

Arquitetura da Incompreensão: Análise Crítica da Leiturabilidade em Documentos das Assistentes Virtuais

Danielle Sanches


*Doutora em História das Ciências pela EHESS/Paris e COC/Fiocruz.
Professora da Escola de Comunicação, Mídia e Informação da Fundação
Getulio Vargas (FGV Comunicação Rio).*

Lara Dalese

*Graduanda de Comunicação Digital na Escola de Comunicação, Mídia e
Informação da Fundação Getulio Vargas (FGV Comunicação Rio).*

Júlia Frazão

*Graduanda de Comunicação Digital na Escola de Comunicação, Mídia e
Informação da Fundação Getulio Vargas (FGV Comunicação Rio).*



Resumo: Este estudo analisa as políticas de privacidade e os termos de uso das assistentes virtuais Siri (Apple) e Alexa (Amazon). A partir de metodologia de análise documental e uso do software ALT, investigou-se o grau de compreensão exigido para interpretar esses textos, evidenciando lacunas na transparência das operações dessas tecnologias, sobretudo no que se refere a métodos de busca e processamento de informações. Os resultados apontam discrepâncias entre o conteúdo dos documentos oficiais e suas aplicações práticas, revelando concentração de poder e vigilância pelas empresas. Identificou-se ainda a presença de áreas ofuscadas, com baixa clareza, que podem favorecer a circulação de desinformação e narrativas revisionistas. Quanto à leitura, os documentos da Siri exigem nível superior ou pós-graduação para plena compreensão, apresentando baixo grau de acessibilidade textual. Já os documentos da Alexa, apesar de avaliados como de alta leitura (adequados a adolescentes de 16 a 17 anos), encontram-se fragmentados em hiperlinks, totalizando mais de cem páginas. Para tanto, há necessidade de maior transparência por parte das Big Techs, bem como da promoção do letramento midiático e digital, via políticas públicas, como instrumentos para participação social ativa e crítica sobre o uso de dados e a circulação de informações.

Palavras-chaves: Assistentes Virtuais de Voz; Privacidade de dados; Bigtechs; Consumo Informacional; Leitura Textual.

The Architecture of Misunderstanding: A Critical Inquiry into the Readability of Virtual Assistants' Documents

Abstract: This study examines the privacy policies and terms of use of the virtual assistants Siri (Apple) and Alexa (Amazon). Employing a documentary analysis methodology and the ALT software, the research investigated the level of comprehension required to interpret these texts, highlighting gaps in the transparency of these technologies' operations, particularly with regard to search methods and information processing. The findings indicate discrepancies between the content of the official documents and their practical applications, revealing the concentration of power and surveillance by these companies. The analysis also identified the existence of obscured areas, marked by low clarity, which may facilitate the dissemination of disinformation and revisionist narratives. In terms of readability, Siri's documents were assessed as requiring higher education or postgraduate training for full comprehension, thus presenting a low level of textual accessibility. In contrast, Alexa's documents, although classified as highly readable (suitable for adolescents aged 16 to 17), are fragmented across hyperlinks, amounting to over one hundred pages. These results underscore the need for greater transparency from Big Tech companies, as well as the promotion of media and digital literacy through public policies, as essential tools to foster more active and critical social participation regarding data usage and the circulation of information.

Keywords: Voice Virtual Assistants; Data Privacy; Big Techs; Information Consumption; Textual Readability

Introdução

A contemporaneidade está marcada pelas realidades sintéticas, ou seja, vivemos em um tempo fluido, onde as suas características têm sido desenhadas por uma sociabilidade mediada por sistemas de inteligência artificial (IA), como, por exemplo, os Assistentes Virtuais Inteligentes (AVIs) ativados por voz (CARDENUTO et al., 2023). Com cerca de 5,3 bilhões de usuários de internet no mundo, sendo 141 milhões apenas no Brasil (CETIC BR, 2024), há uma demanda crescente por informações de toda ordem e os assistentes virtuais surgiram com a função de aprimorar a varredura seletiva que os usuários fazem na web, localizando e apresentando a informação procurada, datificando e segmentando ainda mais as informações solicitadas (CRUZ et al., 2013).

Esses AVIs são softwares projetados para interagir com diversos tipos de usuários em linguagem natural. Utilizam como método principal inputs (entrada), a partir dos dados de comandos de voz ou de texto (MATOS; OLIVEIRA, 2021). No entanto, esses sistemas são artefatos culturais projetados por humanos que expressam e executam um conjunto opaco de valores, práticas e interesses, haja vista a ausência de transparência no que se refere aos documentos oficiais¹ acerca das políticas de privacidade e de uso desses recursos. Nesse aspecto, as políticas de governança de dados dos AVIs deveriam se basear em transparência e responsabilidade pública, para que os cidadãos tivessem uma compreensão clara de onde as informações são provenientes. No entanto, os AVIs ainda são caixas-pretas para os dados sociais; e, como essas práticas de utilização dos dados a partir dos assistentes, ainda são pouco estudadas, as fontes de informação utilizadas pelos Assistentes Virtuais estão mais próximas de hieróglifos do que de barômetros sociais (GILLESPIE, 2014).

¹ Ver documentação da Alexa e Siri: Termos de Uso da Alexa: Ajuda e Serviços de Atendimento ao Cliente. Disponível em: <https://www.amazon.com.br/gp/help/customer/display.html?nodeId=201809740> . Acesso em: 15 outubro 2024. Legal. Siri, Ditado e Privacidade. Disponível em: <https://www.apple.com/br/legal/privacy/data/pt-br/ask-siri-dictation/> Acesso em: 15 outubro 2024.

Sistemas como a Alexa da Amazon ou a Siri da Apple, entre outras, unem robótica, imaginada desde os anos 1920, e algoritmos de aprendizado de máquina, que buscam enriquecer o software com subjetividade semelhante à humana desde os anos 1950, período de desenvolvimento dos sistemas de Inteligência Artificial (GUNKEL, 2017). Portanto, os AVIs são estruturas projetadas com intenções antropomórficas: eles recebem um nome pelo qual são chamados, uma voz humanizada à qual respondem, e, um gênero virtual atribuído (NASS e MOON, 2000; SANTOS, 2022). Essa performidade programada, como se fossem humanos, revitaliza o reconhecimento de que os computadores são atores sociais que refletem as regras da interação humana (LATOUR, 2012; NASS, STEUER, e TAUBER, 1994).

