
Systematic Literature Review in Usability Web: A Tertiary Study

Catalina Astudillo-Rodriguez, Santiago Cedillo-Chacon,
Priscila Verdugo-Cardenas, and Jackelin Verdugo-Cardenas

Universidad del Azuay, Cuenca, AZ 010204, Ecuador

ABSTRACT

Web usability is a topic that several researchers have studied; this is because, with new technological advances, it is possible to improve application interfaces significantly and have usable products that captivate users by being friendly, efficient and easy to use. In this sense, reviewing secondary studies reported in this field is considered appropriate. This review aims to identify the methods and/or usability evaluation and validation methodologies, in what phase of development and software domain it has been evaluated, and the problems reported in literature reviews. The applied method is the one proposed by Kitchenham and Charters (2007) for the presentation of tertiary studies, beginning with the planning phase, which includes the activities: definition of research questions, search strategies, criteria for article selection, procedure for the inclusion or exclusion of studies and, criteria to assess the quality of investigations; followed by the review phase, in which the defined planning is carried out, the activities that include: execution of the selection of studies in the digital libraries, evaluation of the quality of the studies, extraction of relevant data and synthesis of the information; and finally the reporting phase, in which the answers to the research questions posed are evidenced. Regarding the results, we found a total of 106 articles, eliminated the duplicates, reviewed the rest and selected 15 papers, in which we identified the most frequently reported evaluation methods and methodologies, such as the Heuristic method, Observational measurement / Inspection and User test/usability tests. Likewise, we found the most commonly reported problems, such as Accessibility issues, Usability measurement and Methodological problems. Also, we distinguish that usability tests have been reported with a higher incidence in Web applications and Educational. Usability evaluation is applied in the different phases of software development, but with other incidences. Finally, we conclude that there is potential for more SLR about the usability of web applications.

Keywords: Web applications, Usability, Systematic literature review, Tertiary study, SLR

INTRODUCTION

Usability is a line of research within human-computer interaction that has been studied for a long time, which has posed in the last two decades, which has generated a large amount of research on methods and/or methodologies, problems, domains and heuristics on the topic. Companies have identified the importance of developing “usable” products that guarantee the permanence of the most significant number of users in their applications, with friendly, efficient and secure systems with the client (Perurena Cancio and

Moráquez Bergues, 2013). Thus, this work attempts to collect relevant information published in systematic literature reviews and identify the reported problems and the software field in which said tests have been applied.

Related Work

This study is oriented to review secondary studies in web usability. In this way, an investigation of similar works is carried out. After a search in specialised databases, little evidence of tertiary studies in the field of usability has been found. The only work identified and published in this field is that of (Curcio et al., 2019), who categorise secondary studies related to usability and its agile development, the result of which carries out a critical reflection on the quality of the studies found.

This bibliographical review focuses on identifying diverse aspects reported in terms of web usability, methods/methodologies, errors, applied software domain, problem reporting and used software development phase, which is why it differs from the study reported by the preview mentioned author.

METHOD

The methodology applied is proposed by (Kitchenham and Charters, 2007) for preparing tertiary reviews. The phases that were implemented were: Phase 1 - Planning that includes the activities: definition of research questions, search strategies, criteria for selection of articles, the procedure for the inclusion or exclusion of studies, and criteria for evaluating the quality of studies. Phase 2 - Execution of the review, in which the defined planning is executed; the activities include: execution of the selection of studies in the digital libraries, evaluation of the quality of the studies, extraction of relevant data and synthesis of the extracted data. Phase 3 - Report the results and answers to the research questions.

Research Questions

The research questions that drive this research are presented in Table 1.

Search Process

Search strings. For the search, primary and alternative terms were used. Likewise, the logical operator OR was used for alternative terms and AND for connection between main terms. The definition of these terms was based on the guidance provided by (Kitchenham et al., 2010) and (Kitchenham and Charters, 2007), and these are: “systematic literature review”, “systematic mapping”, “systematic review”, and “mapping study”.

