Development and Preliminary Evaluation of a Technology for Assessing Hedonic Aspects of UX in Text-Based Chatbots

Pamella A. de L. Mariano[1][a], Ana Paula Chaves[2][b] and Natasha M. C. Valentim[1][c]
[1]*Federal University of Paraná, Curitiba, Brazil*
[2]*Northern Arizona University, Flagstaff, U.S.A.*

Keywords: Chatbot, Evaluation, Hedonic Aspects, User Experience.

Abstract: For text-based chatbots to achieve a desired level of quality, it is essential to evaluate their performance, particularly focusing on the hedonic aspects of User Experience (UX), which is a crucial quality attribute. A comprehensive evaluation must consider the specifics of the context being assessed. A Systematic Mapping Study (SMS) revealed that no existing UX evaluation technologies address the hedonic aspects of UX in text-based chatbots. The Guidelines to Assess Hedonic Aspects in Chatbots (GAHAC) was developed to address this gap. The guidelines were formulated by selecting and evaluating the hedonic aspects of UX and the evaluation technologies identified in the SMS. Relevant questions from these technologies were filtered and adapted to the context of text-based chatbots. GAHAC aims to provide a context-specific evaluation technology in the form of guidelines encompassing the hedonic aspects of UX. Its primary contribution is providing a structured and accessible method for evaluating hedonic aspects, which have been largely overlooked in UX studies of text-based chatbots. This enables developers and researchers to qualitatively identify opportunities to improve UX in chatbot interactions. A preliminary evaluation conducted with two Human-Computer Interaction experts led to refinements in the guidelines. By offering a dedicated UX evaluation technology for text-based chatbots, GAHAC contributes to improving the quality of such systems.

1 INTRODUCTION

Chatbots are online conversational systems that enable interaction between humans and computers using natural language by text or voice (Ruane et al., 2021; Jia and Jyou, 2021; Veglis et al., 2019). Unlike voice-activated intelligent assistants, they are generally text-based and may include additional interactions, such as point-and-click (Candello and Pinhanez, 2016). Powered by Artificial Intelligence, chatbots can mimic human chat or perform specific tasks, as demonstrated in their use in financial institutions for credit analysis and customer service (Mudofi and Yuspin, 2022).

Chatbots serve multiple purposes and are increasingly used as information providers across various sectors. For example, chatbots are widely used in like education, financial systems, healthcare, and e-commerce. The rise of large language models, such as ChatGPT, has further enhanced the conversational abilities of chatbots, making them more engaging and capable of handling complex interactions and contributing to the chatbot's growing effectiveness in various contexts (Brown et al., 2020). As information systems, chatbots have transformed how industries interact with users, providing efficient, personalized, and scalable solutions.

The interaction between humans and computational systems is crucial for improving the integration of technology into users' daily lives. This is especially important when considering chatbots, as they are increasingly used in various sectors. Ensuring the quality of chatbots is essential for delivering consistent, positive experiences, and continuous evaluation of their performance is necessary. Systems, including chatbots, should be designed to improve users' lives, promote well-being, and meet their needs, which underscores the importance of quality assessment in achieving these objectives (Barbosa et al., 2021).

The model for attractive software systems with a good user experience (UX), proposed by Hassenzahl et al. (2000), divides software quality into two main areas: pragmatic quality, which focuses on usability

[a] https://orcid.org/0009-0006-9919-3448
[b] https://orcid.org/0000-0002-2307-3099
[c] https://orcid.org/0000-0002-6027-3452

Mariano, P. A. L., Chaves, A. P. and Valentim, N. M. C.
Development and Preliminary Evaluation of a Technology for Assessing Hedonic Aspects of UX in Text-Based Chatbots.
DOI: 10.5220/0013289200003929

In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025) - Volume 2*, pages 544-551
ISBN: 978-989-758-749-8; ISSN: 2184-4992

and efficiency, and hedonic quality, which looks at originality, innovation, and aesthetics. The study of Hassenzahl et al. (2000) emphasizes hedonic quality because it relates to the emotional well-being of the user. It recognizes that products satisfying these needs can enhance pleasure and customer loyalty, beyond providing satisfaction Chitturi et al. (2008). Any Information System, such as text-based chatbots, must meet a specific quality standard to be useful to the user. In this regard, Usability and User Experience (UX) evaluations provide essential support to ensure the system's quality (Madan and Kumar, 2012).

