Invisible Web and Academic Research: A Partnership for Quality

Huda Y. Alyami¹ & Eman A. Assiri²

¹ College of Education, Najran University, Kingdom of Saudi Arabia
² Ministry of Education, Yanbu, Kingdom of Saudi Arabia

Correspondence: Huda Y. Alyami¹, College of Education, Najran University, Kingdom of Saudi Arabia. E-mail: hyalyami@nu.edu.sa

Received: November 11, 2017      Accepted: January 4, 2018      Online Published: March 28, 2018
doi:10.5539/ies.v11n4p84                  URL: https://doi.org/10.5539/ies.v11n4p84

Abstract

The present study aims to identify the most significant roles of the invisible web in improving academic research and the main obstacles and challenges facing the use of the invisible web in improving academic research from the perspective of academics in Saudi universities. The descriptive analytical approach was utilized in this study. It covered all faculty members in Saudi universities. It applied a 20-paragrapgh questionnaire to a randomly selected sample of 168 academics. It concluded that the participants agreed on the role of the invisible web in improving academic research, with an arithmetic means of 3.91. They also agreed on the obstacles of using invisible web for the improvement of academic research, with an arithmetic means of 4.107. The study provides ideas that would develop the use of the invisible web in higher education institutions in Saudi Arabia, in particular, and the Arab countries, in general. Furthermore, it is hoped that such results may provide decision-makers, educational designers and programmers with solutions for the development of research engines and academic databases in Arabic.

Keywords: deep web, higher education, invisible web

1. Introduction

Academic research is a key element in the basic system of higher education because it is an active element in higher education processes and outcomes. On reviewing the objectives of higher education in Saudi Arabia, it could be concluded that they reflect a mental orientation towards academic research within the current trends and requirements of academic institutions. Therefore, it has become an essential part of the mission of any higher education institution in Saudi Arabia.

According to Al-Uthman (2010), the difference between wealth and poverty is knowledge and what generates knowledge is the strong academic research. That is, academic research is highly important to higher education institutions. The educational institutions are evaluated in terms of their keenness on academic research and their relevant allocations as well as their usage of technology such as the computer, internet and search engines in obtaining information, conducting relevant qualitative and quantitative analysis and scientifically explaining such information so as to generate knowledge and to serve the community (Al-Mshikah, 2008).

The Internet is fundamental to all fields of academic research and it has become the most preferred tool for academics and researchers to access their sources of information through databases and digital libraries instead of the traditional methods. Su (2009) points out that the Internet connects researchers with academic institutions and provides several tools for electronic access to information sources. Additionally, digital sources provide researchers with quick and direct access to the basic sources of information. Al-Arishi and Al-Mehdar (2011) argue that the Internet is an ideal environment for easily providing various sources of information. It has resulted in web pages, image and audio files, as well as an endless number of scientific and specialized databases. Such databases have qualitative information of higher educational value than those of web pages as they are subject to arbitration and supervision by specialists and academic bodies. Databases are increasing, and their size exceeds five hundred times the size of pages and other web files. Additionally, they differ in their nature, technical and organizational structure and ways of access to information from other web pages. Therefore, they have formed a separate part from the current web part with several terms such as hidden web, invisible web and deep web (Sayyed, 2010).

Nejad and Soleimani (2017) report that information in the invisible network is usually more than the resources
retrieved by search engines in visual web networks. So, they are important for faculty and researchers in terms of quantity and quality of information available. Su (2009) further stresses the significance and uniqueness of information in the invisible web that is closely related to research materials and researchers and assures that many researchers and librarians shall use them for the purposes of academic research.

In order to achieve the optimal utilization of the invisible web in higher education, it is necessary to identify its actual role in improving academic researches, practically by referring to academics in the higher education institutions and identifying the obstacles of using the invisible web in this field.

2. Theoretical Framework

According to Su (2009), the term invisible web emerged in 1997, along with other synonyms including deep web or hidden web. It refers to the content of data that traditional search engines cannot reach (Mayr & Lewandowski, 2006). The data size represents 96% of the total web size (Madhusudan & Poonam, 2017; Kahate & Bangar, 2016). Devine and Egger-Sider (2014) identify the invisible web as a collection of high quality and reliable web pages or files that are often available on the web and are limited to a specific group that search engines cannot discover for technical issues or they require a password to be accessed. Bergman (2001) reports that the invisible web is pages that are dynamically generated as a result of a particular search that traditional search engines cannot see or retrieve their content. Procedurally, Ibrahim (2011) defines it as a set of web pages or websites that rely on databases of different size that cannot be handled by search engines because of the inability of the crawler to deal with the organization mechanisms and to search in its databases. Notably, all definitions agree on the multiplicity of obstacles that hinder full access to the invisible web via traditional methods, restricting its optimal utilization. Long (2006) identifies the most significant differences between the visible and invisible web, as follows:

Table 1. The most significant differences between the visible and invisible web

<table>
<thead>
<tr>
<th>Visible Web</th>
<th>Invisible Web</th>
</tr>
</thead>
<tbody>
<tr>
<td>Available for all users</td>
<td>Limited to its subscribers</td>
</tr>
<tr>
<td>There is no supervision over the available sources. They are used without evaluation.</td>
<td>Sources are prepared by experts in the field and are arbitrated according to certain criteria.</td>
</tr>
<tr>
<td>Can be used in any field</td>
<td>Educational material is limited to a certain group</td>
</tr>
<tr>
<td>Organized, limited and outdated</td>
<td>Has large updated and mostly valid and reliable databases</td>
</tr>
<tr>
<td>Easy to search and use</td>
<td>Requires using keywords followed by a specific mechanism to access the content</td>
</tr>
</tbody>
</table>

The invisible web consists of academic information, medical records, scientific reports, government resources and several valuable web contents that cannot be indexed and accessed by traditional search engines (Madhusudan & Poonam, 2017). Sherman & Price (2007) and Long (2006) agree on the classification of the types of the invisible web according to the validity and accessibility of search engines into four types:

- The opaque web is limited to political areas and is considered a restricted area of a first degree due to its great importance and relevance to the state's affairs.
- The private web is related to a person who manages it himself. It ensures privacy and provides super protection, where ads are blocked and files are not allowed to store in the browser.
- The proprietary web focuses on a specific company to help it organize its schedules and maintain its business and investments at all local and international levels.
- The invisible web is huge repositories that provide an academic content by specialists in a specific field and allow only a certain category to participate in it. They allow additions and modifications according to systems and laws that are subject to control and continuous follow-up.

It is known that each weapon is a double-edged one, which is applicable to the invisible web being a comprehensive knowledge of all political, economic, commercial, health and educational sciences. Ubaid (2015) reported that it is a safe way for obtaining confidential information relating to any company in the banking process and for individuals which highly save their ownership rights. Simultaneously, it is a profitable process that requires the payment of fees according to the adopted system in the payment strategy. Devine and Egger-Sider (2014) provided several proposals on the significance of using the invisible web as a tool for enhancing the way of knowledge sharing. Although it is used to conduct purposeful academic researches, it is a dangerous weapon that may affect the policy of any state internally or externally. Pederson (2013) reported that it is a means of money
laundering, forging credit cards, and identity theft. Paganini (2012) stresses that it provides an ideal environment for an illegal profitable trade and criminal services for selling weapons, drugs and medically prohibited medicines as well as the political files relating to many countries. Nevertheless, serious attempts are underway to eliminate such acts that may devastate people's lives without discrimination.

Several studies sought to identify strategies that explain how to access files or content on the invisible web. For example, Long (2006) developed some strategies, including:

- Independence from traditional search engines.
- Paying subscription fees for digital libraries and specialized databases.
- Having access to some websites such as governmental websites to benefit from their content.
- Maintaining usage to master how to follow the links of dynamic pages.

Su (2009) provided other strategies that expand search options by using keywords along with these ones as a logical search for (database, repository, digital library, archive, data sets and statistics) where many unique links appear.

Banu and Chitra (2014) argue that many attempts are being made to overcome technical barriers or decisions intended to exclude web pages and limit them to a particular group, suggesting that the content of the invisible web will be available in a wider range in order to take advantage of its data.

Academically, Devine and Egger (2004) noted that the invisible web has many advantages that enrich academic libraries and their content. Ibrahim (2011) also notes that digital libraries of the invisible web represent a resource for all academic institutions under the supervision of specialized staff, enabling effective participation in the establishment of projects worldwide by a specific community or any other agency whether educational or commercial.

In light of the growing interest of academics and their keenness on adding quality to the output of their academic research, the need to employ the invisible web is emerging in academic research. Leach (2010) reports that the invisible web has radically changed information in terms of quality and coverage, going beyond the visual web accessible through search engines for general purposes. In addition, Aspinall (2010) stresses the need to know the value of using invisible web resources to access information, warning from the consequences of relying on the web to meet search requirements. This is especially true in the field of health sciences where the cost of not finding the right information can be expensive. Therefore, most international and local universities have made it easier to access a vast number of sources and databases on the invisible web due to its huge characteristics, the most important of which, according to Sherman and Price (2007), are the availability of good scientific content based on the discussion and participation of specialists from experts and scientists in a particular field, limiting amendments and additions to the experienced participants to render the results credible and reliable, maintains intellectual property and encourages researchers to use technology with a high degree of skills in order to save time and effort. Banu and Chitra (2014) believe that the invisible web is difficult to access through traditional search engines, considering it one of the greatest unique characteristics of its documents which maintain accuracy and quality of information.