Online databases. The search string was applied to the journal title only; furthermore, the search was restricted to English-language journals, and the document types were limited to articles, book chapters, conference papers, and journals. The time range of the posts took into account all previous posts up to the present. The digital libraries selected for the automatic search sources were: Scopus, IEEEExplore, ACM portal, Springer Link, Scencedirect, Emerald and Scielo. The final search string used in digital libraries was:

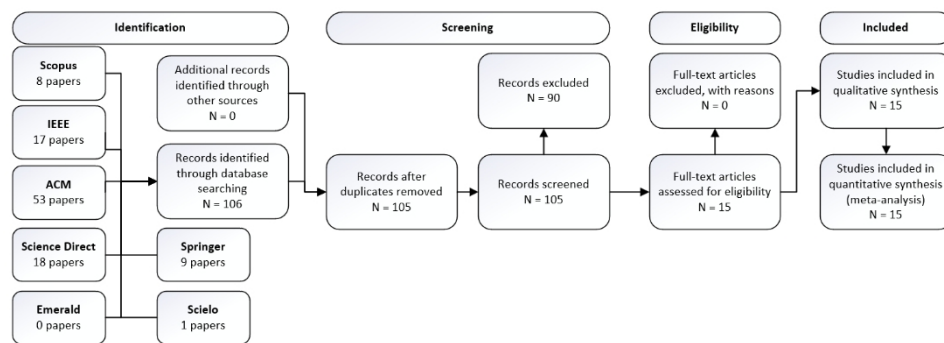


Figure 1: Summarizes the results at each of the review stages.

Table 1. Research questions.

Research question

RQ1. What usability evaluation methods/methodologies have been reported most frequently in web applications?

RQ2. What are the usability issues that have been identified with the highest incidence?

RQ3. What are the software domains in which usability evaluation has been applied?

RQ4. What are the objectives tracked in web application usability evaluation reviews?

Identify where the research interest in usability evaluation in web applications is focused.

RQ5. What usability validation methods/methodologies are most frequently reported in web applications?

“systematic mapping” OR “systematic literature review” OR “systematic review” OR “mapping study” AND “usability” AND “web”.

In some libraries, it was necessary to adjust the search string due to the characteristics of library management.

Study Selection

The inclusion and exclusion criteria that were applied for the selection of articles are summarised. These criteria are established as a review point when reading the title and abstract of each paper.

After applying the search in each library, a total of 106 articles were obtained, and duplicate studies were removed, leaving 105. For each of them, a peer review was considered, in which each researcher reviewed each paper’s abstract and determined whether it was related to the research topic or not, in addition to the application of the established inclusion and exclusion criteria. In those articles where there was a difference between the criteria of the researchers, a joint review was carried out to define their inclusion or

Table 2. Selection criteria.

Inclusion criteria	Exclusion criteria
IC1: The article is related to the research topic	EC1: The article is duplicated
IC2: The article is written in English or Spanish	EC2: The article is not related to the research topic
IC3: The full article is available	EC3: The article is not written in English or Spanish
	EC4: The article is not available

not. After this review, 15 papers were selected. Finally, each one was reviewed in detail to find information that answers one or more of the research questions posed.

Quality Assessment

The quality assessment was carried out according to the recommendations of (Kitchenham and Charters, 2007) and the four criteria presented by DARE <https://www.york.ac.uk/crd/#DARE>. The values considered in the quality evaluation of each question were: 1, 0.5 or 0 according to the criteria of complies, partially complies or does not comply, respectively, according to the specific aspect evaluated, based on what was presented by (Kitchenham et al., 2010). This is detailed in Table 3.

Table 3. Quality evaluation criteria.