Previous studies revealed that no technologies have been explicitly developed for the context of hedonic aspects in text-based chatbots. A two-phased Systematic Mapping Study (SMS) with 52 primary studies categorized the technologies used to evaluate the UX of text-based chatbots (De Souza et al., 2024; Mariano et al., 2024). The study showed that most technologies are designed to evaluate systems, including Information Systems (IS), without considering their specificities. Additionally, the SMS showed that few technologies simultaneously extract and evaluate qualitative data or hedonic aspects of UX.

This paper presents the development of Guidelines to Assess Hedonic Aspects in Chatbots (GAHAC) designed to evaluate text-based chatbots. GAHAC involves three steps: a) the evaluator interacts with the chatbot to perform tasks; b) analyzes the guidelines; and c) documents hedonic issues in a spreadsheet, noting the guideline number that reveals the problem. It can be used during the refinement or design phases, allowing developers to improve prototypes and researchers to adapt the guidelines to specific contexts. The first version of GAHAC includes 75 guidelines in 31 hedonic aspects.

Moreover, two human-computer interaction (HCI) experts conducted a preliminary evaluation of GAHAC, providing valuable feedback on the technology's characteristics. They recommended making the guidelines available as an online platform, including an analysis of "AI hallucinations", tailoring the language for evaluators, and considering cultural differences. They also raised concerns about whether the technology covers all relevant aspects, particularly those not yet addressed in the literature. The results advance the state of the art by introducing an evaluation technology that extracts and evaluates qualitative data and focuses on the hedonic aspects of UX.

This study is included in the socio-technical view of ISs because it considers a better way to evaluate chatbots. In this way, it is not just the use of text-based chatbots, but the relationship between users and this system type, and how these relationships affect the user experience, leading to how the user perceives this interaction. Therefore, thinking of better ways to deliver an IS with value and quality to users (Kujala and Vännänen-Vainio-Mattila, 2009), to improve this interaction, making it more pleasant and satisfying is a way of contributing to the socio-technical view.

2 BACKGROUND

A chatbot, also known as an intelligent bot or digital assistant, is a computer program that uses artificial intelligence and Natural Language Processing to respond to text or voice conversations . The primary motivation for using chatbots is productivity, but they also serve as tools for entertainment, socialization, and novelty (Adamopoulou and Moussiades, 2020). Since early 2020, the AI chatbot market has grown rapidly with the launch of Bard and ChatGPT, which utilize the Transformer Neural Network Architecture to process language effectively, learning patterns from large textual datasets (Al-Amin et al., 2024).

According to Rapp et al. (2021), chatbots can be classified into 4 different types based on their orientation and purpose. Task-oriented chatbots are designed to help users perform specific tasks or solve problems efficiently. In contrast, conversational chatbots focus on providing natural interactions, aiming to maintain high-quality conversations and often establishing a form of relationship with users. There is also a combination of both types (conversation and task-oriented chatbots), where the chatbot seeks to balance task execution and conversational flow. Besides, the chatbot may have an undefined type, without a clear orientation between these three approaches.

For a system to be truly useful to the user, it needs to reach a good level of quality. Usability and UX are good indicators of IS quality (McNamara and Kirakowski, 2006). UX evaluation methods analyze how users interact with existing concepts, design details, prototypes, or final products to understand their experience. These methods assess user interactions and feelings toward the product rather than simply measuring task performance. Since UX is subjective, traditional objective metrics, like task completion time, may not adequately capture the user experience. Instead, UX evaluation methods investigate various subjective qualities, considering factors such as user motivation and expectations. The goal is to ensure that the final product aligns with intended user experience objectives and guides design decisions (Vermeeren et al., 2010).

This work adopts the user experience definition from ISO 9241-210 (2010), which considers the per-

ceptions and responses of a person when using or anticipating the use of a product, system, or service. The approach focuses on the hedonic aspects of experience, recognizing that while pragmatic aspects are important, it is the hedonic attributes that most influence product acceptance (Merčun and Žumer, 2017). This implies that user experience goes beyond functionality, involving emotional and aesthetic elements. Hassenzahl (2018) detail that ergonomic attributes are now seen as pragmatic, while hedonic ones emphasize psychological well-being and have greater pleasure potential. The hedonic function can be subdivided into providing stimulation, communicating identity, and provoking valuable memories Hassenzahl (2018). In summary, a product can be seen as pragmatic in terms of its efficiency and as hedonic for its emotional and memorable effects.

