

Explorando UX em Interfaces Conversacionais: percepções de um estudo preliminar da literatura

Jéssica Favero Ribeiro¹, Leticia dos Santos Machado¹, Marcelo Soares Pimenta¹

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{jfribeiro, leticia.machado, mpimenta}@inf.ufrgs.br

Abstract. *This paper presents a rapid review of the literature on user experience (UX) methods and techniques for evaluating conversational interfaces and identifying current practices. The analysis of 25 articles from the ACM and Scopus databases reveals a diversity of approaches, including usability questionnaires and qualitative methods such as interviews and focus groups. The absence of a universal standard in UX evaluation highlights the need for in-depth studies to consolidate best practices and provide a foundation for future research. This work aims to build a more cohesive and practical field of study, addressing the growing demands for human interactions with complex and intelligent computational systems.*

Resumo. *Este trabalho apresenta uma revisão rápida da literatura sobre métodos e técnicas de experiência de usuário (UX - user experience) para avaliar interfaces conversacionais, identificando práticas atuais. A análise de 25 artigos das bases ACM e Scopus revela a diversidade de abordagens, incluindo questionários de usabilidade, além de métodos qualitativos, como entrevistas e grupos focais. A ausência de um padrão universal na avaliação de UX destaca a necessidade de estudos aprofundados para consolidar melhores práticas e fornecer uma base para futuras pesquisas. Este trabalho visa construir um campo de estudo mais coeso e eficaz, atendendo às crescentes demandas por interações humanas com sistemas computacionais complexos e inteligentes.*

1. Introdução

Interfaces conversacionais são mecanismos que permitem interações entre pessoas e máquinas por meio de linguagem natural, habilitadas por voz, texto ou ambas (multimodais). Estes meios de interação têm sido amplamente utilizadas em diversas áreas, desempenhando um papel crucial para uma experiência de usuário (UX - *user experience*) positiva, eficaz e intuitiva, pois simulam uma conversa humana para oferecer respostas e soluções em tempo real [Fadhil and Gabrielli 2017, Trippas et al. 2020, Schuetzler et al. 2019]. Interações intuitivas ocorrem quando os usuários atingem facilmente um objetivo sem conhecimento prévio do sistema, enquanto a eficiência minimiza o esforço do usuário para realizar uma ação [ISO 2019].

Tendo em vista que a capacidade de interpretar corretamente a subjetividade do usuário e fornecer respostas relevantes pode melhorar significativamente a experiência do usuário, a avaliação adequada de UX para essas interfaces não apenas aprimora a satisfação do usuário, mas também aumenta a eficiência e a confiabilidade do sistema [Guerino et al. 2021, van der Zandt et al. 2021].

Portanto, neste estudo apresentamos uma revisão rápida (RR) [Tricco et al. 2015], abordagem variante da Revisão Sistemática da Literatura (RSL) que simplifica algumas etapas dos procedimentos de revisão, sendo especialmente útil em campos que evoluem rapidamente, para captura de tendências mais recentes e para auxiliar em tomadas de decisões em projetos a um baixo custo e em menos tempo [Cartaxo et al. 2018].

Esta RR foi realizada para investigar métodos de UX comumente utilizados para avaliar Interfaces Conversacionais. Sendo assim, este trabalho teve como objetivo, responder à seguinte pergunta de pesquisa: Quais métodos de *UX Research* são comumente utilizados para avaliar interfaces de usuário conversacionais?. Para tanto, apresentamos os resultados preliminares de uma *rapid review literature* que obteve 441 artigos científicos das bases de dados *ACM* e *Scopus*. Destes, 25 foram selecionados para extração das técnicas utilizadas devido a sua relevância para este trabalho. Através desta revisão rápida, encontramos 36 técnicas, o que demonstra que pesquisadores utilizam a combinação de uma gama variada de técnicas para avaliar a UX dessas interfaces.

Entre as principais, destacaram-se os testes de usabilidade, que envolvem a observação direta dos usuários enquanto interagem com o sistema, e diversos questionários. Além disso, métodos qualitativos, como entrevistas e grupos focais, que fornecem percepções sobre as expectativas dos usuários também foram mencionados. No entanto, dada a vasta quantidade de técnicas disponíveis, nota-se que não existe um padrão de técnicas para avaliar UX em interfaces conversacionais. Tornando necessário estudos para identificar e consolidar as principais metodologias utilizadas atualmente para a avaliação de UX em interfaces conversacionais. Esta necessidade evidencia a importância de uma análise aprofundada e sistemática para estabelecer as melhores práticas no campo.

