Creating Chatbots to talk with humans: HCI evaluations and perspectives

Vinícius F Galvão
Universidade Federal de Mato Grosso, Instituto da Computação
Cuiabá, MT Brasil
vinifgalvao1@hotmail.com

Cristiano Maciel
Universidade Federal de Mato Grosso, Instituto da Computação
Cuiabá, MT Brasil
crismac@gmail.com

Ana Cristina Bicharra Garcia
Departamento de Informática Aplicada, UNIRIO
Rio de Janeiro, Brasil
cristina.bicharra@gmail.com

ABSTRACT

More softwares are with the power of Artificial Intelligence (AI) making their functions, simultaneously more dynamic and capable of attending each users' personal needs. Chatfuel an application that allows users to create chatbots, an AI that can answer messages autonomously. However we question it: How easy chatbot creation is, especially when dealing the complex AI operating in them? In this sense, this exploratory research seeks to understand, through our volunteers' perspectives, with different levels of computational expertise, if this creation process is really as simple as divulged. To conduct the experiments, usability methods were used such as Cognitive Walkthrough, Communicability Evaluation and Attrakdiff. With the study results, among many others, users's noticed certain difficulty while using. the tool and we problematize certain aspects connected to AI, in the context of a human-computer interaction evaluation.

CCS CONCEPTS

• Software design engineering • Empirical studies in interaction design • User studies • Multi-agent systems

KEYWORDS

Chatbots, Artificial Intelligence, Software Engineering, ChatFuel.

ACM Reference format:

Vinícius Galvão, Cristiano Maciel, and Ana Cristina Bicharra Garcia. 2019. Creating Chatbots to talk with humans: HCI evaluations and perspectives. *In Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems (IHC 2019). ACM, New York, NY, USA, 11 pages.* https://doi.org/10.1145/3357155.3358460

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
IHC'19, October, 2019, Vitoria – ES, Brazil
© 2019 Copyright held by the owner/author(s). ACM ISBN 978-1-4503-6971-8/19/10...$15.00
https://doi.org/10.1145/3357155.3358460

1 Introdução

Entre os diversos avanços tecnológicos que têm transformado o mundo, a inteligência artificial (IA) é a que mais se sobressai no presente [43]. IA, artefatos de hardware ou software capazes de realizar tarefas e interagir de forma autônoma com o mundo similarmente como os humanos fariam. Tais artefatos inteligentes liberam o ser humano de tarefas perigosas, rotineiras ou que envolvam grande volume de dados e atenção [1, 2]. Segundo a analista comercial Susan Etlinger [3], o uso de ferramentas computacionais é cada vez mais frequente, seja na organização de agenda, localização de qualquer restaurante em uma cidade, nas respostas de mensagens em um serviço de atendimento ao cliente. Estes são alguns exemplos do quanto à inteligência artificial tem tornado a vida moderna conveniente.

Quanto mais interagimos com software, mais somos influenciados por eles. Grandes empresas como *YouTube*[1], *Netflix*[2], *Amazon*[3], *Walmart*[4] e entre outras; em seus produtos e serviços, utilizam a IA para sugerir vídeos [4, 5], reter usuários [4, 5], vender produtos de acordo com o perfil do usuário [6, 7] e/ou conversar com seus clientes [6]. Vislumbrando uma oportunidade mercadológica, grandes empresas estão na "corrida do ouro da IA" em que uma gama de produtos e serviços cada vez mais da área para expandir ou melhorar os seus negócios [8]. Isso torna oportuno investigar a e avaliar a usabilidade destes softwares ditos inteligentes. Uma das formas populares do uso da IA são os "*chatbots.*" Chatbots são programas de computador especializados em dialogar em linguagem natural com usuários, como se fossem humanos. Criados por volta de 1960, incialmente como sistemas de entretenimento, e estão sendo usados em diversos domínios que precisem interagir com humanos para explicar informações, tais como na educação, comércio eletrônico e entre outros [9].

A criação de chatbots requer conhecimentos de IA. Para minimizar essa necessidade, foram criadas plataformas de

[1] YouTube - Acessado em 25/02/2019. Disponível em: <https://www.youtube.com/intl/pt-BR/yt/about/>
[2] Netflix – Acessado em 25/02/2019. Disponível em: <https://www.netflix.com/br/>
[3] Amazon – Acessado em 25/02/2019. Disponível em: <https://www.aboutamazon.com/working-at-amazon/get-to-know-us>

RIGHTSLINK

configuração de bots para facilitar tal trabalho, como por exemplo, a Chatfuel [10], considerada uma das mais populares e que será foco deste estudo. Atualmente, há várias ferramentas que prometem e permitem qualquer usuário criar o seu próprio bot[5] e usá-lo em redes sociais. Visto isso, a questão deste estudo emerge: ***o quão fácil é a criação desses chatbots, especialmente ao lidar com algo tão complexo que é a inteligência artificial***?

Com este foco, este trabalho busca compreender o processo de construção de um chatbot nessas plataformas, em especial o design da cognição do artefato, e verificar a facilidade e o quão personalizável é a tarefa de criação do bot. Esta avaliação foi realizada através dos métodos de Percurso Cognitivo [34, 35], Teste de Usabilidade [36, 37] e Attrakdiff [38, 39]. Estes métodos buscam, respectivamente, inspecionar a usabilidade da ferramenta com um especialista de tecnologia na área de IHC, avaliar a usabilidade da ferramenta com usuários de vários níveis de expertise e, finalmente, a avaliar a experiência do usuário (UX) com a ferramenta.

Este artigo está estruturado como segue, após a introdução será apresentado o *referencial teórico* em que há discussão dos principais conceitos e conhecimentos necessários para compreensão do artigo. Em seguida, é apresentada a *metodologia* da nossa avaliação, em que destacamos os procedimentos para cada método, e como cada um deles foi aplicado. Após apresentação dos *resultados e sua discussão,* e finalmente as *conclusões* o e as *referências* que tornaram este trabalho possível.

2 Referencial Teórico

Para melhor compreensão deste trabalho é necessário entender a área de inteligência artificial e o que são chatbots através de trabalhos relacionados, e a ferramenta de construção de chatbots ChatFuel [8].

2.1 Inteligência Artificial

Desde o tempo de Aristóteles, o homem procura entender o que nos faz diferente dos outros seres: a inteligência. Para entendê-la, construímos modelos matemáticos que tentam reproduzir a arquitetura do nosso cérebro, nosso processo racional de tomada de decisão ou simplesmente um comportamento que reflita nossa inteligência [12].

Há várias definições para IA. Russel & Norvig [14] apresenta IA juntamente com a busca da própria definição de inteligência humana: "*A humanidade se deu o nome científico de* ***homo sapiens*** *– homem sábio – por causa de nossas capacidades mentais tão importantes para nossas vidas diárias e o nosso sentido de ser. O campo da* ***inteligência artificial****, ou* ***IA****, tenta compreender entidades inteligentes. Portanto, uma razão para estudar é para aprender mais sobre nós mesmos. Mas diferente da filosofia e psicologia, IA busca construir entidades inteligentes e as compreender também.".* Na IA não há um paradigma único que orienta as pesquisas e o entendimento dela, pesquisadores discordam de diversas questões sobre como ela deve ser orientada. Isso faz com que a IA possua significados diferentes para pessoas diferentes [11, 12, 13, 14, 15, 16].

A divisão principal das IAs utilizadas neste estudo estão no domínio das sub-simbólicas, as quais são separadas entre **fraca e forte**. A **IA fraca** é aquela em que há um escopo pré-definido de atuação do sistema, como por exemplo, um sistema de análise de crédito ou um sistema para jogar xadrez. **A IA** forte é aquela em que não há um escopo pré-definido, o sistema consegue se adaptar a contextos diferentes ou mesmo aprender sozinhas coisas diferentes. Como Chatfuel é uma aplicação que permite a criação de bots que reagem as mensagens, *emojis e gifs* do usuário e é limitada ao roteiro de ações **"SE- ENTÃO"** em que é programada; e ainda pode se adaptar e sugerir mudanças baseado nas interações em que participa de forma similar a IA simbólica; o foco neste artigo é no uso da IA Fraca.

