International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 14, Number 9 (2021), pp. 947-952

Proposal of an Automated Tool for Conducting Usability Inspections based on Nielsen Heuristics

Gabriel Elías Chanchí Golondrino[1], Luis Freddy Muñoz Sanabria[2] and Wilmar Yesid Campo Muñoz[3]

[1]Ph.D., Facultad de Ingeniería, Programa de Ingeniería de Sistemas, Universidad de Cartagena, Colombia.

[2]Ph.D., Facultad de Ingeniería, Programa de Ingeniería de Sistemas, Fundación Universitaria de Popayán, Colombia.

[3]Ph.D., Facultad de Ingeniería, Programa de Ingeniería Electrónica, Universidad del Quindío, Colombia.

E-mail: [1]gchanchig@unicartagena.edu.co, [2] lfreddyms@fup.edu.co, [3] wycampo@uniquindio.edu.co

**Corresponding Autor (ORCID: 0000-0002-0257-1988),[2](ORCID:0000-0002-8172-0530),[3](ORCID: 0000-0001-8585-706X)*

Abstract

One of the most important aspects that contribute to the competitiveness of applications in the market is usability. Usability can be used in the software development process, both in the design phase and in the evaluation phase, where one of the most widespread methods to evaluate the usability of a software product are usability inspections. In order to specify the results obtained in a usability inspection and to determine the quantitative and qualitative level of compliance with the usability heuristics, it is necessary to specify the criteria associated with the different heuristics to be considered in an inspection. In this paper we propose as a contribution both the definition of a set of evaluation criteria associated with the Nielsen heuristics, and the development of a tool to automate the process of conducting usability inspections based on the defined criteria. The criteria and the proposed tool are intended to serve as a reference for conducting inspections on applications in different application contexts.

Keywords: Heuristic principles, inspection, Nielsen's heuristics, usability.

I. INTRODUCTION

With the increase in both the number of applications deployed in software repositories and the number of users who download and consume them, one of the key aspects to make applications competitive in the market is usability [1]–[3]. Usability is an attribute that defines software quality [4], [5] and according to ISO 9241 can be defined as the extent to which a software product can be used by specific users to achieve its objectives effectively, efficiently and with satisfaction in a specific context of use [6]–[9]. Similarly, according to ISO 9126, usability can be defined as the ability of a software product to be understood, learned, used and attractive to the user [5], [10].

One of the most widespread methods to evaluate the usability of a software product is the usability inspection, in which a group of evaluators verify compliance with a set of heuristics, so that as a result of this evaluation a set of usability problems associated with the heuristics considered are obtained [11]–[13]. In this sense, in a usability inspection, different types of heuristics can be taken into account depending on the context of the evaluated software, where the Nielsen heuristics are the most widespread [14]. In order to specify the results obtained in a usability inspection and to determine quantitatively and qualitatively the compliance with the usability heuristics, it is necessary to specify the criteria associated with the different heuristics considered in an inspection [2], [15]. In this sense, there is no evidence of the existence of specific evaluation criteria associated with the 10 usability heuristics proposed by Nielsen.

In this paper we proposed as a contribution, the definition of a set of evaluation criteria associated with the 10 Nielsen heuristics, as well as an automated tool for conducting usability inspections, based on the Nielsen heuristic principles and their associated criteria. The proposed tool allows the evaluators to rate the compliance with the criteria of each principle, as well as to perform an automatic statistical and graphic analysis of the evaluation results. The tool built was developed in Java language and has the advantage of being able to load the heuristics and usability criteria from a configuration file, thus enabling the scalability of the application in terms of the use of heuristics and usability criteria different from those of Nielsen. Both the criteria and the proposed tool are intended to serve as a contribution to the design and evaluation of general-purpose software from a usability perspective. Likewise, the defined criteria can be extrapolated or specified to a specific application context (web applications, mobile applications, video games, among others).

The rest of the paper is organized as follows: Section 2 presents the methodology considered for the development of this work. Section 3 shows the results obtained from the present research, which includes the definition of a set of evaluation criteria for each of the Nielsen heuristics, as well as the design and implementation of the proposed inspection tool. This section also presents the development of a case study based on the use of the proposed inspection tool. Finally, Section 4 presents the conclusions and future work derived from this research.