Como observado, os Assistentes Virtuais Inteligentes constituem uma mudança na comunicação humana com implicações sociais de grande alcance, inclusive na esfera educacional e da saúde (BENTES, SANCHES e FONSECA, 2024). Estão baseados na interação entre o poder computacional e a precisão algorítmica, imiscuindo-se com a inferência de padrões a partir de diferentes conjuntos de dados acessados, o que corrobora para o fortalecimento da crença, socioculturalmente construída, da superioridade epistêmica objetiva dos dados (BOYD e CRAWFORD, 2012). Dado que a mediação algorítmica é baseada em tornar os humanos detectáveis, rastreáveis e registráveis, surgem preocupações fundamentais sobre o avanço de uma sociedade de vigilância distribuída em rede, viés algorítmico, insegurança e divisões de gênero que impactam o consumo informacional a partir das respostas informativas que esses AVIs concedem aos seus usuários (ALPAYDIN, 2016; CHUNG et. al, 2017; DELEUZE, 1992; VAN DIJCK, 2014; FOUCAULT, 1987; UNESCO, 2019).

A partir dessa preocupação nos dedicaremos, neste artigo, em analisar o conteúdo dos termos de uso e de privacidade de dados das assistentes virtuais Alexa (Amazon) e Siri (Apple), buscando compreender as implicações da falta de transparência informacional sobre o que é feito com os dados pessoais e informações concedidas, presentes nos documentos disponíveis, por parte das referidas Bigtechs que pode gerar

incompreensibilidade sobre as informações compartilhadas com as assistentes. Para isso, utilizamos a metodologia de índice de leituraabilidade (DUBAY, 2007) , muitas vezes traduzido como legibilidade, que permite a detecção do grau acadêmico indicado para a compreensão do texto examinado. Nosso objetivo é verificar informações disponíveis sobre: os dados dos usuários - o que cada empresa diz fazer e utilizar desses dados; o direito à privacidade por parte dos seus usuários - à compreensão dos usos da voz, dos seus dados pessoais e de terceiros; o rastreamento das informações solicitadas. A preocupação deste artigo se justifica pela necessidade de mapeamento dos efeitos culturais e implicações políticas desses artefatos comunicacionais orientados por sistemas de inteligência artificial, sobretudo em um contexto onde as empresas proprietárias os indica como apoio ao aprendizado e melhoria da experiência do aluno, funcionando como ativadores por voz de ambientes educacionais (ZACCARELLI, 2021).

Nesse sentido, temos como premissa de pesquisa que os termos de uso e documentação sobre privacidade dos dados são promovem, propositalmente, uma ofuscação da informação, em razão de uma escolha de negócio, pois eles protegem mais a privacidade das Bigtechs do que a dos seus usuários-consumidores o que gera riscos para os usuários-consumidores desses produtos. Em razão disso, buscaremos responder a seguinte questão: Quais informações estão de fato públicas para o usuário? Qual grau de letramento digital e midiático é necessário para que o usuário possa ter um conhecimento pleno do que esses documentos revelam? Para isso, dividimos o artigo em quatro partes, além da introdução e conclusão, onde abordamos sobre a (1) documentação da Alexa e da Siri e o que referem sobre os dados pessoais; (2) apresentação da metodologia utilizada; (3) análise de leituraabilidade dos documentos das assistentes de forma comparativa; e, (4) sobre a ausência de transparência das Bigtechs em relação ao uso de dados pessoais que essas assistentes fazem.

Alexa e Siri: como tratam os dados pessoais?

Lançada pela Amazon, globalmente, no final de 2014, a assistente virtual Alexa conquistou um espaço significativo na vida dos brasileiros, sendo usada não apenas para facilitar tarefas cotidianas no trabalho, mas também em momentos de lazer². Por meio de comandos de voz, a Alexa consegue realizar funções que vão desde apagar as luzes via dispositivos *Bluetooth* até realizar compras — algo que, segundo a Amazon Brasil, ocorreu aproximadamente 300 milhões de vezes em 2024 (Amazon, 2023)³. Com o tempo, a assistente deixou de ser apenas uma ferramenta funcional e tornou-se uma forma de interação social para muitos usuários. Estes, muitas vezes, vão além dos comandos básicos (previsão do tempo, ouvir música, ações de internet das coisas, etc) e buscam conversar e interagir, como se fosse uma amiga ou membro da família, com a Alexa, impulsionados pela crescente integração da internet ao dia a dia e pela comodidade trazida pela assistente. No entanto, o avanço da tecnologia muitas vezes é feito sem uma mensuração dos riscos associados, como a observação sobre a privacidade dos dados e a consideração sobre a influência da inteligência artificial na formação da opinião pública. Isto ocorre porque a prioridade dessas tecnologias é, na maioria das vezes, a experiência do usuário, o lucro da empresa proprietária e não o bem-estar coletivo.

Segundo o mesmo relatório produzido pela Amazon, até maio de 2024 mais de 17 milhões de dispositivos de casa inteligente estavam conectados à Alexa no Brasil, isso representa cerca de 24% dos domicílios no Brasil (IBGE, 2022). No site da empresa⁴, estão disponíveis para compra mais de 800 modelos de aparelhos compatíveis com a assistente,

² Ver documentação: 23 Formas de Usar a Alexa para Entreter as Crianças. Disponível em: <https://www.aboutamazon.com.br/noticias/dispositivos/23-formas-de-usar-alexa-para-entreter-as-criancas>. Acesso em: 22 janeiro 2025

³ Ver: Números mostram o quanto a Alexa conquistou os corações dos brasileiros. Disponível em: <https://www.aboutamazon.com.br/noticias/dispositivos/numeros-mostram-o-quanto-alexa-conquistou-os-coracoes-dos-brasileiros>. Acesso em: 22 janeiro 2025.