Quality criteria	Description / Score
QC1: Are the review's inclusion and exclusion criteria described and appropriate?	The inclusion criteria are explicit. Score 1 The inclusion criteria are implicit. Score 0.5 The inclusion criteria are not defined. Score 0
QC2: Is the literature search likely to have covered all relevant studies?	The authors have searched four or more digital libraries and included additional search strategies or identified. Score 1 The authors have searched 3 or 4 digital libraries with no extra search strategies. Score 0.5 The authors have searched up to 2 digital libraries. Score 0
QC3. Did the reviewers assess the quality/validity of the included studies?	The authors have explicitly defined quality criteria and extracted them from each primary study. Score 1 The research question involves quality issues that are addressed by the study. Score 0.5 No detailed quality assessment of individual papers has been attempted, or quality data has been extracted but not used. Score 0
QC4. Were the basic data/studies adequately described?	Information is presented about each paper so that the data summaries can be traced to relevant papers. Score 1 Only summary information is presented about individual papers, e.g., papers are grouped into categories, but it is impossible to link individual studies to each category. Score 0.5 The results of the individual studies are not specified, i.e., the individual primary studies are not cited. Score 0

Data Extraction and Analysis Process

For data extraction, an electronic sheet was used as a support tool. A matrix was used with the data obtained from the articles, which included: code, library name, type of publication, title, authors, keywords, abstract, year of publication, language, and DOI. In this matrix, the necessary annotations were made to indicate duplicate articles, articles included in the study, and quality assessment.

Data Extraction Results

In the 15 selected articles, the quality assessment criteria previously presented in Table 2 were applied, and the results obtained are shown in Table 4.

Table 4. Quality assessment.

Reference	QC1	QC2	QC3	QC4	Total
(Inan Nur, B. Santoso and O. Hadi Putra, 2021)	1	1	1	0.5	3.5
(Salvador, Nakasone and Pow-Sang, 2014)	1	1	1	1	4
(Fernández and Macías, 2021)	1	1	1	1	4
(Paz and Pow-Sang, 2015)	0.5	1	1	1	3.5
(Yusop, Grundy and Vasa, 2017)	1	1	0	0.5	2.5
(Francisco and Benitti, 2014)	1	1	1	0.5	3.5
(Fernandez, Abrahão and Insfran, 2012)	1	1	1	1	4
(Paz and Jose Antonio Pow-Sang, 2014)	1	1	1	1	4
(Rivero, Barreto and Conte, 2013)	0	1	1	1	3
(Doğan, Betin-Can and Garousi, 2014)	0.5	1	1	0.5	3
(Garousi et al., 2013)	1	1	1	1	4
(Insfran and Fernandez, 2008)	1	1	1	1	4
(Fernandez, Insfran and Abrahão, 2011a)	1	1	1	1	4
(Ugras et al., 2016)	0.5	1	1	1	3.5
(Pellizon et al., 2017)	1	1	1	0.5	3.5

RESULTS AND ANSWERS OF RESEARCH QUESTIONS

The detailed review of the 15 selected articles allowed us to answer the research questions posed; the results obtained in each one are presented below.

RQ1. What Usability Evaluation Methods/methodologies Have Been Reported Most Frequently in Web Applications?

According to (Insfran and Fernandez, 2008), a usability evaluation method collects information on the end user's interaction with a software product. Likewise, according to (Garcés and Egas, 2013), a methodology proposes, as the main objective, to establish a set of traditional, modern and agile systems modelling techniques that allow quality software development, which includes construction heuristics and system model comparison criteria.

