

# Legenda Automática do YouTube: essa inteligência artificial atende aos usuários?

Victor Félix Arinos<sup>1</sup>, Karina Furlan<sup>1</sup>, Cristiano Maciel<sup>1</sup>, Milene Selbach Silveira<sup>2</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)  
Av. Fernando Corrêa da Costa, nº 2367 – Bairro Boa Esperança – Cuiabá – MT – Brasil

<sup>2</sup>Pontifícia Universidade Católica do Rio Grande do Sul - Escola Politécnica  
Av. Ipiranga, 6681, Porto Alegre – RS - Brasil.

{crismac, victorfelixar, kaka.furlan.kf}@gmail.com,  
milene.silveira@pucrs.br

***Abstract** The popularization of social media in recent years has increased the number and range of content found on the internet, increasing the dissemination of information. Among these social media, YouTube stands out for the diversity of topics and languages used in its content. Thus, this paper aims to analyze the usability and user experience related to YouTube's content creation tool for automatic captioning through the application of the Cognitive Walkthrough, the AttrakDiff and a Field Study. The user experience was rated and corrections were proposed to the problems found. Furthermore, the results and methods applied were compared, seeking to evaluate their effectiveness and the relationship of user satisfaction and the problems found.*

***Resumo** A popularização das mídias sociais nos últimos anos aumentou o número e a variação dos conteúdos encontrados na internet, facilitando a disseminação. Dentre estas mídias, o YouTube se destaca pela diversidade de temas e linguagens no seu conteúdo. Dessa forma, este artigo tem como objetivo analisar a usabilidade e user experience da ferramenta de Criação de Conteúdo do YouTube para legendas automáticas através da aplicação de um Percurso Cognitivo, AttrakDiff e Estudo de Campo, avaliando as experiências dos usuários e propondo correções para os problemas. Além disso, os resultados e métodos são comparados, avaliando a eficácia dos mesmos e a relação da satisfação dos usuários com as dificuldades.*

## 1. Introdução

Nos últimos anos, as mídias sociais vêm auxiliando no processo de criação, divulgação e interação de conteúdo entre diversas pessoas, gerando novas oportunidades e um novo público alvo de consumidores. Estima-se que 2,2 bilhões de usuário acessem pelo menos 1 vez ao mês o Facebook, sendo que diariamente existem pelo menos 1,45 bilhão de acessos de usuários [Facebook]. Somente no Brasil, existem aproximadamente 129 milhões de usuários ativos no Facebook, estando atrás somente dos Estados Unidos e da Índia. Outro foco dos brasileiros tem sido a plataforma de streaming de vídeos YouTube, com um aumento de aproximadamente 58% de usuários nos últimos anos. Quando colocado em números, pode-se estimar que existam 98 milhões de usuários na plataforma [Costa 2018][Custódio 2019].

Este aumento no número de usuários buscando serviços de streaming de vídeo gera uma grande demanda em inovações tecnológicas nas funcionalidades oferecidas pela plataforma, desde novos algoritmos para indicação de vídeos aos usuários até ferramentas utilizando Inteligência Artificial (IA), como o processo de tradução e legenda automática. No caso das legendas automáticas a plataforma disponibiliza um estúdio de criação de conteúdo. Contudo, devido à alta complexidade envolvida no processo de criação automática de legendas, que vai desde o conhecimento do estúdio de criação de conteúdo até o reconhecimento automático de voz, a qualidade do produto gerado pode ser afetada, assim como a experiência do usuário e a sua satisfação em utilizar o site de reprodução de vídeos online e suas ferramentas. Assim, questiona-se a forma como os sentimentos gerados pela utilização da ferramenta podem afetar o usuário e a sua satisfação com o produto gerado na plataforma. Neste sentido, o presente trabalho apresenta o estudo da usabilidade da interface proposta pela plataforma gerenciada pela Google, além de analisar os impactos que a utilização da ferramenta pode causar nos usuários. Para análise da usabilidade foi aplicado o método de Percorso Cognitivo e realizado um Estudo de Campo, e, para análise da experiência de usuário foi aplicado o AttrakDiff.

O artigo está estruturado em cinco seções, sendo a segunda seção dedicada ao referencial teórico, que abrange questões relacionadas à Usabilidade, Experiência do Usuário e Inteligência Artificial; a terceira, discorrendo sobre a metodologia aplicada, sendo os resultados obtidos apresentados na quarta seção. Por último, tem-se as considerações finais e os encaminhamentos para trabalho futuros na quinta e última seção, seguida da listagem das referências utilizadas para elaboração do trabalho.