In spite of the efforts made by academic institutions to expand the availability of the invisible web content for their employees, the benefit was limited. In this context, Mayr and Lewandowski (2006) state that the use of the invisible web in academic research is very limited due to the lack of guidelines for its usage. Universities also need to hold training courses for their students that show the best ways to access and use available academic resources. Nejad and Soleimani (2017) conclude that faculty members tend to use visible search engines more frequently than the invisible web due to the lack of relevant training. They also stress the necessity of using the web in its visible and invisible forms as a complementary electronic site in order to meet one's needs of general and private knowledge. This result agrees with the present study.

Ehney and Shorter (2016) argue that everyone needs to be educated and informed about the usages and capabilities of the invisible web as the path to knowledge. Despite the ongoing efforts at all levels to support scientific research in academic institutions, the academic community still requires a lot of support, especially in the field of academic research techniques and relevant technical aspects. In this regard, Fadi (2009) emphasizes that the weakness in the outputs of research today is due to the lack of good citation from other qualified researches as a result of the inability to access the invisible web. Kalmbach (2013) reports that the nature of the invisible web requires new tools, methodologies and strategies that researchers have to keep pace with to access information.
3. Method

The study adopted the descriptive approach that studies the phenomenon in reality and describes it accurately in a qualitatively and quantitatively.

3.1 Population and Sampling

The study population comprised all academics in Saudi universities, while the sample of the study included (168) academics who have basic experience and knowledge of the invisible web, from King Abdulaziz University, King Khalid University, Najran University, King Saud University, Princess Noura Bint Abdulrahman University, Dammam University, King Faisal University, Jazan University, Al Baha University, Northern Borders University, Saudi Electronic University and Majmaah University. Subjects of the sample are distributed according to their academic degrees, as follows:

Table 2. Distribution of faculty members according to their academic degrees

<table>
<thead>
<tr>
<th>Academic Degree</th>
<th>Professor</th>
<th>Associate Professor</th>
<th>Assistant Professor</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>9</td>
<td>37</td>
<td>122</td>
</tr>
</tbody>
</table>

3.2 Tool

The study adopted the closed questionnaire as the main tool for collecting the required data to achieve the objectives of the study because it is easy to compile and classify the necessary collected information. A comprehensive survey of the related local, regional and global literature was conducted to identify the role of the invisible web in the improvement of academic research and the obstacles of its usage in this regard. Additionally, the opinions of specialized professors in the field were surveyed through referenda and open interviews. The paragraphs that represent the role of the invisible web in improving the academic research and the obstacles to its usage were prepared, followed by writing and drafting the paragraphs. Then, they were presented to reviewers whose opinions were taken into consideration.

3.2.1 Validity of the Questionnaire

The face validity of the questionnaire was verified by presenting it to a group of (9) specialized reviewers. Based on their observations and suggestions, amendments were made to prepare the questionnaire in its final form.

Its content validity was estimated by calculating the correlation coefficients between each paragraph of the questionnaire and the total score of the domain itself. They ranged from (0.717) and (0.856). They were statistically significant at the level of (α = 0.05). Thus, the questionnaire was considered valid.

3.2.2 Reliability of the Questionnaire

The reliability of the questionnaire was estimated. Its reliability coefficients were calculated using Cronbach's alpha. The reliability coefficient of the tool was (0.866) that was high and emphasized that the tool has a high degree of reliability. Accordingly, the tool is often adequate for the application.

3.2.3. Procedures for Application

The questionnaire was distributed to the study sample of (168) academics in the aforementioned universities. Then, it was collected and data were extracted to be statistically analyzed.
4. Results