⁴ Ver: Amazon.com.br. Disponível em: https://www.amazon.com.br/?&tag=hydrbrabk-20&ref=pd_sl_2symr9h1mr_e&adgrpid=155790195538&hvpones=&hvptwo=&hvadid=677606588095&hvpos=&hvnetw=g&hvrnd=7716158221652962857&hvqmt=e&hvdev=c&hvdvcmdl=&hvlocint=&hvlocphy=9198792&hvtargid=kwd-11510562736&hydadcr=6438_13232876&gad_source=1. Acesso em: 22 janeiro 2025.

que abrange desde aspiradores robôs e lava-louças até alimentadores automáticos de animais (com câmera e microfone), lâmpadas e interruptores de luz inteligentes. Dado o elevado número de brasileiros com acesso direto à essa tecnologia, surge a questão: quantos desses usuários realmente leram os Termos de Uso? E, quanto deles compreenderam como seus dados serão utilizados? Quantos desses fizeram essa leitura antes de adquirir o produto compreendendo a possível posição de risco da sua privacidade?

A pesquisa publicada no *Interacting with Computers Journal* (2018) reforça essa preocupação, pois quase 70% dos indivíduos não costumavam ler manuais de instrução ou rótulos de alimentos, o que dificultava o entendimento das condições e consequências ruins dos produtos adquiridos. Mesmo que essa porcentagem tenha mudado, desde a publicação do relatório até os dias atuais, muitos dos habitantes das 17 milhões de casas brasileiras, conectadas a dispositivos inteligentes, podem não ter pleno conhecimento de quais são as implicações de privacidade ao permitirem que uma empresa de tecnologia acesse seu espaço pessoal, voz e dados de familiares. Seja tanto pela falta de hábito de consulta quanto pela dificuldade de compreensão dos Termos de Uso – somando a página principal e todos os *hyperlinks* que são adicionais – o excesso de informações e a forma de disposição do texto, desses documentos, tornam o acesso bastante limitado.

No que se refere à Siri, lançada pela Apple em 2011 com o iPhone 4S, representa um marco significativo na evolução das interfaces de interação humano-computador. Como assistente virtual integrada ao ecossistema de dispositivos da Apple - incluindo iPhones, iPads, Macs e Apple Watches - emprega tecnologia de inteligência artificial, como processamento de linguagem natural e reconhecimento de voz para executar tarefas e responder aos comandos dos usuários, assim como a Alexa. Porém, é fundamental ressaltar que os produtos da Apple estão presentes entre 87% do público adolescente norte americano (Piper Sandler, 2024), sendo o iPhone a marca de celular utilizada por aproximadamente 9 em cada 10 jovens americanos; como a Siri está inserida nos aparelhos Apple, podemos afirmar que todos esses jovens possuem uma assistente.

Além disso, a pandemia impulsionou um crescimento de 47% no uso de todas as assistentes virtuais por voz. Segundo pesquisa da Ilumeo (2025), o Brasil conta com um cenário expressivo: 90% das pessoas de todas as gerações conhecem assistentes de voz, com 80% reportando uso efetivo da tecnologia. A geração Z se destaca como a principal usuária, embora apresente menor preocupação com questões de privacidade. Esses dados não apenas demonstram a massiva adoção de assistentes virtuais no cotidiano, mas também levantam importantes reflexões sobre o equilíbrio entre conveniência tecnológica e proteção de dados pessoais, especialmente considerando a crescente integração desses sistemas na rotina das pessoas.

A propósito das escolhas metodológicas

Neste artigo utilizamos um levantamento documental para identificar e compreender a ofuscação e o não dizer sobre os dados de privacidade dos usuários dos AVIs. Para isso coletamos todos os documentos referentes a Termos de uso e privacidade de dados presentes nos websites da Amazon e da Apple sobre Assistentes Virtuais Inteligentes e dados de usuários. Embora saibamos que as escolhas que fizemos não possibilitam generalizações, pois os resultados não correspondem a todos os assistentes virtuais existentes, eles são capazes de nos fornecer informações ricas e minuciosas sobre um determinado grupo, assunto, objeto ou caso, permitindo uma compreensão dos fatos a partir da perspectiva de quem ou do quê está sendo observado, promovendo um entendimento profundo e detalhado (SOUZA, KANTORSKI, LUIS, 2011). Para tanto, escolhemos como procedimento metodológico a análise documental, utilizando dessa maneira, técnicas para observação dos Termos de uso e privacidade de dados das assistentes virtuais Alexa e Siri (SÁ-SILVA, ALMEIDA e GUINDANI, 2009).

Após a seleção dos documentos, ao todo mais de 100 (cem) páginas de informações sobre privacidade e retenção de dados da Alexa e da Siri, utilizamos a metodologia de análise de leituraabilidade, também chamada de legibilidade (MORENO et al, 2023). A

leiturabilidade refere-se à característica que influencia a facilidade com que um texto pode ser lido e é criada por *escolhas de conteúdos, estilo, design e organização textual que se adequam ao conhecimento prévio, escolaridade, interesse e motivação do público leitor* (DUBAY, 2007: 6). Existe uma diferença entre a leiturabilidade de um texto e a compreensão do mesmo, pois a compreensão advém de uma relação leitor-conteúdo enquanto a leiturabilidade está diretamente relacionada com o conteúdo e portanto existem formas objetivas de medi-las, a partir da utilização de diversas fórmulas existentes como as *readability formulas* (SMITH; TAFFLER, 1992).

No contexto das impressões físicas e digitais dos textos, a leiturabilidade, trata-se de um atributo essencial que impacta a clareza dos caracteres e a fluidez da leitura, ou seja, para que um material impresso seja considerado legível é fundamental que o leitor consiga interpretar rapidamente os símbolos tipográficos, convertendo-os em significado sem esforço excessivo. Portanto, textos com alta leiturabilidade favorecem a assimilação do conteúdo, enquanto aqueles com baixa leiturabilidade podem comprometer a compreensão e reduzir o engajamento dos leitores, especialmente em ambientes educacionais ou ambientes onde a informação que se pretende transmitir afete diretamente o leitor-usuário (HOLTZ; SANTOS, 2021).