## **2. REFERENCIAL TEÓRICO**

Para a International Standards Organization (ISO) [ISO 2019], usabilidade pode ser definida com uma medida no qual um sistema de computador habilita o usuário, em um dado contexto de uso, a alcançar objetivos específicos com eficácia, gerando concomitantemente sentimentos de felicidade/satisfação. Assim, podemos caracterizar a usabilidade como uma medida capaz de definir o grau de satisfação de um usuário ao utilizar um produto ou sistema, por exemplo.

Assim como a usabilidade interfere no sucesso de uma determinada tarefa sendo executada em um software, as experiências do usuário afetam o resultado de acordo com o que ocorre durante a realização da atividade. Neste sentido, as experiências hedônicas dos usuários, geradas juntamente com o prazer ao suprir as expectativas e necessidades do usuário [Hassenzahl et al 2015], são avaliadas por meio de conceitos da User Experience, identificando as emoções, como felicidade, frustração, interesse, satisfação, etc. [Júnior e Kronbauer 2018]. Estas experiências, quando avaliadas, podem reduzir as emoções negativas em relação ao objeto de estudo, facilitando a adaptação dos usuários ao software, assim como o desejo de utilizar o produto [Júnior e Kronbauer 2018]. Entretanto, métricas de avaliação de usabilidade de inteligências artificiais (IAs) e suas interfaces estão em constante desenvolvimento [Botega, Silva e Murphy 2018][Pyae e Scifleet 2019], sendo definidas de acordo com o escopo estudado, em especial pelas características únicas das IAs, capazes de oferecer suporte ou auxiliar na tomada de decisões, diminuindo os esforços humanos e tempo necessário para determinados processos [Giarrantano e Riley 2005].

Considerando-se a avaliação heurística [Gutwin e Greenber 2000], por exemplo, as heurísticas utilizadas na avaliação podem ser adaptadas conforme o objeto de estudo, devido ao fato de não terem um foco específico ou contexto de uso [Meira et al 2011]. A aplicação destas heurísticas está intrinsicamente ligada com o objeto de estudo. Rusu et al. (2011) anunciaram 16 heurísticas específicas para mundos virtuais, capazes de encontrar problemas com mais facilidade em seu contexto de uso do que as dez heurísticas de contexto geral de Nielsen [Meira et al 2011], demonstrando a eficácia da adaptação dos parâmetros de acordo com o ambiente de aplicação.

Especificamente em relação a interface de sistemas com IA considerando-se o escopo deste trabalho, encontram-se alguns estudos interessantes. Botega, Silva e Murphy (2018), por exemplo, estudaram a interface de um sistema inteligente capaz de oferecer suporte à criatividade com base em respostas fornecidas pelos usuários para as suas perguntas, revelando os aspectos a serem adaptados nas IAs para obtenção de melhores resultados. Já Pyae e Scifleet (2019) investigaram a eficácia do dispositivo da Google Home Smart Speaker comparando a capacidade de reconhecimento de voz com usuários cuja língua nativa é o inglês em relação a outros usuários, demonstrando que o domínio da língua é um fator influente para melhores resultados com uma IA. Strings de busca criadas por voz para uso em sites de pesquisa também são utilizados para estudos com reconhecimento de voz. Shokouhi, Ozertem e Craswell (2016) utilizaram os termos de pesquisa inseridos por áudio como dados de treinamento para reconhecimento de voz, demonstrando que o novo conjunto de dados, extraído dos termos de busca dos usuários, possui influência no processo de reconhecimento de voz, diminuindo os erros.

### **3. METODOLOGIA**

Nesta seção serão apresentados a ferramenta analisada, os métodos aplicados, os roteiros de testes seguidos.

#### **3.1. Criação de legendas automáticas no YouTube**

A plataforma de vídeos YouTube é um meio de propagação de ideais, conhecimentos, informações, notícias, entre outros vários temas que são oferecidos pelo público e para o público, dando a possibilidade de liberdade de expressão e direitos à informação e a oportunidades [YouTube 2019]. Para isso, o YouTube oferece diversas ferramentas que ajudam cada usuário a realizar sua intenção, podendo ver um vídeo aleatório, ou tutoria através de conteúdos educacionais. Dentre as ferramentas de geração de conteúdo que o YouTube possui, existe a opção de legendas automáticas, que são geradas por um processo automático através do reconhecimento de sons. Em 2012, esse recurso ganhou um novo avanço, aumentando seu número de idiomas aceitos para a funcionalidade, incluindo a língua portuguesa [Junqueira 2012].

