

UNIVERSIDADE FEDERAL DO PARANÁ

ANA CAROLINA ROSSI DE SOUZA

TECNOLOGIAS PARA A AVALIAÇÃO DE EXPERIÊNCIA DO USUÁRIO EM CHATBOTS:
UM MAPEAMENTO SISTEMÁTICO DA LITERATURA

CURITIBA PR

2022

ANA CAROLINA ROSSI DE SOUZA

TECNOLOGIAS PARA A AVALIAÇÃO DE EXPERIÊNCIA DO USUÁRIO EM CHATBOTS:
UM MAPEAMENTO SISTEMÁTICO DA LITERATURA

Trabalho apresentado como requisito parcial à conclusão do Curso de Bacharelado em Informática, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Informática Biomédica*.

Orientador: Prof.^ª Dr.^ª Natasha Malveira C. Valentim.

Coorientador: Prof. Me. Guilherme C. Guerino.

CURITIBA PR

2022

AGRADECIMENTOS

Agradeço a Deus, por me capacitar para realizar este trabalho e sempre me abastecer de fé, perseverança e serenidade durante toda a jornada acadêmica.

Sou grata também pelo meus pais, Luiz Henrique e Mariana, que sempre me apoiaram e me incentivaram a nunca desistir e a sempre persistir nos meus sonhos.

Agradeço ao meu namorado, Mateus, por estar ao meu lado durante as conquistas e sempre afirmar que “Vai ficar tudo bem!”.

Agradeço a professora Ana Chaves por assessorar no delineamento do tema do trabalho e auxiliar no descobrimento de trabalhos relacionados, que embasaram a pesquisa.

Agradeço aos meus orientadores, Natasha Valentim e Guilherme Guerino, por trilharem esta etapa acadêmica comigo e por sempre estarem disponíveis para tirar dúvidas, auxiliar no desenvolvimento e me ensinar muito!

RESUMO

Já é evidente a presença de tecnologias no nosso cotidiano profissional, acadêmico e pessoal. Para acompanhar esta crescente, muitos são os estudos que apresentam e avaliam estas tecnologias, desde sistemas de informações, interfaces, até sistemas conversacionais. Um exemplo de tecnologia que vem sendo avaliada com frequência são os *chatbots*, que são interfaces conversacionais que interagem com humanos, em linguagem natural, através de voz ou texto. Uma das formas de avaliar os *chatbots* é por meio da Experiência do Usuário (User eXperience - UX), que é a resposta e percepção do usuário ao utilizar um sistema, produto ou serviço. A partir de tais fatos, neste trabalho de graduação buscou-se realizar um Mapeamento Sistemático da Literatura (MSL) com o objetivo de investigar tecnologias de avaliação de UX com foco em *chatbots*. Os *chatbots* investigados neste trabalho são ativados somente por meio de texto, podendo exibir interações de clicar, vídeos e imagens, além do texto escrito. Além disso, o aspecto da UX explorado é o hedônico, relacionado ao bem-estar psicológico do usuário, e foi realizada apenas a análise quantitativa dos resultados. Palavras-chave: Chatbots. Experiência

do Usuário. Mapeamento Sistemático da Literatura.

ABSTRACT

The presence of technologies in our professional, academic and personal daily life is already evident. To accompany this growth, many studies present and evaluate these technologies, from information systems, interfaces, to conversational systems. An example of technology frequently evaluated is chatbots, which are conversational interfaces that interact with humans, in natural language, through voice or text. One of the ways to evaluate chatbots is through the User Experience (User eXperience - UX), which is the user's response and perception when using a system, product or service. Based on these facts, this undergraduate work sought to carry out a Systematic Mapping Study (SMS) to investigate UX assessment technologies focusing on chatbots. The chatbots investigated in this work are activated only through text, being able to display click interactions, vídeos, and images, in addition to written text. In addition, the aspect of UX explored is hedonic, related to the user's psychological well-being and only the quantitative analysis of the results was performed.

Keywords: Chatbots. User Experience. Systematic Mapping Study.

LISTA DE FIGURAS

2.1	Exemplo do ELIZA. Fonte (Khyani et al., 2020)	13
2.2	Exemplo do Woebot. Fonte (Brodwin, 2018)	14
3.1	Ano de publicação dos artigos selecionados.	23
3.2	Distribuição de artigos por conferência.	24
3.3	Distribuição de artigos por periódicos.	24
4.1	Resultados da SQ1.	27
4.2	Resultados da SQ2.	28
4.3	Tecnologias de Avaliação encontradas	29
4.4	Resultados da SQ3.	29
4.5	Resultados da SQ4.	30
4.6	Resultados da SQ6.	31
4.7	Resultados da SQ8.	31
4.8	Resultados da SQ9.	32
4.9	Resultados da SQ13.	33

LISTA DE TABELAS

3.1	Sub-questões da pesquisa	17
3.2	Palavras e elementos da <i>String</i>	17
3.3	Total de artigos retornados e selecionados no 1º e 2º filtro.	19
3.4	Subquestões de pesquisa e possíveis respostas..	21
4.1	Resultados do MSL para cada uma das subquestões	26
A.1	Tecnologias de avaliação encontradas	41

LISTA DE ACRÔNIMOS

ABCCT	Affective Benefits and Costs of Communication Technology
AMS	Academic motivation scale
CSs	Conversational Systems
GQM	Goal-Question-Metric
IHC	Interação Humano-Computador
IMI	Intrinsic motivation inventory
IRI	Interpersonal Reactivity Index
MIM	Multidimensional Integrative Model
MSL	Mapeamento Sistemático da Literatura
SUS	System Usability Scale
TG	Trabalho de Graduação
UEQ	User Experience Questionnaire
UEQ-S	User Experience Questionnaire - Short
UFPR	Universidade Federal do Paraná
USE	Usefulness, Satisfaction, and Ease of use
UX	User Experience
VAS	Visual Analogue Scale
WoZ	Wizard of Oz

SUMÁRIO

1	INTRODUÇÃO	10
1.1	CONTEXTO	10
1.2	MOTIVAÇÃO	10
1.3	PROBLEMA	11
1.4	OBJETIVO	12
1.5	ORGANIZAÇÃO DO TG	12
2	REFERENCIAL TEÓRICO E TRABALHOS RELACIONADOS	13
3	METODOLOGIA	16
3.1	PLANEJAMENTO	16
3.1.1	Objetivo	16
3.1.2	Questões de Pesquisa	16
3.1.3	Métodos de busca de publicações	17
3.1.4	Escopo da Pesquisa	18
3.1.5	Idiomas	18
3.1.6	Critérios	18
3.2	EXECUÇÃO	18
3.2.1	Procedimento de Seleção	18
3.2.2	Seleção de Artigos	19
3.2.3	Extração de Dados	19
3.3	RELATÓRIO	21
3.3.1	Ano de publicação	22
3.3.2	Locais de Publicação	23
4	RESULTADOS QUANTITATIVOS E DISCUSSÃO	25
4.1	VISÃO GERAL DOS RESULTADOS	25
4.2	SQ1 - ASPECTOS DA UX QUE AS TECNOLOGIAS AVALIAM	26
4.3	SQ2 - TECNOLOGIAS ESPECÍFICAS PARA <i>CHATBOTS</i> OU GENÉRICAS	28
4.4	SQ3 - TECNOLOGIAS CRIADAS PARA O ESTUDO OU EXISTENTES	29
4.5	SQ4 - COLETA DE RESPOSTAS	30
4.6	SQ6 - TIPO DE ANÁLISE	30
4.7	SQ8 - <i>CHATBOTS</i> CRIADOS PARA UM GRUPO ESPECÍFICO	31
4.8	SQ9 - TIPO DOS <i>CHATBOTS</i>	32
4.9	SQ11 - AVALIAÇÃO EMPÍRICA DAS TECNOLOGIAS	32
4.10	SQ13 - ASPECTOS DA SAÚDE EMOCIONAL NA AVALIAÇÃO DE UX	32
4.11	DISCUSSÃO	33

5	CONCLUSÃO E TRABALHOS FUTUROS.	35
	REFERÊNCIAS	36
	APÊNDICE A – TECNOLOGIAS DE AVALIAÇÃO ENCONTRADAS . .	41
	APÊNDICE B – ARTIGOS EXTRAÍDOS NO MSL	42

1 INTRODUÇÃO

1.1 CONTEXTO

As interfaces conversacionais são quaisquer sistemas de inteligência artificial que interagem com humanos através de diálogos. Tecnologias que usufruem da inteligência artificial são treinados, utilizando algoritmos de aprendizado de máquina que entendem entradas abertas, como, por exemplo, frases. A partir das interações o sistema aprende e melhora suas respostas (MarutiTechlab, 2021).

Tais sistemas inteligentes são conhecidos por uma variedade de nomes, como, por exemplo, assistentes pessoais ou assistentes inteligentes (Candello e Pinhanez, 2016). Estas interfaces conversacionais podem ser definidas em quatro categorias (Candello e Pinhanez, 2016): diálogos baseados na fala (*speech-based dialogues*), diálogos baseados no texto (*text-based dialogues*), agentes virtuais interativos (*interactive virtual agents*), e robôs de conversação (*conversational robots*). Em algumas, a entrada de dados é por meio de áudios, como nas interfaces baseadas em voz; já em outras, o *input* é realizado por meio de textos, como, por exemplo, nos *chatbots*.

Os *chatbots* são definidos como sistemas conversacionais *on-line*, em linguagem natural, entre um humano e o computador (Jia, 2003). Diferente dos assistentes inteligentes ativados por voz, os *chatbots* que serão abordados nesta pesquisa, são ativados somente por meio de textos e as interfaces baseadas em diálogos podem apresentar interações de apontar e clicar, assim como imagens e vídeos, além do texto escrito (Candello e Pinhanez, 2016).

A partir da crescente presença de *chatbots* em *websites*, seja para tirar dúvidas de usuários, ou para prestar *customer service* em *e-commerces*, e o gradativo aumento de publicações que examinam e propõe *chatbots* expõem a relevância de avaliar a interação entre os usuários e estes sistemas. Uma das maneiras de se avaliar tal interação é por meio do nível de Experiência do Usuário (User eXperience - UX). De acordo com a ISO 9241 (International Organization for Standardization, 2019), a UX é a resposta e percepção do usuário ao utilizar um sistema, produto ou serviço. Essas percepções do usuário incluem emoções, crenças, preferências, percepções, comodidades, comportamentos e conquistas que podem ocorrer antes, durante e depois da utilização. A UX está relacionada a qualidade do sistema, e, a investigação e avaliação de aspectos da qualidade ocasionam diversos benefícios, como a redução de treinamento dos usuários e menores despesas com o suporte da ferramenta (Barbosa e Silva, 2010).

1.2 MOTIVAÇÃO

A UX é uma consequência da imagem da marca, apresentação, funcionalidade, performance do sistema, comportamento interativo e recursos de assistência de um sistema, produto ou serviço (International Organization for Standardization, 2019). A UX passou a ser relevante e, por consequência, os pesquisadores e profissionais interessaram-se em conhecer como as pessoas se sentem ao utilizar tecnologias (Hassenzahl, 2018).

Para Barbosa e Silva (2010), o estudo da interação entre seres humanos e sistemas computacionais viabiliza a compreensão dos mesmos, assim como, permite a realização de melhorias na construção e introdução das tecnologias no cotidiano dos usuários, com o objetivo de sempre proporcionar uma boa experiência. Os sistemas devem ser desenvolvidos de forma a melhorar a vida dos usuários, proporcionar bem-estar e satisfazer necessidades (Barbosa e

Silva, 2010), considerando frequentemente a qualidade do uso dos sistemas. O investimento na qualidade do sistema faz diversas contribuições, como, por exemplo, aumenta a produtividade do usuário, reduz o custo de treinamento e reduz o custo do suporte técnico. Por tais motivos que se faz tão importante o estudo e avaliação da UX.

De acordo com o modelo para sistemas de software atraentes e com boa UX, proposto por Hassenzahl et al. (2000), os *softwares* são descritos por meio de diferentes dimensões de qualidade, sendo ela dividida em dois grupos: a qualidade pragmática e a qualidade hedônica. A primeira trata da usabilidade e eficiência do sistema, enquanto a segunda considera aspectos não relacionados a tarefa que o usuário deseja realizar e sim a originalidade, inovação e beleza do sistema. O aspecto hedônico da UX também está relacionado ao bem-estar psicológico do usuário (Hassenzahl, 2004) e neste estudo será considerada apenas esta perspectiva da UX, pois é a vertente da qualidade que aborda o bem-estar emocional do usuário, e a importância desta relação será salientada na próxima seção.

Para Hassenzahl e Tractinsky (2006), UX diz respeito às tecnologias que cumprem mais do que somente necessidades instrumentais de uma maneira que seja admitido seu uso como subjetivo, direcionado, complexo e dinâmico. Os autores ainda vão além ao afirmar que “UX é uma consequência do estado interno do usuário, das características do *design system* e do contexto em que a interação se dá”. A partir desta afirmação é possível compreender que a UX é um conjunto de fatores, desde a situação em que a interação se dá, e pode ser influenciada até por particularidades emocionais do usuário.

Como abordado anteriormente, o aspecto hedônico da UX considera o bem-estar psicológico do usuário, assim como as emoções durante a interação com os sistemas. As emoções estão presentes em todos os tipos de interações entre seres vivos e podem impactar na atenção, percepção, memória, comportamento e até cognição. Elas também auxiliam na avaliação do ambiente, identificando se é um ambiente seguro, perigoso, bom ou ruim (Peter e Beale, 2008; Norman e Love, 2004; Piccolo et al., 2010).

1.3 PROBLEMA

As informações a respeito de respostas emocionais elucidam como os usuários expressam suas emoções durante a interação com sistemas de informação (Piccolo et al., 2010). Recentemente, notou-se um crescimento no número de pesquisas que investigam as emoções, e, uma vez que, as emoções influenciam no processo de tomada de decisão (Scherer, 2005; Xavier e de Almeida Neris, 2012), assim como a saúde emocional do usuário, tais aspectos podem impactar na avaliação da experiência do usuário ao utilizar alguma tecnologia.

Foram descobertos alguns Mapeamentos Sistemáticos da Literatura (MSL) a respeito de tecnologias que avaliam a qualidade de sistemas conversacionais. Guerino e Valentim (2020) mapearam as tecnologias de avaliação de usabilidade e UX utilizadas para avaliar sistemas conversacionais que utilizam especificamente a voz. Já Rapp et al. (2021) investigou como se dá a interação humano-computador e *chatbots*. Castro et al. (2019) investigaram tecnologias de avaliação de *chatbots* com foco especificamente no critério de usabilidade. No entanto, não identificamos na literatura um MSL que investigue tecnologias de avaliação com foco nos aspectos hedônicos da UX voltadas para *chatbots*, nos quais o *input* é realizado através de textos.

Dessa forma, foi observada a necessidade de investigar tecnologias de avaliação da UX no contexto de *chatbots* e analisar se durante as avaliações de UX foi considerado algum aspecto da saúde emocional e mental do usuário que está interagindo com o sistema. Portanto, nesta pesquisa temos a seguinte questão: “*Quais tecnologias estão sendo utilizadas para avaliar a UX em chatbots?*”

Para responder esta questão, foi realizado um Mapeamento Sistemático da Literatura (MSL) com o intuito de conhecer o que já foi investigado a respeito do tema. Um MSL é necessário em diversas situações e é executado com o objetivo de conhecer as pesquisas relacionadas a um determinado tópico. O MSL também é executado para compreender a extensão do tema de pesquisa, caracterizar tal tópico, classificá-lo, assim como resumir o que já existe a respeito e, por fim, detectar lacunas existentes no âmbito de pesquisa que podem vir a se tornar projetos (Kitchenham e Charters, 2007).

1.4 OBJETIVO

Este trabalho tem como principal objetivo identificar e caracterizar as tecnologias de avaliação de UX em *chatbots* de texto, e responder à questão principal da pesquisa mostrada na seção anterior.

Como objetivos específicos, temos:

- Definir o objetivo e questão de pesquisa do MSL;
- Definir a *string* de busca;
- Determinar o escopo da pesquisa;
- Estabelecer os critérios de inclusão e exclusão;
- Realizar o procedimento de seleção dos artigos;
- Efetivar a extração de dados;
- Analisar quantitativamente os dados extraídos;

1.5 ORGANIZAÇÃO DO TG

Além deste capítulo, esse trabalho de graduação possui o Capítulo 2 onde são apresentados o referencial teórico e os trabalhos relacionados ao realizado. O Capítulo 3 apresenta a metodologia de pesquisa utilizada. No Capítulo 4 são discutidos os resultados quantitativos obtidos. Por fim, no Capítulo 5 são apresentados a conclusão e os trabalhos futuros.

2 REFERENCIAL TEÓRICO E TRABALHOS RELACIONADOS

Chatbots podem ser conhecidos por diferentes termos, como: sistema de conversação de máquina, sistema de diálogo e *chatterbot* (Shawar e Atwell, 2007). Estes sistemas são projetados para simular uma comunicação inteligente, por meio de texto ou fala (Dahiya, 2017). No presente os *chatbots* são empregados para facilitar vários processos de negócios, como situações relacionadas ao atendimento ao cliente e personalização devido à sua acessibilidade, custo bastante baixo e facilidade de uso para o consumidor final (Przegalinska et al., 2019).

O primeiro chatbot conhecido foi o Eliza, desenvolvido no período entre 1964 e 1966, que utilizava *input* textual e que se comunicava com humanos através de roteiros artesanais (Weizenbaum, 1966). Tais roteiros simulavam um psicoterapeuta rogeriano, e o *chatbot* não entendia a conversa, mas sim procurava pelas respostas ideais por intermédio de correspondência de padrões. A Figura 2.1 apresenta o *chatbot* Eliza, presente no estudo de Khyani et al. (2020).

```

Welcome to
EEEEEE LL      IIII ZZZZZZZ AAAAA
EE      LL      II      ZZ  AA  AA
EEEEEE LL      II      ZZZ  AAAAAA
EE      LL      II      ZZ  AA  AA
EEEEEE LLLLLL IIII ZZZZZZZ AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?

```

Figura 2.1: Exemplo do ELIZA. Fonte (Khyani et al., 2020)

Mais recentemente nota-se um aumento no investimento do desenvolvimento de *chatbots*, agentes virtuais e assistentes pessoais, como a *Siri*, da *Apple*¹ e a *Alexa*, da *Amazon*². Considerando esta crescente, também são muitos os artigos que abordam tais sistemas e apresentam suas diversas aplicações, como Ashktorab et al. (2019) que apresentou um *chatbot* que auxilia no atendimento do cliente no serviço de *help-desk*.

Já Fitzpatrick et al. (2017), apresentou o *chatbot* WOEBOT direcionado ao cuidado com a saúde mental e emocional do usuário. O WOEBOT é um agente de conversação automatizado projetado para fornecer terapia cognitivo-comportamental no formato de conversas breves e diárias, e que também realiza rastreamento de humor (Fitzpatrick et al., 2017). A Figura 2.2 exibe um exemplo de diálogo realizado nesse *chatbot*. Alguns pesquisadores realizaram estudos com esse chatbot, tais como De Nieva et al. (2020), em que os autores investigaram o uso do WOEBOT para aliviar estresse acadêmico e Bae Brandtzæg et al. (2021), onde o *chatbot* foi utilizado no processo de avaliação da percepção de jovens sobre suporte social em *chatbots*.

Este cuidado e atenção com o bem-estar do usuário são tópicos bastante relevantes no aspecto hedônico da UX e existem estudos direcionados somente à relação entre sistemas

¹<https://www.apple.com/ios/siri/>

²<https://developer.amazon.com/alexa/>

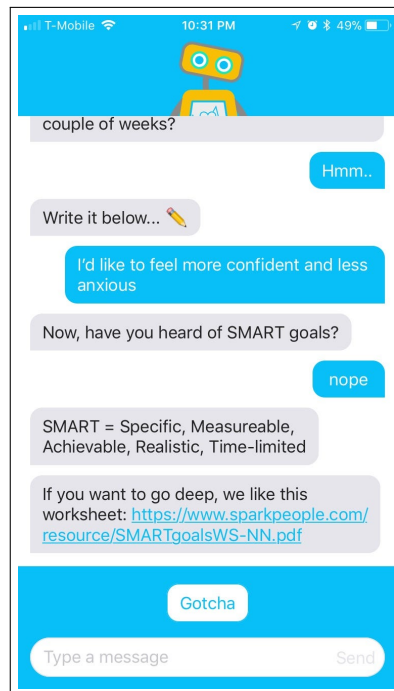


Figura 2.2: Exemplo do Woebot. Fonte (Brodwin, 2018)

computacionais e o estado emocional do usuário, como o realizado por Xavier e de Almeida Neris (2012). Neste trabalho os autores relatam que é relevante considerar os valores emocionais do usuário ao avaliar um sistema, e ainda vão além, ao afirmar que utilizar apenas uma técnica de avaliação de emoção dos usuários pode detectar falsos positivos na identificação do estado emocional. Os autores também observaram que, em alguns casos, usuários idosos podem até ter uma experiência negativa com uma tecnologia, perante padrões comuns de usabilidade, como, tempo de interação, o número de erros ou tarefas não concluídas, e ainda assim avaliar a experiência como positiva, justificando que “Na vida, não podemos ficar tristes”.