In this context, the evaluation methods and methodologies reported in the selected papers have been summarised, ordered by those with the highest frequency of application, outlined in Figure 2. These are the Heuristic method, Observational measurement / Inspection, User test/usability tests, CW

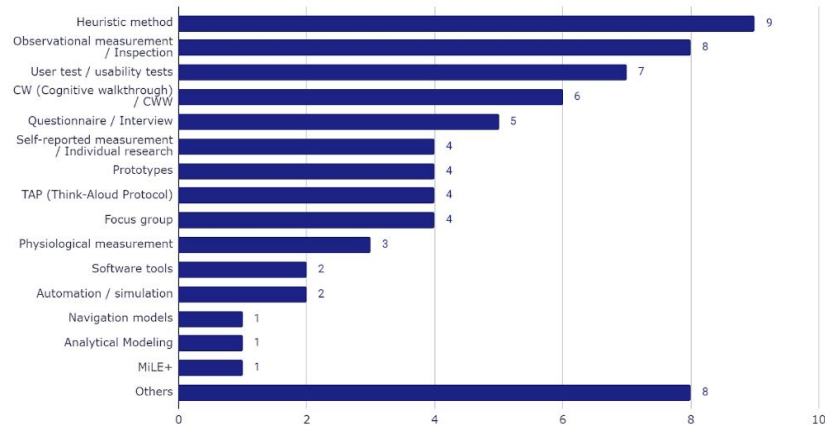


Figure 2: Usability evaluation methods/methodologies.

(Cognitive walkthrough) / CWW, Questionnaire / Interview, Self-reported measurement / Individual research, Prototypes, TAP (Think-Aloud Protocol), Focus group, Physiological measurement, Software tools, Automation/simulation, Navigation models, Analytical Modeling, Mile+, and Others.

RQ2. What Are the Usability Issues That Have Been Identified With the Highest Incidence?

For (Yates and Loaiza, 2003), usability problems are errors or problems in software, web applications, etc., that make them unusable for the end user. Identifying them helps avoid excessively high repair costs and reduces development time. The most frequently reported problems are: Accessibility issues, Usability measurement, Methodological problems, Terminology issues, Search for publications, Publication selection bias, High costs, and Others. The reported frequency is shown in Figure 3.

RQ3. What Are the Software Domains in Which Usability Evaluation Has Been Applied?

The application of usability tests has been reported in different types of software applications; the studies identify Web applications, Educational, Health, Electronic commerce, Social and communication, Entertainment, Industry,

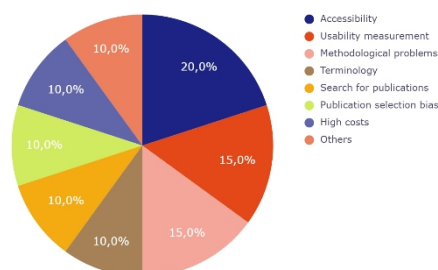


Figure 3: Usability issues identified with the highest incidence.

and Others. The Others are grouped: Culinary Domain, Home Management Domain, Software Development Tools, No Software Information, Training, Government Websites, Search Engines, Prototyping, User Interfaces, and Civil Services. They have been classified by their frequency, as shown in Figure 4.

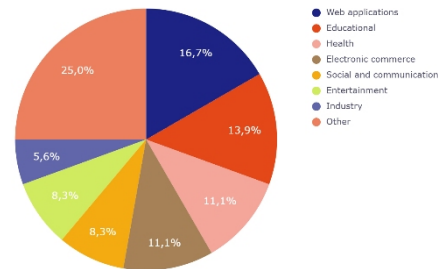


Figure 4: Usability evaluation by software domains.

RQ4. What Are the Objectives Tracked in Web Application Usability Evaluation Reviews?

The reported objectives have been analysed in only five of the selected studies: (Fernandez, Insfran and Abrahão, 2011b, 2011a; Garousi et al., 2013; Paz and Pow-Sang, 2015; Fernández and Macías, 2021). We classify them as Support decision-making in choosing an evaluation technique, Review/Validate a usability evaluation method, Provide usability indices, and Others.

RQ5. What Usability Validation Methods/methodologies Are Most Frequently Reported in Web Applications?