II. METHODOLOGY

For the development of this research, 4 methodological phases were defined: characterization of the Nielsen heuristics, definition of the usability criteria associated with each heuristic,

design and implementation of the automated tool and finally case study (see Figure 1).

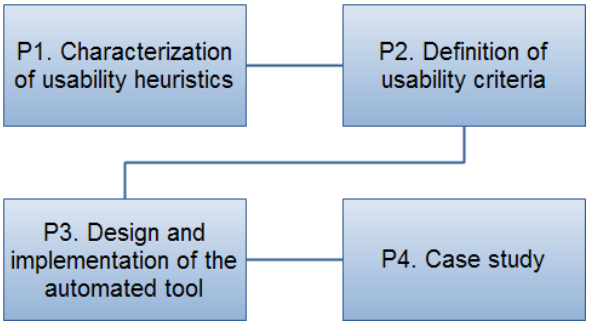

Fig. 1. Methodology considered

In phase 1 of the methodology, the 10 usability principles proposed by Nielsen were explored, in order to identify the objective and scope of each of these principles. Based on the exploration carried out in phase 1, within phase 2, a set of criteria were specified in terms of questions, which will be rated in the inspection by the evaluators on a scale of 1 to 5. Taking into account the defined criteria and the structure of a usability inspection, in phase 3 a tool was designed and implemented for conducting usability inspections based on Nielsen's heuristic principles, which allows the evaluation of the criteria defined for each heuristic, as well as the statistical and graphic analysis of compliance with Nielsen's usability principles. The proposed tool was developed in Java language and allows loading from a configuration file both the Nielsen heuristics and the criteria associated with each one of them. Finally, in phase 4 a case study was developed using the proposed tool, in which the usability of the tool for the creation of algorithms by means of flowcharts: DFD was evaluated.

III. RESULTS AND DISCUSSION

This section presents the results obtained in the development of this research, which includes the definition of a set of criteria associated with the 10 usability heuristics, as well as the design and implementation of the proposed tool and the development of a case study based on the use of the constructed tool. Thus, in the first instance, Table 1 presents a subset of the 55 usability criteria defined for the 10 Nielsen heuristics, which were expressed in terms of questions and seek to contribute to the determination of the scope of each of these heuristics. It is also worth mentioning that these criteria can be rated by the evaluators through the tool on a scale of 0 to 5, where 0 represents non-compliance with the criterion and 5 represents full compliance with the criterion.

Table 1. Example of defined criteria

Heuristic	Criteria
H1.Visibility of system status	To what degree does the system indicate the section where the user is located? To what degree does the system report on the current status of a process or task? To what extent are system response times reasonable and adequate?
H2. Match between system and the real world	To what extent does the system use words, phrases and concepts that are familiar to the user? To what extent are the different options included in the system presented in a simple, natural and intuitive manner? To what degree does the system use easy-to-understand icons that correspond to the user's reality?
H3. User control and freedom	To what extent does the system allow undoing the different actions performed by the user? To what extent when performing a task or process, is there a cancel option? To what extent does the system prevent automatic execution of actions that have not been initiated by the user?
H4. Consistency and standards	To what extent are the icons used similar in the system? To what extent is information structured and presented in the same way throughout the system? To what extent do the system controls and icons follow the established common standards?
H5.Error prevention	To what extent does the system ask for confirmation from the user before performing a critical action (cancel, delete, accept)? To what extent does the system make use of easy-to-understand messages that prevent possible errors? To what extent does the system use option lists or comboboxes as alternatives to manual data entry?
H6. Recognition rather than recall	To what extent are the main options and/or functions easy to find or always visible? To what extent do pre-filled data entry fields keep the information always remembered?

H7. Flexibility and efficiency of use	To what extent does the system have shortcuts to perform frequent tasks? To what extent does the system avoid requesting information from the user that has been previously provided? To what extent does the user perform tasks within the system in an adequate amount of time?
H8. Aesthetic and minimalist design	To what extent is the system interface simple and minimalistic? To what extent is the system not overloaded with information, options or other elements that distract the user? To what extent do the colors used in the system provide adequate contrast?
H9. Help users recognize, diagnose, and recover from errors	To what extent are the error messages understandable to the user and free of technical language? To what extent do the error messages guide the user to solve the problem? To what extent are error messages written without using hostile words?
H10. Help and documentation	To what extent does the system provide simple and clear help? To what extent does the help available in the system have a distinguishable structure? To what extent is the help available in the system contextual?