3 RELATED WORK

Guerino and Valentim (2020) investigate conversational systems that use human voice to perform actions. The study identified 31 assessment technologies focused on usability and user experience (UX) in chatbots. The searches were conducted in the following virtual libraries: Scopus, IEEEXplore, ACM Digital Library, and Engineering Village. The results indicate that the assessment technologies were mostly created specifically for the studies, without empirical evaluation. In addition, most identified chatbots focus on assisting users with daily tasks.

Tubin et al. (2022) examined methods for evaluating the experience with conversational agents to provide more realistic and natural user interaction. The study identified how UX is measured during interactions with these agents. According to the authors, it is essential to evaluate user experience at various stages and apply combined methods to gain insight into aspects such as participants' feelings and behaviors.

Mafra et al. (2024) developed the U2Chatbot inspection checklist, a tool designed to evaluate and detect defects in text-based chatbots, consisting of 107 items and covering a wide range of quality attributes related to usability and UX. Its goal is to provide a more comprehensive assessment than existing tools, ensuring that critical factors affecting chatbot performance are thoroughly addressed.

The technologies identified in the above works extract important hedonic aspects of UX. However, none of the technologies are focused on evaluating the hedonic aspects of text-based chatbots. None of these technologies evaluates a comprehensive set of hedonic aspects and few extract qualitative data. GAHAC aims to fill this gap because it is a technology that helps identify hedonic problems in text chatbots.

4 METHODOLOGY

The methodology for developing GAHAC is divided into 3 phases: (1) the Systematic Mapping Study (SMS), (2) the GAHAC initial definition, and (3) the GAHAC evaluation, as depicted in Figure 1.

In the first phase, we conducted a two-phased Systematic Mapping Study (SMS) (De Souza et al., 2024; Mariano et al., 2024) to analyze and identify research gaps on hedonic aspects in text-based chatbots (Figure 1, Activity 1.1 and 1.2). Through the SMS, we identified 69 evaluation technologies to assess hedonic aspects of UX in text-based chatbots. Among these technologies, most are based on questionnaires and interviews, showing little diversity in formats. The literature lacks empirical studies to evaluate the reliability and consistency of these technologies, which has important implications for the validity of their results. This trend aligns with the findings of Tubin et al. (2022), who noted the frequent use of questionnaires created by the authors for evaluating a specific study. However, only 30% of the technologies are specifically geared toward text-based chatbots, revealing a gap in specificity in the assessments. In comparison to the results of Guerino and Valentim (2020), who found a balance between specific and non-specific technologies for conversational systems, this low specificity is noteworthy. Another important finding is that most evaluation methods employ quantitative approaches, which may limit the depth of the analysis by hindering a detailed view of the user's experience. Additionally, current technologies fail to address the unique characteristics of human-chatbot interaction, such as identity and social interaction, which are essential for a comprehensive evaluation of user experience. This point is emphasized in the study by Ren et al. (2019), which highlights the importance of including the context of use in UX evaluations for chatbots and considering the situation in which the system will be applied.

In the second phase of the methodology, we defined the GAHAC technology based on the findings from the SMS. The GAHAC is structured as a set of guidelines, serving as an investigative technique comprising 31 hedonic aspects and 75 guidelines organized within those aspects.

Activity 2.1 (Figure 1) consisted of synthesizing the hedonic aspects identified in the SMS conducted by De Souza et al. (2024) and its extension by Mariano et al. (2024). The first author was responsible for