2.2 Chatbots

Segundo Shawar e Atwell [9], chatbots foram incialmente criados por diversão, utilizando simples técnicas em que um computador possa dar respostas baseadas em simples entradas dadas por um usuário. Inicialmente o propósito de chatbots era a de iludir usuários, mas chatbots possuem diversas aplicações como entretenimento [22, 23], busca de informações [22, 23], assistentes de e-commerce e negócios [22, 23]. Chatbots segundo [8] são interfaces mais naturais e convenientes para conhecimento de uma empresa, ou produto, ou resolução de dúvidas rápidas através de uma conversação.

Imitar um ser humano é ainda um grande desafio. A maior barreira até o momento é a compreensão da linguagem natural, a que usamos para nos comunicar com outras pessoas. Dependendo do idioma, estrutura verbal, sentimentos e ambiguidades podem ser expressos em uma variedade de forma que ainda são difíceis de uma máquina compreender integralmente [24]. Como mostrado na Figura 2, uma conversa completamente sem sentido usando o Cleverbot[6].

Figura 1: Exemplo de Diálogo do Cleverbot Cleverbot. (Esq. versão original [27], a dir. a versão traduzida).

[4] Walmart – Acessado em 25/02/2019. Disponível em: <https://corporate.walmart.com/our-story>

[5] Bot – Termo encurtado do inglês "robots" é um programa automatizado que executa tarefas na internet. Alguns bots executam de forma automática, enquanto outros executam comandos após receberem uma entrada específica. TechTerms - bot: disponível em: https://techterms.com/definition/bot

[6] <Cleverbot – Acessado em 25/02/2019 Disponível em: www.cleverbot.com/ >

RIGHTSLINK

Uma boa parte dos chatbots disponíveis no mercado ainda não é boa ou intuitiva o suficiente, exigindo uma grande carga cognitiva e compreensão de seus usuários para serem utilizados [8]. No entanto, estão se mostrando cada vez mais necessários, dado o fluxo massivo de acesso a informações que a Internet permite. Segundo [22] um dos fatores que motivam usuário a utilizarem chatbots é o aumento da produtividade em suas tarefas e certo grau de entretenimento na interação. Chatbots conseguem eficientemente ajudar usuários a terem assistência para encontrar e acessar informações específicas.

Por entretenimento, como os chatbots oferecidos pelo site botnerds.com [25], usuários podem ser menos exigentes na qualidade da interação, com o foco recreativo de ter bots tentando realizar atividades estranhas, falando gírias ou frases sem sentido [26], ou mesmo ao terem conversas absurdas [27], de forma a testá-los.

Todavia, no momento, chatbots são ferramentas que podem ser usadas para propósitos benéficos, como a resolução de questões sobre HIV/AIDS [28], no auxílio em moderação de comunidades online [30], ou no fornecimento de ajuda psicológica [29]. Como, também dubiosos como influenciar a opinião política de usuários [32] ou disseminar notícias falsas [31]. E em IHC, Følstad e Brandtzaeg apontam que chatbots são uma forma de tornar a interação com sistemas digitais mais 'humanizados' através de conversações, tornando a interação entre humanos e máquinas similar a como humanos interagem entre si. [44].

2.3 Chatfuel

Chatfuel foi fundado por Dmitry Dumik em 2015 [33], com um interesse por chatbots e querendo tornar o processo de criação o mais simples para qualquer usuário criar o seu próprio chatbot. Na época atual a aplicação alimenta 46% de todos os bots no Facebook [33].

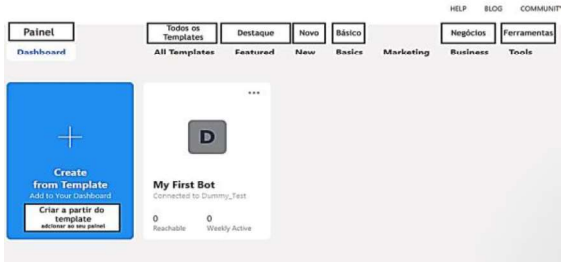

Figura 2: Tela inicial e Criação do Bot no Chatfuel.

A aplicação é online e pode ser utilizada através de navegadores de internet. Possibilitando uso em qualquer dispositivo com acesso a Internet e compatível com navegadores modernos como Google Chrome, Opera, FireFox e similares. A ferramenta atualmente apenas possui o idioma inglês para o seu uso. Ao clicar o botão de mais (+), um robô é criado para ser modificado (Figura 2). O bot pode ser configurado da seguinte forma (Figura 3), o usuário ao clicar na seção destacada pelo retângulo, pode adicionar a trilha de mensagens do dele respostas de texto em que o usuário também pode editar; imagens que o ele pode enviar; e até ícones para indicar que o robô está digitando a mensagem, recursos para manter a ilusão que, realmente esta uma pessoa ao invés do bot. Ao lado do retângulo há também macro blocos, em que o usuário pode organizar da forma que desejar.

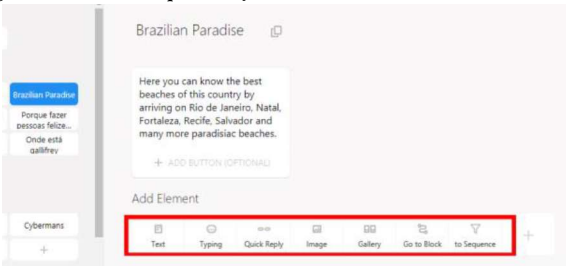

Figura 3: Tela de Configuração do Bot no Chatfuel

Em cada bloco de mensagens o usuário pode customizar quais palavras quais blocos irão reagir. Assim configurando o comportamento do bot (Figura 4). Ao ligar o bot a uma conta de grupo no facebook, o usuário pode testar o seu bot no Messenger do Facebook ou no bot ambiente de teste da plataforma ao clicar no botão "Test your Chatbot" assim verificando se o robô foi criado com sucesso (Figura 5).

Figura 4: Tela de configuração de palavras-chaves

Aplicação atualmente é gratuita para testes O usuário pode definir regras e palavras que o chatbot possa reagir corretamente e facilmente ligar a um grupo do Facebook e testá-lo. A aplicação após interações com o bot sugere mudanças para o usuário o qual pode ser automaticamente aplicadas com permissão do usuário.

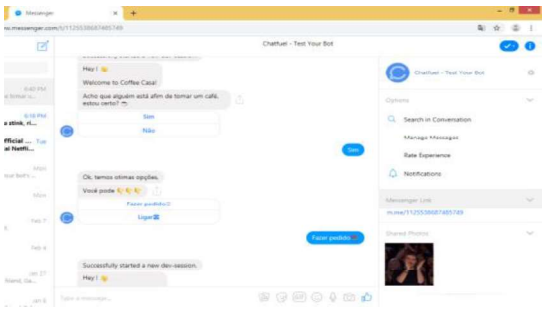

Figura 5: Tela de Teste do Chatbot

RIGHTSLINK

O motivo da escolha desta ferramenta, além do uso gratuito; devido ao seu uso recorrente em estudos para trabalhar com voluntários que possuam pouca ou nenhuma experiência com programação [53], demonstração de modelos de chatbots para propósitos particulares como serviços bancários [52], e interação emocional [54, 56] e o design amigável para criação ágil [55].

3 Metodologia

Para compreensão do funcionamento de IA e o seu processo de criação nas, se tornou necessário delimitar métodos e roteiros para inspecionar a aplicação, testar sua usabilidade e avaliar e medir a experiência do usuário após interação. Consideramos três métodos para avaliação dos chatbots:

1) Método de Inspeção: Percurso Cognitivo (PC) [34, 35].
2) Método de Observação: Teste de Usabilidade (TU). [36, 37]
3) Método de Avaliação da UX: Attrakdiff. [38, 39]

Esta seção apresenta uma breve revisão dos métodos e uma justificativa para o uso desses neste trabalho.

3.1 Sujeitos da Avaliação

Para o PC, um especialista em tecnologia com três anos de experiência em IHC, com formação em Ciência da Computação, determinou as tarefas para avaliação de usabilidade. Por meio deste, ele auxiliou na identificar das principais facilidades/dificuldades que os usuários poderiam encontrar no cumprimento de suas tarefas.