As legendas são geradas com base no idioma selecionado para o vídeo, sendo que o tempo de processamento pode variar de acordo com a complexidade do áudio. Além disso, conforme o tamanho do vídeo aumenta, ou seja, quanto maior for o tempo de duração do vídeo, mais demorado será o processo de geração da legenda automática, podendo variar de horas para dias.

### 3.2. Métodos aplicados

Para possibilitar a análise pretendida, foi aplicado um conjunto de métodos para avaliar o processo de criação de vídeos e utilização da legenda automática proposta pela plataforma, assim como para avaliar a qualidade do produto gerado (vídeo legendado) com base nas palavras acertadas pela IA na legenda. O Percurso Cognitivo [Lewis e Wharton 1997][Bonifácio, Viana e Vieira 2010] foi escolhido para a inspeção da usabilidade da ferramenta, buscando identificar problemas de usabilidade, assim como avaliar a facilidade com que o usuário irá aprender a utilizar o objeto de estudo. Além do Percurso Cognitivo, a usabilidade foi analisada por meio de um estudo de campo, no qual o usuário é observado em seu local de trabalho [Holzinger 2005]. Durante este processo, as ações do usuário foram anotadas de forma imperceptível.

Por fim, foi aplicado o método AttrakDiff, como meio de avaliar a experiência do usuário [Júnior e Kronbauer 2018], por meio de um questionário com três categorias e quatro seções, divididas em 28 pares de palavras (antagônicas). O Percurso Cognitivo foi realizado por 3 especialistas; o estudo de campo foi realizado com 8 usuários voluntários que também responderam ao questionário de experiência do usuário. Além disso, a avaliação do vídeo legendado foi realizada pelos usuários do estudo de campo através da atribuição de uma nota de 0 até 10.

### 3.3 Roteiros de Teste

Tanto no Percurso Cognitivo quanto no Estudo de Campo, os usuários tinham o mesmo objetivo: realizar uma transmissão ao vivo, solicitando, ao final da transmissão, a legendagem automática do vídeo Para o Estudo de Campo (Tabela 1), a ferramenta foi analisada através de seu uso em condições naturais com usuários Universitários. Por estarem habituados com o ambiente da Universidade, tanto para ver, criar e legendar vídeos, os testes foram realizados neste ambiente, variando somente de sala para a escolha de maior conforto dos participantes. Durante todo o processo dois avaliadores estavam presentes no recinto em que ocorriam os testes sem intervir no processo.

**Tabela 1. Passos do Roteiro de Avaliação - Estudo de Campo.**

Sequência	Ação
Passo 1	Logar na sua conta do YouTube.
Passo 2	Transmitir um vídeo ao vivo com a configuração de privacidade como público.
Passo 3	Dizer em voz alta e durante a transmissão diversas frases, caso queira pode utilizar o termo de consentimento como as frases para este passo.
Passo 4	Finalizar a transmissão.
Passo 5	Colocar a legenda do seu vídeo transmitido como geração automática.
Passo 6	Procurar o seu vídeo através do mecanismo de busca da plataforma e tente localiza-lo.
Passo 7	Assistir ao vídeo legendado.

Para o Percurso Cognitivo, os especialistas seguiram um roteiro minimalista (Tabela 2), contendo informação de como prosseguir através da plataforma de streaming de vídeo e do estúdio de criação de conteúdo do YouTube, buscando identificar e

analisar as dificuldades e os tipos de dificuldades que o usuário, em uma situação típica, enfrentaria ao utilizar a ferramenta, além da capacidade da IA do YouTube Legendas no processo de legendar vídeos.