Once the criteria for each of the Nielsen heuristics were defined, the automated inspection tool was designed, taking into account the structure of a conventional usability inspection. Thus, Figure 2 shows a flow chart that includes the different processes that the proposed tool executes in the development of a usability inspection.

As shown in Figure 2, the tool first loads the Nielsen heuristics and the criteria associated to each one of them from a configuration file into the graphical interface, after which the evaluator proceeds to perform the inspection, rating the compliance of each one of these criteria in the software to be evaluated. Once the evaluation process of the heuristics has been completed, the tool automatically calculates the average of each heuristic based on the evaluator's ratings for each criterion and obtains the general average of the Nielsen heuristics. Likewise, the tool generates a set of charts that allow comparing the ratings assigned to each criterion and visualizing the percentage of compliance with the different heuristics in the software evaluated through the proposed inspection tool.

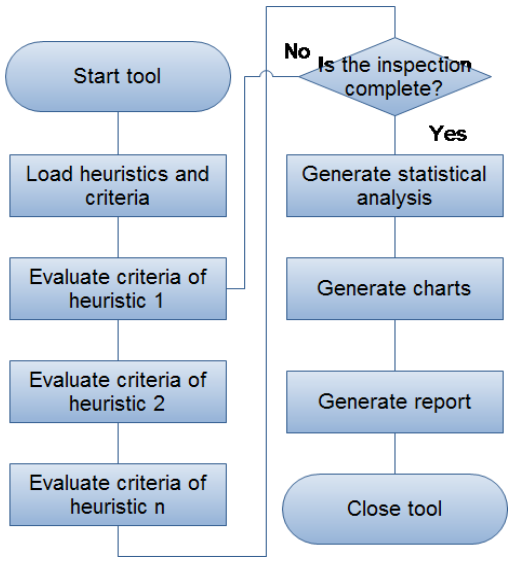

Fig. 2. Tool flow diagram

Taking into account the different processes presented in Figure 2, an inspection tool was implemented in Java language, which consists of 14 tabs, as shown in Figure 3.

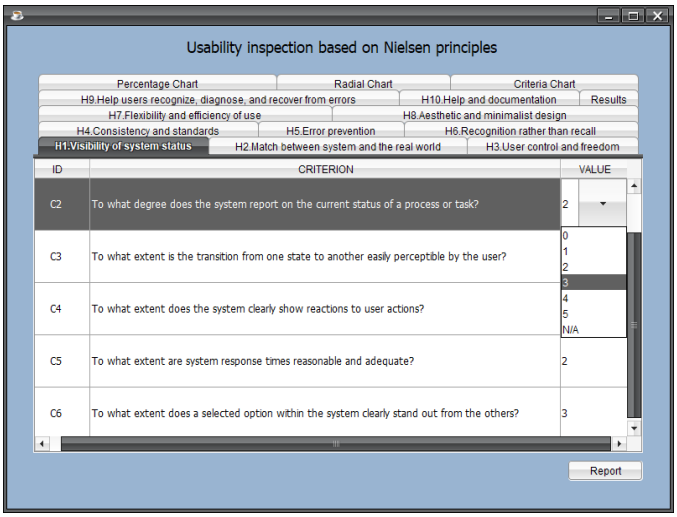

Fig. 3. Interface of the proposed tool

The first 10 tabs of the tool include the criteria associated with the 10 Nielsen heuristics, which are loaded from a configuration file and allow the evaluator to rate these criteria on a scale of 0 to 5. Through tab 11 or the "Results" tab, it is possible to visualize the average obtained in each of the Nielsen heuristics, as well as the total average of the heuristics, as shown in Figure 4. Thus, as an example, from the hypothetical data filled in the first 10 tabs, it can be observed that the heuristic principle H1: "Visibility of system status" obtained an average of 2.0, while the heuristic principle H4: "Consistency and standards" obtained an average of 2.429. Similarly, in the "Results" tab, by clicking on the "Report" button, it is possible to generate a report in a .CSV file with the average of the criteria associated with each of the Nielsen heuristics.

