for students and staff in the university

5.1.4 Recruiting Expert Participants

We sent an invitation email to solicit expert participation. An email was sent to potential participants two weeks before the actual study started such that the participants have enough time to respond to the request. The invitation email contained a description of the purpose of the study, a brief description of the study, and how the participant could contribute to the study. Table 5.4 shows the indicators of the invitation email (an invitation email can be found in Appendix B).

Table 5.4: Indicators of an invitation email for experts' participation.

Indicators	Sample text
The purpose of the study	The aim of the study is to confirm that the criteria in question are good for evaluating the trustworthiness of Web information and to gain a better understanding of the factors that influence the assessment of the trustworthiness of Web information.
Brief description of the study	The measurement tool in this study is a questionnaire, which is created and uploaded to the iSurvey system of University of Southampton. The survey should take approximately 25 minutes or less to complete. You will not be asked for any demographic information.
How the expert's participation contributes to the study	Participation is voluntary. Therefore, refusal to take part in the study involves no penalty or loss of benefits. Participants can withdraw from the study at any time without penalty.
Contact information	If you have any further questions about this study or your rights, or if you wish to lodge a complaint or concern, you may contact the Principal Investigator, Jarutas Pattanaphanchai by email jp11g09@soton.ac.uk. Ethics Reference Number: 2538

5.1.5 Conducting the Validity Study

On the start date of this study, we sent out a second email, which contained a link to the questionnaire, to the experts who agreed to take part. At the beginning of the questionnaire, information regarding the study and ethics information was presented. The participant clicks on the survey link and must tick the box provided on the page to indicate that he or she consented to taking part in the survey. Then, by clicking on the

appropriate radio buttons or filling text boxes, the participant answered the questions regarding the importance of the selected criteria for evaluating the trustworthiness of Web information. After completing the questions, the participant was instructed to click the "Exit" button or to close the browser. The study lasted for a month, and we received 10 responses, which are analysed in the next section.

5.2 Analysis and Results of the Expert Evaluation

The main aim of this study is to identify the importance of a range of selected criteria in assessing the trustworthiness of Web information and the indicators of each criterion that have the greatest influence on the evaluated trustworthiness of Web information. It aims to confirm that the selected criteria are good for evaluating the trustworthiness of Web information and to gain a better understanding of the factors that influence the assessment of the trustworthiness of Web information. Experts were asked to rate the usefulness of the appearance of items and the effect of the absence of that item. In addition, they were asked to rank the three most important elements for assessing the trustworthiness of Web information.

In the questionnaire, we had open-ended questions in Sections 2 and 4 which asked experts to give some suggestions. Consequently, there was a chance that the comments written by experts might contain some words that spell incorrectly. However, we needed to keep these typos in order to reflect the real responses from the experts.

The questionnaire was designed with four sections comprising two types of questions. One type consisted of rating scale questions which asked the participants to rate the usefulness, effect, and importance of items. The second type was open-ended questions which asked for the participants' suggestions. Therefore, we obtained two types of data from our study: quantitative data from rating scale questions and qualitative data from open-ended questions. We discuss the analysis approach adopted for each data type in more detail in sections 5.2.1 and 5.2.2, respectively.

5.2.1 Analysis of the Quantitative Data

Quantitative data gives meaningful numerical results. We used quantitative analysis methods for the quantitative data from the questions in Section 1 (the effect of the presence of each element on the person's confidence in their ability to evaluate the trustworthiness of Web information) and question 1 in Section 2 (the effect of the absence of each element on the person's confidence in the trustworthiness of Web information). We discuss the details in the next section.

5.2.1.1 Identifying a Statistical Test for our Quantitative Data

After we collected the data from the participants, we needed to find a statistical test that fits to the data. There are two main types of tests: parametric tests and non-parametric tests. Parametric tests require the assumptions that the data are normally distributed (i.e. distributed symmetrically around the centre for all scores). Non-parametric tests are referred to assumption-free tests, in which the tests make fewer assumptions about the type of data on which they can be used; however, they make no assumptions about the distribution of data (Field, 2009). We therefore needed to check the normality of our collected data in order to select the statistical test that best suits the data.

We tested the normality of data using the Shapiro-Wilk test, which compares the scores in the sample to a normally distributed set of scores with the same mean[6] and standard deviation[7]. If the difference between the scores in the sample and a normally distributed set of scores with the same mean is not significant (Sig. > 0.05), it means that the distribution of the sample is not significantly different from a normal distribution. In other words, the distribution of the sample is normal. However, if the test is significant (Sig. < 0.05) then the distribution of the data is significantly different from a normal distribution. Therefore, the distribution of data is non-normal (Field, 2009). We used Shapiro-Wilk tests in SPSS in order to test the normality of our collected data from question 1 in Section 1 and question 1 in Section 2. The results of the test are shown in Tables 5.5 and 5.6 respectively. The tables include the elements, the statistic test value (in this test, it is a mean), the degree of freedom (df)[8], which in this test should equal the sample size, and the signification value of the test.

Table 5.5: Tests of distributed normality data of question 1 in Section 1

Elements	Shapiro-Wilk		
	Statistic (Mean)	df	Sig.
The name of the content creator	0.859	10	0.074
Creator's/author's affiliation	0.833	10	0.036
Creator's/author's position	0.594	10	0.000
Creator's/author's title (e.g. Dr, Professor)	0.820	10	0.025
Publication medium	0.802	10	0.015
An overview of the content (e.g. title, abstract, etc.)	0.852	10	0.061

[6]The mean is the measurement of the central of a frequency distribution lies. It is an average of the scores (Field, 2009).

[7]Standard deviation (s) is the average distance between each point in the data and the mean (Field, 2009).

[8]The degree of freedom is the number of observations that are free to vary (Field, 2009).

Table 5.5: Tests of distributed normality data of question 1 in Section 1

Elements	Shapiro-Wilk		
	Statistic (Mean)	df	Sig.
Publication date of content	0.820	10	0.026
The last modification date of content	0.907	10	0.258
The physical address of the organisation	0.832	10	0.035
Brief details of content creator's experience	0.366	10	0.000
The information on the editorial process (e.g. passed peer-review)	0.805	10	0.017
A list of references	0.833	10	0.036
Number of times that the information has been referenced/cited	0.794	10	0.012

From the normality test results of the data in Section 1 in Table 5.5, we found that the name of the content creator, an overview of the content, and the last modification date of the content are normally distributed (Sig. > 0.05). Conversely, five elements; namely, the creator's/author's position, the publication medium, the publication date of content, a brief detail of content creator's experience, and the number of times that the information has been reference or cited, are non-normally distributed (Sig. < 0.05).

Table 5.6: Tests of distributed normality data of question 1 in Section 2

Elements	Shapiro-Wilk		
	Statistic (Mean)	df	Sig.
The name of the content's creator	0.833	10	0.036
Creator's/author's affiliation	0.820	10	0.025
Creator's/author's title (e.g. Dr, Professor)	0.532	10	0.000
A type of content publication	0.640	10	0.012
The content of the creator's experience	0.794	10	0.000
Creator's/author's contact details	0.781	10	0.008

Table 5.6: Tests of distributed normality data of question 1 in Section 2

Elements	Shapiro-Wilk		
	Statistic (Mean)	df	Sig.
Number of times that the information has been referenced/cited	0.802	10	0.015
The information of the editorial process (e.g. passed peer-review)	0.820	10	0.025
Publication date of content	0.731	10	0.002
The last modification date of content	0.805	10	0.017
The title of the content	0.859	10	0.074

From the normality test results of the data in Section 2, shown in Table 5.6, we found that all of the elements showed a non-normal distribution (Sig. <0.05).

According to the normality test, there were some items that are normally distributed but others were not. For the consistency analysis, we decided to choose a non-parametric test in order to analyse our collected data. We selected a non-parametric test because this test requires no assumption about the distribution of the underlying sample. Therefore, it can be used on both items that are normally- and non-normally distributed.

We selected the Wilcoxon signed-rank test[9] for analysis. This test is equivalent to the t-test[10] but it is better suited for use on non-parametric data. According to the non-normal distribution of our data, a probability distribution of data was not symmetric. Therefore, the mean is a poor estimator of the central tendency of the set of data because it is highly influenced by extreme values. Consequently, the median[11] is a more robust estimator inasmuch as it is not influenced by extreme values.

As a result, we compared the median of the importance score of each item with a constant value, which we selected in each section to show whether or not an element has significant importance in evaluating the trustworthiness of Web information.

[9] Wilcoxon signed-rank test is a non-parametric test that is used to test the differences between two related samples (Field, 2009).

[10] t-test is a test statistic in which in this context, it is used to test the differences between two means (Field, 2009).

[11] Median is the middle score of a set of ordered data. When the data consists of an even number of observations, the median is the average of the two scores that are either side of what would be the middle value (Field, 2009).

5.2.1.2 Wilcoxon signed-rank Test Analysis for Section 1 and Section 2 of the Questionnaire

The Wilcoxon signed-rank test in one sample group is used for testing the null hypothesis that the population median of a random variable is equal to a given value M. It is assumed that the variable is symmetrically distributed about its median. We used SPSS to analyse the data and the default test statistic is a two-tailed test.

5.2.1.2.1 Wilcoxon signed-rank test analysis results for Section 1 of the questionnaire

This section investigates the effect of the appearance of elements in one's confidence of one's ability to evaluate the trustworthiness of Web information.

We are interested in the experts' opinions on the elements in both directions (they think items are even more helpful or they think elements are less helpful). Therefore, we are interested in whether the median rating of each element is significantly different from 3, very helpful (either more or less than 3). We set the null hypothesis that the median response is equal to 3. The significance level is 5% (α-level: 0.05). The results are displayed in Table 5.7.

Table 5.7: Wilcoxon signed-rank two-tailed test for a single sample in Section 1

No.	Null Hypothesis	Median	Sig. (Z)	Decision
1	The median of the author's name equals 3.00	3.00	1.000	Retain the null hypothesis
2	The median of the author's affiliation equals 3.00	3.00	0.655	Retain the null hypothesis
3	The median of the author's position equals 3.00	3.00	0.083	Retain the null hypothesis
4	The median of the author's title equals 3.00	2.00	0.010	**Reject the null hypothesis**
5	The median of the publication medium equals 3.00	3.00	0.180	Retain the null hypothesis
6	The median of the content of the title or its abstract equals 3.00	2.50	0.194	Retain the null hypothesis
7	The median of the publication date of the content equals 3.00	3.00	0.705	Retain the null hypothesis

Table 5.7: Wilcoxon signed-rank two-tailed test for a single sample in Section 1

No.	Null Hypothesis	Median	Sig. (Z)	Decision
8	The median of the last modification date of the content equals 3.00	2.50	0.160	Retain the null hypothesis
9	The median of the physical address of the organisation equals 3.00	2.00	0.015	**Reject the null hypothesis**
10	The median of the brief detail of content of the author's experience equals 3.00	2.00	0.003	**Reject the null hypothesis**
11	The median of the information of the editorial process equals 3.00	3.00	0.705	Retain the null hypothesis
12	The median of a list of references equals 3.00	3.00	0.655	Retain the null hypothesis
13	The median of the number of times that the information has been cited equals 3.00	3.00	0.317	Retain the null hypothesis

The results in Table 5.7 show, at the 5% significance level for a two-tailed test, that the median of 10 elements; namely, the author's name, the affiliation, the position, the publication medium, the content of the title or abstract, the publication date of the content, the last modification date of the content, the information regarding the editorial process, the list of references, and the number of times that the information has been cited are equal to 3 with significance levels of Z=1.000, Z=0.655, Z=0.083, Z=0.180, Z=0.194, Z=0.705, Z=0.160, Z=0.705, Z=0.655, and Z=0.317, respectively. This indicates that these items are particularly helpful in assessing the trustworthiness of Web information.

Conversely, the median of three elements, namely the author's title, the physical of the organisation and the brief detail of the author's experience are significantly different from 3, specifically less than 3 with significance levels of Z= 0.010, Z=0.015, and Z=0.003, respectively. This indicates that these three items are not very helpful for assessing the trustworthiness of Web information. In conclusion, we decided to use the 10 elements that retained the null hypothesis in our framework.

5.2.1.2.2 Wilcoxon signed-rank Test Analysis Results for Question 1 of Section 2 of the Questionnaire

This question investigates the effect of the absence of elements in perceiving the trustworthiness of Web information. We set the null hypothesis such that the median response is equal to 3, which means an element will largely decrease the confidence of the trustworthiness of Web information if it is not present. The significance level is 5% (α-level: 0.05). The results are shown in Table 5.8.

Table 5.8: Wilcoxon signed-rank one-tailed test for a single sample of question 1 in Section 2

No.	Null Hypothesis	Median	Sig. (Z)	Decision
1	The median of the name of the content creator equals 3.00	3.00	0.6555	**Retain the null hypothesis**
2	The median of the author's affiliation equals 3.00	3.00	0.414	**Retain the null hypothesis**
3	The median of the author's title equals 3.00	1.00	0.004	Reject the null hypothesis
4	The median of the content of the creator's experience equals 3.00	2.00	0.006	Reject the null hypothesis
5	The median of the author's contact detail equals 3.00	1.50	0.006	Reject the null hypothesis
6	The median of the number of times that the information has been referenced in other document equals 3.00	2.00	0.006	Reject the null hypothesis
7	The median of the information of the editorial process equals 3.00	2.00	0.023	Reject the null hypothesis
8	The median of the publication date of content equals 3.00	3.00	0.059	**Retain the null hypothesis**
9	The median of the last modification date of content equals 3.00	2.00	0.026	Reject the null hypothesis
10	The median of the brief details of content of the author's experience equals 3.00	2.00	0.003	Reject the null hypothesis
11	The median of a type of content publication equals 3.00	2.00	0.014	Reject the null hypothesis

The results in Table 5.8 show, at the 5% significance level for a two-tailed test, that the median of three elements, namely; an author's name, an author's affiliation, and a publication date of content are equal to 3 with significance levels of Z=0.655, Z=0.414, and Z=0.059 respectively. This indicates that the effect of these three elements when they are not present will largely decrease the confidence of the trustworthiness of information on the Web.

Conversely, the median of eight elements; namely, the author's title, the brief detail of the author's experience, the contact details, the number of times that the information has been cited, the information of editorial process, the last modification date of content, the tile of the content matching your needs, and the type of content publication are significantly different from 3, specifically they are all less than 3 with significance levels of Z=0.004, Z=0.006, Z=0.006, Z=0.006, Z=0.023, Z=0.024, Z=0.026, and Z=0.014 respectively. This indicates that the absence of these eight elements does not significantly decrease the trustworthiness of Web information. The Web information can still be trustworthy without it.

5.2.1.3 Analysis of Results for Section 3 of the Questionnaire

This section investigates the importance ranking of each of the proposed trustworthiness indicators in assessing the trustworthiness of Web information. Given the same experimental scenario as before in that the experts are advisors to new undergraduate students who are starting their studies at university, we asked the experts to rank the importance of each of the given indicators for supporting the assessment of the trustworthiness of Web information. We then calculated an importance score of each of the indicators. The importance score was calculated by assigning points to the rank given by the expert for each element, with the highest ranking element receiving the highest number of points. For example, the first place rank is assigned the maximum number of points, *M*, the second place is assigned (*M - 1*), and the third place is assigned (*M - 2*) and so on. In this study, we asked the experts to rank what they considered to be the ***three most important*** elements. Given this, we assigned a first place rank 3 points, a second place rank 2 points, and a third place 1 point. From this, we calculated the importance score of each element as the fraction of its number of importance points given by the experts to the maximum number of importance points it is possible to achieve (i.e. the number of points it would have received had each expert ranked the indicator as being the most important). More formally, it is said that the importance score of an indicator d, I_d is

$$I_d = \frac{3x_{d,1} + 2x_{d,2} + 1x_{d,3}}{3N}$$

where $x_{d,i}$ is the number of votes for the indicator, d, to be in the i^{th} position and N is the number of participants. Thus, I_d represents indicator d's relative importance as a

fraction of the maximum importance. From this, the relative weight of each indicator, W_d, can be calculated as follows:

$$W_d = I_d \left(\sum_{k \in indicators} I_k \right)^{-1}$$

This weighting equation normalised the indicators' scores such that they sum to one. This is beneficial as, assuming the individual indicators scores are of the range [0,1], it allows for a trustworthiness score to be bounded to the range [0, 1]. Thus a 'perfectly' trustworthy piece of information would score 1 where as a completely untrustworthy piece would score 0.

The results of the importance score and weighting factor for each indicator are shown in Table 5.9.

Table 5.9: The importance score and weighting factor of each indicator

Indicators	Importance Score	Weighting factor
Author's affiliation	0.50	0.28
Author's name	0.30	0.17
Editorial process	0.30	0.17
Publication date	0.23	0.13
Publication medium	0.23	0.13
Content of the title or abstract	0.10	0.06
Number of citations	0.07	0.04
Last modification date	0.07	0.04

After finishing the analysis of quantitative data, we analysed the qualitative data which we obtained from open-ended questions in Section 2 (handling missing useful supportive information on the Web) and Section 4 (the process of evaluating the relevance of information and the user's needs) from the questionnaire. We discuss the details of this analysis in the next section.

5.2.2 Analysis of the Qualitative Data

The qualitative data consists of non-numerical results from the study, and they aim to build a subjective understanding of a situation. We used a qualitative analysis method in order to analyse the results from a questionnaire. The qualitative methods and content analysis are recommended when answering "how" and "why" questions in which there is

little control for answering the questions giving rise to unstructured responses from the participants (Yin, 2009). Therefore, we analysed the results from question 2 in Section 2 and from the questions in Section 4 using a qualitative analysis approach. We discuss this in more detail in the next section.

5.2.2.1 Thematic Analysis Approach

As part of the qualitative analysis we performed, we employed thematic analysis, which is a method for analysing classifications and presenting themes (patterns) that relate to the data. Thematic analysis is considered an appropriate approach to discover the relationships between concepts. It can detect and identify factors or variables that influence any ideas or suggestions generated by participants' opinions. Therefore, thematic analysis can help to elicit an appropriate explanation for the participants' responses (Alhojailan, 2012; Braun and Clarke, 2006).

5.2.2.1.1 Thematic analysis methodologies

The methodologies of thematic analysis can be classified into two primary approaches: inductive and deductive methodologies.

- The inductive method, also called the "bottom up" method, is a data driven approach in which the themes will be found from the collected data. It is a process of coding the data without trying to fit them into a pre-existing concept. The coding of the data is a process of selecting the key terms in the response texts that can reflect the behaviour or opinions of the participants and defining each term with the short name or coded name in order to refer it for analysis later. Therefore, the derived themes are strongly linked to the data themselves.
- The deductive method, also known as the "top down" method, is an approach in which the themes will be driven by the researcher's theoretical or analytic interest in the area and then applying these themes to draw conclusions. For example, the researcher set the themes based on the literature review of the topic they are studying. This form of analysis tends to provide less description of the data overall.

In this study, we used the inductive methodology for analysis. We focused on discovering the themes which are strongly linked to the expert's responses. These themes can represent the behaviours of the experts when they evaluate the trustworthiness of Web information when supportive data are not present, and can tell us how the experts judge the relevance of information to their needs.

5.2.2.1.2 Thematic analysis process

Thematic analysis involves finding repeated patterns of activities which are useful to explain behaviours of participants across a data set. The steps to conduct thematic analysis consist of six phases:

- **Phase 1 familiarising yourself with your data:** This phase is about immersing yourself with your data. It involves repeated reading of the data or searching for meaning or patterns while reading the data. In this study, we familiarised ourselves with the data by preparation and organisation of the content of the data using Microsoft Excel and Word. By doing this, we can read through all of the experts' responses and structure them in a way that allows us to easily analyse the responses in more detail. The details of data preparation is explained in section 5.2.2.2.

- **Phase 2 generating initial codes:** This phase involves the generation of the initial codes (the key terms in the response texts that can reflect the behaviour or opinions from the participants and renames them in short form) from the data. These codes provide the meaning or patterns about the processes or patterns at the end of the analysis process. Coding depends on the data that have emerged based on some specific questions that have been set. Coding can be done manually when a person reads through the content and then defines words that can represent an idea or pattern. Moreover, coding can be generated through a software program, where the program counts the frequency with which words appear in the content and defines codes based on the most frequently used words in the content. We used a software package named "NVivo" to help us generate the initial codes, which we discuss in more detail in sections 5.2.2.3 and 5.2.2.4.

- **Phase 3 searching for themes:** After we finished the initial coding, we obtained a list of the different codes which have been identified across the data set. This phase focuses on the analysis at the top level of themes. It involves sorting the different codes into potential themes and collating all the relevant coded data within the identified themes. In this phase, we can use a visual presentation to help in sorting the different codes into the themes. The visual tools that can be used might include tables, mind-maps or charts, etc. This visual representation can show the relationship between codes, between themes, and between different levels of themes. Then, the themes are defined based on the purpose of the study. During this phase, some initial codes may be identified as main themes, whereas others may form sub-themes, and others are discarded. As a result, the collection of candidate themes and sub-themes emerge.

- **Phase 4 reviewing the themes:** This phase refines the candidate themes from the previous phase. It involves two levels of reviewing and refining the themes.

The first level involves reviewing the candidate themes and considering whether they can form a coherent pattern. If the candidate themes pass this level they are considered in the second level, but if they do not match any pattern, they need to be reworked or discarded, or new themes must be created. This second level involves the validation of individual themes in relation to the data sets.

- **Phase 5 defining and naming themes:** This phase involves defining the themes that are presented for analysis. It focuses on identifying the essence of each theme and the aspects of the data that each theme captures. For each individual theme, it is important to write a detailed analysis including the narrative that represents each theme. It needs to fit into the overall story that we are telling about the data in relation to our research question.
- **Phase 6 producing the report:** This phase sets out the detailed narrative of the data and aims to convince the readers of the validity and benefits of the analysis. The report must provide sufficient evidence of the themes within the data. However, it should go beyond simply a description of the data by including an analytical narrative and an argument related to research questions.

However, analysis is not a linear process; it is more likely to be an iterative activity in which the process can be moved back and forth as needed throughout the phases. We adopted this process to analyse our data from experts' answers which are explained in more detail in the next section.

5.2.2.2 Data Preparation for Qualitative Data Analysis

This step corresponds to the first phase of thematic analysis (i.e. becoming familiar with the data). We applied this analysis to data extracted from the answers from Sections 2 and 4 of the questionnaire. When we prepared the data, we could read through all of the answers and we also organised them into a structure that allows for easy querying or analysis in the next step. Other phases of analysis are discussed in more detail in upcoming sections. We used qualitative data analysis software, NVivo version 10, to help us to process and analyse the answers from the experts. It consisted of four main steps as described (Hughes et al., 2010):

1. Specify a unique Identifier (ID) for each respondent: We assigned a unique ID to each of the participants of this survey. In so doing, the participants were named as "Expert" followed by a serial number (for example, Expert001, Expert002, Expert003, and so on). However, we were not focused on any demographic details other than the fact that they are experts in the academic domain. Therefore, we did not collect demographic data from experts.

2. Locate the set of response texts and copy them into Microsoft Excel: We used Microsoft Excel to set up the initial data for analysis. We created a single workbook which consists of a separate Excel worksheet for each open-ended question. We used two columns for recording the response data: one with the unique respondent identifiers, and the other with the response texts. Therefore, each worksheet represented an individual question and the rows in each worksheet represented individual responses.

3. Export the response texts and IDs from Microsoft Excel to Microsoft Word for formatting: We copied the two columns from Microsoft Excel which contain IDs and their associated response texts for each question as shown in Table 5.10.

Table 5.10: An example of preparation of the data in table form

Participants	Comments
Expert001	Investigate whether there are links back to a sponsoring organisation as this may give a clue to origin of the work.
Expert002	Is it published by a commercial publisher or society? If not check citations and links/references.
Expert003	Research the authors. Their online presence and credentials.
Expert004	Searching for the item in relevant search engines to find if the creators name is listed elsewhere.
Expert005	Search the Internet for other sources of this document to see if these details can be found.
Expert006	Don't know.
Expert007	Search more information on it. If it's not available drop the Web information.
Expert008	Googling the title.
Expert009	Look at the web address. Discard if not trustworthy.
Expert010	Critically evaluate the content and check references.

   We converted the columns into a two-column table, and inserted it into a Microsoft Word document. Then, we set the questions to a consistent style using the "Heading style" in Word. Particularly, we used different heading styles for the questions and the IDs of participants (we used "Heading 1" style for questions and "Heading 2" style for IDs). This helps to arrange the data for easier analysis of the data with NVivo later. We converted the table data format to text format using a convert function in Microsoft Word. As a result, Microsoft Word generated a word document similar to that shown in Figure 5.8.

4. Import the response text documents into NVivo: We imported the prepared text documents into NVivo. Completing this preparation step allowed us to explore the

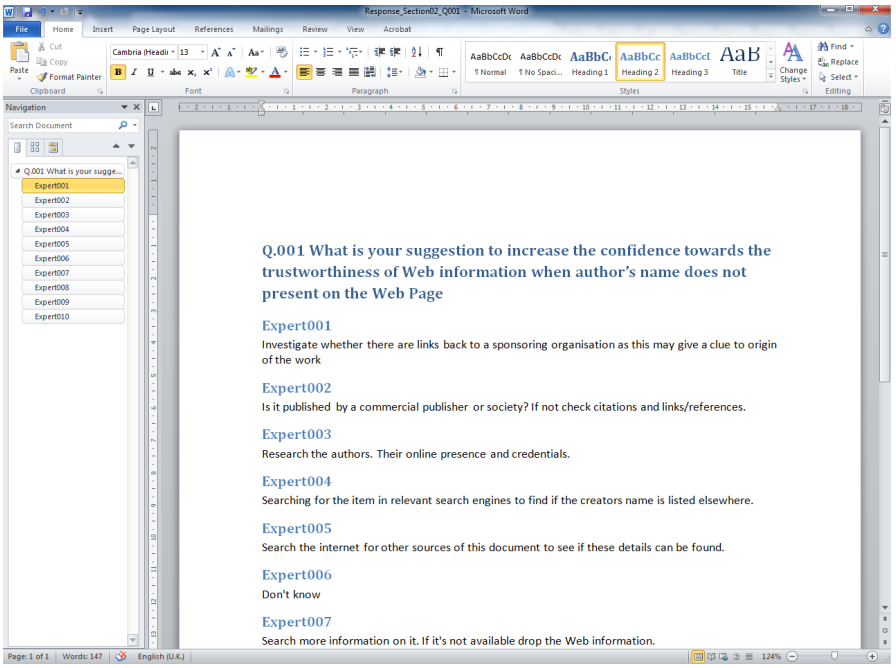


Figure 5.8: An example of the preparation of the data in Microsoft Word

data in more detail by using query functions in NVivo. Completing this preparation step allowed us to explore the data in more detail by using query functions in NVivo.

Other phases of the analysis are discussed in the details in the following sections.

5.2.2.3 Analysis of the Results from Question 2 in Section 2 of the Questionnaire

The second question of Section 2 is an open-ended question which allows experts to freely respond. We asked the experts to give a suggestion regarding how to increase their confidence in a piece of Web information if certain supportive information is not presented alongside it on the Web. We used a thematic analysis approach which was explained in section 5.2.2.1 and the prepared data from the section 5.2.2.2 to analyse the suggestions from the experts. The details of each analysis phase are described below.

5.2.2.3.1 Generating initial codes

In this phase, we developed a coding scheme by using the word frequency tools of the NVivo software. This tool creates a list of the most frequently used words in the response texts. The word frequency provides an early idea of the range of words used and may be used to indicate the most frequently expressed concepts. However, some words might be too general or might not reflect any meaningful concepts. Therefore, we used this tool

to start our coding and then we developed our initial coded nodes from the list manually in order to obtain the coded nodes that are useful for analysis in the next phase.

For NVivo, a word frequency query is a word count function which counts the number of times that each particular word or set of words appears in the responses. We can set a parameter to make the function count the occurrences of words within the response texts. In addition, we can set a parameter to make the function count the occurrences of words within the response texts and we can set a parameter to count occurrences in only the specific piece of text or the specific document in which we are interested. Moreover, we can set a parameter that determines how the results are displayed.

We set parameters for counting the appearance of words including stemmed words (the words that come from the same root of word e.g. sport: stemmed word is sporting) with in selected response texts. For displaying outputs, we set parameters to show the 50 most frequently occurring words of five characters or more in length.

The outputs from this function can be displayed in two ways. One display is a basic list of words with the number of times they have been found in the selected document as shown in Figure 5.9. The columns of the table in Figure 5.9 show words, length of words, the frequency with which that word appears in the response texts, the percentage of the number of times that word appears in the text, and the total number of words in the texts, and any words that are similar to those that the program is matching. The parameter which is set to match the word in this analysis is "including stemmed word", which means the results of the word frequency count function exactly matching with initial word and also the word that matches with stemmed words. Then, the similar words column in the result table shows the words which the program counts.

WF-count all words in survey b

Word	Length	Count	Weighted Percentage (%)	Similar Words
information	11	62	4.44	information, informed
search	6	36	2.58	search, searching
author	6	32	2.29	author, authors
content	7	28	2.01	content, contents
document	8	22	1.58	document, documents
creator	7	17	1.22	creator, creators
affiliation	11	16	1.15	affiliated, affiliation
trustworthiness	15	16	1.15	trustworthiness, trustworthy
website	7	16	1.15	website, websites
check	5	15	1.08	check
title	5	15	1.08	title, titles
available	9	14	1.00	available
google	6	14	1.00	google, googling

Figure 5.9: Word frequency count output-basic list display

Another output from the word frequency count function is a "Tag Cloud" where the words are listed in alphabetical order with the font size proportional to their frequency as shown in Figure 5.10.

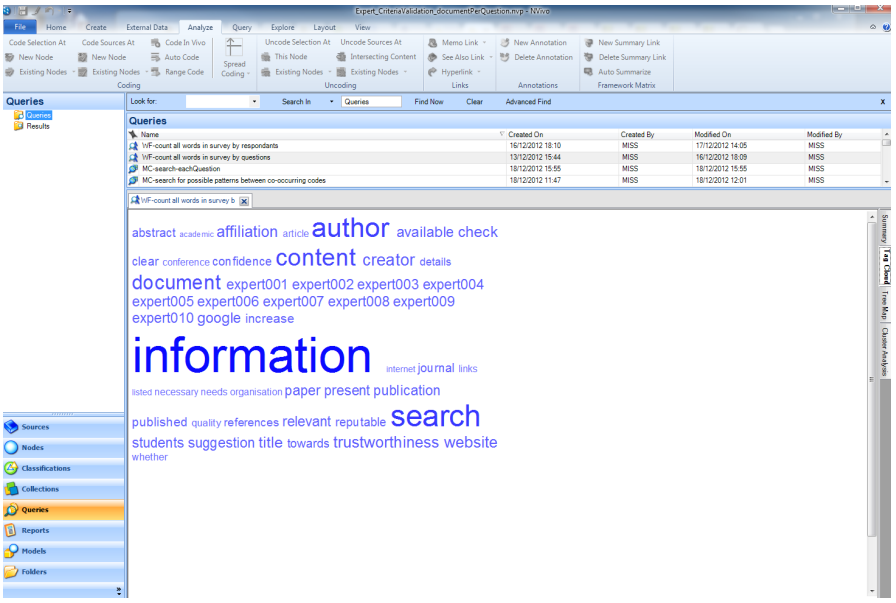


Figure 5.10: Word frequency count output-tag cloud display

Based on the tag cloud presentation, the exploratory word frequency shows five dominant words (i.e. top five words that have a big font size compared with the least dominant words) to appear; namely "information", "search", "author", "content", and "document". However, not all of the values output from this function are meaningful as indications of useful concepts because the word frequency function in NVivo uses a very basic algorithm for counting the frequency. Its algorithm matches the words with its stemmed words. If the word matches, the program increases the frequency of that word. If it does not match, the program searches for the next one, and so on. For example, the first most frequently used word as shown in Figure 5.9 is "information". Obviously, the questions in this section were designed to elicit the experts' suggestions on how to increase confidence in Web information when the specific information stated in the question is not present. Therefore, the word "information" was expected to be mentioned frequently in the responses. As a result, it appeared at the top of the list. Nevertheless, it did not lend itself to any particular thematic code. Similarly, the word "document" appeared within the top five items, but it was mentioned within the corresponding questions. Thus, its presence was determined to be irrelevant for this study and it was excluded from the initial code. By studying the tag cloud and basic list, we considered the words against the response text again as discussed above. Therefore, we excluded some words that were irrelevant to the concepts. As a result, we obtained a list of words that can be defined as initial coded nodes which are meaningful and relevant to the study, as shown in Table 5.11.

Table 5.11: Word frequency table of questions in Section 2

Word	Length	Count	Similar Words
search	6	36	search, searching

Table 5.11: Word frequency table of questions in Section 2

Word	Length	Count	Similar Words
author	6	32	author, authors
content	7	28	content, contents
creator	7	17	creator, creators
affiliation	11	16	affiliated, affiliation
website	7	16	website, websites
check	5	15	check
title	5	15	title, titles
google	6	14	google, googling
published	9	12	published, publisher, publishers
abstract	8	11	abstract, abstracts
clear	5	11	clear
reputable	9	11	reputable, reputation
references	10	10	reference, references
details	7	9	detail, details
quality	7	8	quality
article	7	7	article, articles
conference	10	7	conference, conferences
links	5	7	links
necessary	9	7	necessary
organisation	12	7	organisation, organisations
Internet	8	6	internet

We created the primary coded nodes from Table 5.11 such that the primary coded nodes corresponded to the aim of our study (i.e. what should we do to increase our confidence in Web information when some supportive data are missing). We explored more by reading the response texts to ensure that we had extracted every interesting code. In addition, we were interested to discover any themes that might arise across multiple questions. Consequently, we designed our coded nodes to use a unified coding scheme which is common to all questions in each section. As a result, we obtained the coded nodes displayed in Table 5.12. At the completion of this phase, we generated 55 coded nodes.

Table 5.12: A list of all coded nodes in Section 2

Coded nodes	Sources	References
a copy of the document	1	1
abstract	2	4
affiliation	9	10
author's active year	1	1

Table 5.12: A list of all coded nodes in Section 2

Coded nodes	Sources	References
author's homepage	4	7
author's name	5	12
author's publications	3	4
assume	3	4
background information	1	1
biography	1	1
check	8	15
citations	3	4
conclusion	2	3
confirm	2	2
consult	2	2
contact	1	1
content	4	6
content Of author affiliation	9	9
date of references	2	2
difficult	3	4
discard	4	5
find	6	8
google	7	10
google scholar	2	4
google search engine	6	9
Internet	5	5
introduction	1	1
investigate	2	4
judge	1	1
layout format	1	1
links	5	5
not affect	1	1
not available	11	11
not necessary	7	9
not trustworthy	1	1
organisation page	4	4
other publications on the same website	1	1
other sources	2	4
page information	1	1
profile page	3	3
publisher	4	7

Table 5.12: A list of all coded nodes in Section 2

Coded nodes	Sources	References
read	3	5
references	2	3
research group page	2	2
search	11	38
search engine	6	6
similar papers	1	1
style and tone	1	1
the original publication date	1	1
title	2	3
type of publication	1	1
university website	2	2
Web address	2	2
Web of Science [12]	1	1
website	1	1

Table 5.12 shows the coded nodes, the number of sources (response texts) that each coded node appears in and the number of times that each coded node has been referenced. For example, the coded node "google search engine" appeared six times in response texts (namely in the responses to questions 1, 2, 3, 4, 5, and 8) and it was referenced nine times as shown in Figure 5.11.

5.2.2.3.2 Structure coded nodes scheme

From Table 5.10, we noticed that some coded nodes were an action (i.e. a verb) that the expert would perform in order to increase their confidence in the Web information, whereas others were items that related to the target documents and their metadata (e.g. a noun or an adjective). Other coded nodes included tools that support the action (i.e. a noun), and some were the expert's reaction to the result of the action (i.e. a verb or an adjective). However, there had been a case where some coded nodes can be either a noun or a verb. We defined this type of coded node based on the context in the response answer from the participants. For example in Figure 5.11, the participants mentioned the word "google" in the context of looking for more information. Therefore, the coded node "google" was defined as an action of the participant. Conversely, the participants mentioned the word "google" in conjunction with other words (descriptors), which indicated they were referring to it as a tool that they could use to help search for

[12]Web of Science is an online scientific citation search service that covers multidisciplinary contents of the journals and conference proceedings (Thomson Reuters, 2008)

Name: google search engine

<Internals\\Survey\\Response_Section02_Q001> - § 1 reference coded [1.87% Coverage]

Reference 1 - 1.87% Coverage

Googling the title

<Internals\\Survey\\Response_Section02_Q002> - § 2 references coded [3.30% Coverage]

Reference 1 - 1.65% Coverage

Googling the name

Reference 2 - 1.65% Coverage

Google the author

<Internals\\Survey\\Response_Section02_Q003> - § 2 references coded [3.05% Coverage]

Reference 1 - 1.84% Coverage

Googling the author's name

Reference 2 - 1.20% Coverage

Google the author

<Internals\\Survey\\Response_Section02_Q004> - § 1 reference coded [1.90% Coverage]

Reference 1 - 1.90% Coverage

Google the author

<Internals\\Survey\\Response_Section02_Q005> - § 2 references coded [5.13% Coverage]

Reference 1 - 2.47% Coverage

Googling the author's name

Reference 2 - 2.66% Coverage

Google author or affiliation

<Internals\\Survey\\Response_Section02_Q008> - § 1 reference coded [2.27% Coverage]

Reference 1 - 2.27% Coverage

Googling the author's name

Figure 5.11: The references of the google search engine coded node within the response texts

more information, such as "google search engine" or "google scholar". Thus, we defined the coded node "google search engine" as a tool.

Consequently, we categorised these preliminary coded nodes into four groups; namely, tools, actions, responseToAction, and itemsRelatedDocuments. We grouped the coded nodes based on their type and meaning. We created groups of coded nodes as a hierarchical structure as shown in Figure 5.12.

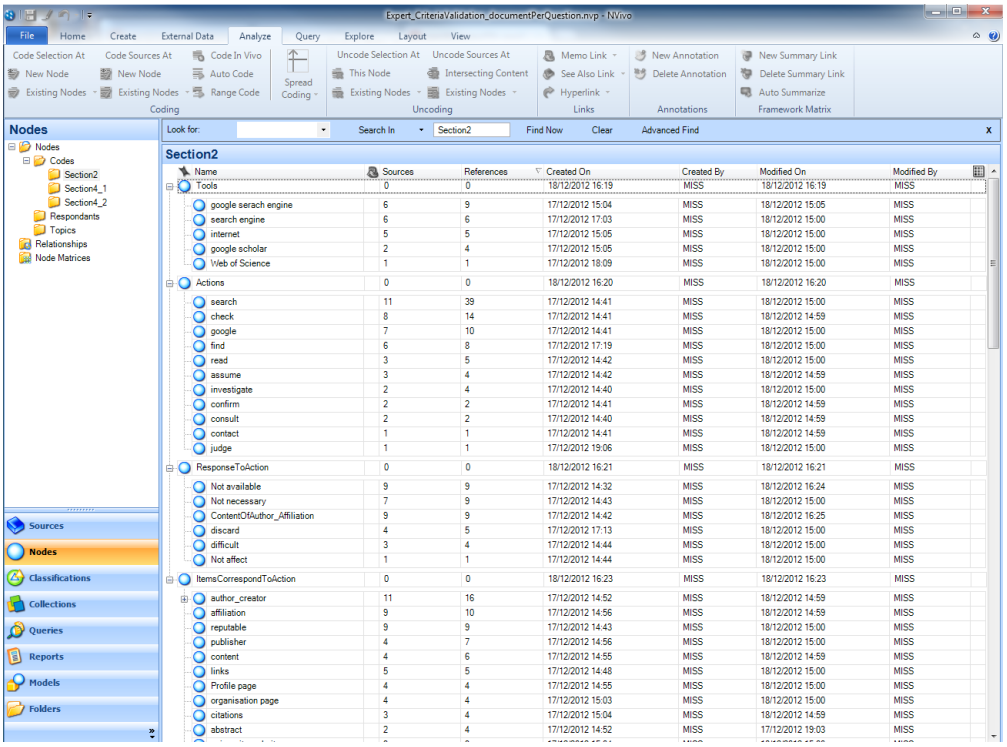


Figure 5.12: A list of all code nodes of question 2 in Section 2

The details of the members in each group are described as below:

- *Tools group*: This group contained coded nodes which pertained to the actual tools that experts would use in order to increase their confidence in the Web information. Table 5.13 shows the coded nodes within this group as well as the number of times each of the coded nodes has been referred to by the experts.

Table 5.13: Coded Nodes and their frequencies that are categorised with the tools group

Coded Nodes	No.Sources	No.References
google search engine	6	9
search Engine	6	6
Internet	5	5
google scholar	2	4
Web of Science	1	1

- *Actions group*: This group consisted of coded nodes which pertained to the experts' performance in order to evaluate the trustworthiness of Web information when the supportive information (e.g. author's name, author's affiliations, editorial process, etc.) is not present. Table 5.14 shows the coded nodes within this group as well as the number of times each of the coded nodes has been referred to by the experts.

Table 5.14: Coded Nodes and their frequencies that are categorised with the actions group

Coded Nodes	No.Sources	No.References
search	11	39
check	8	14
google	7	10
find	6	8
read	3	5
assume	3	4
investigate	2	4
confirm	2	2
consult	2	2
contact	1	1
judge	1	1

- *ResponseToAction*: This group consisted of coded nodes which relate to the experts' reactions to the result of an action. Table 5.15 shows the coded nodes within this group as well as the number of times each of the coded nodes has been referred to by the experts.

Table 5.15: Coded Nodes and their frequencies that are categorised with the ResponseToAction group

Coded Nodes	No.Sources	No.References
content Of author affiliation clear	9	9
difficult	3	4
discard	4	5
not available	11	13
not necessary	8	15
not trustworthy	1	1

- *itemsRelatedDocuments*: This group contained coded nodes that relate to the items that affect the experts' performance in order to increase their confidence in the Web information. Table 5.16 shows the coded nodes within this group as well as the number of times each of the coded nodes has been referred to by the experts.

Table 5.16: Coded Nodes and their frequencies that are categorised with the itemsCorrespondToAction group

Coded Nodes	No.Sources	No.References
a copy of the document	1	1
abstract	2	4
author	1	1
author's active year	1	1
author's affiliation	1	1
author's homepage	5	10
author's name	5	15
author's publications	2	3
background information	1	1
citations	3	5
conclusion	2	3
content	6	7
date of references	2	2
introduction	1	1
layout format	1	1
links	5	5
organisation page	4	4
other publications on the same website	1	1
other sources	11	20
page information	1	2
profile page	0	0
publisher	4	7
references	2	3
research group page	2	2
similar papers	2	2
style and tone	1	1
the original publication date	1	1
title	2	3
type of publication	1	1
university website	0	0
Web address	2	2
website	1	2
a copy of the document	1	1

5.2.2.3.3 Searching for themes

According to the purpose of the study, we must try to identify the patterns that experts use to increase their confidence in Web information when some information is missing. This helps to elicit the process that should be used by our framework to evaluate the trustworthiness of information on the Web for naive users. Therefore, we focus on the "action" coded nodes which indicate the processes used by the experts in order to evaluate the Web information.