Neste contexto, é importante identificar quais tecnologias tem sido utilizadas para avaliar a qualidade de *chatbots* e/ou sistemas conversacionais. Alguns trabalhos e estudos na literatura que mapeiam as tecnologias que avaliam a qualidade de sistemas foram identificados. Por exemplo, no trabalho de Guerino e Valentim (2020) foram investigados sistemas conversacionais que utilizam a voz humana para realizar alguma ação e as tecnologias que avaliam este tipo de sistema. Neste MSL foram realizadas buscas nas seguintes bibliotecas virtuais: Scopus, IEEEExplore, ACM Digital Library e Engineering Village. O objetivo do MSL foi identificar e avaliar as tecnologias utilizadas pelos pesquisadores e profissionais para avaliar Usabilidade e UX de sistemas conversacionais. Foram encontradas 31 tecnologias de avaliação. Contudo, os autores constataram a carência de processos de validação e avaliação empírica das tecnologias que foram criadas pelos autores dos artigos avaliados. Dos *chatbots* encontrados nos estudos identificados a categoria mais comum era a de sistemas para auxiliar em tarefas diárias.

Já o trabalho desenvolvido por Castro et al. (2019) teve como objetivo identificar o estado da arte no uso de chatbot e na aplicação do mesmo como uma técnica de interação humano-computador, com o propósito de analisar como avaliar a usabilidade dos *chatbots*. Os autores realizaram este MSL nas principais bases de dados científicas (Scopus, ACM Digital Library, IEEE Xplorer, SpringerLink e Science Direct) e uma das constatações foi que o procedimento mais realizado para avaliar a usabilidade dos *chatbots* é eleger um grupo de usuários para utilizar

o sistema livremente, ou realizar tarefas determinadas, e então medir a satisfação através do questionário SUS (*System Usability Scale*) (Brooke, 1996).

O estudo realizado por Rapp et al. (2021) é mais um artigo relacionado a interação humano computador e “chatbots”. Foram avaliados 83 artigos e a investigação foi voltada para como os usuários realizam a interação com *chatbots* baseados em texto. A revisão identificou que as pesquisas avaliadas se concentram em como a confiança, o engajamento e a satisfação podem ser aspectos importantes na experiência de interação do usuário. Os pesquisadores também notaram que nos estudos analisados, os autores, utilizam sistemas comerciais, Wizard of Oz (WoZ), e protótipos totalmente desenvolvidos para explorar as experiências, atitudes e comportamentos dos usuários.

Tubin et al. (2021) analisaram como avaliar a experiência com Agentes Conversacionais que estão se tornando populares por oferecerem uma experiência mais realística e natural ao usuário. A revisão da literatura foi feita para identificar como a experiência do usuário é avaliada ao interagir com agentes. Os autores perceberam que é necessário avaliar a experiência em diferentes momentos e utilizar métodos combinados, para compreender aspectos relacionados aos sentimentos e comportamentos dos participantes.

Os artigos apresentados anteriormente possuem suas particularidades, de forma que são abordados diferentes tipos de *chatbots*, a relação com a IHC e os métodos, técnicas e tecnologias para avaliar sistemas conversacionais, seja através da usabilidade ou da experiência do usuário. No trabalho de Guerino e Valentim (2020) o mapeamento realizado considera tanto usabilidade, quanto UX, contudo avalia somente sistemas conversacionais de voz e não faz distinção entre o aspecto hedônico e pragmático da UX. Já no estudo de Rapp et al. (2021) é considerada a IHC como um todo, sem realizar recortes para a experiência do usuário ao utilizar um *chatbot*. Castro et al. (2019) apresentam um MSL generalizado de *chatbots*, sem realizar recortes e delimitações a respeito da natureza dos *chatbots* ou relativo ao aspecto da qualidade considerada no mapeamento, enquanto Tubin et al. (2021), apesar de considerar a UX ao utilizar agentes conversacionais, não fazem distinção do método utilizado para a entrada de dados nos sistemas conversacionais avaliados. Desta forma, o MSL realizado neste trabalho objetiva preencher algumas lacunas encontradas na literatura, como considerar apenas *chatbots* que utilizam entrada textual, examinar somente tecnologias de avaliação de UX com foco na qualidade hedônica e também levantar a questão de se estas avaliações de UX consideram aspectos da saúde mental e emocional do usuário.

3 METODOLOGIA

A metodologia utilizada neste trabalho de graduação se baseia num estudo secundário, que é um estudo que revisa todos os estudos primários relacionados a uma questão de pesquisa específica, e que possui o propósito de integrar evidências relacionadas a uma questão de pesquisa específica (Kitchenham e Charters, 2007). Um dos tipos de estudo secundário é o Mapeamento Sistemático da Literatura (MSL), cujo propósito é averiguar, qualificar e relacionar todas as pesquisas relevantes para um assunto definido (Kitchenham e Charters, 2007). O objetivo do MSL realizado neste trabalho de graduação foi identificar tecnologias de avaliação de UX com foco no aspecto hedônico em *chatbots* de texto. O MSL realizado segue as diretrizes propostas por Kitchenham e Charters (Kitchenham e Charters, 2007), em que a estrutura do projeto é dividida em três etapas: planejamento, execução e relatórios.

Durante a etapa de planejamento, foi estabelecido o protocolo de mapeamento, as questões de pesquisa, as fontes dos dados, a *string* de busca, os critérios de inclusão e os de exclusão. Durante a execução, foram realizadas as buscas nas fontes de dados, a seleção de artigos com base em dois filtros, a extração de dados dos artigos e a análise desses dados. Nos relatórios, os resultados são compartilhados de forma quantitativa através deste trabalho de graduação.

3.1 PLANEJAMENTO

3.1.1 Objetivo

O objetivo deste MSL foi baseado no paradigma GQM (Goal-Question-Metric) (Basili e Rombach, 1988), e está descrito abaixo.

- Analisar **publicações científicas**
- Com o propósito de **caracterizar**
- Com relação a **tecnologias de avaliação de UX com foco no aspecto hedônico em *chatbots* de texto**
- Do ponto de vista de **pesquisadores de Informática Biomédica (IBM) e Ciência da Computação (CC), com expertise nas áreas de Interação Humano-Computador (IHC) e Engenharia de Software (ES)**
- No contexto de **fontes primárias disponíveis nas máquinas de busca ACM Digital Library e IEEE Xplore**

3.1.2 Questões de Pesquisa

A principal questão de pesquisa para este mapeamento é: “Quais tecnologias estão sendo utilizadas para avaliar UX em *chatbots*?”. Além desta questão, foram elencadas sub-questões, com o propósito de responder questionamentos específicos, apresentadas na Tabela 3.1 .

Tecnologia de Avaliação	
SQ1. Aspectos abordados	Quais aspectos da UX a tecnologia avalia?
SQ2. Tecnologia específica ou genérica	A tecnologia é específica para <i>chatbots</i> ou para sistemas em geral?
SQ3. Tecnologia existente	A tecnologia foi criada para o estudo ou foi baseada em um tecnologia já existente?
SQ4. Coleta de respostas	Como a tecnologia de avaliação coleta as respostas dos participantes?
SQ5. Detalhamento da composição da tecnologia	Qual a composição da tecnologia de avaliação?
SQ6. Tipo de Análise	A tecnologia extrai dados quantitativos ou qualitativos?
Avaliação dos <i>Chatbots</i>	
SQ7. Categoria do Software	Qual é a aplicação do <i>chatbots</i> ?
SQ8. Grupo específico	O <i>chatbot</i> foi criado para um grupo específico de pessoas? Qual grupo?
SQ9. Tipo do <i>chatbot</i>	O <i>chatbot</i> é de algum tipo específico? Qual? (Rapp et al., 2021)
SQ10. Tipo de avaliação	De qual forma o <i>chatbot</i> foi avaliado?
Avaliação Empírica da Tecnologia de Avaliação	
SQ11. Avaliação Empírica	A avaliação da tecnologia foi avaliada empiricamente?
SQ12. Tipo de Avaliação	Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?
Avaliação do estado de saúde emocional do usuário	
SQ13. Avaliação Emocional	A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo o <i>chatbot</i> ?
SQ14. Foma de avaliação da saúde mental	De que forma a saúde mental do usuário foi avaliada?

Tabela 3.1: Sub-questões da pesquisa

3.1.3 Métodos de busca de publicações

Para definir a *string* de busca foi utilizado o método PICOC (Kitchenham e Charters, 2007), empregado para formular a questão da pesquisa. A sigla, proveniente da língua inglesa, é referente a: *Population* (P), *Intervention* (I), *Comparison* (C), *Outcome* (O) e *Context* (C), porém, para este MSL serão utilizadas apenas a *Population* (população), a *Intervention* (intervenção) e o *Outcome* (resultado), caracterizado como PIO, visto que as demais são utilizadas em revisões sistemáticas onde há comparação entre os resultados obtidos.

Logo, na *string* de busca a **População** é o *chatbot*, a **Intervenção** são as tecnologias que avaliam a UX em *chatbot*, a **Comparação** não se aplica, o **Resultado** é a UX e o **Contexto** não se aplica. Foram utilizados os conectores *AND* (E) e *OR* (OU) para compor a *string*. Ela foi construída de forma a ser abrangente em questão do resultado e mais restrita quanto a intervenção, resultando na *string*, exposta na Tabela 3.2.

Termos e <i>String</i> de Busca	
População	("chatbot*"OR "conversational agent*"OR "chatbot"OR "artificial conversational entity" OR " conversational interface"OR "conversational system"OR "conversation system" OR "dialogue system"OR "conversational user interface"OR "conversational UI") AND
Intervenção	("tool"OR "framework"OR "technique"OR "method"OR "guideline"OR "pattern"OR "metric"OR "approach"OR "inspection"OR "heuristic"OR "methodology") AND
Resultado	("user experience"OR "UX") AND ("evaluation"OR "assessment"OR "measure"OR "measurement") AND

Tabela 3.2: Palavras e elementos da *String*

3.1.4 Escopo da Pesquisa

A pesquisa foi realizada a partir de bibliotecas virtuais por meio mecanismo de busca avançado. As bibliotecas selecionadas foram: ACM Digital Library e a IEEE Xplore. Foram selecionadas estas bibliotecas, principalmente, por proporcionar um mecanismo de busca competente, por consentir a utilização de termos semelhantes na *string* e também por retornar um número favorável de artigos por conta da abrangência das bases de dados manipuladas. Estas bibliotecas também foram selecionadas porque disponibilizam vários artigos na área da Informática Biomédica e da Ciência da Computação.

3.1.5 Idiomas

Na pesquisa foram aceitas publicações em Português e Inglês. Foi optado pelo português porque é o idioma falado no Brasil, e o inglês porque a maioria das conferências e periódicos são publicados neste idioma.

3.1.6 Critérios

Os Critérios de Inclusão de publicações foram os seguintes:

- **I1:** Publicações que apresentaram tecnologias que avaliam a UX no uso de *chatbots*.
- **I2:** Publicações onde são descritos estudos experimentais de avaliações de UX no uso de *chatbots*.

Os Critérios de Exclusão foram os abaixo:

- **E1:** Não foram selecionadas publicações que não atendam aos critérios de inclusão.
- **E2:** Não foram selecionadas publicações que possuem idioma diferente de Inglês e Português.
- **E3:** Não foram selecionadas publicações que não tinham disponibilidade de conteúdo para leitura e análise dos dados (especialmente em casos onde os estudos são pagos ou não disponibilizados pelas máquinas de buscas).
- **E4:** Não foram selecionadas publicações que fazem parte da literatura cinzenta, como por exemplo relatórios técnicos e trabalhos em progresso.
- **E5:** Não foram selecionadas publicações que já foram incluídas em outra máquina de busca (duplicadas).

3.2 EXECUÇÃO

3.2.1 Procedimento de Seleção

A busca nas bibliotecas digitais foi realizada em outubro de 2021 e a avaliação dos dados que compõe o MSL foi realizada pelos pesquisadores Ana Carolina Rossi de Souza, Natasha Malveira Costa Valentim e Guilherme Corredato Guerino, que participaram do processo de definição do protocolo, execução dos filtros e extração das informações. O emprego de dois ou mais pesquisadores é necessário para preservar a consistência da pesquisa e diminuir o viés (Kitchenham e Charters, 2007). Neste MSL, os procedimentos de seleção se deu da seguinte

forma: o primeiro filtro foi composto da leitura do título e *abstract* dos artigos obtidos e avaliação dos mesmos perante os critérios de inclusão e exclusão definidos anteriormente. Já o segundo filtro compreendeu a leitura completa dos artigos que foram aprovados no primeiro filtro e foi realizada a avaliação com os mesmos critérios. Para realizar a gerência do mapeamento foi utilizada a ferramenta colaborativa Porifera (<https://porifera.app.br>), desenvolvida para gerir estudos secundários (Campos et al., 2022).

3.2.2 Seleção de Artigos

Em concordância com a Tabela 3.3, após a aplicação da *string* nas máquinas de busca selecionadas, foram encontrados 630 artigos. Destes, 91 foram aprovados após a aplicação 1º filtro, de acordo com os critérios de inclusão e exclusão apresentados anteriormente. Um total de 26 artigos foram selecionados após a aplicação do 2º filtro (ver Apêndice B).

Fonte	Total de Artigos Retornados	Artigos eleitos no 1º filtro	Artigos aceitos no 2º filtro
IEEEExplore	8	3	1
ACM Digital Library	622	88	25

Tabela 3.3: Total de artigos retornados e selecionados no 1º e 2º filtro.

Conforme apresentado, foram extraídos os dados dos 26 artigos aceitos no 2º filtro, os quais traçaram os resultados deste MSL e estes são Ceha et al. (2021), El Kamali et al. (2020), Fadhil et al. (2018), Fahn e Riener (2021), Xiao et al. (2019), Elsholz et al. (2019), Kim et al. (2019), Chen et al. (2021), Fiore et al. (2019), Jin et al. (2019), Lee et al. (2021), Völkel e Kaya (2021), Jain et al. (2018), Wald et al. (2021), Park et al. (2021), Yun et al. (2020), Benke et al. (2020), De Nieva et al. (2020), Wambsganss et al. (2021), Bawa et al. (2020), (Portela e Granell-Canut, 2017), Flohr et al. (2021), Liu et al. (2020), Kattenbeck et al. (2018), Bae Brandtzæg et al. (2021), Denecke et al. (2020).

3.2.3 Extração de Dados

A metodologia de extração de dados direcionada a este MSL foi fundamentada em obter respostas para cada sub-questões de pesquisa, expostas na Tabela 3.4. Através deste método é possível garantir os mesmos critérios de extração de dados para todos os artigos avaliados. Além das sub-questões foram elencadas as possíveis respostas que podem ser obtidas a partir das publicações. Também foram extraídos dos artigos as informações: ano de publicação e local de publicação. Estes dados são expostos na subseção de Relatórios.

Subquestão	Possíveis respostas
SQ1. Quais aspectos da UX a tecnologia avalia?	As respostas obtidas variam de artigo para artigo. Aspectos de avaliação de UX podem ser a satisfação do usuário, efetividade, imersão, fadiga, prazer, entre outros.
SQ2. A tecnologia é específica para <i>chatbot</i> ou para sistemas em geral?	Específica: tecnologia de avaliação de UX específica para <i>chatbot</i> . Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software

<p>SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?</p>	<p>Existente: tecnologia é baseada em uma existente. Criada: tecnologia foi criada para o estudo e descrita no artigo.</p>
<p>SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes?</p>	<p>As respostas obtidas variam de artigo para artigo. O objetivo é verificar como as respostas dos usuários foram capturadas ao utilizar a tecnologia de avaliação, como, escala Likert, questões abertas, múltipla escolha, checklist, entre outros.</p>
<p>SQ5. Qual a composição da tecnologia de avaliação?</p>	<p>As respostas obtidas são subjetivas e variam de artigo para artigo. O objetivo é extrair atributos de cada uma das tecnologias, como identificar as questões utilizadas, se a tecnologia for um questionário ou entrevista, identificar as métricas usadas, se é uma tecnologia baseada em avaliação de erro, identificar os cálculos realizados, se é uma tecnologia de análise eficiente, entre outros.</p>
<p>SQ6. A tecnologia extrai dados quantitativos ou qualitativos?</p>	<p>Quantitativos: se a análise da avaliação foi realizada quantitativamente. Qualitativos: se a análise da avaliação foi realizada qualitativamente. Misto: se a análise da avaliação foi realizada de forma qualitativa e quantitativa.</p>
<p>SQ7. Qual é a aplicação do <i>chatbot</i>?</p>	<p>As respostas obtidas são subjetivas e foram identificadas durante a leitura dos artigos. Exemplos de categorias são: saúde, educação, mobilidade urbana, facilidades diárias, entre outros.</p>
<p>SQ8. O <i>chatbot</i> foi criado para um grupo específico de pessoas? Qual grupo?</p>	<p>Sim, o <i>chatbot</i> foi criado para dar assistência a um grupo específico. Exemplos: cegos, surdos, idosos, crianças, entre outros. Não, o <i>chatbot</i> não foi criado para um grupo específico.</p>
<p>SQ9. O <i>chatbot</i> é de algum tipo específico? Qual? (Rapp et al., 2021)</p>	<p>Sim, o <i>chatbot</i> é orientado a tarefas. Significa que o <i>chatbot</i> foi construído para ajudar os usuários a executar uma tarefa ou resolver um problema.</p>

	<p>Sim, o <i>chatbot</i> é orientado a conversação. Significa que o <i>chatbot</i> foi projetado principalmente para manter uma boa qualidade de conversas com humanos ou estabelecer algumas formas de relacionamento com eles.</p> <p>Sim, o <i>chatbot</i> é tanto orientado à tarefas quanto orientado à conversação.</p> <p>Não, é de um tipo indefinido.</p>
SQ10. De qual forma o <i>chatbot</i> foi avaliado?	<p>As respostas obtidas são subjetivas e foram identificadas durante a leitura dos artigos. Exemplos de categorias são: experimento, observação, entre outros.</p>
SQ11. A avaliação da tecnologia foi avaliada empiricamente?	<p>Sim, foi realizado um estudo empírico para avaliar a tecnologia de avaliação;</p> <p>Não, a tecnologia não foi avaliada empiricamente.</p>
SQ12. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	<p>As respostas obtidas são subjetivas e variam de artigo para artigo. Eles podem ser estudos de viabilidade, estudos de observação, estudos de caso, entre outros.</p>
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o <i>chatbot</i>	<p>Sim, são considerados aspectos da saúde mental do usuário;</p> <p>Não, não são considerados aspectos da saúde mental do usuário.</p>
SQ14. De que forma a saúde mental do usuário foi avaliada?	<p>As respostas obtidas são subjetivas e variam de artigo para artigo. A avaliação pode ser realizada através de entrevista, questionário, consulta, entre outros.</p>

Tabela 3.4: Subquestões de pesquisa e possíveis respostas.

3.3 RELATÓRIO

Esta sessão apresenta as informações referentes ao relatório do mapeamento, que contém a avaliação dos anos de publicação dos artigos selecionados e locais de publicação dos mesmos. Os demais resultados quantitativos referentes às subquestões serão apresentados no próximo capítulo.

3.3.1 Ano de publicação

Os artigos seleccionados foram publicados entre os anos de 2017 e 2021. De acordo com o gráfico exposto na Figura 3.1 é possível notar que ocorreu um aumento no número de artigos publicados que abordam avaliação de UX em *chatbots* com o passar dos anos. O ano de 2017 foi o ano com menos publicações, apenas uma, e 2021 é o ano com mais publicações, onze. Nota-se que este tema de pesquisa é recente, dado que o artigo mais antigo é de apenas cinco anos atrás.