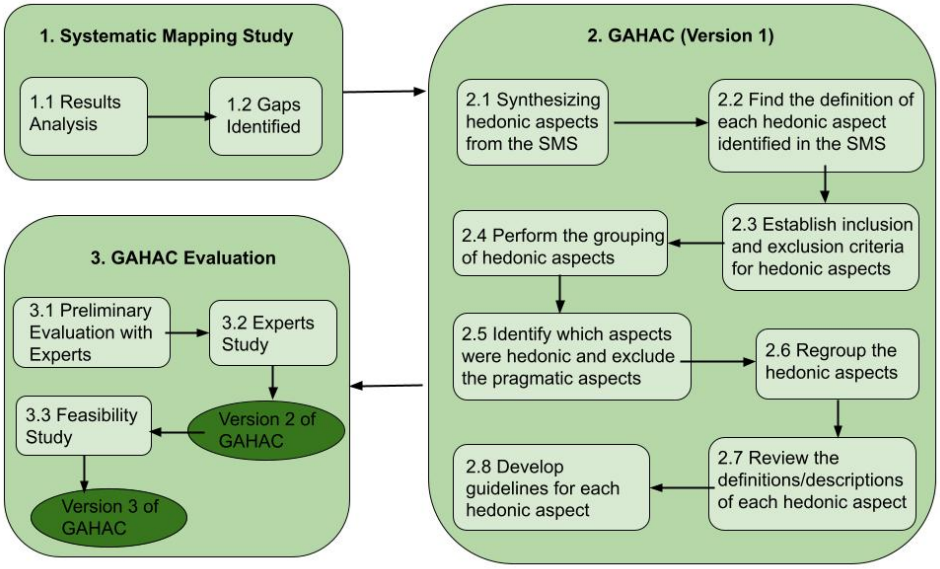

Figure 1: Flowchart of the GAHAC development process.

carrying out this synthesis. This activity resulted in a list of unique hedonic aspects extracted from the literature that guided the subsequent stages and formed the basis of the entire GAHAC technology.

The initial list of hedonic aspects was organized into a spreadsheet, including references, definitions, and a column named "match with" where we listed similar aspects. This process facilitated comparative analysis across studies and technologies, resulting in a detailed organization of the hedonic aspects. We replicated aspects that appear repeatedly across various technologies and studies, enabling us to analyze each aspect according to its application in each context. Therefore, the initial number of aspects exceeds the 188 consolidated ones.

Activity 2.2 (Figure 1) focused on defining each hedonic aspect identified in Activity 2.1. The definitions were extracted from the SMS articles De Souza et al. (2024); Mariano et al. (2024), by collecting definitions mentioned in the articles explicitly or extracting and interpreting explanations of how the aspect is used in the study. Three researchers (one HCI expert and one chatbot expert) gathered weekly to debate and refine the definitions. To ensure the integrity of the process, they scrutinized, discussed, and ultimately agreed upon each definition through consensus. This approach ensured a comprehensive understanding of the hedonic aspects. 37 aspects were removed from the spreadsheet due to lack of definition.

To ensure the quality of the definitions, in Activity 2.3 (Figure 1), we established the inclusion and exclusion criteria for the listed hedonic aspects. We excluded hedonic aspects that did not present clear definitions or elucidative sentences while including those with relevant definitions. For example, we excluded *Affection* because the study where it appears does not provide a clear definition for the term and the aspect does not have a well-established definition in the literature. Likewise, *emotional support* was excluded due to the lack of information about the interview used in the study where the aspect is listed.

The methodological rigor adopted in this research was crucial for consolidating the relevant hedonic aspects. In Activity 2.4 (Figure 1), we grouped the hedonic aspects from different studies into a single construct based on similarities in their definition, removing duplicates from the list. For example, the hedonic aspect of *trust* appeared in multiple studies, such as Fadhil et al. (2018) and Jin et al. (2019), which used questionnaires to assess *trust*, demonstrating the variety of approaches to the same aspect. The initial grouping resulted in 26 remaining aspects.

In Activity 2.5 (Figure 1), we conducted a thorough analysis to exclude pragmatic aspects that remained on the list. For example, after analyzing the definition of *satisfaction*, the aspect was considered pragmatic, following the usability definition proposed by ISO (2018). We removed 121 aspects from the spreadsheet for being pragmatic.

Activity 2.6 (Figure 1) involved reviewing the definitions and regrouping the aspects whenever necessary. For example, in Fadhil et al. (2018), *trust* was evaluated with the item "The agent asked very personal questions. I found the questions very intrusive." After analyzing this item, we considered it aligned with *intimacy* more than *trust*. Therefore, we moved

the item from *trust* to *intimacy*. In this process, 177 aspects were completely merged with other aspects or separated into new constructs, resulting in the final list of hedonic aspects we included in GAHAC.