**Tabela 2. Passos do Roteiro de Inspeção - Percurso Cognitivo.**

<b>Sequência</b>	<b>Ação</b>
Passo 1	Logar na sua conta do YouTube.
Passo 2	Transmitir um vídeo ao vivo com a configuração de privacidade como público.
Passo 2.1	Localizar no canto superior direito o botão com uma câmera, cujo tooltip é “Criar um vídeo ou uma postagem” e selecionar Transmitir ao vivo.
Fluxo alternativo	Caso a conta não esteja verificada, siga a opção primeiros passos para habilitar a opção de transmissão de vídeos ao vivo.
Passo 2.2	Adicionar um título para o vídeo e deixe a opção de privacidade como pública. Clique em próximo.
Passo 2.3	Clicar em transmitir ao vivo.
Passo 3	Dizer em voz alta e durante a transmissão diversas frases, caso queira pode utilizar o termo de consentimento como frases para este passo.
Passo 4	Finalize a transmissão.
Passo 4.1	Localizar o botão vermelho de finalizar transmissão no centro inferior da tela.

Antes da realização dos testes os usuários leram e assinaram o termo de consentimento, contendo as informações sobre o teste e assegurando a privacidade e segurança dos dados, assim como o direito de desistir do processo a qualquer momento caso julgassem necessário. Os tempos gastos na realização dos testes foram registrados, assim como o tempo aproximado gasto para as legendas automáticas estarem disponíveis para uso. Além disso, ao final do teste, uma nota em escala de 0 até 10 foi solicitada aos usuários após a aplicação do AttrakDiff, para definir a qualidade da legenda gerada de acordo com o áudio gravado.

### **3.5. Perfil dos Participantes**

Para participar do Percurso Cognitivo e do Estudo de Campo foram escolhidos 3 especialistas e 8 voluntários. Os especialistas possuem experiências diferentes na área de usabilidade, variando entre 1 a 2 anos de experiência e atuação na área. Os usuários U1, U2, U3 e U4 do Estudo de Campo são alunos de cursos de Computação da Universidade. Os usuários U6 e U7 são alunos de pós-graduação de outras áreas e U5 e U8 analistas de sistemas. Os voluntários do Estudo de Campo foram selecionados com buscando a maior variedade de perfis possível como forma de aumentar a qualidade dos dados a serem obtidos.

## **4. RESULTADOS**

Nas seções a seguir serão os resultados obtidos com a aplicação de cada método.

#### 4.1. Percurso Cognitivo

O Percurso Cognitivo foi realizado com especialistas de diferentes níveis de conhecimento, variando entre 1 a 2 anos de atuação na área de usabilidade, como forma de garantir a extração de dados diferentes. Os teste foram realizados por especialistas com características e experiências diferentes, variando sua especialidade entre Usabilidade (Esp1 e Esp2), Acessibilidade (Esp2) e Privacidade (Esp3). Os participantes realizaram o teste individualmente, seguindo os objetivos especificados no roteiro e avaliando a interface conforme avançavam nos objetivos. Todos os problemas, erros, falhas ou confusões causadas pela interface foram reportadas aos avaliadores ao final do teste. Ao final dos testes do Percurso Cognitivo os avaliadores, autores deste estudo, analisaram e sugeriram soluções para cada problema como forma de eliminar ou minimizar os possíveis impactos destes erros nos usuários. Os problemas encontrados foram organizados na Tabela 3, listando a identificação dos erros na primeira coluna, os problemas na segunda coluna e as possíveis correções para os problemas na coluna final.

**Tabela 3. Tabela com problemas e possíveis correções.**

Erro	Problema	Sugestão de correção
E1	“O botão ao vivo não representa a possibilidade de transmitir vídeos ao vivo”	Alterar o nome do botão para algo mais significativo: “Listar vídeos ao vivo”
E2	“O YouTube não demonstra na ferramenta que existe a aba Publicadas para listar as legendas automáticas”	Manter a aba de legendas “Publicadas” com uma mensagem informativa, como: “Em processamento”
E4	“Não existe informação a respeito do tempo de espera para a geração da legenda”	Adicionar uma imagem informativa, com as etapas do processamento durante o seu andamento. Poderia conter a porcentagem já feita.
E5	“Precisa ter o celular por perto para utilizar a transmissão”	Utilizar o e-mail como passo para verificação.
E8	“Alguns comandos são repetitivos. Ao clicar em editar no estúdio estou querendo editar o vídeo atual. Não há motivo para ir para a listagem de vídeos transmitidos anteriormente”	Alterar o encaminhamento ao clicar no botão “Editar no estúdio”.