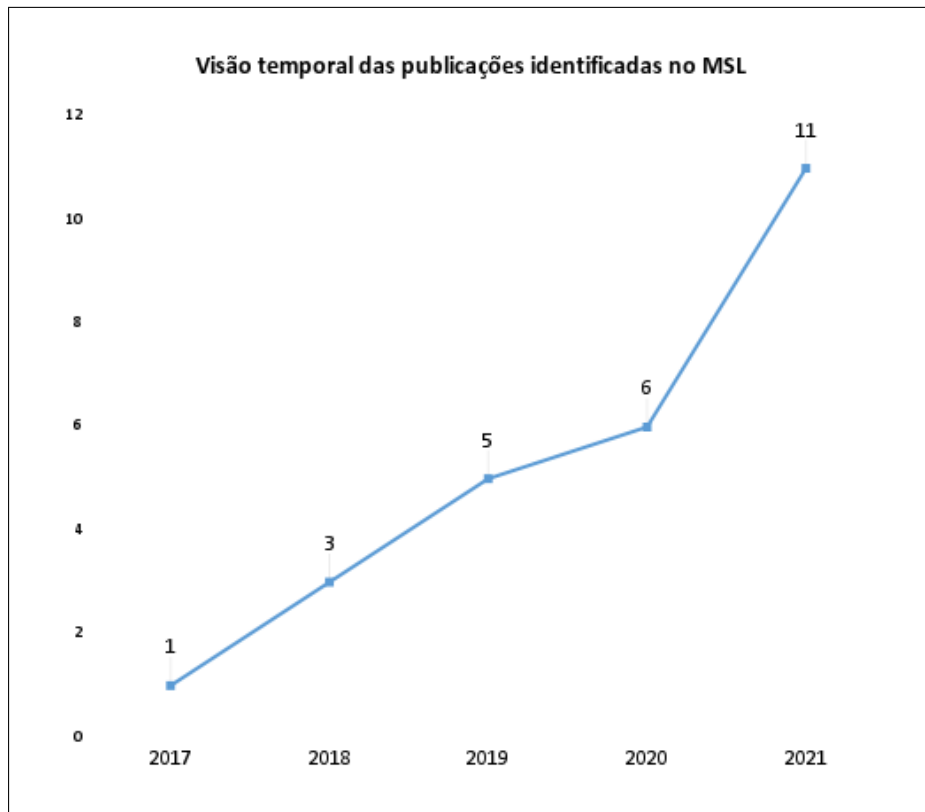


Figura 3.1: Ano de publicação dos artigos selecionados.

3.3.2 Locais de Publicação

A Figura 3.2 provê uma perspectiva global da distribuição dos artigos por conferências. A conferência com o maior número de publicações é a *Conference on Human Factors in Computing Systems - CHI*, com quatro publicações, seguida da *International Conference on Mobile Human-Computer Interaction - Mobile HCI*, *Designing Interactive Systems Conference - DIS* e *Conference on Conversational User Interfaces - CUI*, todas com duas publicações.



Figura 3.2: Distribuição de artigos por conferência.

Já na Figura 3.3 é possível observar a distribuição de artigos em periódicos. O periódico mais frequente é o *Proceedings of the ACM on Human-Computer Interaction*, em que foram retornados 3 artigos, seguido do *Mensch und Computer* e do *IEEE Transactions on Emerging Topics in Computing*, ambos com um artigo.

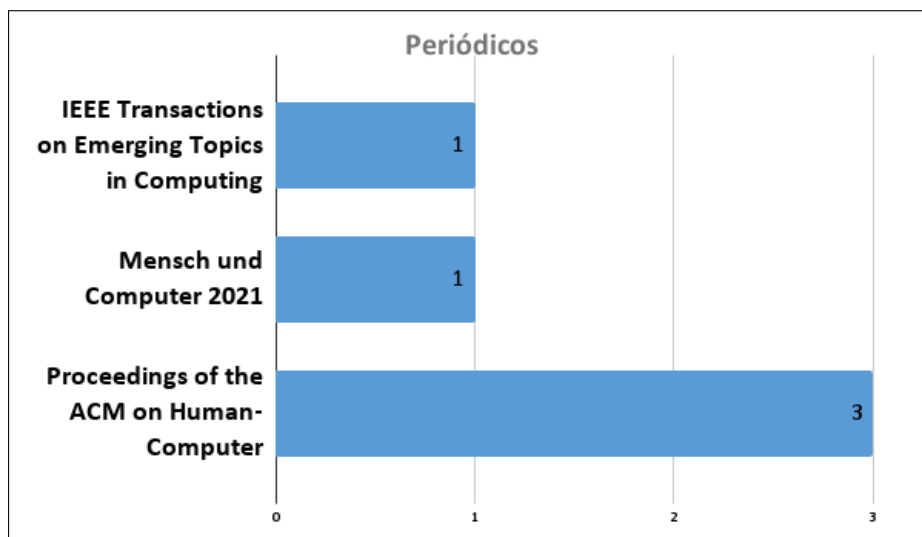


Figura 3.3: Distribuição de artigos por periódicos.

4 RESULTADOS QUANTITATIVOS E DISCUSSÃO

Esta seção apresenta os resultados obtidos a partir da extração dos dados dos artigos selecionados.

4.1 VISÃO GERAL DOS RESULTADOS

Os resultados quantitativos obtidos neste MSL foram gerados a partir das respostas de cada subquestão de pesquisa e alguns deles são expostos na Tabela 4.1. Foram encontradas 32 tecnologias de avaliação nos estudos examinados, que foram analisadas nas subquestões SQ1, SQ2, SQ3, SQ4 e SQ6. As subquestões SQ5, SQ7, SQ10, SQ12 e SQ14 são qualitativas, portanto, as avaliações dos resultados destas subquestões serão apresentados em trabalhos futuros.

As subquestões SQ8 e SQ9 revelam aspectos a respeito dos *chatbots* presentes nos artigos. Já a subquestão SQ11 revela a realização de avaliação empírica das tecnologias e a SQ13 aborda a saúde emocional dos usuários na avaliação de UX.

Vale salientar que serão apresentados somente os resultados quantitativos. A análise qualitativa será realizada em trabalhos futuros.

Subquestões de pesquisa	Possíveis respostas	Resultados	
		Tecnologias	Porcentagem
SQ2. A tecnologia é específica para <i>chatbots</i> ou para sistemas em geral?	Específica	8	25%
	Genérica	24	75%
SQ3 - A tecnologia foi criada para o estudo ou baseada numa já existente?	Existente	22	68,75%
	Criada	10	31,25%
SQ6 – A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos	26	81,25%
	Qualitativos	5	15,63%
	Mistos	1	3,13%
SQ8 – O <i>chatbot</i> foi criado para um grupo específico de pessoas?	Sim	5	19,23%
	Não	21	80,77%
SQ9 – O <i>chatbot</i> é de algum tipo específico? Qual?	Sim, é orientado a tarefa	0	0%
	Sim, o <i>chatbot</i> é orientado a conversação	18	69,23%
	Sim, é tanto orientado a conversação quanto a tarefa	8	30,77%
	Não	0	0%
SQ11 – A avaliação da tecnologia foi avaliada empiricamente?	Sim	0	0%
	Não	32	100%
SQ13 – A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o <i>chatbot</i> ?	Sim	6	23,08%
	Não	20	76,92%

Tabela 4.1: Resultados do MSL para cada uma das subquestões

4.2 SQ1 - ASPECTOS DA UX QUE AS TECNOLOGIAS AVALIAM

Ao todo foram encontrados 66 aspectos de UX que as tecnologias avaliam, sendo os mais frequentes a confiança, analisada em seis diferentes artigos, seguida do prazer, atratividade, eficiência, perspicácia, dependabilidade, estimulação, novidade e engajamento, todos estes surgem em 5 artigos distintos. Por meio do gráfico exposto na Figura 4.1 é possível observar a frequência de cada aspecto encontrado.

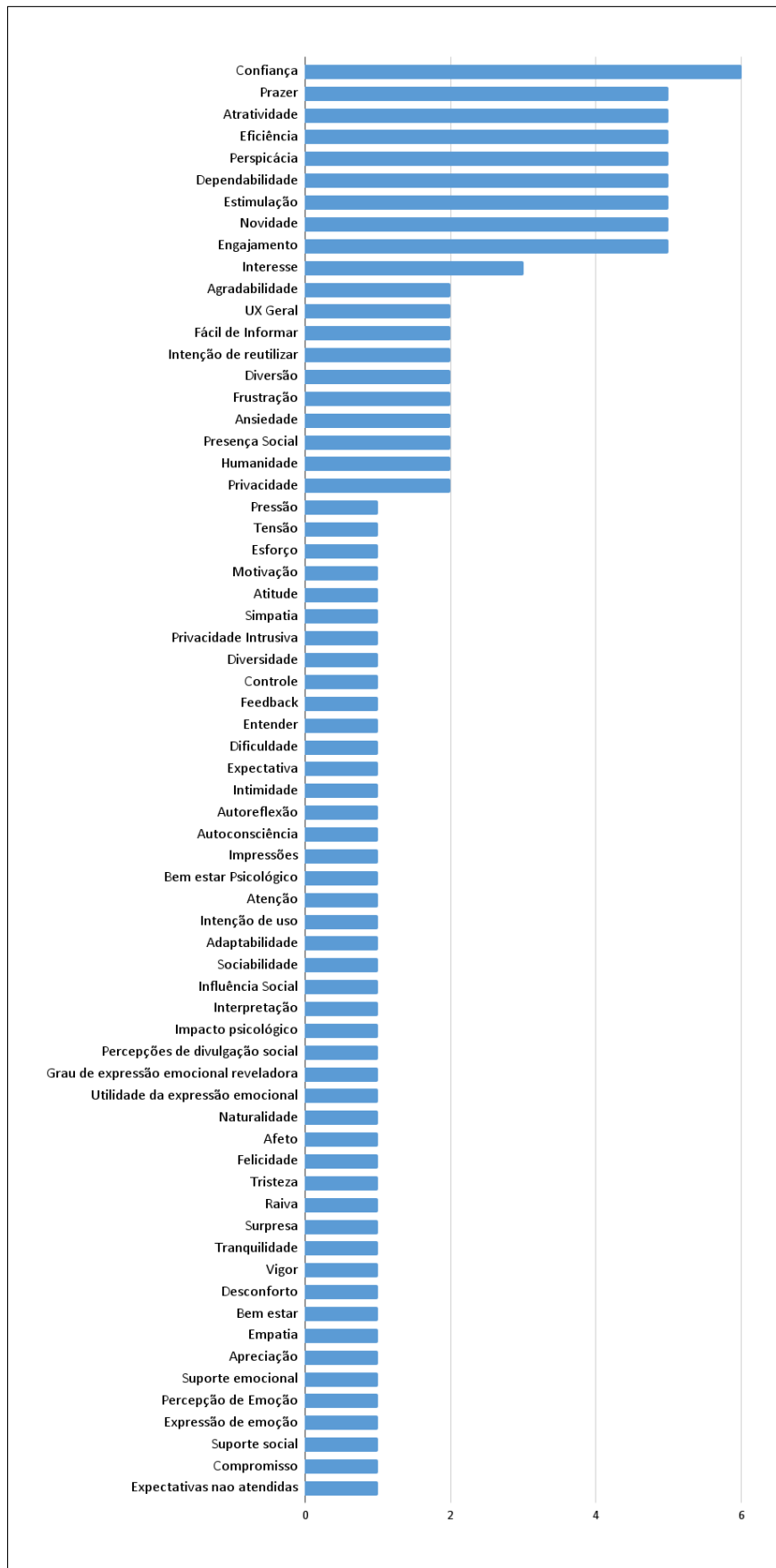


Figura 4.1: Resultados da SQL.

4.3 SQ2 - TECNOLOGIAS ESPECÍFICAS PARA *CHATBOTS* OU GENÉRICAS

Por meio do gráfico exposto na Figura 4.2, é notável que três quartos das tecnologias de avaliação de UX em *chatbots* encontradas são genéricas, equivalente a 75% das tecnologias (N = 24). Estas são tecnologias que podem ser utilizadas para avaliar sistemas em geral. Somente 25% (N = 8) das tecnologias são específicas para *chatbots*, e representam a minoria das tecnologias descobertas. Este resultado indica uma carência em tecnologias para a avaliação de UX específicas para *chatbots*. Esta carência revela que não estão sendo coletadas informações de experiência específicas de *chatbots*, a respeito de características encontradas apenas nestas tecnologias. A ausência destes dados não permite uma análise completa da UX e, conseqüentemente, podem ser perdidas possíveis melhorias de qualidade.

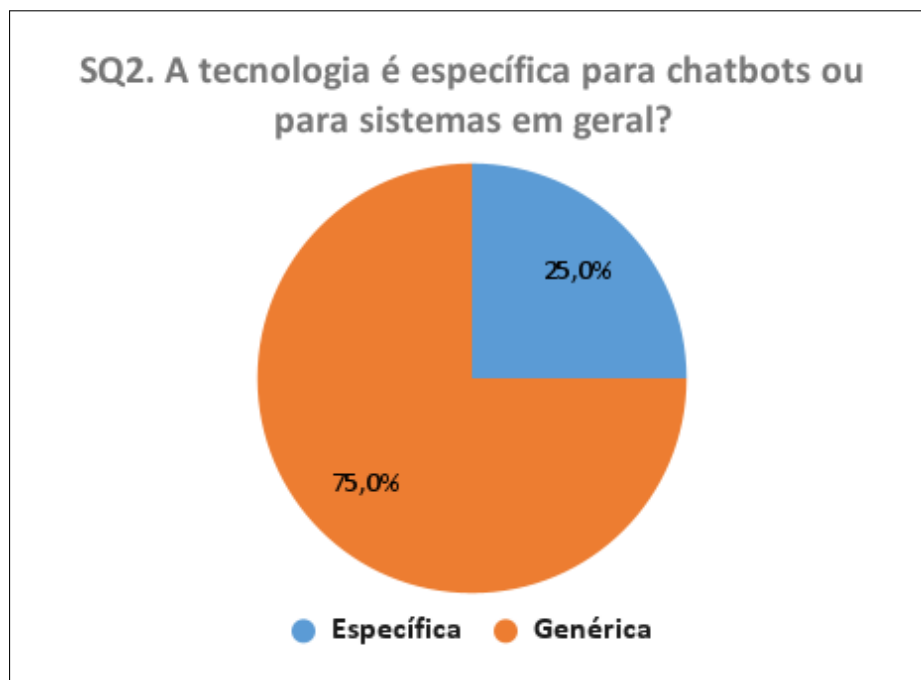


Figura 4.2: Resultados da SQ2.

A respeito da questão de pesquisa “Quais tecnologias estão sendo utilizadas para avaliar UX em *chatbots*?” foram encontradas 32 tecnologias. Destas, a mais utilizada é a de questionários criados para o estudo, presente em 13 artigos, seguida de questionários existentes, que não possuem um nome definido, mas que foram aplicados em outros estudos, seguida de entrevistas criadas para o estudo e dos questionários de experiência do usuário, *User Experience Questionnaire* (UEQ) (Laugwitz et al., 2008) e *User Experience Questionnaire - Short* (UEQ-S) (Schrepp et al., 2017). No gráfico exposto na Figura 4.3 é possível verificar as demais tecnologias encontradas e o Apêndice A apresenta a relação de todas as tecnologias encontradas e as publicações nas quais foram utilizadas.

No artigo de Jin et al. (2019) foi apresentado um questionário com 14 perguntas, desenvolvido especificamente para avaliar o *chatbot* Musicbot. Já no estudo desenvolvido por Xiao et al. (2019) foi realizada uma entrevista com os participantes a respeito do *chatbot* INDIGO, contudo, não era específica para *chatbots*.

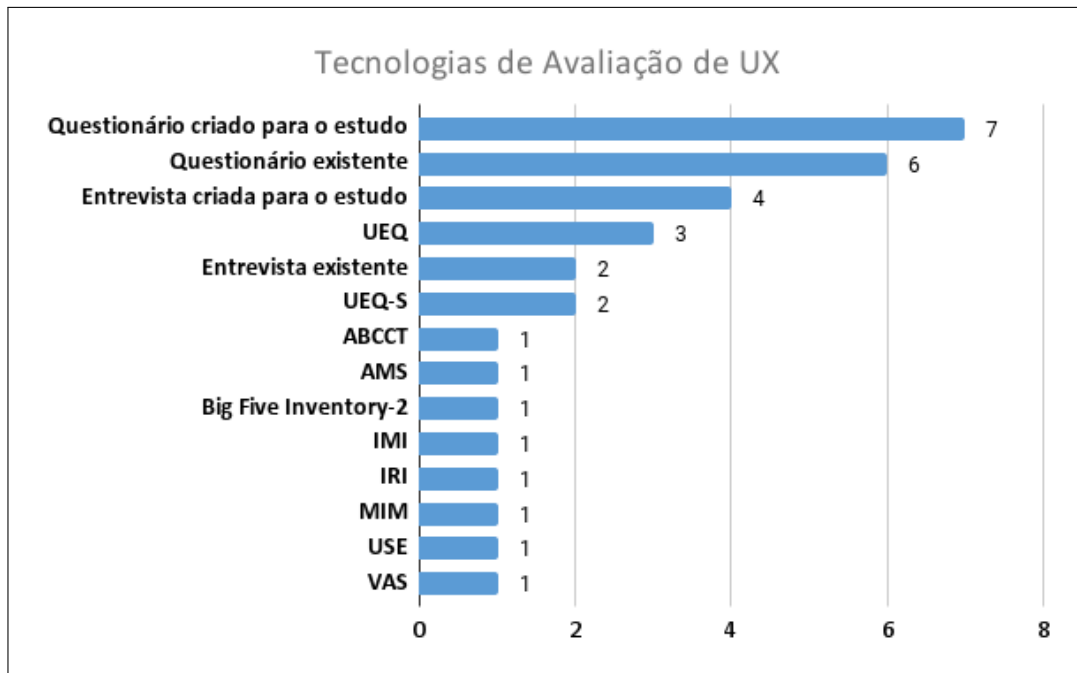


Figura 4.3: Tecnologias de Avaliação encontradas

4.4 SQ3 - TECNOLOGIAS CRIADAS PARA O ESTUDO OU EXISTENTES

O gráfico da Figura 4.4 expõe o dado que a maioria das tecnologias encontradas, equivalente a 65,6% (N = 21), já existiam. As que foram criadas para realizar a avaliação de UX compreendem apenas 34,25% (N = 11) do total. Este resultado sugere que não foram criadas muitas tecnologias para realizar a avaliação de UX em *chatbots*, e indica que é mais comum o emprego de tecnologias já existentes para realizar as avaliações.



Figura 4.4: Resultados da SQ3.

Chen et al. (2021) utilizou um questionário existente, baseado no estudo de Xu et al. (2008), para realizar avaliação de experiência do usuário perante o aspecto de privacidade intrusiva. Já (Kattenbeck et al., 2018) desenvolveu um questionário para apurar a experiência dos participantes com o Airbot.

4.5 SQ4 - COLETA DE RESPOSTAS

A respeito da coleta de respostas, foi possível constatar que o método mais aplicado é a escala Likert de 7 pontos (N = 15), seguida da escala Likert de 5 pontos (N = 10) e de questões abertas (N = 5). No gráfico exposto na Figura 4.5 é possível verificar os demais métodos encontrados.

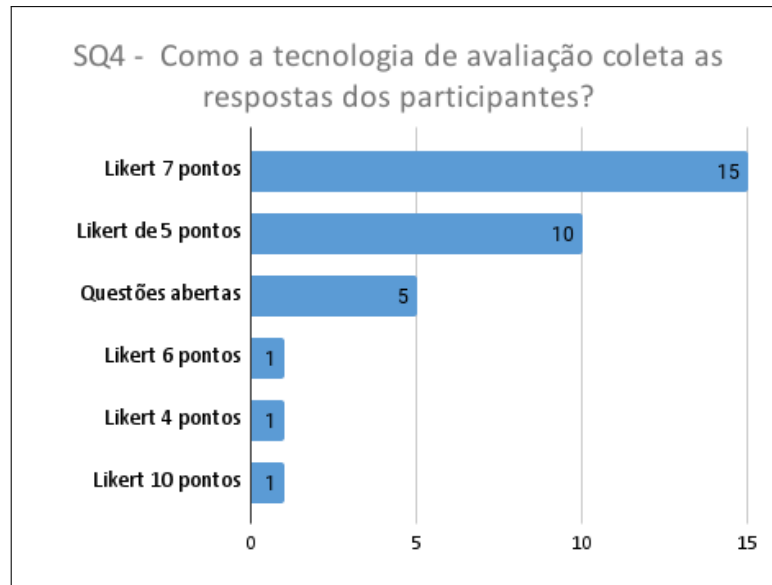


Figura 4.5: Resultados da SQ4.

A escala Likert é um conjunto de afirmações (itens), em que os participantes são solicitados a demonstrar seu nível de concordância de acordo com a afirmação dada (Joshi et al., 2015). No artigo de Park et al. (2021) foram utilizados dois métodos de coleta de respostas em uma das tecnologias de avaliação utilizada: escala Likert de 7 pontos e escala Likert de 5 pontos.

4.6 SQ6 - TIPO DE ANÁLISE

Por meio do gráfico presente na Figura 4.6 observa-se que a maioria das tecnologias extrai dados quantitativos, e compreende a 81,3% (N = 26) das tecnologias, 15,63% (N = 5) das tecnologias extraem dados qualitativos e apenas uma tecnologia extrai dados mistos.