In summary, the initial list of 188 aspects extracted from the SMS was reduced to 31 hedonic aspects that comprise our proposed technology. We systematically evaluated all the aspects and their definitions, ensuring that the aspects listed in the GAHAC represent UX hedonic aspects for text-based chatbots. Additionally, we ensured that the list provides a comprehensive set of items and definitions for these aspects, as found in the literature.

Given the comprehensive list of hedonic aspects we produced in the previous steps, we developed the guidelines for evaluating each aspect using the technology GAHAC. These guidelines were based on the definitions and purpose we found in the literature for each aspect, ensuring that the characteristics of interaction with text-based chatbots are considered to maximize user experience.

Guidelines are general rules commonly observed when designing and evaluating interfaces, based on empirical and theoretical knowledge, aimed at enhancing user experience and usability . The effective application of guidelines depends on the designer's understanding of the domain and the users, requiring a careful assessment of the most suitable guidelines for each design situation (Barbosa et al., 2021).

We followed a detailed methodological process to develop the GAHAC. Activity 2.7 (Figure 1) involved a thorough reading of the definitions we selected for each hedonic aspect. This allowed us to understand the essential elements and characteristics of each aspect. This step was fundamental for tracing how these aspects were being analyzed in each of the technologies identified in the SMS (part 1 and part 2). Thus, the reading was crucial to ensure the guidelines were comprehensive, reflecting the analyses conducted on the different technologies.

Activity 2.8 (Figure 1) consisted in formulating guidelines to measure each hedonic aspect in text-based chatbots. Based on the definitions and descriptions we analyzed in Activity 2.7, we wrote an initial list of guidelines to capture and evaluate the essential elements of each aspect. These guidelines were carefully formulated to ensure that all identified hedonic aspects were adequately measured and analyzed.

In general, we created the GAHAC by conducting reading sessions of the descriptions/definitions of each of the grouped hedonic aspects to understand their meaning and goal. Subsequently, we developed the guidelines based on these definitions to evaluate each identified aspect (Figure 1 Activity 2.8).

We conducted a preliminary evaluation of GAHAC (**3rd phase of the methodology**, Figure 1 - Activity 3.1) with two HCI experts (E1 and E2) with over a decade of experience in their fields.

We introduced the GAHAC to the experts, along with the theoretical background and methodology used for its development. They were given approximately one month to conduct their analysis, after which they provided feedback, and asked clarifying questions about the technology. The feedback was analyzed without a specific process, allowing each author to interpret the results in their own way. Revisions of the GAHAC (versions two and three) will occur after two additional studies to further evaluate the GAHAC: an expert study (Figure 1 - Activity 3.2) and a feasibility study (Figure 1 - Activity 3.3). After analyzing each of these studies, the identified improvements will be implemented in GAHAC, resulting in new versions.

5 GAHAC

GAHAC is designed to evaluate the hedonic aspects of UX in text-based chatbots. It consists of guidelines to extract qualitative data related to the user experience (UX) in Human-Computer Interaction (HCI). These guidelines can be used by evaluators, HCI specialists, chatbot developers, and researchers interested in investigating the hedonic aspects addressed by this technology. Using GAHAC does not require training, but only a solid understanding of the evaluation technology and the assessed chatbot.

Developers and evaluators can benefit from applying the guidelines during the chatbot's refinement stages, ideally between version releases to collect data on evaluators' reactions, helping to pinpoint specific strengths and weaknesses of the chatbot. During the design phase, applying GAHAC allows evaluators to analyze a chatbot's prototype and identify issues before full development.

Researchers can apply GAHAC at any stage and to any chatbot they wish to evaluate. The technology allows customization by omitting hedonic aspects irrelevant to the application context. For instance, a medical appointment scheduling chatbot may not require the entertainment aspect, while banking chatbots might omit novelty. GAHAC contains 75 guidelines grouped into 31 hedonic aspects: *Adaptability*, *Language Use*, *Anger*, *Motivation*, *Attention*, *Novelty*, *Attractiveness*, *Personality*, *Comfort*, *Privacy*, *Competence*, *Reliability*, *Conversation Flow*, *Sadness*, *Disgust*, *Satisfaction*, *Emotions*, *Social Influence*, *Engagement*, *Social Presence*, *Pleasure*, *Surprise*, *Ex-*

pectation, *Trust*, *Fear*, *User Control*, *Generic UX*, *Happiness*, *Humanity*, *Intention to Use*, and *Intimacy*. The full GAHAC version is here [1].

