

18º ERGODESIGN  
& USIHC 2022

## Explorando desafios e oportunidades para aumentar a transparência de Assistentes de Voz (AV)

*Exploring challenges and opportunities to increase the transparency of Voice Assistants (VAs)*

Isabela Motta; Laboratório de Ergodesign e Usabilidade de Interfaces; PUC-Rio  
Manuela Quaresma; Laboratório de Ergodesign e Usabilidade de Interfaces; PUC-Rio

### Resumo

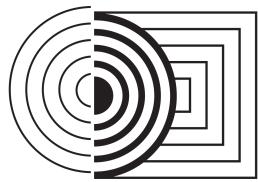
Assistentes de Voz (AV) são cada vez mais populares entre usuários. Porém, evidências mostram que os modelos mentais que usuários constroem sobre os AV não estão alinhados com as reais capacidades do sistema, levando à frustração ou, até mesmo, o abandono parcial ou total desses sistemas pelos usuários. A literatura indica que aumentar a transparência nas respostas que os AV apresentam aos usuários pode facilitar a compreensão sobre o funcionamento do assistente de forma geral. No entanto, ainda é necessário entender quais são as melhores formas de implementar tal solução, identificando os benefícios e desafios associados a ela. Assim, o objetivo deste estudo foi entender as oportunidades e limitações no aumento da transparência do sistema em respostas de AV. Para isso, foram conduzidas entrevistas com profissionais experientes com a pesquisa e/ou o desenvolvimento de interfaces de voz. Os resultados mostraram que os principais desafios para a implementação da proposta são em relação à tecnologia, a aspectos de negócios e de processo de projeto de AV, às características desses sistemas e a usabilidade e experiência dos usuários.

Palavras-chave: Assistentes de Voz; Interfaces de Voz; Transparência; Modelos Mentais;

### Abstract

*Voice Assistants (VAs) are increasingly popular among users. However, evidence shows that users' mental models of VAs are not aligned with these systems' real capabilities, leading to frustration or even partial or complete abandonment of these systems by users. The literature suggests that increasing the transparency in the outputs VAs provide for users might improve their comprehension about the assistants' general functioning. Nevertheless, it is still necessary to understand the best paths to implementing such solution, identifying its benefits and challenges. Thus, this study's goal was to understand the opportunities and limitations in increasing transparency in VA outputs. To this end, we interviewed professionals experienced with research or development of voice interfaces. The results show that the main challenges to implemeting the solution are related to technology, business and design aspects, VAs characteristics, and usability and user experience.*

Keywords: Voice Assistants; Voice User Interfaces; Transparency; Mental Models;



# 18º ERGODESIGN & USIHC 2022

## 1. Introdução

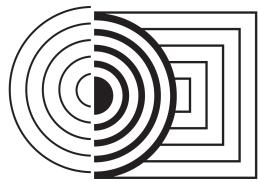
Assistentes de Voz (AV) são assistentes virtuais – agentes de inteligência artificial que realizam diversas tarefas em um sistema – com os quais os usuários podem interagir por meio de uma interface de voz, acompanhada ou não de uma interface visual (WEST; KRAUT; HAN EI, 2019). Siri, Alexa e Google Assistente são alguns dos AV disponíveis no mercado e operam em diferentes dispositivos, como em smartphones, fones de ouvido e smart speakers. O uso da interação por voz possibilitada por essas interfaces trazem muitos benefícios, incluindo vantagens para a experiência do usuário e acessibilidade (BALASURIYA *et al.*, 2018; PEARL, 2016). Estudos estimam que os AV estiveram em uso em mais de quatro bilhões de dispositivos até 2020 (MOAR; ESCHERICH, 2020), indicando o crescimento na popularidade desses produtos.

Apesar do aumento no uso de AV por usuários, ainda existem barreiras para a ampla adoção desses sistemas. A literatura aponta obstáculos como atitudes negativas de usuários em relação à interação por voz, preocupação com a privacidade de dados, percepção de baixa utilidade nos AV e diferentes tipos de falhas no reconhecimento de comandos, que levam à frustração para os usuários (BURBACH *et al.*, 2019; COWAN *et al.*, 2017; KISELEVA *et al.*, 2016; LOPATOVSKA *et al.*, 2019; LUGER; SELLEN, 2016; MAUÉS, 2019; MCLEAN; OSEI-FRIMPONG, 2019; MORIUCHI, 2019; MOTTA; QUARESMA, 2019; PRIDMORE *et al.*, 2019; PURINGTON *et al.*, 2017). Se não solucionados, tais problemas podem comprometer o potencial dos AV em mudar o atual paradigma interacional de interfaces gráficas para interfaces conversacionais (WEST; KRAUT; HAN EI, 2019).