Para conseguir uma avaliação da leiturabilidade textual nos Termos de uso e privacidade de dados das assistentes virtuais inteligentes, utilizamos o software ALT - *Análise de Legibilidade Textual*⁵, que fornece dados quantitativos sobre as pretensões de inteligibilidade e adequação normativa, além disso, ele nos permitiu verificar a frequência com que determinadas palavras-chave foram mencionadas ao longo dos documentos, gerando nuvens de palavras. As escalas empregam dois principais graus de mensuração: uma escala de 0 a 100, na qual o valor máximo indica extrema simplicidade e o valor mínimo denota complexidade elevada, utilizada pelo Teste de Facilidade de Leitura de Flesch e pelo Índice Gulpease; e, a escala de 0 a 20, que correlaciona o valor numérico aos anos de escolaridade necessários para a compreensão adequada do texto,

⁵ Acesse o software: <https://legibilidade.com/>

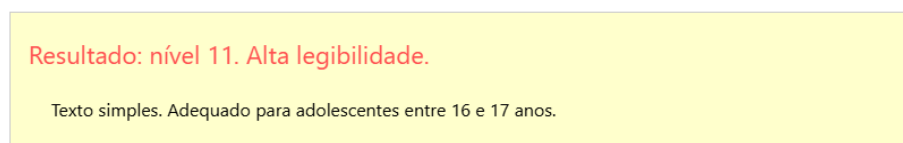
aplicada em métricas como o Nível de Instrução de Flesch-Kincaid, o Índice de Nebulosidade de Gunning, o Índice de Leiturabilidade Automatizado e o Índice de Coleman-Liau, utilizados para verificação do grau de dificuldade de leitura de um texto.

A fim de responder as questões referentes a quais informações estão de fato públicas para o usuário nos termos de uso e qual o grau de letramento digital e midiático se faz necessário para que o usuário possa ter um conhecimento acerca das informações presentes nestes documentos, exploramos o corpus composto pelos termos publicados pela Amazon e Apple, articulando abordagens quantitativas e qualitativas baseadas na combinação entre técnicas computacionais, aplicando o software ATL.

Enquadramos a análise realizada neste artigo nas abordagens de métodos digitais mistos, que combina um conjunto de técnicas quantitativas e qualitativas apoiadas por ferramentas digitais (OMENA, 2019). Na próxima seção, veremos detalhadamente os resultados obtidos no emprego desses métodos.

Análise de Leiturabilidade dos documentos

Ao realizarmos uma análise de leiturabilidade nas oito principais páginas dos Termos de uso da Alexa, das cento e seis páginas coletadas sobre políticas de privacidade de dados da Amazon, constatamos que o vocabulário utilizado ao longo do documento não exige alto grau de escolaridade ou atenção redobrada devido ao uso excessivo de palavras complexas (que atinge apenas 15% do total). À primeira vista, o texto parece ser de leitura simples e acessível para a maior parte dos usuários, com notas consideravelmente boas em diversas métricas das análises aplicadas, a saber: Índice Gulpease, Nível de Graduação de Flesch-Kincaid, Índice de Leiturabilidade Automatizado e Índice de Coleman-Liau. Porém, pela forma como o texto é disponibilizado pela empresa em seu website, a análise não considera um fator essencial neste caso: os *hyperlinks* espalhados ao longo dos Termos de Uso e a fragmentação da informação, que supostamente complementam as indicações que não foram devidamente esclarecidas ao longo das oito páginas.

Figura 1 - Resultados da Análise de Leiturabilidade dos Termos de Uso da Alexa

Métrica	Pontuação
Teste de facilidade de leitura de Flesch	44.2
Índice Gulpease	57.2
Nível de graduação de Flesch-Kincaid	11.4
Índice de nebulosidade de Gunning adaptado	10.3
Índice de legibilidade automatizado (ARI)	10.4
Índice de Coleman-Liau	12.5

Fonte: “Termo de uso Alexa”, ALT. Elaborado pelas autoras.

Com 12 *hyperlinks* distribuídos ao longo do texto, essas ramificações fragmentam a leitura oferecendo apenas explicações curtas sobre informações importantes para o usuário, como o uso dos dados e como “adaptar” a assistente para o alcance de crianças. A necessidade de acessar múltiplas páginas para obter respostas potencialmente completas, além de desestimular o usuário, gera incerteza sobre os limites do serviço da empresa e a privacidade do comprador.

Assim, um documento inicialmente classificado como “Alta Leiturabilidade, texto simples e adequado para adolescentes entre 16 e 17 anos”, como podemos verificar a partir dos resultados da análise realizada (figura 1), na verdade é um conjunto de 106 páginas interconectadas que desafiam a compreensão integral do usuário. Deve-se ressaltar que este total não contempla os adicionais distribuídos pelas 12 “abas” iniciais do site, referente à privacidade, que tornam a leitura ainda mais longa, ou seja, no total são mais

de 106 páginas dedicadas à Alexa, suas funcionalidades e sua relação com seu usuário-consumidor⁶.

A figura a seguir apresenta uma nuvem de palavras, que destaca os termos mais recorrentes ao longo da documentação analisada. "Amazon", a segunda mais frequente após o nome da própria assistente virtual, não aparece apenas para nomear a empresa, mas também como parte do nome de cinco hyperlinks. Outra palavra de destaque é "informações", mencionada 18 vezes ao longo do documento e assim reforçando sua centralidade no discurso da empresa. No entanto, essa repetição contrasta com a real disponibilidade de conhecimentos completos e realmente acessíveis para o usuário.

Figura 2 - Nuvem de palavras com os termos mais repetidos ao longo do documento



Fonte: “Termo de uso Alexa”, ALT. Elaborado pelas autoras.

Em relação aos termos de uso da Siri, é possível perceber que um dos principais problemas é que, apesar de ser um texto curto e de rápida leitura, há um uso frequente

⁶ Ver documentação: Termos de Uso da Alexa. Disponível em: <https://www.amazon.com.br/gp/help/customer/display.html?nodeId=201809740>. Acesso em: 22 janeiro 2025.

de termos técnicos como "criptografia de ponta a ponta", "identificador aleatório" e "processamento local". Algumas recomendações para melhorias seriam simplificar a linguagem técnica, criar seções mais claramente definidas com subtítulos informativos e incluir um glossário para termos específicos necessários, dado que o nível atual de complexidade pode dificultar a compreensão para usuários comuns, além de que muitas pessoas podem acabar aceitando sem entender completamente suas implicações ou até mesmo desencorajá-las de lerem o documento.