TU e Attrakdiff contaram com a participação de 6 usuários voluntários. Apesar de estar dentro do preconizado por Nielsen, o qual determina que 5 voluntários sejam um número adequado para realizar uma TU substancial para estudo qualitativo [42], com objetivo é identificar fatores evidenciados por essas avaliações, mas não ser um estudo conclusivo. Cada participante realizou o TU individualmente, em dias e horários os quais estavam disponíveis em seus locais de trabalho ou moradia. O teste teve como público alvo: adultos, maiores de 18 anos, com variada expertise tecnológica, conforme apresentado no perfil dos participantes. Para organização dos dados dos perfis dos participantes, todos preencheram um formulário on-line no qual eles informaram sua idade, gênero, quais redes sociais usam mais, com que proposto mais usam redes sociais, seus conhecimentos em tecnologia e inglês. Todos assinaram o termo de consentimento para permitir uso e análise de suas gravações durante todo processo.

Dos 6 Voluntários, 4 eram mulheres e 2 eram homens, 5 deles tem idade entre 18 e 25 anos de idade enquanto apenas 1 possui idade entre 46 e 55 anos de idade.

Sobre redes sociais: 4 deles reportaram utilizar mais Whatsapp, Facebook e YouTube respectivamente. Todos os participantes reportaram gastar entre 2 ou mais horas em redes sociais por dia, geralmente com os propósitos de entretenimento, manter contato com amigos e familiares, trabalho e obtenção de notícias. 2 voluntários administram alguma espécie de grupo em redes sociais. Todos já pensaram ou tiveram interesse em automatizar alguma função destas redes sociais.

Sobre conhecimentos em Tecnologia: 3 possuem interesse em desenvolvimento de software; e 2 já tiveram alguma experiência de desenvolvimento. 4 participantes se sentem confiantes em explicar termos avançados de computação como Algoritmos, Messenger, Scripts, Cookies e Redes Neurais.

Sobre conhecimentos em Inglês: 2 participantes se declararam não hábeis na língua inglesa. Estes participantes receberam uma "cheat sheet" (Folha de Trapaça), a qual possuía apenas os termos mais importantes traduzidos para o português do painel principal e da tela de criação do bot.

3.2 Tarefas

Para o PC, **o especialista teve as seguintes tarefas:**

- Identificar as etapas necessárias do software para a criação do bot.
- Estruturar atividades necessárias para o cumprimento das tarefas.
- Marcar possíveis etapas problemáticas durante o processo de criação dos bots. A realização de identificação destas etapas contou com as seguintes questões guias para sua realização:
    1. O usuário consegue criar bots facilmente?
    2. Um usuário que não entenda nada de T.I consegue utilizar o software?
    3. A criação de bot é intuitiva o suficiente via click and play?
    4. O usuário pode reconhecer e usar adequadamente os recursos oferecidos pelo Chatfuel apenas pelo design da aplicação?
    5. O usuário consegue configurar as regras de I.A para o bot reagir a diferentes mensagens?
    6. Houve dificuldade em ligar o bot ao grupo de testes (vulgo dummy)?
    7. Usuário se sente confortável com a linguagem, opções e dados que o chatbot coleta e emite?
    8. Quais possíveis dificuldade ou problemas que podem previstos no momento da interação?

Com os resultados do PC foi possível elaborar o roteiro da TU, juntamente com a coleta da UX de cada participante. A TU contou com as seguintes tarefas:

1. Os participantes, após preencherem o questionário de perfil e conhecimentos sobre tecnologia e redes sociais; escolheriam um tema de uma lista de sugestões, para criar um chatbot que tentará falar sobre ele. Caso eles não se sentissem confortáveis sobre o tema, eles poderiam sugerir um diferente dos pré-selecionados.
2. Eles receberam um termo de consentimento, aprovado pelo comitê de ética da universidade, para autorizar gravação de sua voz e do monitor onde a atividade estraria sendo realizada. (*A gravação do monitor foi realizada um notebook com microfone*)
3. E tentariam cumprir as seguintes tarefas:
    - *Criação do Chatbot*
    - *Instalação das Regras do Chatbot,*
    - *Adição de Funcionalidades Extras: (Galeria de Imagens, Roleta, Envio de Imagens e etc).*
    - *Ligação do bot com conta de testes no facebook para Teste.*
    - *Uso do bot na conta ligada via Messenger no Facebook.*

RIGHTSLINK

4. Eles tiveram 1 hora para realização das atividades. Depois deste tempo dariam seu feedback sobre o uso da ferramenta. Houve um ponto de checagem da atividade após 30 min para verificação do seu progresso. Participantes a qualquer momento após inicio da atividade poderiam desistir ou parar. E, ao final, receberam o formulário Attrakdiff para registro de sua experiência.

3.3 Métodos Utilizados

Nesta sessão, brevemente descreveremos os métodos neste trabalho e seus principais objetivos.

3.3.1 Percurso Cognitivo

Percurso cognitivo (PC) é um método de inspeção de usabilidade que tem o foco principal investigar o qual fácil é a aprendizagem de um sistema através de sua exploração [35]. Busca colocar o inspetor no lugar do usuário e "chutar" através do uso da aplicação onde o usuário poderão ter mais dificuldades para o uso da aplicação. O que se espera é que o percurso revele conflitos entre a concepção do designer e a concepção do usuário [36].

Alguns problemas como falta de feedback, rótulos ou palavras imprecisas e confusas ou até ícones não familiares ou de difícil associação com a tarefa, são revelados nessa etapa. Com este método foi possível elaborar o roteiro e os objetivos dos usuários voluntários para TU bem como quais possíveis etapas poderão ser mais problemáticas aos voluntários. Os resultados da avaliação por PC foram analisados de forma qualitativa, através de entrevistas para obtenção da avaliação do especialista de tecnologia e dos feedbacks relatados que os participantes durante sua experiência com a aplicação.

3.3.2 Teste de Usabilidade

Usabilidade segundo Bevan tem por objetivo descrever a qualidade da interação dos usuários com determinada interface [36]. Segundo Nielsen esta qualidade está associada aos seguintes princípios [37]:

- *Facilidade de aprendizado;*
- *Facilidade de lembrar como realizar uma tarefa após algum tempo;*
- *Rapidez no desenvolvimento de tarefas;*
- *Baixa taxa de erros;*
- *Satisfação subjetiva do usuário*

O Teste de Usabilidade (TU) é responsável por revelar como se estabelece a iteração entre o usuário e o sistema segundo Winckler e Pimenta, há características chaves a serem observados para identificação de problemas de usabilidade [45]:

- **Desempenho do usuário durante realização de tarefas**: a observação direta ou indireta da realização de tarefas por usuários permite a verificação de métricas como: ***Conclusão de tarefas***, ***Tempo de realização das tarefas, Ocorrência de erros***
- **Satisfação Subjetiva do Usuário**: Usabilidade é uma também uma qualidade subjetiva que compreende a opinião e impressões do usuário.
- **Correspondência com objetivos do usuário**: há necessidade de verificar se os objetivos e necessidades do usuário foram atingidos, independente das tarefas suportadas pelo sistema. Esta métrica pode ser quantitativa ou qualitativa, de acordo com os objetivos finais dos usuários.
- **Adequação à padrões** : Grande parte do conhecimento sobre usabilidade é organizado na formas de normas e recomendações. Pode-se verificar e identificar problemas de usabilidade inspecionando a interface considerando tais recomendações

Para realização da TU os áudios e vídeos das sessões de cada participante foram analisados qualitativamente e seus resultados comparados às observações encontradas no percurso cognitivo. Na execução das tarefas, o voluntário da TU possuía uma lista de opções temáticas para o seu chatbot e também as tarefas a serem cumpridas. Ao final, o avaliador marcou se a tarefa foi realizada com sucesso, parcialmente concluída e não concluída em uma tabela. Os resultados foram guias para uma breve entrevista não estruturada quanto ao feedback do voluntário sobre suas experiência com a ferramenta. Cada participante foi anonimizado e identificado via codinome.