Dentre as barreiras mencionadas acima, pode-se identificar problemas relacionados à percepção dos usuários sobre os AV, isto é, os modelos mentais que constroem sobre tais assistentes. Como será discutido a seguir, há evidências de que os modelos mentais dos usuários sobre os AV não são realistas (LUGER; SELLEN, 2016). Visando-se mitigar essa questão, acredita-se que aumentar a transparência nas respostas apresentadas por AV pode melhorar os modelos mentais dos usuários. No entanto, ainda é necessário entender como aplicar essa solução de forma a considerar seus principais benefícios e desafios. Dessa forma, o objetivo deste estudo, de caráter exploratório, foi entender as oportunidades e limitações no aumento da transparência do sistema em respostas de AV.

## 2. Modelos mentais e transparência em AV

Visando-se abordar os desafios mencionados acima, é fundamental que os desenvolvedores de AV entendam o modelo mental que os usuários formam sobre tais sistemas. Modelos mentais são um tipo de modelo conceitual criado pelo usuário que representa como um produto ou sistema funciona (NORMAN, 2013), incluindo uma série de expectativas sobre os componentes do sistema, seu funcionamento e uso apropriado (WICKENS *et al.*, 2014). Esses modelos são



## 18º ERGODESIGN & USIHC 2022

essenciais para os usuários porque influenciam diretamente a forma como as pessoas realizam tarefas, ditando níveis de desempenho (WILSON; RUTHERFORD, 1989).

A formação dos modelos mentais está fortemente relacionada à aspectos de design. Norman (2013) defende que os usuários constroem suas percepções utilizando a imagem do sistema (dicas físicas perceptíveis do produto em si, experiências passadas, comerciais, manuais, etc), tornando as características do produto vitais para a formação de modelos mentais corretos. Similarmente, Wilson e Rutherford (1989, p.3, tradução nossa) argumentam que os modelos mentais são formados com base em “experiências prévias e observações atuais<sup>1</sup>”, indicando que a interação em si serve como base para a construção dos modelos mentais.

Em linha com tais teorias, a literatura sobre interação por voz evidencia que as respostas do sistema – *feedbacks* e *outputs* dos AV que apresentam informações sobre sua operação e capacidades – são um fator de grande influência para as percepções dos usuários de AV. Essas respostas influenciam o entendimento dos usuários sobre o real funcionamento dos AV, incluindo a consciência sobre tarefas disponíveis, o entendimento de ações do sistema e de fontes de erro, o conhecimento sobre termos relevantes para a privacidade de dados e a compreensão dos limites no reconhecimento de voz (PORCHERON *et al.*, 2018; PORCHERON; FISCHER; SHARPLES, 2017).

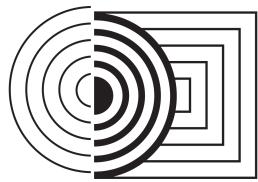
Entretanto, estudos mostram que os usuários têm modelos mentais inadequados, indicando que a forma como tais respostas são projetadas pode ser inapropriada. De forma geral, os usuários apresentam expectativas equivocadas para a inteligência e as capacidades técnicas, sociais e conversacionais dos AV, esperando que o sistema realize tarefas além de sua capacidade (CHO; LEE; LEE, 2019; LUGER; SELLEN, 2016). Além disso, os usuários subutilizam ou abandonam tarefas, reportam confusão sobre fontes de erro e ações do sistema, têm preocupações com privacidade de dados e empregam métodos de aprendizado inefficientes (BENETEAU *et al.*, 2020; CHO; LEE; LEE, 2019; LUGER; SELLEN, 2016; PRIDMORE *et al.*, 2019).

Uma possível solução para alinhar os modelos mentais de usuários de AV com a realidade do sistema pode estar no aumento da transparência nas respostas do sistema. Jordan (1998) define transparência (*explicitness*, no texto original) como “uma maneira de projetar um produto de forma que dicas sejam dadas sobre sua funcionalidade e método de operação<sup>2</sup>” (JORDAN, 1998, p. 38, tradução nossa). Segundo o autor, a transparência pode melhorar a usabilidade dos produtos e sistemas.

Estudos na área de sistemas de inteligência artificial (*Artificial Intelligence - AI*) e aprendizado de máquina reafirmam a possibilidade de que o aumento da transparência pode ser benéfico para produtos que aplicam essas tecnologias, como os AV. A literatura apresenta o conceito de *eXplainable Artificial Intelligence* (XAI), cujo objetivo é projetar sistemas de AI que possibilitem a usuários e desenvolvedores entenderem os motivos por trás de ações automatizadas (OXBOROUGH *et al.*, 2018). Ferreira e Monteiro (2020) identificaram por meio de um

<sup>1</sup> “previous experience as well as current observation” (WILSON; RUTHERFORD, 1989, p. 3)

<sup>2</sup> “Designing a product so that cues are given as to its functionality and method of operation” (JORDAN, 1998, p. 38)



## 18º ERGODESIGN & USIHC 2022

levantamento bibliográfico que aspectos de transparência em AI (XAI) são estudados para melhorar diversos aspectos da experiência do usuário (UX): aceitação, satisfação, engajamento, persuasão, continuidade do uso, aprendizagem, recuperação de falhas e entendimento sobre motivos que levam a erros.