Na análise específica dos termos de uso e documentação da Siri da Apple, disponíveis na seção Legal, em Siri, Ditado e Privacidade, os resultados obtidos nas duas escalas evidenciam um nível crítico de complexidade textual. Na escala de 0 a 100, o Teste de Facilidade de Leitura de Flesch apresentou valor de 44.1 e o Índice Gulpease de 52.1, indicando textos de difícil leitura e compreensão. Na escala de 0 a 20, o Nível de Instrução de Flesch-Kincaid (13.7), o Índice de Nebulosidade de Gunning (27.6), o Índice de Leiturabilidade Automatizado (13.7) e o Índice de Coleman-Liau (11.3) apontam que seria necessária formação superior ou pós-graduação para compreensão adequada do conteúdo.

Figura 4 - Resultados da Análise de Leiturabilidade dos Termos de Uso da Siri



Fonte: “Termo de uso Siri”, ALT. Elaborado pelas autoras.

Estes índices revelam uma significativa falta de transparência por parte da Apple na comunicação com seus consumidores, considerando que a Siri é uma tecnologia de uso cotidiano, presente em dispositivos utilizados por pessoas com diferentes níveis de escolaridade, inclusive crianças. A escolha por uma linguagem técnica, evidenciada pela alta porcentagem de palavras complexas (78.2%) e pela média elevada de palavras por sentença (26.0), sugere uma possível estratégia proposital de obscurecimento das reais implicações do uso da tecnologia, particularmente em aspectos cruciais relacionados à privacidade e ao processamento de dados pessoais. Com isso, fica difícil para os usuários entenderem como suas informações são usadas e quais são seus direitos ao utilizar a Siri.

A nuvem de palavras gerada a partir dos termos de uso da assistente revela que, naturalmente, as palavras de maior destaque são “Siri”, aparecendo 74 vezes, “Apple” e “Ditado”, refletindo os elementos centrais do serviço. Termos como “dispositivos”, “localização”, “privacidade”, “ajustes” e “reconhecimento” aparecem em tamanho intermediário, indicando sua relevância secundária no texto. Particularmente interessante é a presença proeminente de palavras técnicas e relacionadas a configurações, como “transcrições”, “processamento”, “identificador” e “aleatório”, o que corrobora com os resultados obtidos nas análises de leiturabilidade anteriores, que apontaram alta complexidade textual. Esta distribuição visual das palavras reforça a constatação de que o documento prioriza uma linguagem técnica e potencialmente inacessível para usuários sem formação especializada, em detrimento de uma comunicação mais clara e direta sobre o uso e as implicações do recurso.

Figura 5 - Nuvem de palavras com os termos mais repetidos ao longo do documento



Fonte: “Termo de uso da Siri”, ALT. Elaborado pelas autoras.

É evidente que a complexidade textual dos termos de uso da Siri representa um obstáculo significativo à transparência e ao exercício consciente dos direitos dos usuários. A predominância de linguagem técnica, confirmada tanto pelos índices de leiturabilidade quanto pela análise da nuvem de palavras, sugere a necessidade de uma reformulação desses documentos para torná-los mais acessíveis ao público geral, especialmente considerando a ampla base de usuários da Siri, que inclui diferentes faixas etárias e, conseqüentemente, diferentes níveis de escolaridade (Piper Sandler, 2024). A simplificação dessa documentação é, portanto, uma questão de letramento de IA e a garantia de que os usuários compreendam plenamente as implicações do uso dessa tecnologia.

Sobre ofuscação informacional e falta de transparência nos documentos das Assistentes Virtuais Inteligentes

Ao navegar pela aba 'Informações legais', localizada em um menu vertical dos Termos de Uso da Alexa, três páginas estão completamente em branco: “Sobre as Funcionalidades Infantis”, “Diretrizes de Uso do Alexa Communication” e “Notificação de Teste de

Funcionalidade da Alexa”. Além disso, na seção “Ajuda”, apenas os “Termos e Condições de Lembretes da Alexa” estão preenchidos, enquanto tópicos relevantes, como as “Diretrizes de Uso de Ligações e Mensagens Alexa” - que deveriam esclarecer como a assistente lida com chamadas e mensagens - permanecem vazios. A ausência desse tipo de conteúdo, nas seções consideradas essenciais, levanta questionamentos sobre a transparência das informações constantes no documento, que já se mostra de difícil leitura devido à sua extensão e pulverização ao longo do site e reforça a sensação de que, apesar da clareza textual aparente, a leiturabilidade não se traduz em um acesso real às informações fundamentais⁷.

É essencial destacar os impactos da possível invasão de privacidade, por parte da Amazon, na utilização de dados pessoais para modelagem de respostas da assistente e treinamento de modelos e/ou recursos de Inteligência Artificial. Ao longo do documento “Notificação de Privacidade da Amazon”, a empresa aborda que a Alexa melhora o seu serviço a partir do uso, por parte de quem a consome, seguindo um processo de personalização (Amazon, 2020). Por se concentrar em oferecer respostas agradáveis, a Alexa acaba moldando as respostas de acordo com o perfil do indivíduo – construído através da datificação de suas características (Van Dijck, 2014) –, não o desafiando ou oferecendo perspectivas alternativas, a menos que isso seja explicitamente solicitado. Assim, suas respostas podem seguir estereótipos ou abordar as questões de forma limitada por conta do perfil do usuário ou dos dados com os quais aprendeu a acessar, reforçando sem intenção explícita preconceitos e estigmas estruturalmente presentes na sociedade⁸. Isso não só compromete a privacidade e mostra-se difícil de medir os impactos, mas também gera o risco de criar bolhas informativas, prejudicando a formação crítica e o

⁷ Ver documentação: Writing for the Web. Disponível em:

[https://faculty.washington.edu/farkas/HCDE510-](https://faculty.washington.edu/farkas/HCDE510-Fall2012/RedishLettingGoOfWordsIntercom2004DFAnn.pdf)

[Fall2012/RedishLettingGoOfWordsIntercom2004DFAnn.pdf](https://faculty.washington.edu/farkas/HCDE510-Fall2012/RedishLettingGoOfWordsIntercom2004DFAnn.pdf). Acesso em: 28 fev 2025.