Usability validation methods, according to (Fernandez, Insfran and Abrahão, 2011a), are procedures that are made up of a set of well-defined activities to collect usage data related to the end user's interaction with a software product and how specific properties of this software product contribute to achieving a certain degree of usability. Studies report Survey, Experiment, Data extraction form, Empirical validation, Case study, and Others. The detail is observed in Figure 5.

RQ6. In What Development Phases Has the Usability Evaluation in Web Applications Been Reported?

Software development is carried out in several stages; for our study, it is essential to identify the phases where usability tests are currently carried out. Eight of the studies report the software development phase where the usability evaluation was applied; the most mentioned the design and implementation phase. This is shown in Figure 6.

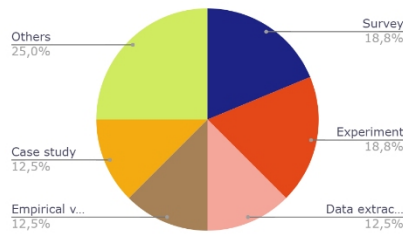


Figure 5: Objectives tracked in web application usability.

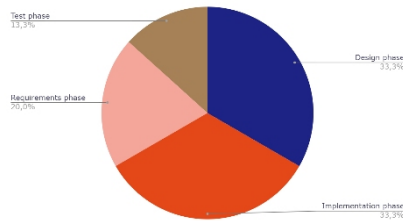


Figure 6: Development phases where applied the usability evaluation.

DISCUSSION OF RESEARCH QUESTIONS

The studies determined that the evaluation methods: Prototypes, User/usability test, and Focus group, present more significant Methodological problems; likewise, the Heuristic method presents Accessibility problems. This is evidenced in Figure 7.

On the other hand, it can be seen in Figure 8 that the domains of educational, health and electronic commerce software present more accessibility problems than other sites. Likewise, the domains of educational, health, e-commerce, and social/communication software present terminology 6 problems.

Both the surveys, the experimentation and the study cases were applied in the design and implementation phase, as shown in Figure 9.

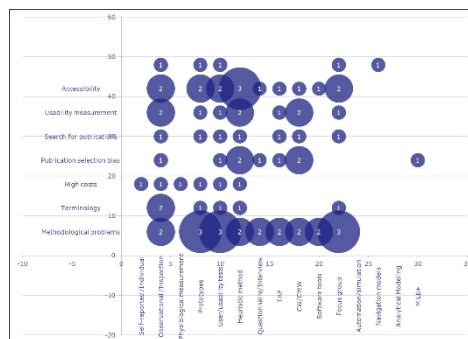


Figure 7: Usability problems by evaluation methods.

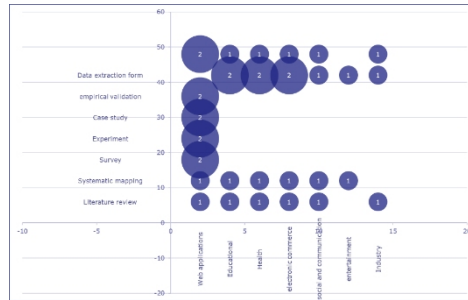


Figure 8: Usability issues in software domains.

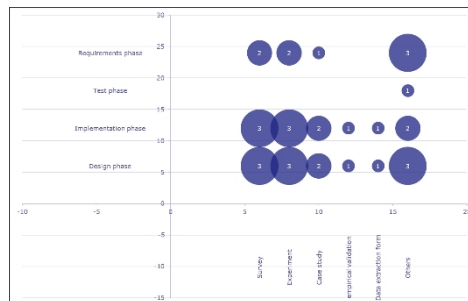


Figure 9: Usability evaluation methodologies in development phases.

CONCLUSION

The objective of our tertiary study was to identify several reported aspects regarding web usability, methods/methodologies, errors, the applied software domain, problem report and the software development phase. Thus, we detected 106 articles between 2008 and 2022, and we selected 15 of these for the final report, which answered our research questions.