3.3.3 Attrakdiff

(UX) compreende os aspectos do sistema realizados na interação com o usuário, seja sua empresa, serviços ou produtos [38]. O mesmo conceito pode ser aplicado para usabilidade, em que o sucesso de uma aplicação é perceptível de como um usuário percebe este software. E um dos métodos para revelar e medir se usuários percebem positivamente ou negativamente uma aplicação é o Attrakdiff.

Com o questionário Attrakdiff [39], é possível avaliar o quão atraente e claro um produto interativo é, oferecendo feedback para otimização e melhoria do sistema. O método não é limitado a somente softwares, os produtos que requerem interação podem ser avaliados segundo esta metodologia. O funcionamento do questionário é simples. Cada categoria possui subcategorias de elementos os quais os usuários marcam um número, em uma escala de -3 a 3, (ou de 1 a 7 em neste estudo); o que reflete melhor sua experiência com o produto. As categorias que Attrakdiff compreende são [40]:

Atratividade (AT): Indica o valor da aplicação buscando a percepção de qualidade estética que ela apresenta.

Qualidade Pragmática (Q-P): Descreve a qualidade prática de uma aplicação e indica o grau de sucesso que usuários alcançam seus objetivos.

Qualidades Hedônicas – São divididas em dois fatores:

- **Estímulo (QH-E)**: Indica até que ponto a aplicação e seu design podem desenvolver e avançar seu uso em termos de originalidade, interesse e estímulo.
- **Identidade (QH-I):** Indica até que ponto a aplicação permite o usuário possa se identificar com ela.

A análise deste questionário se deu por métodos quantitativos. Os participantes, após realização da TU, logo receberam o formulário Attrakdiff via digital para responderem.

RIGHTSLINK

4 Resultados

A realização dos três métodos de avaliação foi bem sucedida. Para melhor leitura e organização, subseções foram organizadas para cada método aplicado.

4.1 Resultados do Percurso Cognitivo

Durante o PC, o foco foi na Tarefa de criação e utilização de um Chatbot. 5 atividades foram identificadas pelo especialista com seus respectivos passos, como descrito na *Tabela 1.*

ATIVIDADES	EXECUÇÃO
1º - Criação do Novo bot	**1** – Clicando o ícone de "mais" (+) e texto azul escrito criar "template" **2** – Através da página de templates o usuário pode selecionar um bot pré-pronto e alterar conforme o necessário.
2º - Instalação das Regras do Bot	**1** –Após os usuários criarem o bot eles são automaticamente mandados para sua página de edição. **2** – Os usuários podem clicar no botão de "mais" (+) ao lado destas caixas para adicionar mais blocos de mensagem. **3 –** Após escrever as mensagens, usuários devem clicar no botão "Set up A.I". E definir quais palavras-chaves irão ativar os respectivos blocos. **4** – Repetir o processo descrito nos passos 2 e 3 até o usuário estiver satisfeito com seu bot.
3º - Adição de Funcionalidades Extras (como galeria, roleta, envio de imagens, vídeos e etc.)	**1 –** Ao final de cada mensagem há um menu horizontal com funcionalidades como roletas, galerias, envio de imagens e entre outras. **2 –** A quais podem ser enviadas no fluxo de mensagens de cada bloco **3** – Iguais ás mensagens, as imagens podem ser ligadas a determinadas palavras que o usuário desejar.
4º - Ligação do Bot a conta Dummy no Facebook para teste.	**1 –** Ao final do processo de criação do bot nas atividades 2 e 3, os usuários podem escolher uma das contas de teste para ligar o seu bot. **2 –** Caso não tenha um grupo o usuário pode criar um grupo no Facebook e conectar automaticamente
5º - Uso e teste do bot via Messenger do Facebook.	**1 –** Ao ter um bot criado e conectado, os usuários podem testar o chatbot criado. **2 –** Tanto enviando para o grupo uma mensagem via Messenger do Facebook ou Clicando o botão "Test This Bot"

Tabela 1: Atividades e os Passos para Respectiva Execução Identificados na PC.

Durante passo a passo de cada atividade identificado pela PC foram também elaboradas perguntas e respostas para cada uma.

P1: Todos os usuários identificarão imediatamente quais opções selecionar para customização do seu bot?

R1: Não, a barreira de linguística, especialmente para aqueles que não têm domínio da língua inglesa pode ser dificultar muito o uso da aplicação.

P2: Um usuário que não conheça nada de T.I pode fazer uso da aplicação?

R2: Sim, a aplicação por ser uma "espécie" de linguagem de programação visual contida em seu software permite qualquer usuário realizar ações em seus chatbots sem necessariamente dominar uma linguagem de programação formal.

P3: Os símbolos no design da aplicação ajudam a identificar as funções da aplicação?

R3: Sim e não. Os símbolos utilizados para indicar Configuração, Divulgação, e Chat ao vivo; por exemplo, são relativamente simples de identificar independente se o usuário entende inglês. O símbolo da *Configuração* é uma engrenagem, *Divulgação* um megafone, e *Chat ao vivo* um balão de mensagens No entanto os símbolos das funções principais do processo de criação não são tão claros O símbolo da automação é um pequeno muro enquanto da I.A. é uma estilização de um cérebro humano. Certas funções não são imediatamente identificadas nos ícones.

P4: Há alguma ajuda que o próprio software possa oferecer para usuários?

R4: Sim. A documentação e tutorias da aplicação estão no canto inferior direito na página, indicado com um grande círculo com ponto de interrogação branco e azul, os usuários podem acessar tutoriais textuais e vídeos indicando o passo a passo da aplicação.

P5: Usuários poderão criar mais mensagens de forma simples?

R5: Sim. No começo de cada bloco vazio há um menu horizontal, indicado pelo retângulo vermelho na figura 5 o qual exibe as ações que podem ser instaladas em seu bot neste bloco. Ao clicar na opção "Text" uma nova caixa de mensagens é adicionada.

P6: Usuários poderão editar as regras de I.A de forma simples?

R6: Sim, com dificuldades. Cada bloco de I.A. é automaticamente criado a cada novo bloco de mensagens. Usuários podem colocar termos longos ou curtos. Toda String de mensagens que tiver espaços apenas terminará quando o usuário apertar "Enter".

P7: Há riscos de um robô ou sua I.A não funcionar?

R7: Sim, se o usuário não configurar adequadamente quais palavras o bot deve reagir, ele não irá seguir a sequencia de seu bloco de mensagens.

Com a PC, foi identificado que os maiores obstáculos para o uso da aplicação são o idioma, o usuário que não compreenda inglês terá dificuldade de compreender a interface de usuário do software. Estes fatores associados, a falta de feedback ativo da aplicação, de símbolos não claros, e não haver descrições rápidas quando a funções delas, pode tornar o uso em um longo processo tentativa e erro. Outro fato notável é que não há configurações de acessibilidade para o uso da aplicação, certos símbolos ou textos e áreas o que pode dificultar usuários com necessidades especiais fazerem uso da aplicação.

RIGHTSLINK

4.2 Resultados do Teste de Usabilidade

A T*abela 2* apresenta os participantes, devidamente anonimizados através de seus respectivos codinomes, os quais serão utilizados para se referir a cada participante ao longo da divulgação dos resultados, e seu desempenho em cada passo da tarefa.

Codinome	Atv 1	Atv 2	Atv 3	Atv 4	Atv 5	Nível de Experiência
Colina Real	S	S	CP	ÑC	S	Alto
Koala	S	ÑC	ÑC	CP	ÑC	Baixo
Terna	S	ÑC	ÑC	S	ÑC	Baixo
Whoviana	S	CP	S	S	CP	Alto
Chapéu de Palha	S	S	S	S	S	Médio
Vulpe Rubra	S	CP	S	S	CP	Médio

Tabela 2: Resultado do Teste de Usabilidade de Cada Voluntário.