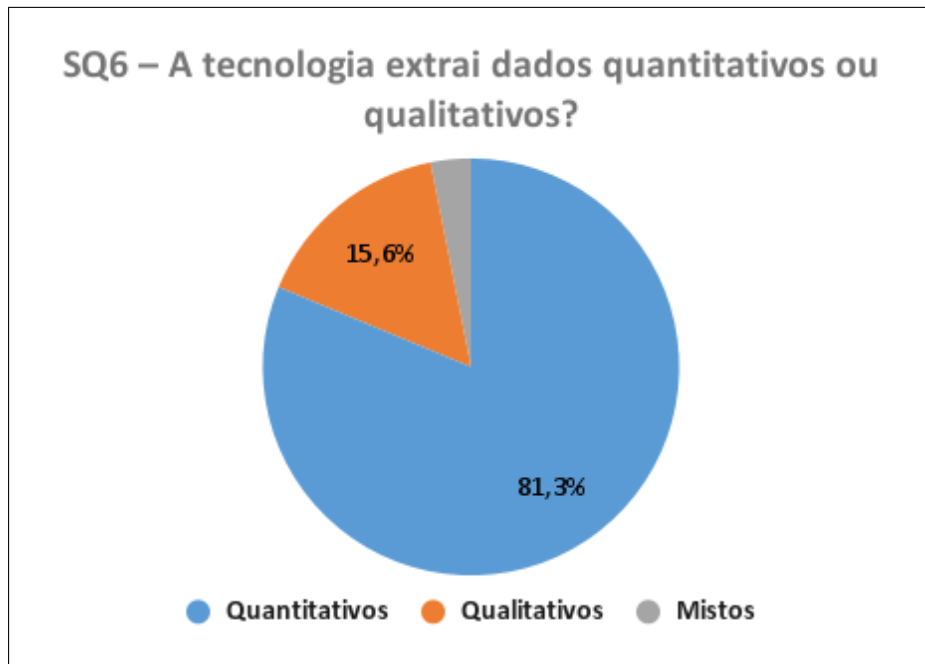


Figura 4.6: Resultados da SQ6.

O resultado é justificado pela questão SQ4, uma vez que os métodos de coleta de respostas mais frequentes são escalas numéricas quantitativas.

4.7 SQ8 - *CHATBOTS* CRIADOS PARA UM GRUPO ESPECÍFICO

Somente 19,2% (N = 5) dos *chatbots* avaliados foram criados para grupos específicos e os demais 80,77% (N = 21) não são direcionados para nenhum grupo de pessoas.

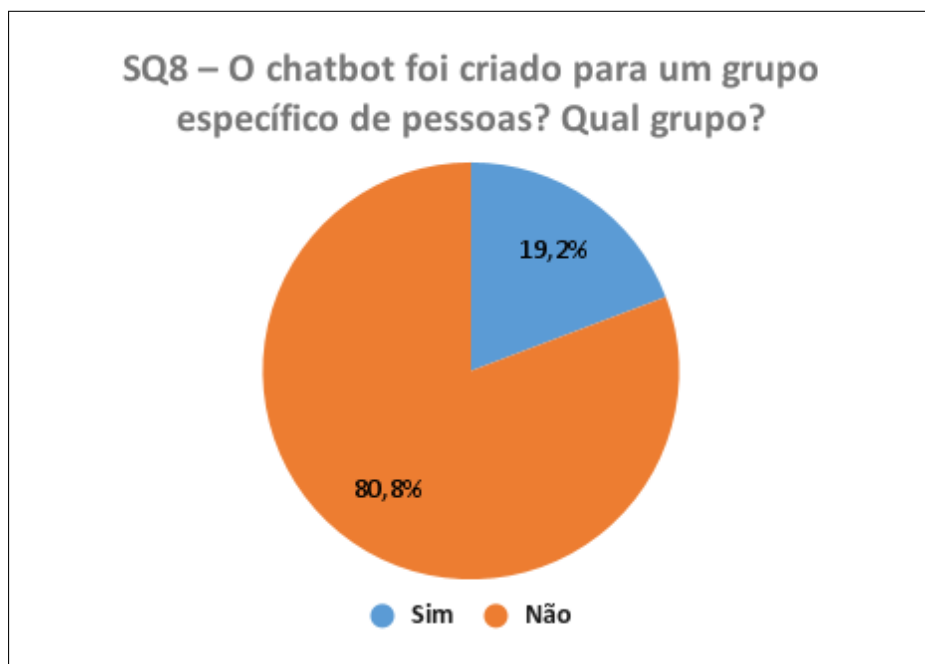


Figura 4.7: Resultados da SQ8.

Dos *chatbots* direcionados para grupos específicos, os grupos encontrados foram: idosos (N = 1), adolescentes na Coreia (N = 1) e alunos (N = 3).

O *chatbot* avaliado por El Kamali et al. (2020) foi desenvolvido para idosos e tem o objetivo de ser um *coach* motivacional. Já o *chatbot* presente no trabalho de De Nieva et al. (2020) é direcionado para alunos de forma a ajudar no alívio do estresse da carga de trabalho acadêmica.

4.8 SQ9 - TIPO DOS *CHATBOTS*

Não foi encontrado nenhum *chatbot* direcionado somente a tarefas ou sem um tipo específico. Contudo, mais da metade, correspondendo a 69,2% (N=18), dos *chatbots* é tanto orientado a conversação quanto a tarefas.

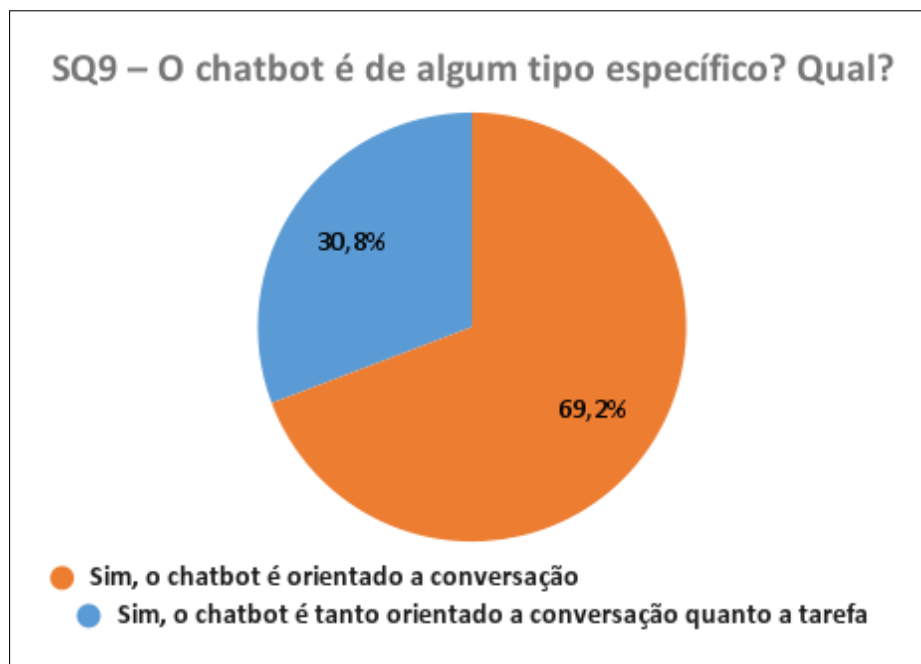


Figura 4.8: Resultados da SQ9.

Um dos chatbots orientados a conversação foi o avaliado no artigo de Denecke et al. (2020). Já o avaliado por Jain et al. (2018) foi desenvolvido tanto para conversação quanto para tarefas.

4.9 SQ11 - AVALIAÇÃO EMPÍRICA DAS TECNOLOGIAS

Os resultados revelaram que nenhuma das tecnologias de avaliação de UX em *chatbots* encontradas foi avaliada empiricamente. Isto se deu porque nenhum dos artigos avaliados era focado na tecnologia de avaliação de UX e sim na avaliação de UX de um, ou mais, *chatbots*.

4.10 SQ13 - ASPECTOS DA SAÚDE EMOCIONAL NA AVALIAÇÃO DE UX

De todas as avaliações de UX em *chatbots* 23,1% (N = 6) considerou algum aspecto da saúde emocional do usuário que está interagindo com o *chatbot*. Dos que consideraram, a maioria aplicou questionários voltados para avaliação da saúde mental, bem-estar físico e psicológico, ou considerou algum aspecto da saúde emocional o usuário durante a avaliação de UX, como, por exemplo, autorreflexão e autoconsciência.

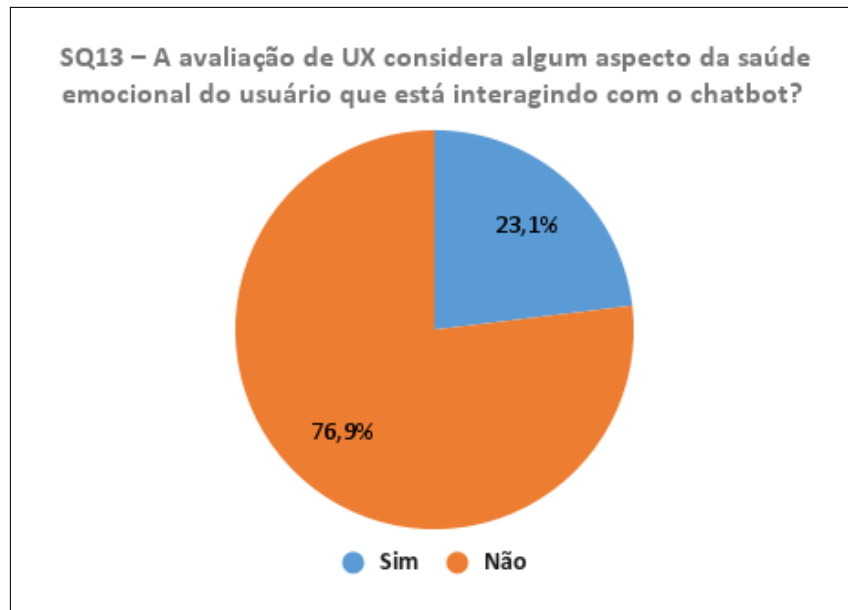


Figura 4.9: Resultados da SQ13.

Na avaliação de UX realizada no estudo de Yun et al. (2020) são avaliados os aspectos: revelação da expressão emocional e utilidade da expressão emocional. Já no artigo de Benke et al. (2020) são considerados os aspectos emoção percebida e expressão da emoção. Ambos refletem a saúde emocional do usuário.

4.11 DISCUSSÃO

Analisando o conjunto de resultados relacionados a tecnologias de avaliação de UX, pode-se verificar que são muitos os aspectos que podem ser avaliados e que os mais frequentes são aqueles presentes em tecnologias de avaliação já consagradas, como, Atratabilidade, Eficiência, Perspicácia, Dependabilidade, Estimulação e Novidade. A respeito disso, foi detectado que na maioria dos artigos a tecnologia utilizada era uma já existente e das onze tecnologias criadas sete são questionários e quatro são entrevistas, ou seja, pouca variedade no formato da tecnologia.

Apenas um quarto das tecnologias são específicas para *chatbots* o que demonstra pouca especificidade nas avaliações deste tipo de sistema. Isto pode impactar no resultado da avaliação, uma vez que, em diversos estudos não são examinadas características particulares deste tipo de sistemas conversacionais.

A respeito da forma que as respostas são coletadas também não há grande variação, uma vez que a maioria das tecnologias utiliza escala Likert, quando quantitativas, ou questões abertas, se qualitativas. Por tal razão, é evidente que a maioria das tecnologias coleta respostas de maneira quantitativa, o que impacta na especificidade das avaliações, uma vez que as respostas qualitativas permitem que o usuário exponha, de forma mais detalhada, como foi sua experiência e também permite que sejam melhor explorados os problemas encontrados nos *chatbots*.

Considerando os *chatbots* que foram avaliados nos artigos selecionados foi reconhecido que a minoria é específica para um grupo de usuários e destes foram identificados apenas três grupos: idosos, alunos e adolescentes da Coreia. Esta descoberta evidencia que quase a totalidade dos sistemas conversacionais encontrados no MSL são projetados e desenvolvidos para qualquer tipo de usuário.

Também identificamos que nenhum dos *chatbots* é direcionado somente a tarefas, isto ocorre porque nos sistemas conversacionais avaliados todos utilizam texto e até para realizar tarefas é necessário o mínimo de conversação, o que torna compreensível a categoria mais frequente ser “*Chatbot* tanto orientado a conversação quanto tarefa”.

Nenhum dos artigos examinados apresentou avaliação empírica das tecnologias e é compreensível, pois todos os artigos tratam da avaliação de UX em *chatbots* e não da avaliação de tecnologias de UX em *chatbots*. Entretanto, nos estudos em que foram criadas tecnologias elas poderiam ter sido minimamente avaliadas, com o propósito de assegurar a qualidade da tecnologia de avaliação proposta.

5 CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho visou descobrir quais tecnologias estão sendo utilizadas para avaliar UX em *chatbots*, através de um mapeamento sistemático da literatura, e foram descobertas 32 tecnologias. Observando os anos de publicação dos estudos avaliados constata-se que o tema, avaliação de UX em *chatbots*, é recente, uma vez que o trabalho mais antigo investigado é de 2017. Este dado nos leva a concluir que ainda há muitas possibilidades de trabalhos futuros que podem explorar os tópicos de tecnologias de avaliação de experiência do usuário para *chatbots* e que as análises ao redor do assunto estão apenas no início.

Considerando os resultados e as análises realizadas é possível concluir que existem lacunas nas tecnologias de avaliação de UX em *chatbots*. Uma delas é a ausência de estudos que avaliam tecnologias de avaliação de UX em *chatbots*, uma vez que, nenhum dos artigos explorados realiza avaliação das tecnologias de avaliação empregadas.

Outra lacuna evidente é a de tecnologias específicas para a avaliação de UX em *chatbots* e tal constatação aponta que em diversos estudos não foram considerados aspectos particulares de *chatbots* no momento de apurar a experiência do usuário. Logo, nota-se a necessidade na criação de tecnologias específicas para a avaliação de experiência do usuário em *chatbots*, de modo que, seja possível extrair resultados direcionados que podem contribuir nas melhorias e aprimoramentos dos *chatbots*.

Um dos resultados mais surpreendentes foi a presença de avaliação de aspectos da saúde emocional na avaliação de UX em seis dos artigos julgados, que representa quase um quarto do total. Este resultado é expressivo, pois demonstra que existiu uma preocupação, por parte dos autores, em examinar o bem-estar psicológico dos usuários, de maneira que a saúde emocional não interferisse na experiência do usuário com o *chatbot*, logo, na avaliação de UX.

Dentre os trabalhos futuros que podem ser desenvolvidos a partir deste TG está a análise qualitativa das informações extraídas através das questões SQ5, SQ7, SQ10, SQ12 e SQ14 e também uma proposta de tecnologia de avaliação de UX específica para *chatbots*, assim como, uma avaliação empírica da mesma.

REFERÊNCIAS

- Ashktorab, Z., Jain, M., Liao, Q. V. e Weisz, J. D. (2019). Resilient chatbots: Repair strategy preferences for conversational breakdowns. Em *Proceedings of the 2019 CHI conference on human factors in computing systems*, páginas 1–12.
- Bae Brandtzæg, P. B., Skjuve, M., Kristoffer Dysthe, K. K. e Følstad, A. (2021). When the social becomes non-human: Young people’s perception of social support in chatbots. Em *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, páginas 1–13.
- Barbosa, S. e Silva, B. (2010). *Interação humano-computador*. Elsevier Brasil.
- Basili, V. e Rombach, H. (1988)). Towards a comprehensive framework for reuse: A re-use-enabling software evolution environment. Relatório técnico, University of Maryland.
- Bawa, A., Khadpe, P., Joshi, P., Bali, K. e Choudhury, M. (2020). Do multilingual users prefer chat-bots that code-mix? let’s nudge and find out! *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1):1–23.
- Benke, I., Knierim, M. T. e Maedche, A. (2020). Chatbot-based emotion management for distributed teams: A participatory design study. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–30.
- Brodwin, E. (2018). I spent 2 weeks texting a bot about my anxiety — and found it to be surprisingly helpful.
- Brooke, J. (1996). Sus: a “quick and dirty’ usability. *Usability evaluation in industry*, 189(3).
- Campos, T., Damasceno, E. e Valentim, N. (2022). Proposal and evaluation of a collaborative is to support systematic reviews and mapping studies. *Simpósio Brasileiro de Sistemas de Informação (SBSI 2022)*.
- Candello, H. e Pinhanez, C. (2016). *Designing Conversational Interfaces*, página <http://www.lbd.dcc.ufmg.br/bdbcomp/servlet/Trabalho?id=25402>.
- Castro, J. W., Ren, R., Acuña, S. T. e Lara, J. d. (2019). Usability of chatbots: A systematic mapping study.
- Ceha, J., Lee, K. J., Nilsen, E., Goh, J. e Law, E. (2021). Can a humorous conversational agent enhance learning experience and outcomes? Em *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, páginas 1–14.
- Chen, J., Chen, C., B. Walther, J. e Sundar, S. S. (2021). Do you feel special when an ai doctor remembers you? individuation effects of ai vs. human doctors on user experience. Em *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, páginas 1–7.
- Dahiya, M. (2017). A tool of conversation: Chatbot. *International Journal of Computer Sciences and Engineering*, 5(5):158–161.

- De Nieva, J. O., Joaquin, J. A., Tan, C. B., Marc Te, R. K. e Ong, E. (2020). Investigating students' use of a mental health chatbot to alleviate academic stress. Em *6th International ACM In-Cooperation HCI and UX Conference*, páginas 1–10.
- Denecke, K., Vaaheesan, S. e Arulnathan, A. (2020). A mental health chatbot for regulating emotions (sermo)-concept and usability test. *IEEE Transactions on Emerging Topics in Computing*, 9(3):1170–1182.
- El Kamali, M., Angelini, L., Lalanne, D., Abou Khaled, O. e Mugellini, E. (2020). Multimodal conversational agent for older adults' behavioral change. Em *Companion Publication of the 2020 International Conference on Multimodal Interaction*, páginas 270–274.
- Elsholz, E., Chamberlain, J. e Kruschwitz, U. (2019). Exploring language style in chatbots to increase perceived product value and user engagement. Em *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*, páginas 301–305.
- Fadhil, A., Schiavo, G., Wang, Y. e Yilma, B. A. (2018). The effect of emojis when interacting with conversational interface assisted health coaching system. Em *Proceedings of the 12th EAI international conference on pervasive computing technologies for healthcare*, páginas 378–383.
- Fahn, V. e Riener, A. (2021). Time to get conversational: Assessment of the potential of conversational user interfaces for mobile banking. Em *Mensch und Computer 2021*, páginas 34–43.
- Fiore, D., Baldauf, M. e Thiel, C. (2019). "forgot your password again?" acceptance and user experience of a chatbot for in-company it support. Em *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, páginas 1–11.
- Fitzpatrick, K. K., Darcy, A. e Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. *JMIR mental health*, 4(2):e7785.
- Flohr, L. A., Kalinke, S., Krüger, A. e Wallach, D. P. (2021). Chat or tap?—comparing chatbots with 'classic' graphical user interfaces for mobile interaction with autonomous mobility-on-demand systems. Em *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*, páginas 1–13.
- Guerino, G. C. e Valentim, N. M. C. (2020). Usability and user experience evaluation of conversational systems: A systematic mapping study. Em *Proceedings of the 34th Brazilian Symposium on Software Engineering*, páginas 427–436.
- Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction*, 19(4):319–349.
- Hassenzahl, M. (2018). The thing and i (summer of' 17 remix). Em *Funology 2*, páginas 17–31. Springer.
- Hassenzahl, M., Platz, A., Burmester, M. e Lehner, K. (2000). Hedonic and ergonomic quality aspects determine a software's appeal. Em *Proceedings of the SIGCHI conference on Human factors in computing systems*, páginas 201–208.

- Hassenzahl, M. e Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & information technology*, 25(2):91–97.
- International Organization for Standardization (2019). ISO 9241-210:2019(en) Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. <https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en>. Online; accessed 16 September 2021.
- Jain, M., Kumar, P., Kota, R. e Patel, S. N. (2018). Evaluating and informing the design of chatbots. Em *Proceedings of the 2018 Designing Interactive Systems Conference*, páginas 895–906.
- Jia, J. (2003). The study of the application of a keywords-based chatbot system on the teaching of foreign languages. *arXiv preprint cs/0310018*.
- Jin, Y., Cai, W., Chen, L., Htun, N. N. e Verbert, K. (2019). Musicbot: Evaluating critiquing-based music recommenders with conversational interaction. Em *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, páginas 951–960.
- Joshi, A., Kale, S., Chandel, S. e Pal, D. K. (2015). Likert scale: Explored and explained. *British journal of applied science & technology*, 7(4):396.
- Kattenbeck, M., Kilian, M. A., Ferstl, M., Alt, F. e Ludwig, B. (2018). Airbot: using a work flow model for proactive assistance in public spaces. Em *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, páginas 213–220.
- Khyani, D., Siddhartha, B., Niveditha, N. e Divya, B. (2020). An interpretation of lemmatization and stemming in natural language processing. *Shanghai Ligong Daxue Xuebao/Journal of University of Shanghai for Science and Technology*, 22:350–357.
- Kim, S., Lee, J. e Gweon, G. (2019). Comparing data from chatbot and web surveys.
- Kitchenham, B. e Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. Relatório técnico, Technical report, Ver. 2.3 EBSE Technical Report. EBSE.
- Laugwitz, B., Held, T. e Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. Em *Symposium of the Austrian HCI and usability engineering group*, páginas 63–76. Springer.
- Lee, Y.-C., Yamashita, N. e Huang, Y. (2021). Exploring the effects of incorporating human experts to deliver journaling guidance through a chatbot. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–27.
- Liu, Y., Kim, D.-j., Miao, T. e Chuang, Y. (2020). Slumberbot: An interactive agent for helping users investigate disturbance factors of sleep quality. Em *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, páginas 1–4.
- MarutiTechlab (2021). How do chatbots work? a guide to chatbot architecture. [urlhttps://marutitech.com/chatbots-work-guide-chatbot-architecture/How_dOChatbotsWork](https://marutitech.com/chatbots-work-guide-chatbot-architecture/How_dOChatbotsWork).