⁸ Ver documentação: O que têm em comum Alexa, Siri, Lu e Bia? Assistentes digitais, sexismo e rupturas de performances de gênero. Disponível em:

<https://www.scielo.br/j/gal/a/Ht93VJ6KQRjQNBvXcGsxY9g/#:~:text=Assistentes%20digitais%2C%20sexismo%20e%20rupturas%20de%20performances%20de%20g%C3%AAnero,-1>. Acesso em:

30 janeiro 2025.

discernimento social. Quando esses objetos são utilizados em ambientes institucionais e/ou escolares, como por exemplo escolas, esse panorama fica ainda mais latente⁹.

Como afirmado anteriormente, mesmo conectada a pelo menos 17 milhões de dispositivos domésticos, a Alexa é utilizada, em grande parte, por pessoas que provavelmente não têm pleno letramento e desconhecem seus impactos ou funções, assim como de grande parte das tecnologias presentes em seu cotidiano. Essa falta de transparência não é apenas um problema de comunicação, mas reflete um modelo de governança algorítmica no qual as decisões automatizadas são justificadas pela Amazon e os direitos do usuário são secundarizados¹⁰. Como discutido por Rosetti e Angeluci (2021), essa prática reforça uma assimetria de poder, em que a empresa estrutura seus termos uso e documentos, voltados para o usuário, de modo a garantir sua liberdade sobre os dados dos usuários-consumidores e sua agência acerca do conhecimento fornecido ao público, sem oferecer alternativas para a compreensão de como essas informações são utilizadas ao longo de todas as suas 106 páginas¹¹.

No caso da Siri, observa-se um padrão semelhante de ofuscação informacional, porém com características distintas. Os “Termos de uso” da assistente virtual da Apple encontram-se diluídos contando com apenas uma aba dentro do extenso site da empresa, o que dificulta o acesso direto às informações sobre privacidade de dados e utilização dos rastros digitais deixados pelos seus usuários. O documento aborda detalhes sobre o processamento de dados, o texto é denso e repleto de tecnicismos que podem confundir o usuário com baixo letramento digital. A fragmentação das informações sobre privacidade, com constantes redirecionamentos para outros *links*, cria barreiras adicionais à compreensão integral de como a Siri e a Apple utilizam e armazenam os

⁹ As Dualidades entre o Uso da Inteligência Artificial na Educação e os Riscos de Vieses Algorítmicos. Disponível em: <https://www.scielo.br/j/es/a/qrTryFvZR9Y9WsRpG5fWGHB/>. Acesso em: 23 fev 2025.

¹⁰ Ver documentação: Ética Algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação. Disponível em: <https://www.scielo.br/j/gal/a/R9F45HyqFZMpOp9BGTfZnyr>. Acesso em: 1 fev 2025.

¹¹ No documento Projeto de Lei (PL) 2338/2023, que propõe uma regulação de IA no Brasil, não existe nenhuma referência a regulação de assistentes virtuais. Ver: Bentes, Sanches e Ferreira (2024).

dados pessoais, como as transcrições de áudio, os contatos e as localizações. Essa estrutura informacional, ainda que aparentemente transparente ao listar as práticas de coleta, acaba reproduzindo a mesma assimetria de poder entre empresa e consumidor como observado na Alexa, onde o usuário raramente compreende completamente as implicações do uso cotidiano da tecnologia.

Apesar dos Termos de uso e privacidade de dados da Siri mencionar medidas de proteção, como a associação dos dados a um identificador aleatório em vez da Conta Apple, é admitido que as transcrições das interações podem ser armazenadas e revisadas por um período indefinido com o objetivo de aprimorar as ações e operações da assistente. Além disso, a Siri coleta informações contextuais de aplicativos de terceiros e do próprio dispositivo Apple, como contatos, histórico de músicas e de preferências, sem esclarecer de forma clara os critérios para essa retenção e o consequente impacto na privacidade do usuário.

É enfatizado, pela Apple, que os usuários têm controle sobre o que compartilham, oferecendo opções de exclusão e ajustes de privacidade, mas a complexidade das configurações e a necessidade de navegação aprofundada nos “Ajustes” podem dificultar esse controle na prática¹². Assim, embora a empresa tenha um discurso voltado à privacidade do usuário, a falta de uma apresentação mais acessível e estruturada de seus documentos pode limitar a compreensão sobre a extensão do monitoramento e do uso de seus dados¹³.

¹² Apple. Termos de usos Siri. Disponível em: <https://www.apple.com/br/legal/privacy/data/pt-br/ask-siri-dictation> . Acesso em: 15/10/2024.

¹³ Ver: Seção “Siri, Ditado e Privacidade”. Disponível em: <https://www.apple.com/br/legal/privacy/data/pt-br/ask-siri-dictation/#:~:text=Acesse%20Ajustes%20%3E%20Siri%20%3E%20Apps%20%3E,%2C%20voc%C3%AA%20pode%20desativ%C3%A1%E2%80%91las>. Acesso em: 11 mar 2025.

Conclusão

Em meio aos avanços dos sistemas, algoritmos e modelos de IA, as assistentes virtuais têm mudado significativamente ao longo dos anos, transformando-se de um mecanismo com funções básicas em um sistema sofisticado profundamente integrado a diferentes dispositivos. No entanto, em contraste com este avanço tecnológico, os termos de uso e documentação desses artefatos apresentam uma complexidade textual que merece atenção especial, no que diz respeito à sua clareza, entendimento e responsabilidade acerca das informações sobre proteção e uso dos dados pessoais dos usuários.