O restante deste trabalho está estruturado da seguinte forma: na seção 2 descrevemos o método de pesquisa, para identificar as principais metodologias de avaliação de UX em interfaces conversacionais utilizadas atualmente. Em seguida, na seção 3 apresentamos as técnicas encontradas e uma breve análise dos resultados. Por fim, a seção 4 discute as implicações desses achados.

2. Método de Pesquisa

2.1. Planejamento da Revisão Rápida (RR)

Essa metodologia foi adaptada para a engenharia de software baseada em evidências, onde algumas etapas do processo de revisão são simplificadas ou omitidas para economizar tempo e custos [Cartaxo et al. 2018]. Para a execução desta RR, elaboramos e seguimos um protocolo de pesquisa, conforme apresentado abaixo na Tabela 1. Além disso, desenvolvemos um formulário de extração para coletar informações essenciais dos estudos selecionados, incluindo o tipo de pesquisa, a natureza da interface conversacional abordada, o objetivo do estudo, os domínios de aplicação e os métodos de avaliação de UX utilizados.

2.2. Execução da Revisão Rápida (RR)

Seguindo as recomendações de [Cartaxo et al. 2018], restringimos o número de fontes de dados e o período de busca para otimizar o processo em termos de custo e tempo. A pesquisa foi realizada nas bases de dados *ACM* e *Scopus* em julho de 2024, abrangendo publicações dos últimos cinco anos (2019-2024). A estratégia de busca resultou em 441

Tabela 1. Protocolo de revisão rápida sobre avaliação de UX de Interfaces Conversacionais

Objetivo	Esta revisão tem como objetivo investigar métodos comumente utilizados para avaliar a UX de sistemas conversacionais.
String de Busca	("UX"OR "User Experience") AND ("Conversational Interfaces"OR "Chatbots"OR "Voice Assistants") AND ("Question-Answer"OR "Q&A"OR "Question-Answering Systems")
Fontes	ACM Digital Library (https://dl.acm.org/) Scopus (http://www.scopus.com) (http://ieeexplore.ieee.org/)?
Critérios de Inclusão (CI)	(CI.1) Estudos abordando a avaliação de UX de Interfaces conversacionais (CI.2) Estudos escritos em Inglês ou Português (CI.3) Estudos publicados nos últimos 5 anos (CI.4) Estudos que avaliam Interfaces Conversacionais
Critérios de Exclusão (CE)	(CE.1) Papers duplicados (CE.2) Não artigos (CE.3) Não totalmente acessível (CE.4) Não relacionado à pesquisa de UX (CE.5) Artigos curtos (menos de 4 páginas)

Fonte: Autoria Própria

estudos. Na fase inicial de seleção, excluímos documentos que não eram artigos científicos, estudos com menos de 4 páginas de conteúdo, estudos que não avaliavam interfaces conversacionais, estudos anteriores com mais de 5 anos, não totalmente acessíveis, assim como aqueles não escritos em inglês ou português, reduzindo o número de estudos para 194. Em seguida, a pesquisadora revisou os títulos e resumos, eliminando 137 estudos que não atendiam aos critérios de inclusão, resultando em 57 estudos. A etapa final envolveu a análise completa do conteúdo desses artigos, o que levou à exclusão de mais 32 estudos que não cumpriam os critérios de aceitação. Isso culminou em 25 artigos que foram selecionados para extração e análise detalhada dos dados. Todo o processo de busca, seleção, extração e análise dos dados foi concluído em aproximadamente 20 dias.

3. Resultados

A análise dos dados obtidos através da revisão rápida revelou 36 técnicas de UX utilizadas para avaliar interfaces conversacionais, uma diversidade significativa, demonstrando a ampla gama de abordagens adotadas por pesquisadores e profissionais. A maioria dos artigos focou em *chatbots* de texto, mencionados em 17 estudos, utilizados em diversos domínios como saúde, Assistência ao cliente e educação. Assistentes de voz foram analisados em 7 artigos, destacando-se em contextos de Saúde, Assistentes de Voz em Diversos Contextos e interação humano-computador. Apenas um estudo apresentou um *chatbot* multimodal, este também foi aplicado no domínio de saúde.

A análise da relação entre o domínio e as técnicas de UX mais utilizadas mostrou que, para *chatbots* de texto, o teste de usabilidade foi a técnica predominante, especialmente no domínio da saúde. Outras técnicas como entrevistas e análise de feedback também foram amplamente empregadas, especialmente em contextos de interação humano-computador (IHC) e assistência ao cliente. No caso dos assistentes de voz, o teste de usabilidade e as entrevistas foram as técnicas mais utilizadas, com uma concentração significativa de estudos nos domínios de educação e saúde. Para as interfaces multimodais, a única técnica mencionada foi o *UMUX-LITE*, aplicada no domínio da saúde.

