

Avaliação da experiência do usuário na utilização de um protótipo de chatbot para o apoio a prevenção do câncer do colo do útero.

Abstract—Este estudo teve como objetivo avaliar a experiência do usuário na utilização do protótipo de chatbot Liz, desenvolvido para apoiar a prevenção do câncer do colo do útero por meio de informações, lembretes e agendamento do exame preventivo. Foi realizada uma pesquisa de abordagem quantitativa e qualitativa com 58 participantes, que utilizaram o chatbot e responderam aos instrumentos System Usability Scale (SUS), Technology Acceptance Model (TAM) e Flow Experience Questionnaire FEQ, além de uma questão aberta. Os dados quantitativos foram analisados por meio dos escores dos instrumentos, enquanto os dados qualitativos foram submetidos à análise lexical no software IRAMUTEQ, utilizando nuvem de palavras, análise de similitude e Análise Fatorial de Correspondência (AFC). Os resultados quantitativos indicaram boa usabilidade, com escore SUS de 75,5 pontos, além de elevada aceitação, utilidade percebida, facilidade de uso e experiência de fluxo. Na análise qualitativa, o termo *agendamento* apresentou maior centralidade, associado a aspectos como facilidade, clareza e funcionalidade. A análise de similitude e a AFC reforçaram a predominância de percepções positivas, embora também tenham identificado dificuldades pontuais na interação. Os resultados indicam que o chatbot Liz apresenta boa usabilidade, aceitação e experiência de interação, demonstrando potencial como ferramenta de apoio à prevenção do câncer do colo do útero.

Index Terms—Chatbot, Interação, Prevenção, Câncer do colo do útero.

I. INTRODUÇÃO

O câncer é um importante problema de saúde pública, caracterizado pelo crescimento desordenado de células que podem invadir tecidos adjacentes ou outros órgãos. No Brasil, estima-se a ocorrência de aproximadamente 781 mil casos novos de câncer para cada ano do triênio 2026–2028, considerando todas as localizações tumorais, incluindo o câncer de pele não melanoma [9], [18].

Entre os tipos de câncer que acometem a população feminina, destaca-se o câncer do colo do útero, associado à infecção persistente por determinados tipos do papilomavírus humano (HPV). Embora a infecção seja frequente e, na maioria dos casos, regrida espontaneamente, algumas alterações celulares podem evoluir para lesões precursoras e, posteriormente, para o câncer. A identificação dessas alterações por meio do rastreamento possibilita a adoção de medidas preventivas e terapêuticas oportunas [17]. Nesse contexto, o exame citopatológico, conhecido como Papanicolau, constitui uma estratégia importante para detectar alterações cervicais e prevenir a progressão da doença [26].

No município de Nova Andradina, localizado na região leste de Mato Grosso do Sul, o rastreamento do câncer do colo

do útero é realizado principalmente no âmbito da Estratégia Saúde da Família (ESF). Essa estratégia busca reorganizar a assistência à saúde, articulando ações de prevenção, promoção e acompanhamento da população [6]. Além disso, a atenção primária constitui um espaço relevante para a organização das ações de rastreamento e para a ampliação do acesso à informação sobre prevenção e cuidados em saúde [29].

Contudo, de acordo com Almeida [2], o câncer de colo do útero afeta principalmente os grupos socioeconômicos mais baixos, e questões sociais influenciam no adoecimento e no acesso ao tratamento oncológico. Diante dessa realidade, foi necessário buscar um meio de comunicação que fosse comum e abrangesse a maioria das pacientes, independentemente de seu nível socioeconômico ou nível de educação.

Neste contexto, o WhatsApp se destaca como o aplicativo de comunicação mais utilizado. Está presente em 99% dos smartphones dos brasileiros, sendo que 93% utilizam o aplicativo diariamente. Dos brasileiros, 90% utilizam o aplicativo para enviar mensagens de texto, e 81% se comunicam por áudio [30]. Na área médica, a utilização do WhatsApp como ferramenta para fortalecer a relação entre profissional de saúde e paciente tem sido positiva, em conformidade com o Parecer CFM N° 14/2017, abrangendo uma ampla gama de aspectos [23].

Nesse cenário, os chatbots, também denominados agentes conversacionais, apresentam potencial para apoiar a disseminação de informações sobre saúde e facilitar o esclarecimento de dúvidas relacionadas à prevenção. Entretanto, a utilização deste tipo de tecnologia na saúde exige avaliações que considerem não apenas seu funcionamento técnico, mas também a qualidade da interação, a usabilidade e a percepção dos usuários [32]–[34]. Na perspectiva da Interação Humano-Computador (IHC), torna-se fundamental investigar como os usuários percebem a utilidade, a facilidade de uso, a clareza, a satisfação e a confiabilidade da interação com esses sistemas. Assim, a avaliação deve considerar tanto aspectos pragmáticos, como a facilidade de encontrar e compreender informações, quanto aspectos subjetivos, como a satisfação durante a interação [32], [37].

Todavia, os recursos humanos sempre são limitados e a comunicação, quando feita diretamente por uma pessoa, pode sofrer interferências, prolongando, atrasando ou até mesmo anulando a efetividade da comunicação entre profissional de saúde e paciente. Dessa forma, uma das alternativas para reduzir esses erros e viabilizar o acompanhamento médico

das pacientes é o uso de Tecnologias da Informação e Comunicação (TIC), especialmente por meio de técnicas de Inteligência Artificial, como os chatbots. Os chatbots são programas de computador que interagem com os usuários usando linguagem natural, podendo ser por texto ou voz. Essa tecnologia teve início na década de 1960, com o objetivo de simular a conversa humana. No entanto, com o avanço no desenvolvimento de chatbots, arquiteturas e capacidades computacionais, seu uso se expandiu consideravelmente [1]. Para compreender a entrada do usuário e fornecer uma resposta relevante, os chatbots utilizam Inteligência Artificial (IA) e técnicas de Aprendizado Profundo (Deep Learning).

Assim, com base nos conceitos e no problema apresentado, foi desenvolvido um protótipo de chatbot chamado Liz [24], como forma de avaliação inicial da proposta. Essa iniciativa contou com a parceria entre o Instituto Federal de Mato Grosso do Sul, campus de Nova Andradina (IFMS/NA), a Prefeitura de Nova Andradina (PMNA) e o apoio da Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT - PICTEC 1 - 02/2021).