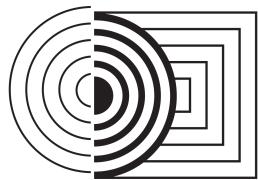
Especificamente, no contexto da interação por voz, evidências indicam que soluções relacionadas ao aumento da transparência podem ser eficazes. Kim, Jeong e Lee (2019) conduziram um experimento no qual motoristas interagiram com uma interface de voz que, no advento de erros, apresentava diferentes tipos de resposta. As variações nas respostas impactaram o entendimento dos participantes sobre a causa das falhas, e as respostas que explicavam como proceder melhoraram a recuperação de erros pelos motoristas. Motta e Quaresma (2022) conduziram testes de usabilidade com usuários de AV e categorizaram diferentes tipos de respostas do sistema e comportamentos dos usuários. As autoras observaram que respostas com mais informações sobre os motivos de erros ou como se recuperar deles (ex: instruções) levaram os participantes a apresentar comportamentos de reparação das falhas ao invés de recomeçar a tarefa do zero. Em linha com esses resultados, Kirschthaler, Porcheron e Fisher (2020) mostraram que aumentar a *discoverability* (explorabilidade) em uma interface de voz (ex: providenciar instruções e explicações) traz melhorias em usabilidade e desempenho de tarefas para os usuários.

Apesar de diversos estudos apontarem para a eficácia da transparência do sistema para melhorar a interação com AV, ainda é necessário compreender as melhores maneiras para implementar tal solução. Para isso, é necessário, primeiramente, entender quais são os principais desafios e oportunidades associados à transparência em respostas de AV. Tal conhecimento é fundamental para auxiliar designers e desenvolvedores na criação de AV que facilitem o entendimento correto do sistema para os usuários.

### 3. Metodologia

Para entender as oportunidades e limitações do aumento da transparência do sistema em respostas de AV, entrevistas semiestruturadas foram conduzidas com pesquisadores e desenvolvedores que trabalham no desenvolvimento de interfaces de voz. A técnica foi considerada adequada porque desejava-se explorar o tema da transparência, e tais profissionais têm experiência e conhecimento prático e teórico sobre o desenvolvimento de interfaces conversacionais. O modelo semiestruturado foi escolhido porque não requer que as perguntas tenham um número ou ordem fixas, permitindo que os participantes se expressem livremente. Assim, a moderadora pôde ajustar as perguntas de acordo com as respostas dos participantes enquanto seguiu uma série de tópicos pré-estabelecidos (GIL, 2008).

Como requisito de participação, os profissionais envolvidos com pesquisa deveriam ter título de Doutor e estarem envolvidos em projetos de pesquisa sobre interação por voz em geral. Já os profissionais ligados ao desenvolvimento dessas interfaces deveriam ter, pelo menos, três anos de experiência na área. Para diversificar as visões desses profissionais o máximo possível,



## 18º ERGODESIGN & USIHC 2022

participantes com formações em áreas diversas foram recrutados. Os participantes foram recrutados por e-mail e redes sociais (LinkedIn) e, no total, oito profissionais (7 brasileiros e 1 estrangeiro; 5 homens e 3 mulheres) foram entrevistados.

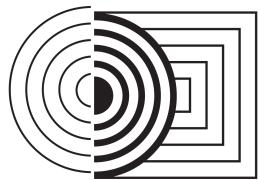
As entrevistas foram feitas de forma remota, pela plataforma Zoom, e o procedimento seguiu alguns passos. Em primeiro lugar, as entrevistas foram marcadas e os participantes receberam, de antemão e por e-mail, o termo de consentimento livre e esclarecido para participar da pesquisa. Então, na entrevista, a pesquisadora apresentou de forma ampla o objetivo do estudo e expôs o tema da interação com AV. A entrevistadora apresentou a questão dos modelos mentais incorretos que usuários criam sobre assistentes de voz e colocou, como hipótese de solução para esse problema, o aumento da transparência em respostas de AV. Foi pedido que os participantes comentassem sua opinião sobre tal hipótese, incluindo seus pontos de vista sobre viabilidade do aumento da transparência e se tal solução seria desejável. Os participantes também foram encorajados a exemplificar suas respostas com situações reais de trabalho. Outras perguntas foram feitas na entrevista, mas não serão abordadas neste artigo. O procedimento durou, em média, 45 minutos e as entrevistas foram gravados em vídeo e áudio. A análise dos dados foi feita a partir da revisão das gravações por uma pesquisadora. Uma análise temática foi feita por meio da identificação de tópicos emergentes nas respostas dos participantes.