Em relação à relação entre o domínio e o tipo de interface conversacional (IC) mais utilizada, observou-se que *chatbots* de texto foram predominantes em uma ampla

variedade de domínios, incluindo saúde, educação, e assistência ao cliente. Os assistentes de voz foram mais comuns nos domínios de saúde e contextos diversos de assistentes de voz, enquanto as interfaces multimodais foram raramente exploradas, com apenas um estudo identificado no domínio da saúde.

Quando se analisou a relação entre o tipo de interface conversacional (IC) e as técnicas de UX mais utilizadas, foi possível identificar que, para *chatbots* de texto, as técnicas mais aplicadas foram o teste de usabilidade, entrevistas e análise de feedback. Para assistentes de voz, além do teste de usabilidade e entrevistas, também se destacou o uso do questionário *System Usability Scale* (SUS). Por outro lado, as interfaces multimodais foram pouco exploradas, com apenas a técnica *UMUX-LITE* sendo mencionada.

Esta diversidade de tipos de interfaces conversacionais, domínios de aplicação e técnicas de avaliação de UX, reforça a necessidade de uma abordagem versátil na avaliação de UX, adaptada às características específicas de cada tipo de interface. A lista dos 25 estudos selecionados na RR com as técnicas e domínios apresentados encontra-se disponível no repositório *Zenodo*¹.

4. Considerações finais

Esta revisão rápida da literatura sobre métodos e técnicas de UX para avaliação de interfaces conversacionais revelou uma diversidade significativa de abordagens adotadas por pesquisadores e profissionais. Tal diversidade de métodos encontrados evidencia a ausência de um padrão universal, indicando que a escolha das técnicas de avaliação deve ser cuidadosamente adaptada às especificidades de cada contexto de uso e tipo de interface conversacional.

Os resultados desta revisão rápida destacam a necessidade de desenvolver um *framework* mais estruturado para a avaliação de UX em interfaces conversacionais. A consolidação de melhores práticas poderia não apenas padronizar os métodos de avaliação, mas também facilitar a comparação de resultados entre diferentes estudos e aplicações. Além disso, a criação de guias e ferramentas específicas para a avaliação de *Chatbots*, assistentes de voz e interfaces multimodais contribuiria significativamente para a evolução da pesquisa e da prática profissional em Interação Humano-Computador.

Como próximos passos, sugerimos a realização de estudos empíricos que testem a eficácia das técnicas de avaliação mais mencionadas na literatura, em cenários reais de uso. Isso permitirá uma compreensão mais aprofundada de suas vantagens e limitações, bem como sua aplicabilidade em diferentes domínios.

Em suma, a revisão rápida aqui apresentada fornece uma base inicial para futuras investigações e desenvolvimentos no campo da UX para interfaces conversacionais. A identificação das técnicas atualmente utilizadas é um passo fundamental para a construção de um campo de estudo mais coeso e eficaz, que atenda às demandas crescentes de interações humanas com sistemas computacionais cada vez mais complexos e inteligentes.

¹<https://zenodo.org/uploads/13522305>

Referências

- Cartaxo, B., Pinto, G., and Soares, S. (2018). The role of rapid reviews in supporting decision-making in software engineering practice. In *Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018*, pages 24–34.
- Fadhil, A. and Gabrielli, S. (2017). Addressing challenges in promoting healthy lifestyles: the al-chatbot approach. In *Proceedings of the 11th EAI international conference on pervasive computing technologies for healthcare*, pages 261–265.
- Guerino, G. C., Silva, W. A. F., Coleti, T. A., and Valentim, N. M. C. (2021). Assessing a technology for usability and user experience evaluation of conversational systems: An exploratory study. In *ICEIS (2)*, pages 463–473.
- ISO (2019). Ergonomics of human-system interaction – part 210: Human-centred design for interactive systems (2nd ed.). Standard, International Organization for Standardization (2019).
- Schuetzler, R. M., Grimes, G. M., and Giboney, J. S. (2019). The effect of conversational agent skill on user behavior during deception. *Computers in Human Behavior*, 97:250–259.
- Tricco, A. C., Antony, J., Zarin, W., Striffler, L., Ghassemi, M., Ivory, J., Perrier, L., Hutton, B., Moher, D., and Straus, S. E. (2015). A scoping review of rapid review methods. *BMC medicine*, 13:1–15.
- Trippas, J. R., Spina, D., Thomas, P., Sanderson, M., Joho, H., and Cavedon, L. (2020). Towards a model for spoken conversational search. *Information Processing & Management*, 57(2):102162.
- van der Zandt, L., van der Stappen, E., and van Turnhout, K. (2021). Towards real-life adoption of conversational interfaces: Exploring the challenges in designing chatbots that live up to user expectations. In *34th British HCI Conference*, pages 306–311. BCS Learning & Development.