Sucesso (S) significa que o voluntário realizar a atividade sem problemas e o seu bot realizou exatamente o que instruiu. **Conlusão Parcial (CP),** o voluntário conseguiu realizar a atividade em termos gerais, no entanto, não conseguiu que o bot se comportasse como instruiu. **Não Conclusão (ÑC),** voluntário não conseguiu iniciar o processo de realização da atividade e/ou não identificar na aplicação onde e realizar a tarefa. O nível de experiência de cada participante relativo a T.I foi considerado. Voluntários com nível **Alto** já tiveram experiência no desenvolvimento de software ou uso de ferramentas avançadas de T.I; **Médio** se o mesmo fosse familiar com ferramentas de T.I, no entanto nunca utilizou/atuou na área; e **Baixo** se o voluntário possui nenhuma experiência ou familiaridade com T.I.

É notável que na primeira tarefa todos os participantes conseguiram criar o bot, como esperado do percurso cognitivo. No entanto nas *atividades 2, 3, 4* e *5*; os resultados variaram para cada participante. A seguir as falas de cada participante foram transcritas como eles as disseram e assim permitindo melhor compreensão das impressões de cada um deles durante a avaliação.

Para *Colina Real*, ao mesmo tempo em que ele achou fácil utilizar a aplicação, sentiu falta de feedback quanto ao uso das funcionalidades extras, a I.A reagiu de acordo com os termos chaves que definiu e em geral um bot funcional foi criado. Em suas palavras, "*Gostei bastante de usar a aplicação, mas senti que o design das funcionalidades, ou até as respostas que o sistema me dá ao colocar uma regra de I.A. não são claros o suficiente (...) Isso pode ser mal ao longo do processo, mas até onde um software intuitivo vai para a criação de um chatbot, essa é uma boa ferramenta.*"

Koala teve problemas similares a *Colina Real.* Mesmo utilizando os tutoriais e vídeos que o próprio software disponibiliza não se sentiu confiante que cumpriu as tarefas adequadamente, citando que: "*O processo escolha das palavras para I.A. reagir, e ligar os blocos de mensagens, aí depois não vejo funcionar, foi frustrante. Eu acho que fiz tudo certo mas nada (do sistema) me disse que faltou ou fiz algo de errado! Eu consigo ler o que cada coisa significa mas eu não entendo que ele (o ChatFuel) está pedindo de mim.*"

Terna sentiu que a barreira linguística foi o maior obstáculo no uso da ferramenta, dizendo que: "*Eu vejo os símbolos, eu tenho uma ideia do que está acontecendo, mas nada parece muito claro para mim, acho que por que não sei inglês, não entendo nada que está acontecendo, mesmo por tentativa e erro, clicando e ver o que está acontecendo nada me chamou a atenção para interagir sabe? (...) Fez a experiência de aprendizado ser mais confusa, mesmo com os tutoriais.*".

Whoviana sentiu que, por já trabalhar na área de T.I., o uso a ferramenta foi relativamente intuitiva, quando comparado com ferramentas as quais está acostumada. Seu bot conseguiu exibir e reagir aos termos específicos durante, no entanto não conseguiu fazer a ligação com o grupo de teste de forma adequada para a I.A do bot poder evoluir. Segundo participante: "*Através de clicar e ver o que cada função trás, foi ficando cada vez mais fácil entender o que aplicação pede para funcionar, mas ainda acho que ela não é tão acessível assim. Ainda falta muito feedback para um usuário comum perceber o que está acontecendo, eu sinto que se tivesse mensagens enquanto o cursor do mouse estivesse por cima ia fazer esse processo (*de aprendizado*), mais fácil.*".

Chapéu de Palha sente que, apesar de não ter domínio completo do inglês, foi relativamente fácil usar a ferramenta, e que após ter seguido os tutoriais, teve uma curva rápida de aprendizado, tendo sucesso na criação e ligação no bot ao grupo de teste, permitindo o a I.A oferecer a criação de blocos de mensagens para palavras que recebeu durante interação e não possuía resposta adequada no Messenger. Segundo participante: "*Eu teria mais dificuldade se só for pelo design, e sinto que não saber inglês foi difícil de poder usar a ferramenta, então tive que estudar um pouco mais e entender o significado dessas palavras na interface e tal. Não saber muito inglês foi uma barreira, mas que dá para passar, e se você possui conhecimentos de computação (...). Se fosse uma pessoa com menos domino, ela provavelmente não saberia o que esta acontecendo.*"

Vulpe Rubra sentiu que usar a ferramenta foi relativamente fácil devido aos seus conhecimentos em inglês, no entanto a difícil compreensão do software e suas funções levaram a uma experiência de uso confusa e frustrante: "*São muitas funções que a aplicação oferece, eu sinto que há a falta de um tutorialzinho depois que cria um robô sabe? Algo bem para "leigo" mesmo, pois eu senti que o processo normal para criar o bot é muito complicado, algo passo a passo assim é importante. (...) Eu senti que não sabia me localizar onde estava no processo ou o que deveria fazer depois. (...) Muita gente não gosta de ser tratado igual um "leigo", mas um tutorial assim passo a passo ia poupar muito tempo, e hoje em dia não tenho tempo para ler manuais ou ás vezes ver vídeos tutoriais e preciso que as coisas funcionem rápido.*" Outro detalhe interessante durante análise dos resultados da TU é como voluntários com nível de experiência **alto** e **médio** T.I. puderam cumprir a maioria das tarefas, enquanto os de **baixo** nível apenas conseguiram criar o bot.

4.3 Resultados do Attrakdiff

Observando o comportamento de cada ponto, no setor Atratividade (**AT),** na Figura 6, percebe-se que, na visão dos usuários, a aplicação é visualmente *Sedutora, Boa e Agradável* de usar. Enquanto ao mesmo tempo é *Rejeitadora, Feia e Pouco*

RIGHTSLINK

Amigável. Em **AT**, os voluntários demostram que a ferramenta em termos de design estético não é atraente, imediatamente amigável e agradável para o usuário.

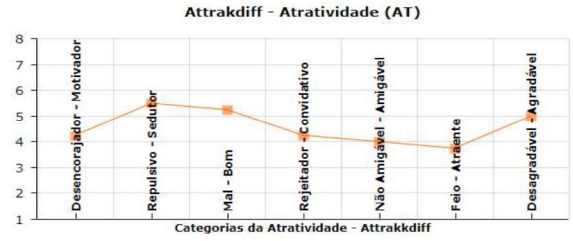

Figura 6: Gráfico resultante do Attrakdiff, setor AT

Em **QH-E,** na figura 7, é notável pelo desenho do ziguezague que os usuários têm opiniões diferentes quanto o seu uso. Usuários concordam ainda que o uso da ferramenta ainda é ligeiramente *Desafiador, Inovador e Criativo*; permitindo usuários criarem seus robôs de varias formas. No entanto, os processos para tornarem um chatbot funcional ainda são *Entediantes, Convencionais e Cautelosos*. Estas caraterísticas refletem que as funcionalidades do Chatfuel, durante o seu uso, os usuários não se sintam estimulados a criar chatbots criativos, inovadores e completamente expressivos da forma que desejem através das funcionalidades oferecidas neste software.

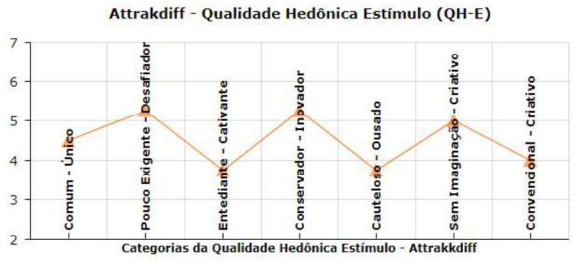

Figura 7: Gráfico resultante do Attrakdiff, setor QH-E.

Enquanto **QH-I,** na figura 8, os participantes realmente sentiram que as possibilidades da ferramenta, como uma extensão do usuário na administração de um grupo do Facebook, respondendo dúvidas, cuidando de vendas, informando sobre o grupo ou produto, são características que os participantes consideraram valiosas. Tanto que as qualidades que o produto reflete, no poder *Aproximar Pessoas,* de ser um software *Apresentável* e*, Integrador* para outros usuários e empresas; e finalmente promover uma experiência que é, na opinião deles, *Superior* a outras aplicações técnicas que a participantes experimentaram tiveram altas pontuações, indicando que os participantes, ao lidarem com softwares de criação similares, terão este uso como uma referência positiva.