### 4. Resultados e discussão

As respostas dos participantes se concentraram em dois temas principais: a viabilidade do aumento da transparência nas respostas do sistema e o quanto certa solução seria desejável. Em termos de viabilidade, os participantes falaram sobre viabilidade tecnológica, projetual e de negócios. Já os tópicos mencionados sobre o quanto desejável é a proposta giraram em torno de usabilidade e adequação aos objetivos dos usuários.

Em primeiro lugar, os profissionais não tiveram opiniões congruentes em relação às limitações tecnológicas que poderiam impedir o aumento da transparência do sistema em respostas de AV. Enquanto alguns mencionaram que a proposta parece tecnicamente viável, outros não tiveram certeza ou colocaram preocupações quanto ao tema. Um participante mencionou que apesar dos grandes avanços na tecnologia de reconhecimento de voz nos últimos anos, erros certamente irão acontecer. Segundo esse participante, para ajudar os usuários a se recuperarem em interações com interfaces de voz, os desenvolvedores devem entender como os humanos fazem isso em conversas reais: "*É importante olhar como as pessoas resolvem problemas de mal-entendidos (...). Elas entendem o que aconteceu e depois tentam consertar*" (P7).

Entretanto, o profissional explicou que, como existem muitas tecnologias envolvidas no funcionamento de um AV, os erros podem ter diversas origens, e o sistema nem sempre é capaz de diagnosticar porquê uma falha aconteceu. O participante exemplificou a situação com outros sistemas de IA: "*Um algoritmo de reconhecimento de imagens não explica porque ele não*



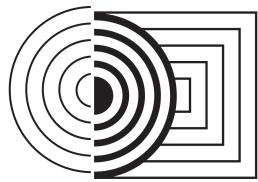
## 18º ERGODESIGN & USIHC 2022

*reconheceu uma imagem, ele só diz que não reconheceu*" (P7). Discussões similares podem ser encontradas na literatura sobre XAIs, que aponta que muitos sistemas – em especial, os de melhor desempenho – são “caixas pretas” incapazes de explicar como chegaram às suas decisões (HOLZINGER, 2018). Assim, pode ser difícil para um AV comunicar ao usuário o motivo de falhas ou como se recuperar - instruções essenciais para a transparência do sistema -, uma vez que o próprio sistema pode não conseguir diagnosticar o problema corretamente em todas as situações.

Além da tecnologia, alguns entrevistados comentaram sobre as limitações do processo de projeto à viabilidade do aumento da transparência. Os entrevistados mencionaram que o projeto de interfaces conversacionais está muito ligado à construção de fluxos de conversação: "*Ainda se pensa no modelo básico de interação, na árvore com as principais jornadas dos usuários. (...) É um fluxo limitado*" (P8); "*Quando a gente faz uma pergunta, a gente está lidando é com um fluxograma*" (P1). Segundo esses profissionais, pode ser difícil prever cenários e projetar respostas mais transparentes para incluí-las nos fluxos pré-estabelecidos. "*A cabeça do usuário é um mundo a parte. Eles interagem com as mesmas ferramentas de formas distintas. O modo como falam o mesmo comando é diferente*" (P8). Essa dificuldade se torna ainda mais saliente em sistemas como AV, que são generalistas e podem desempenhar diversas tarefas. Diferentes de outras interfaces conversacionais, como *voice bots* de atendimento ao consumidor, os AV não têm a operação limitada a um domínio específico (ex: seguro de automóveis).

De forma similar, um desafio relacionado à natureza generalista dos AV está na importância do contexto para o projeto de AV: "*É muito contextual, e é um sistema de aprendizado de máquina, e o contexto influencia muito na parte de conversação. (...) Olha a diferença de domínio de finanças, arte e para pessoas de baixa renda, a diferença de explainability que tem que ser.*" (P3). Logo, a dificuldade de prever contextos de uso também pode limitar o desenvolvimento de respostas mais transparentes. Além disso, no nível da interação, estudos mostram que a produção e o entendimento de explicações por seres humanos, de forma geral, são amplamente dependentes do contexto (MILLER, 2019), de forma que sistemas com XAIs precisam considerar o contexto e informações relevantes para a resposta apresentada (FERREIRA; MONTEIRO, 2020). No entanto, a literatura indica que os AV não são capazes de entender informações contextuais (ex: localizações, comandos anteriores da interação; AMMARI *et al.*, 2019; LUGER; SELLEN, 2016), podendo limitar ainda mais o projeto de respostas nesses sistemas.