The application of GAHAC involves three main steps: a) the evaluator interacts with the chatbot, performing predefined tasks; b) the evaluator reviews and analyzes the GAHAC guidelines; and c) if a hedonic issue is identified, the evaluator documents it in a spreadsheet, noting the corresponding GAHAC guideline that facilitated the identification.

To illustrate how GAHAC works, we selected ChatGPT in its free version and applied specific guidelines to identify potential hedonic issues in the user experience with the chatbot. ChatGPT was chosen due to its widespread adoption, evidenced by its 200 million active users as of November 2024, making it one of the most utilized AI tools globally (Demand Sage, 2024). The analysis focused on three main aspects: privacy, emotions, and language use, all of which compromise the user's hedonic experience during the interaction.

For the privacy aspect, we used guideline 57, which asks if the user felt their privacy might be violated after interacting with the chatbot. We noticed that ChatGPT creates a sense of vulnerability by lacking clear communication about how the user data provided is handled. Although privacy policies are available, the interaction does not actively reinforce them. This can lead to the perception that the information they provide may be shared and used to train the chatbot, which may violate our privacy expectations.

Regarding the emotion aspect, we applied Guideline 15 to assess whether the user felt an emotional connection with the chatbot. We found that interactions with ChatGPT are impersonal. While the chatbot does not provoke negative feelings or disrespect, it also fails to establish an emotional bond. This happens due to its neutral tone and lack of features that foster intimacy or emotional engagement.

Finally, we examined the language use aspect using guideline 45, which suggests checking if the chatbot uses everyday language. We observed that, in several instances, ChatGPT employs formal or technical language, which may feel distant from common vocabulary. This choice of language creates a sense of detachment and reduces the fluidity of the interaction. These examples demonstrate how GAHAC guidelines uncover hedonic issues of the UX. The three selected aspects were chosen purely as examples, and the focus was on highlighting specific issues rather than covering all possible aspects. We do not have a formula for calculating aggregated values for the considered hedonic aspects because our goal is not to assign scores but to identify potential problems. The main objective is to guide discovery, particularly through qualitative evaluations, rather than relying on scales.

[1]Full GAHAC version: https://bit.ly/gahacv1

6 PRELIMINARY EVALUATION

For a preliminary assessment of the technology, we invited experts to provide feedback on the structure and application of the GAHAC guidelines and identify areas for improvement.

The first expert (E1) is a Computer Science tenure-track professor with extensive experience since 2010. She holds a Ph.D. in Electrical Engineering with a focus on Network Engineering, as well as a Master's degree in Computer Science, emphasizing Software Engineering, and a bachelor's degree in Systems Analysis. Her research interests are primarily in Software Engineering, encompassing Requirements Engineering, Databases, Software Systems, Cloud Computing, Usability, and Empirical Methods.

E1 highlighted the good number of guidelines already established (75 guidelines) and suggested that these guidelines be made available on an online platform – whether a website or mobile application – to facilitate access and use of GAHAC. In addition, E1 raised the importance of including a specific analysis of the phenomenon of "AI chatbot's hallucinations" in future versions, addressing the challenges associated with incorrect or invented responses, an essential aspect to improve the reliability of this technology.

The second expert (E2) is a professor and researcher with expertise in Informatics, actively contributing to editorial and academic initiatives. He holds a Ph.D. and a Master's degree in Computer Science and a bachelor's degree in Information Systems. His experience includes serving as an editor for journals focused on Computers in Education and Interactive Systems, as well as participating as a committee member in Human-Computer Interaction initiatives. Currently, he serves as the Editor-in-Chief of a journal on Interactive Systems, an Associate Editor for a journal on Responsible Computing, and the Coordinator of a Graduate Program in Computer Science.

E2 suggested revising the guidelines' language, indicating that they are aimed at evaluators rather than end users. E2 also highlighted that the language should consider cultural aspects, such as regionalisms, ensuring that the guidelines can be applied in different contexts. Finally, they pointed out the need to clarify whether GAHAC is geared towards particular application contexts and whether it covers all possible relevant hedonic aspects.