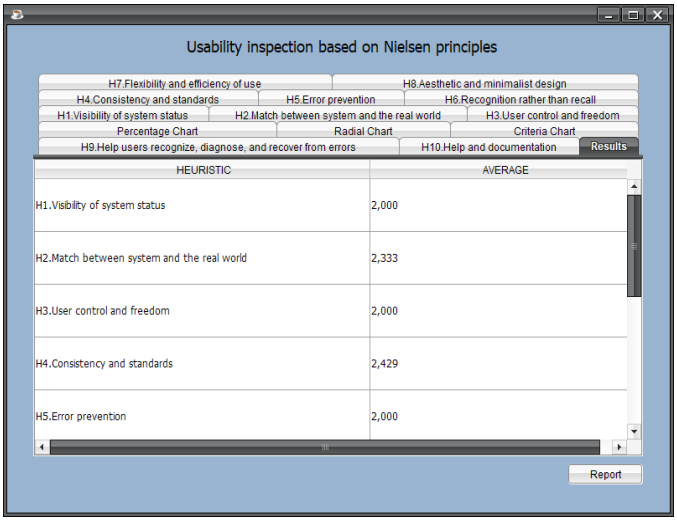

Fig. 4. "Results" tab of the proposed tool

On the other hand, in tabs 12, 13 and 14 it is possible to visualize respectively: a bar chart with the percentage of compliance of the different heuristics (tab "Percentage Chart"), a radial chart showing the average value obtained in each of the heuristics evaluated (tab "Radial Chart") and finally an area chart showing for each heuristic the rating assigned to each criterion by the evaluator (tab "Criteria Chart"), as shown in Figure 5. Thus, as an example, Figure 5 shows in a radial chart the average ratings obtained for the different heuristics, based on the hypothetical data filled in the first 10 tabs.

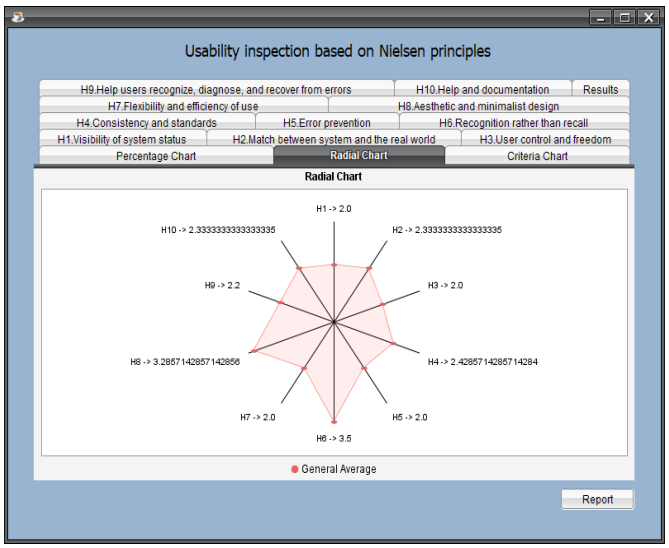

Fig. 5. "Radial Chart" tab of the proposed tool

Similarly, Figure 6 shows, through an area chart, the hypothetical ratings assigned to the 6 criteria defined and corresponding to heuristic H1: "Visibility of system status".

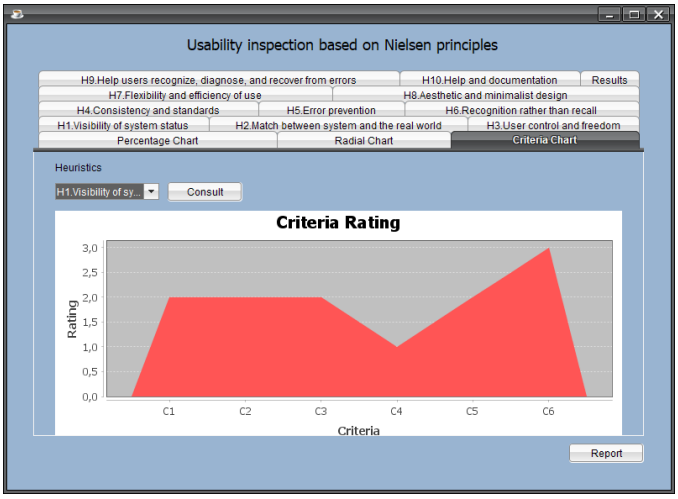

Fig. 6. "Criteria Chart" tab of the proposed tool

Taking into account the tool built and presented in Figures 3,4,5 and 6, a case study was developed on the flowchart algorithm generation tool: DFD (see Figure 7), in order to verify the usefulness of the proposed tool.