Sua função é automatizar a comunicação entre os profissionais da saúde e as pacientes, atuando principalmente no lembrete, agendamento e reagendamento do exame preventivo, além de fornecer informações básicas sobre os problemas e conceitos relacionados ao câncer do colo do útero.

O protótipo, concebido como uma representação preliminar do sistema para fins de prova de conceito, deve ser submetido a uma avaliação centrada no usuário, considerando a experiência das usuárias durante a interação com a solução. Essa etapa deverá permitir analisar aspectos relacionados à usabilidade, facilidade de interação, compreensão das funcionalidades, satisfação e adequação do sistema às necessidades das usuárias e ao contexto de utilização [25]. A avaliação deverá envolver a participação de usuários representativos do público-alvo, que poderão executar tarefas previamente definidas utilizando o protótipo, possibilitando a coleta de dados quantitativos e qualitativos sobre sua experiência de uso [5]. Conforme os princípios de projeto centrado no ser humano estabelecidos pela ISO 9241-210, o feedback dos usuários constitui uma fonte fundamental para identificar limitações e oportunidades de melhoria, permitindo que as soluções sejam progressivamente refinadas a partir de avaliações realizadas em contextos representativos de uso [19]. Dessa forma, os resultados obtidos na avaliação da experiência do usuário poderão subsidiar novas iterações do protótipo, contribuindo para o aprimoramento de suas funcionalidades e para a redução de incertezas relacionadas à usabilidade, aos requisitos e à adequação da solução antes de uma eventual implantação.

Portanto, a proposta deste trabalho foi a avaliação do protótipo já desenvolvido. A avaliação da experiência do usuário constitui uma etapa importante para analisar a usabilidade, a aceitação e a percepção dos usuários em relação a sistemas interativos, sendo especialmente relevante em aplicações na área da Saúde [3]. Diante desse contexto, estabelece-se a seguinte pergunta de pesquisa: como as usuárias avaliam a

usabilidade, a aceitação e a experiência de fluxo durante a interação com o protótipo de chatbot Liz, desenvolvido para apoiar a comunicação e fornecer informações relacionadas à prevenção do câncer do colo do útero?

A hipótese deste trabalho é que o protótipo Liz apresentará resultados positivos na avaliação da usabilidade e da aceitação pelas usuárias, bem como na percepção da experiência de fluxo durante a interação. Espera-se que os resultados indiquem que a solução é compreensível, apresenta facilidade de uso e proporciona uma experiência de interação satisfatória, considerando as necessidades do público-alvo e o contexto de utilização. Essa hipótese será examinada com base nos indicadores coletados durante a avaliação do protótipo, sem pressupor efeitos sobre a adesão ao exame preventivo, uma vez que esse desfecho não foi mensurado no estudo.

II. FUNDAMENTAÇÃO TEÓRICA

III. TRABALHOS RELACIONADOS

No que tange à avaliação de agentes conversacionais, Carvalho et al. [13] conduziram um mapeamento sistemático da literatura com 27 estudos para caracterizar as tecnologias de usabilidade aplicadas a chatbots [13]. Os resultados apontam um predomínio do uso de questionários (86,95%), destacando-se ferramentas consagradas como o *System Usability Scale* (SUS) e o *Technology Acceptance Model* (TAM) [13]. Como principal lacuna na literatura, os autores identificaram que 100% dos estudos avaliados aplicam seus testes de forma tardia, concentrando-se exclusivamente na fase de Verificação, Validação e Teste (VV&T), evidenciando a carência de abordagens de usabilidade voltadas para as etapas iniciais de requisitos e design [13].

Já Silva et al. [28] investigaram os fatores que influenciam a intenção de utilização de chatbots no comércio eletrônico, com ênfase na confiança, risco percebido e experiência de *flow*. O estudo quantitativo contou com 226 participantes europeus, que interagiram com um chatbot desenvolvido no Chatfuel e responderam posteriormente a um questionário. Os dados foram analisados por meio de Modelagem de Equações Estruturais (SEM). Os resultados indicaram que a expectativa de desempenho e, principalmente, a confiança influenciaram positivamente a intenção de uso, enquanto a confiança também reduziu o risco percebido e aumentou a experiência de *flow*. Os demais construtos não apresentaram efeitos diretos significativos sobre a intenção de utilização.

Por fim, Martins et al. [22] desenvolveram, validaram e avaliaram a usabilidade do chatbot ChatLinoHans para apoiar profissionais de saúde na abordagem de lesões da Hanseníase. O estudo metodológico utilizou uma arquitetura híbrida, combinando regras clínicas e GPT-4, e empregou o IRaMuTeQ na organização e análise do corpus utilizado na construção da base de conhecimento. A validação envolveu profissionais de saúde e Tecnologia da Informação, obtendo IVC global de 0,98, concordância de 85,7% a 100% na avaliação da aparência e pontuações de usabilidade entre 77,5 e 100. Os resultados indicaram evidências satisfatórias de validade e usabilidade do chatbot.

Em conjunto, os estudos evidenciam o predomínio de abordagens quantitativas na avaliação de chatbots, com uso de questionários, escalas e métodos estatísticos. Embora o IRaMuTeQ tenha sido empregado por Martins et al. [22], sua aplicação ocorreu na construção da base de conhecimento. Dessa forma, observa-se uma lacuna quanto ao uso de análises qualitativas das interações dos usuários, indicando a necessidade de abordagens que combinem avaliação quantitativa e qualitativa.

IV. METODOLOGIA

A. Aspectos éticos

A Resolução nº 466/2012 do Conselho Nacional de Saúde (CNS) [10] estabelece que todo projeto de pesquisa envolvendo seres humanos deve seguir as diretrizes nela previstas. No entanto, conforme o Artigo 1º, parágrafo único, da Resolução nº 510, de 7 de abril de 2016 [11], certas atividades não precisam ser registradas nem avaliadas pelo sistema CEP/CONEP, como é o caso de pesquisas de opinião pública com participantes não identificados.