- Norman, D. A. e Love, E. D.-W. W. (2004). *Everyday Things*. Basic books.
- Park, S., Thieme, A., Han, J., Lee, S., Rhee, W. e Suh, B. (2021). “i wrote as if i were telling a story to someone i knew.”: Designing chatbot interactions for expressive writing in mental health. Em *Designing Interactive Systems Conference 2021*, páginas 926–941.
- Peter, C. e Beale, R. (2008). *Affect and emotion in human-computer interaction: From theory to applications*, volume 4868. Springer Science & Business Media.
- Piccolo, L. S. G., Hayashi, E. C. e Baranauskas, M. C. C. (2010). The evaluation of affective quality in social software: preliminary thoughts. *II WAIHCWS, IHC2010*, páginas 29–38.
- Portela, M. e Granell-Canut, C. (2017). A new friend in our smartphone? observing interactions with chatbots in the search of emotional engagement. Em *Proceedings of the XVIII International Conference on Human Computer Interaction*, páginas 1–7.
- Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P. e Mazurek, G. (2019). In bot we trust: A new methodology of chatbot performance measures. *Business Horizons*, 62(6):785–797.
- Rapp, A., Curti, L. e 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*, página 102630.
- Scherer, K. R. (2005). What are emotions? and how can they be measured? *Social science information*, 44(4):695–729.
- Schrepp, M., Hinderks, A. e Thomaschewski, J. (2017). Design and evaluation of a short version of the user experience questionnaire (ueq-s). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4 (6), 103-108.
- Shawar, B. A. e Atwell, E. (2007). Chatbots: are they really useful? Em *Ldv forum*, volume 22, páginas 29–49.
- Tubin, C., Mazuco Rodriguez, J. P. e de Marchi, A. C. B. (2021). User experience with conversational agent: a systematic review of assessment methods. *Behaviour & Information Technology*, páginas 1–11.
- Völkel, S. T. e Kaya, L. (2021). Examining user preference for agreeableness in chatbots. Em *CUI 2021-3rd Conference on Conversational User Interfaces*, páginas 1–6.
- Wald, R., Heijselaar, E. e Bosse, T. (2021). Make your own: The potential of chatbot customization for the development of user trust. Em *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, páginas 382–387.
- Wambsganss, T., Küng, T., Söllner, M. e Leimeister, J. M. (2021). Arguetutor: an adaptive dialog-based learning system for argumentation skills. Em *Proceedings of the 2021 CHI conference on human factors in computing systems*, páginas 1–13.
- Weizenbaum, J. (1966). Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45.
- Xavier, R. A. C. e de Almeida Neris, V. P. (2012). A hybrid evaluation approach for the emotional state of information systems users. Em *ICEIS (3)*, páginas 45–53.

- Xiao, Z., Zhou, M. X. e Fu, W.-T. (2019). Who should be my teammates: Using a conversational agent to understand individuals and help teaming. Em *Proceedings of the 24th International Conference on Intelligent User Interfaces*, páginas 437–447.
- Xu, H., Dinev, T., Smith, H. J. e Hart, P. (2008). Examining the formation of individual's privacy concerns: Toward an integrative view.
- Yun, H., Ham, A., Kim, J., Kim, T., Kim, J., Lee, H., Park, J. e Jang, J. (2020). Chatbot with touch and graphics: An interaction of users for emotional expression and turn-taking. Em *Proceedings of the 2nd Conference on Conversational User Interfaces*, páginas 1–5.

APÊNDICE A – TECNOLOGIAS DE AVALIAÇÃO ENCONTRADAS

Tecnologia de Avaliação	Publicações
Questionário criado para o estudo	Fadhil et al. (2018) Jin et al. (2019) Jain et al. (2018) Yun et al. (2020) De Nieva et al. (2020) Bawa et al. (2020) Kattenbeck et al. (2018)
Questionário existente	Elsholz et al. (2019) Lee et al. (2021) Wald et al. (2021) Park et al. (2021) Wambsganss et al. (2021)
Entrevista criada para o estudo	Xiao et al. (2019) Benke et al. (2020) Liu et al. (2020) Bae Brandtzæg et al. (2021)
User Experience Questionnaire - UEQ	Fiore et al. (2019) Flohr et al. (2021) Denecke et al. (2020)
Entrevista existente	Lee et al. (2021) Park et al. (2021)
User Experience Questionnaire - Short -UEQ-S	El Kamali et al. (2020) Fahn e Riener (2021)
Affective Benefits and Costs of Communication Technology - ABCCT	Benke et al. (2020)
Academic motivation scale - AMS	Ceha et al. (2021)
Big Five Inventory-2	Völkel e Kaya (2021)
Intrinsic motivation inventory - IMI	Ceha et al. (2021)
Interpersonal Reactivity Index - IRI	Portela e Granell-Canut (2017)
Multidimensional Integrative Model - MIM	Portela e Granell-Canut (2017)
Usefulness, Satisfaction, and Ease of use - USE	Kim et al. (2019)
Visual Analogue Scale - VAS	Portela e Granell-Canut (2017)

Tabela A.1: Tecnologias de avaliação encontradas

APÊNDICE B – ARTIGOS EXTRAÍDOS NO MSL

009 - Can a Humorous Conversational Agent Enhance Learning Experience and Outcomes?

Jessy Ceha, Ken Jen Lee, Elizabeth Nilsen, Joslin Goh, and Edith Law

Maio 2021

CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems

Tecnologia de Avaliação Tecnologia 1: Online IMI (Intrinsic motivation inventory)
Tecnologia 2:AMS (Academic motivation scale)

SQ1. Quais aspectos da UX a tecnologia avalia?	Tecnologia 1: interesse/prazer, pressão/tensão e o esforço Tecnologia 2: Motivação
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Ambas Genéricas: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Ambas Existentes: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Tecnologia 1: Através de escala de 1 até 7 Tecnologia 2: Através de escala de 1 até 7
SQ5. Qual a composição da tecnologia de avaliação?	IMI Robert J Vallerand, Marc R Blais, Nathalie M Brière, and Luc G Pelletier. 1989. Construction et validation de l'échelle de motivation en éducation (EME). Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement 21, 3 (1989), 323. https://doi.org/10.1037/h0079855 Interesse\Prazer <ul style="list-style-type: none">• Gostei muito de fazer essa atividade;• Esta atividade foi divertida de fazer;• Achei essa atividade chata. ®• Essa atividade não me chamou a atenção. (R)• Eu descreveria esta atividade como muito interessante.• Achei muito legal essa atividade.• Enquanto fazia esta atividade, pensava no quanto gostava dela. Pressão/Tensão <ul style="list-style-type: none">• Eu não me senti nervoso ao fazer isso. (R)

- Senti-me muito tenso ao fazer esta atividade.
- Eu estava muito relaxado em fazer isso. (R)
- Eu estava ansioso enquanto trabalhava nesta tarefa.
- Eu me senti pressionado ao fazer isso.

Esforço

- Eu me esforcei muito nisso.
- Eu não me esforcei muito para me sair bem nesta atividade. (R)
- Eu me esforcei muito nessa atividade.
- Era importante para mim fazer bem nesta tarefa.
- Eu não coloquei muita energia nisso.

AMS

Robert J Vallerand, Luc G Pelletier, Marc R Blais, Nathalie M Briere, Caroline Senecal, and Evelyne F Vallieres. 1992. The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement* 52, 4 (1992), 1003–1017. <https://doi.org/10.1177/0013164492052004025>

1. Porque com apenas o ensino médio eu não encontrar um emprego bem remunerado mais tarde.
2. Porque sinto prazer e satisfação enquanto aprende coisas novas.
3. Porque eu acho que uma educação universitária vai me ajudar me preparar melhor para a carreira que escolhi.
4. Pelos sentimentos intensos que experimento quando estou comunicar as minhas próprias ideias aos outros.
5. Sinceramente, não sei; Eu realmente sinto que estou desperdiçando meu tempo na escola.
6. Pelo prazer que sinto ao superar eu mesmo em meus estudos.
7. Para provar a mim mesmo que sou

	<p>capaz de completar meu Grau acadêmico.</p> <p>8. Para obter mais tarde um emprego de maior prestígio.</p> <p>9. Pelo prazer que sinto quando descobro coisas novas nunca vistas antes.</p> <p>10. Porque eventualmente me permitirá entrar no mercado de trabalho em uma área que eu gosto.</p> <p>11. Pelo prazer que sinto ao ler autores interessantes.</p> <p>12. Certa vez tive bons motivos para ir para a faculdade; no entanto, agora me pergunto se devo continuar.</p> <p>13. Pelo prazer que sinto enquanto estou superando eu mesmo em uma de minhas realizações pessoais.</p> <p>14. Por causa do fato de que quando eu tiver sucesso na faculdade Eu me sinto importante.</p> <p>15. Porque eu quero ter "uma vida boa" mais tarde.</p> <p>16. Pelo prazer que sinto em ampliar meus conhecimentos sobre assuntos que me atraem.</p> <p>17. Porque isso me ajudará a fazer uma escolha melhor em relação à minha orientação profissional.</p> <p>18. Pelo prazer que sinto quando me sinto completamente absorvido pelo que certos autores escreveram.</p> <p>19. Não consigo ver por que vou para a faculdade e, francamente, Eu não poderia me importar menos.</p> <p>20. Pela satisfação que sinto quando estou no processo de realizar atividades acadêmicas difíceis.</p> <p>21. Para me mostrar que sou uma pessoa inteligente.</p> <p>22. Para ter um salário melhor depois.</p> <p>23. Porque meus estudos me permitem continuar aprendendo sobre muitas coisas que me interessam.</p> <p>24. Porque acredito que alguns anos adicionais de educação melhorará minha competência como trabalhador.</p> <p>25. Para a sensação "alta" que experimento ao ler sobre vários assuntos interessantes.</p>
--	---

	<p>26. Não sei; Eu não consigo entender o que eu sou fazendo na escola.</p> <p>27. Porque a faculdade me permite experimentar uma satisfação pessoal na minha busca pela excelência nos meus estudos.</p> <p>28. Porque quero me mostrar que posso ter sucesso nos meus estudos.</p>
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	<p>Tecnologia 1: Quantitativos.</p> <p>Tecnologia 2: Quantitativos.</p>
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	É utilizado para ensinar a classificar rochas.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é tanto orientado à tarefas quanto orientado à conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi avaliado por meio de experimento controlado, com 58 participantes, 35 mulheres, 21 homens, 2 não-binários; idades entre 18 e 34 anos, média 24.4 anos e mediana 25 anos. Todos os voluntários responderam questionário demográfico, responderam um quiz sobre rochas, assistiram um vídeo sobre como o chatbot funciona e qual a tarefa deles. Utilizaram o modo de ensinar por 40 minutos e após isso responderam os questionários sobre a percepção. Foi dado um total de 90 minutos para realizar a sequência descrita acima.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

012 - Multimodal Conversational Agent for Older Adults' Behavioral Change

Mira El Kamali, Leonardo Angelini, Denis Lalanne, Omar Abou Khaled, Elena Mugellini

Outubro 2020

ICMI '20 Companion: Companion Publication of the 2020 International Conference on Multimodal Interaction

Tecnologia de Avaliação: UEQ-s

SQ1. Quais aspectos da UX a tecnologia avalia?	Atratabilidade, Eficiência, Perspicácia Dependabilidade, Estimulação Novidade
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de diferencial semântico. É medida numa escala de 1 até 7: -3 (muito ruim) até 3 (muito bom).
SQ5. Qual a composição da tecnologia de avaliação?	UEQ-s Foram avaliados os itens <ul style="list-style-type: none">• claro / confuso• ineficiente / eficiente• complicado / fácil• obstrutivo / solidario• chato / animador• não é interessante / interessante• convencional / inventivo• usual / de ponta
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Coach motivacional para idosos, chamado NETSTORE
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Sim. Idosos
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado à conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi realizado um experimento controlado com sete participantes; 3 mulheres e 4 homens; Idades entre 53 e 78 anos. Produziram 5 vídeos retratando cada cenário e uma pesquisa na web foi elaborada e difundida para potenciais idosos usuários do sistema NESTORE. A

	<p>pesquisa apresentou primeiro o sistema NESTORE e as duas interfaces de conversação. Em seguida, apresentou cada cenário, acompanhado de um questionário UEQ-S. No final da pesquisa, os participantes foram convidados a classificar os 5 cenários de interação, com base em sua preferência pessoal. Eles também tiveram que explicar em duas perguntas abertas o que mais gostaram da melhor interface e o que não gostaram da pior.</p>
<p>Avaliação Empírica da Tecnologia de Avaliação</p>	
<p>SQ12. A avaliação da tecnologia foi avaliada empiricamente?</p>	<p>Não</p>
<p>SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?</p>	<p>N.A.</p>
<p>Avaliação do Estado de saúde emocional do usuário</p>	
<p>SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?</p>	<p>Sim</p>
<p>SQ14. De que forma a saúde mental do usuário foi avaliada?</p>	<p>O usuário interage por meio do questionário HAPA (Health Action Process Approach), utilizado no projeto NETSTORE, para avaliar o impacto da intervenção de mudança comportamental em diferentes domínios do bem-estar.</p>

015 - The Effect of Emojis when interacting with Conversational Interface Assisted Health Coaching System

Ahmed Fadhil, Gianluca Schiavo, Yunlong Wang, Bereket A. Yilma

Maio 2018

Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare.

Tecnologia de Avaliação: Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	Prazer, Atitude e Confiança.
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Específica: tecnologia de avaliação de UX específica para chatbots.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Coleta através de questionário e utiliza escala Likert de 6 pontos (de "concordo totalmente" a "discordo totalmente").
SQ5. Qual a composição da tecnologia de avaliação?	Experiência de interação: prazer, atitude e confiança. A interação foi avaliada por meio de três escalas, ou seja, engajamento, atitude e confiança, cada uma composta por quatro itens com uma escala de classificação Likert de 6 pontos. Questões: Eu gostava de conversar com o conversador agente durante a interação Senti-me à vontade para responder às perguntas Quanto mais eu interagia com o agente, mais mais gostei da experiência Eu gostaria de conversar novamente com este convertido agente sazonal no futuro Encontrei uma espécie de "conexão emocional" entre mim e a conversa agente Encontrei o diálogo com a conversa agente para ser realista Notei uma mudança emocional negativa no eu mesmo durante a interação * Achei o diálogo coerente. Em outros palavras, a sequência de respostas do

	<p>parceiro de conversa fez sentido</p> <p>O agente perguntou questões muito pessoais</p> <p>Eu confio no agente de conversação</p> <p>Achei as perguntas muito intrusivas *</p> <p>Respondi as perguntas com sinceridade</p>
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	<p>É um chatbot que coleta dados de bem estar mental e físico, através de dois questionarios:</p> <p>General Health Questionnaire - GHQ [21] and the Global Physical Activity Questionnaire - GPAQ [2].</p>
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	<p>Foi realizado experimento, com 58 participantes aleatórios; 28 homens e 30 mulheres; Idades entre 18 e 60 anos; Metade respondeu a um questionário escrito e metade participou de entrevista. Participantes foram instruídos como realizar o teste e foi solicitado que respondessem ao questionário para mensuração da autoconfiança e confiança individual. Após o preenchimento do questionário, os participantes iniciaram o primeiro diálogo com o chatbot usando o aplicativo de desktop Telegram. Ao final de cada interação os participantes preencheram um questionário para estimar engajamento, atitude e confiança no agente conversacional. Após a última interação, os participantes foram responderam às questões comparando a experiência com os dois diálogos. Todo o procedimento durou cerca de 20 minutos, com tempo médio de 4 minutos de interação por diálogo.</p>
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do	N.A

experimento conduzido para avaliar a tecnologia de avaliação?	
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Sim
SQ14. De que forma a saúde mental do usuário foi avaliada?	Utilizam dois questionários: General Health Questionnaire - GHQ [21] e o Global Physical Activity Questionnaire - GPAQ [2].

032 - Time to Get Conversational: Assessment of the Potential of Conversational User Interfaces for Mobile Banking

Vanessa Fahn, Andreas Riener

Setembro 2021

MuC '21: Mensch und Computer 2021

Tecnologia de Avaliação: User Experience Questionnaire (UEQ-S)

SQ1. Quais aspectos da UX a tecnologia avalia?	Atratividade, Perspicuidade, Eficiência, Confiabilidade, Estimulação, Novidade
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de diferencial semântico. É medida numa escala de 1 até 7: -3 (muito ruim) até 3 (muito bom).
SQ5. Qual a composição da tecnologia de avaliação?	Questionário User Experience Questionnaire (UEQ-S). Foi utilizada uma escala de 1 até 7 e foram avaliados os itens <ul style="list-style-type: none">• claro / confuso• ineficiente / eficiente• complicado / fácil• obstrutivo / solidario• chato / animador• não é interessante / interessante• convencional / inventivo• usual / de ponta
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos: se a análise da avaliação foi realizada quantitativamente.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	É um chatbot para realizar transações e tratativas bancárias.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é tanto orientado à tarefas quanto à conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi realizado um estudo de usuários com 18 participantes, empregados ou estudantes, 12 mulheres e 6 homens; Idades entre 19 e 58 anos, com média de 27.7 anos. Foi explicado para os participantes o tema do estudo, a

	<p>declaração de proteção de dados foi entregue aos participantes para assinarem e em seguida, eles preencheram o questionário preliminar em um tablet fornecido. Depois, foi explicado a eles a política PSD2 e como ela é realizada nos três protótipos. Cada tarefa durou cerca de 30 segundos a um minuto. Após cada tentativa, o questionário pós-tarefa composto por TAM, escala de confiança, segurança percebida e UEQ-S foi preenchido pelos voluntários em um tablet. Depois de executar todas as três tarefas, os sujeitos foram solicitados a classificar os três CUIs. Para mais insights qualitativos, foi realizada uma entrevista semiestruturada final sobre as interações bancárias conversacionais. O procedimento durou entre 35 e 45 minutos.</p>
<p>Avaliação Empírica da Tecnologia de Avaliação</p>	
<p>SQ12. A avaliação da tecnologia foi avaliada empiricamente?</p>	<p>Não</p>
<p>SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?</p>	<p>N.A.</p>
<p>Avaliação do Estado de saúde emocional do usuário</p>	
<p>SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?</p>	<p>Não</p>
<p>SQ14. De que forma a saúde mental do usuário foi avaliada?</p>	<p>N.A.</p>

033 - Who Should Be My Teammates: Using A Conversational Agent to Understand Individuals and Help Teaming

Ziang Xiao, Michelle X. Zhou, Wat-Tat Fu

Março 2019

IUI '19: Proceedings of the 24th International Conference on Intelligent User Interfaces

Tecnologia de Avaliação: Entrevista estruturada

SQ1. Quais aspectos da UX a tecnologia avalia?	Simpático, útil, agradável
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de escala de 1 (Nada) até 5 (Muito)
SQ5. Qual a composição da tecnologia de avaliação?	Não descrito.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	É um chatbot utilizado para estudar e potencializar os esforços de uma equipe de alunos em um ambiente educacional, chamado INDIGO (Diferenças Individuais para Otimização de Grupo)
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Sim, alunos.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi avaliado através de experimento controlado, por 184 estudantes. Foi realizada uma entrevista pré-curso que incluiu cinco seções. A primeira seção foi uma conversa de aquecimento, durante a qual INDIGO e um aluno se apresentaram e conversaram sobre seus hobbies e filme favorito. A segunda seção incluiu um questionário de 20 itens da Escala de Gerenciamento de Impressões (IM) para medir como os alunos se favorecem conscientemente para impressionar os outros. Na terceira seção, o INDIGO fez ao

	<p>aluno um conjunto de perguntas abertas sobre sua experiência de equipe e preferências de equipe. Na quarta parte da entrevista, o INDIGO discutiu com o aluno sobre as suas próprias características, como os seus pontos fortes e fracos. A última seção foi solicitar o feedback do aluno sobre o INDIGO. O aluno foi solicitado a expressar sua impressão do INDIGO e avaliar o INDIGO em várias escalas, como sua utilidade e simpatia. No final do semestre, o INDIGO entrevistou novamente cada aluno para obter os seus pontos de vista e opiniões sobre a experiência geral da equipa.</p>
<p>Avaliação Empírica da Tecnologia de Avaliação</p>	
<p>SQ12. A avaliação da tecnologia foi avaliada empiricamente?</p>	<p>Não</p>
<p>SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?</p>	<p>N.A.</p>
<p>Avaliação do Estado de saúde emocional do usuário</p>	
<p>SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?</p>	<p>Não</p>
<p>SQ14. De que forma a saúde mental do usuário foi avaliada?</p>	<p>N.A.</p>

036 - Exploring Language Style in Chatbots to Increase Perceived Product Value and User Engagement

Ela Elsholz, Jon Chamberlain, Udo Kruschwitz

Março 2019

CHIIR '19: Proceedings of the 2019 Conference on Human Information Interaction and Retrieval

Tecnologia de Avaliação: Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	UX geral
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Escala Likert de 7 pontos e questão aberta.
SQ5. Qual a composição da tecnologia de avaliação?	Questionário Pós-Tarefa Você gostaria de assistir a uma apresentação de uma peça de Shakespeare? Você estaria interessado em ver uma peça no teatro apresentado na seção anterior? Quanto você estaria disposto a pagar por uma produção profissional de Shakespeare no teatro apresentada na seção anterior? Como você aproveitou sua experiência de chatbot na seção anterior? Quão fácil foi para você completar a reserva? Comentários sobre o chatbot da seção anterior
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Mistos
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Chatbot de venda de ingressos.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação e tarefa.
SQ10. De qual forma o chatbot foi avaliado?	Foi avaliado através de experimento estruturado, com 169 participantes, 111 mulheres, 58 homens. Responderam dois questionários, um antes da tarefa e um

	depois. A tarefa constava de utilizar um dos chatbots disponíveis e utilizar o serviço de reservar um ingresso para uma peça teatral.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

037 - Comparing Data from Chatbot and Web Surveys

Soomin Kim, Joonhwan Lee, Gahgene Gweon

Maio 2019

CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems

Tecnologia de Avaliação: Questionário USE

SQ1. Quais aspectos da UX a tecnologia avalia?	Prazer
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Os itens relativos ao prazer em utilizar o chatbot foram medidos em escala Likert de 1 até 5.
SQ5. Qual a composição da tecnologia de avaliação?	Composto da questão: "É divertido usar o sistema de pesquisa. "
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	É um questionário que coleta informações demográficas e questões sobre o uso de internet.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Sim, adolescentes da Coreia.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi avaliado através de experimento controlado, com 106 participantes, porém 11 não finalizaram o experimento. Os participantes tiveram que responder a um questionário composto por informações demográficas e perguntas sobre o comportamento de uso da Internet que foi desenvolvido pela Agência Nacional da Sociedade da Informação da Coreia. Após o preenchimento do questionário principal, foram solicitadas perguntas de acompanhamento sobre a usabilidade e percepção do sistema de pesquisa.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi	Não

avaliada empiricamente?	
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

039 - Do You Feel Special When an AI Doctor Remembers You? Individuation Effects of AI vs. Human Doctors on User Experience

Jin Chen, Cheng Chen, Joseph B. Walther, S. Shyam Sundar

Maio 2021

CHI EA '21: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems

Tecnologia de Avaliação: Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?