Outro ponto colocado pelos participantes se relacionou a limitações na equipe de desenvolvimento das interfaces ou nos clientes que demandam esses projetos. Um entrevistado mencionou como desafio o fato de que há muitos profissionais envolvidos no desenvolvimento de AV, fazendo com que o sistema seja resultado de uma combinação de percepções desses projetistas: "*Tem dois modelos [mentais]: o do usuário e o do sistema. E o modelo do sistema pode parecer estranho, porque aí você vai lá no designer. Mas não é [só] o designer, porque é tão mais complexo o sistema de conversação, é tanta gente que mexe naquilo, tantos*



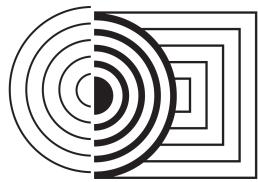
## 18º ERGODESIGN & USIHC 2022

*curadores*" (P3). Em linha similar, um entrevistado mencionou que é importante que a equipe tenha o conhecimento correto sobre projetar essas interfaces: "*O único impedimento vem da falta de conhecimento do time. A pessoa precisa saber estruturar o conteúdo da resposta usando práticas linguísticas*" (P6). Finalmente, diversos participantes mencionaram dificuldade na comunicação com gerentes ou demandantes do projeto, que muitas vezes não entendem os benefícios do aumento da transparência nas respostas do sistema. O exemplo real colocado por um participante ilustra esse tema: "*A pessoa ligava e caía numa URA [Unidade de Resposta Automática] (...) Então, a pessoa tinha consciência de que ela tava falando com um robô porque era DTMF a tecnologia ali de interação<sup>3</sup> (...). Aí, a URA falava assim: 'tá bom, eu já estou te transferindo para um de nossos atendentes', o que não deixava claro que seria um outro robô (...). Eles [a empresa] eram muito contra se apresentar como um robô e eu era muito contra não se apresentar [como robô] (...). A expectativa deles [consumidores] era ser atendido por um humano mas um bot atendia eles, então, talvez, enganava muita gente*" (P2).

Os participantes também foram perguntados sobre o quanto desejável seria o aumento na transparência do sistema, partindo do ponto de vista projetual, de negócios, do usuário, etc. A maioria dos participantes considerou que a solução pode ser vantajosa, principalmente, em relação à usabilidade dos AV. De forma geral, aqueles que foram a favor comentaram que o aumento da transparência pode resultar em menos erros e deixar os usuários mais seguros e confiantes: "*Deixa o sistema falar o que pode entender ao invés de deixar a pessoa descobrir sozinha e errar*" (P6). Alguns entrevistados consideraram vantajoso deixar o usuário entender os limites da máquina: "*Faz sentido ser honesto com o usuário. (...) As pessoas compreendem que tá, é uma máquina, mas não vai fazer tudo. (...) Acho que é interessante no sentido do usuário porque eu sei até onde eu posso ir naquele sistema. (...) Claro que tem a questão da expectativa e a realidade, mas nem sempre a realidade é algo frustrante.*" (P5). Além disso, foi comentado que aumentar a transparência dos AV pode ajudar não só usuários, mas também os atores envolvidos nos projetos a entender melhor as limitações do próprio sistema: "*Vale a pena para o entendimento de todo mundo que está nesse processo. Seja para o usuário ou para pessoas que projetam, porque para a gente também é um pouco complicado às vezes.*" (P5). Deixar explícito para os usuários as funcionalidades de uma interface, também pode ajudar na identificação de lacunas no produto: "*Isso acaba saindo até mais barato para empresa porque aí você já tem o mapeamento de aquilo que a pessoa tá sentindo falta*" (P5).

De forma contrária, outros participantes expressaram preocupação quanto à possíveis impactos negativos do aumento excessivo da transparência na usabilidade dos AV. Os entrevistados mencionaram que apresentar explicações demais sobre o sistema pode tornar as interações mais lentas, chatas ou confusas, e até mesmo prejudicar a facilidade da interação por voz: "*Porque um dispositivo que fica te explicando o tempo todo como você deve interagir é chato, é moroso, demora. A gente pouco tolera isso. (...) [a interação] deixa de explorar um dos aspectos positivos dessas interações que é o fácil aprendizado que as pessoas já tem uma ideia de que aquilo funciona*" (P4). Em linha similar, um participante questionou, inclusive, se projetar uma

<sup>3</sup> Interações que utilizam a tecnologia DTMF (Dual Tone Multi Frequency) são aqueles que pedem para o usuário clicar em um botão no teclado telefônico para selecionar um comando (ex: "para falar com um atendente, aperte 1").



## 18º ERGODESIGN & USIHC 2022

interface que apresente tantas informações sobre si mesma está alinhada com os objetivos dos usuários ao desempenharem suas tarefas: "Será que o público tá esperando essa complexidade de trabalho? De informações? Será que é isso que o público quer? (...) O público quer um companheiro nessa hora [quando realiza uma tarefa]? O público quer um Google nessa hora? O que me parece é que o público quer um Google." (P1).

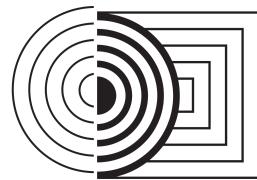
Os aspectos de usabilidade mencionados pelos entrevistados estão alinhados com a literatura, que aponta que, apesar do aumento na transparência melhorar o desempenho nas tarefas, o uso excessivo de instruções pode não agradar usuários e deixar as interações mais lentas (KIRSCHTHALER; PORCHERON; FISCHER, 2020; MOTTA; QUARESMA, 2022). Desenvolvedores também recomendam que o uso de instruções, confirmações ou outros *outputs* cujo objetivo é direcionar a interação sejam utilizados com cautela para evitar repetições (GOOGLE, 2017).