When crossing information between the results obtained about RQ1 and RQ2, we were able to observe the studies determined that the evaluation methods: Prototypes, User/usability test, and Focus group, present more significant Methodological problems; Similarly, the Heuristic method presents Accessibility problems.

Observing the results of RQ3, we can conclude that the software domain in which the most significant number of usability evaluations are applied is web applications and educational applications, which suggests that this area is where the demand for software and its usability evaluation is topmost. The opposite is reported with industry-oriented software, which reveals a space for research in that sector.

We believe that knowing the objective of applying usability evaluations is essential since it gives us a vision of what can be improved in usability aspects and guarantee the correct development of future applications; however, in RQ4 of the 15 papers, only 5 of these present clearly defined objectives, represented in investigate the usability methods, validation of the methods and identification of the usability evaluation techniques.

About RQ5 and RQ6, we have found that the most frequent methods for software validation are applied in the design and implementation phases through surveys and experiments.

Finally, we believe that the number of tertiary study SLRs being published regarding web usability seems to be increasing. However, it is still the case that many of these reviews do not report the applied methodology.

REFERENCES

- Curcio, K. et al. (2019) 'Usability in agile software development: A tertiary study', *Computer Standards & Interfaces*, 64, pp. 61–77. Available at: <https://doi.org/10.1016/j.csi.2018.12.003>.
- Doğan, S., Betin-Can, A. and Garousi, V. (2014) 'Web application testing: A systematic literature review', *Journal of Systems and Software*, 91, pp. 174–201. Available at: <https://doi.org/10.1016/j.jss.2014.01.010>.
- Fernández, J. and Macías, J. A. (2021) 'Heuristic-Based Usability Evaluation Support: A Systematic Literature Review and Comparative Study', in *Proceedings of the XXI International Conference on Human Computer Interaction. Interacción '21: XXI International Conference on Human Computer Interaction, Málaga Spain: ACM*, pp. 1–9. Available at: <https://doi.org/10.1145/3471391.3471395>.
- Fernandez, A., Abrahão, S. and Insfran, E. (2012) 'A systematic review on the effectiveness of Web usability evaluation methods', in *16th International Conference on Evaluation & Assessment in Software Engineering (EASE 2012). 16th International Conference on Evaluation & Assessment in Software Engineering (EASE 2012), Ciudad Real, Spain: IET*, pp. 52–56. Available at: <https://doi.org/10.1049/ic.2012.0007>.
- Fernandez, A., Insfran, E. and Abrahão, S. (2011a) 'Usability evaluation methods for the web: A systematic mapping study', *Information and Software Technology*, 53(8), pp. 789–817. Available at: <https://doi.org/10.1016/j.infsof.2011.02.007>.
- Fernandez, A., Insfran, E. and Abrahão, S. (2011b) 'Usability evaluation methods for the web: A systematic mapping study', *Information and Software Technology*, 53(8), pp. 789–817. Available at: <https://doi.org/10.1016/j.infsof.2011.02.007>.
- Francisco, L. and Benitti, F. B. V. (2014) 'Usability evaluation in Brazil: A systematic mapping', in *2014 9th Iberian Conference on Information Systems and Technologies (CISTI). 2014 9th Iberian Conference on Information Systems and Technologies (CISTI), Barcelona, Spain: IEEE*, pp. 1–7. Available at: <https://doi.org/10.1109/CISTI.2014.6877047>.
- Garcés, L. and Egas, L. M. (2013) 'Evolución de las Metodologías de desarrollo de la Ingeniería de software en el proceso la Ingeniería de Sistemas Software.', *Revista Científica y Tecnológica UPSE*, 1(3). Available at: <https://doi.org/10.26423/rctu.v1i3.29>.
- Garousi, V. et al. (2013) 'A systematic mapping study of web application testing', *Information and Software Technology*, 55(8), pp. 1374–1396. Available at: <https://doi.org/10.1016/j.infsof.2013.02.006>.
- Inan Nur, A., B. Santoso, H. and O. Hadi Putra, P. (2021) 'The Method and Metric of User Experience Evaluation: A Systematic Literature Review', in *2021 10th International Conference on Software and Computer Applications. ICSCA 2021: 2021 10th International Conference on Software and Computer Applications, Kuala Lumpur Malaysia: ACM*, pp. 307–317. Available at: <https://doi.org/10.1145/3457784.3457832>.