Neste artigo, buscamos observar os documentos referentes à Alexa e Siri no que tange a privacidade dos dados dos seus usuários e constatamos que estes artefatos não se configuram apenas como facilitadores de acesso informacional, mas também trazem à tona desafios éticos e jurídicos associados ao seu uso cotidiano. A análise dos “Termos de Uso”, por exemplo, revelou a falta de transparência e a fragmentação das informações, o que demonstra como a governança algorítmica pode reforçar desigualdades de poder entre a empresa e o seu usuário-consumidor. Esse cenário dificulta o entendimento das práticas de coleta e uso dos dados, tornando mais difíceis de serem entendidos os direitos digitais dos consumidores e os impactos de tais tecnologias no cotidiano de cada usuário. Com a crescente integração dos assistentes virtuais inteligentes em milhares de domicílios e instituições brasileiras, fica evidente a urgência de um marco regulatório mais robusto, que, além de promover a inovação tecnológica, também assegure que os usuários tenham acesso claro e transparente às informações (BENTES, SANCHES e FONSECA, 2024).

Como abordado por Regina Rossetti e Alan Angeluci (2021), a privacidade e a responsabilidade são duas das sete questões éticas centrais relacionadas à presença de algoritmos, pois levantam discussões acerca da proteção dos dados pessoais e a quem cabe a responsabilidade pelas ações dos mesmos. Nesse caso, quem seria responsável pelo resultado informacional concedido pelas Assistentes? O CEO ou os acionistas da empresa

proprietária desse sistema de inteligência artificial? O desenvolvedor/funcionário responsável pela construção e pelo algoritmo que acessa a informação da web? O usuário-consumidor que fez a pergunta e circulou as respostas concedidas em outros fóruns? A máquina, no caso, a própria assistente?

Essa lacuna de responsabilidade cria um ambiente de insegurança informacional onde todos, enquanto sociedade, estamos sujeitos a indagar sobre a verdade dos fatos e do que ouvimos e vemos. A desconfiança informacional e o questionamento acerca da verdade, cria uma espécie de Realidade Paralela (2023) onde não sabemos distinguir mais o que é material e o que é sintético. Somado a isso, as empresas de tecnologia têm adotado um discurso ambíguo sobre os direitos dos usuários no uso de seus serviços e produtos, apontando que protegem os dados dos usuários - anonimizando-os - mas utilizam as informações para aprimoramento dos seus sistemas e *gadgets*¹⁴. Mesmo com o avanço de regulamentações, sobretudo na Europa, algo que vem sendo cobrado acerca das BigTechs, muitas companhias ainda parecem relutantes em comunicar de forma clara as implicações de suas práticas de coleta e uso das informações pessoais dos seus usuários. Essa postura cria uma hierarquia de conhecimento entre a empresa e o indivíduo, no que tange ao que se é utilizado, dificultando que os usuários tomem decisões informadas sobre as informações que estão instruindo as suas assistentes. Uma abordagem mais transparente e centrada no respeito aos direitos digitais do usuário, poderia contribuir para o estabelecimento de relações de confiança mútua e consciência dos usos que a empresa faz com as informações pessoais dos seus consumidores.

Ao indagarmos quais informações estão de fato públicas para o usuário notamos uma sequência de informações fragmentadas, com pouca clareza, mas a afirmação de que os dados são utilizados para personalização e melhoria de desempenho das assistentes. Ao

¹⁴ A ambiguidade do discurso se revela nos próprios documentos da empresa quando, por exemplo, a Apple afirma que se preocupa com os dados e informações do usuário, assegurando anonimato ao mesmo, porém utiliza seus dados para treinamento dos modelos e utilização de melhoria do mecanismo de respostas da Siri. Ver: Apple. Termos de uso Siri. Disponível em: <https://www.apple.com/br/legal/privacy/data/pt-br/ask-siri-dictation> . Acesso em: 15/10/2024.

utilizarmos a metodologia de análise de leiturabilidade, a partir do software ATL, constatamos que o grau de leiturabilidade dos documentos referentes à Alexa e à Siri são diferentes. A Alexa tem leiturabilidade alta e possibilita a leitura de seus termos por qualquer pessoa alfabetizada, porém com a documentação fragmentada ao longo de dezenas de “abas” do site da Amazon e em diferentes hyperlinks, contando cento e seis páginas, fica quase impossível para um usuário não pesquisador a compreensão de todos os documentos. No caso da Siri, a disposição dos documentos é facilitada e disponibilizada de forma sintética, acessível para ser encontrada pelo usuário, no entanto o grau de leiturabilidade de seus documentos é alto, um tanto quanto complexo, o que dificulta a leitura por parte dos indivíduos-clientes. Esses resultados, como vimos, apontam que os utilizadores dessas assistentes virtuais inteligentes ou AVIs, precisam ter um letramento digital, e, também midiático, elevado para a compreensão de todas as nuances circundantes a esses textos (Azevedo et al., 2018).

No entanto, como se sabe a maioria dos usuários, mesmo utilizadores assíduos de produtos tecnológicos, não possuem capacidade analítica, compreensão dos ambientes digitais e de como são estruturadas as arquiteturas computacionais das Bigtechs. Essa superficialidade no uso tecnológico revela que entender os ambientes e artefatos digitais vai muito além de saber navegar em interfaces ou utilizar dispositivos; significa compreender as possíveis implicações sociais e jurídicas dos espaços virtuais. Assim, a aceitação automática dos “Termos de uso” das assistentes, bem como a sua utilização, sem qualquer tipo de exame crítico, não é um ato de familiaridade tecnológica, mas sim uma demonstração da ausência de letramento digital que capacite os usuários a entenderem criticamente os contextos e as condições a que estão se submetendo.