Para o primeiro especialista (ESP1) a realizar a Inspeção, o processo, apesar de simples, pode ser frustrante para o usuário, devido ao grande número de informações omitidas pela plataforma durante o processamento da linguagem, gerando impactos como confusão (E1 e E2), mal entendimento por duplo sentido (E3) e impaciência (E4). Em outro comentário temos: “Não existe nenhuma informação a respeito do início do processo da legenda automática”.

Quanto a qualidade da legenda oferecida pela plataforma através da IA, ESP1 definiu como ótimo o resultado, tendo significado lógico e contextual as frases geradas. Para ele a legenda gerada foi avaliada como 7, devido a problemas na transcrição de algumas palavras. Para o segundo especialista (ESP2), a legenda foi de péssima qualidade, com nota 1, devido à falta de sentido na legenda gerada pela IA. Contudo, o

terceiro especialista (ESP3) avaliou como ótima a performance do gerador de legenda automática, atribuindo nota 9, com apenas 3 erros durante todo o áudio transmitido.

## 4.2. Estudo de Campo

O Estudo de Campo ocorreu igualmente com 8 usuários em ambientes e momentos separados. O estudo tinha como condição inicial ter uma conta no YouTube e estar verificada, contudo, nenhum usuário possuía estas condições satisfeitas no momento de início do teste. Durante os testes foram anotados o tempo de execução de cada tarefa juntamente com as dificuldades, além da nota atribuída para a legenda gerada. Os problemas foram divididos pelas seguintes classificações: 3 categorias (Pragmático, Hedônico e Atrativo) e 9 seções (Alto impacto, Médio impacto e Baixo impacto). A tabela 4 e 5 encontram-se os problemas Pragmáticos e Hedônicos, sendo os atrativos discutidos ao final. Nos erros classificados como pragmáticos, o usuário teria dificuldade para alcançar o seu objetivo, os hedônicos geram sentimento negativos no usuário, podendo alterar a nota do produto final, e os atrativos interferem na qualidade final. Os seguintes fatores foram levados em conta para avaliar os impactos:

- **Alto impacto** – pode influenciar no resultado e na nota atribuída pelo usuário.
- **Médio impacto** – não influencia no resultado, mas pode influenciar na nota atribuída pelo usuário.
- **Baixo impacto** – não influencia no resultado e tem pouca influência na nota final.

A Tabela 4 elenca os problemas pragmáticos, demonstrando os obstáculos encontrados durante a tentativa de transmitir e legendar os vídeos. Alguns comentários durante os testes deixaram claro a dificuldade em gerar as legendas automáticas, como para o Usuário 8 (U8), que disse que seria necessário um guia na plataforma para entender como legendar automaticamente: “Não existe nenhuma informação de que está legendando... Eu teria que pesquisar no Google para saber”. Outro comentário, dessa vez dito pelo Usuário 7 (U7), afirma que o processo de legendar não é intuitivo: “Só após transmitir é possível ver a opção de legendar o vídeo, em momento algum antes ele me diz que tem essa opção”.

**Tabela 4. Problemas pragmáticos.**

Tipo	Problemas	
Pragmático – Alto Impacto	“Estou tendo dificuldades para legendar o vídeo.”	“Não deixa claro que existe a opção de legenda automática. Não existe aviso e nem pergunta se deseja ela”.
	“O processo não é intuitivo. Não sei aonde eu peço a legenda.”	“Deveria informar como legendar. Não tem informação, nem no início nem no fim”.
Pragmático – Baixo Impacto	“A tela de criação de conteúdo tem muita informação, algumas até repetidas”	“O símbolo sonoro não é fiel a realidade. Estou falando, sai som na transmissão e ele não muda”.

A Tabela 7 apresenta os problemas Hedônicos encontrados na plataforma. Todos os problemas localizados possuem, no mínimo, médio impacto, afetando a qualidade do resultado e gerando sentimentos negativos, sendo eles confusão e frustração. Para U6 o processo não apresenta nenhum tipo de retorno para o usuário, como demonstrado no seguinte comentário: “Não fala nada na tela de transmissão ao vivo, nem no estúdio de criação, nem na reprodução do vídeo. Eu preciso saber se ele está legendando ou fico caçando feito um burro”.