- Insfran, E. and Fernandez, A. (2008) 'A Systematic Review of Usability Evaluation in Web Development', in S. Hartmann, X. Zhou, and M. Kirchberg (eds) *Web Information Systems Engineering – WISE 2008 Workshops*. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 81–91. Available at: https://doi.org/10.1007/978-3-540-85200-1_10.
- Kitchenham, B. and Charters, S. (2007) 'Guidelines for performing Systematic Literature Reviews in Software Engineering'.
- Kitchenham, B. et al. (2010) 'Systematic literature reviews in software engineering – A tertiary study', *Information and Software Technology*, p. 14.
- Paz, F. and Jose Antonio Pow-Sang (2014) 'Current Trends in Usability Evaluation Methods: A Systematic Review', in 2014 7th International Conference on Advanced Software Engineering and Its Applications. 2014 7th International Conference on Advanced Software Engineering and Its Applications (ASEA), Hainan Island, China: IEEE, pp. 11–15. Available at: <https://doi.org/10.1109/ASEA.2014.10>.
- Paz, F. and Pow-Sang, J. A. (2015) 'Usability Evaluation Methods for Software Development: A Systematic Mapping Review', in 2015 8th International Conference on Advanced Software Engineering & Its Applications (ASEA). 2015 8th International Conference on Advanced Software Engineering & Its Applications (ASEA), Jeju: IEEE, pp. 1–4. Available at: <https://doi.org/10.1109/ASEA.2015.8>.
- Pellizon, L. H. et al. (2017) 'Software Analytics for Web Usability: A Systematic Mapping', in O. Gervasi et al. (eds) *Computational Science and Its Applications – ICCSA 2017*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 246–261. Available at: https://doi.org/10.1007/978-3-319-62407-5_17.
- Perurena Cancio, L. and Moráquez Bergues, M. (2013) 'Usabilidad de los sitios Web, los métodos y las técnicas para la evaluación', *Revista Cubana de Información en Ciencias de la Salud*, 24(2), pp. 176–194.
- Rivero, L., Barreto, R. and Conte, T. (2013) 'Characterizing Usability Inspection Methods through the Analysis of a Systematic Mapping Study Extension', *CLEI Electronic Journal*, 16(1). Available at: <https://doi.org/10.19153/cleiej.16.1.11>.
- Salvador, C., Nakasone, A. and Pow-Sang, J. A. (2014) 'A systematic review of usability techniques in agile methodologies', in *Proceedings of the 7th Euro American Conference on Telematics and Information Systems - EATIS '14*. the 7th Euro American Conference, Valparaiso, Chile: ACM Press, pp. 1–6. Available at: <https://doi.org/10.1145/2590651.2590668>.
- Ugras, T. et al. (2016) 'Research Trends in Web Site Usability: A Systematic Review', in A. Marcus (ed.) *Design, User Experience, and Usability: Design Thinking and Methods*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 517–528. Available at: https://doi.org/10.1007/978-3-319-40409-7_49.
- Yates, R. B. and Loaiza, C. R. (2003) 'Ubicuidad y usabilidad en la web', *Gerencia Tecnológica Informática*, 2(2), pp. 42–52.
- Yusop, N. S. M., Grundy, J. and Vasa, R. (2017) 'Reporting Usability Defects: A Systematic Literature Review', *IEEE Transactions on Software Engineering*, 43(9), pp. 848–867. Available at: <https://doi.org/10.1109/TSE.2016.2638427>.