As a consequence, we investigated 11 coded nodes in the action group (details in Table 5.14) using a cluster analysis which is a technique used to explore how similar the coded nodes are. In this analysis, we used Jaccard's coefficient (Jaccard, 1901) to compare the similarity of words in coded nodes in the action group. Jaccard's coefficient measures the similarity between two sets, and is defined as the cardinality of the intersection of the two sets divided by the cardinality of the union of those two sets. We measured the difference between words by treating those words as sets of letters. We are interested in whether there are some coded nodes that are similar. If any coded nodes are similar, they are be clustered together. The results are displayed as a horizontal dendrogram where coded nodes that are similar are clustered together on the same branch and less similar coded nodes are further apart as shown in Figure 5.13.

Figure 5.13: Nodes clustered by coding similarity

From the results of the cluster analysis based on code similarity, we found that the coded nodes "find" and "google" were similar. Likewise, the coded nodes "check" and "search" were also similar. Those four coded nodes referred to the act of looking up more information, which corresponds to the coded node "investigate" which was a higher level of similarity. Therefore, we categorised these coded nodes into a group called "investigate". Correspondingly, other coded nodes which were clustered together were

also grouped; namely, "confirm", "assume", and "read". In summary, we identified four main groups of action: investigate, confirm, assume, and read. The theme, "investigate", consisted of sub-actions such as check, search, find, and google, as shown in Figure 5.14. We explored each case of missing data in more detail to discover any sub-themes that

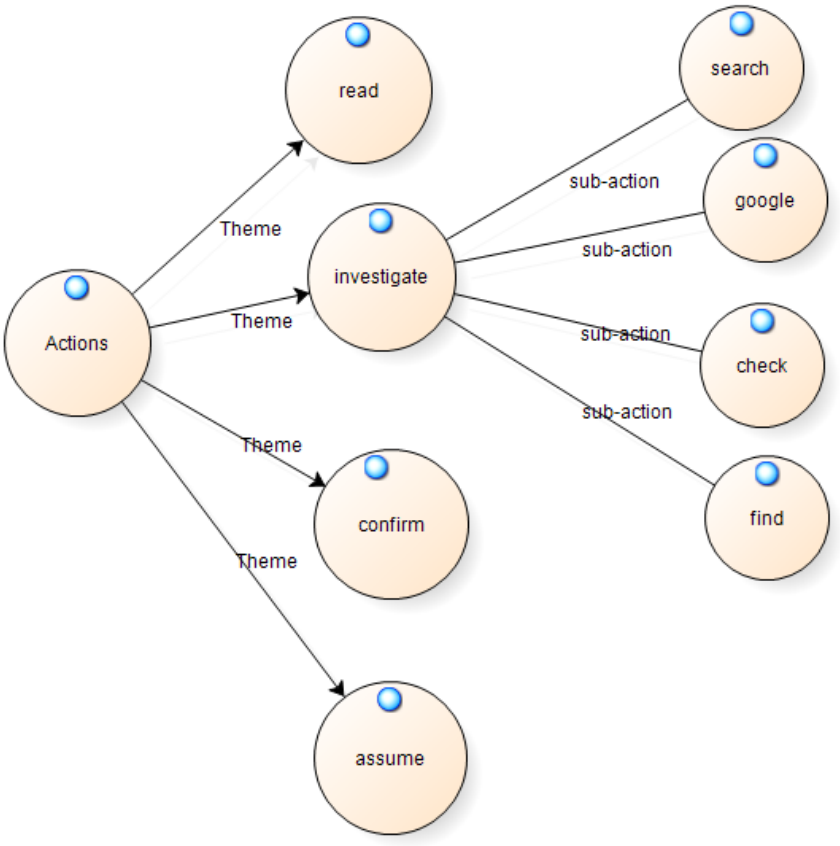


Figure 5.14: The initial main themes

might exist. We discuss this in the next section.

5.2.2.3.4 The trustworthiness evaluation process when supportive information from the trustworthiness criterion is not present on the Web

We investigated the process of gaining confidence in information when some data from the trustworthiness criterion are not available. The investigation process consisted of three steps:

1. Modelling the coded nodes: we analysed the response texts in each case of missing supportive data from the trustworthiness criterion using a model function of NVivo which is a tool that can present coded nodes in a visual way in order to help to identify patterns and relationship across of the response texts. The modelling displayed all of the coded nodes in the form of a circle shape and its codes.

We used our classification groups (discussed in section 5.2.2.3.2) to categorise the coded nodes. For example, in the case that the supportive data from the trustworthiness criterion "author/creator's name", was not present, the modelling can be illustrated in Figure 5.15. The coded nodes were grouped into four groups according to the structure node schema. The actions group mentioned the coded nodes "search", "check", "investigate", and "google". The responseToAction group mentioned the coded nodes "not available", "not trustworthy" and "discard". The tools group referred to the coded nodes "google search engine", "general search engine", and "internet". In addition, the itemsRelatedDocuments mentioned the publisher, content, title, links, other sources, Web address, citations, references, organisation page, and author's homepage.

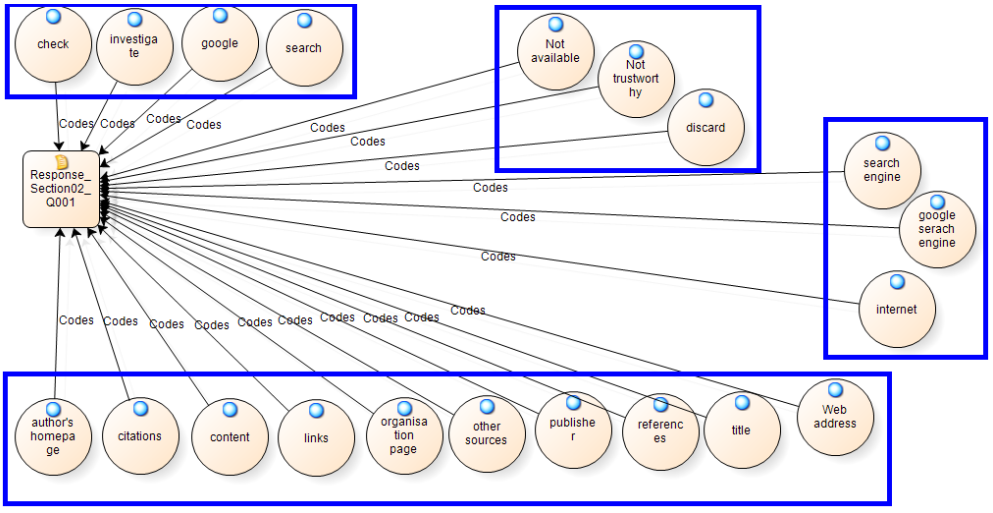


Figure 5.15: A model of coding related to question 1: supportive data from the trustworthiness criterion "author/creator's name" is not present

2. Checking with the initial main themes: we matched the groups from the modelling results in the previous previous step with the initial main themes we defined in Figure 5.14 in order to investigate the themes which occur in each case. For instance, from the modelling result in Figure 5.15, according to the initial main themes discussed in section 5.2.2.3.3, the actions which were mentioned in this question fall under the same theme as the proposed "investigate" theme. Therefore, we applied the theme 'investigation' into this question.

3. Investigating the patterns: we investigated in more detail what experts look for when they investigate for more information to increase their confidence in the trustworthiness of Web information. We used a Matrix coding query in NVivo to help us to compare how the different actions correspond to items. This helps to indicate the patterns that experts use. The Matrix coding query compares between the coded nodes from different groups in order to find the relationships between them. There are a number of relationship types: one coded node appears next

Table 5.17: The results of Matrix coding query of question 1 in Section 2

	other-source	refer-ences	dis-card	auth-or's	home-page	cita-tions	con-tent	links	organ-isat-ion page	title	web ad-dress	not avail-able	not trust-wor-thy	pub-lisher
check	0	2	1	0	1	0	1	0	0	1	0	1	0	0
google	0	0	0	0	0	0	0	0	1	0	0	0	0	0
invest-igate	0	0	0	0	0	1	0	1	0	0	0	0	0	0
search	3	0	1	1	0	0	0	0	0	0	1	0	0	0
sum	3	2	2	1	1	1	1	1	1	1	1	1	0	0

to another coded node; one coded node appears in front of another, or one coded node is surrounded by another codes node. These type of relationships help to indicate the patterns that might emerge in the process of assessing trustworthiness of Web information suggested by the experts. Matrix coding queries create tables to compare multiple pairs of specific items in a matrix. Each cell in the matrix represents a coded node containing the content coded at the intersection of the row and column. For example, the query results of the case in Figure 5.15 are shown in Table 5.17. The column represents the elements which the expert suggested to look for when the author's or creator's name is not provided on the Web. The row represents the expression of the expert's action. The number in each cell is the number of coding references at the intersection of an action and an element.

Table 5.17 shows the top elements of investigation when the author's or creator's name is not present on the page as "other sources". This means that experts suggested searching for the author's name or the author's details from another source. For example, expert004 mentioned "Searching for the item in relevant search engines to find if the creators name is listed elsewhere." The second is references, which can help to ensure that a piece of information has good evidence to support its content. For instance, expert002 mentioned, "Is it published by a commercial publisher or society? If not check citations and links/references." Similarly, citations and links on the Web can help to indicate the trustworthiness of Web information. The author's homepage and organisation page are another source that can be used to look for the author's name in order to gain more confidence in this piece of information. For example, expert 003 suggested, "Research the authors. Their online presence and credentials"; and expert001 said, "Investigate whether there are links back to a sponsoring organisation as this may give a clue to origin of the work." Expert008 suggested, "Googling the title" which means the expert recommended using the title of the information or article to find out the author's name.

In addition, we noticed that the coded nodes "search" and "check" have been coded with the surrounding coded nodes "discard", "not available" and "not trustworthy". Therefore, we explored each coded node in more detail, and find that the experts had a specific process of searching whereby, if there is no information

from other sources, the information was discarded. For example, expert007 said, "Search more information on it. If it's not available drop the Web information." Similarly, expert009 cautioned "Look at the web address. Discard if not trustworthy." This indicates that if the Web address is not familiar or it does not have a good reputation, the information should be discarded. Therefore, if the additional information that is used to support the decision is not available or not reputable, the piece of information should be discarded. As a result, we obtained suggested patterns of assessing the trustworthiness of Web information when the author's name is missing, as displayed in Figure 5.16.

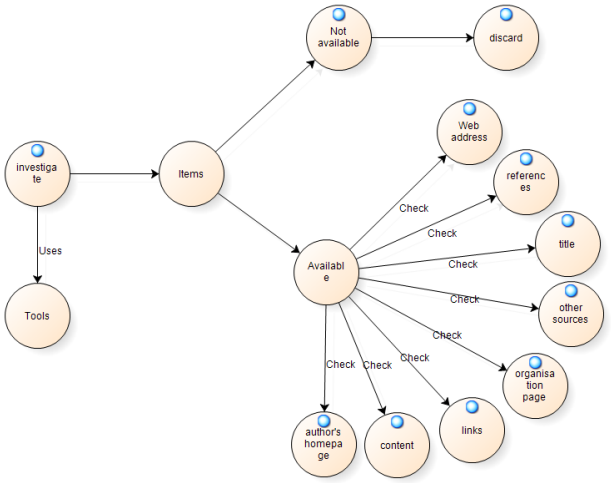


Figure 5.16: The investigation process to increase the user's confidence in the Web information when the author's name is missing

4. Reviewing the patterns: after we obtained the patterns, we investigated in more detail by checking the response texts against the patterns again in order to assure that we discovered any interesting patterns that might be presented. Completing this step, we might be able to find a new pattern which can develop to become sub-themes. For example, referring to the patterns discovered in a previous section (see Figure 5.16),we found that if the users could find the other elements that can help them support their evaluation, they assessed the trustworthiness of Web information based on that elements. However, if they could not find any of them or if the source does not have a good reputation, they discard the information. Therefore, we set the "discard" action as a sub-theme of the "investigate" theme as displayed in Figure 5.17.

Broadly following the same steps as above, we investigated each of the next ten elements of question 2 which asked experts' suggestions for gaining confidence in the trustworthiness of Web information when supportive information of that item is not available. The patterns derived from each item are:

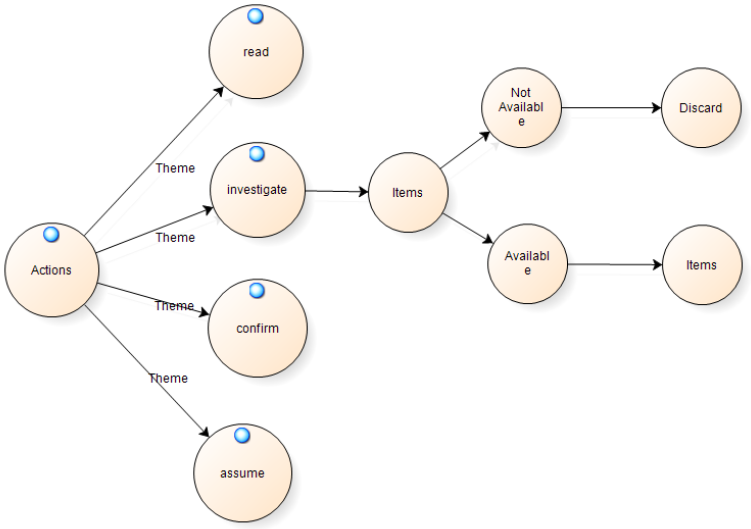


Figure 5.17: The initial main themes and discarded sub-theme

- **Author's/creator's affiliation**: The results from the model function consisted of the same groups as discussed in section 5.2.2.3.2. In addition, in the action group, we found that it fell into two main themes as we proposed in section 5.2.2.3.3, which were "confirm" and "investigate". In addition, the results of the Matrix query showed that the first element to investigate when the author's or creator's affiliation is not present on the page was the author's or creator's name. Most of the experts suggested searching for the affiliation of the author using the author's name. For example, expert004 said, "Searching using the creator's name in relevant search engines to find their affiliation listed elsewhere" and expert008 mentioned, "Googling the name." Other information which can be used to increase the confidence in the Web information when author's affiliation is missing was the author's homepage. This provided other supportive information including the author's publication lists and links, and their organisation to help judge the trustworthiness of Web information. This question showed that experts evaluated the trustworthiness of Web information when author's or creator's affiliation is missing using two approaches; investigation and confirmation. This corresponds to the initial main themes we discussed in section 5.2.2.3.3. However, for the "investigate" theme, we could not find a new sub-theme. The process of investigation to increase the user's confidence in the Web information when the author's affiliation is missing displayed in Figure 5.18.

- **Author's/creator's title**: the result from the Model function still consisted of four groups of coded nodes. However, we found new coded nodes in the responseToAction group which were "not necessary" and "content of author affiliation clear." We investigated in more detail by using a Matrix coding query. The result of the query showed that searching for the title or for more information about the author using their name was the top suggestion from the experts. Others

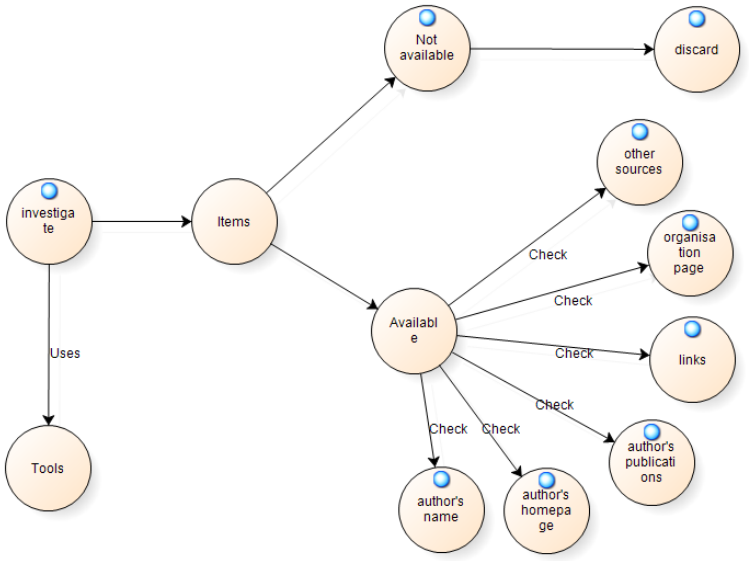


Figure 5.18: The investigation process to increase the user's confidence in the Web information when the author's affiliation is missing

were the author's homepage or the author's organisation page which could provide supportive information. For example, expert 001 mentioned, "A quick internet search on the author's name may shed some light on this."; while expert004 said, "Searching using the creators name in relevant search engines to find their title listed elsewhere." Similarly, expert008 suggested "Googling the author's name." In addition, expert002 mentioned about looking for more information from the author's homepage by saying "Try to find the author's details (e.g. profile pages on university websites). This is easier for less common names!"

We identified an issue, however, that was raised by the experts in the case of author's title being missing. They suggested that we could find more information to gain confidence but it was not necessary to do it. For example, expert003 mentioned, "Search but not vital"; while expert001 said,

> A quick internet search on the author's name may shed some light on this. If the website is affiliated to an organisation then a search of that organisation's website may help discover with the author has an appropriate academic qualification. It's worth noting that merely having an academic title does not automatically give the information a stamp of approval: the qualification should be in an appropriate area for the information concerned and it needs to be borne in mind that people without a PhD can write reliable good quality information - just as professors can write ill-informed biased misinformation.

As a result, if the author's title is missing, we can look for more information by searching via their name. Nevertheless, if we cannot find information about

the author's title, it is not necessarily a problem. We still can use this piece of information as long as the author's name and their affiliation are stated clearly on the page as mentioned by expert007: "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable". The process of investigation to increase the confidence of Web information when the author or creator's title is not provided can be seen in Figure 5.19.

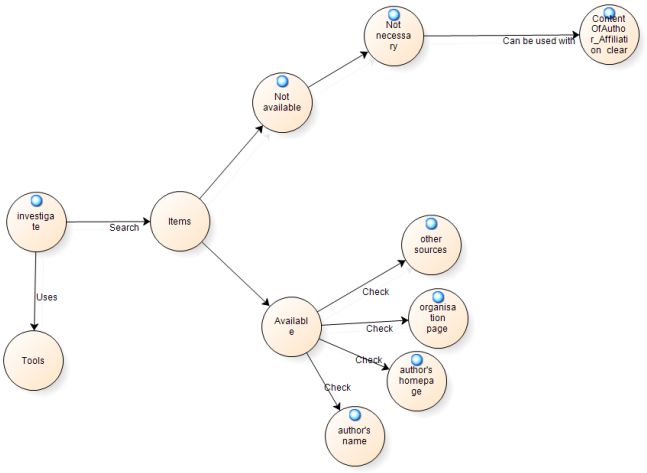


Figure 5.19: The investigation process to increase the user's confidence in the Web information when the author's title is missing

At the end of this process, we identified a new sub-theme which should be added in to our main themes. The new theme was "not necessary" which means the information can still be used so long as the author's name and author's affiliation are provided, as shown in Figure 5.20.

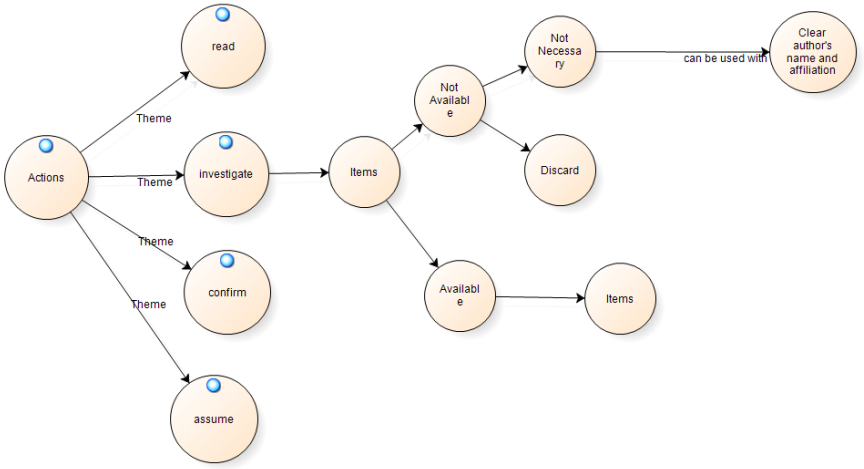


Figure 5.20: The initial main themes and "not necessary" sub-theme

- **The content of the author's/creator's experience**: The result of model function was composed of four groups of coded nodes which represent the actions, the response of the action, the elements corresponding to actions, and the tools used within the action. We then used a Matrix coding query to explore the relationship between elements. In the case that the details of the author's experience were not shown on the page, the experts suggested looking for more information or background information using the author's name. This could turn lead to other sources of information about the author which could provide information on their level of experience. For example, expert004 said, "Searching using the creators name to find a blog or biography that might reflect their experience", and expert010 mentioned, "Google the author". Alternatively, experts also suggested that it is not necessary to have this information and that if the information could not be found, a piece of information can still be used if the author's name and affiliation are stated on the page. For example, expert006 said, "Not necessary", and expert007 mentioned, "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable". The process of investigation to increase the confidence of Web information when the details of the author's experience are not provided is shown in Figure 5.21.

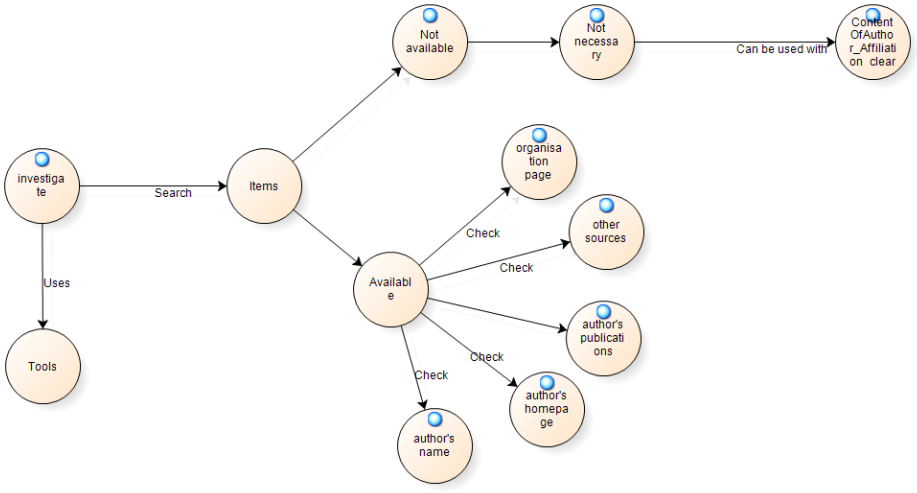


Figure 5.21: The investigation process to increase the user's confidence in the Web information when the details of the author's or creator's experience are missing

- **Author's/creator's contact details**: the result of model function showed that the coded nodes related to this question were still categorised into four groups corresponding to the groups from section 5.2.2.3.3. We then explored the relationships between action nodes and elements using a Matrix coding query. We found that using the author's name to search for their contact details was the most mentioned suggestion from the experts. For example, expert004 said, "Searching

using the creators name in relevant search engines to find their contact details listed elsewhere", and expert008 said "Googling the author's name". In addition, the author's homepage was a useful source that could provide information on author's contact detail as mentioned by expert001, who said "I would suggest that the student does a search on the internet for the author's name and also check's the author's home institution website (if one is given)"; and expert002 also said "Try to find the author's details (e.g. profile pages on university websites)". However, according to some experts, the author's contact details did not necessarily have to be confirmed. For example, expert007 said, "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable". The process of investigation to increase the confidence of Web information when the author's contact details are not provided is shown in Figure 5.22.

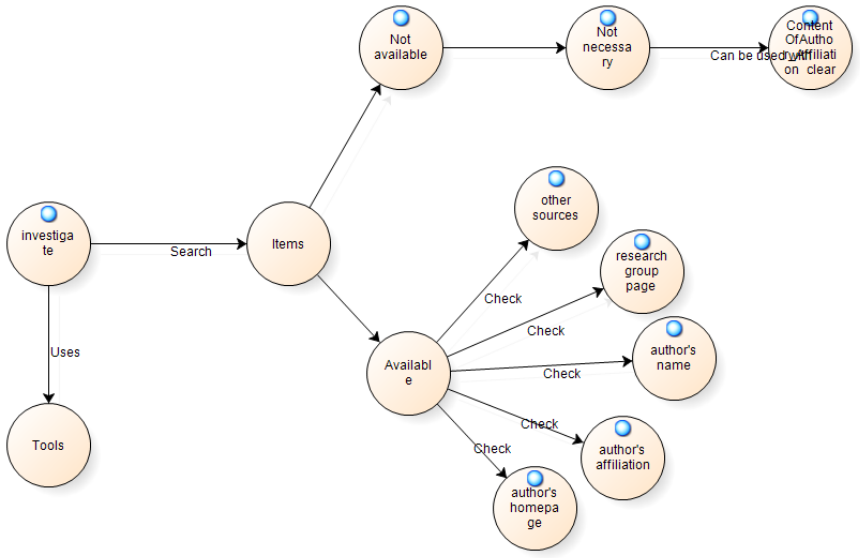


Figure 5.22: The investigation process to increase the user's confidence in Web information when the content of the authors or creator's contact details is missing

- **The number of citations**: the results of model function showed that the expert suggested new tools (Google Scholar and Web of Science) that could help the user to search for more information on the number of times that information has been cited. The results of Matrix query showed that the experts suggested looking for the citations that can be retrieved from specific tools (namely Google Scholar or Web of Science). For example, expert002 said, "Check citations in Web of Science and Google Scholar"; expert004 said, "Use Google scholar to get an impression of how often the item has been referenced"; and expert005 said, "Google scholar or other databases might contain this information". In addition, the experts also recommended searching for other sources or similar information

that could corroborate the information. Expert005 said, "I'd advise the student to look for similar papers and see how well it corroborates with information from those papers", and expert010 said, "Use Web search to find work referencing the content". However, it was quite difficult to search for other information to support the judgment unless it was stated as expert001 said; "Unless the information in question is a formally published academic paper then I don't see how they can do this (if it is of course Web of Science is the place to look)", expert005 said "Google scholar or other databases might contain this information. If nowhere contains this information nothing can be assumed", and expert008 said "Difficult". Still, expert007 recommended that the information can still be used even if it does not have information regarding how often it is cited as long as the author's name and affiliation are provided. As expert007 mentioned, "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable". Alternatively, expert006 mentioned "Not necessary". That is, it was not necessary to search for more information to support the decision if the number of citations is missing. The process of investigation to increase the user's confidence of Web information when the number of times that the information has been referenced is not provided is shown in Figure 5.23.

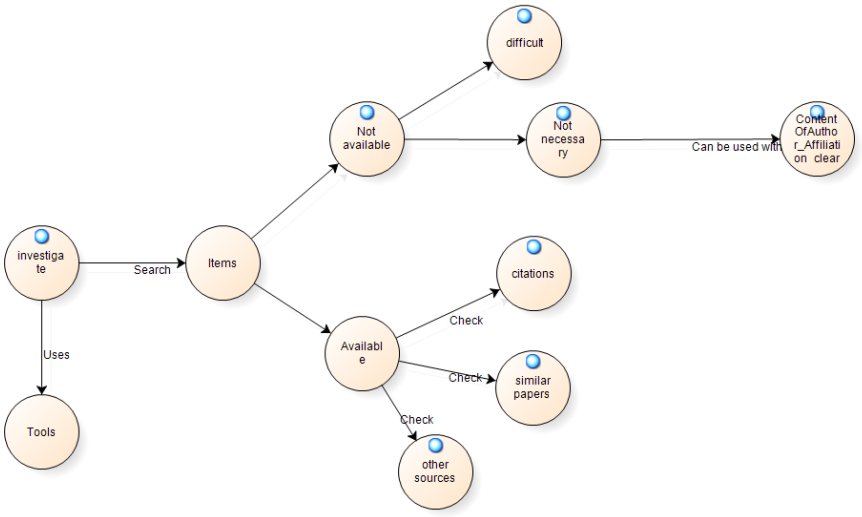


Figure 5.23: The investigation process to increase the user's confidence in the Web information when the number of citations is missing

- **The editorial process**: From the Model function, coded nodes were grouped into four groups as discussed in section 5.2.2.3.3. In addition, we found that the actions group showed an action which related to the "assume" theme, which was an action that sets up some assumption about the information. Also, it made reference to the the theme "confirm" which was a suggestion to consult the Website to find out about the editorial process. The experts' top recommendation for finding information regarding the editorial process is to search the publisher's Website. In

this regard, expert010 said, "Check publication venue and the editorial processes they employ", and expert004 suggested, "Look up the conference or journal of where the item was published for details on their editorial process." Moreover, the experts suggested searching for this information on the site on which the content was published. For example, expert002 said; "For articles and books look on publishers website. For other material look at the item and the website it belongs to". However, it seemed to be difficult to search for this kind of information from other sources unless it was stated in the page. Alternatively, the editorial process can be inferred based on the type of information; as expert005 suggested, "This can be inferred from the mode of publication (journal, conferences etc. are almost certain to be peer-reviewed whereas self-published documents are most likely not peer reviewed)". Nevertheless, if the editorial process information could not be located, it still could be used as long as the author's name and affiliation are presented, as expert007 recommended "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable". The process of investigation to increase the user's confidence of Web information when the editorial process content is not provided is shown in Figure 5.24.

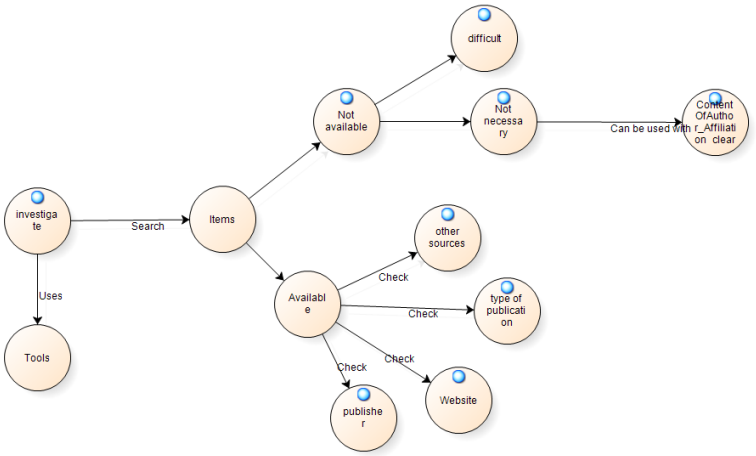


Figure 5.24: The investigation process to increase the user's confidence in the Web information when the editorial process content is missing

In addition, the experts mentioned actions in two of the main themes proposed in section 5.2.2.3.3. We explored and found the activity that should be a sub-theme of the "assume" theme (setting an assumption about the information) which was when an expert assumed the editorial process based on the type of publication. Therefore, we named the additional information such as mode of publication as a piece of background information and set it as a sub-theme of "assume" theme as shown in Figure 5.25. Moreover, the expert mentioned the "confirm" theme and suggested that consulting the publishing source such as conference- or journal's

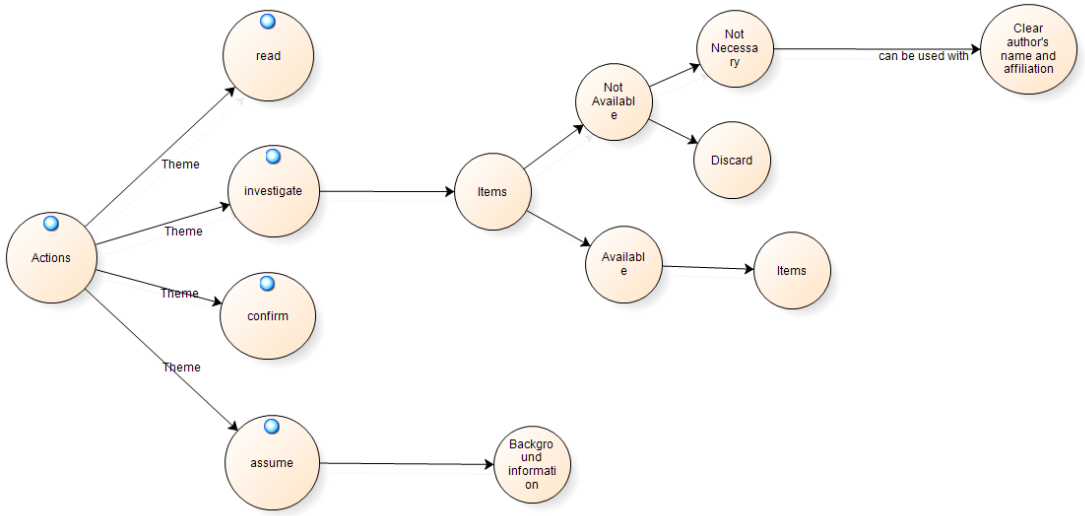


Figure 5.25: The initial main themes and background information sub-theme

website would be a sub-process of the confirm action. Therefore, we set consulting source as a sub-theme of the confirm theme, as shown in Figure 5.26.

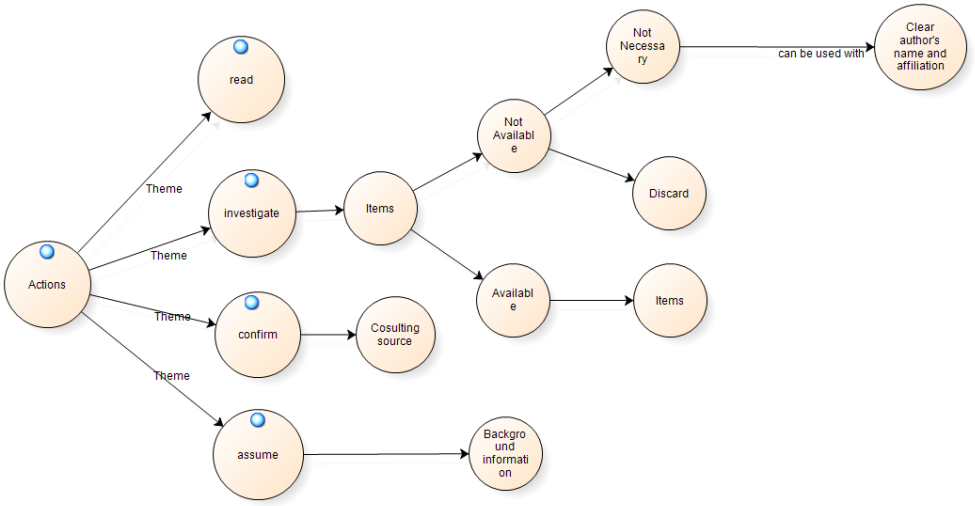


Figure 5.26: The initial main themes and consulting source sub-theme

- **The publication date of content**: the results of Model function showed that four groups of coded nodes related to this question. The Matrix query result showed that the experts suggested searching for more information from other sources by using the author's name and checking the publisher. As expert004 said, "Find the date either by searching using relevant search engines or consulting the journal/conference of publication", expert007 said, "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable", and expert008 said "Googling the author's name". In addition, checking the links or references that were provided

in the information could be of help to indicate the publication date of information. Expert002 suggested, "Check date of references. Check links work" and expert005 recommended, "You can get an idea of a timespan in which a paper was published by looking for similar papers from the same author or finding the years in which the author was active". However, it was difficult to find this information if it was not stated. Expert001 said, "Unless stated this information will be almost impossible to obtain accurately". On the other hand, some experts suggested that it was not necessary to find this information if it was not provided. Expert007 also said "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable". In particular, one expert recommended that we could gain confidence about how recent information was by checking the content. As expert010 said, "Verify that the information given might not have become outdated". The process of investigation in order to increase the confidence of Web information when the publication date of the content is not provided is shown in Figure 5.27.

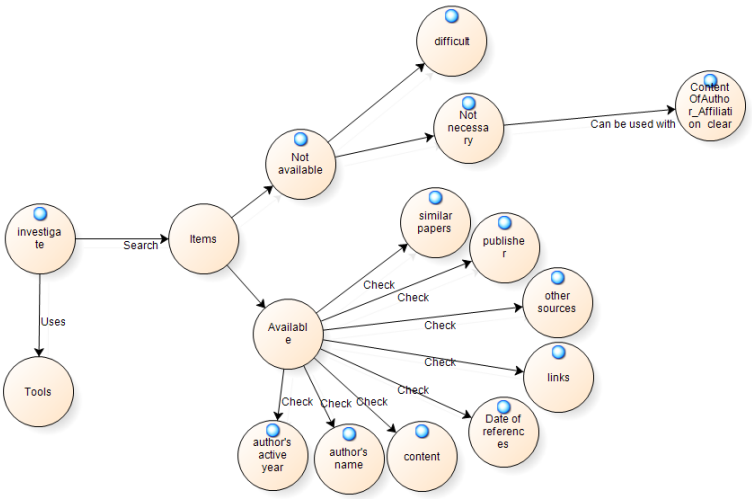


Figure 5.27: The investigation process to increase the user's confidence in the Web information when the publication date of the content is missing

- **The last modification date of content**: the result of the Model function showed that the coded nodes in this question corresponded to the main themes which we discussed in section 5.2.2.3.3. It did not state any particular tool to use in the process. However, we could refer to the previous processes which suggested using a search engine to find more information to support the decision. The result from the Matrix query showed that the experts suggested that the page that hosts the data was the best source to get information about the last modification date. As expert001 suggested, "Check the page information on your web browser". Similarly, the experts recommended that the last modification date could be assumed by the date on which the information is published and thus it can be checked from

the page information. As expert006 mentioned, "Assume it was last modified on the publication date". In addition, links and the date of the references in the information could indicate how recently the information has been changed as expert002 said; "Check date of references. Check links work". Similarly, a copy of the document or the original publication date could be used to determine whether the document has been changed since it was peer-reviewed. Expert005 said, "Find a copy of the document that has not been modified since the peer review process" and expert010 recommended "Check that the original publication date is recent or that the content is unlikely to be outdated". However, it was not necessary to look for other information to support the judgement; as expert007 suggested, "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable", and expert008 said "Not important". The process of investigation to increase the user's confidence in the Web information when the last modification date of the content is not provided is shown in Figure 5.28.

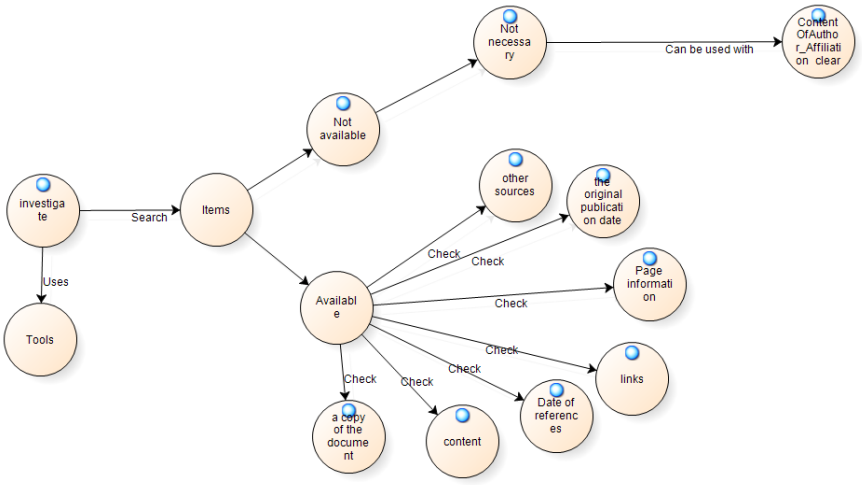


Figure 5.28: The investigation process to increase the user's confidence in the Web information when the last modification date of the content is missing

In addition, we found a new sub-theme that should be added to the "assume" theme based on an expert's suggestion that the process of assuming might have a basic function such as expert006 recommended, "Assume it was last modified on the publication date". Therefore, we set "assumption" as a sub-theme of the "assume" theme as shown in Figure 5.29.

- **The title of the content**: in this question, the experts did not mention tools that they used to help them gather other information. We explored in detail the relationship between action coded nodes and element nodes using the Matrix query. The results showed that the experts mentioned that abstract, conclusion, introduction and content of the information are other options to look at in order to judge the relevance of information in case the title of the content does not

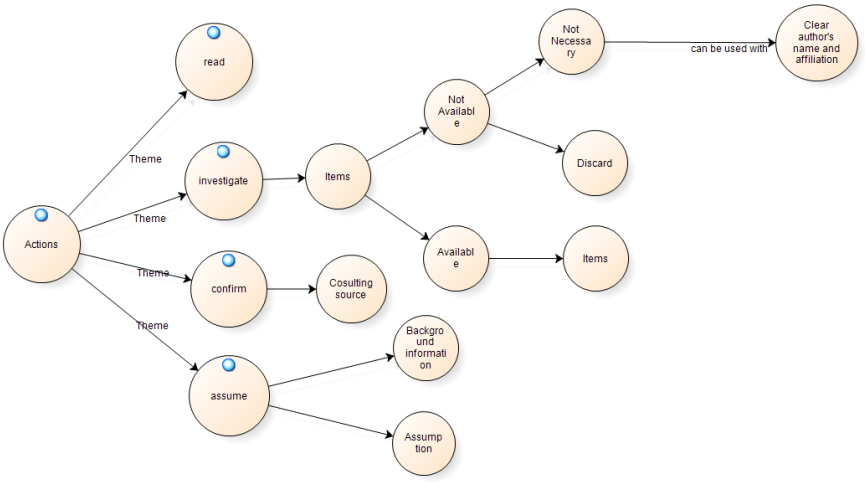


Figure 5.29: The initial main themes and the assumption sub-theme

provide this. Expert002 suggested, "Read abstract/introduction and conclusions", expert004 said, "Consult the contents abstract or listed summary to ensure this is the document needed", and expert001 suggested "Read through the information you have at hand and critically appraise it". However, there was a conflict among the experts on the issues of whether or not the title matches the user's needs affects the trustworthiness of the information. Expert005 considered that lack of attention to naming the information properly means the information was likely to be untrustworthy:

> If the title doesn't match your needs then the content of the paper probably doesn't either so find another document? If the content does in fact match your needs but the title doesn't it suggests the paper wasn't carefully written and therefore another paper should be found as this one isn't likely to be reliable.

On the other hand, some experts mentioned that it was not necessary to have a title that matches the user's needs. The information could be used if the author's name and affiliation was clear as expert007 suggested, "Search more information on it. If it's not available the students could use the information as long as content creator and affiliation are clear and reputable", expert010 said "This shouldn't affect trustworthiness" and expert006 mentioned "Not necessary". The process of investigation to increase the user's confidence in the Web information when the title of the content is not provided can be seen in Figure 5.30. In addition, most of the experts suggested reading or checking the abstract, introduction, or conclusion to increase the confidence of the trustworthiness of the Web. This action corresponded to the "read" theme which is one of the proposed main themes. Therefore, we set these key areas as a new sub-theme into the "read" theme, as shown in Figure 5.31.