Neste estudo, essa situação é descrita na seção referente ao percurso metodológico, especificamente na etapa 3 — aplicação do teste. Nessa fase, não houve identificação dos respondentes, tampouco coleta de dados que pudessem associar respostas a pessoas específicas, garantindo assim a privacidade e o sigilo dos participantes e proporcionando um ambiente de segurança e confidencialidade. Ainda assim, em conformidade com o item II.3 da referida normativa, foi assegurada assistência aos participantes durante todo o processo, mesmo com o anonimato mantido.

A Resolução nº 466/2012 também destaca, em seu item II.23, que o respeito à dignidade humana exige a obtenção do consentimento livre e informado dos participantes. Essa etapa está detalhada no percurso metodológico: os participantes foram devidamente informados sobre os objetivos do estudo e cientes de que sua participação seria voluntária. Para isso, foi utilizado o Termo de Consentimento Livre e Esclarecido (TCLE) antes do preenchimento do formulário, garantindo que o participante tivesse conhecimento completo sobre a pesquisa antes de iniciar sua participação.

Por fim, cabe ressaltar que, em conformidade com a Resolução nº 466/2012 [10], não houve envolvimento de menores de idade e que todos os dados coletados foram utilizados exclusivamente para os fins previstos neste estudo, sem qualquer uso adicional.

B. Tipo da Pesquisa

O estudo adotou uma abordagem de caráter quantitativo, utilizando procedimentos de quantificação tanto na etapa de coleta das informações quanto na análise dos dados por meio de técnicas estatísticas [27] e qualitativa buscando a compreensão detalhada dos significados e características situacionais apresentados [27].

A pesquisa teve natureza aplicada, buscando produzir conhecimentos voltados à utilização prática, orientados para a solução de problemas específicos. Quanto aos objetivos, foi

classificada como exploratória, por procurar ampliar a compreensão sobre o problema e torná-lo mais claro. Como procedimento metodológico, empregou-se o levantamento, com a intenção de reunir dados organizados que possibilitassem uma análise mais aprofundada [16].

C. Percurso metodológico

A avaliação do chatbot seguiu a estrutura metodológica proposta por Barbosa et al. [3], que organiza o processo em cinco atividades fundamentais, garantindo o rigor científico e a validade dos resultados coletados.

1. Contextualização e Análise de Domínio: Inicialmente, realizou-se um estudo sobre a situação atual para compreender o domínio do problema, os perfis dos usuários e o contexto de uso do sistema. Esta etapa permitiu identificar os objetivos e as tarefas críticas que serviriam de base para os cenários de avaliação.

2. Planejamento e Preparação: Nesta fase, foram definidos os objetivos específicos da avaliação e o seu escopo. O planejamento incluiu o recrutamento de participantes, a preparação da infraestrutura tecnológica e dos materiais de apoio, como termos de consentimento livre e esclarecido (TCLE) e roteiros. Como meios para a avaliação quantitativa foram utilizados o Modelo de Aceitação da Tecnologia (TAM), teste de usabilidade System Usability Scale (SUS) e o Questionários de experiência de fluxo (FEC). Já para a parte qualitativa foi adicionada uma pergunta aberta, na qual o usuário poderia expor sua opinião após utilizar o software. Previamente à execução principal, foi realizado um Teste Piloto para validar o planejamento e os materiais, permitindo ajustes antes da coleta definitiva de dados [4].

2. Coleta de Dados e Observação: A execução seguiu os preceitos éticos, com o acolhimento cordial dos participantes e a coleta dos dados mediante a aceitação do termo de consentimento. Como condução os participantes foram estimulados a utilizar todas as funcionalidades da aplicação por meio de um roteiro, como pode ser visto na Figura 1. Após o uso, estes seguiam para responder o questionário.

Os dados foram registrados por meio de formulários eletrônicos sendo estes mantidos para análise posterior.

4. Interpretação de Dados: Os dados registrados foram submetidos a um processo de análise seguindo rigorosamente as especificações de cada técnica, onde os avaliadores buscaram atribuir significado aos dados coletados. Para os dados qualitativos foi utilizado o software Irameteq. Esta etapa focou em identificar rupturas na comunicação e responder às perguntas de pesquisa formuladas no planejamento, podendo incluir análises manuais ou automáticas, dependendo do tipo de dado registrado.



Fig. 1. Roteiro para uso do Chatbot Liz

5. Consolidação e Relato dos Resultados: Por fim, os achados individuais foram consolidados em uma análise para identificar recorrências e tendências de problemas que afetam grupos de usuários. As descobertas foram organizadas em um relatório final detalhando os objetivos, o método aplicado, o perfil dos participantes, o sumário dos dados, a lista de problemas identificados e, crucialmente, sugestões fundamentadas de reprojetos da interface e interação

V. ANÁLISE DOS RESULTADOS

O estudo contou com 58 participantes, com idades entre 18 e 64 anos. Embora o chatbot seja direcionado ao agendamento do exame para o público feminino, não houve necessidade, nesta etapa, de separação por gênero, pois a avaliação concentrou-se exclusivamente na experiência de uso do sistema. Estudos mais específicos poderão ser conduzidos em fases posteriores e estão disponíveis mediante solicitação.

A. Análise quantitativa

1) Usabilidade: Os resultados da aplicação do SUS, apresentados na Figura 2, indicam uma percepção predominantemente positiva dos participantes em relação à usabilidade do sistema. Observa-se elevada concordância nos itens formulados positivamente, especialmente nos itens 3 e 7, que registraram 83% e 86% de respostas favoráveis, respectivamente. Esses resultados sugerem que os participantes consideraram o sistema fácil de utilizar e de aprender.

Nos itens formulados negativamente, também se verificou uma tendência favorável, evidenciada pela predominância de respostas de discordância. Destacam-se os itens 4 e 10, nos quais 86% e 85% dos participantes, respectivamente, discordaram das afirmações negativas. De forma semelhante, 69% discordaram do item 2 e 64% do item 8, indicando baixa percepção de complexidade, dificuldade ou inconsistência na utilização do sistema.