### 5. Conclusão

A popularidade de Assistentes de Voz (AV) é crescente, mas ainda existem barreiras para sua ampla adoção, como as incorreções nos modelos mentais criados por usuários sobre esses sistemas. Uma possível solução para essa questão é aumentar a transparência nas respostas apresentadas pelos AV aos usuários para explicitar suas formas de funcionamento, limitações e capacidades. Para entender como implementar tal solução adequadamente, o objetivo deste estudo, de caráter exploratório, foi entender as oportunidades e limitações no aumento da transparência do sistema em respostas de AV. Para isso, oito profissionais experientes na pesquisa ou desenvolvimento de interfaces conversacionais por voz foram entrevistados.

Os resultados apontaram desafios existentes para o aumento da transparência em respostas de AV. Devido à complexidade da tecnologia envolvida nesses sistemas, ainda pode ser difícil explicar para os usuários sobre o funcionamento do produto, especialmente, em casos de erros. Além disso, a própria natureza dos AV, que não operam em domínio específico e devem dar suporte à uma ampla gama de comandos, pode dificultar o projeto de respostas mais transparentes, que são fortemente ligadas à contextos e fluxos pré-planejados de conversa. Ainda, obstáculos como limitações de recursos em equipes e exigências de outros *stakeholders* envolvidos no projeto podem impedir a implementação de determinadas soluções. Finalmente, ainda é um desafio entender como projetar o aumento da transparência de forma balanceada para melhorar a usabilidade e a experiência dos usuários, uma vez que, se implementadas incorretamente, essas respostas podem ser prejudiciais.

Essa pesquisa teve como limitação principal uma pequena amostra de entrevistados. No entanto, como o estudo teve caráter exploratório, ainda serão conduzidos estudos futuros com uma amostra maior para ampliar os resultados apresentados. Como desdobramentos, ainda é necessário gerar hipóteses de recomendações para a aplicação correta de respostas mais transparentes em AV e conduzir estudos para testar tais presunções.



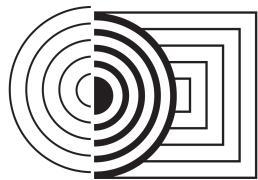
# 18º ERGODESIGN & USIHC 2022

## Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001; e da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

## 6. Referências Bibliográficas

- AMMARI, T.; KAYE, J.; TSAI, J. Y.; BENTLEY, F. Music, Search, and IoT: How people (really) use voice assistants. **ACM Transactions on Computer-Human Interaction**, vol. 26, no. 3, 2019. DOI 10.1145/3311956. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065127771&doi=10.1145%2F3311956&partnerID=40&md5=959c2e865e4d177d99df994284680aef>.
- BALASURIYA, S. S.; SITBON, L.; BAYOR, A. A.; HOOGSTRATE, M.; BRERETON, M. Use of voice activated interfaces by people with intellectual disability. 4 Dec. 2018., cited By 6. **Proceedings of the 30th Australian Conference on Computer-Human Interaction** [...]. New York, NY, USA: ACM, 4 Dec. 2018. p. 102–112.
- BENETEAU, E.; GUAN, Y.; RICHARDS, O. K.; ZHANG, M. R.; KIENTZ, J. A.; YIP, J.; HINIKER, A. Assumptions Checked: How Families Learn About and Use the Echo Dot. **Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.**, New York, NY, USA, vol. 4, no. 1, Mar. 2020. DOI 10.1145/3380993. Available at: <https://doi.org/10.1145/3380993>.
- BURBACH, L.; HALBACH, P.; PLETTENBERG, N.; NAKAYAMA, J.; ZIEFLE, M.; CALERO VALDEZ, A. “Hey, Siri”, “Ok, Google”, “Alexa”. Acceptance-Relevant Factors of Virtual Voice-Assistants. 2019-July., Jul. 2019. **2019 IEEE International Professional Communication Conference (ProComm)** [...]. [S. l.]: IEEE, Jul. 2019. vol. 2019-July, p. 101–111. DOI 10.1109/ProComm.2019.00025. Available at: <https://ieeexplore.ieee.org/document/8804568/>.
- CHO, M.; LEE, S.; LEE, K.-P. Once a Kind Friend is Now a Thing. 18 Jun. 2019. **Proceedings of the 2019 on Designing Interactive Systems Conference** [...]. New York, NY, USA: ACM, 18 Jun. 2019. p. 1557–1569. DOI 10.1145/3322276.3322332. Available at: <https://doi.org/10.1145/3322276.3322332>.
- COWAN, B. R.; PANTIDI, N.; COYLE, D.; MORRISSEY, K.; CLARKE, P.; AL-SHEHRI, S.; EARLEY, D.; BANDEIRA, N. “What can i help you with?”: Infrequent users’ experiences of intelligent personal assistants. 2017. **Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2017** [...]. [S. l.: s. n.], 2017. DOI 10.1145/3098279.3098539.
- FERREIRA, J. J.; MONTEIRO, M. S. What Are People Doing About XAI User Experience? A Survey on AI Explainability Research and Practice. [S. l.: s. n.], 2020. p. 56–73. DOI 10.1007/978-3-030-49760-6\_4. Available at: [http://link.springer.com/10.1007/978-3-030-49760-6\\_4](http://link.springer.com/10.1007/978-3-030-49760-6_4).