Privacidade Intrusiva

SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?

Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software

SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?

Existente: tecnologia é baseada em uma existente.

SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes

É medida numa escala de 1 até 7.

SQ5. Qual a composição da tecnologia de avaliação?

Privacidade intrusiva foi avaliada por meio de 6 itens adaptados do estudo. Os quatro itens que constam na referência são:

- Eu sinto que, como resultado do meu uso desses sites, outras pessoas sabem sobre mim mais do que eu me sinto confortável.
- Acredito que, como resultado do uso desses sites, as informações sobre mim que considero privadas agora estão mais disponíveis para outras pessoas do que eu gostaria.
- Sinto que, como resultado do uso desses sites, as informações sobre mim estão disponíveis e, se usadas, invadirão minha privacidade.
- Sinto que, como resultado do uso desses sites, minha privacidade foi invadida por outras pessoas que coletam todos os dados sobre mim

Heng Xu, Tamara Dinev, H. Jef Smith, and Paul Hart. 2008. Examining the formation of individual's information privacy concerns: Toward an integrative view. Proceedings of 29th Annual International Conference on Information Systems (ICIS 2008), Paris, France.

SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Três chatbots foram programados para simular a primeira visita a um médico online, enquanto dois foram programados para a segunda visita.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	<p>Foi realizado um estudo de duas fases, com um total de 295 participantes da primeira fase do nosso estudo, e 223 deles participaram da segunda fase, com idades entre 19 e 76 anos. O primeiro estudo foi enquadrado como uma consulta médica pela primeira vez em uma plataforma de e-saúde, uma interface web do chat. Os participantes foram aleatoriamente designados para interagir com um dos três chatbots. Em seguida, direcionados para uma página da Web separada, onde preencheram um formulário de saúde (“Por favor, preencha este formulário para que nossos médicos on-line possam fornecer melhores conselhos de saúde para você!”).</p> <p>Após a conclusão, eles retornaram à interface de bate-papo e continuaram a interação. Os chatbots foram programados para fazer oito perguntas sobre sintomas e comportamentos do Covid-19. No final do bate-papo, os participantes receberam um diagnóstico e recomendações do Covid-19, modelados após o CDC Coronavirus Self-Checker (https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/coronavirus-self-checker.html). Cerca de dez dias após a primeira interação, os participantes foram notificados e foram designados para interagir com o médico da mesma identidade com a qual interagiram anteriormente. E eles foram aleatoriamente designados para a condição de individualização ou não-individualização usando dois chatbots. Após o segundo bate-papo,</p>

	os participantes foram direcionados a um questionário Qualtrics onde avaliaram o médico e sua interação.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

048 - "Forgot Your Password Again?" - Acceptance and User Experience of a Chatbot for In-Company IT Support

Dario Fiore, Matthias Baldauf, Christian Thiel

Novembro 2019

MUM '19: Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia

Tecnologia de Avaliação: Questionário UEQ

SQ1. Quais aspectos da UX a tecnologia avalia?	Atratividade, Perspicácia, Eficiência, Dependabilidade, Estimulação e Novidade.
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de diferencial semântico. É medida numa escala de 1 até 7: -3 (muito ruim) até 3 (muito bom).
SQ5. Qual a composição da tecnologia de avaliação?	Foi realizado o questionário UEQ, com escala de 1 até 7 irritante / agradável não compreensível/compreensível criativo / maçante fácil de aprender/difícil de aprender valioso / inferior chato / emocionante não é interessante/interessante imprevisível/previsível rápido lento inventivo/convencional obstrutivo / de suporte bom mau complicado/fácil desagradável / agradável usual/borda de ataque desagradável / agradável seguro/não seguro motivando / desmotivando atende às expectativas/não atende às expectativas ineficiente / eficiente claro/confuso impraticável / prático organizado/desordenado

	atraente / não atraente amigável não amigável conservador/inovador
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	É um chatbot para Suporte de TI de uma empresa
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é tanto orientado à tarefas quanto orientado à conversação.
SQ10. De qual forma o chatbot foi avaliado?	<p>Foi realizado experimento controlado, com 12 participantes, com idades entre 21 e 58 anos. Foi respondido um questionário inicial para colher informações sobre o conhecimento atual dos participantes sobre chatbots. Assinaram termo de consento.</p> <p>Em seguida utilizaram um chatbot de e-commerce por 5 minutos. Após isso, os participantes tiveram que resolver tarefas (redefinir senha, responder perguntas frequentes, liberar anexos de e-mail da quarentena), por 10 minutos. Em seguida responderam um questionário sobre cada caso de uso, também, responderam um questionário para indicar casos de uso adicionais que considerassem adequados para esse chatbot de solução de problemas de TI. Além disso, eles foram solicitados a avaliar sua disposição geral (independentemente de um caso de uso concreto) de usar um respectivo chatbot para suporte de TI e classificar diferentes canais de comunicação com o suporte de TI da empresa em relação à sua preferência pessoal. Por fim, foi pedido aos participantes que avaliassem sua experiência geral do usuário no User Experience Questionnaire (UEQ) de Laugwitz et al.</p>
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não.
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.

Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

065 - MusicBot: Evaluating Critiquing-Based Music Recommenders with Conversational Interaction

Yucheng Jin, Wanling Cai, Li Chen, Nyi Nyi Htun, Katrien Verbert

Novembro 2019

CIKM '19: Proceedings of the 28th ACM International Conference on Information and Knowledge Management

Tecnologia de Avaliação Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	Interesse; Diversidade; Fácil de informar; Controle; Confiança; Feedback; Entender; Dificuldade; Expectativa; Intenção de reutilizar; Satisfação;
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Específica: tecnologia de avaliação de UX específica para chatbots.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia foi criada para o estudo e descrita no artigo.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de questionário com Escala Likert de 7 pontos, variando de discordo fortemente até concordo fortemente.
SQ5. Qual a composição da tecnologia de avaliação?	Itens de perguntas relacionadas a Ux (Q1, Q4 -Q14) Q1: Os itens recomendados para mim corresponderam aos meus interesses. Q4: As músicas recomendadas para mim são diversas. Q5: Achei fácil informar ao sistema se não gosto/gosto do música recomendada. Q6: Eu me senti no controle de modificar meu gosto usando o MusicBot. Q7: Estou confiante de que vou gostar das músicas recomendadas para mim. Q8: Gosto de dar feedback sobre a música que estou ouvindo. Q9: Este chatbot de música pode ser confiável. Q10: Achei o sistema fácil de entender

	<p>nesta conversa.</p> <p>Q11: Nesta conversa, eu sabia o que poderia dizer ou fazer a cada ponto do diálogo.</p> <p>Q12: O sistema funcionou da maneira que eu esperava nesta conversa.</p> <p>Q13: Vou usar este chatbot de música novamente.</p> <p>Q14: No geral, estou satisfeito com o chatbot. "</p>
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	É um recomendador de música baseado em crítica híbrida MusicBot, que usa um chatbot para permitir que os usuários interajam com as recomendações por meio de texto e voz. O sistema oferece duas técnicas principais de crítica, a crítica iniciada pelo usuário (UC) e a crítica sugerida pelo sistema (SC) para refinar a recomendação.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é tanto orientado à tarefas quanto orientado à conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi realizado um experimento controlado, com 45 participantes. A tarefa experimental foi encontrar cinco músicas que melhor correspondessem ao cenário apresentado e à preferência musical do usuário. Cada participante precisava realizar essa tarefa em dois cenários diferentes: pegar o metrô depois da aula/trabalho e organizar a festa de aniversário de um amigo. O procedimento contém as seguintes etapas. O procedimento teve cinco etapas: Leitura de um tutorial, montar um perfil de preferências musicais, responder a um questionário pré- estudo sobre dados demográficos e características pessoais, experimentar o sistema, utilizar o sistema de maneira a cumprir a tarefa dada pelo mesmo, no caso era necessário para reagir às músicas recomendadas uma a uma aceitando/pulando a música atual ou fazendo uma crítica, e por fim, responder

	<p>mais um questionário, baseado em uma estrutura de avaliação centrada no usuário para sistemas de recomendação que incluiu 14 perguntas para medir a percepção do usuário de sistemas de recomendação. Os usuários foram capazes de fornecer comentários de texto livre no final.</p>
<p>Avaliação Empírica da Tecnologia de Avaliação</p>	
<p>SQ12. A avaliação da tecnologia foi avaliada empiricamente?</p>	<p>Não</p>
<p>SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?</p>	<p>N.A.</p>
<p>Avaliação do Estado de saúde emocional do usuário</p>	
<p>SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?</p>	<p>Não</p>
<p>SQ14. De que forma a saúde mental do usuário foi avaliada?</p>	<p>N.A.</p>

085 - Exploring the Effects of Incorporating Human Experts to Deliver Journaling Guidance through a Chatbot

Yi-Chieh Lee, Naomi Yamashita, Yun Huang

Abril 2021

Proceedings of the ACM on Human-Computer Interaction

Tecnologia de Avaliação: Tecnologia 1: Questionário

Tecnologia 2: Entrevista

SQ1. Quais aspectos da UX a tecnologia avalia?	Tecnologia 1: Confiança percebida, engajamento percebido e Intimidade percebida, Autorreflexão e Autoconsciência. Tecnologia 2: Engajamento e Impressões.
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Ambas Genéricas: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Ambas Existentes: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Tecnologia 1: Foi utilizada uma escala likert de 1 até 7, sendo 1=discordo totalmente a 7=concordo totalmente Tecnologia 2: questões abertas.
SQ5. Qual a composição da tecnologia de avaliação?	Confiança percebida: nove itens adaptados da literatura. Os itens presentes na referência são: Confiança – Competência 1. Este consultor virtual é como um verdadeiro especialista na avaliação de câmeras digitais. 2. Este consultor virtual tem experiência para entender minhas necessidades e preferências sobre câmeras digitais. 3. Este consultor virtual tem a capacidade de entender minhas necessidades e preferências sobre câmeras digitais. 4. Este orientador virtual possui bons conhecimentos sobre câmeras digitais. 5. Este consultor virtual considera minhas necessidades e todos os atributos importantes das câmeras digitais ferrugem – Benevolência 1. Este consultor virtual coloca meus

	<p>interesses em primeiro lugar.</p> <ol style="list-style-type: none"> 2. Este orientador virtual mantém meus interesses em mente. 3. Este consultor virtual quer entender minhas necessidades e preferências. <p>Confiança – Integridade</p> <ol style="list-style-type: none"> 1. Este consultor virtual fornece recomendações imparciais de produtos. 2. Este consultor virtual é honesto. 3. Considero este consultor virtual íntegro. <p>(Izak Benbasat and Weiquan Wang. 2005. Trust in and adoption of online recommendation agents. Journal of the association for information systems 6, 3 (2005), 4.)</p> <p>(SeoYoung Lee and Junho Choi. 2017. Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. International Journal of Human-Computer Studies 103 (2017), 95–105.,)</p> <ol style="list-style-type: none"> 1. Tenho fé no que o CA está me dizendo. 2. O CA fornece comigo um filme imparcial e preciso recomendações. 3. O CA é honesto. 4. A CA é confiável 5. O CA quer conhecer e entender minhas necessidades e preferências. 6. O CA quer lembrar meus interesses. 7. Acredito que a CA presta um serviço confiável. 8. Posso confiar na CA com meus informação pessoal. 9. Posso confiar nas informações fornecidas pela CA. <p>Os itens 5 e 6 foram excluídos da análise fatorial.</p> <p>(Yi-Chieh Lee, Naomi Yamashita, and Yun Huang. 2020. Designing a Chatbot as a Mediator for Promoting Deep Self-Disclosure to a Real Mental Health Professional. Proceedings of the ACM on Human-Computer Interaction (2020).)</p>
--	--

	<p>Engajamento: Os 12 itens de medição para a dimensão de engajamento da pesquisa foram adaptados da literatura</p> <p>FA-S.1 Eu me perdi nessa experiência. FA-S.2 O tempo que passei usando o Aplicativo X passou rápido. FA-S.3 Fiquei absorto nessa experiência.</p> <p>PU-S.1 Senti-me frustrado ao usar este Aplicativo X. PU-S.2 Achei este Aplicativo X confuso de usar. PU-S.3 Usar este Aplicativo X era cansativo.</p> <p>AE-S.1 Esta Aplicação X era atrativa. AE-S.2 Esta Aplicação X era esteticamente atraente. AE-S.3 Esta Aplicação X atraiu meus sentidos.</p> <p>RW-S.1 Valeu a pena usar o Aplicativo X. RW-S.2 Minha experiência foi gratificante. RW-S.3 Me interessei por essa experiência.</p> <p>(Heather L O'Brien, Paul Cairns, and Mark Hall. 2018. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. International Journal of Human-Computer Studies 112 (2018), 28–39.) relatados em frases.</p> <p>Intimidade: seis itens adaptados da referencia</p> <ol style="list-style-type: none"> 1. Sinto-me próximo do CA. 2. Sinto que o CA é meu amigo íntimo. 3. Sinto-me emocionalmente próximo do CA. 4. Acho que a CA afetará minha seleção de mídia conteúdo. 5. A CA usa declarações de apoio para conquistar mim. 6. Desenvolvi um senso de familiaridade com o CA. <p>(SeoYoung Lee and Junho Choi. 2017. Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity.</p>
--	--

	<p>International Journal of Human-Computer Studies 103 (2017), 95–105)</p> <p>Autorreflexão e Autoconsciência: doze itens adaptados das referencias:</p> <ol style="list-style-type: none"> 1. Neste momento, estou bem ciente de tudo em meu ambiente. 2. Neste momento, estou consciente dos meus sentimentos interiores. 3. Neste momento, estou preocupado com a forma como me apresento. 4. Neste momento, tenho consciência de minha aparência. 5. Neste momento, estou consciente do que está acontecendo ao meu redor. 6. Neste momento, estou refletindo sobre minha vida. 7. Neste momento, estou preocupado com o que as outras pessoas pensam de mim. 8. Neste momento, estou ciente dos meus pensamentos mais íntimos. 9. Neste momento, estou consciente de todos os objetos ao meu redor. <p>(John M Govern and Lisa A Marsch. 2001. Development and validation of the situational self-awareness scale. Consciousness and cognition 10, 3 (2001), 366–378.)</p> <p>Eu não costumo pensar em meus pensamentos</p> <p>Raramente passo tempo em auto-reflexão</p> <p>Eu frequentemente examino meus sentimentos</p> <p>Eu realmente não penso sobre por que me comporto da maneira que me comporto</p> <p>Eu frequentemente tomo tempo para refletir sobre meus pensamentos</p> <p>Muitas vezes penso sobre como me sinto sobre as coisas</p> <p>Não estou realmente interessado em analisar meu comportamento</p> <p>É importante para mim avaliar as coisas que faço</p> <p>Estou muito interessado em examinar o que penso sobre</p> <p>É importante para mim tentar entender o que meus sentimentos significam</p> <p>Eu tenho uma necessidade definida de</p>
--	--

	<p>entender a maneira como minha mente funciona</p> <p>É importante para mim ser capaz de entender como meus pensamentos surgem (Anthony M Grant, John Franklin, and Peter Langford. 2002. The self-reflection and insight scale: A new measure of private self-consciousness. Social Behavior and Personality: an international journal 30, 8 (2002), 821–835.)</p> <p>Entrevista: Para capturar diferenças na forma como os entrevistados responderam às orientações específicas do chatbot, pedimos que descrevessem seus sentimentos sobre essas sugestões, inclusive se achavam que valia a pena seguir; se se sentiram à vontade para recebê-los; o que aprenderam após segui-los/ignorá-los; e se tais sentimentos mudaram ao longo do tempo. Além disso, perguntamos aos participantes quanto esforço eles despenderam para aprender com o sistema durante a Fase 2; e o quanto eles revelaram ao chatbot/coach ao praticar as habilidades de journaling que aprenderam. Além de suas impressões gerais sobre o chatbot/coach, perguntamos como eles definiam sua relação com ele; se confiaram nela, e por quê; e se perceberam que se tornaram dependentes dele/dela ao aprender e praticar as habilidades de diário. Por fim, pedimos aos entrevistados que refletissem sobre se sua participação no estudo como um todo afetou sua autorreflexão ou autoconsciência na vida cotidiana e se ainda estavam usando alguma das habilidades de registro em diário que aprenderam durante o experimento.</p>
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Tecnologia 1: quantitativos. Tecnologia 2: qualitativos
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	O objetivo do sistema de chatbot era orientar os participantes a praticar habilidades de escrita jornalística.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.

SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi realizado um experimento com 35 participantes, com idades entre 20 e 29 anos. Todos os participantes estavam familiarizados com as plataformas de mensagens. O estudo foi realizado em quatro semanas, com um dia para os participantes se familiarizarem com chatbot, e os 27 dias restantes foram divididos em três fases de tempo iguais. Os usuários tiveram que utilizar o chatbot por 4 semanas. Após o período de quatro semanas de interação com o chatbot, 34 participantes participaram de entrevistas individuais presenciais, cada uma com duração entre 45 e 60 minutos. Foi conduzido um survey no final da fase 2.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Sim
SQ14. De que forma a saúde mental do usuário foi avaliada?	Foi avaliada através da Autorreflexão e Autoconsciência.