Eles também concordam que a experiência de criação do bot, não é **Profissional, Elegante** e até **Isoladora** pelo qual o indivíduo tem que "se virar" para aprender o seu uso, ou descobrir o que houve de errado. Algo que corresponde as falas dos participantes sobre a falta de feedback claro sobre os erros ou acertos do participante, o que é, principalmente, destacado nas falas de *Vulpe Rubra, Koala* e *Terna.*

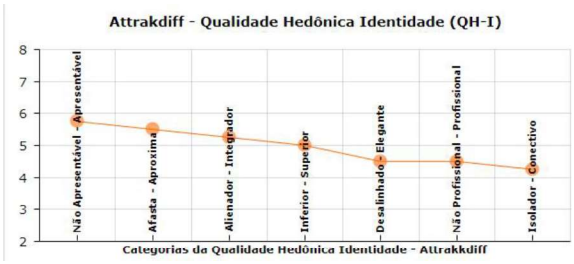

Figura 8: Gráfico resultante do Attrakdiff, setor QH-I.

E finalmente **QP** teve a média mais baixa. Um pouco surpreendente, pois intuitividade implica em facilidade de uso. Os usuários sentiram que a aplicação tem se conteúdo e opções que são *Gerenciáveis e Apresentáveis*. Mas que seu uso ou funcionalidades para o chatbot ainda são *Confusas, não tão Práticas, Complicados* e *Técnicos* demais para o usuário comum. Tanto que os participantes que sentiram mais a vontade para usar a aplicação como *Colina Real, Whovian* e *Chapéu de Palha* tinham mais familiaridade com a T.I. tiveram mais sucesso comparado aos outros.

Figura 9: Gráfico resultante do Attrakdiff, setor QH-P.

Em termos gerais, segundo a UX dos usuários, ChatFuel é adequada para sua proposta de ter uma criação de bots fácil e acessível para diversos tipos de usuários, no entanto não é intuitiva o suficiente para todos os tipos de usuários.

4.4 Discussões e Reflexões

Via PC, foi possível observar que, teoricamente, o processo de criação de um chatbot via ChatFuel é simplificado o máximo possível para usuários, sem conhecimentos aprofundados de T.I, fazerem uso. Enquanto na PC foi possível detectar possíveis conflitos que os usuários enfrentariam, os quais em parte, foram confirmados durante TU. Apesar dos problemas encontrados pelos participantes, a maioria conseguiu criar seus bots e ter uma UX relativamente satisfatória. Após aplicação dos métodos e análise dos resultados 4 insights importantes para o design de softwares em relação a I,A.

1) Intuitividade é um conjunto de fatores em software, como interface, funcionalidades, feedback e entre outros similares, em harmonia. Segundo Kreitzberg [46]: "*Uma interface intuitiva é uma que funciona do jeito que o usuário espera. Em outras palavras, achamos que algo é intuitiva quando não temos que pensar sobre ela.* (tradução livre)". O papel da intuitividade em software para a interação com os sistemas é chave para tornar aplicações mais compreensíveis para qualquer pessoa [47, 48, 49]. Chatfuel e sua IA permite compreender "pontos cegos" no

RIGHTSLINK

bot. Interação que o bot participou, e que não teve respostas adequadas e também não foram consideradas pelo usuário. No entanto a IA, neste tipo de software, não é um fator critico para sua intuitividade. O design de sua interface e feedback visual são mais importantes para guiar o usuário durante o processo de criação afim de evitar frustrações que foram evidentes em todos os nossos voluntários.

2) Feedback é importante. Pois é a forma do software "falar" com o usuário durante o uso. Como demonstrado em nossa TU, a falta deste componente conduziu uma experiência de uso frustrante para os participantes como esperado por Winkler e Pimenta [45], Nielsen [37] e Bevan [36]. A falta deste elemento foi particularmente evidenciada por *Koala, Terna e Vulpe Rubra*, pois não conseguiram fazer um bom uso da ferramenta já que não sabiam o que ela estava necessitando ou em que etapa do processo de criação o participante se localizava.

3) O processo de criação da I.A é bem diferente da sua utilização. A I.A do Chatfuel apenas reage enviando mensagens confome as respectivas palavras-chaves definidas pelo usuário durante o processo de criação. Apesar do uso da aplicação ter sido limitado apenas uma hora por usuário, a I.A é capaz de identificar termos os quais o usuário, durante sua criação, não havia considerado, e sugerir alterações para melhorar interações futuras. Esta funcionalidade permite auxiliares usuários a compreender e trabalhar junto a I.A para criação de um bot satisfatório. O uso de uma IA no Chatfuel, para criar uma IA no bot torna este processo mais simples do que o usual. Apesar do software não ser completamente intuitivo, o processo de criação auxiliado pelo design possui forte intuitividade como descrita por Kreitzberg [46].

4) A barreira linguística dificulta o uso do software para alguns participantes. Há evidências como nos estudos de Bajas e Meyers [51] e Ko et al [50] que conforme softwares se tornam cada vez mais utilizados internacionalmente, o sucesso de sua utilização deve também considerar superação desta barreira. A IA pode complicar este processo. Especialmente como Chatfuel é em inglês e usa mecanismos dependentes do idioma para sugerir mudanças e melhorias do bot para o usuário. O uso de símbolos pode permitir uma compreensão mais universal do feedback da ferramenta com usuários. A falta deste elemento como discutido por *Chapéu de Palha, Colina Real e Whoviana,* não impede totalmente o uso do software, no entanto torna o uso da ferramenta mais vagarosa e complexa do que realmente é.

Com o avançar da tecnologia, o design de softwares que comuniquem eficientemente suas funções será cada vez mais necessários. Nem sempre usuários poderão demandar tempo para aprender toda aplicação ou realizar cursos de informática com focos específicos. O sucesso do software atualmente depende, principalmente, do quão eficientemente ele consegue comunicar com o usuário.

Para o sucesso no uso do software, é necessário considerar formas mais claras de como apenas o design pode "falar" muito para qualquer usuário, sem necessariamente explicitar isto através de longos menus, manuais e vídeos. Um conceito sugerido é que softwares, similares a bons filmes, adotem a filosofia "Mostrar, não falar" [41]. Ao mostrar as funcionalidades com clareza e objetividade, permite usuários cumprirem seus objetivos de forma rápida e satisfatória. Características que conduzem ao sucesso de uma aplicação, tirando vantagem de imagens e textos claros para "narrativa" de seu uso ser memorável e usuários se tornarem especialistas do uso de um software em particular ferramenta e suas atividades.

Estudos sobre a avaliação de chatbots tendem a focar mais em seu design e performance, como através de um estudo semiótico [58]; do quão natural é conversação que eles promovem [57]; os papeis que eles podem realizar em design de serviços [60, 52]; e a experiência emocional que eles promovem para usuários [56, 54]. Enquanto não é possível direta comparação destes estudos com as experiências de nossos voluntários, é possível notar que há estratégias semióticas, fora do escopo deste estudo, que tornou possível a criação de chatbots com voluntários com baixo nível de experiência. Porém a configuração e customização dos bots, através da experiência dos usuários necessita de clareza, mesmo usuários com níveis de experiência alto, sentiram falta de feedback claro da ferramenta, assim não conseguiram aproveitar tanto a experiência de criação quanto as possibilidades criativas que a ferramenta permite ou e também se localizarem melhor no processo para realizar o melhor uso possível da ferramenta.

Estes resultados devem ser considerados dentro das limitações deste estudo, sendo realizado com voluntários pequeno, sem separação ou grupo de controle, homogeneidade, experimento com tempo de apenas 1 hora, tornam estes resultados circunstanciais, limitados e de difícil generalização. Bem como também, apesar dos chatbots terem sido criados pelos participantes, a falta de engajamento com eles online não permitiu um estudo mais aprofundado de sua IA.

Embora muito interessantes os resultados realizados neste estudo é necessário replica-lo com um número maior de participantes, um grupo de controle para comparar melhor as diferenças dos participantes com experiência dos que não, e as diversos outros softwares criadores de chatbots no mercado como Botmock[7], FlowXO[8], Botsify[9] e entre outros similares.