7 CONCLUSIONS AND FUTURE WORK

This paper presents the motivation and methodology for the initial version of a technology designed to address the gaps identified in a SMS (De Souza et al., 2024; Mariano et al., 2024). The SMS identifies a lack of evaluation techniques for text-based chatbot, particularly to assess aspects of hedonic UX and qualitative data collection. Based on these findings, we developed GAHAC, a technology that includes 75 qualitative guidelines covering 31 aspects of hedonic UX.

In a preliminary evaluation, two experts analyzed GAHAC and provided key insights. They emphasized the importance of analyzing "hallucinations" in chatbots, a phenomenon involving fabricated or incorrect responses that can impact trust and user experience. E1 emphasized the importance of addressing this issue in future versions of GAHAC to increase the trustworthiness of chatbot technologies. In addition, the experts recommended refining the language of the guidelines to ensure clarity for evaluators. E2 specifically suggested tailoring the language to evaluators rather than end users. Cultural considerations were also highlighted as crucial for broad applicability; E2 highlighted the importance of avoiding regionalisms and adapting the guidelines to diverse cultural contexts. Finally, they emphasized the need to verify that all relevant hedonic aspects are included, with E2 asking for clarification on whether the GAHAC was designed for specific application contexts and ensuring comprehensive coverage of hedonic dimensions.

The research team will conduct two additional studies with GAHAC. In the first study, we will work with HCI and chatbot experts to evaluate GAHAC. Experts will participate in an initial orientation meeting with the researcher. After reviewing GAHAC, they will join a follow-up meeting for a semi-structured interview to discuss their perceptions of the guidelines. The team will use the findings from this evaluation to create a revised version of GAHAC.

The second study will test GAHAC's feasibility. The researchers will divide a graduate class into two groups: one group will use GAHAC to identify hedonic UX issues, while the other will use the U2CHATBOT (Mafra et al., 2024) checklist for the same purpose. Both groups will use the ChatGPT application, following specific instructions to identify hedonic aspects of UX issues and record their findings in a spreadsheet. Afterward, participants will complete a post-assessment questionnaire based on TAM (Technology Acceptance Model) (Venkatesh and Bala, 2008), covering Ease of Use, Perceived Usefulness, and Intention to Use in the Future. The questionnaire also includes open-ended questions for detailed feedback and improvement suggestions

We also intend to address the limitations of this study by conducting a correlation analysis between the aspects to gain deeper insights into their interdependencies. Additionally, we aim to further explore the overlaps identified among certain aspects, which may share attributes or evaluation criteria.

ACKNOWLEDGEMENTS

This research was funded Coordination for the Improvement of Higher Education Personnel (CAPES) - Program of Academic Excellence (PROEX).

REFERENCES

Adamopoulou, E. and Moussiades, L. (2020). An overview of chatbot technology. In Maglogiannis, I., Iliadis, L., and Pimenidis, E., editors, *Artificial Intelligence Applications and Innovations*, pages 373–383, Cham. Springer International Publishing.

Al-Amin, M., Ali, M. S., Salam, A., Khan, A., Ali, A., Ullah, A., Alam, M. N., and Chowdhury, S. K. (2024). History of generative artificial intelligence (ai) chatbots: past, present, and future development. *arXiv*. Submitted on 4 Feb 2024.

Barbosa, S. D. J., da Silva, B. S., Silveira, M. S., Gasparini, I., Darin, T., and Barbosa, G. D. J. (2021). *Interação Humano-Computador e Experiência do Usuário*. Elsevier.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.

Candello, H. and Pinhanez, C. (2016). Designing conversational interfaces. *Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais*.

Chitturi, R., Raghunathan, R., and Mahajan, V. (2008). Delight by design: The role of hedonic versus utilitarian benefits. *Journal of marketing*, 72(3):48–63.

De Souza, A. C. R., Mariano, P. A. D. L., Guerino, G. C., Chaves, A. P., and Valentim, N. M. C. (2024). Technologies for hedonic aspects evaluation in text-based chatbots: A systematic mapping study. In *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems*, IHC '23, New York, NY, USA. Association for Computing Machinery.

Demand Sage (2024). ChatGPT Statistics (November 2024) – 200 Million Active Users. Accessed: 18 November 2024.

Fadhil, A., Schiavo, G., Wang, Y., and Yilma, B. A. (2018). The effect of emojis when interacting with conversational interface assisted health coaching system. In *Proceedings of the 12th EAI international conference on pervasive computing technologies for healthcare*, pages 378–383.