092 - Examining User Preference for Agreeableness in Chatbots

Sarah Theres Völkel, Lale Kaya

Julho 2021

CUI '21: CUI 2021 - 3rd Conference on Conversational User Interfaces

Tecnologia de Avaliação: Big Five Inventory-2 (BFI-2)

SQ1. Quais aspectos da UX a tecnologia avalia?	Agradabilidade
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de escala Likert (escala de 1=desagradável a 5=agradável)
SQ5. Qual a composição da tecnologia de avaliação?	Amabilidade Itens de compaixão É compassivo, tem um coração mole. Pode ser frio e indiferente É útil e altruísta com os outros Sente pouca simpatia pelos outros Itens de respeito É respeitoso, trata os outros com respeito É educado, cortês com os outros Às vezes é rude com os outros Inicia discussões com outras pessoas Itens de confiança Assume o melhor das pessoas Tem uma natureza perdoadora Tende a encontrar defeitos nos outros Suspeita das intenções dos outros (Daniel Danner, Beatrice Rammstedt, Matthias Bluemke, Lisa Treiber, Sabrina Berres, Christopher Soto, and Oliver John. 2016. Die deutsche Version des Big Five Inventory 2 (BFI-2). GESIS, Mannheim, Germany. https://doi.org/10.6102/zis247)
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Foi implementado um sistema de recomendação de filmes, baseado nas preferências e contexto dos usuários. A base de dados consistiu em 30 filmes, ou

	seja, os três filmes mais bem avaliados para dez gêneros comuns, conforme informado por um site de recomendação de filmes alemão.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi realizado um experimento controlado com 30 participantes, com idades entre 18 e 58 anos. Os participantes interagiram com três versões de um mesmo chatbot. (Uma versão mais agradável, uma neutra e uma desagradável). Em todas as três interações do chatbot, os participantes foram solicitados a perguntar ao chatbot sobre uma recomendação de filme. Em cada execução, foi variada ligeiramente a descrição da tarefa para evitar a repetição monótona da tarefa. Essa variação da descrição da tarefa referia-se à empresa, com a qual o usuário está assistindo ao filme. Tanto a ordem dos chatbots quanto as descrições das tarefas foram contrabalançadas usando um quadrado latino. Após cada interação, pedimos aos participantes que descrevessem sua impressão sobre o chatbot, primeiro, os participantes especificaram sua percepção da agradabilidade do chatbot, preenchendo a agradabilidade e em segundo lugar, os participantes indicaram o quanto gostariam de interagir novamente com este chatbot
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não.
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

102 - Evaluating and Informing the Design of Chatbots

Mohit Jain, Pratyush Kumar, Ramachandra Kota, Shwetak N. Patel

Junho 2018

DIS '18: Proceedings of the 2018 Designing Interactive Systems Conference

Tecnologia de Avaliação Questionário Pós Estudo

SQ1. Quais aspectos da UX a tecnologia avalia?	Diversão e frustração
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia foi criada para o estudo e descrita no artigo.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Escala likert de 5 pontos.
SQ5. Qual a composição da tecnologia de avaliação?	Não consta.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos

Avaliação dos Chatbots

SQ7. Qual é a aplicação do Chatbot?	Foram avaliados oito chatbots: Alterra - viagens, Call of Duty - Entretenimento, chatShopper - Shopping, CNN- Notícias, Hi Poncho - Utilidades, Pandorabots - Bate papo, Swelly - Social, Trivia blast - jogo
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é tanto orientado à tarefas quanto orientado à conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi realizado um estudo com 16 participantes, com idades entre 23 e 45 anos, Os participantes receberam uma lista de 8 chatbots (Alterra, Call of Duty, chatShopper, CNN, Hi Poncho, Pandorabots, Swelly, Trivia Blast) em uma ordem aleatória. Eles foram solicitados a interagir com cada chatbot por pelo menos 3-5 minutos diariamente por três dias. Os participantes não foram instruídos sobre como interagir com os chatbots, sobre o que eram os chatbots ou que tipo de tarefas realizar usando os chatbot. Após os três dias de interação com os chatbots, o

	participante teve uma entrevista semiestruturada presencial com o facilitador. No início da entrevista, os participantes foram solicitados a classificar os chatbots e classificá-los em relação a diferentes métricas, incluindo curva de aprendizado, nível de frustração e diversão de usar.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não.
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

105 - Make your own The Potential of Chatbot Customization for the Development of User Trust

Rebecca Wald, Evelien Heyselaar, Tibor Bosse

Junho 2021

UMAP '21: Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization

Tecnologia de Avaliação: Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	Confiança
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de escala likert de 1(nada) até 7(completamente)
SQ5. Qual a composição da tecnologia de avaliação?	O quanto você confia no sistema. Bickmore, T., Vardoulakis, L., Jack, B., Paasche-Orlow, M. (2013). Automated Promotion of Technology Acceptance by Clinicians Using Relational Agents. In Aylett R., Krenn B., Pelachaud C., Shimodaira H. (eds) Intelligent Virtual Agents. IVA 2013, 68-78.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Chatbot Mitsuku que imita uma conversação humana num contexto não específico, baseada em informações recuperadas de entradas livres dos usuários.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Após dar consentimento para participar deste estudo, as informações demográficas e pessoais (por exemplo, interesses pessoais) dos participantes foram coletadas na primeira parte do questionário online. Dependendo da condição atribuída, o

	<p>questionário pedia ainda ao participante que personalizasse seu chatbot com opções de múltipla escolha ou o apresentava ao chatbot genérico. Após ser direcionado para a sala de chat por meio de uma URL externa, cada participante teve uma conversa de dez minutos com o chatbot sobre interesses pessoais, que o pré-teste identificou como uma duração razoável e tópico de conversa. Após dez minutos, o chatbot agradeceu ao participante pela conversa e deu instruções sobre como retornar ao questionário online, que avaliou a experiência do chatbot do participante por meio de autorrelatos compostos por respostas nas escalas de medição.</p>
<p>Avaliação Empírica da Tecnologia de Avaliação</p>	
<p>SQ12. A avaliação da tecnologia foi avaliada empiricamente?</p>	<p>Não</p>
<p>SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?</p>	<p>N.A.</p>
<p>Avaliação do Estado de saúde emocional do usuário</p>	
<p>SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?</p>	<p>Não</p>
<p>SQ14. De que forma a saúde mental do usuário foi avaliada?</p>	<p>N.A.</p>

<p>111 - “I wrote as if I were telling a story to someone I knew.”: Designing Chatbot Interactions for Expressive Writing in Mental Health</p> <p>SoHyun Park, Anja Thieme, Jeongyun Han, Sungwoo Lee, Wonjong Rhee, Bongwon Suh</p> <p>Junho 2021</p>	
<p><i>DIS '21: Designing Interactive Systems Conference 2021</i></p>	
<p>Tecnologia de Avaliação Tecnologia 1: Questionário Tecnologia 2: Entrevista semi estruturada</p>	
<p>SQ1. Quais aspectos da UX a tecnologia avalia?</p>	<p>Tecnologia 1: Bem estar Psicológico, Engajamento, Ansiedade, Atenção, Intenção de Uso, Adaptabilidade Percebida, Prazer, Sociabilidade, Utilidade, Influência social, Presença Social, confiança e interpretação.</p> <p>Tecnologia 2: Experiencia geral, impacto psicológico, percepções de divulgação social.</p>
<p>SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?</p>	<p>Ambas Genéricas: tecnologia de avaliação de UX não é restrita a tipos específicos de software</p>
<p>SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?</p>	<p>Ambas Existentes: tecnologia é baseada em uma existente.</p>
<p>SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes</p>	<p>Tecnologia 1: Likert de 1 até 5. e Likert de 1 até 7.</p> <p>Tecnologia 2: Questões abertas</p>
<p>SQ5. Qual a composição da tecnologia de avaliação?</p>	<p>Tecnologia 1</p> <p>Bem estar Psicologico</p> <ol style="list-style-type: none"> 1. Dada a minha condição física atual, estou satisfeito com o que posso fazer. 2. Tenho confiança na minha capacidade de manter relacionamentos importantes. 3. Sinto-me esperançoso quanto ao meu futuro. 4. Muitas vezes estou interessado e animado com as coisas da minha vida 5. Sou capaz de me divertir. 6. Geralmente estou satisfeito com minha saúde psicológica. 7. Sou capaz de me perdoar pelos meus fracassos. 8. Minha vida está progredindo de acordo com minhas expectativas. 9. Sou capaz de lidar com conflitos com os

	<p>outros.</p> <p>10. Tenho paz de espírito.</p> <ol style="list-style-type: none"> 1. Tenho me sentido terrivelmente sozinho e isolado 2. Tenho me sentido tenso, ansioso ou nervoso 3. Sinto que tenho alguém a quem recorrer para apoio quando necessário 4. Eu me senti O.K. sobre mim 5. Tenho me sentido totalmente desprovido de energia e entusiasmo 6. Fui fisicamente violento com os outros 7. Eu me senti capaz de lidar com as coisas erradas 8. Tenho sido incomodado por dores, dores ou outros problemas físicos 9. Já pensei em me machucar 10. Falar com as pessoas foi demais para mim 11. A tensão e a ansiedade me impediram de fazer coisas importantes 12. Fiquei feliz com as coisas que fiz 13. Fui perturbado por pensamentos e sentimentos indesejados 14. Senti vontade de chorar 15. Senti pânico ou terror 16. Eu fiz planos para acabar com minha vida <p>Engajamento</p> <p>Quão pessoais, em geral, os escritos eram? Quanto a escrita expressou seus sentimentos?</p> <p>Quão difícil foi escrever nos últimos quatro dias?</p> <p>Quão significativo foi o estudo?</p> <p>Ansiedade</p> <ol style="list-style-type: none"> 1. Se eu usasse o robô, teria medo de cometer erros com isso 2. Se eu usasse o robô, ficaria com medo de quebrar alguma coisa 3. Acho o robô assustador 4. Acho o robô intimidador <p>Atitude</p> <ol style="list-style-type: none"> 5. Acho uma boa ideia usar o robô 6. o robô tornaria minha vida mais
--	---

	<p>interessante</p> <p>7. É bom fazer uso do robô</p> <p>FC 8. Tenho tudo o que preciso para fazer bom uso do robô.</p> <p>9. Conheço o robô o suficiente para fazer bom uso dele.</p> <p>Intenção de uso</p> <p>10. Acho que vou usar o robô nos próximos dias</p> <p>11. Estou certo de que usarei o robô durante os próximos dias</p> <p>12. Estou planejando usar o robô durante os próximos dias</p> <p>Adaptabilidade percebida,</p> <p>13. Acho que o robô pode se adaptar ao que preciso</p> <p>14. Acho que o robô só fará o que eu preciso naquele momento</p> <p>15. Eu acho que o robô vai me ajudar quando eu achar necessário</p> <p>Prazer percebido</p> <p>16. Eu gosto do robô falando comigo</p> <p>17. Gosto de fazer coisas com o robô</p> <p>18. Acho o robô agradável</p> <p>19. Acho o robô fascinante</p> <p>20. Acho o robô chato</p> <p>Percepção de facilidade de uso</p> <p>21. Acho que saberei rapidamente como usar o robô</p> <p>22. Acho o robô fácil de usar</p> <p>23. Acho que posso usar o robô sem ajuda</p> <p>24. Acho que posso usar o robô quando houver alguém por perto para me ajude</p> <p>25. Acho que posso usar o robô quando tenho um bom manual.</p> <p>Sociabilidade Percebida</p> <p>26. Considero o robô um parceiro de conversação agradável</p> <p>27. Acho agradável interagir com o robô</p> <p>28. Sinto que o robô me entende.</p> <p>29. Eu acho o robô legal</p> <p>Utilidade percebida</p> <p>30. Acho que o robô é útil para mim</p> <p>31. Seria conveniente para mim ter o robô</p> <p>32. Acho que o robô pode me ajudar em</p>
--	--

	<p> muitas coisas</p> <p> IS 33. Acho que a equipe gostaria que eu usasse o robô</p> <p> 34. Acho que daria uma boa impressão se eu usasse o robô.</p> <p> Presença Social</p> <p> 35. Ao interagir com o robô, senti como se estivesse falando com um pessoa</p> <p> 36. Às vezes parecia que o robô estava realmente olhando para mim</p> <p> 37. Posso imaginar o robô como uma criatura viva</p> <p> 38. Costumo pensar que o robô não é uma pessoa real.</p> <p> 39. Às vezes o robô parece ter sentimentos reais</p> <p> Confiar</p> <p> 40. Eu confiaria no robô se ele me desse conselhos.</p> <p> 41. Eu seguiria o conselho que o robô me dá</p> <p> Tecnologia 2: Não descrito.</p>
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	<p> Tecnologia 1: Quantitativos.</p> <p> Tecnologia 2: Qualitativos.</p>
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	<p> Foram implementados dois chatbots de conversação que instruem o usuário a escrever sobre algumas de suas experiências mais difíceis na vida. O Básico - que entregar as instruções de escrita expressiva de Pennebaker et al. [68] em uma interface de chatbot, traduzida para uma linguagem mais casual e conversacional; O Responsivo - no entanto, ele acompanha a escrita do usuário com 5 prompts de perguntas abertas situados. As perguntas foram elaboradas para ajudar o usuário a revisar o que escreveu em resposta às instruções iniciais; baseiam-se na combinação de três técnicas terapêuticas que podem ajudar o usuário a refletir sobre sentimentos passados, relações sociais ou circunstâncias situacionais e sobre si mesmo.</p>
SQ8. O chatbot foi criado para um grupo	<p> Não, o chatbot não foi criado para um grupo</p>

específico de pessoas? Qual grupo?	específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Recrutaram 30 participantes (14 homens, 16 mulheres; Mediana = 28 anos, min = 23, max = 41). Todos eles eram coreanos nativos, principalmente estudantes de graduação e pós-graduação de diferentes disciplinas. O estudo durou 4 dias para manter a configuração original de escrita expressiva. Para obter feedback sobre suas experiências, os participantes deveriam preencher pesquisas todos os dias antes e depois da redação, e antes e depois do estudo. Também realizamos uma entrevista post-hoc para investigar mais profundamente suas experiências. A totalidade das atividades durou cerca de 120 minutos por participante. Todas as sessões de estudo ocorreram em um ambiente de laboratório para manter a configuração experimental consistente para todos os participantes e para responder a qualquer emergência potencial imediatamente. Além disso, para facilitar a escrita e a proteção dos dados, foi utilizado um computador portátil.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Sim
SQ14. De que forma a saúde mental do usuário foi avaliada?	Para avaliar o bem-estar psicológico dos usuários pré e pós-estudo, usamos a Escala de Resultados de Schwartz de 7 pontos (SOS-10; por exemplo, "Sinto-me esperançoso em relação ao meu futuro.") e Resultado Clínico de 5 pontos na Rotina Avaliação (CORE-10; por exemplo, "Na última semana, me senti infeliz."). Ambos são comumente usados para medir o bem-estar em um período de tempo

	relativamente curto. Os escores que computamos para análise estatística para encontrar qualquer diferença pré e pós-estudo.
--	---

157 - Chatbot with Touch and Graphics: An Interaction of Users for Emotional Expression and Turn-taking

Hyeonggeun Yun Auejin Ham, Jongrae Park, Jin Kim Taeyeong Kim, Jinkyu Jang, Jeongeun Kim Haechan Lee

Julho 2020

CUI '20: Proceedings of the 2nd Conference on Conversational User Interfaces

Tecnologia de Avaliação: Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	Grau de expressão emocional reveladora, utilidade da expressão emocional, naturalidade.
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Específica: tecnologia de avaliação de UX específica para chatbots.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia foi criada para o estudo e descrita no artigo.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Escala Likert de 7 pontos para todos os aspectos
SQ5. Qual a composição da tecnologia de avaliação?	Não descrito.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Foi desenvolvido um sistema de bate-papo móvel baseado na web. No sistema, as pessoas podem conversar e interagir com um chatbot de várias maneiras inclusive de formas não verbais, para promover a interação entre o chatbot e o usuário.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foram recrutados 28 participantes interessados em conversar com o chatbot. O experimento foi conversar com um chatbot que tem três condições diferentes (interação baseada em texto, interação baseada em texto e toque e interação baseada em texto e TwG) em relação às preocupações dos participantes relacionadas a três dos quatro tópicos (aprendizagem, carreira, personalidade e relacionamento interpessoal. Antes do experimento, os participantes escolheram

	<p>três tópicos entre os quatro disponíveis. Depois disso, explicamos a interação de cada condição e os participantes podiam praticar a interação até que estivessem familiarizados com a interação. Após a prática, os participantes se comunicaram com o Oz (o chatbot) por 10 minutos por condição. O Oz escolheu as respostas entre o conjunto predefinido de backchannels e perguntas. Após terminar a conversa, eles responderam à pesquisa sobre a experiência do usuário na comunicação.</p>
<p>Avaliação Empírica da Tecnologia de Avaliação</p>	
<p>SQ12. A avaliação da tecnologia foi avaliada empiricamente?</p>	<p>Não</p>
<p>SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?</p>	<p>N.A.</p>
<p>Avaliação do Estado de saúde emocional do usuário</p>	
<p>SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?</p>	<p>Sim</p>
<p>SQ14. De que forma a saúde mental do usuário foi avaliada?</p>	<p>A saúde mental foi avaliada por meio dos aspectos da revelação da expressão emocional e da utilidade da expressão emocional</p>

185 - Chatbot-based Emotion Management for Distributed Teams: A Participatory Design Study

Ivo Benke, Michael Thomas Knierim, Alexander Maedche

Outubro 2020

Proceedings of the ACM on Human-Computer Interaction

Tecnologia de Avaliação Tecnologia 1 Entrevista
Tecnologia 2: ABCCT

SQ1. Quais aspectos da UX a tecnologia avalia?	Tecnologia1 Percepção de emoção Tecnologia 2: expressão da emoção, engajamento, presença social, suporte social, compromisso, expectativas não atendidas e privacidade.
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Tecnologia 1: Específica: tecnologia de avaliação de UX específica para chatbots. Tecnologia 2: Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Tecnologia 1: Criada: tecnologia foi criada para o estudo e descrita no artigo. Tecnologia 2: Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Tecnologia 1: Questões abertas Tecnologia 2: Escala likert 5 pontos
SQ5. Qual a composição da tecnologia de avaliação?	Tecnologia 1: Como foi a percepção do chatbot? e Como foi a sua experiência com a aparência do chatbot? Tecnologia 2: Comunicar-se com <X> usando <M> me ajuda a saber como <X> está se sentindo naquele dia. Comunicar-me com <X> usando <M> me ajuda a deixar <X> saber como estou me sentindo. Comunicar-me com <X> usando <M> me ajuda a ver o quanto <X> se importa comigo. Sinto que o contato comigo usando <M> é atraente para <X>. Estou animado para usar <M> com <X>. Eu me divirto com <X> enquanto uso <M> Comunicar-me com <X> usando <M> me ajuda a me sentir mais próximo de <X>. Depois que terminamos de nos comunicar, ainda fico pensando em algo que <X>

	<p>compartilhou usando <M>.</p> <p>Comunicar-me com <X> usando <M> me ajuda a me sentir mais conectado com <X>.</p> <p>A comunicação com <X> usando <M> me ajuda a fornecer suporte social a <X>.</p> <p><X> me faz sentir especial em nosso contato usando <M>.</p> <p>Comunicar-se comigo usando <M> ajuda <X> a estar lá para mim quando eu precisar.</p> <p>Comunicar-me com <X> usando <M> quando estou tendo um dia ruim me ajuda a me sentir melhor</p> <p>Comunicar-me com <X> usando <M> me ajuda a me sentir menos preocupado com alguma coisa.</p> <p>Eu me preocupo que <X> se sinta obrigado a me contatar usando <M>.</p> <p>Eu tenho que falar com <X> usando <M> mesmo que eu não queira.</p> <p>Eu me sinto culpado se não responder a um contato que <X> faz usando <M>.</p> <p>Eu tenho que responder quando <X> tenta entrar em contato comigo usando <M> mesmo que eu não queira.</p> <p>Fico triste quando <X> não está por perto quando tento entrar em contato com <X> usando <M></p> <p>Fico triste quando <X> demora muito para responder quando tento entrar em contato com <X> usando <M>.</p> <p>Eu me preocupo por não estar atendendo às expectativas de <X> para nosso contato usando <M>.</p> <p>Fico triste quando <X> não presta atenção suficiente em mim quando usamos <M>.</p> <p>Eu me preocupo que <X> possa aprender algo usando <M> que eu queira manter em segredo.</p> <p>Eu me preocupo com minha privacidade enquanto <X> e eu estamos usando <M> juntos.</p> <p>Eu me preocupo que outras pessoas possam ouvir ou ver algo que <X> e eu compartilhamos usando <M>.</p> <p>Temo estar violando a privacidade de <X> durante nosso contato usando <M>.</p>
--	---

SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Tec 1: Qualitativos Tec 2: Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	O chatbot foi criado para equipes com o propósito de auxiliar em desafios previamente identificados.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi conduzido um um experimento de laboratório com três condições (as três instâncias de projeto NBT, SBT e ABT) como variáveis independentes. Cada sessão durou duas horas. Para obter insights mais profundos sobre as experiências com os chatbots, foram coletados dados qualitativos e quantitativos. No foco da avaliação estão as experiências dos participantes com EM baseado em chatbot, a percepção do design e as implicações para o indivíduo e para a equipe. Foram recrutados 27 participantes para o estudo.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	Não
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Sim
SQ14. De que forma a saúde mental do usuário foi avaliada?	Foi avaliada através da emoção percebida e da expressão da emoção.