Apesar da tendência geral positiva, o item 6 apresentou maior dispersão nas respostas, com 41% de discordância, 31% de neutralidade e 29% de concordância, constituindo o resultado menos homogêneo entre os dez itens avaliados. Esse comportamento sugere a existência de um aspecto específico da experiência de uso que merece investigação adicional. De

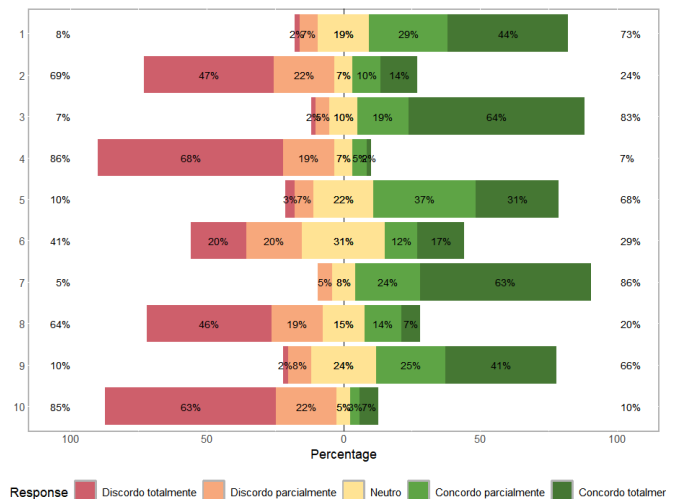


Fig. 2. Resultados do SUS para o Chatbot Liz

modo geral, entretanto, a distribuição das respostas evidencia uma avaliação favorável da usabilidade do sistema, especialmente no que se refere à facilidade de uso, aprendizagem e ausência de complexidade percebida.

A partir da distribuição das respostas, foi calculado o escore global. Considerando o procedimento padrão de cálculo do instrumento, o sistema obteve um escore estimado de aproximadamente **75,5 pontos**.

2) Aceitação: O questionário de aceitação foi desenvolvido com base nos indicadores do Modelo de Aceitação de Tecnologia (TAM), proposto por Venkatesh e Davis [3]. Esse modelo busca compreender os fatores que influenciam a adoção e o uso de novas tecnologias, permitindo a coleta de métricas relacionadas à aceitação e intenção de uso por parte dos usuários.

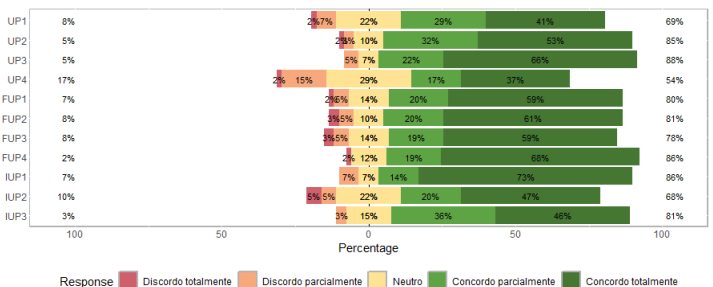


Fig. 3. Resultados do Questionário de Aceitação (TAM)

A análise dos resultados do TAM evidencia uma forte tendência de concordância entre os participantes, especialmente nos itens relacionados à utilidade percebida (UP) e à facilidade de uso percebida (FUP). Os itens UP1, UP2 e UP3 apresentaram percentuais elevados de concordância, indicando que os usuários perceberam o chatbot como uma ferramenta útil para o agendamento do exame. De maneira semelhante, os itens FUP1, FUP2 e FUP3 demonstraram que

os participantes consideraram o sistema simples de operar, reforçando a percepção de facilidade de uso.

Os itens associados à intenção de uso percebida (IUP) também apresentaram resultados positivos, com predominância de respostas favoráveis. Isso sugere que os participantes demonstraram predisposição para utilizar novamente o chatbot, reconhecendo seu valor prático e sua eficiência no contexto proposto. A baixa incidência de discordância e a presença moderada de neutralidade reforçam que, embora alguns aspectos possam ter sido percebidos com menor intensidade, o conjunto da experiência foi amplamente satisfatório e alinhado aos princípios do TAM.

3) **Experiência de Fluxo:** No que diz respeito à Experiência de Fluxo, o FEQ é utilizado para avaliar a imersão e o engajamento, fundamentado na teoria do flow, definida por Csikszentmihalyi e Csikszentmihalyi [14], que corresponde ao estado mental de um participante em relação à atividade realizada, caracterizado por elevado nível de concentração durante o envolvimento. Jackson e Eklund [20] desenvolveram e validaram a *Flow State Scale* (FSS), que possibilita mensurar o estado de flow em diferentes contextos. A aplicação do Questionário de Experiência de Fluxo teve como objetivo avaliar as sensações psicológicas dos participantes durante a utilização da ferramenta. Os participantes foram questionados sobre os itens apresentados na Tabela 14. Ressalta-se que, nesta pesquisa, foi utilizada a versão *short* da FSS, que avalia nove dimensões essenciais do modelo original proposto por Csikszentmihalyi [15].

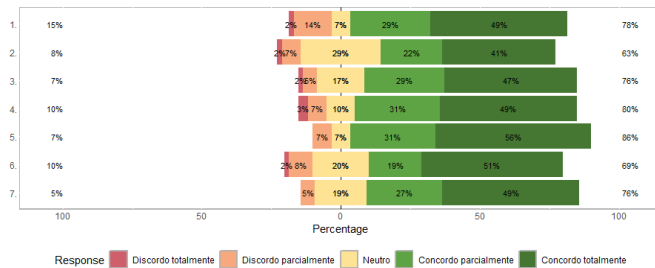


Fig. 4. Resultados do Questionário de Experiência de Fluxo

Os resultados evidenciam uma tendência clara de concordância elevada entre os participantes, indicando que a ferramenta proporcionou uma experiência alinhada aos princípios centrais da teoria do flow. A predominância de respostas positivas — especialmente nas categorias *Concordo parcialmente* e *Concordo totalmente* — revela que os usuários vivenciaram estados de imersão, concentração e envolvimento compatíveis com as dimensões avaliadas pela versão *short* da FSS. A baixa incidência de discordância e a presença moderada de neutralidade reforçam que, embora alguns aspectos possam ter sido percebidos com menor intensidade, o conjunto da experiência foi amplamente satisfatório e favoreceu a sensação de controle, clareza de objetivos e prazer intrínseco durante a atividade. Esses achados sugerem que a ferramenta foi eficaz em promover condições psicológicas propícias ao estado de flow, fortalecendo sua adequação ao contexto investigado.