## 18º ERGODESIGN & USIHC 2022

GIL, A. **Métodos e técnicas de pesquisa social**. 6th ed. São Paulo: Editora Atlas, 2008.

GOOGLE. **Instilling User Confidence Through Confirmations and Acknowledgements**.

Disponível em: <<https://developers.google.com/actions/design/instilling-user-confidence>>.

Acesso em: 5 set. 2017

HOLZINGER, A. From Machine Learning to Explainable AI. Aug. 2018. **2018 World Symposium**

**on Digital Intelligence for Systems and Machines (DISA)** [...]. [S. I.]: IEEE, Aug. 2018. p. 55–66.

DOI 10.1109/DISA.2018.8490530. Available at:

<https://ieeexplore.ieee.org/document/8490530/>.

JORDAN, P. W. **An Introduction to Usability**. Philadelphia, PA: Taylor & Francis, 1998. vol. 352, .

DOI 10.1016/B978-0-240-81203-8.00002-7. Available at:

<http://www.useit.com/alertbox/20030825.html>.

KIM, J.; JEONG, M.; LEE, S. C. “Why did this voice agent not understand me?” 21 Sep. 2019., e1;

i1. **Proceedings of the 11th International Conference on Automotive User Interfaces and**

**Interactive Vehicular Applications: Adjunct Proceedings** [...]. New York, NY, USA: ACM, 21 Sep.

2019. p. 146–150. DOI 10.1145/3349263.3351513. Available at:

<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85073617376&doi=10.1145%2F3349263.3351513&partnerID=40&md5=b5d0d815cd5258eb6f55464559b0f517>.

KIRSCHTHALER, P.; PORCHERON, M.; FISCHER, J. E. What Can i Say?: Effects of Discoverability in VUIs on Task Performance and User Experience. 2020., cited By 2. **ACM International**

**Conference Proceeding Series** [...]. [S. I.: s. n.], 2020. DOI 10.1145/3405755.3406119. Available at:

<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85090136336&doi=10.1145%2F3405755.3406119&partnerID=40&md5=778305888d0b9419d24885855e1c7d05>.

KISELEVA, J.; WILLIAMS, K.; JIANG, J.; HASSAN AWADALLAH, A.; CROOK, A. C.; ZITOUNI, I.;

ANASTASAKOS, T. Understanding User Satisfaction with Intelligent Assistants. 2016. **Proceedings**

**of the 2016 ACM on Conference on Human Information Interaction and Retrieval** [...]. New

York, NY, USA: Association for Computing Machinery, 2016. p. 121–130. DOI

10.1145/2854946.2854961. Available at: <https://doi.org/10.1145/2854946.2854961>.

LOPATOVSKA, I.; RINK, K.; KNIGHT, I.; RAINES, K.; COSENZA, K.; WILLIAMS, H.; SORSCHE, P.;

HIRSCH, D.; LI, Q.; MARTINEZ, A. Talk to me: Exploring user interactions with the Amazon Alexa.

**Journal of Librarianship and Information Science**, vol. 51, no. 4, p. 984–997, 7 Dec. 2019. DOI

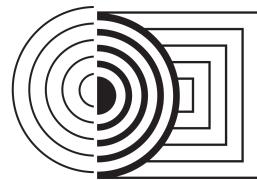
10.1177/0961000618759414. Available at: <https://doi.org/10.1177/0961000618759414>.

LUGER, E.; SELLEN, A. “Like Having a Really Bad PA”: The Gulf between User Expectation and Experience of Conversational Agents. 2016. **Proceedings of the 2016 CHI Conference on Human**

**Factors in Computing Systems** [...]. New York, NY, USA: Association for Computing Machinery,

2016. p. 5286–5297. DOI 10.1145/2858036.2858288. Available at:

<https://doi.org/10.1145/2858036.2858288>.



## 18º ERGODESIGN & USIHC 2022

MAUÉS, M. P. **Marcela Pedroso Maués Um olhar sobre os assistentes virtuais personificados e a voz como interface.** 2019. Pontifical Catholic University of Rio de Janeiro, 2019.