5 Conclusões

Após análise dos resultados desta pesquisa, percebe-se que é possível tornar a criação de um artefato complexo como um chatbot em uma tarefa para usuários, sem necessitar de profundos conhecimentos sobre I.A. No entanto, é discutível se o software em questão, ChatFuel é fácil de ser utilizado por qualquer usuário.

Apesar das limitações em que o estudo foi realizado, com participantes com um perfil relativamente homogêneo, pouco tempo para uso da ferramenta em cada teste, as barreiras linguísticas da aplicação, as sutilezas do funcionamento da IA da aplicação às quais não foram evidenciadas em todas as TU's e entre outras similares.

Softwares, especialmente quanto ao seu design, precisam ser desenvolvidos e avaliados em quesitos que vão além de sua técnica e eficiência computacional. A necessidade de preocupação com os símbolos, linguagem e organização visual utilizada, são necessários para promover a comunicação e uso claro do usuário com a ferramenta. E quando há uso da IA como uma funcionalidade é necessária cautela aos detalhes, pois o papel que desempenha nem sempre é óbvio ou ativo, mas é

[7] Botmock. Disponível em: https://botmock.com/
[8] FlowXO, Disponível em: https://flowxo.com/
[9] Botsify, Disponível em: https://botsify.com/

RIGHTSLINK

significante em entender como softwares se tornam mais dinâmicos ao auxiliar usuários atingirem seus objetivos de uso.

Nas perspectivas dos voluntários, é notável que mesmo que a tecnologia esteja cada vez mais presente no dia-a-dia, nem todos os usuários terão conhecimentos específicos de computação como I.A. Tais conceitos são cada vez mais necessários na construção de software mais eficientes e capazes de atender as necessidades dos usuários. Ao mesmo tempo isso, exige ainda cautela dos desenvolvedores tornarem seus softwares intuitivos, através do uso do software, como este conceito é utilizado em seu produto. Tal ensinamento deve fazer parte de um software bem desenhado.

Visto isso, futuros estudos também são necessários na compreensão de como desenvolvimento de tais aplicações priorizam o uso da IA, e com o papel social que softwares desempenham nos dispositivos e nas vidas diárias dos usuários deve ser investigado como a IHC pode melhorar e orientar o desenvolvimento destes softwares para atenderem necessidades humanas. Também considerar, futuramente, replicar o estudo considerando um grupo de controle para comparar as diferenças de habilidade entre usuários iniciantes vs experientes, as diferenças de avaliação de usabilidade em uma gama maior de ferramentas, bem como uma avaliação mais das IAs de cada bot criado, considerando após maior tempo de engajamento com as ferramentas, não só no tempo de criação pelos usuários, mas também com os seus bots criados. Para assim tornar este processo de criação tão simples e compreensível quanto possível.

Com a crescente interação com Chatbots através de diversos de meios tecnológicos, compreender o seu processo de criação e funcionamento, permite insights de como os construir de maneira melhor e mais compreensiva as necessidades do usuário.

AGRADECIMENTOS

Agradecemos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de Iniciação Científica executada via Pró-reitoria de Pesquisa (PROPeq) da UFMT; a esta instituição de ensino pelo apoio parcial que viabilizou a participação no evento e aos participantes do grupo focal pelas ricas discussões.

REFERÊNCIAS

[1]. S. L. dos Santos & F.G.Teixeira. (2010). Design de uma interface de interação tridimensional com foco na usabilidade e no desempenho gráfico. *In: Design & tecnologia [recurso eletrônico]*. Vol. 1, n. 1, p. 39-50. Porto Alegre, RS, Brazil.

[2] N. Statt (2018) - The AI boom is happening all over the world, and it's accelerating quickly. *The Verge 12 Dez 2018.* Disponível em: <https://www.theverge.com/2018/12/12/18136929/artificial-intelligence-ai-index-report-2018-machine-learning-global-progress-research > - Acesso em 25 de fev. 2019.

[3]. S. Etlinger. (2017) The Age of AI--How Artificial Intelligence is Transforming Organizations. Altimeter, São Francisco, USA,

[4] D. Meyer. (2018) YouTube's AI is so good at finding offensive content it needs more staff to keep up. *World Economic Forum, 25 Abr 2018.* Disponível em: <https://www.weforum.org/agenda/2018/04/ai-is-now-youtube-s-biggest-weapon-against-the-spread-of-offensive-videos/> - Acesso em 25 de fev. 2019.

[5]. M. Gavira. (2019) Here's How Netflix Uses AI and Data to Conquer the World. *VibeProjects the power of purpose.* Disponível em: https://vibeprojects.com/heres-how-netflix-uses-ai-and-data-to-conquer-the-world/ - Acesso em 25 de fev. 2019.

[6] B. Marr. (2018) Machine Learning In Practice: How Does Amazon's Alexa Really Work? *Forbes, 5 Out 2018.* Disponível em: <https://www.forbes.com/sites/bernardmarr/2018/10/05/how-does-amazons-alexa-really-work/#422cc34e1937> - Acesso em 25 de fev. 2019.

[7] B. Marr. (2017). How Walmart is Using Machine Learning AI, IoT and Big Data to Boost Retail Performance. *Forbes, 29 Ago 2017.* Disponível em: <https://www.forbes.com/sites/bernardmarr/2017/08/29/how-walmart-is-using-machine-learning-ai-iot-and-big-data-to-boost-retail-performance/#3aee77fa6cb1> - Acesso em 25 de fev. 2019.

[8] A. J. Ask, M. Facemire & A. Hogan. (2016) The State of Chatbots-Pilot Chatbots as Part of your App + Mobile Strategy. *Forrester Research.* Disponível em: <https://www.forrester.com/report/The+State+Of+Chatbots/-/E-RES136207> Acessado em: 5 Mar de 2019.

[9] A. B. Shawar & E. Atwell. (2007). Chatbots: are they really useful?. *In: Ldv forum. p. 29-49.*

[10] CHATFUEL. *Site Oficial da Empresa* Disponível em:: <https://chatfuel.com/about-us.html> - Acessado em 25 de Fev de 2019.

[11] J. F. Teixeira. (2014). Inteligência artificial. *Pia Sociedade de São Paul*, Editora Paulus.

[12] S. Russell & P. Norvig. (2004). Inteligência Artificial. Ed. Campus. Rio de Janeiro, 2004.

[13] L. M. Pereira. (1988) Inteligência Artificial: mito e ciência. *In Revista Colóquio-Ciências*, v. 3, p. 1-13.

[14] S. Russell & P. Norvig. (2016) Artificial intelligence: a modern approach. Pearson Education Limited, Malaysia.

[15] J. Gordon, E. Edward & H. Shortliffe. (1984) H. Rule-Based Expert Systems. Chapter The Dempster–Shafer Theory of Evidence.

[16] E. Castillo, J. M. Gutierrez, M. Jose & A.S. Hadi. (2012). Expert systems and probabilistic network models. Springer Science & Business Media.

[17] R. R. Yager & L. A. Zadeh (Ed.). (2012). An introduction to fuzzy logic applications in intelligent systems. Springer Science & Business Media.

[18] F. Zambonelli, R. N. Jennings & M. Wooldridg. (2003) Developing multiagent systems: The Gaia methodology. *In ACM Transactions on Software Engineering and Methodology (TOSEM)*, v. 12, n. 3, p. 317-370.

[19] Y. Shoham & K. Leyton-Brown. (2008) Multiagent systems: Algorithmic, game-theoretic, and logical foundations. Cambridge University Press.

[20] J. Kolodner. (2014) Case-based reasoning. Morgan Kaufmann.

[21] Â. G. B. Lima. (2003) Interfaces entre a ciência da informação e a ciência cognitiva. *In Ciência da Informação*, v. 32, n. 1.

[22] P. B. Brandtzaeg. & A. Følstad. (2017) Why people use chatbots. *In: International Conference on Internet Science.* Springer, Cham. p. 377-392.

[23] C. F. De C. Júnior et al. (2018). Chatbot: uma visão geral sobre aplicações inteligentes. *Revista Sítio Novo*, v. 2, n. 2, p. 68-84.