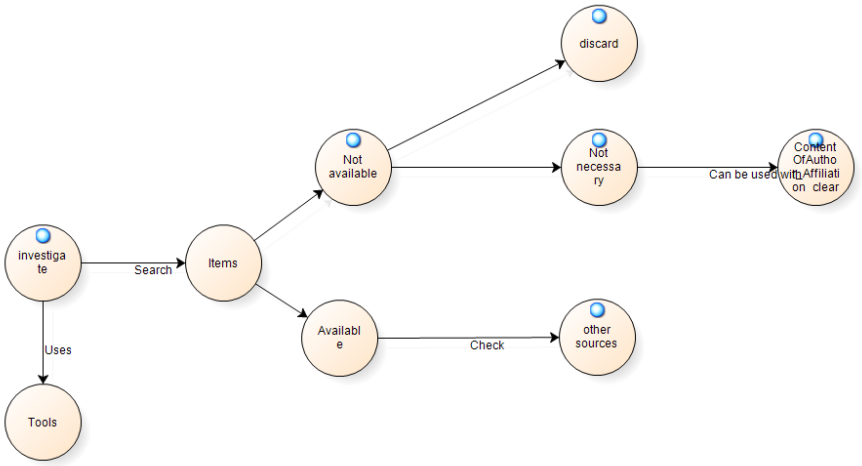


Figure 5.30: The investigation process to increase the user's confidence in the Web information when the title of the content is missing

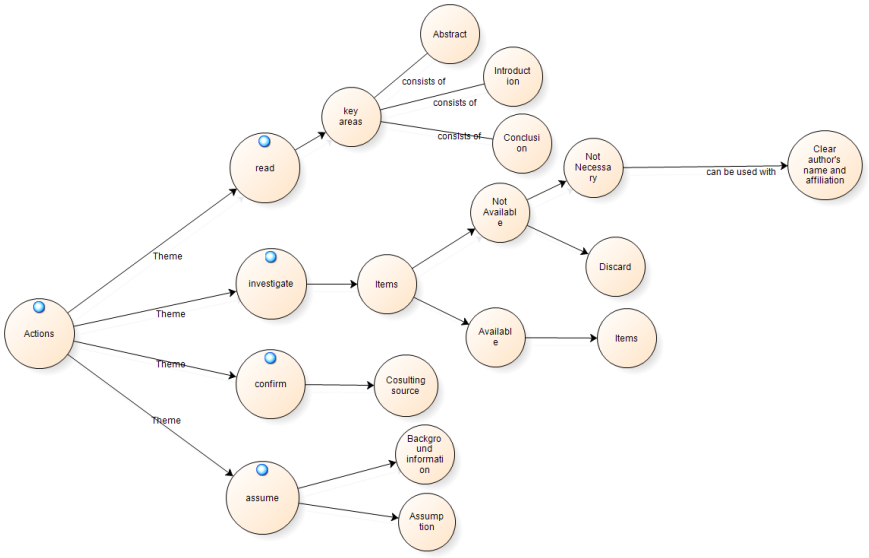


Figure 5.31: The initial main themes and key areas sub-theme

- **The type of content publication**: in this question, experts suggested actions in three themes; namely, investigate, read, and assume. The experts mentioned that citations, other publications on the same website, links, and the publisher could be used to support the evaluation process when the type of information required is not provided. Expert005 mentioned, "If it isn't clear from this searching the internet for how others have cited the document should provide the information", expert010 said "Try to establish this from the context and from other publications on the same website", and expert004 suggested "Look up the conference or journal of where the item was published for details on type of publication". The process of investigation to increase the user's confidence in the Web information when the content's publication type is not provided can be seen in Figure 5.32. Although,

the response texts from experts mentioned the "assume" and "read" themes, we could not find any new sub-themes. The sub-themes that we discussed in previous sections can still be used to explain the process.

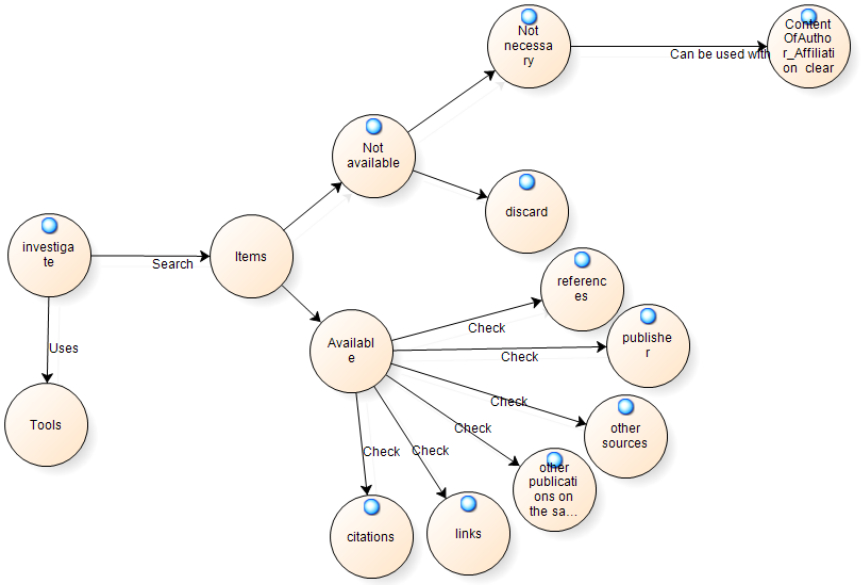


Figure 5.32: The investigation process to increase the user's confidence in the Web information when the publication type of the content is missing

Concluding this process, we obtained the collection of candidate themes and sub-themes as shown in Figure 5.33. Then, we refined our themes; this is discussed in the next section.

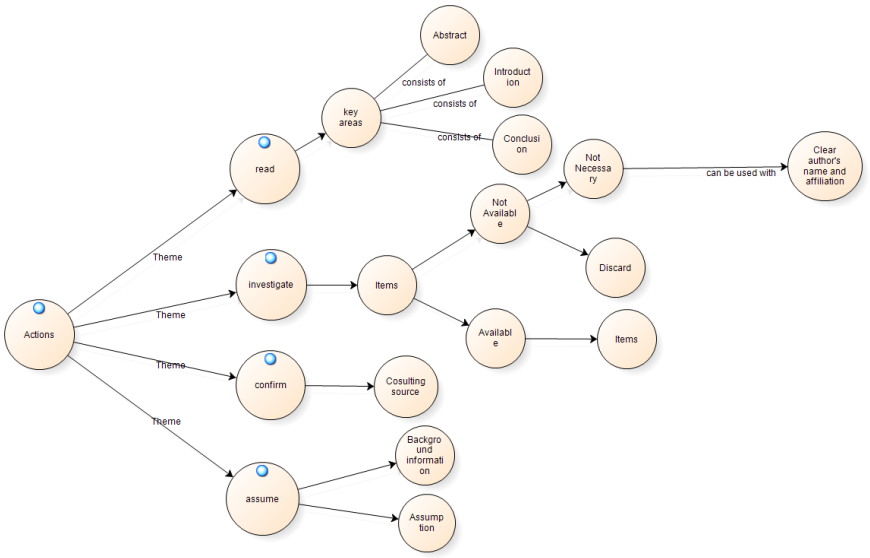


Figure 5.33: The initial main themes and sub-themes for the evaluation process when supportive items are missing

5.2.2.3.5 Reviewing the themes

In this phase, we collated the derived themes from the previous process with the response texts from the experts manually. We found that the themes form a coherent pattern of processes in order to evaluate the trustworthiness of Web information when the suggested supportive information is not present.

Considering the 11 patterns of investigation processes from previous sections, we found that in general the pattern of investigation could be categorised into two main methods, one of which was finding the missing information using other supportive information provided. The key item used to find additional supportive information was the author's name (except for the case in which the author's name is itself missing). The other method was to use the provided supportive information itself to assess the trustworthiness of Web information. The main supportive information from the experts' recommendations was the author's homepage, which might provide links to the organisation or research group's Web page. This information could be obtained by a Web address (URL). In addition, the publisher, the type of information, the references in the information, and page information could be used to help to evaluate the trustworthiness of Web information when some supportive information was missing. The outcome of the investigation into the trustworthiness of the information could lead to one of three possible responses.

- First to accept the information which was being considered.
- Second to discard the information because it lack trustworthy supportive information. In particular, if the author's name or the title of information was missing, the outcome of the process was more likely to be to discard the information.
- Third to ignore the missing information because it did not affect the trustworthiness of the information. Thus, the information could be used so long as the author's name and affiliation were stated clearly.

Therefore, we refined the "investigate" theme as shown in Figure 5.34.

The pattern of the "read" theme assessed the trustworthiness of Web information based on the supportive information. However, it focused on reading through the supportive information itself in order to estimate the trustworthiness of Web information. Alternatively, a pattern that emerged to increase the confidence of the trustworthiness of Web information was to consult the source in order to confirm the information on the page. In addition, making assumptions also emerged as a pattern to evaluate the trustworthiness of Web information. Nevertheless, the "assume" and "confirm" themes themselves could be considered as an investigate theme because they were based on the assumption that the background information was checked and confirmed by finding more information from the source. Therefore, we merged these two themes into the "investigate"

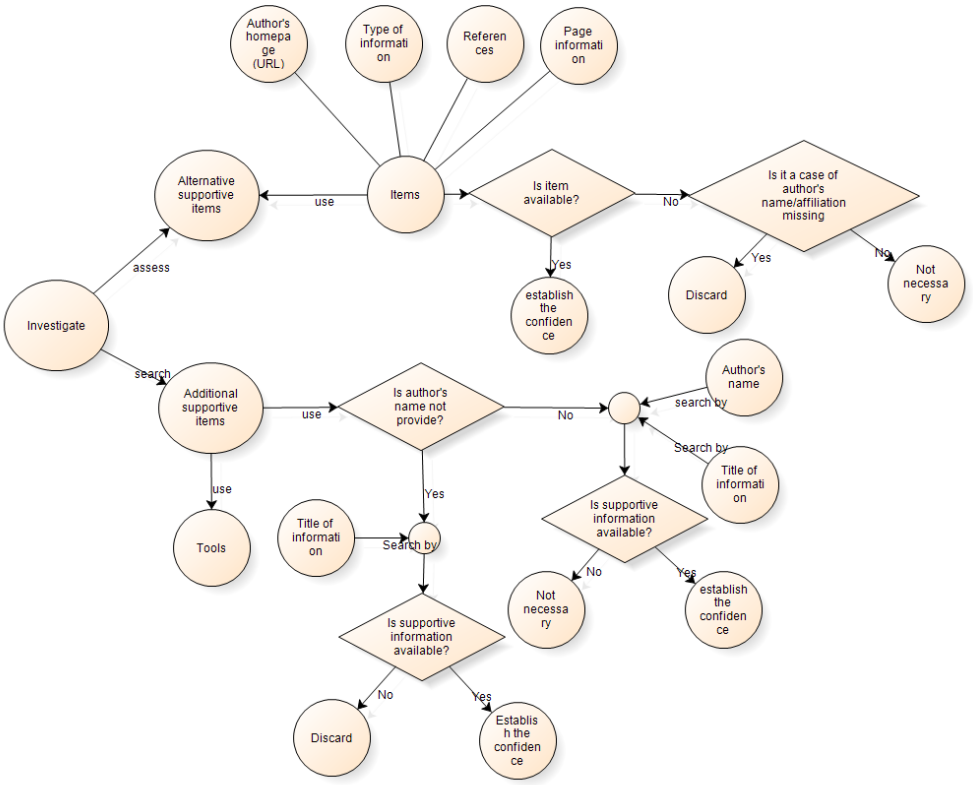


Figure 5.34: The refined "investigate" theme

theme. As a result, we refined our themes and merged them into the evaluation process as shown in Figure 5.35.

5.2.2.4 Analysis of the Results from Questions 1 and 2 in Section 4 of the Questionnaire

In Section 4 of the questionnaire, we aimed to explore any additional elements that should be considered in order to help users to improve their process of determining the trustworthiness of Web information. In addition, we were interested in the process of assessing the relevance of information to the experts' needs. We used an approach which was explained in section 5.2.2.1.2 and the prepared data from the section 5.2.2.2 to analyse the suggestions from the experts. The detail of each analysis phase based on thematic analysis is described as below:

5.2.2.4.1 Other features of a Web document apart from the proposed supportive information listed in previous sections that affect trust in information

This question was designed to allow the experts to raise any additional factors that might affect the trustworthiness of Web information apart from the proposed items. We

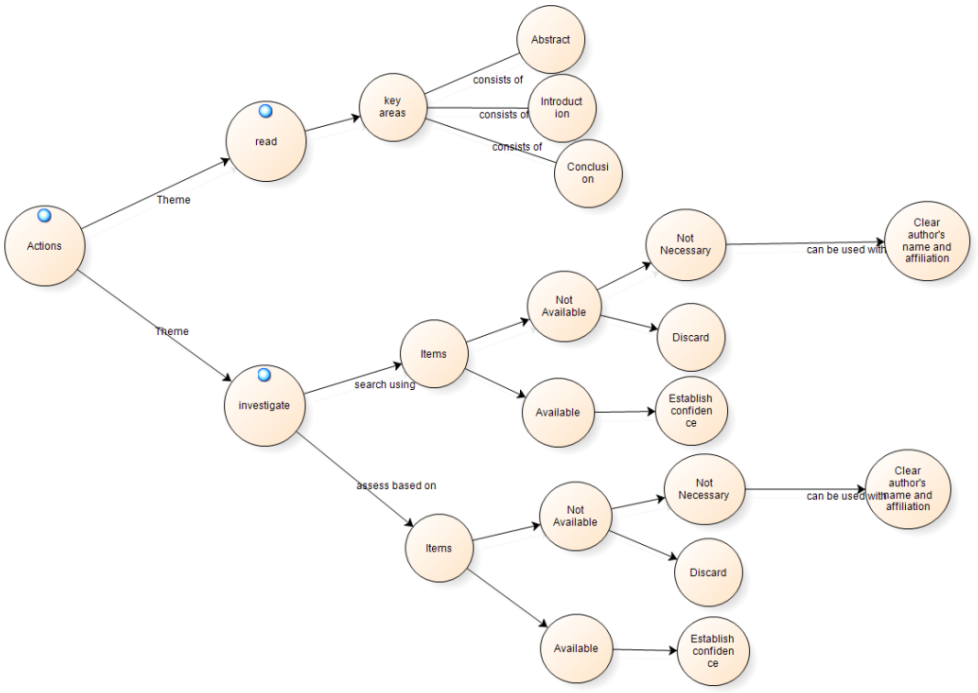


Figure 5.35: The refined main- and sub-themes

explain the details of our analysis of this question in the following section.

- **Generating initial codes**: we developed a coding scheme following the same step as discussed in section 5.2.2.3.1. The result of the word frequency query is displayed in Table 5.18.

Table 5.18: Word frequency table of question 1 in Section 4

Word	Length	Count	Similar Words
document	8	10	document, documents
quality	7	5	quality
trust	5	5	trust
mathematical	12	3	mathematical, mathematics
recommendation	14	3	recommendation, recommendations, recommended
references	10	3	references
support	7	3	support, supporting
website	7	3	website, websites
article	7	2	article, articles
author	6	2	author
claiming	8	2	claiming, claims
content	7	2	content
information	11	2	information

Table 5.18: Word frequency table of question 1 in Section 4

Word	Length	Count	Similar Words
journal	7	2	journal
location	8	2	location
presented	9	2	presented, presents
proof	5	2	proof, proofs
published	9	2	published
readability	11	2	readability
results	7	2	results
trustworthy	11	2	trustworthy
academic	8	1	academic
advertise	9	1	advertise
affect	6	1	affect
affiliation	11	1	affiliation
agreement	9	1	agreement
amount	6	1	amount
analysed	8	1	analysed
apart	5	1	apart
applicable	10	1	applicable
arguments	9	1	arguments
balanced	8	1	balanced
besides	7	1	besides
books	5	1	books
clarity	7	1	clarity
colleagues	10	1	colleagues
conference	10	1	conference
course	6	1	course
critical	8	1	critical
demonstrations	14	1	demonstrations
depends	7	1	depends
depth	5	1	depth
discussions	11	1	discussions
domain	6	1	domain
either	6	1	either
encountered	11	1	encountered
entire	6	1	entire
evidence	8	1	evidence
example	7	1	example

By studying the list of word frequencies in Table 5.18, we considered these words

against the response texts again. We created the primary coded nodes from the word frequency list by selecting the words that referred to a noun which reflected elements that might affect the trustworthiness of Web information in addition to the proposed elements. We excluded some words that were not relevant to the elements that might affect the trustworthiness of Web information. For example, the word "document" appeared at the top of the word frequency list, but it was mentioned within the corresponding questions. Thus, its appearance in the results is determined irrelevant for this study. Consequently, we obtained a list of coded nodes as shown in Table 5.19.

Table 5.19: A list of potential coded nodes from experts' responses

Word	Length	Count	Similar Words
quality	7	5	quality
mathematical	12	3	mathematical, mathematics
recommendation	14	3	recommendation, recommendations, recommended
references	10	3	references
website	7	3	website, websites
author	6	2	author
content	7	2	content
location	8	2	location
proof	5	2	proof, proofs
readability	11	2	readability
results	7	2	results
affiliation	11	1	affiliation
agreement	9	1	agreement
amount	6	1	amount
arguments	9	1	arguments
balanced	8	1	balanced
clarity	7	1	clarity
colleagues	10	1	colleagues
demonstrations	14	1	demonstrations
discussions	11	1	discussions
domain	6	1	domain
evidence	8	1	evidence
example	7	1	example

We considered coded nodes from Table 5.19 within the response texts and generated coded nodes. The list of frequently mentioned coded nodes allows us to identify the potential coded nodes that could refer to the other features that lead

to trust in information. In addition, we also considered the response texts themselves to discover others ideas mentioned by the experts which might not frequently stated in the response texts but which were still an interesting issue. However, some of the coded nodes which were frequently mentioned refer to the supportive information proposed in the previous question, such as author's credentials and author's affiliation. Therefore, we excluded these coded nodes from the potential coded nodes. As a result, we obtained a list of coded nodes and the number of times each coded nodes was referenced, as shown in Table 5.20.

Table 5.20: A list of initial coded nodes of question 1 in Section 4 of the questionnaire

Coded Nodes	Sources	References
recommendation	1	3
publisher	1	2
quality	1	2
references	1	2
website	1	2
arguments	1	1
clarity	1	1
demonstrations	1	1
discussions	1	1
evidence	1	1
mathematical	1	1
methodology	1	1
readability	1	1
results	1	1
style and tone	1	1

- **Structure code scheme**: from the list of initial coded nodes shown in Table 5.20, we investigated in detail to discover any relationships between the coded nodes by using Matrix coding. The results from the Matrix coding query showed that the coded nodes "quality" was surrounded by coded nodes "methodology" and "results". Therefore, we explored the response texts in more detail and found that when experts mentioned quality they gave a specific idea of what qualities the information should have. For example, expert005 suggested,

  > Quality and readability of results and the depth to which the results are explained and analysed also affect the amount I trust a document. Moreover the quality readability and reasonableness of the methodology used in the paper is a large factor in how trustworthy I find the document.

Expert002 recommended, "Quality of links and references - are they high quality information (books, journal, articles, government publications, etc.) recent and supporting the arguments in the web document." As a result, we grouped the coded nodes "results", "references", and "methodology" into a sub-node of the "quality" coded node. Similarly, the coded node "evidence" was coded surrounding by the coded node "mathematical". These coded nodes inferred that the quality of the mathematics (including proofs) could be used to support the trustworthiness of Web information. For example, expert005 suggested,

> Its quality and clarity of mathematics. Documents with mathematical proofs I find to be much more trustworthy as they tend to be more robust (of course this depends on how mathematical the subject of the document is).

Consequently, we merged the "mathematical" coded node as a sub-node of "evidence" node. Completing this phase, we obtained a list of coded nodes as shown in Table 5.21.

Table 5.21: A list of coded nodes from question 1 in Section 4

Coded Nodes	Sources	References
quality	1	6
evidence	1	3
recommendation	1	3
publisher	1	2
website	1	2
arguments	1	1
demonstrations	1	1
discussions	1	1
readability	1	1
style and tone	1	1

Table 5.21 showed the coded nodes that represented additional features suggested by experts for helping to evaluate the trustworthiness of Web information (i.e. those features that can be used to evaluate the trustworthiness of Web information but were not proposed in our previous questions). The "quality" coded nodes consisted of sub-coded nodes; namely, references, methodology, and result, as shown in Table 5.22.

Table 5.22: A list of sub-coded in "quality" coded nodes

Coded Nodes	Sources	References
references	1	2
methodology	1	1

Table 5.22: A list of sub-coded in "quality" coded nodes

Coded Nodes	Sources	References
results	1	1

In addition, the "evidence" coded node had "mathematical" coded node as a sub-code node as displayed in Table 5.23.

Table 5.23: A sub-coded of "evidence" coded node

Coded Nodes	Sources	References
mathematical	1	2

Therefore, the feature that was the most recommended by experts for helping to assess the trustworthiness of Web information was quality, which included quality of references, methodology and results. The second feature was evidence that could support the content provided such as mathematical proof. The third one was recommendations from the colleagues and people you trust. Particularly, the experts also mentioned about the publisher and website that could indicate the trustworthiness of Web information. The remaining features related to the style and tone of the content.

As a result, we considered adding information about the publisher in terms of including the URL of the publisher in our proposed criteria. This allowed the user to trace to the source who distributed the information. In addition, we considered providing links to the content of the supportive information (i.e. PDF file) for users such that they could use the information to support their assessment.

5.2.2.4.2 The process of assessment the relevance of information with a user's need

The aim of this question was to discover the process of evaluating the relevance of information to an expert's needs. We then adopted this process into our framework to refine our framework. The process of analysis is discussed in the following section:

- **Generating the initial codes**: we developed a coding scheme by using the word frequency tool of the NVivo software. This followed the same process as was described in section 5.2.2.3.1. As a result we obtained a list of words that could be used to generate the potential coded nodes as shown in Table 5.24.

Table 5.24: A list of frequency words used in question 2 of Section 4

Word	Length	Count	Similar Words
abstract	8	9	abstract, abstracts

Table 5.24: A list of frequency words used in question 2 of Section 4

Word	Length	Count	Similar Words
title	5	6	title
content	7	4	content, contents
check	5	3	check
conclusions	11	2	conclusions
introduction	12	1	introduction
first paragraph	9	1	paragraph
reading	7	1	reading
references	10	1	references
search	6	1	search
skimming	8	1	skimming

The next step was to create the coded nodes based on the list in Table 5.24 and to explore the response texts in order to generate coded nodes. The initial coded nodes that used to analyse the process of evaluating the relevance of information and the experts' needs are shown in Table 5.25.

Table 5.25: A list of sub-coded in "quality" coded nodes

Coded Nodes	Sources	References
abstract	1	8
check	1	3
conclusion	1	3
content	1	4
first paragraph	1	1
introduction	1	1
keywords	1	1
match	1	1
read	1	9
references	1	1
search	1	1
skim	1	3
title	1	6

- **Structure codes scheme**: from Table 5.25, we found that some coded nodes referred to the action of evaluating the relevance of the information. Others were the elements that experts suggested to use for assessing the relevance of the information. Therefore, we categorised the coded nodes into two groups; namely, "actions" and "items". The details of elements in each group are shown in Table 5.26 and Table 5.27.

Table 5.26: Coded nodes and their frequencies that are categorised within the actions group

Coded Nodes	Sources	References
check	1	3
match	1	1
read	1	9
search	1	1
skim	1	3

Table 5.27: Coded nodes and their frequencies that are categorised within the items group

Coded Nodes	Sources	References
abstract	1	8
article	1	1
conclusion	1	3
content	1	4
first paragraph	1	1
introduction	1	1
keywords	1	1
references	1	1
title	1	6

- **Searching for themes**: the purpose of this question was to find out the patterns that the experts used to evaluate the relevance of information with their needs. This helped to refine the relevance criterion of our framework to evaluate the trustworthiness of information on the Web for naive users. We focused on the "action" coded nodes which indicated which process to use in order to evaluate the relevance of Web information. As a consequence, we investigated five coded nodes in the action group; namely, "check", "match", "read", "search", and "skim". We used a cluster analysis function in NVivo focused on coding similarity. The result is displayed in Figure 5.36.

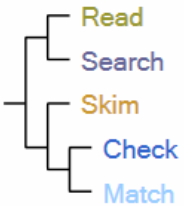


Figure 5.36: Coded nodes clustered by coding similarity in question 2 of Section 4

Figure 5.36 showed that the coded nodes "check" and "match" were similar and they were in the same cluster as the coded node "skim". In addition, the coded nodes "read" and "search" were similar. We then explored the response texts in more detail based upon this clustering. We discovered that the coded nodes "skim", "check", and "match" referred to the action of the expert having a quick look through the details of the content or the title of information: as expert010 suggested, "I check the title and skim over the content". Similarly, the coded nodes "read" and "search" were mentioned in terms of a process that considers the content or the details of the information more carefully. In this respect, expert004 recommended, "Initially by consulting an abstract or summary and then by consulting the document itself"; while expert002 suggested, "Read title, abstract and conclusions", and expert001 mentioned, "you would read the title and then (if you need more information to help you decide whether it's relevant) the abstract". However, the cluster "skim" was a similar process to that of assessing by reading because skimming through the content was essentially just a process of quick reading. Therefore, we combined these five actions into the same cluster. As a result, we obtained the initial main theme of the process of assessing the relevance of information to the experts' needs as the theme "read" as shown in Figure 5.37. After completing this phase, and having created the initial main themes, we ex-

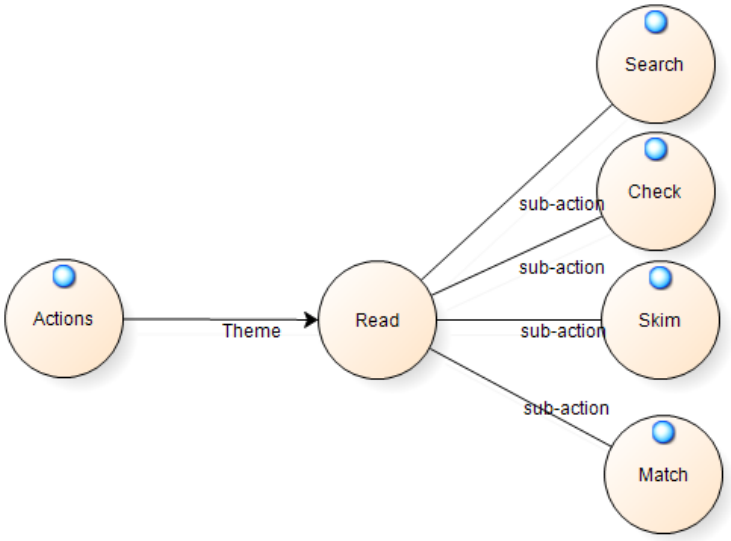


Figure 5.37: The initial themes of question 2 in Section 4

plored the response texts again to discover any relationships between coded nodes using the Matrix coding query. The result showed that the most mentioned elements which experts used for assessing the relevance of information to their needs was the abstract. For example, expert004 suggested, "Initially by consulting an abstract", expert003 said, "Reading the abstracts if available" and expert010 recommended, "When available I read the abstract". The abstract was a section that

explained the overall concept of the information. Therefore, it could indicate the relevance of the information to the needs of the user. In addition, the title was another important item that was used to assess the relevance of information. It was also the first spot that the experts looked for; as expert001 mentioned, "With journal articles you would read the title and then (if you need more information to help you decide whether it's relevant) the abstract", expert002 said "Read title, abstract and conclusions", expert005 mentioned "I'll read the title and the first paragraph and/or abstract", and expert006 suggested "Check the title and any abstract or introduction". Similarly, the content itself was the area that users can use to estimate the relevance between the information and their needs; as expert004 mentioned "Initially by consulting an abstract or summary and then by consulting the document itself". In summary, the experts suggested assessing the relevance by reading the data from the key areas of the article; namely the, title, abstract (first paragraph), and conclusion. These key areas were the main sections that could indicate the overview of the concepts discussed by the Web information. Interestingly, one expert (Expert007) mentioned the option to evaluate the relevance by matching the keywords (if available) with the user's needs and recommended, "Search key words & match their usage to my needs". As a result, we obtained the theme for assessing the relevance of information and the experts' needs as shown in Figure 5.38. Consequently, we adopted this process into the

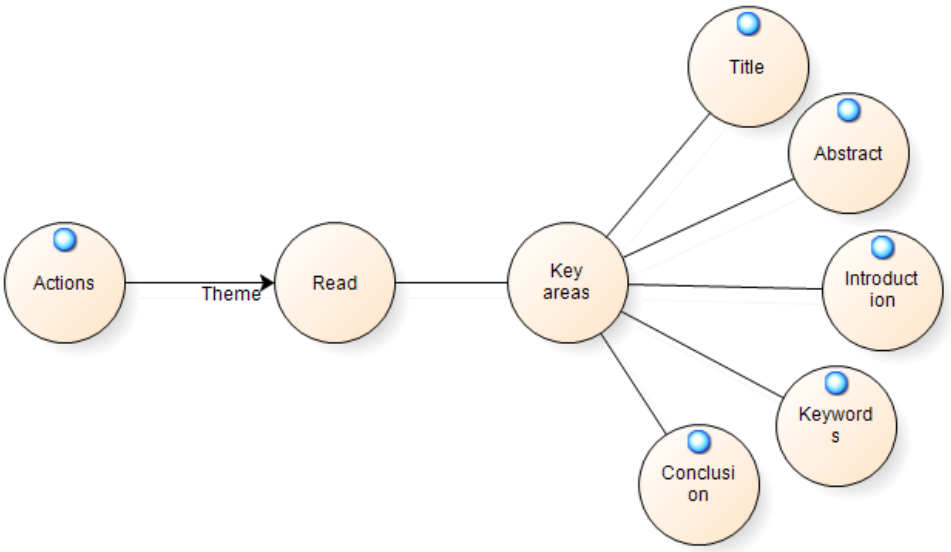


Figure 5.38: The theme of a process of evaluated the relevance

relevance criterion of our framework. We matched the search terms from the user of the framework with the key areas that we gathered from the other sources as the supportive information.

5.3 Conclusion

In this chapter, we discussed the process for validating the criteria used in the TWINE framework. We undertook a survey to ask five academic researchers and five members of the university's library team about the helpfulness of supportive information provided on the Web page and the resultant effect when some key, or supportive information was missing. In addition, we encouraged them to suggest other features that we should consider using in order to assess the trustworthiness of Web information. Moreover, we asked the experts about the process of evaluating the relevance of information to their needs. We then analysed their answers by using both quantitative and qualitative analysis approaches.

Based on this case study, the quantitative analysis results suggested that ten elements (namely, the author's name, the author's affiliation, the author's position, the publication medium, the title or abstract, the publication date of content, the last modification date of content, the information of editorial process, a list of references, and the number of times that information has been cited) were useful for helping to evaluate the trustworthiness of information. All of these elements helped support judgments about the trustworthiness criteria of the TWINE framework; namely, authority, accuracy, recency, and relevance. The qualitative analysis results suggested that users need to search for other supportive information or assess the trustworthiness of Web information based on other supportive information when some of these elements were missing. This suggestion confirms our idea of providing the metadata could affect the decision of users whether or not to trust Web information. In addition, the key areas such as the title, abstract, and conclusion were important to help in assessing the relevance of information to the users.

Consequently, we refined our framework based on the results from the survey. The next chapter demonstrates an application of the TWINE framework through the development of *the trustworthiness of Web information evaluation prototype* which is implemented based on the proposed and refined framework.

Chapter 6

Development of the Prototype

In this chapter, we explain the process of implementing a prototype based on the TWINE framework proposed in Chapter 3, called "Twine". Prototyping helps to illustrate how the framework can be adapted for use in a real-life scenario. As our case study, we chose to build a prototype that implemented the TWINE framework to search for academic publications. The prototype architecture is presented in section 6.1. Then, we address the development process of the prototype in section 6.2. In section 6.3, we discuss the usability test for the implemented prototype. Finally, we summarise the chapter in section 6.4.

6.1 The System Architecture of the Twine Prototype

Based on the functional architecture of our proposed framework in section 3.1, we design the system architecture of our prototype as displayed in Figure 6.1.

As can be seen, the prototype consists of four main functions; namely, input, generating an HTML page, metadata integration, and output functions. Moreover, as part of the metadata integration function, an integrated metadata graph is generated. In turn, this metadata graph is used to create JSON data to display the results using the output function. We discuss the prototype data model in section 6.1.1 and the details of each function in section 6.1.2.

6.1.1 Twine Prototype Data Model

We employ the named graphs data model (Carroll et al., 2005a) to represent our metadata as a metadata graph. Moreover, we use the Semantic Web Publishing Vocabulary (SWP) (Bizer, 2006) to express the basic provenance information of the gathered metadata. In this prototype, the provenance information states the authorising relationship

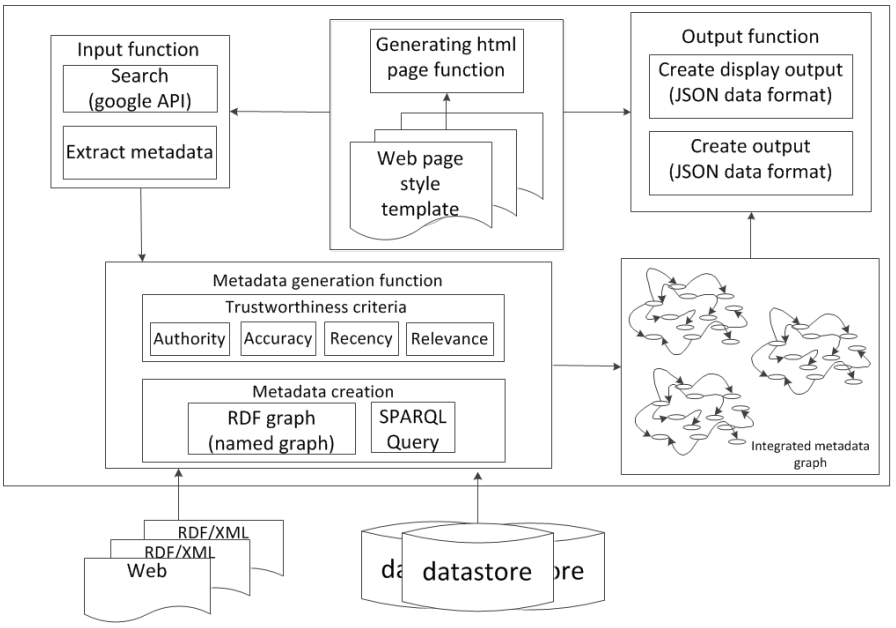


Figure 6.1: The systems architecture of the Twine prototype

between a named graph and an authority in the form of a warrant. The authorising relationship indicates a commitment between the authority and the content of the graph, and represents the properties by stating by whom the gathered metadata is asserted or quoted or the information's validity. We use the TriX syntax (Carroll and Stickler, 2007) to describe our named graph.

When the prototype gathers metadata from the RDF links which are returned from search results or when it queries additional metadata from RDF data stores, the prototype creates a new named graph for each search result. It extracts provenance information from the RDF documents from the search results' RDF links and attaches that information to the metadata graph. An example of a statement describing a metadata graph which is stored in the prototype is displayed in Figure 6.2. This metadata graph represents information about one of the search results from the University of Southampton's ePrints repository together with the recorded provenance information.

Figure 6.2 shows an example of a graph set which uses the SWP vocabulary for representing provenance information in our metadata graph about the authorising relationship. The graph in Figure 6.2 describes the graph which is named as "`http://eprints.soton.ac.uk/id/eprint/265992`"; the graph is asserted by a warrant with the authority "`http://id.ecs.soton.ac.uk/UoS/ECS`" (lines 7-16). The provenance information of each graph can help to determine the accuracy of the metadata itself.

Figure 6.3 illustrates a data model which is used in our prototype. A graph which is built from this data model is called the "metadata graph". Within the context of Twine, a metadata graph presents about the academic publications which are built in the form of graphs. A data model consists of three types of named graphs:

```
1. <TriX
2. xmlns:swp="http://www.w3.org/2004/03/trix/swp-1/"
3. xmlns="http://www.w3.org/2004/03/trix/trix-1/"
4. >
5.  <graph>
6.    <uri>http://eprints.soton.ac.uk/id/eprint/265992</uri>
7.              <triple>
8.     <uri>http://eprints.soton.ac.uk/id/eprint/265992</uri>
9.      <uri>http://www.w3.org/2004/03/trix/swp-1/authority</uri>
10.       <uri>http://id.ecs.soton.ac.uk/UoS/ECS</uri>
11.    </triple>
12.    <triple>
13.      <uri>http://eprints.soton.ac.uk/id/eprint/265992</uri>
14.      <uri>http://www.w3.org/2004/03/trix/swp-1/assertedBy</uri>
15.      <uri>http://eprints.soton.ac.uk/id/eprint/265992</uri>
16.    </triple>
17. </graph>
```

Figure 6.2: An example of a metadata graph in TriX syntax using the Semantic Web Publishing vocabulary for representing authorising relationships

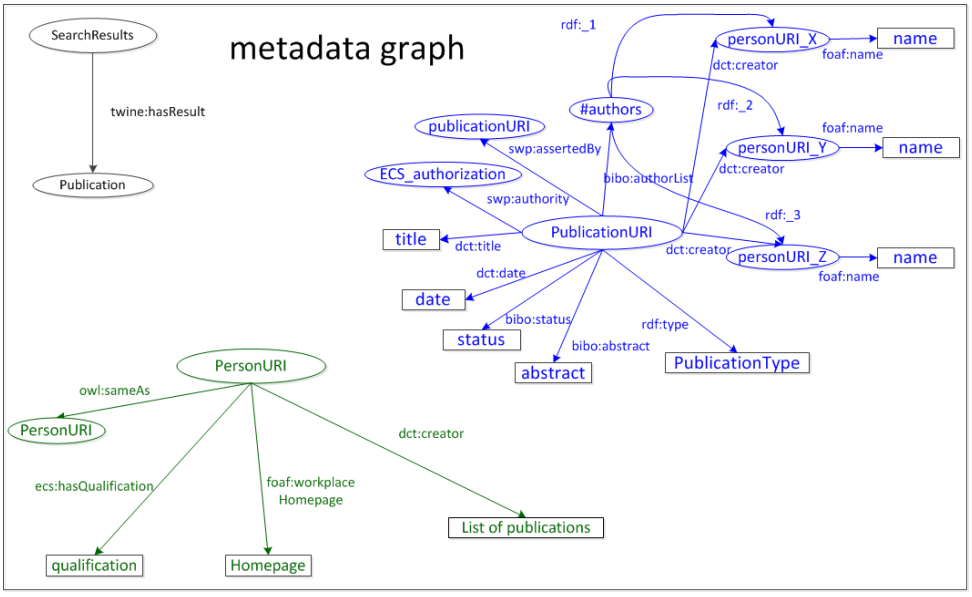


Figure 6.3: A data model of the Twine prototype

- Named graph of search results: This describes all the publications from the search results for which the user is interested in assessing the trustworthiness. It describes each search result by its publication URI.

- Named graph of publication metadata: This describes the information regarding the publications; namely, the authors of the publication (one author or multiple authors), the date that it was published, the title of the publication, the status of the publication, the type of publication, its abstract, and the provenance information of it authority.

- Named graph of metadata about the author: This represents the author's credentials and expertise such as their list of publications or projects, their qualifications or the URL of their homepage .

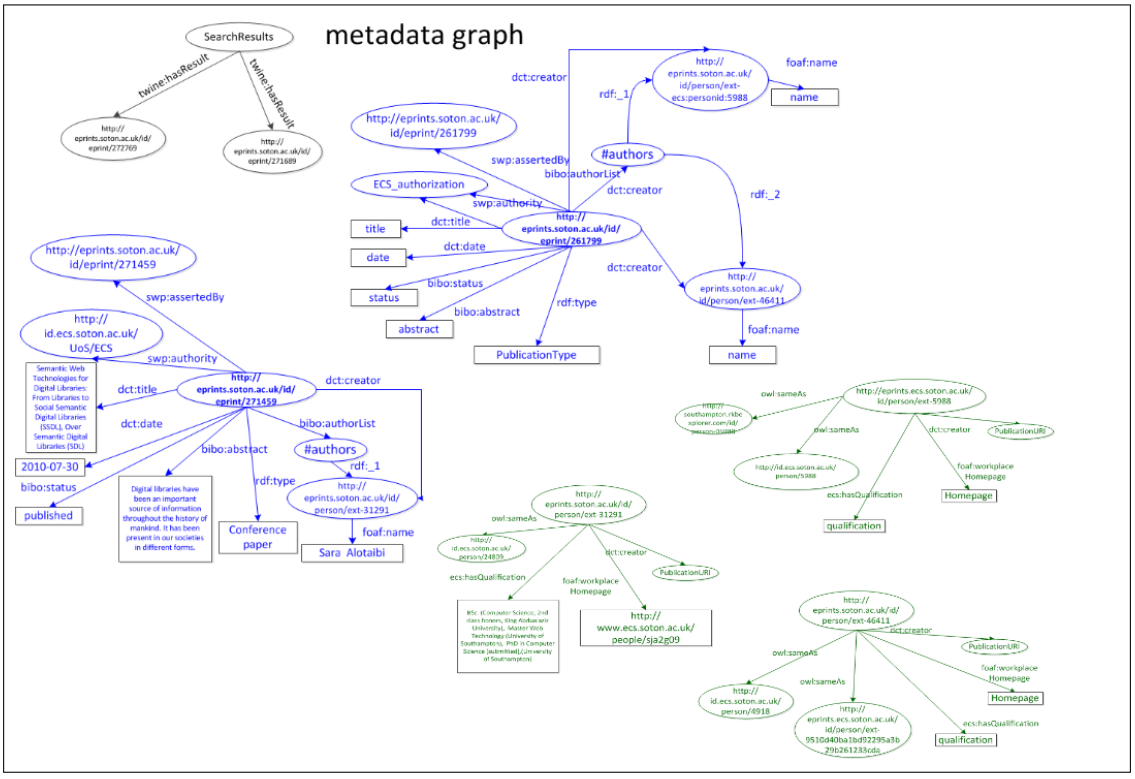


Figure 6.4: An example of a metadata graph

Figure 6.4 shows an example of a metadata graph which is constructed based on the data model shown in Figure 6.3. The metadata graph consists of a main graph, named "searchResults", from which two results are derived: publication URI "`http://eprints.soton.ac.uk/id/eprint/271459`" and "`http://eprints.soton.ac.uk/id/eprint/261799`". The metadata graph of publication "`http://eprints.soton.ac.uk/id/eprint/271459`" was created by one author "`http://eprints.soton.ac.uk/id/person/ext-31291`" whose name is "Sara Alotaibi". This publication was asserted by a warrant with the authority "`http://id.ecs.soton.ac.uk/UoS/ECS`". The title of the publication was "Semantic Web Technologies for Digital Libraries: From Libraries to Social Semantic Digital Libraries (SSDL), Over Semantic Digital Libraries (SDL)" which was created on 2010-07-30. Its status was published, and the type of publication was "conference paper". Moreover, the metadata graph described the detail of the abstract of the publication.

In addition, the named graph about the author (Sara Alotaibi) shows the details of her credentials. For example, the author with URI "`http://eprints.soton.ac.uk/id/person/ext-31291`" has a BSc. (Computer Science, 2nd class honours, King Abdualaziz University), and a Masters in Web Technology (University of Southampton), and has submitted for a PhD in Computer Science (University of Southampton). Furthermore, her workplace homepage is `http://www.ecs.soton.ac.uk/people/sja2g09`. In addition, the named graph displays a list of her publications by URI and the details of each publication.