4.1 Explanation of the Questionnaire’s Results to Clarify the Role of the Invisible Web in Improving Academic Researches

Table 3. The role of the invisible web in improving academic researches

<table>
<thead>
<tr>
<th>No.</th>
<th>First Domain</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Undecided</th>
<th>Agree</th>
<th>Strongly agree</th>
<th>Weighted average</th>
<th>Standard deviation</th>
<th>Trend</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>It represents a main sector of information sources on the Internet among the academic community.</td>
<td>1</td>
<td>8</td>
<td>26</td>
<td>76</td>
<td>57</td>
<td>4.07</td>
<td>0.859</td>
<td>Agree</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>It represents a more efficient source of academic research, increasing supporting results.</td>
<td>2</td>
<td>5</td>
<td>19</td>
<td>41</td>
<td>33</td>
<td>3.98</td>
<td>0.966</td>
<td>Agree</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>It provides high-quality sources with standards subject to the supervision of academic agencies.</td>
<td>1</td>
<td>5</td>
<td>23</td>
<td>38</td>
<td>33</td>
<td>3.96</td>
<td>0.928</td>
<td>Agree</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>It provides highly credible data with the scientific nature.</td>
<td>2</td>
<td>4</td>
<td>20</td>
<td>46</td>
<td>28</td>
<td>3.95</td>
<td>0.897</td>
<td>Agree</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>It supports effective participation with specialized people.</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>25</td>
<td>3.82</td>
<td>0.905</td>
<td>Agree</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>It serves as a leading source of new academic information on the Internet.</td>
<td>1</td>
<td>7</td>
<td>51</td>
<td>63</td>
<td>45</td>
<td>3.86</td>
<td>0.904</td>
<td>Agree</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>Comprehensive coverage of all fields of knowledge and fields with various forms (ppt, mdb, pdf).</td>
<td>2</td>
<td>7</td>
<td>51</td>
<td>63</td>
<td>45</td>
<td>3.85</td>
<td>0.909</td>
<td>Agree</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>It provides a set of useful tools for academic researches.</td>
<td>2</td>
<td>15</td>
<td>44</td>
<td>55</td>
<td>42</td>
<td>3.83</td>
<td>1.007</td>
<td>Agree</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>It overcomes obstacles of the invisible web through allowing for tracking all changes in the sources.</td>
<td>5</td>
<td>7</td>
<td>42</td>
<td>73</td>
<td>41</td>
<td>3.82</td>
<td>0.950</td>
<td>Agree</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>It represents a main sector of information sources on the Internet among the academic community.</td>
<td>3</td>
<td>4</td>
<td>25</td>
<td>44</td>
<td>24</td>
<td>3.81</td>
<td>0.945</td>
<td>Agree</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 3 shows that the general mean of the total responses scored (3.91), reflecting an agreement in the responses to the paragraphs of this domain, given that the paragraphs leveled (Agree) with means ranged (3.82-4.07).

Additionally, the paragraph “it represents a main sector of information sources on the Internet among the academic community” scored the highest means of agreement (4.07), while other paragraphs were lower with various degrees of agreement. While “it represents a more efficient source of academic research, increasing supporting results” was ranked second (3.98), “it provides high-quality sources with standards subject to the supervision of academic agencies” was ranked third (3.96) and “it provides highly credible data with the scientific nature” was
ranked fourth (3.95). “It provides highly specialized sources” which ranked fifth (3.92) and “It supports effective participation with specialized people” was ranked sixth (3.89). “It serves as a leading source of new academic information on the internet” was ranked seventh (3.86), “comprehensive coverage of all fields of knowledge and fields with various forms (ppt, mdb, pdf) was ranked eighth (3.85) and “it provides a set of useful tools for academic researches” was ranked ninth (3.83). “It overcomes obstacles of the invisible web through allowing for tracking all changes in the sources” was ranked last with arithmetic means of (3.82).

4.2 Explanation of the Results of the Questionnaire Applied to Clarify the Obstacles of the Invisible Web for the Improvement of Academic Researches

Table 4. Obstacles of using the invisible web for the improvement of academic researches