**Tabela 5. Problemas Hedônicos**

<b>Tipo</b>	<b>Problemas</b>	
Hedônico – Alto Impacto	Frustração – “Porque só posso transmitir um vídeo após 24 horas?”.	Frustração – “A legenda não foi gerada na linguagem selecionada. Está em espanhol e selecionei português”.
	Frustração – “O sistema não tem informações de como colocar a legenda automática. Não deveria ter uma opção de legendas automáticas no botão Adicionar Legendas?”.	Frustração – “Não posso transmitir algo imediatamente. Qual a necessidade de esperar 24 horas se eu já confirmei por mensagem?”
Hedônico – Médio Impacto	Confusão – “Não deixa claro quando está ocorrendo a legenda do vídeo”.	Confusão – “Como ele legendou se eu não pedi para ele legendar automaticamente?”.
	Confusão – “A ideia de transcrever e sincronizar automaticamente da a impressão de gerar a legenda automática, mas não é essa a sua função”.	Confusão – “Achei que eu tinha que apertar um botão para ele gerar a legenda automática”.

Os erros atrativos percebidos na plataforma, em especial no resultado, afetaram a qualidade da legenda, como no comentário de U3 “Não está de acordo com o que foi ditto no video, apesar de possuir algum sentido”. Em outro caso, U2 afirma “A legenda não está de acordo com o idioma do video. Elas estão em espanhol, sendo que falei e selecionei português”. Alguns outros casos, em relação ao idioma, foram reportados durante os testes, como traduções em idiomas diferentes do selecionado pelo usuário, assim como traduções para o inglês dos botões da interface.

### 4.3. Discussão

Um dos grandes problemas identificados e que foi recorrente durante todos os testes está na característica de verificação da conta na plataforma de streaming de vídeos. Todos os participantes não possuíam uma conta verificada, resultando numa espera de 24 horas antes da realização do teste. Podemos perceber que, apesar de interessante, a escolha da plataforma em verificar a conta pode limitar e frustrar o usuário. Além disso, observou-se que grande parte das dificuldades foram oriundas da falta de sentido dos botões, podendo ter significado duplo dependendo do contexto (como o principal exemplo que foi o Ao Vivo e o Transcrever e Sincronizar automaticamente). O Estudo de Campo comprovou os problemas apontados na aplicação do Percurso Cognitivo, demonstrando casos de confusão e frustração durante o uso da plataforma.

Além disso, os usuários que experimentaram sentimentos negativos, decorrente de erros ou pela espera da legenda, durante os testes conseguiram localizar mais erros durante o procedimento, resultando em notas baixas, além das reações negativas.



## 5. CONSIDERAÇÕES FINAIS

As mídias sociais deixaram de ser somente um meio de comunicação a distância, ocupando diferentes papéis na atualidade. Neste cenário, tratar questões como a legenda automática se tornam mais necessárias devido ao árduo processo de transcrição de áudio, sendo preciso estudar e analisar o quão útil para o usuário é o serviço oferecido.

A aplicação dos diferentes métodos permitiu a identificação de problemas e decorrentes da interface do YouTube, principais objetivos dos métodos aplicados. Considerando a análise de usabilidade, podemos notar que ambas as técnicas possuem eficácia quanto a avaliação da interface, complementando as lacunas que possam aparecer em cada uma delas. O Estudo de Campo revelou mais problemas conforme os usuários realizavam o teste, enquanto o Percorso Cognitivo apresentou resultados satisfatórios com os diversos especialistas devido às suas diferentes experiências na área, sendo que somente no último especialista ocorreram alguns problemas repetidos. Além disso, as experiências de usuário demonstraram que grande parte dos problemas afetam o usuário causando o sentimento de frustração, em especial devido à falta de informação sobre o processo de legenda da plataforma.

Por fim, podemos perceber que as legendas geradas pelo YouTube estão atreladas a qualidade do áudio da transmissão, complexidade do conteúdo e o tema do vídeo (caso seja um vídeo mais usual do YouTube as legendas serão geradas com maior facilidade). Para muitos usuários a ferramenta de legenda automática poderia ser melhorada já que não possui informação quanto ao estado da legenda (se está sendo realizada ou se já está finalizada). Além disso, as legendas criadas receberam, em média, nota 5, demonstrando uma satisfação mediana quanto ao seu desempenho. Apesar de possuir palavras erradas, alguns usuários demonstraram satisfação, afirmando que a ferramenta, apesar dos erros, satisfaz as necessidades básicas dos usuários.