De forma integrada, os resultados obtidos por meio dos três instrumentos de avaliação — SUS, TAM e FEQ — revelam que o chatbot apresenta desempenho sólido em diferentes dimensões da experiência do usuário. O resultados também conversam com os trabalhos de Silva et al. [28] e Camargo et al. [12], na qual apresentam resultados satisfatórios na aplicação das ferramentas de mensuração. Os achados do SUS indicam que o sistema é amplamente percebido como fácil de usar, aprender e operar, com baixa percepção de complexidade ou inconsistência. Complementarmente, o questionário baseado TAM demonstra que os participantes reconhecem a utilidade prática da ferramenta, consideram sua interação simples e manifestam intenção positiva de utilizá-la novamente, evidenciando forte aceitação da tecnologia. Por fim, os resultados da FEQ mostram que o chatbot é capaz de promover estados de imersão, concentração e envolvimento compatíveis com a teoria do *flow*, reforçando que a experiência proporcionada é não apenas funcional, mas também psicologicamente engajadora. Em conjunto, esses instrumentos apontam que o sistema oferece boa usabilidade, é bem aceito pelos usuários e favorece uma interação envolvente, sustentando sua adequação ao contexto de aplicação e seu potencial para adoção contínua.

B. Análise qualitativa

Para a análise qualitativa foi utilizado o Software IRAMUTEQ. O IRAMUTEQ é um software gratuito, desenvolvido sob a lógica open source e licenciado pela GNU GPL (v2). Ele funciona ancorado no ambiente estatístico do software R, permitindo a realização de diferentes tipos de análise de dados textuais. Entre suas funcionalidades estão desde análises simples, como a lexicografia básica (cálculo de frequência de palavras), até procedimentos multivariados, como a classificação hierárquica descendente e análises de similitude. Além disso, o programa organiza a distribuição do vocabulário de forma clara e visualmente compreensível, por meio de recursos como nuvem de palavras. Nas análises lexicais clássicas, o IRAMUTEQ identifica e reformata as unidades de texto, transformando Unidades de Contexto Iniciais (UCI) em Unidades de Contexto Elementares (UCE), calcula frequência média, número de hapax (palavras com ocorrência única), realiza lematização e cria dicionários de formas reduzidas, distinguindo entre formas ativas e suplementares [12].

1) **Núvem de Palavras:** A nuvem de palavras, como pode ser visto na Figura 5, as agrupa e as organiza graficamente em função da sua frequência. É uma análise lexical mais simples, porém graficamente bastante interessante, na medida em que possibilita rápida identificação das palavras-chave de um corpus [12].

A análise evidencia as palavras em maior destaque, como *agendamento*, *chatbot*, *mensagem*, *muito*, *achar*, *bem* e *vez*, indicam que os usuários associaram fortemente o uso da ferramenta ao processo de marcar consultas e à comunicação automatizada. Outros termos relevantes, como *cadastro*, *confirmar*, *informação* e *atendimento*, reforçam a percepção de clareza e praticidade na interação. Já palavras como *fácil*, *bom*



Fig. 5. Núvem de Palavras

e *usuário* apontam para uma avaliação positiva da usabilidade, enquanto termos como *erro* e *pessoa* sugerem dificuldades pontuais. Dessa forma, a nuvem de palavras complementa os resultados quantitativos, revelando que o chatbot foi associado a uma experiência prática, clara e predominantemente positiva [12].

2) *Análise de similitude*: A análise de similitude, fundamentada na teoria dos grafos, permite identificar as coocorrências entre os termos de um corpus textual. Seu resultado evidencia as conexões existentes entre as palavras, auxiliando na compreensão da estrutura do texto e destacando tanto os elementos comuns quanto as especificidades relacionadas às variáveis ilustrativas consideradas na análise. Esse procedimento contribui para revelar padrões de associação semântica e a organização interna do discurso, oferecendo uma visão mais detalhada das relações entre os conteúdos expressos [12].

A análise de similitude, como vista na Figura 6, evidencia que o termo *agendamento* ocupa posição central na rede semântica, funcionando como núcleo das interações relacionadas pelos participantes. As conexões mais fortes estabelecidas com palavras como *mensagem*, *vez*, *bem* e *achar* indicam que o processo de marcar consultas foi o aspecto mais recorrente e estruturador da experiência de uso do chatbot. Além disso, ramificações secundárias que incluem termos como *erro*, *pessoa* e *fácil* revelam tanto percepções positivas de simplicidade quanto relatos de dificuldades pontuais. A presença de palavras como *cadastro* e *confirmação* reforça a importância dos elementos funcionais da interação, enquanto *ideia* e *problema* apontam para reflexões críticas sobre o uso da ferramenta. Em conjunto, essa análise qualitativa complementa os resultados quantitativos, demonstrando que o chatbot Liz foi fortemente associado ao ato de agendar exames, mas também trouxe à tona nuances ligadas à clareza das mensagens, à facilidade de uso e às limitações percebidas pelos usuários.

3) **Análise Fatorial de Correspondência:** Por fim, a Análise Fatorial de Correspondência (AFC) apresentada na Figura 7, permite visualizar a distribuição das palavras nos diferentes quadrantes do plano cartesiano, evidenciando

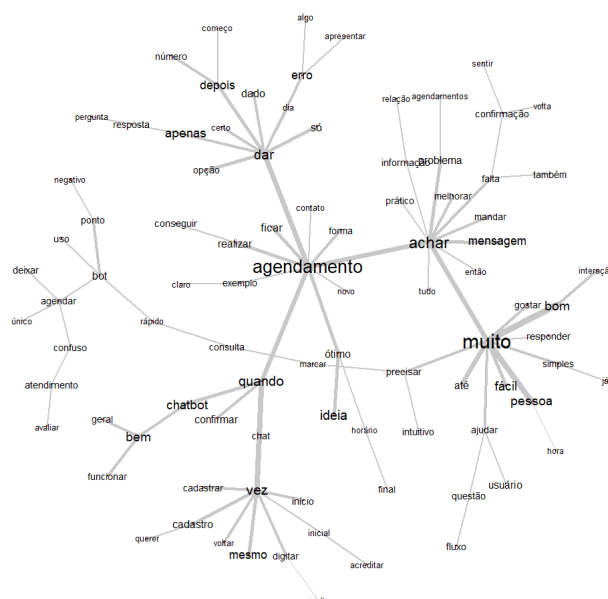


Fig. 6. Arvoré de Similitudes

padrões de associação e oposição entre os termos. De acordo com [21], a interpretação da AFC depende da posição das palavras em relação aos eixos fatoriais, de modo que termos próximos tendem a compartilhar significados ou contextos semelhantes, enquanto aqueles distantes indicam menor relação semântica. No presente corpus, observa-se que palavras como *ótimo*, *intuitivo*, *geral* e *pessoa* se agrupam em um mesmo quadrante, sugerindo avaliações positivas e percepções relacionadas à experiência geral de uso. Em contraste, termos como *digital*, *único*, *responder* e *até* aparecem mais dispersos, indicando especificidades ou menor associação ao núcleo semântico predominante.