<table>
<thead>
<tr>
<th>Second domain</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Undecided</th>
<th>Agree</th>
<th>Strongly agree</th>
<th>Weighted average</th>
<th>Standard deviation</th>
<th>Trend</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>The different design of databases and their information retrieval</td>
<td>-</td>
<td>5</td>
<td>21</td>
<td>65</td>
<td>77</td>
<td>4.27</td>
<td>0.794</td>
<td>Strongly agree</td>
<td>1</td>
</tr>
<tr>
<td>Most of the resources available through the invisible web is in English.</td>
<td>2</td>
<td>3</td>
<td>12%</td>
<td>39%</td>
<td>46%</td>
<td>4.26</td>
<td>0.862</td>
<td>Strongly agree</td>
<td>2</td>
</tr>
<tr>
<td>The restrictions and conditions set to benefit from the contents of the invisible web.</td>
<td>2</td>
<td>3</td>
<td>11%</td>
<td>38%</td>
<td>47%</td>
<td>4.24</td>
<td>0.830</td>
<td>Strongly agree</td>
<td>3</td>
</tr>
<tr>
<td>The efforts required for searching on the invisible web for the sub-domains in its pages</td>
<td>-</td>
<td>7</td>
<td>21</td>
<td>67</td>
<td>73</td>
<td>4.23</td>
<td>0.824</td>
<td>Strongly agree</td>
<td>4</td>
</tr>
<tr>
<td>The need for advanced research computer skills</td>
<td>1</td>
<td>7</td>
<td>27</td>
<td>60</td>
<td>73</td>
<td>4.17</td>
<td>0.889</td>
<td>Agree</td>
<td>5</td>
</tr>
<tr>
<td>The lack of knowledge about the research strategies and policies in the invisible web</td>
<td>-</td>
<td>5</td>
<td>29</td>
<td>72</td>
<td>62</td>
<td>4.16</td>
<td>0.857</td>
<td>Agree</td>
<td>6</td>
</tr>
<tr>
<td>The need for specific software to access the available information</td>
<td>-</td>
<td>5</td>
<td>29</td>
<td>72</td>
<td>62</td>
<td>4.14</td>
<td>0.804</td>
<td>Agree</td>
<td>7</td>
</tr>
<tr>
<td>The lack of Arabic (references and manuals) in the topic of the invisible web</td>
<td>5</td>
<td>3</td>
<td>29</td>
<td>65</td>
<td>66</td>
<td>4.10</td>
<td>0.949</td>
<td>Agree</td>
<td>8</td>
</tr>
<tr>
<td>The need for technical support in addressing technical obstacles</td>
<td>3%</td>
<td>5</td>
<td>30</td>
<td>74</td>
<td>50</td>
<td>3.92</td>
<td>0.979</td>
<td>Agree</td>
<td>9</td>
</tr>
<tr>
<td>The nature of the invisible web may include immoral activities like hacking and malware</td>
<td>10</td>
<td>22</td>
<td>35</td>
<td>63</td>
<td>38</td>
<td>3.58</td>
<td>1.150</td>
<td>Agree</td>
<td>10</td>
</tr>
<tr>
<td>The general mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.107</td>
<td>0.894</td>
<td>Agree</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 shows that the general mean of the responses is (4.107), reflecting an agreement in the responses to the paragraphs of this domain given that the paragraphs leveled (Agree) with means of (3.58-4.27). It also shows that the paragraph “the different design of databases and their information retrieval” was ranked first with a means of (4.27), while other paragraphs were ranked lower means with various degrees of agreement. “Most of the
resources available through the invisible web is in English” was ranked second (4.26), “the restrictions and conditions set to benefit from the contents of the invisible web” was ranked third (4.24), followed by “the efforts required for searching on the invisible web for the sub-domains in its pages” (4.23) and “the need for advanced research computer skills” which ranked fifth (4.17). While “the lack of knowledge about the research strategies and policies in the invisible web” was ranked sixth (4.16), “the need for specific software to access the available information” ranked seventh (4.14), “the lack of Arabic (references and manuals) in the topic of the invisible web” ranked eighth (4.10) and “the need for technical support in addressing technical obstacles” was ranked ninth (3.92). “The nature of the invisible web may include immoral activities like hacking and malware” was ranked last with an arithmetic means of (3.58).

Devine and Egger-Sider (2014), Ibrahim (2011) and Mayr (2006) agreed that the knowledge of how to identify the mechanisms and features of search engines is the basis for the quality of academic research. In addition, working as one team in various countries enhances knowledge and addresses several problems and difficult issues by reviewing and examining them by specialists in the field, which in turn guarantees credibility, quality and objectivity.

The findings of the present study are consistent with Lossau and Summann (2004), Long (2006) and Nejad and Soleimani (2017) that highlighted the significance of using the invisible web in academic research and providing access to it in a reliable manner that ensures the quality of the academic content.

5. Recommendations and Conclusion

In light of the findings, the study provides the following recommendations:

1) Expanding the use of the invisible web in academic research.
2) Producing of training media (electronic and traditional) to develop the skills of academics in the use of the invisible web in education and academic research.
3) Preparing guidelines on strategies and mechanisms of using the invisible web for academics.
4) Discussing ways of cooperation with individuals and academic institutions to increase and develop the Arabic content in the invisible web.

The American University of Sharjah (2012) quoted a statement by Jonathan Winters “If your ship doesn't come in, swim out to it” as published by the American University of Sharjah. Therefore, academics and researchers are required to get through the huge amount of accumulated information throughout the whole web in order to maintain the quality of their researches. In this context, the present study stresses the roles of the invisible web in improving academic research and the obstacles of its use in this regard. Despite the sheer volume of data available on the web, it has not been actively invested by academics in academic institutions.

References


Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).