188 - Investigating Students' Use of a Mental Health Chatbot to Alleviate Academic Stress

Johan Oswin De Nieva, Jose Andres Joaquin, Chaste Bernard Tan, Ruzel Khyvin Marc Te, Ethel Ong

Outubro 2020

CHluXiD '20: 6th International ACM In-Cooperation HCI and UX Conference

Tecnologia de Avaliação Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	Humanidade e Afeto
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia foi criada para o estudo e descrita no artigo.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Escala likert de 4 pontos
SQ5. Qual a composição da tecnologia de avaliação?	Não descrita.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Woebot é um chatbot desenvolvido para ajudar os alunos do ensino médio a aliviar o estresse da carga de trabalho acadêmica.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Sim. Alunos.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	53 alunos participaram do estudo, porém somente 25 responderam a todos os questionários. Para a coleta de dados foi utilizada uma série de instrumentos que foram administrados por meio do Google Forms. Os participantes tiveram que responder a um avaliação de sofrimento psicológico (PDA), avaliação do nível de estresse pré-teste e pós-teste (SLA), avaliação diária de conversação (DCA) e formulário de avaliação do chatbot.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do	N.A.

experimento conduzido para avaliar a tecnologia de avaliação?	
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

209 - ArgueTutor: An Adaptive Dialog-Based Learning System for Argumentation Skills

Thiemo Wambsganss, Tobias Küng, Matthias Söllner, Jan Marco Leimeister

Maio 2021

CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems

Tecnologia de Avaliação Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	Prazer
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Específica: tecnologia de avaliação de UX específica para chatbots.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Escala Likert de 1 a 5
SQ5. Qual a composição da tecnologia de avaliação?	"Tecnologia1: questionário pós teste Prazer: "A interação com a ferramenta de aprendizagem foi emocionante" e "É divertido interagir com a ferramenta de aprendizagem"
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.

Avaliação dos Chatbots

SQ7. Qual é a aplicação do Chatbot?	O chatbot é uma ferramenta de aprendizado conversacional em combinação com mineração de argumentação. O ArgueTutor (abreviação de Argumentation Tutor), um sistema de tutoria baseado em diálogo adaptativo que fornece aos alunos feedback adaptativo e instantâneo, entrada teórica e orientação passo a passo durante seu processo de escrita.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Sim, Estudantes
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Recrutaram 55 estudantes para participar do experimento. Contabilizaram 31 resultados válidos no tratamento e 24 no grupo controle. Os participantes do grupo de tratamento tinham idade média de

	<p>22,75, 17 eram do sexo masculino, 14 do sexo feminino. No grupo controle, a idade média dos participantes foi de 23,52 (DP= 2,81), 11 eram do sexo masculino e 13 do feminino. O experimento foi conduzido em três etapas: 1) Pré-teste: O experimento começou com uma pré-pesquisa de 14 questões. Onde testaram três construções diferentes para avaliar se a randomização foi bem-sucedida.</p> <p>2) Exercício de escrita: Na parte de escrita dos experimentos, pediram aos participantes que realizassem tarefas de escrita persuasiva, simulando um típico trabalho de casa de redação do aluno. Pediram aos alunos que escrevessem uma revisão sobre a discussão da pré-pesquisa.</p> <p>3) Pós-teste: No pós-pesquisa, mediram o nível percebido de prazer dos alunos, uma vez que o prazer durante um processo de aprendizagem tem uma grande influência na adoção de ferramentas de TI e no sucesso de aprendizagem dos alunos.</p>
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

240 - Do Multilingual Users Prefer Chat-bots that Code-mix? Let's Nudge and Find Out!

Anshul Bawa, Pranav Khadpe, Pratik Joshi, Kalika Bali, Monojit Choudhury

Maio 2020

Proceedings of the ACM on Human-Computer Interaction

Tecnologia de Avaliação: Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	Humanidade
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Específica: tecnologia de avaliação de UX específica para chatbots.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia foi criada para o estudo e descrita no artigo.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Escala likert de 7 pontos
SQ5. Qual a composição da tecnologia de avaliação?	Habilidade de falar como um humano.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Não mencionado.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Os participantes foram questionados primeiro sobre questões demográficas, sobre sua proficiência em hindi e inglês, sua formação educacional, cidades em que moraram e sua exposição anterior a chatbots. Em seguida, os participantes interagem com o chatbot monolíngue em inglês. Isso dura 15 turnos de conversação ou 15 minutos, o que for satisfeito antes. Eles são solicitados a fornecer feedback sobre essa interação. Em seguida, os participantes interagem com um dos dois chatbots do CM. Novamente, essa interação dura 15 turnos ou 15 minutos. Os participantes então concluem o estudo fornecendo feedback sobre a segunda interação e feedback comparativo entre os dois

Avaliação Empírica da Tecnologia de Avaliação

SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

254 - A new friend in our Smartphone? Observing Interactions with Chatbots in the search of emotional engagement

Manuel Portela, Carlos Granell-Canut

Setembro 2017

Interacción '17: Proceedings of the XVIII International Conference on Human Computer Interaction

Tecnologia de Avaliação Tecnologia 1: VAS
Tecnologia 2: MIM
Tecnologia 3: IRI

SQ1. Quais aspectos da UX a tecnologia avalia?	Tecnologia 1: Felicidade, Tristeza, Raiva, Surpresa, Ansiedade, Tranquilidade e Vigor; Tecnologia 2: Desconforto, Frustração/raiva/impotência, Interesse, Bem-estar; Tecnologia 3: Empatia;
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Todas Genéricas: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Todas Existentes: Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Tecnologia 1: Escala likert de 10 pontos; Tecnologia 2: Escala likert de 5 pontos; Tecnologia 3: Escala likert de 5 pontos;
SQ5. Qual a composição da tecnologia de avaliação?	Tecnologia 1: Felicidade, Tristeza, Raiva, Surpresa, Ansiedade, Tranquilidade e Vigor; Tecnologia 2: Desconforto, Frustração/raiva/impotência, Interesse, Satisfação, Bem-estar; Tecnologia 3: Empatia;
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Foram desenvolvidos dois chatbots de conversação com o objetivo de entender a natureza do engajamento emocional entre a mentalidade psicológica individual e um chatbot durante uma conversa.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.

SQ10. De qual forma o chatbot foi avaliado?	<p>Foi conduzido um experimento com 13 participantes, com idades entre 18 e 50 anos. Antes da interação com os bots, foi dado aos participantes um conjunto de pesquisas. Em seguida, também realizaram uma entrevista semiestruturada para entender o conhecimento prévio e uso dos participantes em relação à IA, Agentes de Conversação e chatbots Assistentes Virtuais , e os questionou sobre sua disposição de usar essas interfaces em um futuro próximo.</p> <p>O experimento consistiu em uma sessão individual onde cada participante recebeu um celular com o Telegram instalado e pronto para conversar com os dois chatbots. O tempo esperado para cada chatbot era de 10 minutos, mas algumas das conversas terminaram mais cedo a pedido dos participantes.</p> <p>Após a interação com ambos os sistemas (CH1 e CH2), os participantes foram solicitados a preencher um questionário. Seguiu-se uma entrevista onde os usuários explicaram sua experiência e destacaram se seu preconceito mudou sobre os agentes conversacionais. "</p>
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Sim
SQ14. De que forma a saúde mental do usuário foi avaliada?	<p>Para medir seu estado emocional utilizaram uma Escala Visual Analógica (EVA) para contabilizar 7 atitudes ou variáveis (Alegria, Tristeza, Raiva, Surpresa, Ansiedade, Tranquilidade e Vigor)</p>

283 - Chat or Tap? – Comparing Chatbots with ‘Classic’ Graphical User Interfaces for Mobile Interaction with Autonomous Mobility-on-Demand Systems

Lukas A. Flohr, Sofe Kalinke, Antonio Krüger, Dieter P. Wallach

Setembro 2021

MobileHCI '21: Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction

Tecnologia de Avaliação: UEQ

SQ1. Quais aspectos da UX a tecnologia avalia?	Eficiência, Atratividade, Novidade, Estimulação, Perspicácia, Dependabilidade
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de diferencial semântico de 7 pontos, vai de -3 até 3
SQ5. Qual a composição da tecnologia de avaliação?	<p>UEQ irritante / agradável</p> <p>não compreensível/compreensível</p> <p>criativo / maçante</p> <p>fácil de aprender/difícil de aprender</p> <p>valioso / inferior</p> <p>chato / emocionante</p> <p>não é interessante/interessante</p> <p>imprevisível/previsível</p> <p>rápido lento</p> <p>inventivo/convencional</p> <p>obstrutivo / de suporte</p> <p>bom mau</p> <p>complicado/fácil</p> <p>desagradável / agradável</p> <p>usual/borda de ataque</p> <p>desagradável / agradável</p> <p>seguro/não seguro</p> <p>motivando / desmotivando</p> <p>atende às expectativas/não atende às expectativas</p> <p>ineficiente / eficiente</p> <p>claro/confuso</p> <p>impraticável / prático</p> <p>organizado/desordenado</p> <p>atraente / não atraente</p> <p>amigável não amigável</p>

	conservador/inovador
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Chatbot desenvolvido para planejar e reservar um percurso.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é tanto orientado à tarefas quanto orientado à conversação.
SQ10. De qual forma o chatbot foi avaliado?	<p>No início da sessão de estudo, os participantes receberam um briefing detalhado sobre o objetivo e o procedimento do estudo e os possíveis efeitos colaterais do simulador. Em seguida, assinaram um formulário de consentimento de participação informativo e preencheram um pré-questionário (demografia, ATI-S). Os participantes foram incentivados a pensar em voz alta e fazer perguntas sempre que quisessem e – caso se sentissem desconfortáveis – pausar ou desistir do estudo a qualquer momento sem consequências. Para o teste baseado em tarefas, os participantes experimentaram duas jornadas AMoD na configuração baseada em laboratório . Sua tarefa era viajar com um amigo para um parque público (viagem 1) e voltar para casa (viagem 2) usando o serviço AMoD. Portanto, eles tiveram que inserir o respectivo local de destino e horário de partida e selecionar uma das três opções fornecidas pelo aplicativo.</p> <p>Após a reserva, os participantes usaram a funcionalidade de navegação integrada do aplicativo, fizeram check-in em uma interface de usuário de porta prototipada em papel usando o código QR do bilhete recebido e fizeram o passeio no simulador AV. Na jornada do ‘caminho feliz’, os participantes não receberam instruções explícitas sobre o que fazer durante o passeio. Ou seja, eles estavam livres para, por exemplo, monitorar a exibição de informações ou apenas aproveitar o passeio (simulado). Na condição CoP,</p>

	<p>ocorreram dois CoPs. (1) A alteração da hora de partida de uma viagem já reservada implicava o cancelamento da reserva inicial antes da partida e a reserva de uma nova viagem. (2) Uma mudança do destino alvo durante o passeio exigia que os participantes abortassem o passeio em andamento e reservassem um passeio de conexão para o novo destino (um restaurante).</p> <p>Quando chegaram ao destino ('caminho feliz'; após um passeio de 14 minutos), respectivamente a parada especial (CoP; após um passeio de 7 minutos), os participantes fizeram o check-out e avaliaram o serviço usando o aplicativo (Fig. 3).</p> <p>Após cada viagem, os participantes avaliaram a aceitação, UX e confiança usando um questionário digital e relataram suas mudanças emocionais durante a jornada desenhando uma curva de emoção.</p> <p>Ao final da sessão, os participantes avaliaram a percepção de presença, bem-estar e mal estar do simulador com o questionário digital e participaram de uma entrevista semiestruturada. Cada sessão durou cerca de 75 a 90 minutos e foi conduzida por um experimentador e um anotador."</p>
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

333 - SlumberBot: An Interactive Agent for Helping Users Investigate Disturbance Factors of Sleep Quality

Yizhou Liu, Da-jung Kim, Ting Miao, Yaliang Chuang

Outubro 2020

NordiCHI '20: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society

Tecnologia de Avaliação: Entrevista

SQ1. Quais aspectos da UX a tecnologia avalia?	Engajamento
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia foi criada para o estudo e descrita no artigo.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Entrevista
SQ5. Qual a composição da tecnologia de avaliação?	Questões abertas.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Qualitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Slumber-Bot é uma abordagem semiautomática para coletar medições objetivas do sono.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi realizado um estudo de campo de quatro semanas com cinco participantes com SlumberBot. O experimento dividido em três fases na linha do tempo. Na primeira fase (Semana 1), pedimos aos participantes que relatassem sua qualidade diária de sono ao acordar, usando o papel impresso CSD para criar a linha de base deste estudo. Na segunda fase (Semana 2), os participantes deveriam interagir com o SlumberBot todos os dias durante uma semana e participar de uma entrevista semiestruturada ao final desta fase. O objetivo desta entrevista foi coletar e comparar a experiência dos usuários entre

	<p>as duas ferramentas de autorrelato (ou seja, CSD e SlumberBot) e o feedback dos usuários sobre como o SlumberBot influencia o processo de autorreflexão dos usuários sobre seus problemas de sono durante um curto período de tempo (ou seja, uma semana). Na última fase (da semana 3 à semana 4), os participantes foram convidados a continuar usando o SlumberBot por mais duas semanas e participaram de uma entrevista final no final do estudo. O objetivo da entrevista final era entender a influência do SlumberBot por um período relativamente longo (ou seja, duas semanas).</p>
<p>Avaliação Empírica da Tecnologia de Avaliação</p>	
<p>SQ12. A avaliação da tecnologia foi avaliada empiricamente?</p>	<p>Não</p>
<p>SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?</p>	<p>N.A.</p>
<p>Avaliação do Estado de saúde emocional do usuário</p>	
<p>SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?</p>	<p>Não</p>
<p>SQ14. De que forma a saúde mental do usuário foi avaliada?</p>	<p>N.A.</p>

529 - Airbot: using a work flow model for proactive assistance in public spaces

Markus Kattenbeck, Florian Alt, Melanie Pflamminger, Bernd Ludwig, Matthias Ferstl

Setembro 2018

MobileHCI '18: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct

Tecnologia de Avaliação: Questionário

SQ1. Quais aspectos da UX a tecnologia avalia?	Divertido, informativo, usaria novamente.
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Específica: tecnologia de avaliação de UX específica para chatbots.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia foi criada para o estudo e descrita no artigo.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Numa escala Likert de 1 até 5
SQ5. Qual a composição da tecnologia de avaliação?	Não descrito.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Quantitativos.
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	É um chatbot que auxilia usuários num aeroporto.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é tanto orientado à tarefas quanto orientado à conversação.
SQ10. De qual forma o chatbot foi avaliado?	Foi feito um experimento com 101 passageiros, em que, pesquisa para avaliar o assistente pessoal foi projetada para ser aplicável tanto aos participantes do grupo de tratamento (usuários do AIRBOT) quanto aos participantes do grupo de controle (não usuários do AIRBOT) e teve duas partes. A primeira parte foi apresentada a todos os usuários. Em seguida responderam um questionário sobre a experiência no aeroporto. Enquanto os passageiros do grupo de controle só precisavam responder a perguntas sobre esses quatro fatores, os usuários do AIRBOT foram solicitados a responder a perguntas adicionais na segunda parte.

	Essas perguntas avaliaram a experiência dos usuários com o AIRBOT.O experimentador explicou o tipo de assistência que o AIRBOT foi capaz de dar e pediu aos participantes que se lembrassem de responder à pesquisa final logo antes da decolagem ou logo após o pouso em seu destino.
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não.
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.

533 - When the Social Becomes Non-Human: Young People's Perception of Social Support in Chatbots

Petter Bae Brandtzaeg, Marita Skjuve, Kim Kristoffer Dysthe Asbjørn Følstad

Maio 2021

CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems

Tecnologia de Avaliação: Entrevista

SQ1. Quais aspectos da UX a tecnologia avalia?	Confiança, privacidade, apreciação e suporte emocional
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Específica: tecnologia de avaliação de UX específica para chatbots.
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Criada: tecnologia foi criada para o estudo e descrita no artigo.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Questões abertas.
SQ5. Qual a composição da tecnologia de avaliação?	Não descrito.
SQ6. A tecnologia extrai dados quantitativos ou qualitativos?	Qualitativos.

Avaliação dos Chatbots

SQ7. Qual é a aplicação do Chatbot?	O Woebot conduz a CBT – uma abordagem baseada em evidências para o tratamento de vários problemas psicológicos – e é projetada para que os princípios da CBT sejam utilizados de maneira amigável e envolvente. O chatbot se assemelha a um amigo que verifica o humor do usuário e dá conselhos sobre experiências e preocupações. O Woebot também oferece questionários e vídeos que ajudam os usuários a descobrir padrões de pensamento habituais ou automatizados que afetam seu bem-estar e saúde mental.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação
SQ10. De qual forma o chatbot foi avaliado?	"Foi feito um experimento com 16 participantes. Os participantes foram convidados a usar o Woebot todos os dias por 14 dias e poderiam se retirar do estudo a qualquer momento. Após 14 dias de uso, eles participaram de uma entrevista

	<p>individual, realizada presencialmente.</p> <p>Como parte da sessão de entrevistas, os participantes também experimentaram um protótipo de chatbot, o Ungbot, direcionado aos jovens. Finalmente, dois meses depois, foram realizadas entrevistas por e-mail com os participantes. As entrevistas foram gravadas em áudio e transcritas. A maioria das entrevistas durou aproximadamente uma hora ou mais (média: 1 hora e 12 minutos)</p> <p>As entrevistas constaram de três partes. A primeira parte das entrevistas concentrou-se em questões abertas sobre as experiências gerais dos entrevistados com avaliação, apoio emocional, informativo e instrumental. A segunda parte consistiu em uma pergunta aprofundada sobre as experiências e reflexões dos participantes sobre os chatbots em geral, e o Woebot em particular, como fonte de apoio social. A terceira parte, depois de dois meses, deram continuidade a um estudo de e-mail. Perguntamos se eles usaram o Woebot ou chatbots semelhantes após a entrevista. Se eles respondessem “não”, foi pedido que explicassem o motivo e, se respondessem “sim”, foi pedido que compartilhassem suas experiências e motivação para usar chatbots”</p>
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Sim
SQ14. De que forma a saúde mental do usuário foi avaliada?	Foi avaliado o suporte emocional do chatbot, através de expressões de amor, empatia, confiança e preocupação.

626 - A Mental Health Chatbot for Regulating Emotions (SERMO) - Concept and Usability Test

Kerstin Denecke, Sayan Vaaheesan, and Aaganya Arulnathan

Fevereiro 2020

IEEE Transactions on Emerging Topics in Computing

Tecnologia de Avaliação: UEQ

SQ1. Quais aspectos da UX a tecnologia avalia?	Atratividade, Perspicuidade, Eficiência, Confiabilidade, Estimulação, Novidade
SQ2. A tecnologia é específica para chatbots ou para sistemas em geral?	Genérica: tecnologia de avaliação de UX não é restrita a tipos específicos de software
SQ3. A tecnologia foi criada para o estudo ou foi baseada numa já existente?	Existente: tecnologia é baseada em uma existente.
SQ4. Como a tecnologia de avaliação coleta as respostas dos participantes	Através de diferencial semântico de 7 pontos, vai de -3 até 3
SQ5. Qual a composição da tecnologia de avaliação?	UEQ irritante / agradável não compreensível/compreensível criativo / maçante fácil de aprender/difícil de aprender valioso / inferior chato / emocionante não é interessante/interessante imprevisível/previsível rápido lento inventivo/convencional obstrutivo / de suporte bom mau complicado/fácil desagradável / agradável usual/borda de ataque desagradável / agradável seguro/não seguro motivando / desmotivando atende às expectativas/não atende às expectativas ineficiente / eficiente claro/confuso impraticável / prático organizado/desordenado atraente / não atraente amigável não amigável conservador/inovador
SQ6. A tecnologia extrai dados quantitativos	Quantitativos.

ou qualitativos?	
Avaliação dos Chatbots	
SQ7. Qual é a aplicação do Chatbot?	Chatbot que auxilia o usuário na regulação de suas emoções.
SQ8. O chatbot foi criado para um grupo específico de pessoas? Qual grupo?	Não, o chatbot não foi criado para um grupo específico.
SQ9. O Chatbot é de algum tipo específico? Qual?	Sim, o chatbot é orientado a conversação.
SQ10. De qual forma o chatbot foi avaliado?	Os testes foram feitos por seis tarefas. Os usuários foram solicitados a realizar as tarefas para testar as funcionalidades específicas e fornecer feedback se conseguiram concluir a tarefa (sim/não) ou se e quais problemas ocorreram. As tarefas incluíam Defina um objetivo, Entre em um estado de espírito, Insira um evento atual, Escolha uma atividade agradável, Faça um exercício de atenção plena, Converse com SERMO por pelo menos 1 minuto. Na segunda parte, os participantes tiveram que julgar aspectos concretos da experiência do usuário. Para isso, foi aplicado o questionário de experiência do usuário (UEQ).
Avaliação Empírica da Tecnologia de Avaliação	
SQ12. A avaliação da tecnologia foi avaliada empiricamente?	Não
SQ13. Se sim, qual foi o tipo do experimento conduzido para avaliar a tecnologia de avaliação?	N.A.
Avaliação do Estado de saúde emocional do usuário	
SQ13. A avaliação de UX considera algum aspecto da saúde emocional do usuário que está interagindo com o chatbot?	Não
SQ14. De que forma a saúde mental do usuário foi avaliada?	N.A.