Os eixos fatoriais, que explicam 24,5% e 17,8% da variância, respectivamente, organizam duas dimensões principais do discurso: uma relacionada à avaliação da experiência e outra à operacionalidade da interação. Considerando a distância entre os termos e sua separação entre quadrantes, é possível inferir relações de dependência moderada entre palavras próximas e dependência baixa ou nula entre aquelas posicionadas em lados opostos dos eixos. Assim, a AFC revela tanto agrupamentos semânticos consistentes quanto dispersões que refletem a diversidade de percepções dos usuários sobre o chatbot.

A integração dos três métodos qualitativos aplicados — nuvem de palavras, análise de similitude e Análise Fatorial de Correspondência (AFC) — evidencia um padrão consistente nas percepções dos participantes sobre o uso do chatbot Liz, compatível com os apresentados por Martins et al. [22]. A nuvem de palavras destacou a centralidade do termo *agendamento*, acompanhado de palavras como *mensagem*, *vez* e *bem*, indicando que o processo de marcação de consultas foi o

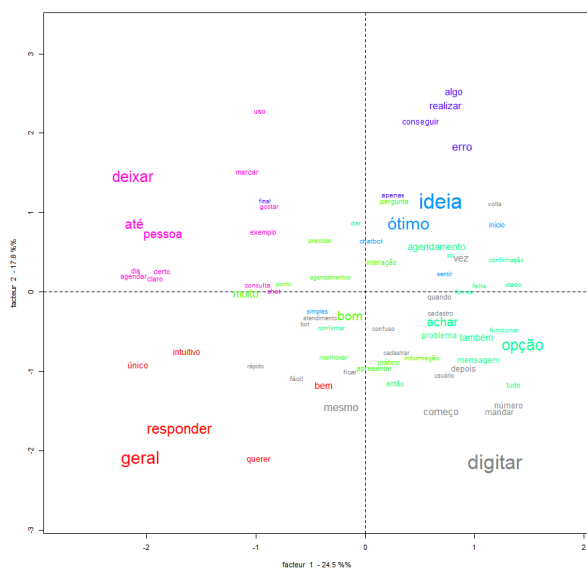


Fig. 7. Análise Fatorial de Correspondência

elemento mais recorrente no discurso. A análise de similitude reforçou essa centralidade ao posicionar *agendamento* como núcleo da rede semântica, conectado a termos que expressam tanto funcionalidade quanto experiência de uso, revelando dependências moderadas entre palavras relacionadas à clareza e à praticidade da interação. Complementarmente, a AFC mostrou a organização dessas relações em dimensões fatoriais que explicam 24,5% e 17,8% da variância, agrupando termos positivos como *ótimo*, *intuitivo* e *geral* em um quadrante, enquanto palavras como *digitar* e *responder* surgiram em posições mais distantes, indicando menor associação ao núcleo semântico predominante. Em conjunto, os três métodos convergem ao demonstrar que o chatbot foi amplamente associado à facilidade de uso e à efetividade no agendamento, embora também tenham emergido nuances relacionadas a dificuldades pontuais e à diversidade de experiências dos usuários.

VI. LIMITAÇÕES DO PROJETO

Embora os resultados obtidos na validação inicial do protótipo Liz tenham sido positivos, algumas limitações precisam ser destacadas para orientar futuras pesquisas e aprimoramentos:

- **Tamanho da amostra:** O estudo contou com apenas 58 participantes, o que restringe a generalização dos resultados para populações maiores e mais diversas.
- **Contexto controlado:** A aplicação foi realizada em ambiente específico e supervisionado, não refletindo integralmente as condições reais de uso em diferentes cenários de atendimento. Outro fato foi que as pessoas possuíam familiaridade com celulares e mensageiros, não refletindo casos de baixa escolaridade ou analfabetismo.
- **Escopo funcional:** O chatbot foi projetado para funções básicas de lembrete e agendamento. Recursos mais

avancados, como integração com sistemas de prontuário eletrônico ou suporte a múltiplos idiomas, ainda não foram implementados.

- **Dependência tecnológica:** O uso do WhatsApp como principal canal de comunicação pode limitar o alcance em grupos populacionais que não utilizam o aplicativo ou que possuem restrições de acesso à internet.
- **Avaliação longitudinal:** O estudo não contemplou análises de longo prazo sobre adesão contínua, impacto na cobertura do exame preventivo ou redução efetiva de casos de câncer do colo do útero.

Essas limitações indicam a necessidade de novos estudos com amostras ampliadas, em diferentes contextos de aplicação e com funcionalidades adicionais, de modo a fortalecer a robustez dos resultados e ampliar o potencial de impacto do protótipo Liz na saúde pública.

VII. CONSIDERAÇÕES FINAIS

O desenvolvimento e a validação do protótipo de chatbot Liz demonstraram o potencial das tecnologias de Inteligência Artificial como ferramentas de apoio à saúde pública, especialmente no contexto da prevenção do câncer do colo do útero. Os resultados obtidos evidenciaram que o sistema apresenta boa usabilidade, elevada aceitação pelos usuários e capacidade de promover uma experiência de interação envolvente, reforçando sua adequação ao cenário de aplicação.

A análise realizada por meio dos instrumentos SUS, TAM e FSS-short mostrou que o chatbot é percebido como simples de utilizar, útil para o agendamento de exames e capaz de proporcionar estados de imersão e concentração compatíveis com a teoria do *flow*. Esses achados indicam que a solução proposta pode contribuir para reduzir barreiras de comunicação entre profissionais de saúde e pacientes, além de favorecer a adesão às práticas preventivas.