Guerino, G. C. and Valentim, N. M. C. (2020). Usability and user experience evaluation of conversational systems: A systematic mapping study. In *Proceedings of the XXXIV Brazilian Symposium on Software Engineering*, pages 427–436.

Hassenzahl, M. (2018). The thing and i: Understanding the relationship between user and product. In Blythe, M. and Monk, A., editors, *Funology 2: From Usability to Enjoyment*, pages 301–313. Springer International Publishing, Cham.

Hassenzahl, M., Platz, A., Burmester, M., and Lehner, K. (2000). Hedonic and ergonomic quality aspects determine a software's appeal. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 201–208.

ISO (2010). Ergonomics of human-system interaction — part 210: Human-centred design for interactive systems. Standard ISO 9241-210, International Organization for Standardization.

ISO (2018). Iso 9241-11:2018 ergonomics of human-system interaction — part 11: Usability: Definitions and concepts. Relatório técnico, International Organization for Standardization, Geneva, Switzerland.

Jia, M. and Jyou, L. (2021). The study of the application of a keywords-based chatbot system on the teaching of foreign languages. *Journal of Intelligent & Fuzzy Systems*, pages 1–10.

Jin, Y., Zhang, X., and Wang, W. (2019). Musicbot: Evaluating critiquing-based music recommenders with conversational interaction. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2):1–22.

Kujala, S. and Vännänen-Vainio-Mattila, K. (2009). Value of information systems and products: Understanding the users' perspective and values. *Journal of Information Technology Theory and Application (JITTA)*, 9(4):4.

Madan, A. and Kumar, S. (2012). Usability evaluation methods: a literature review. *International Journal of Engineering Science and Technology*, 4.

Mafra, M. G. S., Nunes, K., Rocha, S., Braz Junior, G., Silva, A., Viana, D., Silva, W., and Rivero, L. (2024). Proposing usability-ux technologies for the design and evaluation of text-based chatbots. *Journal on Interactive Systems*, 15(1):234–251.

Mariano, P. A. d. L., Souza, A. C. R. d., Guerino, G. C., Chaves, A. P., and Valentim, N. M. C. (2024). A systematic mapping study about technologies for hedonic aspects evaluation in text-based chatbots. *Journal on Interactive Systems*, 15(1):875–896.

McNamara, N. and Kirakowski, J. (2006). Functionality, usability, and user experience: three areas of concern. *Interactions*, 13(6):26–28.

Merčun, T. and Žumer, M. (2017). Exploring the influences on pragmatic and hedonic aspects of user experience. *Published Quarterly by the University of Borås, Sweden*, 22(1). Proceedings of the Ninth International Conference on Conceptions of Library and Information Science, Uppsala, Sweden, June 27-29, 2016.

Mudofi, L. N. H. and Yuspin, W. (2022). Evaluating quality of chatbots and intelligent conversational agents of bca (vira) line. *Interdisciplinary Social Studies*, 1(5):532–542.

Rapp, A., Curti, L., and Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, 151:102630.

Ren, R., Castro, J. W., Acuña, S. T., and Lara, J. d. (2019). Usability of chatbots: A systematic mapping study. In *Proceedings of the 31st International Conference on Software Engineering and Knowledge Engineering*, SEKE2019. KSI Research Inc. and Knowledge Systems Institute Graduate School.

Ruane, E., Farrell, S., and Ventresque, A. (2021). User perception of text-based chatbot personality. In Følstad, A., Araujo, T., Papadopoulos, S., Law, E. L.-C., Luger, E., Goodwin, M., and Brandtzaeg, P. B., editors, *Chatbot Research and Design*, pages 32–47, Cham. Springer International Publishing.

Tubin, C., Mazuco Rodriguez, J. P., and de Marchi, A. C. B. (2022). User experience with conversational agent: A systematic review of assessment methods. *Behaviour & Information Technology*, 41(16):3519–3529.

Veglis, A., Maniou, T. A., et al. (2019). Chatbots on the rise: A new narrative in journalism. *Studies in Media and Communication*, 7(1):1–6.

Venkatesh, V. and Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2):273–315.

Vermeeren, A., Law, L.-C., Roto, V., Obrist, M., Hoonhout, J., and Väänänen, K. (2010). User experience evaluation methods: Current state and development needs. pages 521–530.