Apesar dos diferentes métodos de avaliação aplicados, algumas características, como a percepção do usuário sobre a IA do YouTube e seu funcionamento, além da forma como ela pode ser influenciada pelas ações do utilizador não foram perceptíveis. Este fato nos faz questionar a adequação dos testes em relação a avaliação de interface e a Inteligência Artificial utilizada na geração de legendas automáticas. Mudanças no método, como testes em etapas com repetições, para avaliar a evolução da IA no processo, poderiam gerar novos dados relacionados com reconhecimento de voz e legenda automática, assim como com a satisfação e experiência do usuário. Além disso, como trabalho futuro pode ser, também, avaliada a forma como a qualidade das legendas é afetada pela dicção dos usuários, assim como pelas variações linguísticas.

## 6. Referências

- Holzinger, A. (2005) “Usability engineering methods for software developers”, In: Communications of the ACM, p. 71-74.
- Pyae, A. e Scifleet, P. (2019) “Investigating the Role of User's English Language Proficiency in Using a Voice User Interface: A Case of Google Home Smart Speaker”, In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, p. LBW0282.
- Bonifácio, B., Viana, D., Vieira S., Araújo C. e Conte, T. (2010), “Aplicando técnicas de inspeção de usabilidade para avaliar aplicações”, In: Proceedings of

- the IX Symposium of Human Factors in Computing Systems, 189-192.
- Gutwin, C. e Greenberg, S. (2000) “The mechanics of collaboration: Developing low cost usability evaluation methods for shared workspaces”, In: Proceedings IEEE 9th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, p. 98-103.
- Lewis, C. e Wharton, C. (1997), “Cognitive walkthroughs”, In: M. Helander, T. K. Landauer, & P. Prabhu (Eds.), Handbook of human-computer interaction, p. 717-732.
- Custódio, M. (2019), “Conheça as 10 redes sociais mais usados no Brasil”.  
<https://resultadosdigitais.com.br/blog/redes-sociais-mais-usadas/>.
- Rusu, C., Roncagliolo, S., Rusu, V. e Collazos, C. (2011), A Methodology to Establish Usability Heuristics, In: ACHI 2011: The Fourth International Conference on Advances in Computer-Human Interactions, p. 59-62.
- Facebook. (2018) “Facebook Q1 2018 Results”,  
[https://s21.q4cdn.com/399680738/files/doc\\_financials/2018/Q1/Q1-2018-Earnings-Presentation-\(1\).pdf](https://s21.q4cdn.com/399680738/files/doc_financials/2018/Q1/Q1-2018-Earnings-Presentation-(1).pdf).
- International Organization for Standardization. (2019). ISO 9241-210:2019,  
<https://www.iso.org/standard/52075.html>
- Júnior, J. B. da S. e Kronbauer, A. H. (2018). “A Study of Hedonic Experience Related to UX Capture Techniques”, In: Proceedings of the 17th Brazilian Symposium on Human Factors in Computing Systems, p. 4.
- Girrantano, J. e Riley, C. (2005), Expert Systems: Principles and Programming. 4. Thomson Course Technology.
- Botega, L. F. de C., Silva, J. C. e Murphy, G. R. (2018). “Usability study on the interface of an artificial intelligence system for creativity support”, In: Human Factors in Design, p. 041-060.
- Hassenzahl, M., Wiklund\_Engblom , A., Bengs, A., Hägglund , S. e Diefenbach , S. (2015), “Experience-oriented and product-oriented evaluation: psychological need fulfillment, positive affect, and product perception”. In: International journal of human-computer interaction, p. 530-544.
- Shokouhi, M., Ozertem, U. e Craswell, N. (2016). “Did you say u2 or youtube?: Inferring implicit transcripts from voice search logs”, In: Proceedings of the 25th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, p. 1215-1224.
- Costa, T. (2018) “Quais as redes sociais mais usadas?”.  
<https://rockcontent.com/blog/redes-sociais-mais-usadas-no-brasil/>, Agosto.
- Meira, S. R. D. L., Costa, R. A., Jucá, P. M. e Silva, E. D. (2011). “Redes sociais, In: Sistemas Colaborativos”, p. 53-64.
- Junqueira, D. (2012) “YouTube agora coloca legendas automáticas nos vídeos em português”, <https://gizmodo.uol.com.br/legendas-automaticas-youtube-portugues/>, Agosto.
- Youtube. <https://www.youtube.com/intl/pt-BR/yt/about/>, Agosto.