The prototype can query the metadata based on URIs using a SPARQL query in order to obtain data from the graph itself or further information about the publications. For instance, the query shown in Figure 6.5 queries all of the results from the search results in order to retrieve data regarding publication "`http://eprints.soton.ac.uk/id/eprint/272769`". Specifically, the query attempts to obtain the title and the authors of the publication, and the qualifications of the authors. The results of this query are shown in Figure 6.6.

```
SELECT distinct ?pubid ?title ?author_name ?homepage
WHERE
{
GRAPH ?g1 {
        ?search Twine:hasResult ?pubid .
        }
GRAPH ?g2 {
        <http://eprints.soton.ac.uk/id/eprint/271459> dct:title ?title .
        <http://eprints.soton.ac.uk/id/eprint/271459> dct:creator ?authors .
        ?authors foaf:name ?author_name .
        }
GRAPH ?g3 {
        ?authors foaf:workplaceHomepage ?homepage .
        }
}
```

Figure 6.5: An example of a SPARQL query for retrieving metadata about publication "`http://eprints.soton.ac.uk/id/eprint/272769`"

```
http://eprints.soton.ac.uk/id/eprint/271459
Semantic Web Technologies for Digital Libraries: From Libraries to
    Social Semantic Digital Libraries (SSDL), Over Semantic Digital
    Libraries (SDL)
Sara Alotaibi
http://www.ecs.soton.ac.uk/people/sja2g09
```

Figure 6.6: A result of the metadata graph query of publication "`http://eprints.soton.ac.uk/id/eprint/272769`" specified in Figure 6.5

6.1.2 Twine Prototype Functions

As discussed above, the Twine prototype is composed of four functions: an input function, a metadata integration function, an HTML generation function, and an output function. We discuss each function in more detail in the following sections.

6.1.2.1 Input Function

The input function receives the search terms from the users and it calls the Google API to search for information based on the search terms. In this research, we focus on academic publications as a case study. In particular, we select the publications which

are stored on a research repository using the ePrints[1] repository system. We select ePrints because it is an open access[2] (self-archiving), Web-based repository for research literature and it also provides metadata about published research which can be used to create the supportive information in order to help Web users assess the trustworthiness of data they are consuming. In addition, ePrints is used worldwide for repositories of an extensive volume of online research literature such as UAL research online[3], University of Southampton EPrints[4], CaltechTHESIS (University of California)[5], eRA - Department of Employment, Economic Development & Innovation (DEEDI) (Australia)[6], and the Policy Documentation Center[7]. These repositories are implemented based on the ePrints repository system. The metadata on the system are generated with the same data model. Consequently, a system that can work with one ePrints application will be able to adapt to cooperate with other systems which follow the same basic pattern of the ePrints data model. ePrints is designed to store research literature with a well-defined metadata structure. Therefore, we can use this feature to our benefit when developing a system that can provide good additional information to support Web users' decisions on whether or not to trust Web information. However, in today's Web environment, it may difficult to retrieve good metadata from the WWW. In this case a technique called web scraping can be used to extract unstructured data on a Web page into structured data (metadata) (Scrapy Developers, 2008).

In our work, we select academic publications from the University of Southampton ePrints as our sample case. We refine our search arguments to search for publications from eprints.soton.ac.uk. Then, from the results returned by Google, the input function extracts the URL and URI of the search results for use in the metadata integration function.

6.1.2.2 HTML Page Generation Function

This function generates the html page which is used as the interface between the users and the prototype. It does this from a page layout template, which is written using

[1] ePrints are repositories of electronic copies of research literature (e.g. journal articles, book chapters, conference papers), scientific data, theses, reports, and multimedia. The details about the research publications are available online and, for some of these items, the full text can be accessed and used in accordance with copyright and end-user permissions (EPrints, 2000).

[2] Open Access "means immediate, permanent, free online access to the full text of all refereed research journal articles" (Harnad, 2005). The two most common ways to provide open access are self-archiving (green), where authors provide open access to their own published articles in their central institutional repository and journal-publishing (golden), where journals provide open access to their articles on the publisher's Website (Harnad, 2000).

[3] `http://ualresearchonline.arts.ac.uk/`

[4] `http://eprints.soton.ac.uk/`

[5] `http://thesis.library.caltech.edu/`

[6] `http://www2.dpi.qld.gov.au/extra/era/index.html`

[7] `http://pdc.ceu.hu/`

the Mako language. Mako[8] is a template library written in Python, which provides a non-XML syntax that can be compiled into Python modules.

6.1.2.3 Metadata Integration Function

Based on the indicators of the trustworthiness criteria, the metadata integration function starts with building the basic metadata graph which describes the basic information of publications such as the title, the date, and the types of publication. These metadata are gathered from the page itself based on the indicators of the four basic criteria in the trustworthiness criteria. Then, the metadata integration function retrieves further metadata of each publication by querying these additional metadata from the ePrints RDF data store using the publications' URIs and authors' URIs. The collected metadata are aggregated in order to build a metadata graph based on the data model discussed in section 6.1.1. This metadata graph is used in the output function to create the output to be displayed to the users.

6.1.2.4 Output Function

The output function uses the metadata graph from the metadata integration function to create the results to display to the users. We use a TriQL.P query (Bizer, 2004) to query the metadata from the metadata graph based on the authority criterion which explicitly indicates the quality of published data from the publisher. For example, we query the metadata graphs of all search results from Electronics and Computer Science, the University of Southampton.

```
SELECT ?publication
WHERE
{
 GRAPH ?graph1 {?result Twine:hasResult ?publication}
 GRAPH ?graph2
 { ?publication swp:assertedBy ?warrant.
  ?warrant swp:authority <http://id.ecs.soton.ac.uk/UoS/ECS>
 }
}
```

Figure 6.7: An example of a TriQLP query for querying the metadata graphs of all search results from the University of Southampton

Specifically, in this function, we use the weight factors and rating scores of the usefulness of each accepted indicator from the expert validation results in section 5.2.1.2.1 and section 5.2.1.3 to calculate the suggested trustworthiness score of each search result. Then, this score is used to rank the order in which the results are to be displayed on the result page.

[8]Mako Template for python: http://www.makotemplates.org/

The equation for calculating the suggested trustworthiness score of Web information is the score from the authority criterion and the sum of the product of the usefulness score of the indicators in three criteria (from the experts' rating score in section 5.2.1.2.1) with the weighting value of these suggested indicators. Therefore, the suggested trustworthiness score of the i^{th} result is given by

$$T_i = S_{A,i} + \sum_{d \in D} U_d \cdot W_d \cdot P_{i,d} \tag{6.1}$$

where

$$P_{i,d} = \begin{cases} 1, & \text{if } d \text{ is matched in result } i \\ 0, & \text{otherwise} \end{cases} \tag{6.2}$$

and where

- D is the set of indicators in accuracy, currency, and relevance criteria
- U_d is the usefulness score of indicators, d
- W_d is the weighting value of indicators, d.

The value of D, U_d, and W_d are defined in Table 6.1.

Table 6.1: The usefulness score and weighting value of the indicators in accuracy, recency and relevance criteria

Indicators, $d \in D$	Usefulness Score, (U_d)	Weighting value, (W_d)
Editorial process	3.10	0.17
Publication date	2.90	0.13
Publication medium	3.30	0.13
Content of the title or abstract	2.50	0.06
Number of citations	2.80	0.04
Last modification date	2.50	0.04

Further to the above equations, $S_{A,i}$ is the score representing the combined *authority* of the authors of the paper, i. In this research, we take importance of and broad impact of authors on the research area into account because the expertise and good reputation of authors in the community can indicate the quality of information they produced. We considered the expertise of authors by considering how often they are cited using the h-index[9]. However, an issue with the h-index is that it is unbounded. Therefore, we need to bound the effect of the h-index on the score in order to control the effect of

[9]The h-index "gives an estimate of the importance, significance, and broad impact of a scientist's cumulative research contributions."(Hirsch, 2005)

the h-index, which might dominate the score in the authority criterion. As a result, the score of author in the author list is computed based on the sum of the individual author scores multiplied by the usefulness score and weighting value of the indicators in authority criterion and the bounded h-index value of each author. The equation to calculate the authority score for the set of authors of paper i, A_i, is given by

$$S_{A,i} = \frac{1}{|A_i|} \sum_{a \in A_i} \left[1 - \left(\frac{1}{1 + h_a} \right) + \sum_{k \in K} U_k \cdot W_k \cdot P_{k,a} \right] \tag{6.3}$$

where

$$P_{k,a} = \begin{cases} 1, & \text{if indicator } k \text{ matches for author } a \\ 0, & \text{otherwise} \end{cases} \tag{6.4}$$

where h_a is h-index of author a, and A_i is the set of authors of result i, and where K, U_k and W_k are defined in Table 6.2.

Table 6.2: The usefulness score and weighting value of each indicator

Indicators, $k \in K$	Usefulness Score, (U_k)	Weighting value, (W_k)
Author's affiliation	2.90	0.28
Author's name	3.00	0.17

The suggested trustworthiness score and the other supportive information are stored and made available in JSON format, which then are interpreted and shown to the users in a way that is easy to understand. The results of the prototype are displayed in order of decreasing trustworthiness. Furthermore, the displayed results are shown as a combination of textual and visual elements such as bar charts and scales. We discuss the implementation process of the prototype in the next section.

6.2 Twine Prototype Implementation

The implementation of the prototype which employed the TWINE framework was developed based on a method for scenario-based usability engineering (Rosson and Carroll, 2002). We discuss the implementation process in more detail in the following sections.

6.2.1 Defining the Activity of the Prototype

We developed a preliminary list of the activities that our prototype must support.

6.2.1.1 Activity 1: Searching for Interesting Publications

In this situation, the main purpose of the activity was for the user to find the publications which relate to their topic of interest and for them to evaluate the trustworthiness of those publications easily. Providing the supportive metadata would help to support this assessment. Regarding the frequency of this activity's occurrence, we expected the users to do this often because the process of studying frequently requires the users to search for information. Users might have experience with similar search tools such as Google, Bing, or their own university Website. However, those tools might only generate limited metadata or not provide a useful way to solidly support the user's evaluation of the trustworthiness of the information the search engines return. By providing useful supportive metadata, the search engine would help to increase the user's confidence in the information they find.

6.2.1.2 Activity 2: Selecting the Publications

In this situation, users would like to select the publications that related to their work or interests. Also, the publications they selected should be the ones they feel confident about with regards to the information's validity and trustworthiness (for example because it is written by renowned authors). This scenario was a very common occurrence when searching for information. The general process was to read the title and abstract of the information, following which the users would make a decision about how much the information relates to their area of interest without necessarily ranking it in terms of relevance and trustworthiness. By having a ranking of the results that were most likely to be relevant and trustworthy, it was made easier for the user to evaluate and select information, which would in turn help to save time.

6.2.1.3 Hierarchical Activity Analysis

From the activities discussed above, we analysed and synthesised a sequence of activities as shown in shown in Figure 6.8.

Figure 6.8 illustrates the hierarchical activities of the prototype. The activities are divided into five levels:

1. Top level: This level is a start point of the activity. It starts with the fact that users would like to search for publications in which they are interested.
2. Generating level: Users generate the search terms, and then select the area of interest (in this case it is research papers). Then, they indicated a number of search results to display per page.

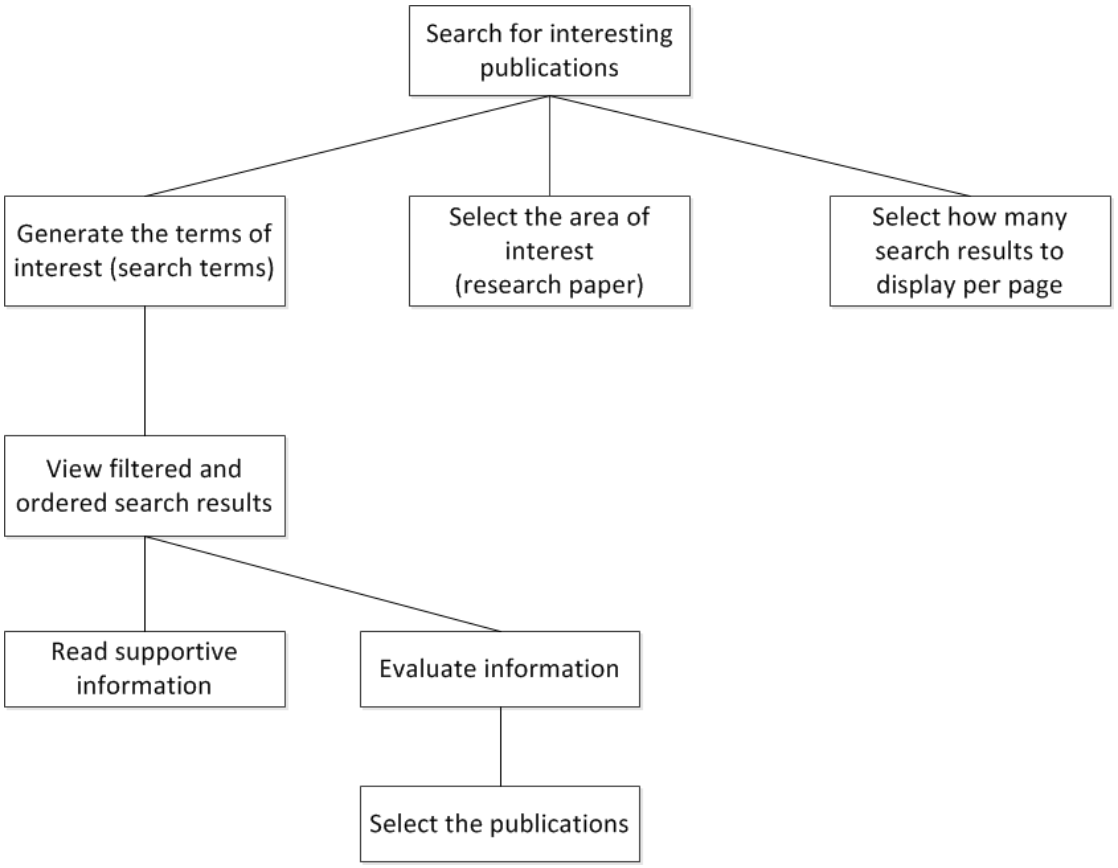


Figure 6.8: Hierarchical activity of prototype

3. Viewing level: Users reviewed all of the search results which have been filtered and ranked based on the suggested trustworthiness of information displayed by the prototype.

4. Assessment level: Users evaluate the trustworthiness of information based on the provided supportive information.

5. Decision making level: Users make a decision and select the publications to use in their work.

In the next section, we describe the interactions which occurs based upon the discussed activities.

6.2.2 Defining the Interactions of the Prototype

In this section, we developed a preliminary list of the interactions that our prototype must support. Based on activities in which the prototype should support users, we designed a site path diagram of our prototype that showed the range of interactions that the prototype must support, as shown in Figure 6.9.

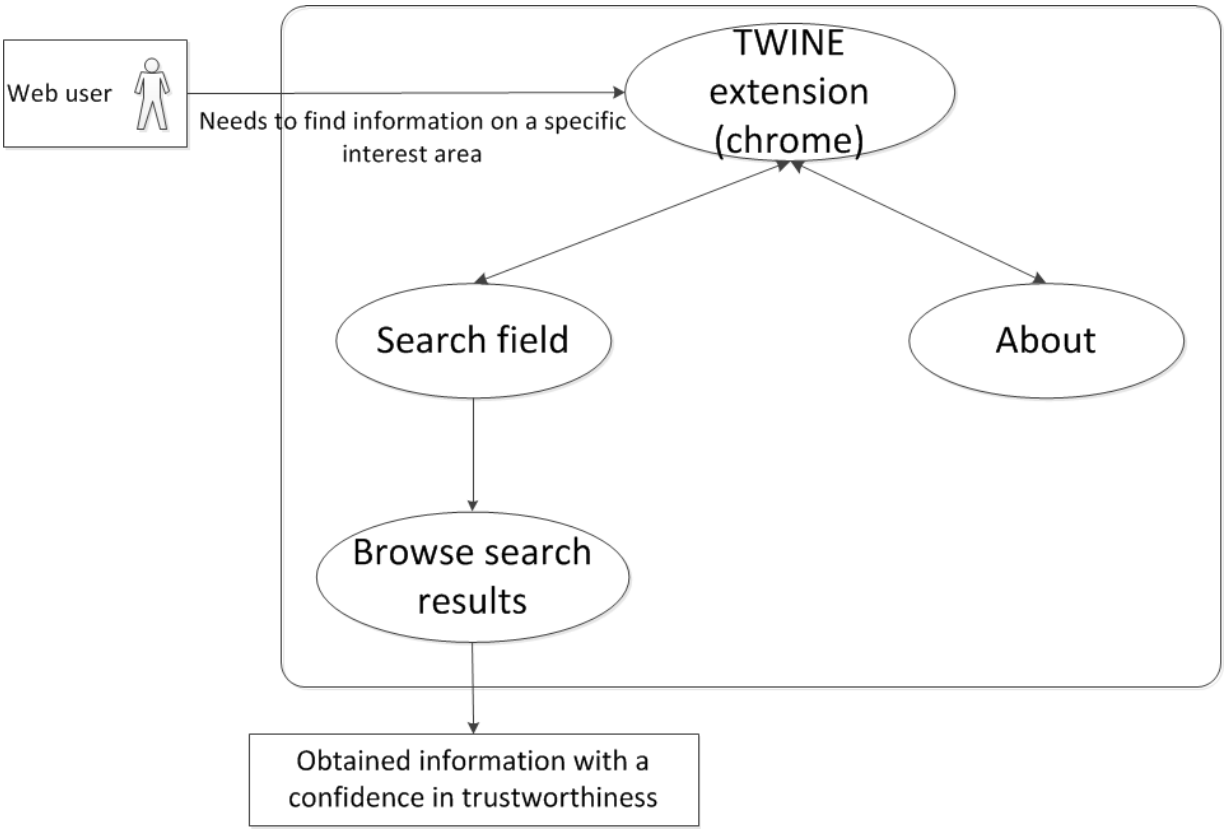


Figure 6.9: A site path of Twine prototype

6.2.2.1 Basic Interactions

In this section, we described interactions that would take place in the site path diagram shown in Figure 6.9 in detail, as discussed below:

- Finding information: The process, shown in Figure 6.10, describes steps taken when users would like to search for information that they are interested in.

  The flowchart in Figure 6.10 shows the different stages of interaction that takes place between the users and the prototype. We describe the details of the interaction as shown in Table 6.3.

Users' need	Users need to find information about some topics. Users have a rough idea of what they are interested in and in which area.
Users' attitude	Users do not have a specific definition regarding which information they want. Moreover, they have no idea how much they can trust it.

Interaction	User Action	Prototype Response
	(1) Click on Twine extension	
		(2)Go to Twine page
		(3)Show the main page of the system
	(4)Fill in the search terms	
	(5)Select the desired domain of interest from the list	
	(6)Select how many search results to show on each page	
		(7)Display the search results and supportive information
	(8)Review each search result	
	(9)Click the supportive information link	
		(10)Select to use a piece of information or move to the other results
	(11)Review supportive information	
	(12)Select to use a piece of information or move to the other results	

Table 6.3: A list of interactions for finding information

- Learning the prototype: The process describes steps taken when the user would like to find out more about the prototype that is displayed in Figure 6.11.

  The flowchart in Figure 6.11 is explained in greater detail in Table 6.4.

Users' need	Users need to learn about the system or how to use the system.

<table>
<tr><td>Users' attitude</td><td colspan="2">Users can find out how to use the system but they might like to be sure about some specific requirement. On the other hand, they might like to know what the system is about.</td></tr>
<tr><td rowspan="7">Interaction</td><td>User Action</td><td>Prototype Response</td></tr>
<tr><td>(1)Click on the about link</td><td></td></tr>
<tr><td></td><td>(2)Go to the explain page</td></tr>
<tr><td></td><td>(3)Show the details of the system and how to use it</td></tr>
<tr><td>(4)Read the information</td><td></td></tr>
<tr><td>(5)Click back to the main page</td><td></td></tr>
<tr><td></td><td>(6)Show the main page of system</td></tr>
</table>

Table 6.4: A list of interactions for learning about the prototype

6.2.2.2 Interaction Design Essentials

There are three essential issues for designing the system in order to ensure that the user can perform an efficient task (i.e. be able to operate the system faster); namely, current status, feedback (tell what has happened on the system), and the user's control (the system should give the users the feeling that they do a task in their own way) (Rosson and Carroll, 2002).

We now discuss how we could incorporate the essential factors regarding user interaction that we described above into the prototype:

- An indication of the current status: We use the status on the page title to indicate to the users where they were. In addition, each page provides a consistent navigation link. At the bottom of each page, there is a back button which links users back to the previous page or homepage. Also users can use a link from the navigation area to return to the homepage.
- An indication of the status of the interactions: The results page displays search results based on the selection from a user. The number of results per page are fixed and displayed on the page, as will the search terms and the interest domain. In addition, the current working page number is displayed in red in the paginator. In this way, feedback from the system will remind the user of what they are looking

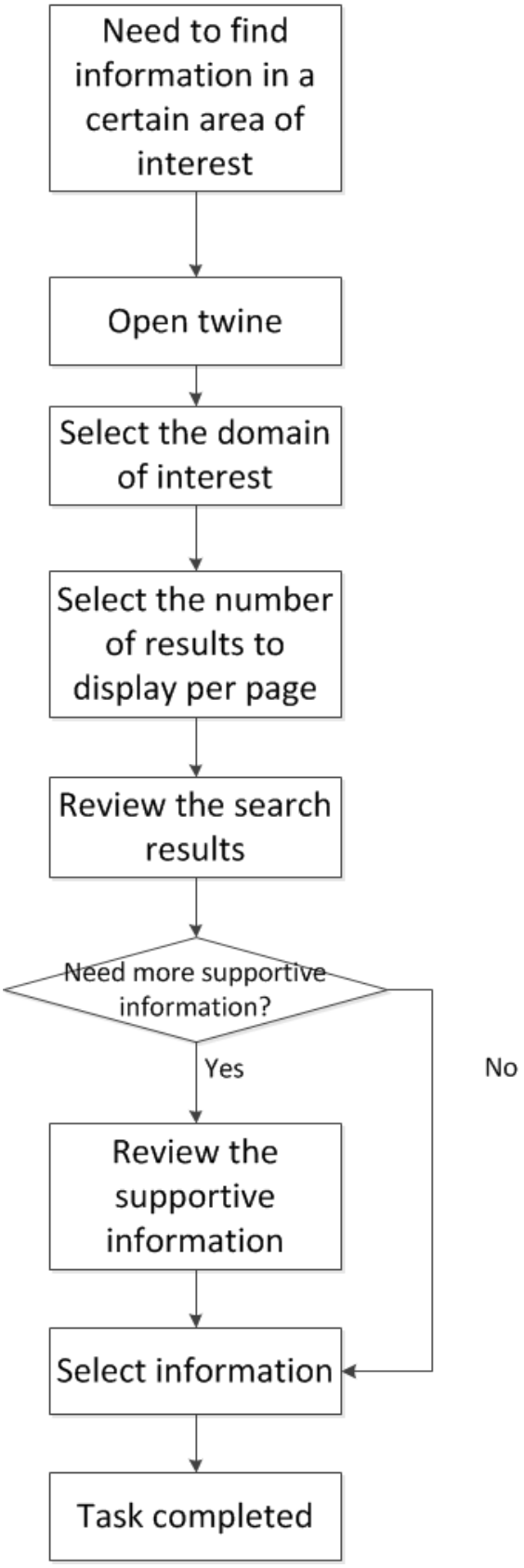


Figure 6.10: Searching for information flowchart

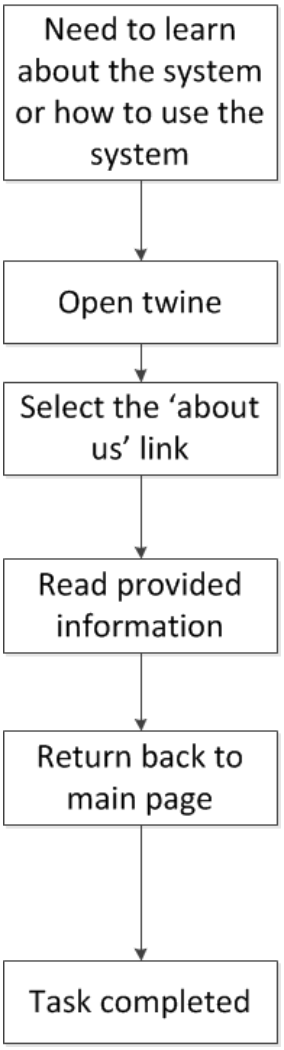


Figure 6.11: Learning about the prototype flowchart

for and which page they are on. A user can change the number of search results displayed per page by changing the number from the drop-down list and clicking the search button.

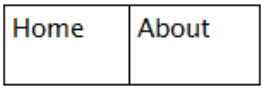
- Control of the interactions: Users are able to manage the input interaction of the prototype in three ways: specifying search terms, selecting the scope of interest, and selecting the number of search results per page. Then, the prototype generates results and display them on the page. The output interaction allows users to click on the links for viewing the supportive information. Consequently, the prototype provides a text field that allows users to enter search terms they want. In addition, the prototype provides options for users to choose the scope of interest and the number of search results per page. These options make the users feel in control of their query. As a result, they will feel more comfortable about using the prototype, and in return, can perform a task faster. However, the system also provides default

values of each option to show users how the system works and how they should interact with it.

6.2.2.3 Prototype Development

In this section, we developed a prototype based on the interactions that we designed in the previous section. In doing so, we converted an interaction into an interface. We designed our interface to be as simple as possible. We designed our prototype to have a text field that users can use to fill their search terms in order to retrieve the information they want. In addition to basic search results, the prototype included the necessary supportive information to help them evaluate the trustworthiness of Web information. We designed our prototype to have some options such that users can feel that they can control the system. Therefore, it would be good to have options for users to specify the area in which they are interested and how many pieces of information they can manage in one page. We explain the layout of the prototype in the next section:

- The navigation on the prototype: Our navigation consisted of two menu items; namely, the home and about items. The layout is displayed as below:



- A wire frame of the prototype: We designed a wire frame to display the sequence of Web pages that implement the interaction between the user and the prototype as discussed in section 6.2.2.1:

  1. Twine extension: The Twine prototype was implemented as a chrome extension. After installing the extension application, the Twine icon appeared on the page as shown in Figure 6.12 corresponding to step (1) in the user action column in the finding information use case in Table 6.3 :

Figure 6.12: A wire frame of the Twine extension icon

  2. Twine homepage: The Twine icon linked to the Twine page and then it displayed the main page of the prototype (steps 2-3 in the prototype response column in Table 6.3). The homepage consisted of the header, navigations,

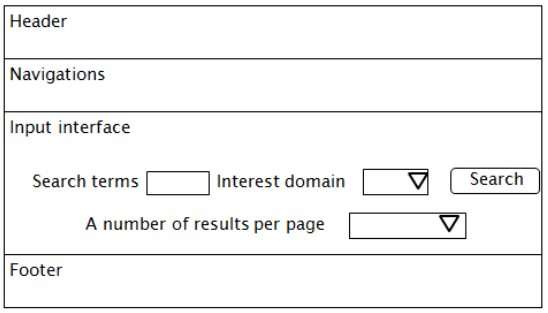


Figure 6.13: A wire frame of the Twine homepage

input interface, and footer. The input interface was designed to have a text field for accepting search terms. In addition, users could select the area of interest and a number of search results to display per page. This interaction corresponds to steps 4-6 in the user action column in Table 6.3.

3. Output display page: The prototype displayed the search results and supportive information to the users (step 7 in the prototype response column in Table 6.3) as shown in Figure 6.14.

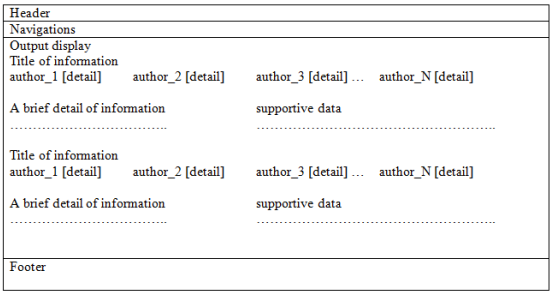


Figure 6.14: A wire frame of the Twine output display

When a user clicked on the supportive information link, the prototype displayed the details as shown in Figure 6.15.

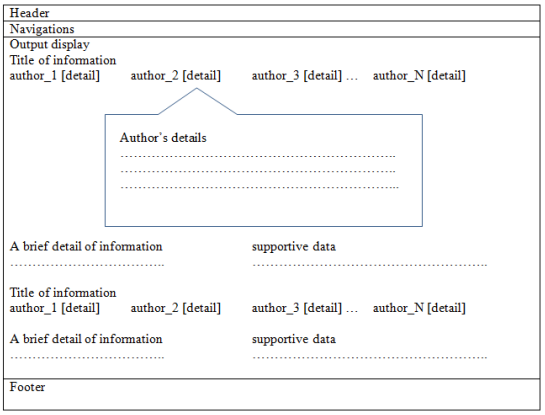


Figure 6.15: A wire frame of the Twine output display when the user clicks on the supportive information link

From the interaction design, we created a homepage screen as displayed in Figure 6.16.

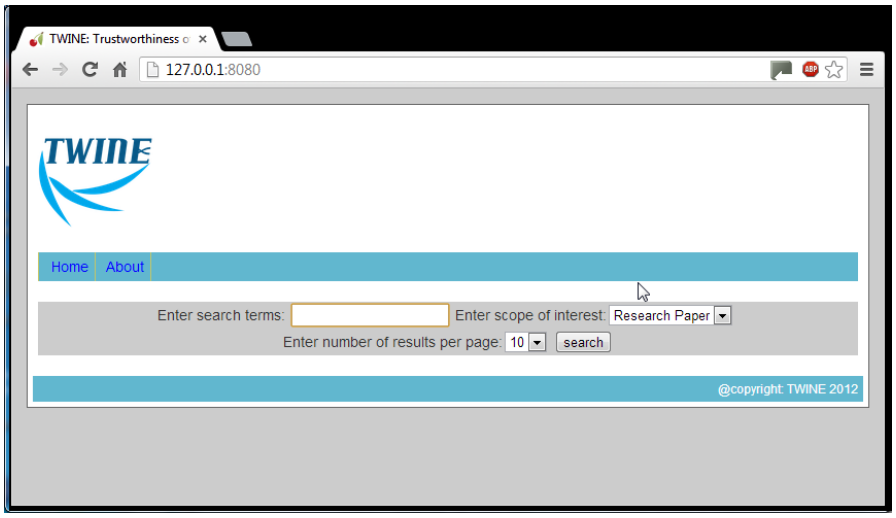


Figure 6.16: The design of the initial Twine interface page

6.3 Usability Test

In this section, we tested a prototype that we developed based on our proposed framework. The goal of the framework is to help Web users to evaluate the trustworthiness of Web information by providing critical supportive metadata about Web information.

6.3.1 Purpose

The purposes of our usability test were to ensure that the trustworthiness of Web information evaluation prototype, Twine, could provide a useful service to its users and for us to gain a better understanding of the factors that influence the usability of the prototype. Specifically, we would like to verify whether the user can find the information that they are interested in, and to assess whether they could select the information they want based on their evaluation of the trustworthiness of Web information. The results from the study were used to refine the prototype.

6.3.2 Test Plan

During the weeks of July 1 - August 1, 2013, we tested our prototype with five postgraduate students. We elected to use five students based on the suggestion of Nielsen and Landauer (1993) who posited that the best results of a usability test come from testing no more than five candidates; in which case we can identify the usability problems of

the design which occur most often, within optimal time and minimum consumption of resources.

The participants were postgraduate students ranging from 23 to 55 years of age who studied in the school of Electronics and Computer Science, University of Southampton. The participants were selected based on the state they were at in their studies. In more detail, the participants should be just starting their studies or they should have studied for no more than 2.5 years.

We administered an entrance question before each test, and asked the participants to sign a release form giving their permission for notes to be taken and used for data-gathering purposes. One facilitator led each session, which included one participant and a note taker (the facilitator). Users were asked to complete a task read aloud to them by the facilitator. In addition, the users were asked to think aloud while performing the task.

Our goal is to determine what is or is not working successfully on the Twine prototype from the users' perspective. We look for information such as -

- Do the users understand what the prototype is for?
- Do the users complete each task successfully?
- Is the provided supportive information useful?
- Are the users satisfied with the service prototype provided?
- Where do they stumble? What problems do they have? Where do they get confused?

After each session, we included an open-ended general discussion period where users could share their thoughts on any aspect of the prototype or testing with us.

We employed a task-based think-aloud protocol, in which we asked users to communicate their thought processes verbally while they perform the task. We asked them to vocalise what steps of work they undertake to complete a task, what questions they had while they interact with the prototype, and what surprised or confused them as they went through the prototype. After users finished their task, we asked open-ended neutral questions, such as "What do you think overall?" When users identify a problem, we asked them how they would fix it. We observed body language and facial expressions as well. These expressions might help to indicate the participant's opinion when they did not say it verbally. We noted these expressions along with the participant's explanation.

All users used a laptop (Lenovo Thinkpad X201i) on which the prototype is installed as a chrome extension.

6.3.3 Usability Test Design

As discussed previously, we used five postgraduate students as our test subjects to identify any usability problems. The consent form for this experiment is reproduced in Appendix C. In addition, we provided them with our prototype and a task for them to complete. The screen shot of our prototype is shown in Figures 6.17 and 6.18.

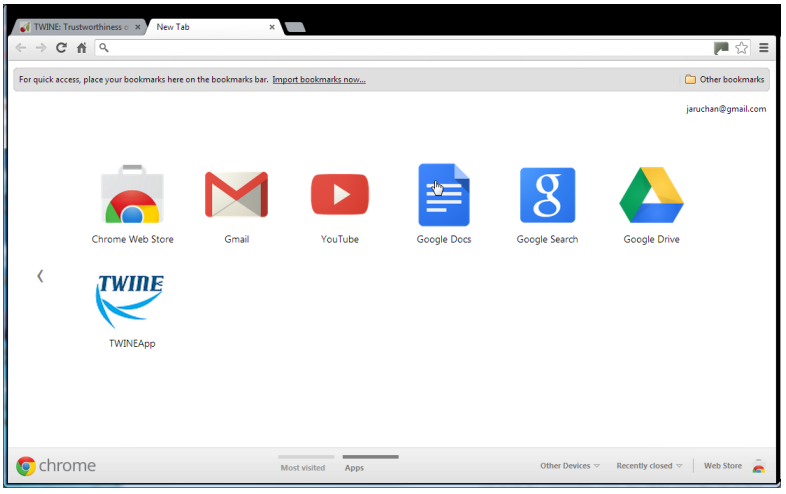


Figure 6.17: The Twine chrome extension

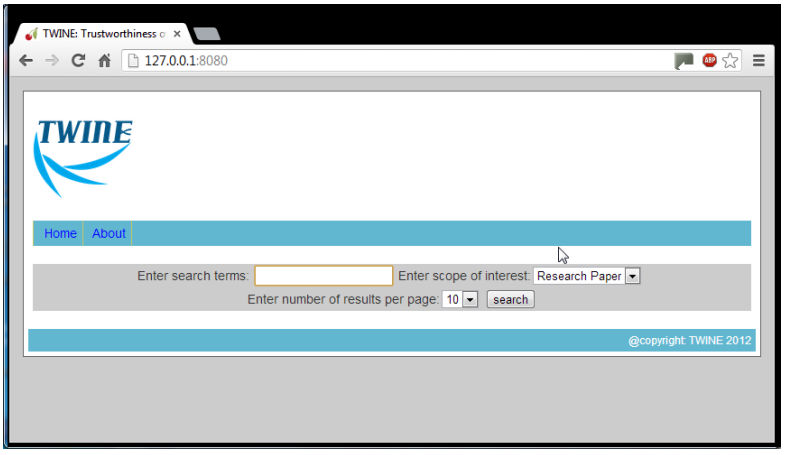


Figure 6.18: The Twine input interface

6.3.3.1 Reaction Test

For the first part of the Test, we tested whether the user understood the purpose of the system, how it works and how it is organised. By doing this, we asked the user to click on the Twine extension on the chrome browser. Then, when the browser displayed the Twine homepage, we asked the user to look at the page and we asked them questions and record their responses. Overall, the participants understood that the Twine prototype

is a search tool which could be used to search for information they want based on search terms. The participants' answers can be found in full in Appendix D.

6.3.3.2 Key Task Test

For the second part of the exercise, we gave the user a task to perform. By doing this, we developed a task scenario of the interaction described in the test plan in section 6.3.2. We set a scenario in which the participants were interested in research on the topic of 'privacy' and they needed to select papers in this area to reference in their report. Their task was to search for research papers in privacy topics and to select papers that were most likely to be trustworthy and relevant to their research interest. We asked them to use our prototype in order to complete this task.

Then, we read the task scenario and hand it to the users and asked them to perform the task. While performing the task, we asked the user to think aloud, and then we wrote down our observation in terms of the usability measures that we described in the test plan section. A summary of the usability problems which we observed is described in section 6.3.3.3. In addition, the overall opinions of the users' satisfaction are discussed in section 6.3.3.4. The participants' answers can be found in Appendix E.

6.3.3.3 The Usability Problems

After we finished from the usability test section, we listed down all of the main problems we observed during the test. A summary list of these usability problems is shown below:

- The logo of the prototype did not indicate its meaning to the users
- The home, about, and back buttons were not obvious enough
- The colour of the home and about buttons were too close to the colour of the navigation area
- The search terms input area was too narrow
- The icons of the author's details and author's name were not obvious enough
- The explanation of how to read the results was not detailed enough

6.3.3.4 Strengths and Weaknesses

- Strengths: Overall, users felt the Twine prototype was easy to use. The design and layout were clean and simple. They greatly appreciated the supportive information that was provided, which helped them to evaluate the trustworthiness of Web information.

- Weaknesses: Users provided feedback for improving the Twine prototype. The following items were not be included in this prototype, and this should definitely be rectified in future Twine versions.
  - A user suggested that the Twine prototype should have an option to sort the search results based on the user's preference or focus; for example by date of publishing.
  - A user suggested that if the Twine prototype highlighted the search terms in the abstract, it would help the users to spot how relevant the piece of information was to their interest.

6.3.3.5 Recommendations

From the usability problems described in the previous section, we made recommendations in response to the usability test results (e.g. How would we fix the problems that we observed? What could be changed to make the system more usable?).

6.3.3.5.1 High Priority Twine Recommendations

Items in this section could significantly improve the usability of the Twine which is shown in Table 6.5.

Table 6.5: Itemised high-priority recommendations

Items	Usability of the Twine	
	Problem	Suggested solution(s)
Logo of the Twine	Users mentioned that the logo is the first thing they noticed and it would be good to get an idea from it to say what Twine is.	Explain on the logo what Twine means or stands for.
Home and About Button	Users did not find the Home and About button to be obvious.	- Change the colour of the text "Home" and "About" to white. - Increase the font size of text "Home" and "About". - Make the "hit area" larger.
Back Button	Users reported that the Back button was not initially obvious.	-Increase the font size of text "Back" on the button - Make the Back button just go to the previous screen

Table 6.5: Itemised high-priority recommendations

Items	Usability of the Twine	
	Problem	Suggested solution(s)
The search terms layout	The users felt that the search terms area is too narrow. They felt like something was missing from the page.	- Increase the area of the search terms layout.
The publication author's name	Users reported that the author's name was not initially obvious.	- Make the font of the author's name bold. - Change the colour of the author's name to a bright colour
Author's details badge icon	Users reported that the author's details badge was not initially obvious.	- Increase the size of the icon. - Change to a more obvious colour (for example, blue).
Explanation of how to read the results	Some users were not sure how to interpret the results from the prototype. Sometimes, they were not clear what each element on the page was. They would like more explanations about how to read or interpret the displayed results.	- Expand the explanation on how to read the results, such as adding the template of the display results page then point out each element and then describes what is each element is.

6.3.3.5.2 Second Priority Twine Recommendations

These recommendations did not greatly enhance the usability of the Twine for this pilot. They can be considered if there is time, otherwise they should be reviewed prior to any future Twine projects.

- Change the shape of the navigation bar area to have more curves at the end of each corner.

In summary, we mainly focused on interface interaction usability as it is an important factor that allow users to perform their task successfully because if the users could not understand how to use the prototype, they might not be able to use the provided supportive information to help them assess the trustworthiness of Web information.

6.3.3.6 The Refined Twine Prototype

We employed the high- priority recommendations from the usability test results to refine our prototype. The new layout of the Twine prototype is displayed in Figures 6.19 and 6.20.

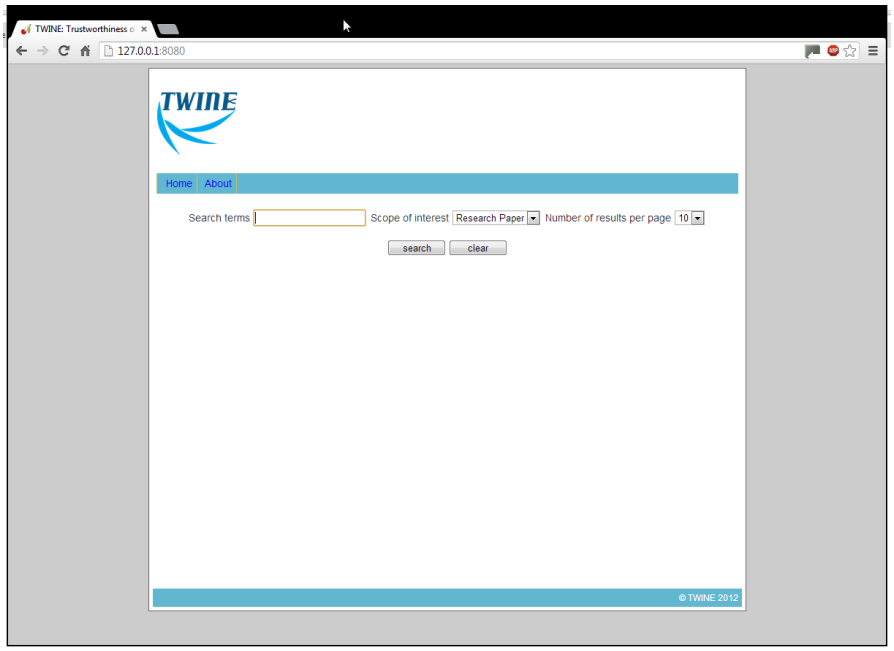


Figure 6.19: The refined Twine input interface

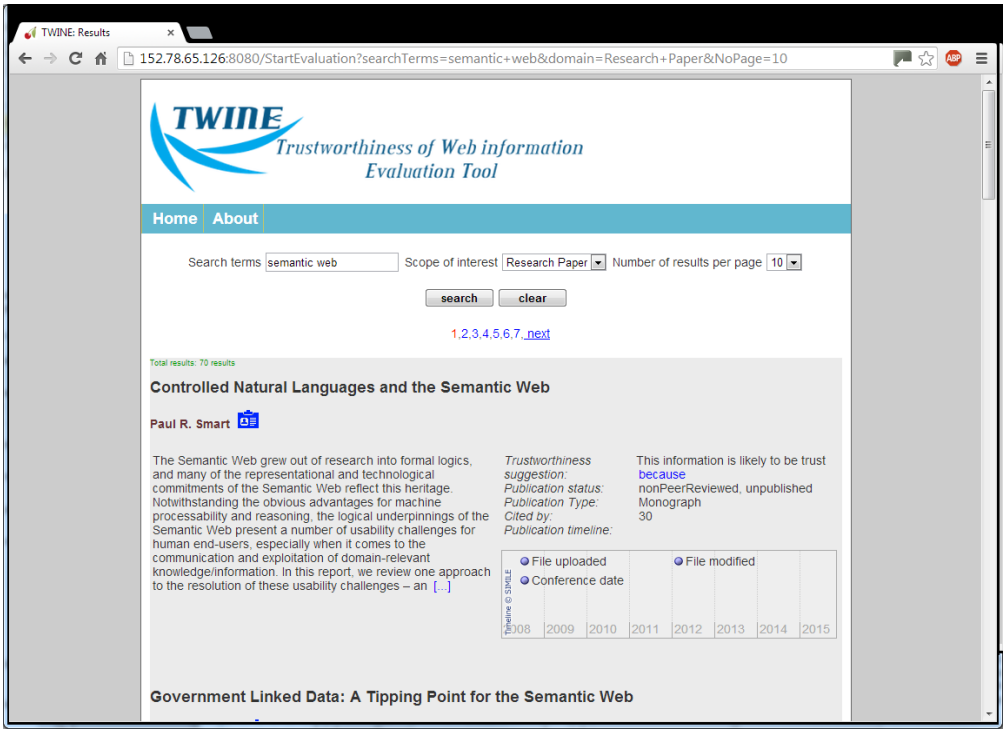


Figure 6.20: The refined Twine output interface

6.4 Conclusion

In this chapter, we discussed the implementation of a Twine prototype which is used to illustrate how we can adopt the proposed framework in practice. We presented the prototype architecture and the process of developing the prototype. In addition, we performed a usability test study to ensure that the proposed prototype can provide a useful service to the users. The results from the study showed that the users felt the Twine prototype was easy to employ. Moreover, they appreciated the provided supportive information which helped them to evaluate the trustworthiness of Web information. In the next chapter, we discuss a process of evaluation the TWINE framework based on the implemented prototype from this chapter.