Os resultados qualitativos complementam e aprofundam essas evidências. A nuvem de palavras revelou a centralidade do termo *agendamento*, indicando que o processo de marcação de consultas foi o elemento mais recorrente no discurso dos participantes. A análise de similitude reforçou essa percepção ao posicionar *agendamento* como núcleo da rede semântica, conectado a termos que expressam funcionalidade, clareza e facilidade de uso. Já a Análise Fatorial de Correspondência (AFC) evidenciou agrupamentos semânticos coerentes, com palavras associadas a avaliações positivas — como *ótimo*, *intuitivo* e *geral* — concentradas em um mesmo quadrante, enquanto termos relacionados a dificuldades pontuais, como *digitar* e *responder*, apareceram mais dispersos. Em conjunto, esses três métodos qualitativos convergem ao demonstrar que os usuários perceberam o chatbot como uma ferramenta eficaz, clara e funcional, embora tenham emergido nuances que refletem diferentes experiências individuais.

Apesar dos resultados positivos, é importante destacar que esta etapa corresponde a uma validação inicial, realizada em ambiente controlado. Estudos futuros deverão ampliar a amostra de participantes, explorar diferentes contextos de uso e avaliar o impacto do chatbot em indicadores concretos de

saúde, como aumento da cobertura do exame preventivo e redução da taxa de abandono. Além disso, melhorias contínuas na interface e nos recursos de interação poderão potencializar ainda mais a eficácia da ferramenta.

VIII. AGRADECIMENTOS

Agradecemos à Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT) pelo suporte financeiro e incentivo à pesquisa, fundamentais para a realização deste estudo. Manifestamos também nossa gratidão ao Instituto Federal de Mato Grosso do Sul (IFMS) pelo apoio institucional e pela infraestrutura disponibilizada, que contribuíram decisivamente para o desenvolvimento e a consolidação deste trabalho.

IX. DECLARAÇÃO SOBRE O USO DE INTELIGÊNCIA ARTIFICIAL

Em conformidade com o Art. 2º, inciso III, do Manual de Publicações da SBC, os autores declaram que o uso de Inteligência Artificial (IA) generativa foi realizado apenas para a revisão textual.