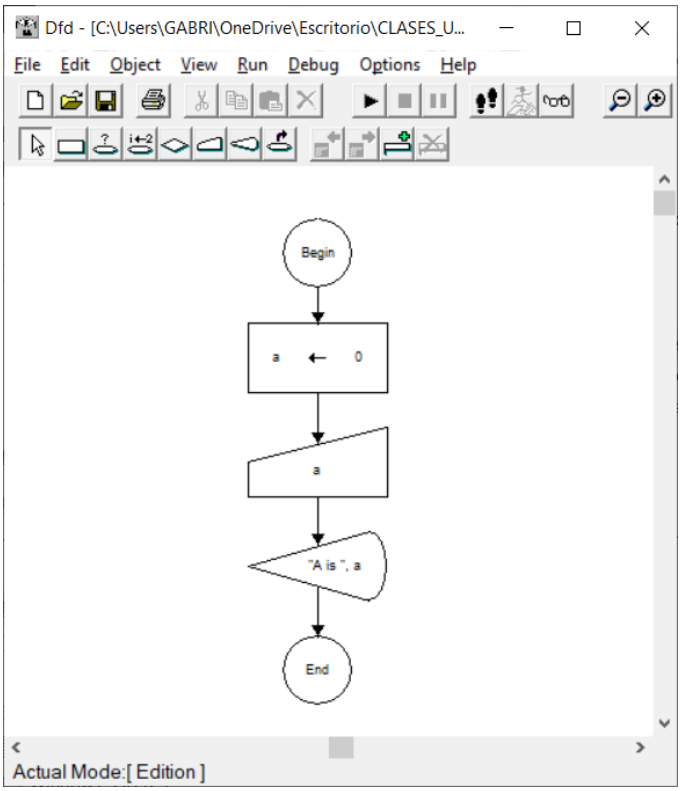

Fig. 7. DFD Tool

Once the usability inspection of the DFD tool was carried out by the researchers of this work through the use of the proposed tool, the results presented in Table 2 were obtained, which shows the average rating of the criteria associated with the 10 Nielsen heuristics.

Table 2. Results of the case study inspection

Heuristic	Average
H1.Visibility of system status	4.667
H2. Match between system and the real world	3.333
H3. User control and freedom	0.667
H4. Consistency and standards	3.429
H5.Error prevention	2.667
H6. Recognition rather than recall	3
H7. Flexibility and efficiency of use	3.667
H8. Aesthetic and minimalist design	4.143
H9. Help users recognize, diagnose, and recover from errors	4.2
H10. Help and documentation	0.250
Total Average	3.002

From the results presented in Table 2, it can be seen how the tool obtained the best averages in the heuristics: H1: "Visibility of system status" and H8: "Aesthetic and minimalist design", considering that the DFD tool offers feedback to the different actions performed by the user and has a simple and minimalist interface. Similarly, it is possible to see how the tool obtained the worst averages in the heuristics: H10: "Help and documentation" and H3: "User control and freedom", considering that the tool does not include a help option to guide the user in the creation of flowcharts and does not have an option that allows the user to undo the actions performed in the process of creating the algorithms. On the other hand, it is possible to observe from Table 1, how the total average of the heuristics in the evaluated tool is 3.002, which corresponds to an acceptable level, so taking into account the criteria that obtained a low score in the inspection, it is possible to increase the usability of the tool to a good or excellent level. The results presented in Table 1 can be seen more clearly in the percentage graph generated by the inspection tool (see Figure 8), which shows the percentage of compliance associated with each of the 10 Nielsen heuristic principles. Figure 8 shows how 6 of the heuristics (H1, H2, H4, H7, H8, H9) have a percentage of compliance greater than or equal to 60%, 2 of the heuristics (H5, H6) have a percentage of compliance between 50% and 60%, while the remaining 2 heuristics (H3, H10) have a percentage of compliance lower than 20%. Finally, if the compliance percentages of the 10 heuristics are averaged, an overall compliance percentage of 60% is obtained for the DFD tool, which corresponds to an acceptable level of usability.