Chapter 7

The Evaluation of the TWINE Framework based on the Prototype

In the previous chapter, a prototype was implemented based on the TWINE framework introduced in Chapter 3. This prototype allows us to evaluate the TWINE framework. A usability test was conducted to validate the prototype in terms of the ability of the prototype to provide an easy-to-understand method of use in order to provide useful, supportive information to the users. In this chapter, we discuss the evaluation process adopted to assess the TWINE framework. In section 7.1, we describe the design of a study used to evaluate the TWINE framework based on the prototype developed in the previous chapter. In section 7.2, we discuss the method through which the survey designed in section 7.1 is carried out. Then, in section 7.3, we analyse and discuss the results from the study. Finally, in section 7.4, we summarise the results of the survey and draw the chapter to a close with some conclusions from these results.

7.1 Study Design for Evaluating the TWINE Framework

In this section, we discuss the design of our study, which aims to evaluate our proposed framework. In order to evaluate the TWINE framework, we developed a prototype based upon it. We then evaluate the prototype to show that the framework can be used to implement a tool and also that the implemented tool helps users to increase their ability to evaluate the trustworthiness of Web information. We address the hypotheses of our research in section 7.1.1. Then, in section 7.1.2, we discuss our evaluation plans.

7.1.1 Research Hypotheses

The top-level hypothesis for our research is "A framework (such as TWINE) with properties of gathering, integrating and presenting supportive information using Semantic Web technologies helps users to more effectively evaluate the trustworthiness of Web information." More specifically, we use academic publications as a case study and we divide this hypothesis into the following sub-hypotheses.

Using our framework:

- The users increase their confidence in their judgment of the trustworthiness of the Web information that they find.
- The users increase the number of pieces of trustworthy information which they select to use.
- The users are satisfied with the supportive information insofar as it helps them to evaluate the trustworthiness of Web information.

7.1.2 Evaluation Plan

In order to evaluate our framework, we recruited a set of participants to take part in an evaluation study. The participants are postgraduate students who are just starting their studies, or postgraduate students who have been studying for no more than 2.5 years. We then assigned tasks to the participants for them to complete. The tasks were designed to evaluate the TWINE framework using two types of Twine tool environment:

- A control Twine prototype, called **Twine 1**, which only provides basic information about the search results such as the title of the publication, the authors' names, the abstract of the information, and the number of times that the publication has been referenced. These data are normally available to users when they use the search engine on the Web to find some information.
- An experimental Twine prototype, called **Twine 2**, which provides the basic information available in the control prototype, but also provides additional supportive information which is more detailed about the publications and authors; namely, the details of authors (e.g. position, workplace, qualifications, number of publications, projects, etc.), the editorial process, the status of the publications, the type of publication, the date of publishing, and the explanation for why this publication would be trustworthy for the users. The precise details regarding which supportive data are used, and the reasoning behind selecting said data, are discussed in section 4.3.

The tasks assigned to participants were designed to test our hypotheses. We asked the participants to search for information in the assigned topics using the two prototype environments described above. We constructed a counterbalance experiment to run our studies. The counterbalance experiment is the process of systematically varying the order of the experimental conditions in the conducted study (Field, 2009). That is, the participants receive their tasks in a different order and on different topics. This approach helps to remove systematic bias caused by practice effects or boredom effects. In addition, in our study we set the same participants to perform the tasks with both prototypes and different topics. This is to investigate the differences between users' opinions and decisions depending on whether they are given just the basic information or provided with additional information. We discuss the details of our study regarding our hypotheses in the following sections.

7.1.2.1 Sub-hypothesis 1

The first of our sub-hypotheses is, "When using the TWINE-based prototype, the users increase their confidence in their judgment of the trustworthiness of the Web information that they find." From this, we obtained the research question:

- Research Question 1: Does the supportive information provided by the TWINE-based prototype affect the users' confidence in their ability to evaluate the trustworthiness of Web information?
- Study 1: To investigate research question 1, we proposed the following study. We divided the experiment into four sessions, as illustrated in Figure 7.1. In addition, we divided the participants randomly into four groups. Each group of participants was randomly assigned to participate in each experimental session. Therefore, the study consists of four sessions of experiments and four different groups of participants. We discuss this in more detail in the upcoming paragraphs.
  - **Experiment session 1**: We assigned two search topics to the participants of group 1 for them to perform:
    - Firstly, we asked them to search for publications relating to the topic of "A" (where A is either "privacy" or "semantic web") using the control prototype. In addition, we asked the participants to set the number of results to display per page as ten, as this number is perceived as easier to scan and find information that users want at one time (suggested manageable amount of information that users can handle) (Bernard et al., 2002; Höchstötter and Lewandowski, 2009). Then, we asked them to rate the trustworthiness of the top ten search results and their confidence in their given score as a percentage for each of the top ten search results on the first page. After that, we asked them to wait for two minutes before

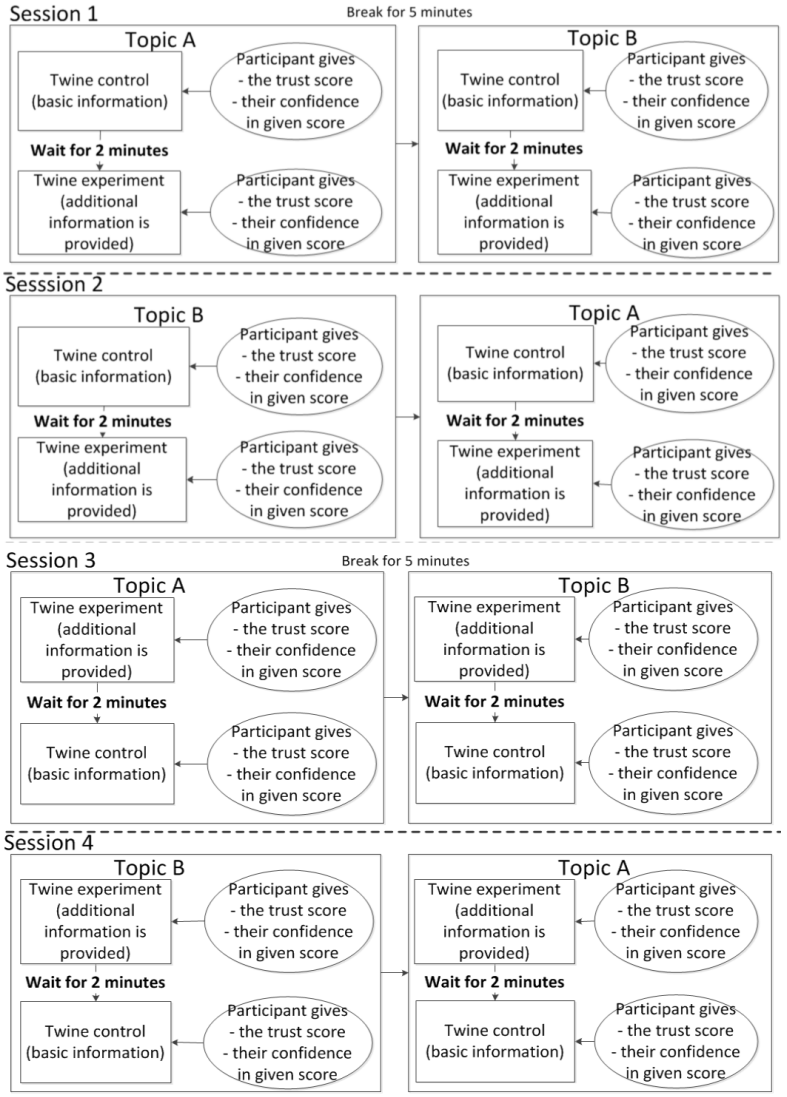


Figure 7.1: The design plan of study 1

starting the task again using the experiment prototype with the same topic.

* Secondly, after a five minute break, we asked the participants to search for publications relating to the topic of "B" (where B is the remaining item of the two identified in the previous paragraph) using the control prototype. Then, as before, they were asked to rate the trust score for each of the top ten search results on the first page. After that, the participants wait for two minutes and then start to search for publications on the same topic but this time they were asked to use the experiment prototype.

– **Experiment session 2**: In this session, we used the same procedure as in Experiment 1, except that the participants were from group 2, and we began with the participants searching for information on the topic of "B" using

the control prototype first, followed by the experiment prototype, and then repeating the process searching for information on topic "A".

- **Experiment session 3**: This follows broadly the same pattern as above. However, this time, participants from group 3 first search for information on topic "A" using the *experiment* prototype, followed by the control prototype; repeating this process after a five minute break for topic "B".
- **Experiment session 4**: This is performed as above, except with participants of group 4 and topics A and B transposed.

7.1.2.2 Sub-hypothesis 2

The second of our sub-hypotheses is, "When using the TWINE-based prototype, the users' confidence in their ability to assess the trustworthiness of information based on the supportive information provided increases and they will accept more pieces of information to use." From this, we obtained the research question:

- Research Question 2: Do the users increase the number of pieces of information they would select when they obtain supportive information?
- Study 2: To investigate research question 2, we extended study 1 by asking extra questions. Specifically, the participants must answer how many publications they would select to use in their work, which ones they would use, and why they decided to select them for use in their work from the top ten search results. The design of the study is displayed in Figure 7.2.

7.1.2.3 Sub-hypothesis 3

The third of our sub-hypotheses is, "The users are satisfied with the supportive information provided by the TWINE-based prototype insofar as it helps the users to evaluate the trustworthiness of Web information." From this, we obtained the research question:

- Research Question 3: Is the participant satisfied with the usefulness of the supportive information provided by the TWINE-based prototype?
- Study 3: To investigate research question 3, we proposed the following study. After the participants completed the tasks in study 1 and study 2, we asked them to rate their level of satisfaction of the supportive information provided to help them to assess the trustworthiness of Web information. The study plan is shown in Figure 7.3.

In the next section, we discuss the process of conducting the survey based on our evaluation plan.

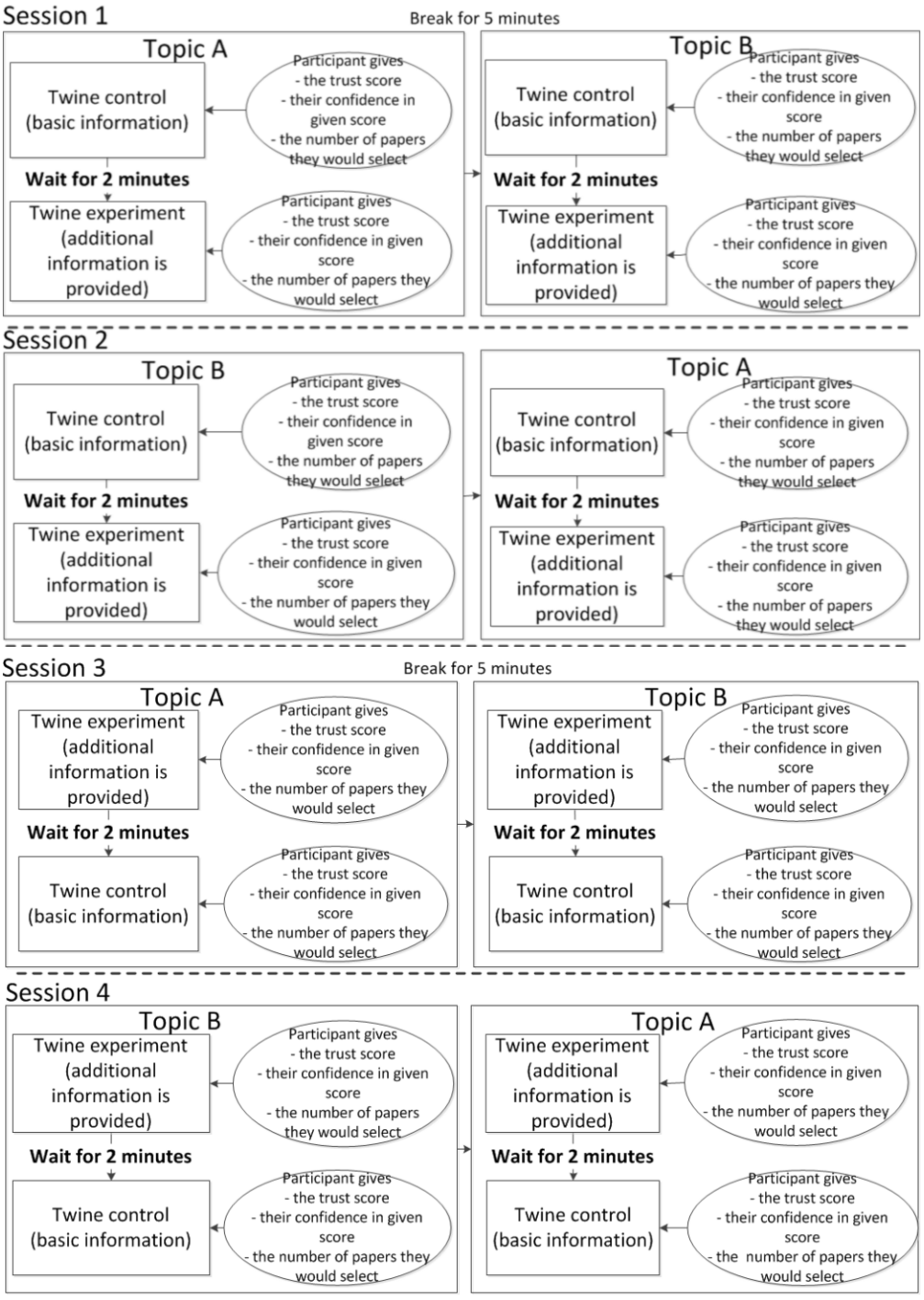


Figure 7.2: An extended design plan of study 1 used in study 2

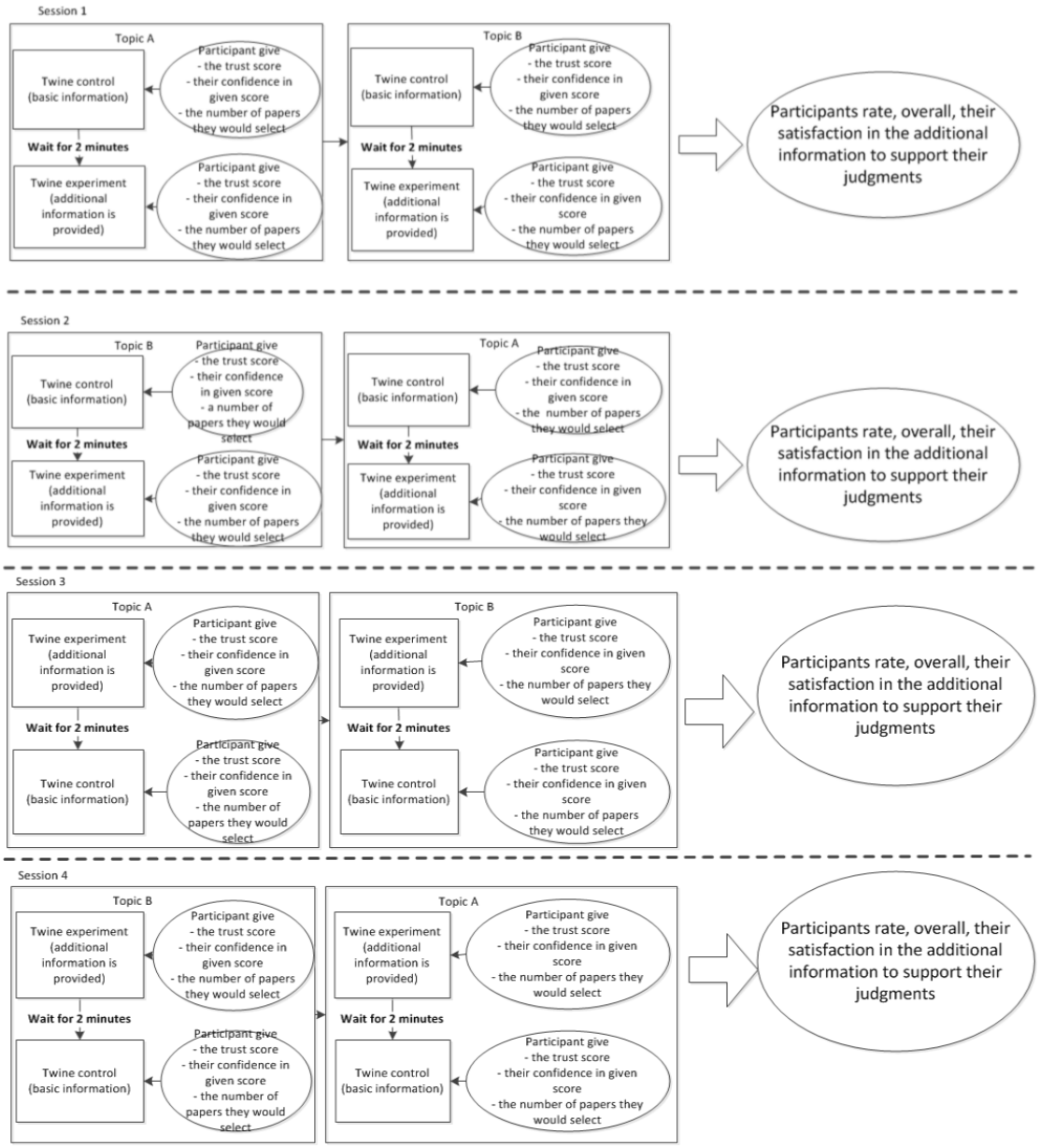


Figure 7.3: The complete study design plan of the evaluation of the TWINE framework

7.2 Framework Evaluation Study

The objectives of this study are to investigate and to assess the capacity of the TWINE framework to help users to evaluate the trustworthiness of Web information. In the previous section, we described the design of a study to perform such an evaluation. In this section, we explain the process of conducting said study.

7.2.1 Designing the Questionnaire

We designed a questionnaire to elicit responses from the participants based on the study plan in section 7.1.2. The purpose of this questionnaire is to allow the participants to rate the trustworthiness of the Web information they are consuming (in this case, the study focuses on research publications), and to assess their confidence in their ratings, and their level of satisfaction with the additional information provided by TWINE in order to support their assessment. In addition, the participants were asked to give the number of search results of research publications they would select to use in their report.

The questionnaire aims to verify our proposed framework, which is designed to help users assess the trustworthiness of Web information more critically based on the supportive information provided. There are two parts in this questionnaire:

- Part 1 consists of four tests; each test asked the participant to state the topic they were searching for and how familiar the participant is with the topic. In addition, each test asked the participants to rate the trustworthiness of Web information they were considering and to state how much confidence they had in that rating. Moreover, they were asked to answer how many papers they would select for use in their report and why they would or would not choose the publications.
- Part 2 asked for the participants' overall opinion concerning their level of satisfaction of the supportive information provided by the TWINE framework.

7.2.2 Identifying the Sample Size of the Potential Participants

We used a *priori* power analysis as discussed in section 5.1.3.1 to define the number of participants taking part in the study.

The designed study as discussed in section 7.1.2 was divided into four sessions. In each session, the same group of participants was asked to perform a search task using two environments: Twine 1 and Twine 2. Therefore, we tried to compare the difference between the level of confidence felt by the users when they evaluated the trustworthiness of academic publications available on the Web when they were given basic information (as in Twine 1) and when they were given both the basic and some additional information (as is the case in Twine 2). It was important to note, though, that we did not immediately assume that providing extra data would *improve* the user's confidence. Therefore, we conducted a two-tailed analysis of the survey results.

In this study, we chose α-level as 0.10 (based on Fisher's suggestion in (Fisher and Bennett, 1973)) because we wished to explore the possibility that the framework could increase the users' confidence when assessing trustworthiness of Web information. In addition, we set β as 0.2, which was the suggested maximum acceptable probability of

a Type II error based on the suggestion of Cohen (1992a). Moreover, we set a desired statistical power as 0.8 (1-β) because we wanted at least an 80% chance of detecting a statistically significant effect from the study results. We set the effect size (d) as 1.0 (large effect), according to the effect size conventions of Cohen (1992b), as we wished to be able to detect whether the difference of the user's confidence was statistically significant. This would mean that providing additional information helped the users to increase their confidence when they made a decision on whether or not to trust the information they were consuming.

Given the effect size, the α-level, and the statistical power as mentioned above, the minimum sample size of participants we needed to recruit is calculated as shown in Figure 7.4.

t tests – Means: Difference between two dependent means (matched pairs)			
Analysis:	A priori: Compute required sample size		
Input:	Tail(s)	=	Two
	Effect size d	=	1.0
	α err prob	=	0.1
	Power (1–β err prob)	=	0.8
Output:	Noncentrality parameter δ	=	2.8284271
	Critical t	=	1.8945786
	df	=	7
	Total sample size	=	8
	Actual power	=	0.8150305

Figure 7.4: A *priori* power analysis function in the G*Power to calculate the sample size of the study

As a result, the sample size for each session in the experiment (the experiment has four sessions) was estimated as eight participants, with α-level of 0.05 and the power of a statistical as 0.8. Therefore, in total, we needed 32 participants to participate in the study.

7.2.3 Defining the Panel of Potential Participants

For our study, we recruited a sample of 32 respondents ranging from 23 to 55 years of age from the University of Southampton, who have been studying at the postgraduate level for no longer than 2.5 years. The list of potential participants and contact details (e-mail address) were obtained from suggestions and introductions from the researcher's colleagues in different research groups and subject areas. Moreover, a poster requesting participants to sign up to the study, which included the e-mail address of the investigator was posted in the school building. We chose to use new postgraduate students as our sample group because they could be considered new to research, and to have less

experience in assessing the trustworthiness of Web information. Therefore, they needed a tool that could help them to evaluate the trustworthiness of Web information they would like to consume. Moreover, as a demographic, they were the most likely to benefit from using a tool developed using the TWINE framework, as TWINE would help them to critically assess the trustworthiness of Web information using the provided supportive information during the course of their research studies.

7.2.4 Materials

The study required a computer on which the Twine chrome extension was installed. In addition, a set of three types of documents were placed next to each console prior to the start of the experiment. The set of documents consisted of

- A consent form, which the participant must complete and return to the researcher before starting the experiment (see Appendix F)
- A set of instructions for the task the participant would perform (see Appendix G). We selected the topics for searching academic publications as "privacy", which was a generally term that postgraduate students in any subject area would know, and "Semantic Web", which was quite specific to some subject areas of study. We chose one topic which was fairly general and another topic which was quite specific because we wanted to investigate how the familiarity with the topic affects the user's confidence in evaluating the trustworthiness of Web information.
- A questionnaire, containing a set of questions that a participant needed to answer (see Appendix H).

7.2.5 Procedure

Three weeks before the actual study commenced, an invitation email, including a participant information sheet, was sent to the postgraduate student mailing lists of the Electronics and Computer Science Department at the University of Southampton and to the suggested participants such that the participants had enough time to respond to the request. Moreover, a poster requesting participation was posted in the school building. Those agreeing to participate were sent an e-mail which contained a link to an online scheduling system. This allowed the participants to select the date and time best suited to their schedule during which the study would be carried out. The experiment was divided into four sessions and the procedure for the experiment in each session was as follows.

7.2.5.1 Introduction (five minutes)

On the day of the study, at the start of the session, the participants were given instructions about the general nature of the experiment and the tasks they would perform in the session. In addition, the participants were asked to read and sign a consent form. After completing and returning the consent form, the participants started their tasks.

7.2.5.2 Performing tasks (20-45 minutes)

Each session ranged from 20 minutes to 45 minutes in length and the procedure is as follows:

1. The participant searched for papers on the topic given in the instruction sheet and they were asked to give an opinion on the perceived trustworthiness of the information contained within the top ten search results from each prototype environment. They were also asked how many papers they would select to use in their report.
2. After completing the search for relevant papers, the participant were asked to rate their satisfaction with the additional supportive information provided by the experiment prototype over the basic information provided by the control prototype.
3. The participant returned the questionnaire.

7.2.5.3 Ending (10 minutes)

At the end of each session, the participants were debriefed to inform and assure them about our ethical practice, and they were provided the opportunity to ask any questions they might have. We also used this opportunity to thank the participants for their participation.

7.3 Analysis and Results of the Study

This study investigated whether there were any significant differences in a user's ability to evaluate the trustworthiness of Web information between when they were provided with basic information about the Web information and when they were provided with additional supportive information. We wished to investigate whether using supportive information which was gathered and integrated by Semantic Web technologies allowed the user to improve their ability to evaluate the trustworthiness of Web information with confidence. In this section, we analysed the results from the participants' answers with regards to our hypotheses as discussed in section 7.1.1

7.3.1 Hypothesis 1: The users increase their confidence in their judgment of the trustworthiness of the Web information that they find

As part of the evaluation of this hypothesis, participants were asked to rate their confidence when assigning a trustworthiness value to the Web information they are consuming when using the Twine 1 prototype (which provides basic information) and the Twine 2 prototype (which provides additional supportive information). The sample size was more than 30; thus the sampling distribution tends to be normal (Field, 2009). A t-test[1] was chosen to explore the difference between the means of the users' reported levels of confidence in these two prototypes. Specifically, we used a paired-samples t-test because both groups of prototypes comprised the same participants. The details of the results of this analysis are discussed in section 7.3.1.1. Furthermore, in section 7.3.1.2, we investigate the significant differences in the changes in confidence of the users between when they were using the control prototypes (Twine 1) and when they were using the experiment prototype (Twine 2). This analysis was preformed over two topics in order to check whether or not the topic of the search had any effect on the users' confidence when making a judgment of the trustworthiness of Web information. In addition, we were also interested in the correlation between the familiarity of the topic of the users and their confidence in evaluating the trustworthiness of the Web information they consumed in that topic. This is discussed in section 7.3.1.3. Finally, in section 7.3.1.4, we analyse the variability of the participants' trustworthiness score when they perform the tasks with the control and experiment prototype.

7.3.1.1 A t-test for the Users' Confidence in assessing the Trustworthiness of Web Information between Two Prototypes

A paired-samples t-test was conducted to compare the confidence levels of users' evaluations of the trustworthiness of Web information when they used Twine 1 and when they were asked to evaluate the trustworthiness of Web information using Twine 2. Results from the t-test data analysis showed that there was a statistically significant difference in the confidence level of the users when they were given only basic information and when they were given additional supportive information.

On average, participants had significantly higher confidence in their ability to assess the trustworthiness of Web information based on the given supportive information (M = 71.42%, SE = 1.54) than to assess the trustworthiness of Web information based on only the basic information provided (M = 58.91%, SE = 1.68), $t(63)$ = -9.00, $p < 0.05$.

[1] A t-test is "a test statistic which is used to test whether the differences between two means are significantly different" Field (2009)

These results suggested that the confidence level of the users when evaluating the trustworthiness of Web information really did increase if they obtained useful supportive information about that Web information. Specifically, our results suggested that when users evaluated the trustworthiness of Web information using a tool which also provides supportive information along with the Web information, the mean of their confidence level in their judgments increased by 12.51 percentage points. A summary of the paired-samples t-test and a profile plot can be seen in Table 7.1, Table 7.2, and Figure 7.5.

Table 7.1: A paired samples statistics of the confidence level

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Confidence with basic information	58.91%	64	13.46%	1.68%
	Confidence with additional information	71.42%	64	12.32%	1.54%

Table 7.2: A paired samples test of the confidence level

		Paired Differences					t	df	Sig.(2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Confidence with basic information Confidence with additional information	-12.52%	11.12%	1.39%	-15.29%	-9.74%	-9.00	63	0.000

7.3.1.2 A t-test comparing the Change in Users' Confidence within Two Topics between Two Prototypes

In this analysis, we investigated whether the topic of the information for which the users was searching have any effect on the change of the users' confidence levels when users obtained additional supportive information compared to when they did not. In order to achieve this, a paired-samples t-test was conducted to compare the difference between the change in users' confidence levels with regards to the evaluation of the trustworthiness of Web information when 1) only basic information is provided, and 2) their confidence in doing the same when additional supportive information was provided with one topic and the changes in users' confidence in using two prototypes in another topic.

Results from the t-test data analysis showed that on average, the changing confidence of participants using two prototypes was higher in one topic (M = 12.84%, SE = 2.17) than when they assessed the trustworthiness of Web information using two prototypes on another topic (M = 12.19%, SE = 1.77). However, the difference in the mean was not significant $t(31)$ = -0.314, p > 0.05.

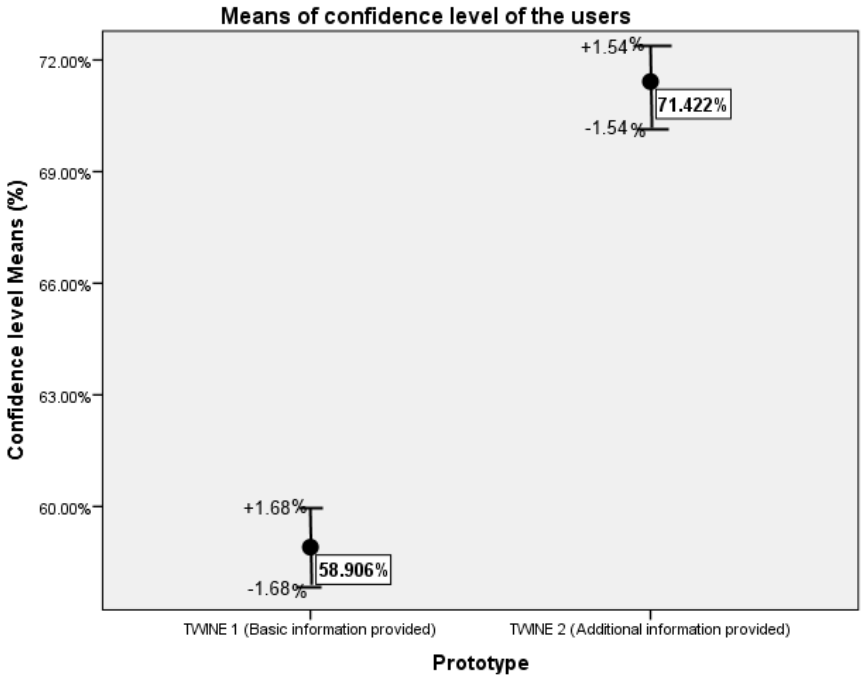


Figure 7.5: A profile plot of the mean confidence levels of users (plus and minus one standard error) in the control and experimental prototypes

These results suggested that the topic had no effect in increasing the confidence level of the users. There was no difference in the increase of confidence. A summary of the paired-samples t-test and a profile plot can be seen in Table 7.3, Table 7.4, and Figure 7.6.

Table 7.3: A paired samples statistics of the changes in confidence level

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Difference of confidence level first	12.19%	32	10.01%	1.77%
	Difference of confidence level second	12.84%	32	12.29%	2.17%

Table 7.4: A paired samples test of the changes in confidence level

		Paired Differences					t	df	Sig.(2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Difference of confidence level first Difference of confidence level second	-0.66%	11.81%	2.09%	-4.91%	3.60%	-0.314	31	0.755

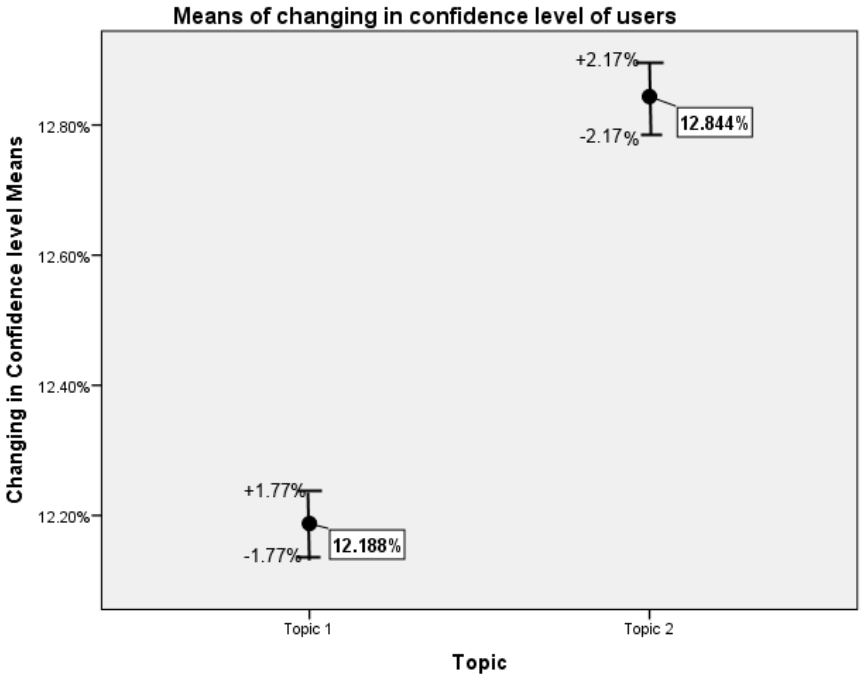


Figure 7.6: A profile plot of the means of the change in confidence level of users (plus and minus one standard error) in controls and experiment prototype

7.3.1.3 Correlations Analysis

A Pearson's (r) analysis was computed to assess the relationship between the user's familiarity with the topic being searched and the user's confidence in evaluating the trustworthiness of Web information when provided with basic information. In addition, we investigated how the changing intervals in users' confidence levels performs when basic information was provided and when they were provided with the additional information.

The results from correlation analysis showed that there was no correlation between the familiarity of the topic and the confidence level of users when the basic information is provided with r = 0.231, n = 64, $p > 0.05$ as shown in Table 7.5. However, the correlation between the increase in the users' confidence levels when provided with basic information and when provided with additional information showed that there was a negative correlation between the two conditions, r = -0.511, n =64, $p < 0.05$ as shown in Table 7.6. That was, when users already had a high confidence in the trustworthiness of a paper with little supportive information, providing the user with additional supportive information generated less of an increase in confidence.

The scatterplot in Figure 7.7 summarised these results. It showed that overall, there was a negative correlation between the confidence level of the users when providing them with basic information and the increase in their confidence level when providing them with supportive information.

Table 7.5: Correlations between the familiarity of the topic and the confidence level of users when provided with basic information

		Familiarity	Confidence with basic information
Familiarity	Pearson Correlation	1	0.231
	Sig. (2-tailed)		0.066
	N	64	64
Confidence with basic information	Pearson Correlation	0.231	1
	Sig. (2-tailed)	0.066	
	N	64	64

Table 7.6: Correlations of the change in users' confidence levels between when they are provided with basic information and when they are provided with additional information

		Confidence with basic information	Difference of confidence level
Confidence with basic information	Pearson Correlation	1	-0.511^{**}
	Sig. (2-tailed)		0.000
	N	64	64
Difference of confidence level	Pearson Correlation	-0.511^{**}	1
	Sig. (2-tailed)	0.000	
	N	64	64
** Correlation is significant at the 0.01 level (2-tailed)			

7.3.1.4 A User's given Trustworthiness Score Variance Analysis

In the questionnaire, we asked the participants to allocate a score to the trustworthiness of the Web information they were consuming. The results showed that there were cases in which the trustworthiness score given by participants increased when they obtained the additional information. However, there were also situations in which providing the user with additional information had no significant effect on their confidence. Furthermore, there were also even cases in which the participants gave a lower trustworthiness score when they saw more information from supportive information. Accordingly, we investigated the variability of the trustworthiness score given by participants using the F-ratio, which is calculated by equation 7.1 in order to define the case in the study:

$$F = \frac{\text{Var(TSbasicinformation)}}{\text{Var(TSadditionalinformation)}} \tag{7.1}$$

where *Var(TS basic information)* is the variance of the trustworthiness score given by participants when they obtain basic information and *Var(TS additional information)* is the variance of the trustworthiness score given by participants when they obtain additional information.

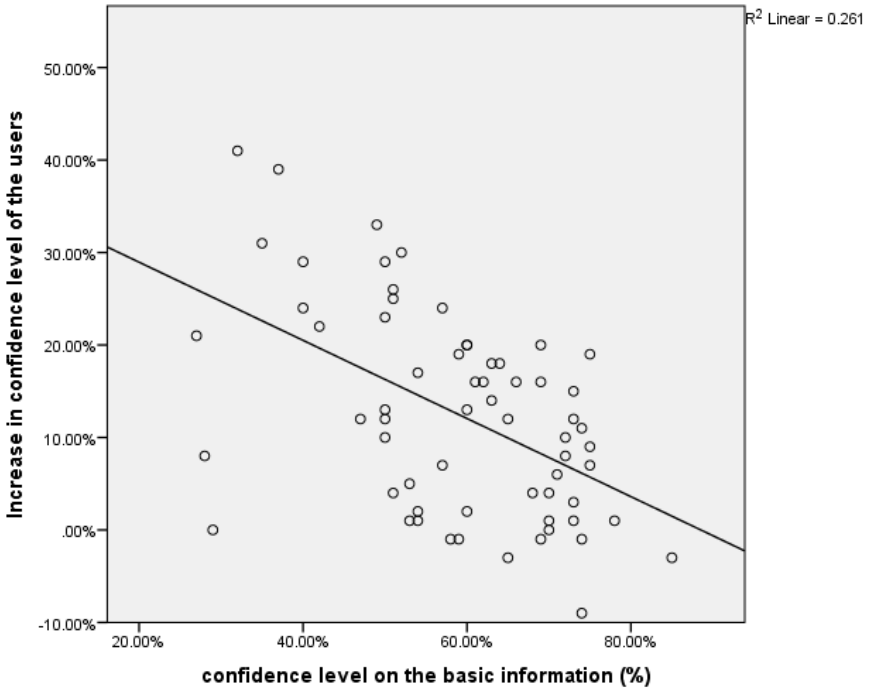


Figure 7.7: A scatterplot of the confidence levels of users when provided with basic information versus the increase in their confidence level when provided with additional information

We calculated the standard deviation of the trustworthiness score in the control and experiment prototypes using a paired-samples t-test in SPSS as shown in Table 7.7.

Table 7.7: Descriptive statistics of the trustworthiness score between control and experiment prototypes

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Trust score with basic information	4.56	64	0.814	0.102
	Trust score with additional information	5.22	64	0.629	0.079

Consequently, we computed the F-ratio using the standard deviation value from Table 7.7 as shown below:

$$F = \frac{(0.814)^2}{(0.629)^2} = \frac{0.669}{0.396} \approx 1.67, df = (63, 63) \tag{7.2}$$

From the table of critical values for the F-Distribution at the $p = 0.05$ (see Appendix I), the critical F-value with (63, 63) degrees of freedom was 1.53. The obtained F-ratio from the equation 7.2 was 1.67. Therefore, because the F-ratio was larger than the F-value, the variance of the trustworthiness score given by the participants when given only basic information was significantly larger than that given by the participants when they were given the additional information as shown in Figure 7.8. This means that

when participants obtained only basic information, they tended to give highly variable trustworthiness scores (i.e. some gave very high scores, some gave very low scores). As a result, the average of the trustworthiness score was on the scale of a neutral score (neither untrustworthy or trustworthy), whereas when participants obtained basic information and also received supportive information, they were more likely to give consistent trustworthiness scores. In addition, from Figure 7.8, on average participants allocated five score points to trustworthiness score (somewhat trustworthy).

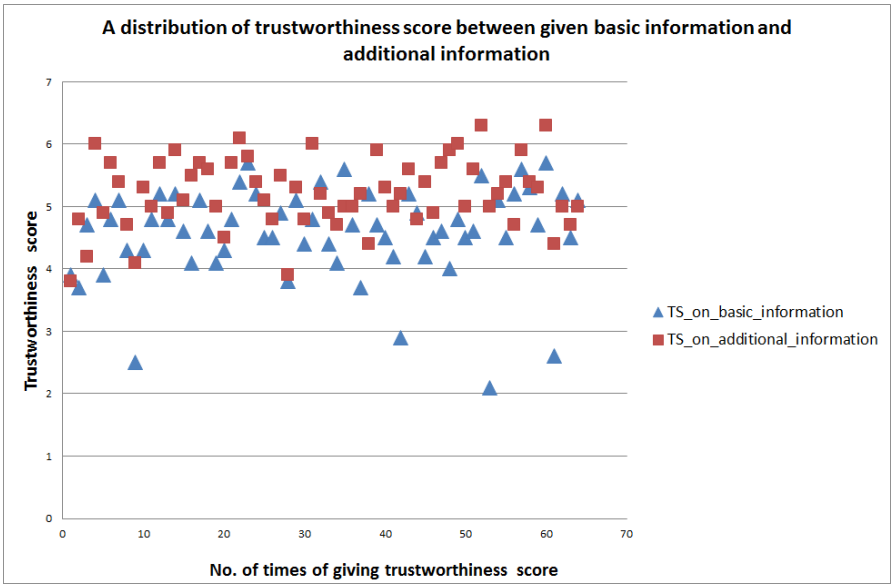


Figure 7.8: A scatter plot of the distribution of trustworthiness score in which the trustworthiness score scale, range from 1 = "Very untrustworthy" to 7 = "Very trustworthy"

7.3.2 Hypothesis 2: The users increase the number of pieces of trustworthy information which they select to use.

As part of the evaluation of this hypothesis, the participants were asked to select the Web information they would choose to use in their own work. In addition, they were asked to give brief details of which factors encourage them to make a decision to select any of the Web information which was being displayed. Nevertheless, it might be the case that the participants would not select any Web information. In this case, they were also asked to give the reason why they would *not* select any of the Web information available.