[24] Natural Language Processing: Why The Machines Still Struggle To Understand Us? – *Intellectsoft, 3 Dez 2015.* Disponível em: <https://www.intellectsoft.net/blog/natural-language-processing-why-the-machines-still-struggle-to-understand-us/>, Acessado em: 25 de fev de 2019.

[25] Types of Bots: Na Overview. *Botnerds.* Disponível em: <http://botnerds.com/types-of-bots/> Acessado em 25 de fev de 2015.

[26] S. R. Fussell et al. (2008). How people anthropomorphize robots. *In: 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE*, p. 145-152.

[27] How to have great sex with cleverbot? *Yahoo! Answers* Disponível em: <https://answers.yahoo.com/question/index?qid=20111014235033AAqeJOY> Acessado em: 25 de fev 2015.

[28] J. Brixey et al. (2017). Shihbot: A facebook chatbot for sexual health information on hiv/aids. *In: Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue.*. p. 370-373. Pathfinding Discourses of Self in Social Network Sites Ariza Geogalou.

[29] H. S. de A. Souza. (2017) A robô Eurekka: desenvolvendo um chatbot de psicoeducação para o messenger do Facebook. *In Automating power: Social bot interference in global politics.* Samuel C. Woolley.

[30] R. D. Vukovic & I. M. Dujlovic. (2016). Facebook messenger bots and their application for business. *In: 2016 24th Telecommunications Forum (TELFOR). IEEE. p. 1-4.*

[31] Y. A. De Lima, A. C Chares & J. S. De Oliveira. (2018) Identificação de fake news: uma abordagem utilizando métodos de busca e chatbots. I*n: 7º Brazilian Workshop on Social Network Analysis and Mining (BraSNAM 2018).* SBC.

[32] M. A. Ruediger et al. (2017) Robôs, redes sociais e política no Brasil: estudo sobre interferências ilegítimas no debate público na web, riscos à democracia e processo eleitoral de 2018.

[33] Crunch Base | ChatFuel Disponível: <https://www.crunchbase.com/organization/chatfuel> Acessado em: 25 de fev de 2019.

[34] P. Polson, C. Lewis, J. Rieman & C. Wharton. (1992). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. *In: International Journal of ManMachine Studies*, 36 (5). p. 741-773.

[35] C. Wharton, J. Rieman, C. Lewis & P. Polson. (1994). The cognitive walkthrough method: A practitioner's guide. *In Nielsen, J., and Mack, R. L. (Eds.), Usability Inspection Methods.* John Wiley & Sons. 105-140. New York, USA.

[36] Bevan, N. (1995). Usability is quality of use. In *Advances in Human Factors/Ergonomics* (Vol. 20, pp. 349-354). Elsevier.

RIGHTSLINK

[37] -Nielsen, J. (1994). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). ACM. Boston. Massachusetts, USA.

[38] D. Norman & J. Nielsen. (2016). Avaliando a Experiência do Usuário e a Usabilidade de um Aplicativo Web Móvel: um relato de experiência. *In The definition of user experience (UX)*. Nielsen Norman Group Publication, v. 1.

[39] AttrakDiff. *Site Oficial*. Disponível em: http://attrakdiff.de/index-en.html. Acessado em 25 de Fev de 2019

[40] N. M. C. Valentim, W. Silva & T. Conte. (2015). Avaliando a Experiência do Usuário e a Usabilidade de um Aplicativo Web Móvel: um relato de experiência. *In: XVIII Congresso Ibero-Americano em Engenharia de Software, 2015, Congresso Ibero-Americano em Engenharia de Software (CIbSE 2015). 2015.* p. 788-801. Lima. Peru.

[41] A. Macaulay. (2003) Don't Tell Me, Show Me: Directors Talk about Acting. Currency Press. Sydney.

[42] J. Nielsen. (2012). How Many Test Users in a Usability Study? *Nielsen Norman Group*, 04 de Junho de 2012. Tópico: User Testing. Fremont, Califórnia, USA. Disponível em: <https://www.nngroup.com/articles/how-many-test-users/> Acessado em: em 25 de Fev de 2019.

[43] J. Vincent. (2019). The State of AI in 2019 *The Verge*, 28 de Janeiro de 2019. Manhattan, New York USA. Disponível em: <https://www.nngroup.com/articles/how-many-test-users/> Acessado em: em 03 de Jun de 2019. P. B. Brandtzaeg. & A. Følstad. (2017) Chatbots and the New World of HCI. *In: interactions,* v. 24, n. 4, p. 38-42. *DOI: https://doi.org/10.1145/3085558*

-Winckler, M., & Pimenta, M. S. (2002). Avaliação de usabilidade de sites web. *In: ESCOLA REGIONAL de Informática*. SBC, p.1-54, Porto Alegre:Brazil.

[46] Kreitzberg C. (2017). The Intuitive Interface. User Experience Office. *Princenton University.* 3 Fev 2017. Disponível em: <https://ux.princeton.edu/designing-experience/intuitive-interface> Acessado em: 1 de Maio de 2019.

[47] Gasparini, I., Kimura, M. H., & Pimenta, M. S. (2013, October). Visualizando 15 anos de IHC. *In Proceedings of the 12th Brazilian Symposium on Human Factors in Computing Systems* (pp. 238-247). Brazilian Computer Society. Manaus. Brazil.

[48] Pereira, R. (2013). IHC na Especialização em Engenharia de Software: discussões sobre uma experiência prática. In *WEIHC - Workshop de Ensino de IHC*. (pp. 1-6) Manaus, Brazil. .

[49] Ferreira, A. L. S., de Miranda, L. C., & de Miranda, E. E. C. (2012). Interfaces cérebro-computador de sistemas interativos: estado da arte e desafios de IHC. In *Proceedings of the 11th Brazilian Symposium on Human Factors in Computing Systems* (pp. 239-248). Brazilian Computer Society. Cuiabá. Brazil

[50] Ko, A. J., Myers, B. A., & Aung, H. H. (2004, September). Six learning barriers in end-user programming systems. In *2004 IEEE Symposium on Visual Languages-Human Centric Computing* (pp. 199-206). IEEE. Rome. Italy

[51] Barajas, M., & Owen, M. (2000). Implementing virtual learning environments: Looking for holistic approach. *Journal of Educational Technology & Society*, v.*3, n*(3) p39-53. .[52] Teixeira, F. A. A. (2008). Chatbot para serviços bancários. Faculdade de Ciências, Universidade do Porto. Cidade do Porto, Portugal.

[53] ValÊrio, F. A. M., Guimar, T. G., Prates, R. O., & Candello, H. (2018). Chatbots Explain Themselves: Designers' Strategies for Conveying Chatbot Features to Users. *SBC Journal on Interactive Systems*, *9*(3), 61-79.

[54] Smestad, T. L. (2018). *Personality Matters! Improving The User Experience of Chatbot Interfaces-Personality provides a stable pattern to guide the design and behaviour of conversational agents* (Master's thesis, NTNU).

[55] Razaque, A., & Yang, G. (2018, November). Chatbot: Efficient and Utility-Based Platform. In *Intelligent Computing: Proceedings of the 2018 Computing Conference* (Vol. 1, p. 109). Springer.

[56] Duijst, D. (2017). *Can we improve the User Experience of Chatbots with Personalisation* (Doctoral dissertation, M. Sc Thesis, University of Amsterdam).

[57] Hung, V., Elvir, M., Gonzalez, A., & DeMara, R. (2009, October). Towards a method for evaluating naturalness in conversational dialog systems. In *2009 IEEE International Conference on Systems, Man and Cybernetics* (pp. 1236-1241). IEEE.

[58] Valério, F. A., Guimarães, T. G., Prates, R. O., & Candello, H. (2017, October). Here's What I Can Do: Chatbots' Strategies to Convey Their Features to Users. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems* (p. 28). ACM.

[59] Jain, M., Kumar, P., Kota, R., & Patel, S. N. (2018, June). Evaluating and informing the design of chatbots. In *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 895-906). ACM.

[60] Jylkäs, T., Aijälä, M., Vuorikari, T., & Rajab, V. (2018). AI Assistants as Non-human Actors in Service Design.

RIGHTSLINK