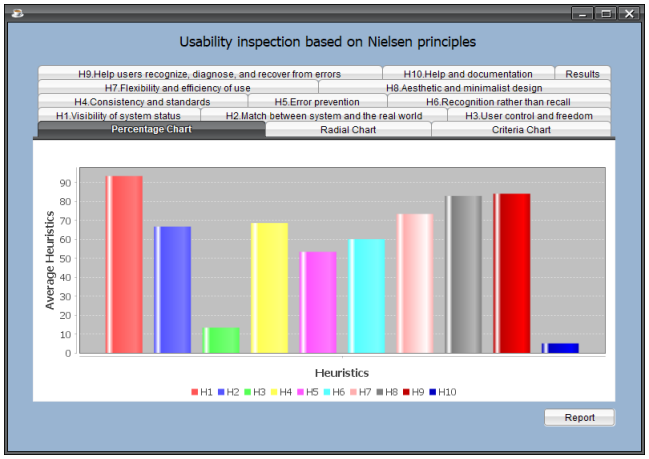

Fig. 8. Percentage of compliance with heuristics

Finally, Table 3 presents a set of recommendations that can be included in the DFD tool, which were extracted from the lowest rated criteria.

Table 3. Recommendations obtained

Heuristic	Recommendation
H2	It is recommended to use conventional icons for the creation of flowcharts, since these icons do not correspond to the symbols established to represent a flowchart.
H3	It is recommended to improve the understanding of the menu icons by incorporating the conventional symbols used in the structuring of a flowchart.
H4	Se recomienda mejorar la comprensión de los iconos del menú, a partir de la incorporación de los símbolos convencionales usados en la estructuración de un diagrama de flujo.
H5	Although the error messages are free of technicalities, the accuracy of the messages used in the tool's dialog boxes can be improved. It is recommended that the fields used in the structuring of a flowchart include default values to improve the understanding of the tool.
H7	It is recommended to give the user the possibility to customize the tool interface (colors, font size, etc.).
H8	It is recommended to avoid the use of abstract icons and replace them with the conventional icons used in the creation of flowcharts.
H9	It is recommended to present more precise error messages to guide the user in troubleshooting.
H10	It is recommended to include help within the tool and to include context-sensitive help functionality.

VI. CONCLUSIONS

In this paper we proposed as a contribution the specification of a set of usability criteria, associated to Nielsen's 10 heuristics, which can be used both in the design and evaluation of interactive applications. Similarly, we proposed as a contribution the development of an automated tool for conducting usability inspections based on the defined criteria.

The tool proposed in this paper has the advantage of allowing the evaluators to rate the criteria associated with the Nielsen heuristics, as well as to obtain statistical and graphic analysis of the results obtained in the evaluation. Similarly, the tool allows loading heuristics and criteria from a configuration file, so it can be adapted to work with usability principles different from those of Nielsen.

The case study developed allowed to demonstrate the usefulness of the inspection tool, in order to determine the usability aspects to be improved in the DFD tool. In this sense, the inspection based on the defined criteria made it possible to determine more precisely the different elements of the DFD tool interface that should be prioritized to improve its usability.

As a future work derived from the present research, it is intended to include in the inspection tool the functionality to generate automatic recommendations based on the criteria that obtained a worse rating by the evaluators.

REFERENCES

[1] K. Radle and S. Young, "Partnering usability with development: how three organizations succeeded," *IEEE Softw.*, vol. 18, no. 1, pp. 38–45, Jan. 2001, doi: 10.1109/52.903164.

[2] G. E. Chanchí, M. C. Gómez Álvarez, and W. Y. Campo Muñoz, "Criterios de usabilidad para el diseño e implementación de videojuegos," *Rev. Ibérica Sist. e Tecnol. Informação*, no. E26, pp. 461–474, 2020.