As a result, we selected a paired-samples t-test to analyse the difference of means of the number of selected items of Web information when using the control prototype and when using the experiment prototype as discussed the detail in section 7.3.2.1. In section 7.3.2.2, we analyse and discuss the participants' evaluation processes based on their answers regarding whether they would select each item of Web information or not.

7.3.2.1 A t-test Compares Difference Level of Means of Selected Web Information between Two Prototypes

A paired-samples t-test was conducted to compare the number of selected pieces of Web information in a basic information prototype and an additional supportive information prototype that were provided. Results from the t-test data analysis showed that on average, participants selected a significantly higher number of pieces of Web information when additional supportive information was provided ($M = 3.31$, $SE = 0.27$) than they did when only basic information was provided ($M = 2.75$, $SE = 0.24$), $t(63) = -2.55$, $p < 0.05$.

These results suggested that, on average, the number of selected pieces of Web information did increase if they obtain supportive information about that Web information. Specifically, our results suggested that when users evaluated the trustworthiness of Web information which also provided supportive information along with the Web information, the number of pieces of Web information selected by the users increased on average by 0.56 percentage points. A summary of the paired-samples t-test and a profile plot can be seen in Table 7.8, Table 7.9, and Figure 7.9.

Table 7.8: A paired samples statistics of the number of selected items of Web information

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	No. selected papers with basic information	2.75	64	1.919	0.240
	No. selected papers with additional information	3.31	64	2.181	0.273

Table 7.9: A paired samples test of the number of selected items of Web information

		Paired Differences					t	df	Sig.(2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	No. selected papers with basic information - No. selected papers with additional information	-0.563	1.763	0.220	-1.003	-0.122	-2.553	63	0.013

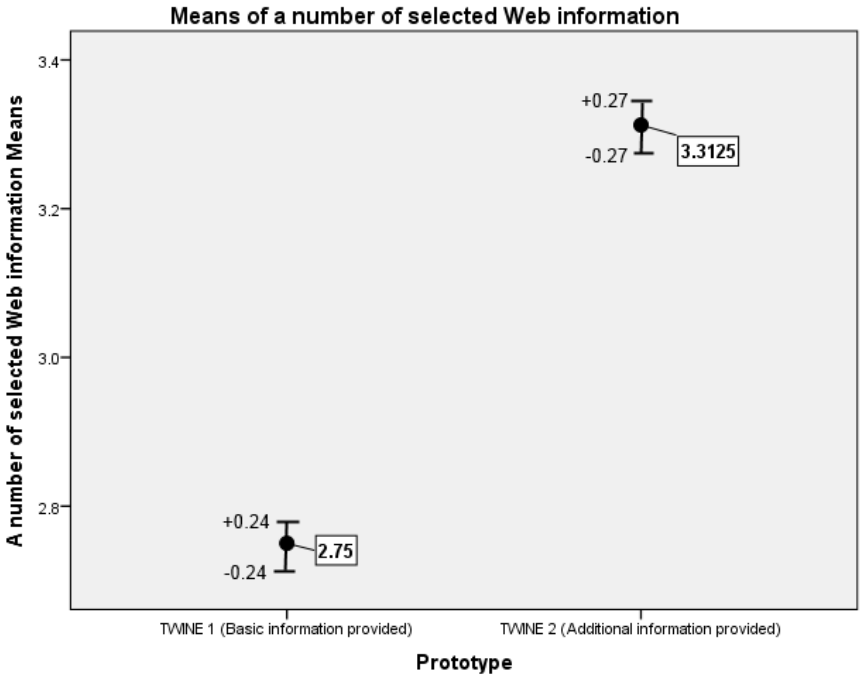


Figure 7.9: A profile plot of the means of the number of selected items of Web information (plus and minus one standard error) in the control and experiment prototypes

7.3.2.2 Thematic Analysis of the Participant Behaviour of Evaluation of the Trustworthiness of Web Information

We used the inductive methodology as mentioned in section 5.2.2.1 for analysis of the answers from the participants in order to investigate the themes related to their evaluation process. In particular, we were interested in the themes that represented the behaviours of the participants. The details of the thematic analysis process can be seen in sections 5.2.2.1.2 and 5.2.2.2. We used NVivo version 10 to help us processed and analysed the responses from the participants.

7.3.2.2.1 Preparing data for analysis

We used the iSurvey system of the University of Southampton to record the answers from the designed questionnaire in section 7.2.1 (see Appendix H). The system created a unique ID for each participant. Then, we gathered the answers from each participant when they performed the search in each prototype; Twine 1 (control) and Twine 2 (experiment). We asked them to give the reasons for why they did or did not choose to select a piece of information. Their responses were then recorded into a separate Excel worksheet. After that, we transferred the data from Microsoft Excel to Microsoft Word in which we set the participant's ID with Heading 2 style and set the answers as

normal style. Finally, we converted the table data format to text format. As a result, Microsoft Word generated a word document which contained the participants' IDs and their answers, as shown in Figure 7.10 and Figure 7.11. Then, we import these response text documents into NVivo as shown in Figure 7.12.

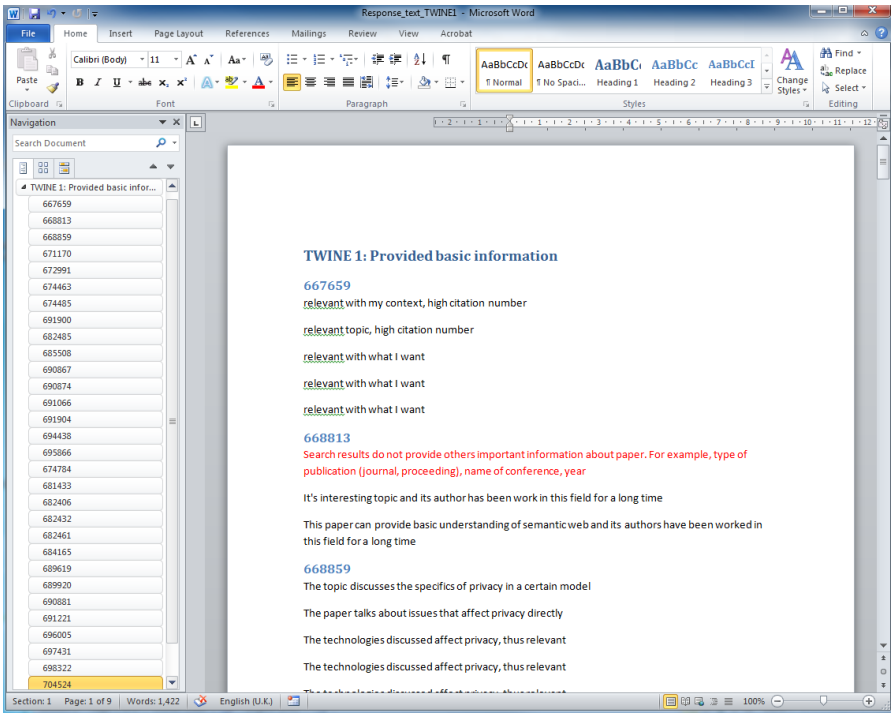


Figure 7.10: An example of the preparation of the response text of the participants when they use the Twine 1 prototype in Microsoft Word

Completing this preparation step allowed us to explore the data in more detail by using query functions in NVivo. We discuss the details in the following section.

7.3.2.2.2 Generating initial codes

We used the word frequency tool of the NVivo software to develop a coding scheme which were used to create the structure and theme in the next phases. We set parameters for counting the appearance of words including stemmed words and showing the 100 most frequently occurring words of four characters or more in length. The results from the function can be displayed in Figure 7.13.

We considered the words from the word frequency list against the response text because not all of the values' output from this function were meaningful as indications of useful concepts. For example, words might be too general and therefore not expressed any specific pattern, such as the word "paper", which participants might have mentioned within the corresponding question. These words did not lend themselves to any particular thematic code. Consequently, we manually excluded these words that were irrelevant

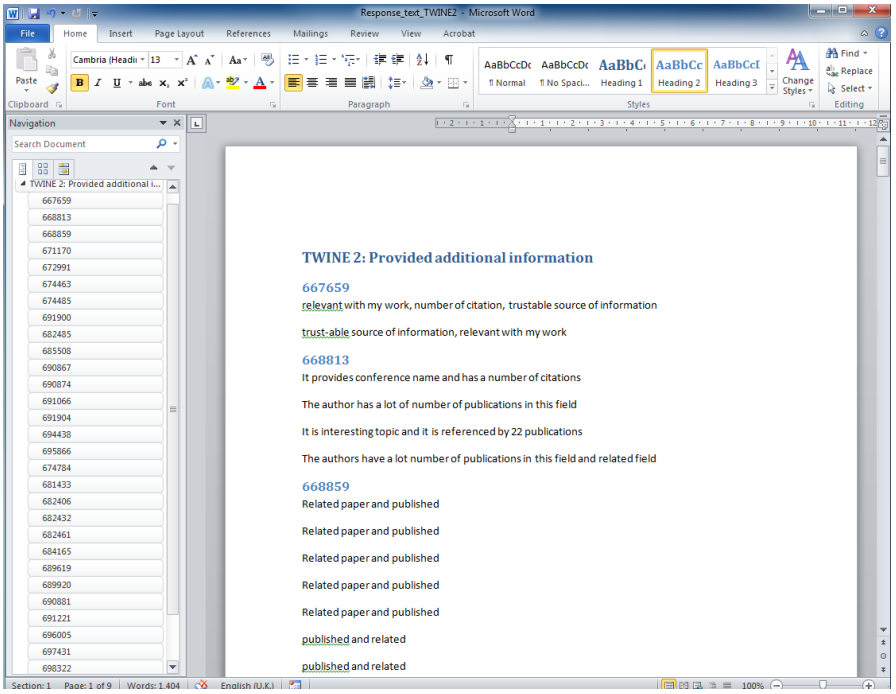


Figure 7.11: An example of the preparation of the response text of the participants when they use the Twine 2 prototype in Microsoft Word

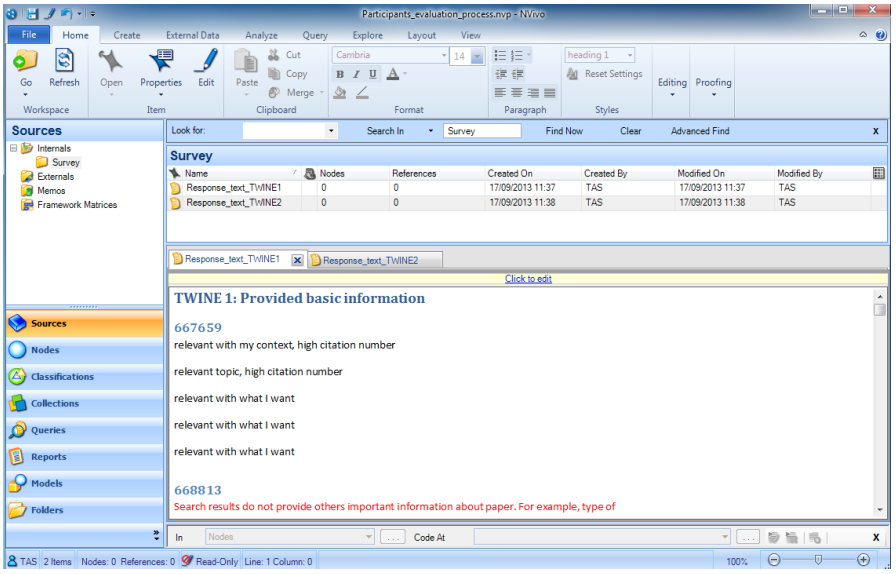
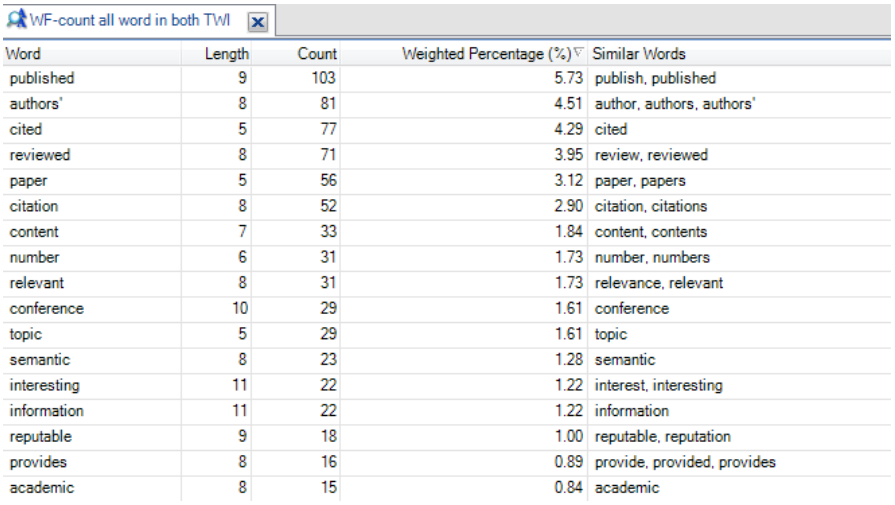


Figure 7.12: An example of importing response texts regarding the participants' decisions on whether or not to select Web information into NVivo



WF-count all word in both TWI

Word	Length	Count	Weighted Percentage (%)	Similar Words
published	9	103	5.73	publish, published
authors'	8	81	4.51	author, authors, authors'
cited	5	77	4.29	cited
reviewed	8	71	3.95	review, reviewed
paper	5	56	3.12	paper, papers
citation	8	52	2.90	citation, citations
content	7	33	1.84	content, contents
number	6	31	1.73	number, numbers
relevant	8	31	1.73	relevance, relevant
conference	10	29	1.61	conference
topic	5	29	1.61	topic
semantic	8	23	1.28	semantic
interesting	11	22	1.22	interest, interesting
information	11	22	1.22	information
reputable	9	18	1.00	reputable, reputation
provides	8	16	0.89	provide, provided, provides
academic	8	15	0.84	academic

Figure 7.13: An example of the word frequency count output of participants' responses basic list display

to understanding a participant's behaviour when evaluating the trustworthiness of information. As a result, we obtained a list of words that defined a set of initial coded nodes which are meaningful and relevant to the study as shown in Table 7.10.

Table 7.10: A word frequency table of potential initial codes

Word	Length	Count	Similar Words
published	9	103	publish, published
authors'	8	81	author, authors, authors'
cited	5	78	cite, cited
reviewed	8	71	review, reviewed
citation	8	52	citation, citations
content	7	33	content, contents
relevant	8	31	relevance, relevant
conference	10	29	conference
topic	5	29	topic
interesting	11	22	interest, interesting
reputable	9	18	reputable, reputation
know	4	16	know
academic	8	15	academic
abstract	8	14	abstract
report	6	12	report
source	6	12	source
seems	5	11	seem, seems
book	4	10	book
known	5	10	known
like	4	8	like, likely

Table 7.10: A word frequency table of potential initial codes

Word	Length	Count	Similar Words
make	4	8	make, makes
recent	6	8	recent
thesis	6	8	thesis
type	4	8	type
introduction	12	7	introduction
name	4	6	name
words	5	6	word, words
detail	6	5	detail, details
experience	10	5	experience
might	5	5	might
sounds	6	5	sound, sounds
status	6	5	status
style	5	5	style
consider	8	4	consider, considered, considering
overview	8	4	overview
still	5	4	still
title	5	4	title
unpublished	11	4	unpublished
understand	10	3	understand, understanding

We found that some coded nodes have the same meaning, such as “cited” and “citation” (both of which refer to the piece of the information having been mentioned or referenced by other publications). Thus, we merged these coded nodes into the same group of coded nodes. Consequently, we created the primary code from Table 7.10 such that the coded nodes corresponded to the purpose of our investigation. In addition, we were interested in discovering any themes that might arise across the prototypes. Therefore, we designed our codes to use a unified coding scheme which was common to all questions in each prototype. Consequently, we obtained the primary coded nodes displayed in Table 7.11.

Table 7.11: A list of primary codes

No.	Coded Nodes	Sources	References
1	Abstract	2	14
2	Academic	2	15
3	Author	2	81
4	Book	1	10
5	Citation	2	130
6	Conference	2	29

Table 7.11: A list of primary codes

No.	Coded Nodes	Sources	References
7	Consider	2	4
8	Content	2	44
9	Detail	2	5
10	Experience	1	5
11	Interesting	2	22
12	Known	2	26
13	Make sense	2	8
14	Might	2	5
15	Peer-reviewed	1	71
16	Published	2	103
17	Recent	2	8
18	Relevant	2	46
19	Report	2	12
20	Reputable	2	18
21	Seems	2	19
22	Sounds	2	5
23	Source	2	12
24	Status	1	5
25	Style	1	5
26	Thesis	1	8
27	Title	2	4
28	Topic	2	29
29	Type of publication	2	8
30	Understand	1	3
31	Unpublished	1	4
32	Words	1	6

Table 7.11 shows the coded nodes, the number of sources (response texts) in which each code appears, and the number of times that each code has been referenced. As a consequence, the completion of this phase generated 32 coded nodes.

7.3.2.2.3 Structure node scheme

From Table 7.11, we found that some coded nodes were adjective words and expressed an attitude such as 'seems', 'sounds', and 'interesting'. Therefore, we defined these coded nodes as an attitude of participants to describe their opinions, whereas others were items that *relate* to the attitude. Other coded nodes were items which influence the

participants' judgment of the trustworthiness of Web information. Moreover, we found that there were two main actions which the participants performed with the information on papers they were looking at. Consequently, we categorised these preliminary coded nodes into three groups; namely, attitudes, items, and actions. Then, we created groups of nodes as a hierarchical structure using the "Tree nodes" functions of Nvivo as shown in Figure 7.14.

Name	Sources	References
Actions	0	0
Reject	0	0
Select	0	0
Attitudes	0	0
Consider	2	4
Interesting	2	22
Known	2	26
Make sense	2	8
Might	2	5
Seems	2	19
Sounds	2	5
Understand	1	3
Items	0	0
Abstract	2	14
Academic	2	15
Style	1	5
Words	1	6
Author	2	81
Detail	2	5
Experience	1	5

Figure 7.14: An example of a tree nodes

The details of the members in each group are described as below:

- **Action group** This group contained coded nodes which represented the actions of the participants when they were evaluating the trustworthiness of Web information. We obtained these coded nodes by deriving them from the responses of the participants to questions 3 and 4 in each section of the questionnaire. Therefore, we defined two main coded nodes - select and reject - regarding the action the participants perform in the study.

- **Attitudes group** This group consisted of nodes that relate to the participants' attitudes toward Web information or the items in the Web information. Table 7.12 shows the coded nodes within this group.

Table 7.12: Coded nodes and their frequencies that are categorised for the attitudes group

Coded Nodes	No. Sources	No. References
Consider	2	4
Interesting	2	22
Known	2	26
Make sense	2	8
Might	2	5
Seems	2	19
Sounds	2	5
Understand	1	3

- **Items group** This group contained coded nodes that related to the items that affected the participants' decisions on whether or not to select a piece of information. Table 7.13 shows the coded nodes within this group and the number of times each nodes has been referred to by the participants.

Table 7.13: Coded nodes and their frequencies that are categorised for the items group

Group	Coded Nodes	No. Sources	No. References
Author	Detail	2	5
	Experience	1	5
	Reputable	2	18
Citation		2	130
Content	Abstract	2	14
	Brief Introduction	2	7
	Overview	2	4
Recent		2	8
Relevant		2	31
Source		2	12
Status	Peer-reviewed	1	71
	Published	2	103
	Unpublished	1	4
Style	Words	1	6
Title		2	4
Topic		2	29
Type	Book	1	10
	Conference	2	29
	Report	2	12
	Thesis	1	8

7.3.2.2.4 Searching for themes

We aimed to explore the pattern of evaluation followed by users when evaluating the trustworthiness of Web information. Therefore, we focused on the actions and the items that might affect the participants' decisions. In addition, we were interested in the attitudes of the participants when they were evaluating the Web information with the basic information and with additional supportive information. Therefore, we defined two main themes of the patterns generated by the participants' responses, and the action corresponding to each theme as shown in Figure 7.15.

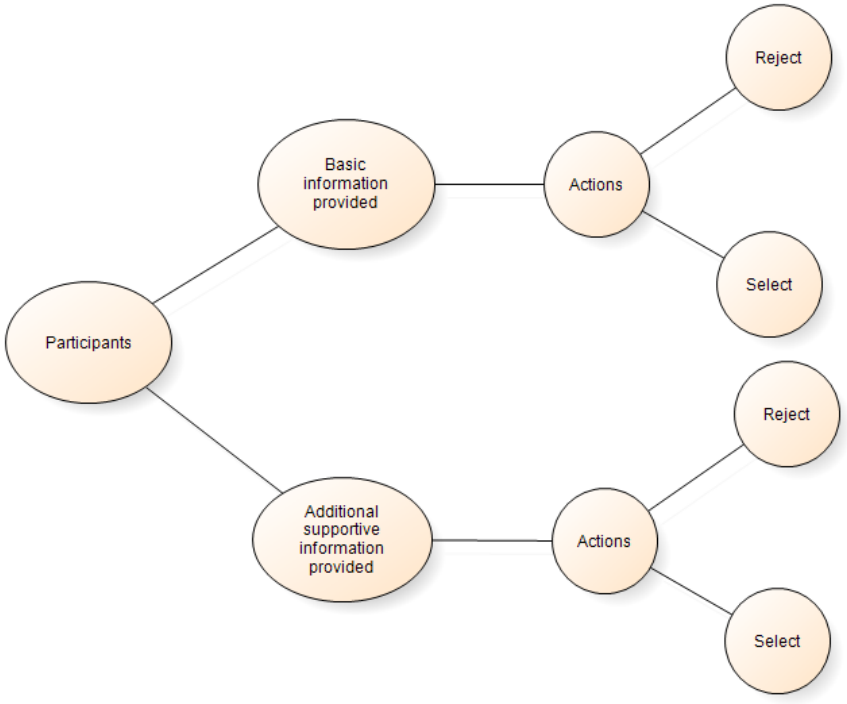


Figure 7.15: The initial main themes of the participants' evaluation process

Figure 7.15 shows the process of assessing the trustworthiness of Web information of participants which could be divided into two main themes; namely, when they acquired the provided basic information and when they acquired the additional information. Each theme consisted of corresponding actions. In addition, reject and select actions were sub-actions that happened when participants evaluated the trustworthiness of Web information they were looking at.

We explored to find the items which might affect the participants' judgments and the reasons why they would not select a piece of information in more detail in the next section.

7.3.2.2.5 The process of evaluating the trustworthiness of Web information based on the basic information provided

In this section, we investigated the process of evaluating the trustworthiness of Web information followed by participants based on the provided basic information about the piece of information they are looking for. We analysed data collected from the responses when the participants used the Twine 1 prototype in both topics (privacy and Semantic Web) using the model functions of NVivo. The model built from the function can be seen in Figure 7.16.

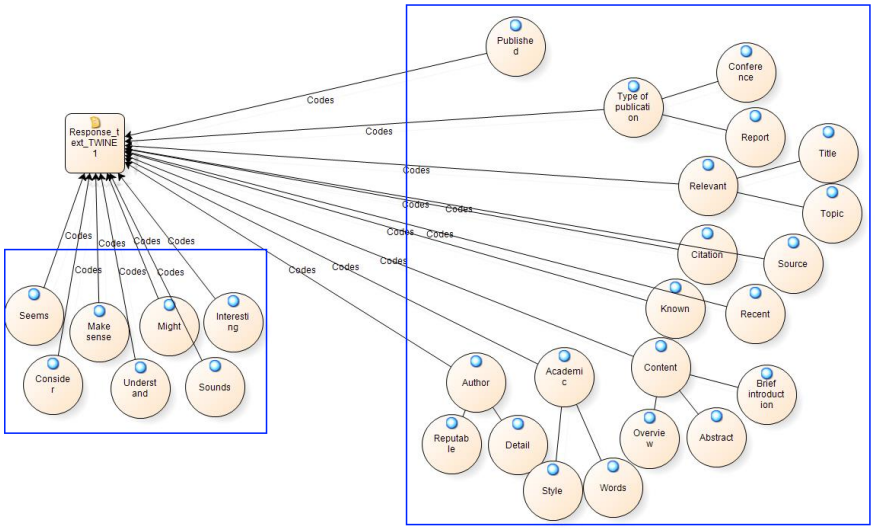


Figure 7.16: A model of coding related to the participants' responses when they are using the Twine 1 prototype

From Figure 7.16, coded nodes can be grouped into two groups based on the structure node schema discussed in section 7.3.2.2.3. The attitudes group mentioned the nodes "seems", "make sense", "consider", "understand", "might", "sounds", and "interesting". The items group mentioned the "author" node which has sub-nodes of "reputable" and "detail" nodes. In addition, it mentioned the nodes "academic", "content", "know", "recent", "source", "relevant", "type of publication", and "published". The details of sub-nodes can be seen in Figure 7.16.

According to the initial main themes discussed in section 7.3.2.2.4, we explored the response texts of the participants in more detail. We found that the patterns of evaluating the trustworthiness of Web information based on the provided basic information would be one of two cases; either selected at least one publication from the top ten search results or rejected all the top ten search results. The participants would select that information because it provided data based on the items shown in Figure 7.16. However, they would reject the information because the information lacks supportive information

that could support their decisions. Examples of statements from the participants (directly quoted from participants) to describe their reasons for why they would reject the information are shown belows:

- participant 668813 mentioned that "*Search results do not provide others important information about paper. For example, type of publication (journal, proceeding), name of conference, year.*"
- participant 691904 said "*None shows very high citation count & some come across as partisan in their description; i.e. they sound more like opinion pieces than serious scientific discussions.*"
- participant 689920 suggested that "*There is not enough information for me to select anything given my comparative ignorance of the topic. I would try and find high cited papers first to give me more information on other worthwhile papers.*"

In addition, the participant had less confidence to make a decision on whether to trust a piece of information when there was limited information provided. For example,

- participant 674784 stated that "*I would not be able to judge the content of the papers as there is no evidence of conference, author or even the paper itself being trustworthy, such as citations, impact factors, and author's h-index.*"
- participant 698322 mentioned that "*There is no information to prove all those paper. For instance, the number of citation.*"

As a result, there was a case that the participant rejected to use that piece of information. Therefore, we added the "**lack of supportive information**" as the sub-theme of the reject theme.

Alternatively, when they made a decision to select a piece of information, they based their decision on mixed criteria between *objective* criteria which could help to identify the trustworthiness of a piece of information such as the citation or the type of publication (e.g. journal, proceeding) and *subjective* criteria which were not as useful because they could be biased or disguised, like writing style and the familiarity with the names of the authors. Therefore, we added "**objective criteria**" and "**subjective criteria**" as sub-themes of the "select" theme. In addition, the participants' responses showed that when only basic information was provided, the participants had less confidence in evaluating the trustworthiness of information. Accordingly, we added "**less confidence**" as a sub-themes of the "subjective criteria" theme. Consequently, we obtained a process of evaluating the trustworthiness of Web information when basic information is provided as shown in Figure 7.17

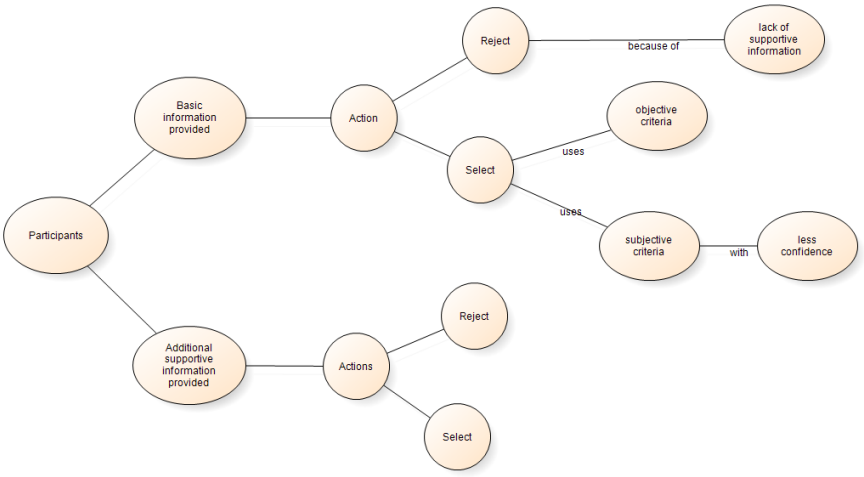


Figure 7.17: The evaluation process initial main- and sub-themes when provided with basic information

7.3.2.2.6 The process of evaluating the trustworthiness of Web information based on the additional supportive information provided

Similar to the analysis process in section 7.3.2.2.5, we started by exploring the overall nodes related to the responses of the participants when they used the Twine 2 prototype. The result of the model function is shown in Figure 7.18.

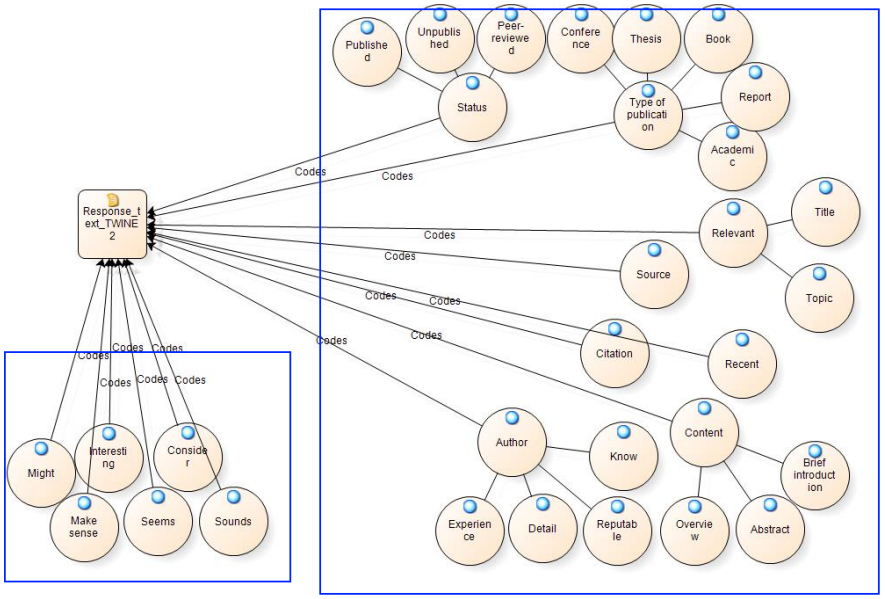


Figure 7.18: A model of coding related to the participants' responses when they are using the Twine 2 prototype

Figure 7.18 shows that participants used the same group of items as when they evaluated the trustworthiness of Web information when basic information was provided (detail in section 7.3.2.2.5). However, coded nodes in the attitudes group appear less frequently

in the response texts of the participants who used the Twine 2 prototype than when the participants used the Twine 1 prototype. For example, only one participant mentioned their reasons to select a piece of information because the information seems relevant to their search terms which shows their uncertain decision using the term "seems relevant".

In addition, the response texts of the participants when they used the Twine 2 prototype showed that the participants made a decision to select the information based on *objective* criteria to support their judgment and to give them more confidence. For instance,

- participant 682485 mentioned that "*It provides scope of work and published status.*"
- participant 690867 said "*Sufficient information are provided including considerably number of citation comparing to the others.*"
- participant 694438 suggested that "*It might be presented at a workshop but it is published and cited so many times!*"
- participant 682461 mentioned that "*The author has much experience and this work was published with peer-review.*"
- participants 691221 said "*This paper gives briefly meaning of semantic web with its evaluation. It is relevant to semantic web and also has evaluation to guarantee the trust.*"
- participant 704544 mentioned that "*This is a good and clear meaning of semantic web that I want to find. Also, the file was up to date and published in the book section.*"

The comments from participants implied that they gained more confidence in the trustworthiness of Web information based on objective criteria such as published status, number of citations, or whether the work was peer-reviewed before publication. Therefore, we added the "**objective criteria**" node as a new sub-theme of the "select" theme of the main theme, and "additional supportive information provided" and, "**more confidence**" nodes as new sub-themes of "objective criteria" as shown in Figure 7.19.

7.3.2.2.7 Reviewing the themes

After completing the process of developing the themes, we manually reviewed the candidate themes that we obtained with the response texts. We found that, using the Twine 2 prototype the participants were always able to select at least one publication to use in their tasks with confidence and using objective criteria. Therefore, we refined our theme by deleting the "**reject**" action from the theme "additional supportive information provided". As a result, we refined the candidate theme and defined it as the themes of the participant's evaluation process as shown in Figure 7.20.

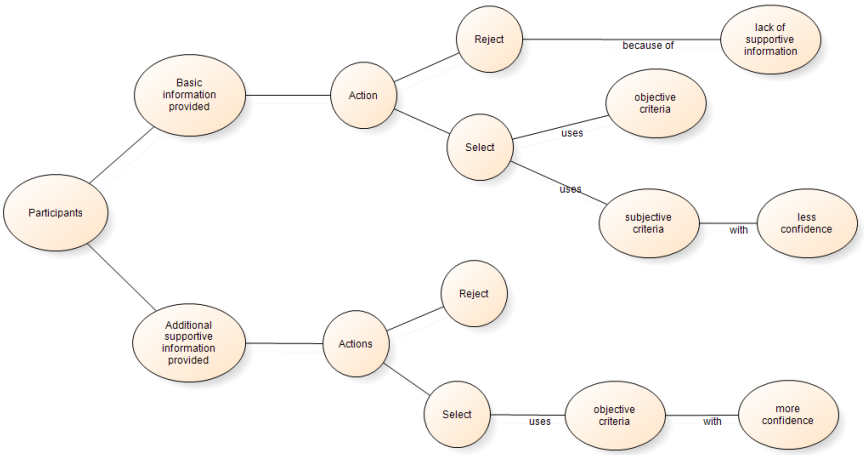


Figure 7.19: The evaluation process initial main- and sub-themes when the users were provided with additional information

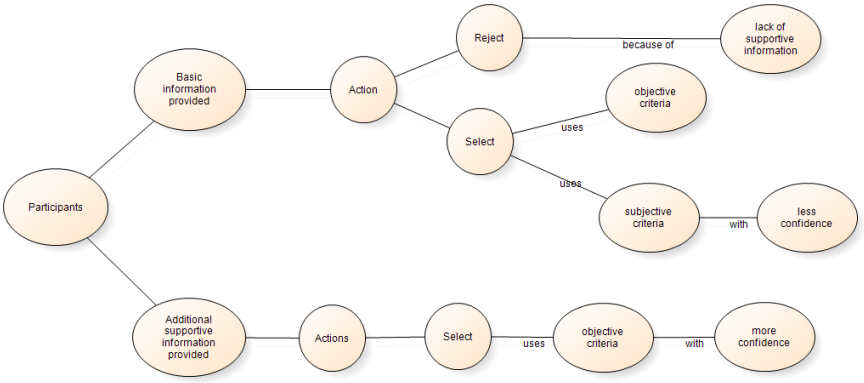


Figure 7.20: The refined main- and sub-themes of the participants evaluation process

In summary, when participants used the Twine 1 prototype, which provided only basic information about the publication, they would select any publications based on a combination of subjective and objective criteria. Specifically, they seemed to be unsure about their decisions to select a publication if they based their assessment on subjective criteria alone. In addition, they would not select any publications because they were unable to find strong supportive information, whereas when the participants obtained basic information and additional supportive information to support their decision (used Twine 2 prototype), they tended to gain more confidence in their ability to evaluate the trustworthiness of Web information, as they were always able to make a decision and select at least one publication to use in their work. Moreover, they would base their assessment on objective criteria.

7.3.3 Hypothesis 3: The users are satisfied with the supportive information insofar as it helps them to evaluate the trustworthiness of Web information.

As part of the evaluation of this hypothesis, the participants were asked to rate their level of satisfaction with the supportive information and its effect on their evaluation process. In addition, participants were asked to rate the influence the information has on their judgment using a Likert scale[2]. In addition, we also asked the participants to state their opinion regarding the perceived level of usefulness (participants think the supportive information is a good indicator to identify the trustworthiness of Web information) and helpfulness (participants think the supportive information helps to support their judgment) of the supportive information the framework provided. However, the sample size was more than 30; thus the sampling distribution tends to be normal (Field, 2009). Therefore, our collected data met the assumptions of parametric tests.

As a result, we used a one-sample t-test[3] in order to investigate whether the participants were satisfied with the supportive information provided by the framework. We also used a one sample t-test to assess the influence of the supportive information on the participants' judgment, and the participants' overall opinions of the helpfulness and usefulness of the supportive information. The details of each analysis are discussed in the following sections.

7.3.3.1 The Satisfaction of the Participants with the Additional Information provided by the Framework

A one-sample t-test was conducted to compare the satisfaction level of the participants towards the supportive information provided. In this analysis, a constant value, set as 1, represents the opinion of "*not at all satisfied*", and the value 5 represents the opinion of "*extremely satisfied*". Results from the t-test data analysis showed that the mean satisfaction level of the participants (M=3.69, SD=0.69) was statistically significantly larger than the "*not at all satisfied*"; $t(31)$=21.95, $p < 0.05$.

These results suggested that, on average, the participants were satisfied with the supportive information provided. The scale used in this analysis has five points. They are not at all satisfied (1), slightly satisfied (2), moderately satisfied (3), very satisfied (4), and extremely satisfied (5). Consequently, our results suggested that the participants were moderately satisfied with the supportive information provided (M=3.69). A summary of a one-sample t-test and a histogram can be seen in Table 7.14, Table 7.15, and Figure 7.21.

[2]Likert scale is a psychometric response scale which is used in questionnaires in order to obtain a participant's preferences or degree of agreement with a statement or set of statements (Likert, 1932)

[3]A one-sample t-test is a statistical procedure for testing the difference between the mean value of a sample and a hypothesis value (a constant value) (IBM Corporation, 2012).

Table 7.14: A one sample statistics of the satisfaction level of the participants

	N	Mean	Std. Deviation	Std. Error Mean
Satisfaction score	32	3.69	0.693	0.122

Table 7.15: A one sample test of the satisfaction level of the participants

	Test value = 1 (not at all satisfied)					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Satisfaction score	21.947	31	0.000	2.688	2.44	2.94

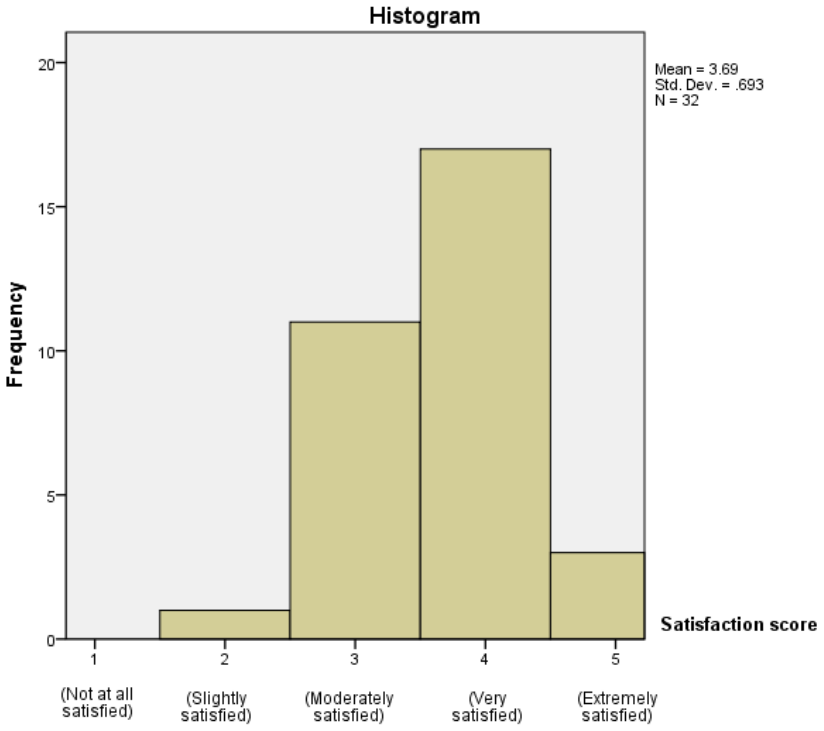


Figure 7.21: A histogram of the satisfaction score

7.3.3.2 The Influence of the Supportive Information Provided on the Participants' Judgment of the Trustworthiness of Web Information

A one-sample t-test was conducted to compare the influence level of the supportive information provided from the framework affects on the participants' judgment of the trustworthiness of Web information. Similar to before, a constant value, set as 1, represents the opinion of "*not at all influential*" whereas the value 5 represents "*extremely influential*". Results from the t-test data analysis showed that the mean of the influence on the participants (M=3.72, SD=0.68) was statistically significantly larger than "*not at all influential*"; $t(31)$=22.51, $p < 0.05$.

These results suggested that on average the participants were influenced by the supportive information provided. The scale used in this analysis consists of five points: not at all influential (1), slightly influential (2), somewhat influential (3), very influential (4), and extremely influential (5). Consequently, our results suggested that the participants were influenced by the supportive information provided (with the mean influence being M=3.72, or "somewhat influential"). A summary of the one-sample t-test and a histogram can be seen in Table 7.16, Table 7.17, and Figure 7.22.

Table 7.16: One sample statistics of the influence level of the supportive information on participants' judgment

	N	Mean	Std. Deviation	Std. Error Mean
Influence provided supportive data	32	3.72	0.683	0.121

Table 7.17: One sample test of the influence level of the supportive information on the participants' judgments

	Test value = 1 (not at all influential)					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Influence provided supportive data	22.512	31	0.000	2.719	2.47	2.97

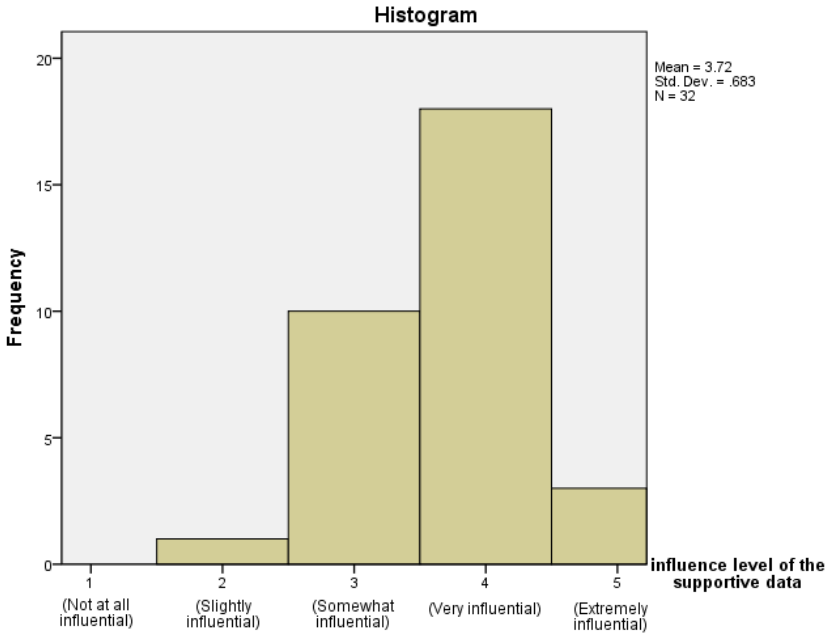


Figure 7.22: A histogram of the influence level of the supportive information on participants' judgments

7.3.3.3 The Additional Information about the Authors is Helpful to Support the Participants' Evaluation Process

A one-sample t-test was conducted to compare the level of agreement of the participants over the helpfulness of the additional information about the authors. In this analysis, a constant value, set as 3, represents the neutral opinions (neither agree nor disagree). The score "5" represents "*strongly agree*" and conversely, the score "1" represents "**strongly disagree**". Results from the t-test data analysis showed that the mean of the participants' agreement over the helpfulness of the additional information about the authors (M=4.75, SD=0.62) was statistically significantly larger than the "*neither agree nor disagree*"; $t(31)$=15.91, $p < 0.05$.