REFERENCES

- [1] ABU SHAWAR, Bayan; ATWELL, Eric. Chatbots: Are they Really Useful?. *Journal for Language Technology and Computational Linguistics*, v. 22, n. 1, 2007.
- [2] DE ALMEIDA, Karina Isabel Vieira. Desigualdade social e câncer do colo do útero: uma revisão sistemática. *Anais do XVI Encontro Nacional de Pesquisadores em Serviço Social*, v. 16, n. 1, 2018. Disponível em: <https://periodicos.ufes.br/abepss/article/view/23409>. Acesso em: 21 maio 2026.
- [3] ARNOLD, Donald M. et al. The design and interpretation of pilot trials in clinical research in critical care. *Critical Care Medicine*, v. 37, suppl., p. S69–S74, 2009.
- [4] BAILER, Cyntia; TOMITCH, Lêda Maria Braga; D'ELY, Raquel Carolina Souza Ferraz. O planejamento como processo dinâmico: a importância do estudo piloto para uma pesquisa experimental em linguística aplicada. *Intercâmbio*, v. 24, 2011.
- [5] BARBOSA, Simone Diniz Junqueira et al. *Interação humano-computador e experiência do usuário*. 1. ed. Rio de Janeiro: Simone Diniz Junqueira Barbosa, 2021. E-book.
- [6] BRASIL. O que é Atenção Primária?. [S. l.]: [s. d.], 2020. Disponível em: <https://aps.saude.gov.br/smp/smpoquee>.
- [7] BRASIL. Instituto Nacional do Câncer (INCA): Tipos de Câncer. Ministério da Saúde. [S. l.], 2022. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/tipos/>. Acesso em: 3 maio 2026.
- [8] BRASIL. Instituto Nacional do Câncer (Inca). Ministério da Saúde. *Estimativa 2023: incidência de câncer no Brasil*. [S. l.]: [s. d.], 2023. Disponível em: <https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/estimativa-2023.pdf>. Acesso em: 11 mar. 2026.
- [9] INSTITUTO NACIONAL DE CÂNCER (INCA). *Estimativa 2026: incidência de câncer no Brasil*. Rio de Janeiro: INCA, 2026. 168 p. ISBN 978-65-88517-52-9.
- [10] BRASIL. Resolução n.º 466, de 12 de dezembro de 2012: Dispõe sobre diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. [S. l.]: [s. d.], 2013.
- [11] BRASIL. Resolução n.º 510, de 7 de abril de 2016. Brasília: Ministério da Saúde, 2016.
- [12] CAMARGO, Brígido Vizeu; JUSTO, Ana Maria. IRAMUTEQ: um software gratuito para análise de dados textuais. *Temas em Psicologia*, Ribeirão Preto, v. 21, n. 2, p. 513–518, dez. 2013. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1413-389X2013000200016&lng=pt&nrm=iso. Acesso em: 2 set. 2026. DOI: <https://doi.org/10.9788/TP2013.2-16>.
- [13] CARVALHO, Lucas; LEIFHEIT, Bhruno Roan; RIVERO, Luis; NAKAMURA, Walter; GUERINO, Guilherme; VALLE, Pedro Henrique Dias; SOUZA, Renato; SILVA, Williamson. What Do We Know About Usability Evaluation for Chatbots?: A systematic mapping study. In: *Proceedings of the XX Brazilian Symposium on Information Systems (SBSI '24)*. Juiz de Fora, Brazil, 2024. ACM, New York, NY, USA. DOI: <https://doi.org/10.1145/3658271.3658324>.
- [14] CSIKSZENTMIHALYI, M.; CSIKSZENTMIHALYI, I. S. (org.). *Beyond boredom and anxiety: The experience of play in work and games*. San Francisco: Jossey-Bass, 1975.
- [15] CSIKSZENTMIHALYI, M. *Finding flow: The psychology of engagement with everyday life*. New York: Basic Books, 1997.
- [16] GERHARDT, Margô; MONTAGNER, Graziela. *Pesquisa em Ciências Sociais: métodos e técnicas*. Porto Alegre: Editora UFRGS, 2009.
- [17] INCA. Instituto Nacional do Câncer - INCA. O que é câncer?. [S. l.]: [s. d.], 2026. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/o-que-e-cancer/>. Acesso em: 4 jun. 2026.
- [18] INSTITUTO NACIONAL DE CÂNCER (INCA). O que é câncer?. [S. l.], 2026. Disponível em: <https://www.gov.br/inca/pt-br/assuntos/cancer/o-que-e-cancer/>. Acesso em: 4 set. 2026.
- [19] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems*. ISO 9241-210:2019. Geneva: ISO, 2019.
- [20] JACKSON, S. A.; EKLUND, R. C. Assessing flow in physical activity: The Flow State Scale-2 and Dispositional Flow Scale-2. *Journal of Sport and Exercise Psychology*, v. 24, n. 2, p. 133–150, 2002.
- [21] MARTINS, I.; LIMA, V. M. R.; AMARAL-ROSA, M. P.; MOREIRA, L. J.; RAMOS, M. G. Handcrafted and software-assisted procedures for Discursive Textual Analysis: analytical convergences or divergences?. In: COSTA, A. P.; REIS, L. P.; MOREIRA, A. (Org.). *Computer Supported Qualitative Research*. Cham: Springer Nature Switzerland, 2020. v. 1068, p. 189–205.
- [22] MARTINS, R. M. G.; ALVES, D. A.; ALVES, S. A. A.; MARTINS, F. J. G.; LIMA, L. G. A.; PEREIRA, M. L. D. Desenvolvimento e validação de chatbot para abordagem de lesões da hanseníase. *Ciência & Saúde Coletiva*, v. 31, n. 5, e01402026, 2026. DOI: <https://doi.org/10.1590/1413-81232026315.01402026>.
- [23] NAZARETH, Rodrigo Trisoglinio; DE ALMEIDA, José Júlio Gonçalves; BASTOS, Alder Thiago. Utilização do WhatsApp e o parecer CFM n.º 14/2017. *Revista da Universidade Ibirapuera*, 2020. Disponível em: <https://www.ibirapuera.br/seer/index.php/rev/article/view/222>. Acesso em: 21 maio 2026.
- [24] PICOLO, L.; TOMBINI, G. F.; DOS SANTOS, R. A. A.; DE SOUZA, C.; ENZ, L. de G. UMA Proposta de Chatbot para Apoio à Consentização e Prevenção do Câncer do Colo do Útero no Contexto do Sistema de Saúde Pública do Município de Nova Andradina. *Encontro Internacional de Gestão, Desenvolvimento e Inovação (EIGEDIN)*, v. 7, n. 1, maio 2024. Disponível em: <https://periodicos.ufms.br/index.php/EIGEDIN/article/view/20590>.
- [25] PREECE, Jennifer; ROGERS, Yvonne; SHARP, Helen. *Design de interação: além da interação humano-computador*. 3. ed. Porto Alegre: Bookman, 2013.
- [26] RAMA, C. et al. Rastreamento anterior para câncer de colo uterino em mulheres com alterações citológicas ou histológicas. *Revista de Saúde Pública*, v. 42, n. 3, p. 411–419, 2008. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102008000300004&lng=pt&tlng=pt. Acesso em: 11 mar. 2026.
- [27] RICHARDSON, Roberto J. *Pesquisa social: métodos e técnicas*. 4. ed. São Paulo: Atlas, 1999.
- [28] SILVA, S. C.; DE CICCIO, R.; VLAČIĆ, B.; ELMASHHARA, M. G. Using chatbots in e-retailing: how to mitigate perceived risk and enhance the flow experience. *International Journal of Retail & Distribution Management*, v. 51, n. 3, p. 285–305, 2023. DOI: <https://doi.org/10.1108/IJRDM-05-2022-0163>.
- [29] VALE, Diama Bhadra Andrade Peixoto do et al. Avaliação do rastreamento do câncer do colo do útero na Estratégia Saúde da Família no Município de Amparo, São Paulo, Brasil. *Cadernos de Saúde Pública*, v. 26, n. 2, p. 383–390, 2010. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-311X2010000200017&lng=pt&tlng=pt. Acesso em: 22 jun. 2026.
- [30] VENTURA, Felipe. WhatsApp chega a 99% dos celulares no Brasil; Telegram cresce. [S. l.], 2020.

Disponível em: <https://tecnoblog.net/noticias/2020/02/27/whatsapp-chega-a-99-por-cento-celulares-brasil-telegram-cresce/>. Acesso em: 4 jun. 2026.

- [31] REDAÇÃO WISHBOX. Protótipo: Entenda o Que é, Para Que Utilizar e Como Fazer. [S. l.], 2017. Disponível em: <https://www.wishbox.net.br/blog/prototipo/>. Acesso em: 4 jun. 2026.
- [32] DENECKE, K.; MAY, R. Usability Assessment of Conversational Agents in Healthcare: A Literature Review. In: *Challenges of Trustable AI and Added-Value on Health*. Studies in Health Technology and Informatics, 2022. DOI: <https://doi.org/10.3233/SHTI220431>.
- [33] LARANJO, L. et al. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, v. 25, n. 9, p. 1248–1258, 2018. DOI: <https://doi.org/10.1093/jamia/ocy072>.
- [34] MILNE, D. et al. The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review. *Journal of Medical Internet Research*, 2020. DOI: <https://doi.org/10.2196/20346>.
- [35] WORLD HEALTH ORGANIZATION. *Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*. Geneva: World Health Organization, 2021. ISBN: 978-92-4-002920-0. Disponível em: <https://www.who.int/publications/i/item/9789240029200>.
- [36] HUA, Y. et al. Standardizing and Scaffolding Health Care AI-Chatbot Evaluation: Systematic Review. *JMIR AI*, v. 4, 2025. DOI: <https://doi.org/10.2196/69006>.
- [37] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *ISO 9241-210:2019: Ergonomics of Human-System Interaction – Part 210: Human-Centred Design for Interactive Systems*. Geneva: ISO, 2019. Disponível em: <https://www.iso.org/standard/77520.html>.