MCLEAN, G.; OSEI-FRIMPONG, K. Hey Alexa ... examine the variables influencing the use of artificial intelligent in-home voice assistants. **Computers in Human Behavior**, vol. 99, p. 28–37, 2019. DOI <https://doi.org/10.1016/j.chb.2019.05.009>. Available at: <http://www.sciencedirect.com/science/article/pii/S0747563219301840>.

MILLER, T. Explanation in artificial intelligence: Insights from the social sciences. **Artificial Intelligence**, vol. 267, p. 1–38, Feb. 2019. DOI [10.1016/j.artint.2018.07.007](https://doi.org/10.1016/j.artint.2018.07.007). Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0004370218305988>.

MOAR, J.; ESCHERICH, M. Hey Siri, how will you make money? 2020..

MORIUCHI, E. Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty. **Psychology & Marketing**, vol. 36, no. 5, p. 489–501, 15 May 2019. DOI [10.1002/mar.21192](https://doi.org/10.1002/mar.21192). Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mar.21192>.

MOTTA, I.; QUARESMA, M. OPPORTUNITIES AND ISSUES IN THE ADOPTION OF VOICE ASSISTANTS BY BRAZILIAN SMARTPHONE USERS. **Ergodesign & HCI**, vol. 7, no. Especial, p. 138, 31 Dec. 2019. DOI [10.22570/ergodesignhci.v7iEspecial.1312](https://doi.org/10.22570/ergodesignhci.v7iEspecial.1312). Available at: <http://periodicos.puc-rio.br/index.php/revistaergodesign-hci/article/view/1312>.

MOTTA, I.; QUARESMA, M. Users' Error Recovery Strategies in the Interaction with Voice Assistants (VAs). [S. l.: s. n.], 2022. p. 658–666. DOI [10.1007/978-3-030-74614-8\\_82](https://doi.org/10.1007/978-3-030-74614-8_82). Available at: [https://link.springer.com/10.1007/978-3-030-74614-8\\_82](https://link.springer.com/10.1007/978-3-030-74614-8_82).

NORMAN, D. **The design of everyday things**. Revised an. New York, NY, USA: Basic books, 2013.

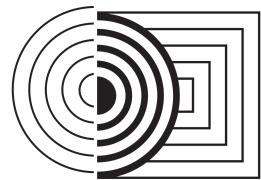
OXBOROUGH, C.; CAMERON, E.; RAO, A.; BIRCHALL, A.; TOWNSEND, A.; WESTERMANN, C. **Explainable AI: driving business value through greater understanding**. [S. l.: s. n.], 2018.

PEARL, C. **Designing Voice User Interfaces: Principles of Conversational Experiences**. [S. l.]: O'Reilly, 2016.

PORCHERON, M; FISCHER, J. E.; REEVES, S.; SHARPLES, S. Voice interfaces in everyday life. 2018-April., 2018., i2. **Conference on Human Factors in Computing Systems - Proceedings** [...]. [S. l.: s. n.], 2018. vol. 2018-April, . DOI [10.1145/3173574.3174214](https://doi.org/10.1145/3173574.3174214). Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85046968214&doi=10.1145%2F3173574.3174214&partnerID=40&md5=dd0f40084357cf979ce1233b688ae026>.

PORCHERON, Martin; FISCHER, J. E.; SHARPLES, S. "Do Animals Have Accents?" 25 Feb. 2017., E. **Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing** [...]. New York, NY, USA: ACM, 25 Feb. 2017. p. 207–219. DOI [10.1145/2998181.2998298](https://doi.org/10.1145/2998181.2998298). Available at: <https://doi.org/10.1145/2998181.2998298>.

PRIDMORE, J.; ZIMMER, M.; VITAK, J.; MOLS, A.; TROTTIER, D.; KUMAR, P. C.; LIAO, Y. Intelligent Personal Assistants and the Intercultural Negotiations of Dataveillance in Platformed



## 18º ERGODESIGN & USIHC 2022

Households. **Surveillance & Society**, vol. 17, no. 1/2, p. 125–131, 31 Mar. 2019. DOI 10.24908/ss.v17i1/2.12936. Available at: <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/12936>.

PURINGTON, A.; TAFT, J. G.; SANNON, S.; BAZAROVA, N. N.; TAYLOR, S. H. "Alexa is my new BFF": Social roles, user satisfaction, and personification of the Amazon Echo. **Conference on Human Factors in Computing Systems - Proceedings**, vol. Part F1276, p. 2853–2859, 2017. <https://doi.org/10.1145/3027063.3053246>.

WEST, M.; KRAUT, R.; HAN EI, C. **I'd blush if I could.** [S. l.: s. n.], 2019. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=1>.

WICKENS, C. D.; LEE, J. D.; LIU, Y.; GORDON-BECKER, S. **Introduction to human factors engineering.** 2nd ed. London, England: Pearson, 2014.

WILSON, J. R.; RUTHERFORD, A. Mental models: Theory and application in human factors. **Human Factors**, vol. 31, no. 6, p. 617–634, 1989. <https://doi.org/10.1177/001872088903100601>.