In more detail, we used a five-point scale which values were as follows; strongly disagree (1), disagree (2), neither agree nor disagree (3), agree (4), and strongly agree (5). Therefore, the results suggested that, on average, the participants significantly agree that the additional information about the authors was helpful. A summary of the one-sample t-test and a histogram can be seen in Table 7.18, Table 7.19, and Figure 7.23.

Table 7.18: One sample statistics for an agreement over the helpfulness of the additional information regarding the authors

	N	Mean	Std. Deviation	Std. Error Mean
The additional information about the authors is helpful	32	4.75	0.622	0.110

Table 7.19: One sample test for an agreement over the helpfulness of the additional information regarding the authors

	Test value = 3 (Neither agree nor disagree)					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
The additional information about the authors is helpful	15.911	31	0.000	1.75	1.53	1.97

7.3.3.4 The Additional Information about the Publication is Helpful to Support the Participants' Evaluation Process

A one-sample t-test was conducted to compare the level of agreement of the participants over the helpfulness of the additional information about the publications. The same scale was used here as was used in the previous analysis. Results from the t-test

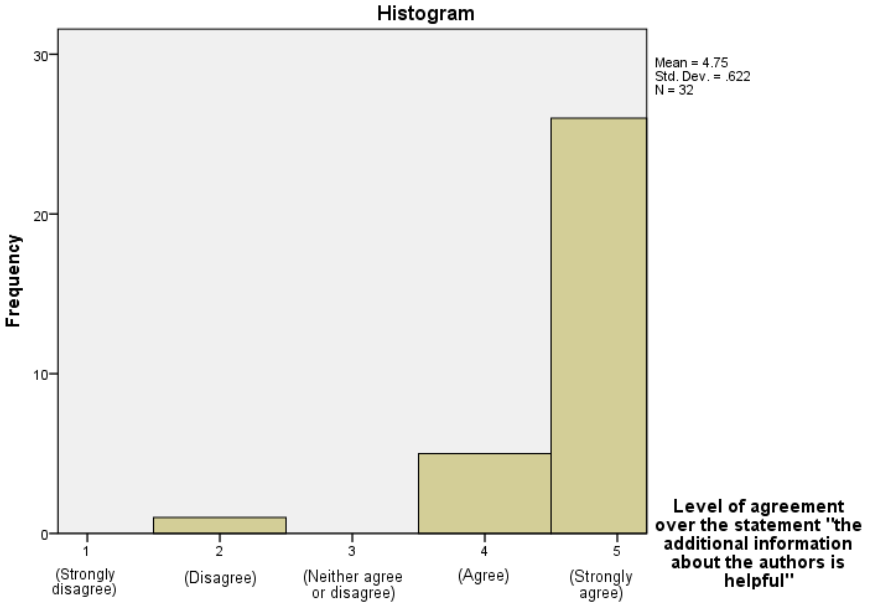


Figure 7.23: A histogram graph showing the participants responses to the statement that "the additional information of the authors is helpful"

data analysis showed that the mean of the participants' agreement over the helpfulness of the additional information about the publication (M=4.59, SD=0.56) was statistically significantly larger than the "*neither agree nor disagree*"; $t(31)$=16.10, $p < 0.05$. Thus, the results suggested that, on average, the participants significantly agree that the additional information about the authors was helpful. A summary of the one-sample t-test and a histogram can be seen in Table 7.20, Table 7.21, and Figure 7.24.

Table 7.20: One sample statistics for level of agreement over the helpfulness of the additional information of the publication

	N	Mean	Std. Deviation	Std. Error Mean
The additional information about the publications is helpful	32	4.59	0.560	0.099

Table 7.21: One Sample Statistics test for level of agreement over the helpfulness of the additional information of the publication

	Test value = 3 (Neither agree nor disagree)					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
The additional information about the publications is helpful	16.102	31	0.000	1.594	1.39	1.80

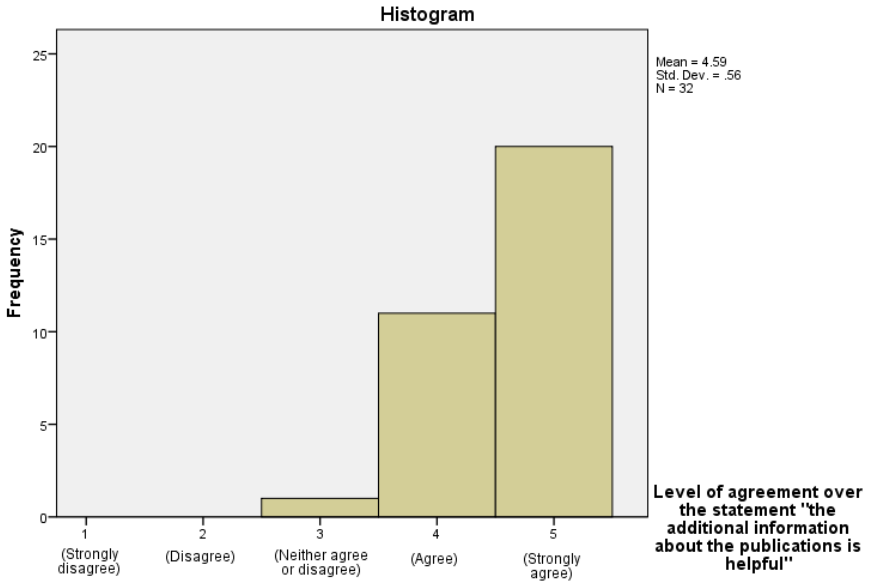


Figure 7.24: A histogram graph showing the participants responses to the statement that "The additional information about the publications is helpful"

7.3.3.5 The explanation for why the piece of information should be trusted provided by the prototype is helpful

A one-sample t-test was conducted to compare the level of agreement of the participants over the helpfulness of the explanation provided by the prototype on why the piece of information should be trusted. The same scale was used here as used in the previous analysis. Results from the t-test data analysis showed that the mean of the participants' agreement over the helpfulness of the explanation (M=4.00, SD=0.76) was statistically significantly larger than the "*neither agree nor disagree*"; $t(31)$=7.42, $p < 0.05$. Therefore, the results suggested that, on average, the participants significantly agree that the explanation for why the piece of Web information should be trusted was helpful. A summary of the one-sample t-test and a histogram can be seen in Table 7.22, Table 7.23, and Figure 7.25.

Table 7.22: One sample statistics for level of agreement over the helpfulness of the explanation

	N	Mean	Std. Deviation	Std. Error Mean
The explanation of the prototype is helpful	32	4.00	0.762	0.135

Table 7.23: One sample test for level of agreement over the helpfulness of the explanation

	Test value = 3 (Neither agree nor disagree)					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
The explanation of the prototype is helpful	7.424	31	0.000	1.000	0.73	1.27

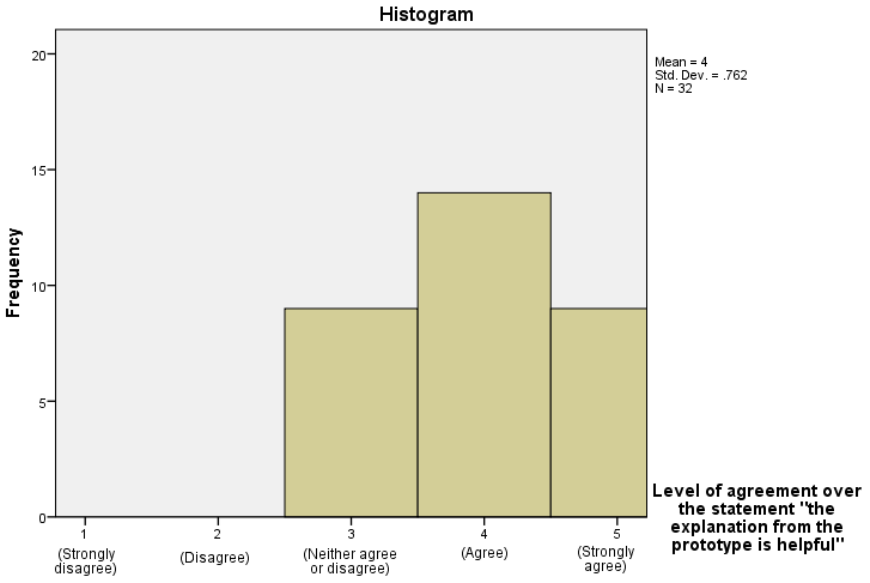


Figure 7.25: A histogram graph showing the participants responses to the statement that "The explanation of the prototype is helpful"

7.3.3.6 The Additional Information about the Authors is Useful to Support the Participants' Evaluation Process

A one-sample t-test was conducted to compare the level of agreement of the participants over the usefulness of the additional information about the authors. The same scale was used here as was used in the previous analysis. Results from the t-test data analysis showed that the mean of the participants' agreement to the usefulness of the additional information about the authors (M=4.22, SD=0.70) was statistically significantly larger than the "*neither agree nor disagree*"; $t(31)$=9.76, $p < 0.05$. Thus, the results suggested that, on average, the participants significantly agree that the additional information about the authors was useful. A summary of the one-sample t-test and a histogram can be seen in Table 7.24, Table 7.25, and Figure 7.26.

Table 7.24: One sample statistics for level of agreement over the usefulness of the additional information of the authors

	N	Mean	Std. Deviation	Std. Error Mean
The additional information about authors is useful	32	4.22	0.706	0.125

Table 7.25: One sample test for level of agreement over the usefulness of the additional information of the authors

	Test value = 3 (Neither agree nor disagree)					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
The additional information about authors is useful	9.760	31	0.000	1.219	0.96	1.47

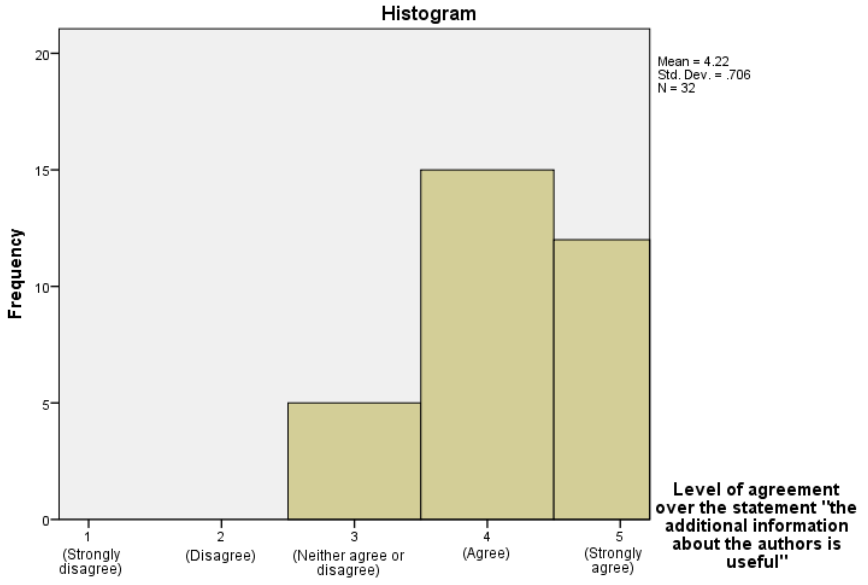


Figure 7.26: A histogram graph showing the participants' responses to the statement that "the additional information of the authors is useful"

7.3.3.7 The Additional Information about the Publication is Useful to Support the Participants' Evaluation Process

A one-sample t-test was conducted to compare the level of agreement of the participants over the usefulness of the additional information about the publication. The same scale is used here as was used in the previous analysis. Results from the t-test data analysis showed that the mean of the participants' agreement to the usefulness of the

additional information about the publication (M=4.50, SD=0.62) was statistically significant larger than the "*neither agree nor disagree*"; $t(31)$=13.64, $p < 0.05$. Therefore, the results suggested that, on average, the participants significantly agree that the additional information about the publication was useful. A summary of the one-sample t-test and a histogram can be seen in Table 7.26, Table 7.27, and Figure 7.27.

Table 7.26: One sample statistics for level of agreement over the usefulness of the additional information of the publication

	N	Mean	Std. Deviation	Std. Error Mean
The additional information about the publications is useful	32	4.50	0.622	0.110

Table 7.27: One sample test for level of agreement over the usefulness of the additional information of the publication

	Test value = 3 (Neither agree nor disagree)					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
The additional information about the publications is useful	13.638	31	0.000	1.500	1.28	1.72

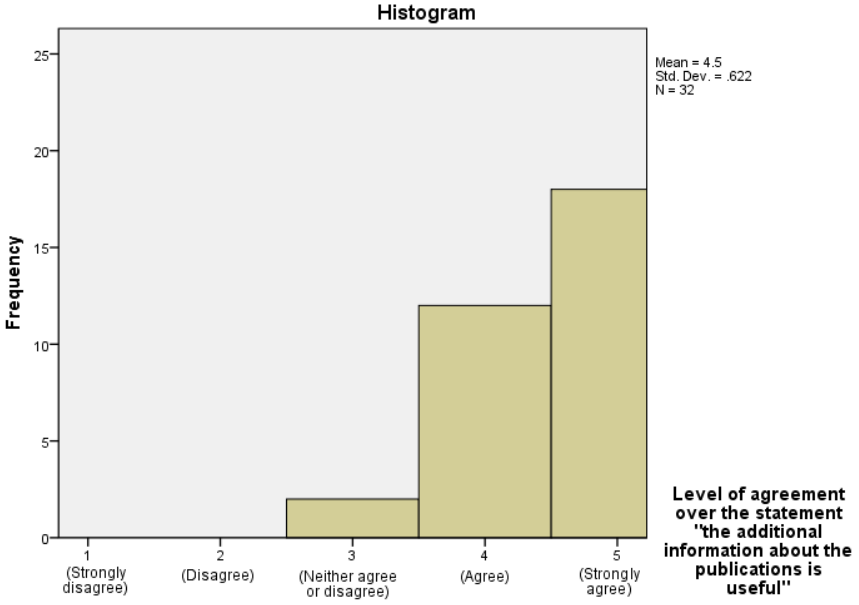


Figure 7.27: A histogram graph showing the participants' responses to the statement that the additional information about the publications is useful

7.3.3.8 The Explanation for Why the Piece of Information should be Trusted provided by the Prototype is Useful

A one-sample t-test was conducted to compare the level of agreement of the participants over the usefulness of the explanation provided by the prototype for why the piece of information should be trusted. The same scale is used here as was used in the previous analysis. Results from the t-test data analysis showed that the mean of the participants' agreement over the usefulness of the explanation (M=3.69, SD=0.90) was statistically significant larger than the "*neither agree nor disagree*"; $t(31)$=4.34, $p < 0.05$. Thus, the results suggested that, on average, the participants significantly agree that the explanation for why the piece of Web information should be trusted was useful. A summary of a one-sample t-test and a histogram can be seen in Table 7.28, Table 7.29, and Figure 7.28.

Table 7.28: One sample statistics for level of agreement over the usefulness of the explanation

	N	Mean	Std. Deviation	Std. Error Mean
The explanation of the prototype is useful	32	3.69	0.896	0.158

Table 7.29: One sample test for level of agreement over the usefulness of the explanation

	Test value = 3 (Neither agree nor disagree)					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
The explanation of the prototype is useful	4.342	31	0.000	0.688	0.36	1.01

7.4 Conclusion

In this chapter, we presented the process of the evaluation of the proposed framework based on the prototype. The proposed framework aims to provide an application framework that can be used to implement tools in order to help Web users to evaluate the trustworthiness of Web information with confidence. Therefore, we designed a study to undertake an experiment with 32 postgraduate student participants. The participants were asked to search for research publications on an assigned topic using the provided prototypes (one of which was a control and the other the experiment prototype). They were then asked to rank the trustworthiness score of the first ten search results and their

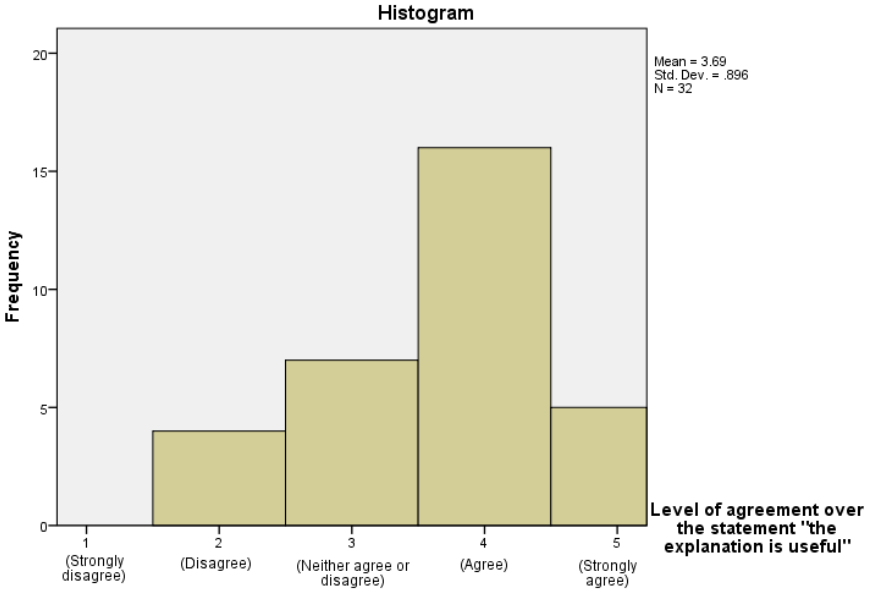


Figure 7.28: A histogram graph showing the participants responses to the statement that "The explanation is useful"

level of confidence over the given score within each prototype. In addition, the participants were asked to select the research publications they would like to use along with the reasons why they did or did not choose each publication. Moreover, they were asked to rate their satisfaction with the supportive information provided by the prototype.

We analysed the collected data using quantitative and qualitative analysis approaches. The results suggested that the participants did experience an increase in their confidence in their ability to evaluate the trustworthiness of Web information, regardless of the topic they were looking for and the familiarity of the topic. In addition, the number of selected pieces of Web information does increase if the participants obtained useful supportive information about that Web information. Moreover, thematic analysis results advised that the participants were likely to improve their judgment of the trustworthiness of Web information when they obtained useful additional information to support their decision. Finally, overall, the participants were satisfied with the supportive information provided to support their evaluation of the trustworthiness of Web information. In addition, the participants significantly agreed that the additional information about the authors, publication and the explanation provided by the prototype were helpful and useful to support their evaluation of the trustworthiness of Web information.

In the next chapter, we summarise the findings of our research, state the contributions, and identify directions for future work.

Chapter 8

Conclusions and Future Work

This thesis discussed the problem and challenge of evaluating the trustworthiness of Web information using Semantic Web technologies to integrate supportive data and to provide this useful information to users. This chapter draws the thesis to a close by discussing the conclusions from this work. This thesis has developed, validated, and implemented the Trustworthiness of Web Information Evaluation framework (TWINE) in a research scenario, specifically in a case study of academic research publications. We summarise the work completed during the doctoral research in section 8.1. In section 8.2, we discuss the contributions of this work. Finally, in section 8.3, we discuss the possible directions of future work that can incorporate and extend the work of this thesis.

8.1 Trustworthiness of Web Information Evaluation Framework (TWINE): A Conclusion

The Trustworthiness of Web Information Evaluation (TWINE) Framework is a conceptual application framework. The TWINE framework was developed in order to help Web users to evaluate the trustworthiness of the information which they are consuming. The framework provides trustworthiness criteria that can be used to collect data to support users' judgments.

This research was completed in five main phases; (1) development of the TWINE framework, (2) development of the trustworthiness criteria used in the framework, (3) validation of the trustworthiness criteria, (4) development of the twine prototype, and (5) evaluation of the TWINE framework based on the prototype.

8.1.1 Development of the TWINE Framework

The potential criteria and components in Chapter 4 were developed to construct the trustworthiness criteria module for TWINE. This supports the TWINE framework in two ways: as a foundation of the criteria to use when evaluating the trustworthiness of Web information and as a basis to determine which data need to be gathered for the framework. The advantage of our trustworthiness criteria is that they can guide the developers to design an application which is suited to the developers' cases. In addition, the trustworthiness criteria were generated based on analysis from across the information quality and web credibility research areas. Therefore, the criteria are not bound to any implementation technique or technology. TWINE has been designed for use in any domain.

Moreover, the TWINE framework consists of two main layers: the presentation- and logic-layers. The criteria and integration module is set in the logic layer in order to work as a backend of the framework; gathering and integrating useful supportive metadata. Its output is a metadata graph. The presentation layer is a front end which interacts with the end users in order to search for information and to interact with and display the results from the system.

8.1.2 Development of the Trustworthiness criteria

Following a literature review of relevant work in information quality and web credibility research in Chapter 2, a set of potential criteria was summarised. Then, we analysed and synthesised these criteria in order to create a list of potential trustworthiness criteria which are used as a foundation of the proposed framework. We generate the criteria and components based on the potential trustworthiness criteria. Finally, we obtained a set of criteria and components which are used as a base component in our framework. These criteria are authority, accuracy, recency, and relevance.

8.1.3 Validation of the Trustworthiness Criteria

To ensure that the proposed framework provides useful supportive information to the users, we validated our proposed trustworthiness criteria, which are the foundation of our framework. The criteria were validated through a panel of experts which consisted of librarians and academic researchers. The questionnaire was designed to find out how useful the criteria and the indicators - which were representative of each criterion - were to the evaluation of the trustworthiness of Web information. In this research, we focused on academic publications, which is a challenging area insofar as there are an extensive range of publications of differing quality and research areas. Specifically, we have addressed the challenges by selecting publications from the ePrints institutional repository

of the University of Southampton as our case study. The ePrints collections of the University of Southampton is representative of an online repository of academic publishing from a variety of research areas. In addition, the pattern of presenting information from ePrints can be extended to other storage systems of online publications in general as it provides basic information about the publication, in the same way as other academic publication websites. However, ePrints also provides other metadata to help Web users in evaluating Web information.

The results of the expert validation suggested that ten indicators were useful for helping to evaluate the trustworthiness of information, which we can discover easily. Moreover, from the qualitative analysis which explored the pattern of the evaluation of the trustworthiness of Web information when some supportive information is missing, it suggested that users needed to search for other supportive information which could help them to evaluate the trustworthiness of Web information instead of the missing one. For example, if the Web information did not provide information about a type of publication, users would try to look for other information that could help them make a judgment on whether or not to trust that information. For instance, they might look for where this information was published or the status of the publication in which this information was published. This supportive information helped them to evaluate the trustworthiness of Web information in place of information regarding the type of publication, which was missing. Therefore, this suggestion supports the idea that providing supportive information affects the user's decision on whether or not to trust the information.

8.1.4 Development of the Twine prototype

We developed a prototype in order to show how TWINE can be used to develop an application for helping users to evaluate the trustworthiness of Web information. In Chapter 6, we presented the process of implementing the prototype TWINE application, which was implemented as a chrome extension. As our case study, we focused on academic publications; specifically the publications that were provided by the ePrints system of the University of Southampton. The data model of the prototype was implemented using named graphs which attached the basic provenance information into the RDF graph.

The prototype worked as a search engine in which users can search for publications of interest (by specifying search terms). The framework returned the publication search results including the supportive information for each publication. This section of work demonstrated how the framework would be used realistically.

In addition, we also tested our prototype by conducting usability testing. The objective of testing was to ensure that the TWINE prototype provided a useful service to its users

and for us to gain a better understanding of the factors that influence the usability of the prototype. The results from the usability test were used to refine the prototype.

8.1.5 Evaluation of the TWINE framework based on the Prototype

The evaluation of the framework was conducted in order to assess our proposed framework through the proposed prototype. In Chapter 7, we designed a study for evaluating the framework based on our research hypothesis which is "A framework (such as TWINE) with properties of gathering, integrating and presenting supportive information using Semantic Web technologies helps users to more effectively evaluate the trustworthiness of Web information." We divided our hypothesis into three sub-hypotheses as follows:

- Using our framework, users increase their level of confidence in their judgment of the trustworthiness of the Web information that they find.
- Using our framework, users increase the number of pieces of trustworthy information which they select to use.
- Using our framework, users are satisfied with the supportive information insofar as it helps the users to evaluate the trustworthiness of Web information.

The evaluation plan was designed to use two prototypes: a control TWINE prototype (Twine 1) and an experiment TWINE prototype (Twine 2). The control prototype was a prototype which only provides basic information about the search results and the experiment prototype provided that same basic information accompanied by additional supportive information. A sample of 32 postgraduate students from the University of Southampton was recruited to take part in the study. They were randomly divided into four sessions in which they were asked to perform a search task for academic publications in two topics using the control prototype and then the experiment prototype for each topic. The order of the assigned topics and prototypes was different in order to reduce systematic bias caused by practice effects or boredom effects. The participants answered questions in a questionnaire regarding their opinions based on their experience with the prototype.

Quantitative and qualitative analysis were used to assess the responses of the participants. The results suggested that the confidence of the participants in their ability to evaluate the trustworthiness of Web information increased, regardless of the topic they were looking for and their familiarity with the topic. In addition, they tend to improve their evaluation of the trustworthiness of Web information when they acquired useful additional information. This was evidenced by the increase in confidence in their ability to evaluate the trustworthiness of Web information and increasing the number of pieces

of information they would select to use in their work when using the experiment prototype over the control prototype. Moreover, the participants improved their evaluation of the trustworthiness of Web information by using objective criteria which was more effective in identifying the trustworthy information. The results showed that the users were satisfied with the supportive information provided to support their evaluation of the trustworthiness of Web information.

8.2 Contributions of the Research

The main contribution of this research is a framework for helping Web users to critically evaluate the trustworthiness of Web information with confidence. In addition, the following related contributions were made:

- The evaluation and selection of the criteria used for supporting the evaluation of the trustworthiness of Web information.
- The calculation of the weights of each criterion for use in calculating the total trustworthiness score of the information. The calculation of such a score allows search results to be ranked based on their relative trustworthiness.
- The integration of metadata using Semantic Web technologies to gather metadata based on our trustworthiness criteria in order to build a data model that can be used in the prototype which provides this information to the users.
- The interpretation of the data model in order to produce an explanation of the trustworthiness of Web information in a human-readable form to users for supporting their decisions.
- A prototype tool, which is implemented based on the proposed framework.

8.3 Future Work

The work presented in this thesis has proven, that the TWINE framework can be used to implement tools for helping Web users to evaluate the trustworthiness of Web information with more confidence. However, there are still some challenging issues that- if addressed- can improve the framework.

The framework is designed for use in the general domain. Each criterion is a conceptual criterion, that should be used for evaluating the trustworthiness of Web information. However, in certain domains there might be the need for special components, which can help to provide more precise information to help users further. In chapter 6 an

example was presented, implementing an application based on the proposed framework. Academic publications from the ePrints institutional repository of the University of Southampton were used as a case study. The components used in this prototype were designed to be specific to the publication domain. Results presented in chapter 7 suggest, that the framework helps the users to be able to evaluate the trustworthiness of Web information with more confidence. However, when the framework is adapted to a new domain, some components will probably have to be added or modified. The framework works well with well-structure data (e.g. RDF), which provides useful metadata to generate the supportive information. In today's Web environment, it may difficult to retrieve good metadata from a Web page. Web scrapy (Scrapy Developers, 2008) can be used to address this issue. Web scraping is a technique to extract unstructured data on a Web page and turn it into structured data (metadata). A study of users' behaviour after using the framework could be performed, in order to learn whether there is any change in their evaluation process after they have used the framework.

Another direction of further research is to adopt a recommendation or a review from a social media for the framework in order to provide more useful supportive information to the users. For example, health Web site, apart from using objective criteria as presented in the TWINE framework, reviews of trustworthy users can be used. Reviews from users of social media, such as twitter or linkedIn can be used to determine who reads or follows the suggestion from a particular web site about certain issue, to increase the confidence of the trustworthiness of the information provided by the corresponding Web sites. Alternatively, trust and privacy accountability (i.e. the use of personal information) can be combined by looking at the data that is used by the people users trust (e.g. TrustLayers (Weitzner and Towvim, 2014)). This information can be integrated with the criteria provided by the framework, in order to make a decision whether or not to trust the information.

Appendix A

An expert evaluation into the importance of the trustworthiness evaluation criteria in assessing the trustworthiness of Web information

The purpose of this questionnaire is to allow you to rate the effect of evaluation factors on the evaluation of the trustworthiness of Web information. It aims to gain a better understanding of the factors that influence the assessment of the trustworthiness of Web information. The results from this will be used to refine our designed framework.

Our framework is designed to help the novice Web user to assess the trustworthiness of information found on the Web. Our focus is the academic domain. There are four sections in this questionnaire:

Section 1: The effect of the appearance of each element in one's confidence of one's ability to evaluate the trustworthiness of Web information.

Section 2: The effect of the absence of each element in perceived the trustworthiness of Web information.

Section 3: The importance of the elements in assessing the trustworthiness of Web information.

Section 4: Additional elements which should be considered.

Section 1: The effect of the appearance of each element in the list below in one's confidence of one's ability to evaluate the trustworthiness of Web information.

The situation we would like you to put yourself in, is that of an academic advisor to new undergraduate students who are starting their studies at the university. They have come to you for advice about things to look for on the Web to indicate that they can trust the information on the Web. Given the following list of items, how useful would you advise your student each item is in order to evaluate the trustworthiness of Web information?

Criteria:

Description	Definition
Not helpful	The presence of this item does not affect one's evaluation of the trustworthiness of Web information.
Somewhat helpful	While the presence of this item helps to build one's confidence in the evaluation of trustworthiness of Web information, its absence does not seriously detract from one's confidence.
Very helpful	This item is needed to be truly confident of one's evaluation of the trustworthiness of the Web information. However, without this item, one can still have ***some*** confidence in one's evaluation.
Critical helpful	This item is essential in order to evaluate the trustworthiness of Web information. Without it one cannot have any confidence in one's evaluation of the trustworthiness of said information.

For each statement below, please rate each item on an effect scale of 1 to 4 by ticking √ in the appropriate box.

Items	Not helpful	Somewhat helpful	Very helpful	Critical helpful
The name of the content creator (e.g. author's name or a name of organization)				
Creator/author's affiliation				
Creator/author's position				
Creator/author's title (e.g. Dr, Professor)				
Creator/author's contact detail				
Publication medium (e.g. book, journal article, blog, facebook, etc.)				
An overview of the content (e.g. title, abstract, etc.)				
The publication date of content				
The last modification date of content				
The physical address of organization				
The brief detail of content creator's experience				
The information of editorial process (e.g. passed peer-review or reviewed from others)				
A list of references				
Number of times that the information has been referenced/cited				

Section 2: The effect of the absence of each element in the list below in perceived the trustworthiness of Web information.

Given the same situation as before, in that you are an advisor to new undergraduate students who are starting their studies at the university. Below is our list of items that people have suggested may be important. This time we ask you, as a supervisor:

1. Given the following list of items, by how much would the absence of each item decrease your confidence in the trustworthiness of Web information?

Criteria:

Description	Definition
No change	The absence of the item does not decrease my confidence in the trustworthiness of the Web information.
Small decrease	The absence of this item will decrease the trust I place in the Web information, but the Web information can still be trustworthy without it.
Large decrease	The absence of this information is damaging to the Web information's trustworthiness. However, other features of the Web information may redeem ***some*** trust.
Destroys confidence	If this information is not present, one cannot place any trust in the Web information.

For each statement below, please rate each item on an effect scale of 1 to 4 by ticking √ in the appropriate box.

Items	No change	Small decrease	Large decrease	Destroys confidence
The name of the content creator (e.g. author's name or a name of organization)				
Creator/author's affiliation				
Creator/author's title (e.g. Dr, Professor)				
The content creator's experience				
Creator/author's contact detail				
Number of times that the information has been referenced in other document				
The editorial process (e.g. peer-review)				
The publication date of content				
The last modification date of content				
The title of content match your needs				
A type of content publication (e.g. book, journal article, personal homepage)				

2. Given the list of items, how would you recommend your student go about increasing their confidence in Web information when the item is not present?

Items	Recommendations
The name of the content creator (e.g. author's name or a name of organization)	
Creator/author's affiliation	
Creator/author's title (e.g. Dr, Professor)	
The content creator's experience	
Creator/author's contact detail	
Number of times that the information has been referenced in other document	
The editorial process (e.g. peer-review)	
The publication date of content	
The last modification date of content	
The title of content match your needs	
A type of content publication (e.g. book, journal article, personal homepage)	

Section 3: The importance of the elements in assessing the trustworthiness of Web information.

Given the same situation as before, in that you are an advisor to new undergraduate students who are starting their studies at the university. This time we ask you to rank the three most important elements from the following list for supporting assesses the trustworthiness of Web information for a novice user.

ID	Elements
A	The name of the content creator
B	Creator/author's affiliation
C	Creator/author's title (e.g. Dr, Professor)
D	The content creator's experience
E	Creator/author's contact detail
F	The resource locator (URL)
G	Number of times that the information has been referenced in other document
H	The editorial process (e.g. peer-review)
I	The publication date of content
J	The last modification date of content
K	The title of content match your needs
L	A type of content publication (e.g. book, journal article, personal homepage)

Please write the IDs of the three most important elements in decreasing orders of importance in the boxes below.

Rank	ID
1	
2	
3	

Section 4: Additional elements which should be considered.

1. What other features of a Web document apart from the list above would lead you to trust it?

---

2. How do you assess whether a Web document is relevant to you?

Appendix B

An invitation email for experts

I would like to invite you to participate in a survey to assess the value of various trustworthiness evaluation criteria in assessing the trustworthiness of Web information. The aim of the study is to confirm that the criteria in question are good for evaluating the trustworthiness of Web information and to gain a better understanding of the factors that influence the assessment of the trustworthiness of Web information. The measurement tool in this study is a questionnaire, which is created and put on the iSurvey system of University of Southampton.

The survey should take approximately 25 minutes or less to complete. You will not be asked for any demographic information.

Participation is voluntary. Therefore, refusal to take part in the study involves no penalty or loss of benefits. Participants can withdraw from the study at any time without penalty.

You can find more details about this study from the participant information sheet attached to this e-mail.

If you agree to take part in this study, please reply to this mail before 23.59 pm on August 13th, 2012. Then, a second e-mail, which will contain a link to the survey, ethics information and a consent form, will be sent to you on August 14th, 2012 and the questionnaire can be completed between August 15th, 2012 and September 15th, 2012.

If you have any further questions about this study or your rights, or if you wish to lodge a complaint or concern, you may contact the Principal Investigator: Jarutas Pattanaphanchai by email (jp11g09@ecs.soton.ac.uk). This study is supervised by Professor Dame Wendy Hall (wh@ecs.soton.ac.uk) and Dr. Kieron O'Hara (km@ecs.soton.ac.uk). You may also contact the Research Governance office (rgoinfo@soton.ac.uk) or Dr Martina Prude, Head of Research Governance (02380 595058, mad4@soton.ac.uk). Ethics Reference Number: 2538

Appendix C

A consent form for a usability test of TWINE prototype

UNIVERSITY OF Southampton

CONSENT FORM

Study title: A usability test of the trustworthiness of web information evaluation (TWINE) prototype

Researcher name: Miss Jarutas Pattanaphanchai
Ethics reference number: 6526

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet [30/05/2013] [Version 2] and have had the opportunity to ask questions about the study.	☐
I agree to take part in this research project and agree for my data to be used for the purpose of this study	☐
I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected	☐
I am happy to be contacted regarding other unspecified research projects. I therefore consent to the University retaining my personal details on a database, kept separately from the research data detailed above. The 'validity' of my consent is conditional upon the University complying with the Data Protection Act and I understand that I can request my details be removed from this database at any time.	☐

Data Protection
I understand that information collected about me during my participation in this study will be stored on a password protected computer and that this information will only be used for the purpose of this study. All files containing any personal data will be made anonymous.

Name of participant (print name).....

Signature of participant.....

Date.....

Appendix D

Reaction test of the usability test study record

Participant 1

Question	Answer
What do you think this page is used for?	Participant1 thinks it is some kind of search engine from its looks. Moreover, Participant1 thinks it is a specific search engine in which the user can define the scope of their search and it can filter information based on the scope.
What is the first thing you notice?	Participant1 spots the logo of the prototype first but the participant does not understand what TWINE at the logo is.
What do you think you would click on first?	1. Home 2. Tab on the input field or some other thing on the page 3. Look at the options 4. Do not pay attention on how many results per page
Are there anything you like or don't like?	Like - Logo - Design of prototype Dislike - The input layout is not clear how the system works and the input area is too narrow - It has too many options to choose - The layout is not simple enough

Participant 2

Question	Answer
What do you think this page is used for?	Participant2 thinks this prototype is a custom search engine in which you can select the categories of the information you want.
What is the first thing you notice?	Participant notices straight away that the prototype is search engine
What do you think you would click on first?	1. Fill some keyword in the text field 2. Click on the search to see what the results is
Are there anything you like or don't like?	Like - simple interface Dislike - The input layout is too narrow, it seems something missing from the page - Input option should be in one line rather than two lines because the participant is familiar with google search engine and quite likes the one line option

Participant 3

Question	Answer
What do you think this page is used for?	Participant3 thinks it is some kind of search engine which can define a scope of interest and limit of search results display per page. Participant mentioned that it can spot easily that this is a search engine.
What is the first thing you notice?	Logo of the prototype
What do you think you would click on first?	The text field of search terms because the cursor is set on it
Are there anything you like or don't like?	Like - Layout: it is clean and simple Dislike - Navigation bar looks too solid. Participant does not like its shape (rectangular) because it does not match with the style of logo - Colour of home and about link are too similar with the colour of navigation bar. This makes link too difficult to find and click on it. Participant thinks change a colour of link will help to make its clear (suggested white colour)

Participant 4

Question	Answer
What do you think this page is used for?	Participant thinks it is search tool that can search from a given scope of interest (for example, research paper, news, and business).
What is the first thing you notice?	Logo and title of the logo. This tells the user what this tool is.
What do you think you would click on first?	Scope of interest, because participant would like to know what kind of information can be searched.
Are there anything you like or don't like?	Like - Design which is very simple and straight forward Dislike - Nothing in particular. Everything looks fine.

Participant 5

Question	Answer
What do you think this page is used for?	Participant thinks it is a tool that can help to search for something.
What is the first thing you notice?	The blinking cursor in the text field of the search terms.
What do you think you would click on first?	The text field
Are there anything you like or don't like?	Like - Design is straight forward and simple - Colour of page is nice and easy to look Dislike - Participant prefers to have a separate line of options. One line for search terms, another line for the scope for interest and another line of the number of results to display per page because it can indicate the sequence in which data should be input in the prototype.

Appendix E

Key task test of the usability test study record

Participant1

Task scenario	You are interested in research on privacy topic and you need to select papers in this area to reference in your report. Your task is to search research papers in privacy topics and select papers that most likely to be trustworthy and relevant to your research interest.
Observation	1. Participant types "privacy" into the text field provided 2. Then, chooses the scope of interest which is "research paper" 3. Selects to display 20 results per page 4. When the prototype returns the results, the participant looks from the top to the bottom of the page - Read the title of each paper - Click on the information icon because the participant thought it would provide an explanation of the paper - Select the paper that has received a high suggested score - Click on the link of the title to see the real data of the paper - Back to the search results page of prototype 5. Might stop at the top two results but if a paper is not interesting enough, the participant will carry on to the next results based on the suggested score. Note: 1. The participant did not use any supportive information provided on the system. May be participant did not see the supportive information or it is not obvious enough. 2. The participant mentioned the order of the search results affected the participant's decision 3. The participant does not pay attention on publication status but focuses on citation and publication type. 4. The participant suggested showing less of the abstract may help to improve the results display by making the layout looks cleaner. 5. If the prototype can highlight the search terms in the abstract, it can help the users spot out the relevance of a piece of information to their interest easily.

Participant2

Task scenario	You are interested in research on privacy topic and you need to select papers in this area to reference in your report. Your task is to search research papers in privacy topics and select papers that most likely to be trustworthy and relevant to your research interest.
Observation	1. Participant types the keyword 2. Participant is not familiar with the topic so the participant would like to know the definition of the topic by use basic keyword 3. Participant wants to find research papers so the participant selects the scope of interest as "research paper" 4. After the search results return, the participant looks at the total number of search results first 5. "Number of citations" is the first piece of supporting information to which the participant pays attention 6. Then, the participant looks at the publication date 7. Type of publication is the next piece of information the participant looks at 8. If the participant could not find a paper that the participant wants, they would change the keywords Note: - The icon of author's detail is not clear enough - If the participant can see that it has some supportive information, "affiliation" is the first factor that participant will pay intention to - The publication's detail is presented in a way that is easy to read and notice - May be sort the publications based on certain participant's focus such as by date of publication, citation number, etc.

Participant3

Task scenario	You are interested in research on privacy topic and you need to select papers in this area to reference in your report. Your task is to search research papers in privacy topics and select papers that most likely to be trustworthy and relevant to your research interest.
Observation	1. Participant types "privacy" in the search terms text field 2. Checks the scope of interest (what options it has) 3. Clicks the list button to display scope of interest list 4. Selects the research papers because participant need to find papers to use in participant's report 5. Checks what is the maximum number of results that can be displayed per page 6. Notices that the maximum number is 20 results per page 7. Decide to use the default value which is 10 results per page 8. Waits for the search results 9. After seeing the results, the participant thinks that the search results are sorted by suggested trustworthiness 10. Participant starts to check the detail of search results 11. Participant spots the author's detail icon and is curious as to what this icon is, then participant clicks it - The author's details are displayed on the page - Participant is surprised that the details of author that are shown 12. Participant notices the trustworthiness suggestion score and wonder where this score comes from. Therefore, the participant clicks the information icon behind the score bar - Participant reads through the explanation details 13. Then, participant notices the link to the original file of the

	paper so the participant clicks it - The original web page is shown - Participant scans the page and sees it is an original page of a paper - Clicks back 14. After checking a couple of items of the provided supportive information, the participant realizes that this tool provides the brief details of each of the papers. If the participant wants more detail, the participant can click on publication link provided 15. Then, the participant is interested in timeline scale, which, at first glance, the participant does not understand what it is - The participant tries to click on an icon on the timeline - It shows the details of date of publication - Thus, the participant knows this supportive information is a timeline scale of this paper 16. Checks on the other pages 17. Scans through each page 18. Finally, the participant selects the first order paper on first page because it has the highest score of suggested trustworthiness

Participant4

Task scenario	You are interested in research on privacy topic and you need to select papers in this area to reference in your report. Your task is to search research papers in privacy topics and select papers that most likely to be trustworthy and relevant to your research interest.
Observation	1. Participant types "privacy" into the search terms text field 2. Then, the participant chooses the scope of interest as "research paper" 3. The participant looks at the number of results per page option. However, the participant does not bother much, the participant just uses the default value of the prototype 4. After the search results returns, participant checks a number of the total results 5. Then, participant scrolls down to scan all of the results on the first page 6. Participant spots on the timeline and score bar at first glance. These two visual displays attract the participant's attention but the participant does not read the detail of these two options yet. 7. The next item that the participant pays attention is an abstract of a paper 8. The participant sees the trustworthiness score bar but it is not quite clear what the trustworthiness score bar means. Then, participant sees there is an information icon after the score bar. Therefore, the participant clicks on it. The information icon displays the explanation of the score, the participant likes this option because it helps participant to understand what the trustworthiness score is. 9. Then, participant looks for the type of publication, the date of the publication, the status of publication, and the number of citations of publication respectively 10. After scan through publications' supportive details, the participant notices an author badge icon. However, the participant is not sure what it is. Thus, participant clicks it.