[3] D. Hering, T. Schwartz, A. Boden, and V. Wulf, "Integrating usability-engineering into the software developing processes of SME: A case study of software developing SME in Germany," in *Proceedings - 8th International Workshop on Cooperative and Human Aspects of Software Engineering, CHASE 2015*, Jul. 2015, pp. 121–122, doi: 10.1109/CHASE.2015.22.

[4] M. A. Kabir, M. U. Rehman, and S. I. Majumdar, "An analytical and comparative study of software usability quality factors," *Proc. IEEE Int. Conf. Softw. Eng. Serv. Sci. ICSESS*, vol. 0, pp. 800–803, Jul. 2016, doi: 10.1109/ICSESS.2016.7883188.

[5] C. Santos, T. Novais, M. Ferreira, C. Albuquerque, I. H. De Farias, and A. P. C. Furtado, "Metrics focused on usability ISO 9126 based," *Iber. Conf. Inf. Syst. Technol. Cist.*, vol. 2016-July, Jul. 2016, doi: 10.1109/CISTI.2016.7521437.

[6] J. L. Gozález-Sánchez, F. Montero-Simarro, and F. L. Gutiérrez-Vela, "Evolución del concepto de usabilidad como indicador de calidad del software. José-Luis González-Sánchez, Francisco Montero-Simarro, Francisco-Luis Gutiérrez-Vela," *Rev. Int. Inf. Doc. Bibl. y Comun.*, vol. 21, no. 5, pp. 529–536, 2012, Accessed: Oct. 17, 2020. [Online]. Available: http://www.elprofesionaldelainformacion.com/contenidos/2012/septiembre/13.html.

[7] V. F. Martins, M. De Paiva Guimaraes, and A. G. Correa, "Usability test for Augmented Reality applications," 2013, doi: 10.1109/CLEI.2013.6670668.

[8] K. Finstad, "The usability metric for user experience," *Interact. Comput.*, vol. 22, no. 5, pp. 323–327, 2010, doi: 10.1016/J.INTCOM.2010.04.004.

[9] G. E. Chanchi, M. A. Ospina, and M. E. Monroy, "Aplicación de la lógica difusa en el análisis de inspecciones heurísticas de usabilidad," *Rev. Espac.*, vol. 41, no. 27, pp. 159–173, 2020, Accessed: Aug. 08, 2020. [Online]. Available: https://www.revistaespacios.com.

[10] S. Rochimah, H. I. Rahmani, and U. L. Yuhana, "Usability characteristic evaluation on administration module of Academic Information System using ISO/IEC 9126 quality model," *2015 Int. Semin. Intell. Technol. Its Appl. ISITIA 2015 - Proceeding*, pp. 363–368, Aug. 2015, doi: 10.1109/ISITIA.2015.7220007.

[11] V. F. Martins, A. G. D. Corrêa, and M. De Paiva Guimarães, "Inspeção de Usabilidade de uma Smart TV," 2014, doi: 10.1109/CISTI.2014.6877010.

[12] N. M. C. Valentim and T. Conte, "Improving a usability inspection technique based on quantitative and qualitative analysis," *Proc. - 28th Brazilian Symp. Softw. Eng. SBES 2014*, pp. 171–180, Oct. 2014, doi: 10.1109/SBES.2014.23.

[13] J. G. Enriquez and S. Casas, "Usabilidad en aplicaciones móviles," *Inf. Científico Técnico UNPA*, vol. 5, no. 2, pp. 25–47, 2013, Accessed: Sep. 08, 2020. [Online]. Available: https://dialnet.unirioja.es/servlet/articulo?codigo=5123524&info=resumen&idioma=SPA.

[14] F. Paz, F. A. Paz, J. A. Pow-Sang, and L. Collantes, "Usability heuristics for transactional web sites," *ITNG 2014 - Proc. 11th Int. Conf. Inf. Technol. New Gener.*, pp. 627–628, 2014, doi: 10.1109/ITNG.2014.81.

[15] G. Chanchí, M. C. Gómez Álvarez, and W. Y. Campo, "Propuesta de una herramienta de inspección según los atributos de usabilidad de Nielsen," *Rev. Ibérica Sist. e Tecnol. Informação*, no. E26, pp. 448–460, 2020.