Referências

- ALPAYDIN, E. **Machine Learning**. Cambridge, Massachussets: The MIT Press, 2016.
- AZEVEDO, D. S. et al. **Letramento digital: uma reflexão sobre o mito dos "nativos digitais"**. Novas Tecnologias na Educação, v. 16, n. 2, 2018.
- Bentes, A.; SANCHES, D.; FONSECA, P. **Assistentes Virtuais Inteligentes e saúde mental: debates regulatórios no Brasil**. Reciis: Revista Eletrônica de Comunicação, Informação & Inovação em Saúde, v. 18, n. 3, 2024. p. 538-553.
- BOYD, D.; CRAWFORD, K. **Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon**. Information, Communication & Society, vol. 15, n. 5, 2012. p. 662–679.
- CARDENUTO, J. P. et al. **The Age of Synthetic Realities: Challenges and Opportunities**. APSIPA Trans. on Signal and Information Processing, 2023.
- CELLARD, A. A análise documental. In: POUPART, J. et al. **A pesquisa qualitativa: enfoques epistemológicos e metodológicos**. Petrópolis, Vozes, 2008.
- CETIC.BR. **Pesquisa TIC Domicílios 2023: Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros**. São Paulo: Comitê Gestor da Internet no Brasil, 2023. Disponível em: <https://www.cetic.br>. Acesso em: 10 mar. 2025.
- CHUNG, H. et al. **"Alexa, Can I Trust You?"**. Computer, vol. 50, n. 9, 2017. p. 100-104.
- DELEUZE, G. **Postscript on the Societies of Control**. October, 59, 1992. p. 3-7.
- DUBAY, W. H. **Smart Language: Readers, Readability, and The Grading of Texts**. Costa Mesa, CA: Impact Information, 2007.
- FOUCAULT, M. **Vigiar e Punir: nascimento da prisão**. Ramallete. Petrópolis, Vozes, 1987.
- GILLESPIE, T. # trendingistrending: When algorithms become culture. In: SEYFERT, R.; ROBERGE, J. (Orgs.). **Algorithmic cultures: essays on meaning, performance and new technologies**. 1 Edition. London ; New York: Routledge, Taylor & Francis Group, 2016, p. 52–75.
- GUNKEL, D. J. Comunicação e inteligência artificial: novos desafios e oportunidades para a pesquisa em comunicação. Galaxia, n. 34, jan-abr., 2017. p. 05-19.
- HEDMAN, J.; SRINIVASAN, N.; LINDGREN, R.. **Digital Traces of Information Systems: Sociomateriality Made Researchable**. In: Proceedings of the 34th International

Conference on Information Systems. ICIS 2013Design. Università Bocconi, Milan, Italy: Association for Information Systems. AIS Electronic Library (AISeL), 2013. Disponível em: <https://research.cbs.dk/en/publications/digital-traces-of-information-systems-sociomateriality-made-resea>. Acesso em: 26 abr. 2022. p. 1069

HOLTZ, L.; SANTOS, O. M. **Legibilidade das notas explicativas das empresas brasileiras de capital aberto**. Enfoque: Reflexão Contábil, v. 39, n. 1, 2020. p. 57-73.

ILUMEO. **Assistentes virtuais por voz em 2025**. São Paulo: Ilumeo, 2025. Disponível em: <https://ilumeo.com.br/estudos/assistentes-de-voz/>. Acesso em: 27 fev 2025.

LATOUR, B. **Reagregando o social**: uma introdução à teoria do ator-rede. Salvador-Bauru: EDUFBA-EDUSC, 2012.

MORENO, G. C. L et al. **ALT**: um software para análise de legibilidade de textos em língua portuguesa. Policromias: Revista de Estudos do Discurso, Imagem e Som, vol. 8, n. 1, 2023. p. 91-128.

NASS, C., e MOON, Y. **Machines and Mindlessness**: Social Responses to Computers. Journal of Social Studies, vol. 56, n.1, 2000. p. 81-103.

NASS, C., STEUER, J., e TAUBER, E. R. **Computers are Social Actors**. Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, New York, 1994.

OMENA, J. **Métodos digitais**: teoria-prática-crítica. Lisboa: Livros ICNOVA, 2019.

PIPER SANDLER. **Taking Stock with Teens Survey**. Minneapolis: Piper Sandler Ed., 2024. Disponível em: <https://www.pipersandler.com/teens>. Acesso em: 10 mar. 2025.

ROSSETTI, R.; ANGELUCI, A. **Ética Algorítmica**: questões e desafios éticos do avanço tecnológico da sociedade da informação. Galáxia, São Paulo, n. 46, 2021. p. 1-18.

SANTOS, L. Alexa, você é uma mulher?": gênero e tecnologia no caso das assistentes digitais. Revista Fronteiras – Estudos Midiáticos, vol. 24, n. 3, 2022. p. 35-46.

SÁ-SILVA, J. R.; ALMEIDA, C. D. de; GUINDANI, J. F. **Pesquisa documental**: pistas teóricas e metodológicas. Revista Brasileira de História & Ciências Sociais, Rio Grande, v. 1, n. 1, 2009. p.1-15.

SMITH, M.; TAFFLER, R. **Readability and understandability**: differents measures of the textual complexity of accounting narrative. Accounting, Auditing, Accountability Journal, vol. 5, n. 4, 1992. p. 84-98.

SOUZA, J.; KANTORSKI, L. P.; LUIS, M. A. V. **Análise documental e observação participante na pesquisa em saúde mental**. Revista Baiana de Enfermagem, Salvador, v. 25, n. 2, 2011. p. 221-228.

TUDOCELULAR. **iPhone: 9 a cada 10 adolescentes nos EUA possuem um;** IA motiva compra do modelo mais recente, diz pesquisa. Disponível em: <https://www.tudocelular.com/mercado/noticias/n227150/iphone-9-cada-10-adolescentes-eua-ia-motiva-compra.html>. Acesso em: 14 fev 2025.

UNESCO. **I'd Blush If I Could:** Closing Gender Divides in Digital Skills through Education. Paris: UNESCO, EQUALS Skills Coalition, 2019.

VAN DIJCK, J. **Datafication, dataism and dataveillance:** Big Data between scientific paradigm and ideology. *Big Data Surveillance*, vol. 12, n.2, 2014.

ZACCARELLI, S. **A utilização de assistentes virtuais no apoio ao aprendizado e melhoria da experiência do aluno.** AWS Blog Brasil, 02 jul. 2021. Disponível em: <https://aws.amazon.com/pt/blogs/aws-brasil/a-utilizacao-de-assistentes-virtuais-no-apoio-ao-aprendizado-e-melhoria-da-experiencia-do-aluno/>. Acesso em: 10 mar 2025.

ZUBOFF, S. **A era do capitalismo de vigilância:** a luta por um futuro humano na nova fronteira de poder. Rio de Janeiro: Intrínseca, 2020.