	- The details of author is displayed on the page, the participant then again is impressed on the detail displayed - The participant thinks it is useful information in that the participant can see the expertise of the author based on their interest, and projects list. 11. The timeline is the next thing that the participant pays attention to. Similarly with the score bar, it draws the participant's attention and the participant thinks it is some kind of timeline of something but it is not clear. Therefore, the participant clicks the icon on the timeline, the icon displays information about the date of publication and the date of publication was uploaded 12. Finally, the participant looks at the suggested trustworthiness score bar. However, the participant is not clear on whether this score is a percentage of relevance of the paper to the search terms or a percentage of a paper's trustworthiness 13. Participant reads the abstract of a paper again before selecting the first paper on the first page. Note: 1. The participant is not clear about some the provided information at the result page. Therefore, it would be helpful to have an about page to explain how to read or use the results. 2. In addition, participant could not see the about menu clearly when participant saw it, it then suddenly attracts the participant's attention. 3. The participant is not familiar with the topic of the task thus the participant needed to read the abstract again before they can make a decision on whether to select a paper or not. - The participant is not clear whether trust in this situation means privacy, copy right or the attitude toward a paper

Participant5

Task scenario	You are interested in research on privacy topic and you need to select papers in this area to reference in your report. Your task is to search research papers in privacy topics and select papers that most likely to be trustworthy and relevant to your research interest.
Observation	1. Participant types search terms "privacy" and clicks search button 2. After the results return, the participant scans through the whole of the first page to have a look at what results have been received 3. Then, the participant reads the title of each paper on the page 4. After that, the participant clicks on the next page to check the other search results. 5. On each result page, the participant will scan through all of the results in each page by reading the title - While participant checks the results on each page, the participant is confused as to which page the participant is on, so the participant checks by looking at the page number and can see that the current page is displayed by red. This makes the participant know which page participant is working on. 6. After that, participant navigates back to the first page of the search results and then starts reading through the abstract (as shown on the page) of each paper 7. Finally, the participant selects the second paper on the first page because the title and abstract mention about regulation which the participant thinks it is important to get started about researching on privacy. Note: 1. The participant prefers to have a large area for displaying the abstract because an abstract will indicate the relevance of a paper to the participant's interest. The participant does not bother to click on expand icon if the shown abstract is not interesting or does not attract the participant's attention. 2. The author's name should have a different colour to make stand out.

Appendix F

A consent form for an evaluation of the TWINE framework

UNIVERSITY OF Southampton

CONSENT FORM (version 3)

Study title: An Evaluation of the Trustworthiness of Web Information Evaluation (TWINE) Framework

Researcher name: Miss Jarutas Pattanaphanchai

Ethics reference number: ERGO/FoPSE/6800

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (25-07-2013/ Version 3) and have had the opportunity to ask questions about the study.	☐
I agree to take part in this research project and agree for my data to be used for the purpose of this study	☐
I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected	☐

Data Protection

I understand that information collected during my participation in this study will be stored on a password protected computer and that this information will only be used for the purpose of this study.

Name of participant (print name).....

Signature of participant.....

Date.....

Appendix G

An example of instruction sheet for a participant

An Evaluation of the Trustworthiness of Web Information Evaluation (TWINE) Framework

Instruction Sheet

This study consists of 2 parts in which the part 1 has 4 tests that you will be asked to perform and the part 2 you will be asked to answer the questions. In each test of part 1, you will be given a scenario which you must search for academic publications in given topic. **Before starting the tasks, please read page 1 of the questionnaire. Then, turn on to page 2 of the questionnaire and start do the task in part 1.**

PART 1: Test 1

Your task in this test is to search research papers in **privacy** topics.

1. Use the questionnaire, please state the search topic in **question 1 of section 1**
2. Use the questionnaire, please indicate how familiar you are with the search topic in **question 2 of section 1**
3. Then, use the machine for starting the search task
4. Open a google chrome browser
5. Click the **TWINEApp_1** icon 

on the browser
6. On the main page of the TWINE application
   5.1 Fill the word "**privacy**" in the search terms field
   5.2 Select the scope of interest as "Research Paper"
   5.3 Select the number of results per page as 10 results
   5.4 Click the search button

6 From the ten search results on the first page, please answer **question 1-4 of section 2** in the questionnaire

7 Please take a break for **2 minutes** before starting the next test

PART 1: Test 2

Your task in this test is to search research papers in **privacy** topics.

1. Use the questionnaire, please state the search topic in **question 1 of section 1**
2. Use the questionnaire, please indicate how familiar you are with the search topic in **question 2 of section 1**
3. Then, use the machine for starting the search task
4. Open a new google chrome browser
5. Click the **TWINEApp_2** icon 

on the browser

6. On the main page of the TWINE application
   5.1 Fill the word "**privacy**" in the search terms field
   5.2 Select the scope of interest as "Research Paper"
   5.3 Select the number of results per page as 10 results
   5.4 Click the search button
6 From the ten search results on the first page, please answer **question 1-4 of section 2** in the questionnaire
7 Please take a break for **5 minutes** before starting the next test

PART 1: Test 3

Your task in this test is to search research papers in **semantic web** topics.

1. Use the questionnaire, please state the search topic in **question 1 of section 1**
2. Use the questionnaire, please indicate how familiar you are with the search topic in **question 2 of section 1**
3. Then, use the machine for starting the search task
4. Open a new google chrome browser
5. Click the **TWINEApp_1** icon  on the browser
6. On the main page of the TWINE application
   5.1 Fill the word "**semantic web**" in the search terms field
   5.2 Select the scope of interest as "Research Paper"
   5.3 Select the number of results per page as 10 results
   5.4 Click the search button
7. From the ten search results on the first page, please answer **question 1-4 of section 2** in the questionnaire
8. Please take a break for **2 minutes** before starting the next test

PART 1: Test 4

Your task in this test is to search research papers in **semantic web** topics.

1. Using the questionnaire, please state the search topic in **question 1 of section 1**
2. Using the questionnaire, please indicate how familiar you are with the search topic in **question 2 of section 1**
3. Then, use the machine for starting the search task
4. Open a new google chrome browser
5. Click the **TWINEApp_2** icon  on the browser
6. On the main page of the TWINE application
   5.1 Fill the word "**semantic web**" in the search terms field

5.2 Select the scope of interest as “Research Paper”
5.3 Select the number of results per page as 10 results
5.4 Click the search button

7. From the ten search results on the first page, please answer **question 1-4 of section 2** in the questionnaire
8. Please take a break for **2-3 minutes** before starting the next test

PART 2

After finishing the tasks in part 1, please answer the question 1-5 of part 2

THANK YOU

Appendix H

The TWINE evaluation questionnaire

The purpose of this questionnaire is to allow the participants to rate the trustworthiness of Web information they are consuming (in this case study is the research publications), their confidence of a given score, their satisfactions of the proposed framework. In addition, the participants will be asked to state the research publications they select to use in their report.

The questionnaire aims to validate our proposed framework which is designed to help users assess the trustworthiness of Web information more critically based on the provided supportive information. There are two parts in this questionnaire:

Part 1 consists of 4 tests in which each test will ask the participant to state the topic they are searching for and how familiar of the participant with the topic. In addition, each test will ask the participants to rate the trustworthiness of Web information they are considering and how much confidence on the given score. Moreover, they will be asked to answer which papers they would select and reasons for selecting them or not selecting any of them.

Part 2 will ask the overall opinion of the study

How to complete this questionnaire

For each question please **circle** your answer that can indicate your opinion. Sometimes you may need to tick one box or may be asked to write in your answer. If you change your mind about one of your answers, or you have selected the wrong box by mistake, simply shade in the old box completely and then put a circle in the box that you want, as shown in the example below.

1. Please indicate how familiar you are with the topic

Not at all familiar	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar
1	●	3	(4)	5

2. Think about your level of confidence in your ability to evaluate the trustworthiness of Web information of the assigned topic. Please, indicate the degree of confidence you have in the given score of trustworthiness of information using the following scale:

No confidence					Moderate Confidence					Complete Confidence
0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
				●			✓			

PART 1: Test 1

Section 1: PLEASE TELL US A LITTLE ABOUT THE TOPIC

1. Search topic: ……

2. Please indicate how familiar you are with the topic.

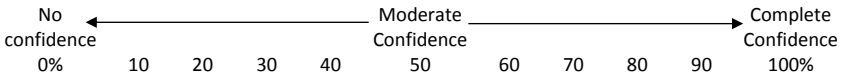
Not at all familiar	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar
1	2	3	4	5

Section2: TWINE 1

1. On a scale of being very untrustworthy to being very trustworthy, please rate the trustworthiness of the following search results

Search results	Very untrustworthy	Mostly untrustworthy	Somewhat untrustworthy	Neither untrustworthy or trustworthy	Somewhat trustworthy	Mostly trustworthy	Very trustworthy
The 1st							
The 2nd							
The 3rd							
The 4th							
The 5th							
The 6th							
The 7th							
The 8th							
The 9th							
The 10th							

2. Think about your level of confidence in your ability to evaluate the trustworthiness of Web information of the following search results in the assigned topic. Indicate the degree of confidence you have in the given score in the question 1 using the following scale:



Search results	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
The 1st											
The 2nd											
The 3rd											
The 4th											
The 5th											
The 6th											
The 7th											
The 8th											
The 9th											
The 10th											

3. Please tell us about your decision on **<u>whether or not to select</u>** any of the publications from the ten search results on the first page by ticking in the "**Select**" boxes (select as many as apply) and please give us brief reasons of selecting them**. <u>If you would not select any publications to this question, please answer question 4. Otherwise, please skip question 4 and continue from the later questions.</u>**

Search results	Select	Reasons of selecting
The 1st		
The 2nd		
The 3rd		
The 4th		

Search results	Select	Reasons of selecting
The 5th		
The 6th		
The 7th		
The 8th		
The 9th		
The 10th		

4. Briefly, could you tell us why you would not select any papers to use in your report?

……

……

……

……

……

……

……

……

……

……

……

……

……

……

PART 1: Test 2

Section1: PLEASE TELL US A LITTLE ABOUT THE TOPIC

1. Search topic: ……

2. Please indicate how familiar you are with the topic.

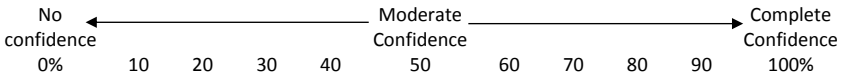
Not at all familiar	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar
1	2	3	4	5

Section 2: TWINE 2

1. On a scale of being very untrustworthy to being very trustworthy, please rate the trustworthiness of the following search results.

Search results	Very untrustworthy	Mostly untrustworthy	Somewhat untrustworthy	Neither untrustworthy or trustworthy	Somewhat trustworthy	Mostly trustworthy	Very trustworthy
The 1st							
The 2nd							
The 3rd							
The 4th							
The 5th							
The 6th							
The 7th							
The 8th							
The 9th							
The 10th							

2. Think about your level of confidence in your ability to evaluate the trustworthiness of Web information of the following search results in the assigned topic. Indicate the degree of confidence you have in the given score in the question 1 using the following scale:



Search results	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
The 1st											
The 2nd											
The 3rd											
The 4th											
The 5th											
The 6th											
The 7th											
The 8th											
The 9th											
The 10th											

3. Please tell us about your decision on **<u>whether or not to select</u>** any of the publications from the ten search results on the first page by ticking in the "**<u>Select</u>**" boxes (select as many as apply) and please give us brief reasons of selecting them. **<u>If you would not select any publications to this question, please answer question 4. Otherwise, please skip question 4 and continue from the later questions.</u>**

Search results	Select	Reasons of selecting
The 1st		
The 2nd		
The 3rd		
The 4th		

Search results	Select	Reasons of selecting
The 5th		
The 6th		
The 7th		
The 8th		
The 9th		
The 10th		

4. Briefly, could you tell us why you would not select any papers to use in your report?

……

……

……

……

……

……

……

……

……

……

……

……

……

……

PART 1: Test 3

Section 1: PLEASE TELL US A LITTLE ABOUT THE TOPIC

1. Search topic: ……

2. Please indicate how familiar you are with the topic.

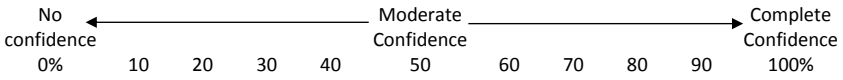
Not at all familiar	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar
1	2	3	4	5

Section 2: TWINE 1

1. On a scale of being very untrustworthy to being very trustworthy, please rate the trustworthiness of the following search results.

Search results	Very untrustworthy	Mostly untrustworthy	Somewhat untrustworthy	Neither untrustworthy or trustworthy	Somewhat trustworthy	Mostly trustworthy	Very trustworthy
The 1st							
The 2nd							
The 3rd							
The 4th							
The 5th							
The 6th							
The 7th							
The 8th							
The 9th							
The 10th							

2. Think about your level of confidence in your ability to evaluate the trustworthiness of Web information of the following search results in the assigned topic. Indicate the degree of confidence you have in the given score in the question 1 using the following scale:



Search results	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
The 1st											
The 2nd											
The 3rd											
The 4th											
The 5th											
The 6th											
The 7th											
The 8th											
The 9th											
The 10th											

3. Please tell us about your decision on **<u>whether or not to select</u>** any of the publications from the ten search results on the first page by ticking in the "**<u>Select</u>**" boxes (select as many as apply) and please give us brief reasons of selecting them. **<u>If you would not select any publications to this question, please answer question 4. Otherwise, please skip question 4 and continue from the later questions.</u>**

Search results	Select	Reasons of selecting
The 1st		
The 2nd		
The 3rd		
The 4th		

Search results	Select	Reasons of selecting
The 5th		
The 6th		
The 7th		
The 8th		
The 9th		
The 10th		

4. Briefly, could you tell us why you would not select any papers to use in your report?

……

……

……

……

……

……

……

……

……

……

……

……

……

……

PART 1: Test 4

Section 1: PLEASE TELL US A LITTLE ABOUT THE TOPIC

1. Search topic: ……

2. Please indicate how familiar you are with the topic.

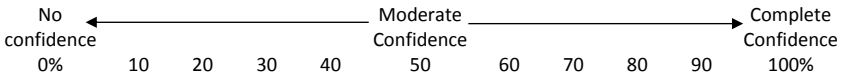
Not at all familiar	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar
1	2	3	4	5

Section2: TWINE 2

1. On a scale of being very untrustworthy to being very trustworthy, please rate the trustworthiness of the following search results.

Search results	Very untrustworthy	Mostly untrustworthy	Somewhat untrustworthy	Neither untrustworthy or trustworthy	Somewhat trustworthy	Mostly trustworthy	Very trustworthy
The 1st							
The 2nd							
The 3rd							
The 4th							
The 5th							
The 6th							
The 7th							
The 8th							
The 9th							
The 10th							

2. Think about your level of confidence in your ability to evaluate the trustworthiness of Web information of the following search results in the assigned topic. Indicate the degree of confidence you have in the given score in the question 1 using the following scale:



Search results	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
The 1st											
The 2nd											
The 3rd											
The 4th											
The 5th											
The 6th											
The 7th											
The 8th											
The 9th											
The 10th											

3. Please tell us about your decision on **<u>whether or not to select</u>** any of the publications from the ten search results on the first page by ticking in the "**<u>Select</u>**" boxes (select as many as apply) and please give us brief reasons of selecting them. **<u>If you would not select any publications to this question, please answer question 4. Otherwise, please skip question 4 and continue from the later questions.</u>**

Search results	Select	Reasons of selecting
The 1st		
The 2nd		
The 3rd		
The 4th		

Search results	Select	Reasons of selecting
The 5th		
The 6th		
The 7th		
The 8th		
The 9th		
The 10th		

4. Briefly, could you tell us why you would not select any papers to use in your report?

……

……

……

……

……

……

……

……

……

……

……

……

……

……

PART 2

PLEASE TELL US ABOUT YOUR OPINION

1. From the task that you did in part 1, please tell us, what are items that might affect to your decision whether to select the papers to use in your report. Please do this by ticking the appropriate boxes below (select as many as apply)?

   ☐ Author's name
   ☐ Author's position (e.g. research staff, lecturer, academic staff etc.)
   ☐ Author's qualification
   ☐ Author's workplace
   ☐ Author's past projects
   ☐ Author's current projects
   ☐ Type of publication (e.g. conference, book, technical report etc.)
   ☐ Status of publication (e.g. peer-reviewed, published etc.)
   ☐ The publication date of content
   ☐ The date of file uploaded
   ☐ The date of uploaded file has been modified
   ☐ A number of times that paper has been referenced
   ☐ Title of paper match with search terms
   ☐ Abstract has search terms appear in it

2. Please indicate whether you agree or disagree with the following statement about the evaluation the trustworthiness of Web information you just did in part 1:

statement	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
I find the additional information about the authors is **helpful** to support my evaluation of the trustworthiness of Web information					
I find the additional information about the publication is **helpful** to support my evaluation of the trustworthiness of Web information					
I find the explanation why the piece of information should be trust is **helpful**					

statement	Strongly disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
I find the additional information about the authors is **useful** to increase my confidence on the trustworthiness of Web information					
I find the additional information about the publication is **useful** to increase my confidence on the trustworthiness of Web information					
I find the explanation why the piece of information should be trust is **useful**					
The provided supportive information is easily to read					
The provided supportive information is easily to understand					

3. Please rate your overall satisfaction you are with the provided additional information to support your evaluation of the trustworthiness of Web information you did in part 1:
   - ☐ Not at all satisfied
   - ☐ Slightly satisfied
   - ☐ Moderately satisfied
   - ☐ Very satisfied
   - ☐ Extremely satisfied

4. Please rate how influence of the provided supportive information to your judgment of the trustworthiness of Web information
   - ☐ Not at all influential
   - ☐ Slightly influential
   - ☐ Somewhat influential
   - ☐ Very influential
   - ☐ Extremely influential

5. Please rate how influence of the search results ordered to your judgment of selecting the information
   - ☐ Not at all influential
   - ☐ Slightly influential
   - ☐ Somewhat influential
   - ☐ Very influential
   - ☐ Extremely influential

Appendix I

Table of critical values for the F-Distribution (0.05 level)

df		df for (v_h)											
(v_e)	$1-\alpha$	1	2	3	4	5	6	7	8	9	10	11	12
22	.90	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.88	1.86
	.95	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	22.3
	.99	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12
24	.90	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.85	1.83
	.95	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.21	2.18
	.99	7.82	561	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03
26	.90	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.84	1.81
	.95	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
	.99	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	3.02	2.96
28	.90	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	184	1.81	1.79
	.95	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12
	.99	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.96	2.90
30	.90	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.79	1.77
	.95	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09
	.99	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84
40	.90	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.73	1.71
	.95	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00
	.99	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66
60	.90	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.68	1.66
	.95	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
	.99	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50
120	.90	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.62	1.60
	.95	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.87	1.83
	.99	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.40	2.34
200	.90	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.60	1.57
	.95	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80
	.99	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.27
∞	.90	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.57	1.55
	.95	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75
	.99	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18

df		df for (v_h)													df
(v_e)	$1-\alpha$	15	20	24	30	40	50	60	100	120	200	500	$1-\alpha$	(v_e)
22	.90	1.81	1.76	1.73	1.70	1.67	1.65	1.64	0.61	1.60	1.59	1.58	1.57	.90	22
	.95	2.15	2.07	2.03	1.98	1.94	1.91	1.89	1.85	1.84	1.82	1.80	1.78	.95	
	.99	2.98	2.83	2.75	2.67	2.58	2.53	2.50	2.42	2.40	2.36	2.33	2.31	.99	
24	.90	1.78	1.73	1.70	1.67	1.64	1.62	1.61	1.58	1.57	1.56	1.54	1.53	.90	24
	.95	2.11	2.03	1.98	1.94	1.89	1.86	1.84	1.80	1.79	1.77	1.75	1.73	.95	
	.99	2.89	2.74	2.66	2.58	2.49	2.44	2.40	2.33	2.31	2.27	2.24	2.21	.99	
26	.90	1.76	1.71	1.68	1.65	1.61	1.59	1.58	1.55	1.54	1.53	1.51	1.50	.90	26
	.95	2.09	1.99	1.95	1.90	1.85	1.82	1.80	1.76	1.75	1.73	1.71	1.69	.95	
	.99	2.81	2.66	2.58	2.50	2.42	2.36	2.33	2.25	2.23	2.19	2.16	2.13	.99	
28	.90	1.74	1.69	1.66	1.63	1.59	1.57	1.56	1.53	1.52	1.50	1.49	1.48	.90	28
	.95	2.04	1.69	1.91	1.87	1.82	1.79	1.77	1.73	1.71	1.69	1.67	1.65	.95	
	.99	2.75	2.60	2.52	2.44	2.35	2.30	2.26	2.19	2.17	2.13	2.09	2.06	.99	
30	.90	1.72	1.67	1.64	1.61	1.57	1.55	1.54	1.51	1.50	1.48	1.47	1.46	.90	30
	.95	2.01	1.93	1.89	1.84	1.79	1.76	1.74	1.70	1.68	1.66	1.64	1.62	.95	
	.99	2.70	2.35	2.47	2.39	2.30	2.55	2.21	2.13	2.11	2.07	2.03	2.01	.99	
40	.90	1.66	1.61	1.57	1.54	1.51	1.48	1.47	1.43	1.42	1.41	1.39	1.38	.90	40
	.95	1.92	1.84	1.79	1.74	1.69	1.66	1.64	1.59	1.58	1.55	1.53	1.51	.95	
	.99	2.52	2.37	2.29	2.20	2.11	2.06	2.02	1.94	1.92	1.87	1.83	1.80	.99	
60	.90	1.60	1.54	1.51	1.48	1.44	1.41	1.40	1.36	1.35	1.33	1.31	1.29	.90	60
	.95	1.84	1.75	1.70	1.65	1.59	1.56	1.53	1.48	1.47	1.44	1.41	1.39	.95	
	.99	2.35	2.20	2.12	2.03	1.94	1.88	1.84	1.75	1.73	1.68	1.63	1.60	.99	
120	.90	1.55	1.48	1.45	1.41	1.37	1.34	1.32	1.27	1.26	1.24	1.21	1.19	.90	120
	.95	1.75	1.66	1.61	1.55	1.50	1.46	1.43	1.37	1.35	1.32	1.28	1.25	.95	
	.99	2.19	2.03	1.95	1.86	1.76	1.70	1.66	1.56	1.53	1.48	1.42	1.38	.99	
200	.90	1.52	1.46	1.42	1.38	1.34	1.31	1.28	1.24	1.22	1.20	1.17	1.14	.90	200
	.95	1.72	1.62	1.57	1.52	1.46	1.41	1.39	1.32	1.29	1.26	1.22	1.19	.95	
	.99	2.13	1.97	1.89	1.79	1.69	1.63	1.58	1.48	1.44	1.39	1.33	1.28	.99	
∞	.90	1.49	1.42	1.38	1.34	1.30	1.26	1.24	1.18	1.17	1.13	1.08	1.00	.90	∞
	.95	1.67	1.57	1.52	1.46	1.39	1.35	1.32	1.24	1.22	1.17	1.11	1.00	.95	
	.99	2.04	1.88	1.79	1.70	1.59	1.52	1.47	1.36	1.32	1.25	1.15	1.00	.99	

Original source: Ahrens, H. (1968), Pearson, E. S., and H. O. Hartley (Edit.): Biometrika Tables for Statisticians Vol. I, 3. Auflage. University Press, Cambridge 1966. XVI + 264 S. Preis 35 s. Biom. J., 10: 226.

References

M. Alhojailan. Thematic Analysis: A critical review of its process and evaluation. In *WEI International European Academic Conference*, pages 8–21, Zagreb, Croatia, 2012.

A. Anastasi and S. Urbina. *Psychological testing.* Upper Saddle River, NJ: Prentice Hall, 7th edition, 1997.

Associates Princeton Survey Research. A matter of trust: what users want from web sites. `http://consumersunion.org/research/a-matter-of-trust-what-users-want-from-web-sites/`, 2002. [online; accessed: 28-10-2013].

D. Beckett. N-Triples: A line-based syntax for an RDF graph. `http://www.w3.org/TR/2013/NOTE-n-triples-20130409/`, 2013. [online; accessed: 28-10-2013].

D. Beckett and T. Berners-Lee. Turtle-Terse RDF Triple Language. `http://www.w3.org/TeamSubmission/turtle/`, 2011. [online; accessed: 14-11-2013].

M. Bernard, R. Baker, and M. Fernandez. Paging vs. scrolling: Examining ways to present search results. In *The human Factors and Ergonomics Society Annul Meeting*, pages 1296–1299, 2002.

T. Berners-Lee. Semantic Web - XML2000. `http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html`, 2000. [online; accessed: 28-10-2013].

T. Berners-Lee. Artificial Intelligence and the Semantic Web. `http://www.w3.org/2006/Talks/0718-aaai-tbl`, 2006a. [online; accessed: 06-05-2014].

T. Berners-Lee. Design issues: Linked data. `http://www.w3.org/DesignIssues/LinkedData.html`, 2006b. [online; accessed: 06-05-2014].

T. Berners-Lee, R. Cailliau, A. Luotonen, H. F. Nielsen, and A. Secret. The world-wide web. *Communications of the ACM*, 37(8):76–82, 1994.

T. Berners-Lee and D. Connolly. Hypertext Markup Language. `http://www.rfc-editor.org/rfc/rfc1866.txt`, 1995. [online; accessed: 04-12-2013].

T. Berners-Lee, R. Fielding, and H. Frystyk. Hypertext Transfer Protocol – HTTP/1.0. `http://www.rfc-editor.org/rfc/rfc1945.txt`, 1996. [online; accessed: 04-12-2013].

T. Berners-Lee, R. Fielding, and L. Masinter. RFC2396 - Uniform Resource Identifiers (URI): Generic Syntax. `http://www.ietf.org/rfc/rfc2396.txt`, 1998. [online; accessed: 25-10-2013].

T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific american*, 284 (5):28–37, 2001.

J. Bettels and A. Bishop. Unicode: A universal character code. *Digital Technical Journal*, 5(3):21–31, 1993.

C. Bizer. TriQL.P - A Query Language for Querying Named Graphs Published by Untrustworthy Sources. `http://wifo5-03.informatik.uni-mannheim.de/bizer/triqlp/Spec/`, 2004. [online; accessed: 07-05-2014].

C. Bizer. Semantic Web Publishing Vocabulary (SWP) User Manual. Technical report, 2006. [online; accessed: 07-05-2014].

C. Bizer and R. Cyganiak. Quality-driven information filtering using the WIQA policy framework. *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(1):1–10, January 2009. ISSN 15708268.

C. Bizer and R. Cyganiak. RDF 1.1 TriG. `http://www.w3.org/TR/2013/CR-trig-20131105/`, 2013. [online; accessed: 14-11-2013].

C. Bizer, R. Cyganiak, T. Gauss, and O. Maresch. The TriQL. P browser: Filtering information using context-, content-and rating-based trust policies. In *Proceedings of the Semantic Web and Policy Workshop at the 4th International Semantic Web Conference*, volume 7, pages 12–20, 2005.

C. Bizer, T. Heath, and T. Berners-Lee. Linked data-the story so far. *International Journal on Semantic Web and Information Systems*, 5(3):1–22, 2009.

K. Blomqvist. The many faces of trust. *Scandinavian journal of management*, 13(3): 271–286, 1997.

J. Bohannon. Who's Afraid of Peer Review? *Science*, 342(6154):60–65, 2013.

V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.

J. J. Carroll, C. Bizer, P. Hayes, and P. Stickler. Named graphs, provenance and trust. In *Proceedings of the 14th international conference on World Wide Web*, pages 613–622, New York, New York, USA, 2005a. ACM Press. ISBN 1595930469.

J. J. Carroll and P. Stickler. Trix: RDF Triples in xml. `http://www.hpl.hp.com/techreports/2004/HPL-2004-56.pdf`, 2007. [online; accessed: 28-10-2013].

J.J. Carroll, C. Bizer, P. Hayes, and P. Stickler. Named graphs. *Web Semantics: Science, Services and Agents on the World Wide Web*, 3(4):247–267, 2005b.

Cheskin Research and Studio Archetype/Sapient. eCommerce trust study. Technical report, Cheskin Research, 1999. [online; accessed: 07-05-2014].

J. Cohen. A power primer. *Psychological Bulletin*, 112(1):155–159, 1992a.

J. Cohen. Statistical power analysis. *Current directions in psychological science*, 1(3): 98–101, 1992b.

J. Cook and T. Wall. New work attitude measures of trust: Organizational commitment and personal need non-fulfilment. *Jounal of Occupational Psychology*, 53(1):39–52, 1980.

C. L. Corritore, B. Kracher, and S. Wiedenbeck. On-line trust: concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, 58(6):737–758, June 2003. ISSN 10715819.

P. Crowder and D. A. Crowder. The Basics of Building Web Pages and Sites. In *Creating web sites bible*, chapter Chapter 1, pages 3–10. John Wiley & Sons, 1 edition, 2008.

S. C. Currall and T. A. Judge. Measuring Trust between Organizational Boundary Role Persons. *Organizational Behavior and Human Decision Processes*, 64(2):151–170, November 1995. ISSN 07495978.

R. Cyganiak and A. Jentzsch. The Linking Open Data cloud diagram. `http://lod-cloud.net/`, 2011. [online; accessed: 06-05-2014].

S. Decker, P. Mitra, and S. Melnik. Framework for the semantic Web: an RDF tutorial. *Internet Computing, IEEE*, 4(6):68–73, 2000. ISSN 10897801.

W. H. DeLone and E. R. McLean. Information Systems Success: The Quest for the Dependent Variable. *Information Systems Research*, 3(1):60–95, March 1992. ISSN 1047-7047.

EPrints. ePrints. `http://www.eprints.org/`, 2000. [online; accessed 19/12/2013].

E. Erdfelder, F. Faul, and A. Buchner. GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, 28(1):1–11, March 1996. ISSN 0743-3808.

A. Field. *Discovering statistics using SPSS*. SAGE Publications Ltd, 3rd edition, 2009. ISBN 9781847879066.

R. A. Fisher and J. H. Bennett. *Statistical Methods, Experimental Design, and Scientific Inference*. Oxford University Press, 1973.

FOAF. The Friend of a Friend (FOAF) project. `http://www.foaf-project.org/`, 2000. [online; accessed 13/11/2013].

B. J. Fogg, J. Marshall, A. Osipovich, C. Varma, O. Laraki, N. Fang, J. Paul, A. Rangnekar, J. Shon, P. Swani, and M. Treinen. Elements that Affect Web Credibility: Early Results from a Self-Report Study. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems*, pages 287–288, New York, NY, USA, 2000. ACM.

B. J. Fogg, C. Soohoo, D. Danielson, L. Marable, J. Stanford, and E. R. Tauber. How do people evaluate a Web site's credibility? Results from a large study. Technical report, 2002. [online; accessed 07/05/2014].

B. J. Fogg, C. Soohoo, D. R. Danielson, L. Marable, J. Stanford, and E. R. Tauber. How do users evaluate the credibility of Web sites?: a study with over 2,500 participants. In *Proceedings of the 2003 conference on Designing for user experiences*, pages 1–15, 2003.

B.J. Fogg, J. Marshall, T. Kameda, J. Solomon, A. Rangnekar, J. Boyd, and B. Brown. Web credibility research: a method for online experiments and early study results. *CHI'01 extended abstracts on Human factors in computing systems*, pages 295–296, 2001a.

B.J. Fogg, J. Marshall, O. Laraki, A. Osipovich, C. Varma, N. Fang, J. Paul, A. Rangnekar, J. Shon, P. Swani, and M. Treinen. What makes Web sites credible?: a report on a large quantitative study. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 61–68, 2001b.

F. Fukuyama. *Trust: The social virtues and the creation of prosperity*. New York: Free Pres, 1995.

J. Golbeck. *Trust on the World Wide Web: A Survey*. Foundations and Trends in Web Science, 1 edition, 2006.

J. Golbeck, B. Parsia, and J. Hendler. *Trust networks on the semantic web*. Springer Berlin Heidelberg, 2003.

T. Grandison and M. Sloman. A survey of trust in internet applications. *Communications Surveys & Tutorials, IEEE*, 3(4):2–16, 2000.

S. Harnad. What is Open Access. `http://www.eprints.org/openaccess/`, 2000. [online; accessed: 19-12-2013].

S. Harnad. The Implementation of the Berlin Declaration on Open Access. *D-Lib Magazine*, March 2005.

A. Harth, M. Janik, and S. Staab. Semantic Web architecture. *Handbook of Semantic Web Technologies*, pages 43–75, 2011.

A. Harth, A. Oberweis, A. Wagner, C. R. Hagen, D. Seese, F. Dengler, G. Qi, H. Schmeck, H. Lewe, M. Krötzsch, P. Cimiano, P. Sorg, R. Studer, S. Blohm, S. Rudolph, T. Högler, Duc T. Tran, U. Hellinger, and W. Stucky. Semantic Web. `http://semanticweb.org/wiki/Main_Page`, 2009. [online; accessed: 28-10-2013].

T. Heath and C. Bizer. Linked data: Evolving the web into a global data space. In James Hendler and Frank Van Harmelen, editors, *Synthesis Lectures on the Semantic Web: Theory and Technology*, pages 1–136. Morgan & Claypool publishers, 1 edition, 2011.

I. Herman, H. Halpin, S. Hawke, E. Prud'hommeaux, D. Raggett, and R. Swick. W3C Semantic Web Activity. `http://www.w3.org/2001/sw/`, 2008. [online; accessed: 28/10/2013].

J. E. Hirsch. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46):16569–16572, 2005.

N. Höchstötter and D. Lewandowski. What users see Structures in search engine results pages. *Information Sciences*, 179(12):1796–1812, May 2009. ISSN 00200255.

G. Hughes, A. Lewins, and C. Silver. Analysing Open-Ended Survey Question Data in CAQDAS Packages: Data Preparation for NVIVO (V7 and V8). `http://www.surrey.ac.uk/sociology/research/researchcentres/caqdas/support/analysingsurvey/analysing_openended_survey_question_data_in_caqdas_packages_data_preparation_for_nvivo_v7_v8.htm`, 2010. [online; accessed: 28-10-2013].

T. Huynh, N. R. Jennings, and N. R. Shadbolt. An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 13(2):119–154, March 2006. ISSN 1387-2532.

IBM Corporation. Statistics Base option. `ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM$\delimiter"026E30F$_SPSS$\delimiter"026E30F$_Statistics$\delimiter"026E30F$_Base.pdf`, 2012. [online; accessed: 28-10-2013].

P. Jaccard. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37:547–579, 1901.

B. Kaliski. A survey of encryption standards. *Micro, IEEE*, 13(6):74–81, December 1993. ISSN 0272-1732.

K. Kelton, K. R. Fleischmann, and W. A. Wallace. Trust in digital information. *Journal of the American Society for Information Science and Technology*, 59(3):363–374, 2008.

S. A. Knight and J. M. Burn. Developing a framework for assessing information quality on the World Wide Web. *Informing Science: International Journal of an Emerging Transdiscipline*, 8(5):159–172, 2005.

J. D. Lewis and A. Weigert. Trust as a social reality. *Social forces*, 63(4):967–985, 1985.

R. Likert. A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22 (140):5–55, 1932.

M. MacDonald. Preparing for the Web. In *Creating a Website: The Missing Manual*, pages 7–17. O'Reilly Media, Inc., 2011. ISBN 9781449301729.

F. Manola, E. Miller, and B. McBride. RDF primer. *W3C recommendation*, 10:1–107, 2004.

Marriam-webster. Merriam-Webster Online Dictionary. `http://www.merriam-webster.com/`, 2011. [Online; accessed 05-12-2013].

S. Marsh and M. R. Dibben. The role of trust in information science and technology. *Annual Review of Information Science and Technology*, 37(1):465–498, January 2003. ISSN 00664200.

B. Matthews. Semantic web technologies. *E-learning*, 6(6):8, 2005.

D. L. McGuinness and F. Van Harmelen. OWL web ontology language overview. *W3C recommendation*, 10(2004-03):10, 2004.

D. H. McKnight and N. L. Chervany. The meanings of trust. Technical report, Carlson School of Management, University of Minnesota, 1996.

H. Miyamori, S. Akamine, Y. Kato, K. Kaneiwa, K. Sumi, K. Inui, and S. Kurohashi. Evaluation data and prototype system WISDOM for information credibility analysis. *Internet Research*, 18(2):155–164, 2008.

F. Naumann. Quality-driven query answering for integrated information systems. *Database Management & Information Retrieval*, 2261:3–10, 29–50, 2002.

W. Nejdl, M. Wolpers, and C. Capelle. The RDF schema specification revisited. In *Workshop Modellierung 2000*, 2000.

J. Nielsen and T. K. Landauer. A mathematical model of the finding of usability problems. *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 206–213, 1993.

N. Noorderhaven. Trust and transactions: toward transaction cost analysis with a differential behavioral assumption. *Tijdschrift voor Economie en Management*, 15(1): 5–18, 1995.

OED Online. The Oxford English Dictionary 2nd ed. 1989,. `http://www.oed.com/`, 2011. [online; accessed: 28-10-2013].

K. O'Hara. A general definition of trust. [online; accessed: 07-05-2014], 2012.

J. A. Oluwatayo. Validity and Reliability Issues in Educational Research. *Journal of Educational and Social Research*, 2(2):391–400, 2012.

J. R. Otto, J. H. Cook, and Q. B. Chung. Extensible markup language and knowledge management. *Journal of Knowledge Management*, 5(3):278–285, 2001.

Persuasive Technology Lab. Stanford Web Credibility Research. `http://credibility.stanford.edu/`, 2007. [online; accessed: 28-10-2013].

L. Pipino, R. Wang, D. Kopcso, and W. Rybold. Developing Measurement Scales for Data-Quality Dimensions. *ME Sharpe, New York*, 1:37, 2005.

J. T. Pollock. *Semantic Web for Dummies*. John Wiley and Sons Ltd, 2009.

B. Prajapati, M. Dunne, and R. Armstrong. Sample size estimation and statistical power analyses. *Optometry Today*, 16(7), 2010.

S. Ramachandran, S. Paulraj, S. Joseph, and V. Ramaraj. Enhanced Trustworthy and High-Quality Information Retrieval System for Web Search Engines. *Internation Journal of Computer Science*, 5:38–42, 2009.

S. Y. Rieh and N. Belkin. Interaction on the Web: Scholars' Judgement of Information Quality and Cognitive Authority. *Proceedings of the annual meeting-american society for information science*, 37:25–38, 2000.

S. Y. Rieh and N. J. Belkin. Understanding judgment of information quality and cognitive authority in the WWW. In *Proceedings of the 61st annual meeting of the american society for information science*, volume 35, pages 279–289. Citeseer, 1998.

RIF Working Group. RIF. `http://www.w3.org/2005/rules/wiki/RIF_Working_Group`, 2005. [online; accessed: 28-10-2013].

M. B. Rosson and J. M. Carroll. *Usability Engineering: Scenario-based Development of Human-computer Interaction*. Morgan Kaufmann Publishers Inc., 2002.

F. D. Schoorman, R. C. Mayer, and J. H. Davis. An integrative model of organizational trust: Past, present, and future. *The Academy of Management Review*, 32(2):344–354, 2007.

Scrapy Developers. Scrapy. `http://scrapy.org/`, 2008. [online; accessed: 19-12-2013].

G. B. Shelly, H. A. Napier, and O. N. Rivers. Types of Web Sites. In *Web Design: Introductory Concepts and Techniques*, chapter 1, pages 18–20. Course Technology, Cengage Learning, third edition, 2008. ISBN 9781423927181.

J. Sklar. Planning the Site. In *Principles of Web Design*, chapter 3, pages 64–65. Course Technology, Cengage Learning, fourth edition, 2008.

M. E. Smid and D. K. Branstad. The Data Encryption Standard Past and Future. *Proceedings of the IEEE*, 76(5):550–559, 1988.

B. E. Smith. Web Publishing Basics. In *Creating Web Pages For Dummies*, chapter Chapter 1, pages 9–23. Wiley Publishing, Inc., 9th edition, 2008.

SPARQL Working Group. SPARQL. `http://www.w3.org/2009/sparql/wiki/Main_Page`, 2009. [online; accessed: 28-10-2013].

B. Stvilia, L. Gasser, M. B. Twidale, and L. C. Smith. A framework for information quality assessment. *Journal of the American Society for Information Science and Technology*, 58(12):1720–1733, 2007.

M. A. Tate. *Web Wisdom: How To Evaluate and Create Information Quality on the Web*. CRC Press, 2nd edition, 2010.

G. K. Tayi and D. P. Ballou. Examining data quality. *Communications of the ACM*, 41 (2):54–57, 1998.

R. S. Taylor. *Value-added processes in information systems*. Norwood, NJ: Ablex Publishing Corporation, 1986.

Thomson Reuters. Web of science. `http://thomsonreuters.com/products/ip-science/04_062/wos-next-gen-brochure.pdf`, 2008. [online; accessed: 07-05-2014].

S. Tseng and B. J. Fogg. Credibility and computing technology. *Communications of the ACM*, 42(5):39–44, 1999.

M. Uschold. Where are the Semantics in the Semantic Web? *AI Magazine*, 24(3):25, 2003.

W3C. RDF Primer. `http://www.w3.org/TR/2004/REC-rdf-primer-20040210/`, 2004. [Online; accessed 24-10-2013].

W3C. SPARQL Query Language for RDF. `http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/`, 2008. [online; accessed: 28-10-2013].

W3C SWEO Community Project. Linking Open Data project. `http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData`, 2013. [online; accessed: 06-05-2014].

W. Wang, G. Zeng, and D. Tang. Using evidence based content trust model for spam detection. *Expert Systems with Applications*, 37(8):5599–5606, August 2010. ISSN 09574174.

C. N. Wathen and J. Burkell. Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, 53 (2):134–144, 2002. ISSN 1532-2882.

D. Weitzner and A. Towvim. TrustLayers. `http://trustlayers.com/`, 2014. [online; accessed: 07-05-2014].

R. K. Yin. *Case study research design and methods.* Thousand Oaks, CA: Sage Publications, third ed. edition, 2009.