

Assistentes Virtuais Inteligentes e saúde mental: debates regulatórios no Brasil

Intelligent Virtual Assistants and mental health: regulatory debates in Brazil

Asistentes Virtuales Inteligentes y salud mental: debates regulatorios en Brasil

Anna Bentes^{1,a}

annabentes@gmail.com | <https://orcid.org/0000-0003-4167-6093>

Danielle Sanches^{1,b}

danielle.sanchesx@gmail.com | <https://orcid.org/0000-0001-9916-167X>

Paulo Fonseca^{2,c}

pfonseca@ufba.br | <https://orcid.org/0000-0003-3403-0753>

¹ Fundação Getulio Vargas, Escola de Comunicação, Mídia e Informação da Fundação Getulio Vargas. Rio de Janeiro, RJ, Brasil.

² Universidade Federal da Bahia, Instituto de Ciência, Tecnologia e Inovação. Camaçari, BA, Brasil.

^a Doutorado em Comunicação e Cultura pela Universidade Federal do Rio de Janeiro.

^b Doutorado em História das Ciências pela École des Hautes Études en Sciences Sociales e Fundação Oswaldo Cruz.

^c Doutorado em Sociologia pela Universidade de Coimbra.

RESUMO

A inserção dos Assistentes Virtuais Inteligentes na vida cotidiana representa um marco na história da comunicação entre humanos e máquinas. Devido às suas características interativas, estes estão sendo cada vez mais apropriados e desenvolvidos para fins de cuidado, especialmente no âmbito da saúde mental. Este artigo visa compreender se e como o debate regulatório brasileiro oferece instrumentos para lidar com os desafios e as preocupações desses sistemas de Inteligência Artificial em relação à saúde mental. A partir de uma análise documental, mapeamos exemplos de aplicação dos Assistentes Virtuais Inteligentes em saúde mental, a fim de identificar riscos a direitos dos usuários e avaliar, na legislação brasileira vigente e em discussão, se há proteção suficiente para lidar com eles. Por meio de uma abordagem crítica, salientamos a insuficiência da legislação brasileira atual e a necessidade de ampliação do debate sobre como equilibrar possíveis riscos e benefícios dessas tecnologias.

Palavras-chave: Assistentes Virtuais Inteligentes; Inteligência Artificial; Saúde mental; Regulação.

ABSTRACT

The integration of Intelligent Virtual Assistants into everyday life marks a milestone in the history of human-machine communication. Due to their interactive characteristics, they are increasingly being appropriated and developed for caregiving purposes, especially in the field of mental health. This article aims to understand whether and how the Brazilian regulatory debate provides tools to address the challenges and concerns of these Artificial Intelligence systems concerning mental health. Through a document analysis, we map examples of Intelligent Virtual Assistants's applications to mental health to identify risks to users' rights and evaluate whether the current and the proposed Brazilian legislation offer sufficient protection to address these risks. Through a critical approach, we highlight the inadequacy of current Brazilian legislation and the need to expand the debate on how to balance the potential risks and benefits of these technologies.

Keywords: Intelligent Virtual Assistants; Artificial Intelligence; Mental health; Regulation.

RESUMEN

La inserción de los Asistentes Virtuales Inteligentes en la vida cotidiana representa un hito en la historia de la comunicación entre humanos y máquinas. Debido a sus características interactivas, cada vez son más apropiados y desarrollados para fines de cuidado. Este artículo tiene como objetivo comprender si y cómo el debate regulatorio brasileño ofrece instrumentos para abordar los desafíos y preocupaciones de estos sistemas de Inteligencia Artificial en relación con la salud mental. A partir de un análisis documental, mapeamos ejemplos de la aplicación de los Asistentes Virtuales Inteligentes a la salud mental, con el fin de identificar riesgos para los derechos de los usuarios y evaluar, en la legislación brasileña vigente y en discusión, si hay protección suficiente para abordarlos. Destacamos la insuficiencia de la legislación brasileña actual y la necesidad de ampliar el debate sobre cómo equilibrar los posibles riesgos y beneficios de estas tecnologías.

Palabras clave: Asistentes Virtuales Inteligentes; Inteligencia Artificial; Salud mental; Regulación.

INFORMAÇÕES DO ARTIGO

Contribuição dos autores:

Concepção e desenho do estudo: Anna Bentes, Danielle Sanches.

Coleta de dados: Anna Bentes, Danielle Sanches, Paulo Fonseca.

Análise dos dados: Anna Bentes, Danielle Sanches, Paulo Fonseca.

Interpretação dos dados: Anna Bentes, Danielle Sanches, Paulo Fonseca.

Todos os autores são responsáveis pela redação e revisão crítica do conteúdo intelectual do texto, pela versão final publicada e por todos os aspectos legais e científicos relacionados à exatidão e à integridade do estudo.

Declaração de conflito de interesses: não há.

Fontes de financiamento: Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (Faperj).

Considerações éticas: não há.

Agradecimentos/Contribuições adicionais: não há.

Histórico do artigo: submetido: 2 abr. 2024 | aceito: 4 jul. 2024 | publicado: 27 set. 2024.

Apresentação anterior: não houve.

Licença CC BY-NC atribuição não comercial. Com essa licença é permitido acessar, baixar (*download*), copiar, imprimir, compartilhar, reutilizar e distribuir os artigos, desde que para uso não comercial e com a citação da fonte, conferindo os devidos créditos de autoria e menção à Reciis. Nesses casos, nenhuma permissão é necessária por parte dos autores ou dos editores.

INTRODUÇÃO: ASSISTENTES VIRTUAIS E A INTELIGÊNCIA ARTIFICIAL

Os Assistentes Virtuais Inteligentes (AVI) ativados por voz e por texto estão na fronteira da comunicação contemporânea. Sistemas como Alexa (da Amazon), Assistente Google ou Siri (da Apple) são agentes de *software* de conversação concebidos para atuarem como interlocutores digitais, mudando a nossa relação com os meios de comunicação e com a nossa própria representação como sujeitos. Contudo, mais do que simples canais ou *gadgets*, essas tecnologias são artefatos culturais bastante opacos em seu funcionamento, seus riscos e efeitos socioculturais e subjetivos.

Os AVI são dispositivos híbridos que combinam tecnologias avançadas de Processamento de Linguagem Natural (PLN) para compreender e responder aos comandos por voz ou por texto. São técnicas de aprendizado de máquina para realizar e melhorar as respostas ao longo do tempo e promover a integração com outros serviços digitais para executar tarefas diversas, desde definir alarmes até controlar dispositivos domésticos inteligentes. Essas funcionalidades variam conforme a plataforma e a integração com outros sistemas e aplicativos, proporcionando uma experiência personalizada e adaptativa para os usuários. Desse modo, ao combinar diferentes técnicas de Inteligência Artificial (IA), os AVI são capazes de entender, interpretar e reagir à linguagem humana, mas também de aprender, a partir de dados e de suas formas de interação.

A capacidade dos AVI de interagir, por meio de linguagem natural, com um ecossistema tecnológico mais amplo os torna ferramentas versáteis que podem se inserir na rotina diária dos usuários. Os AVI já estão presentes em casas (por exemplo, em caixas de som, TVs e geladeiras interativas), em veículos (integrados em sistemas de navegação), em pulsos (*smartwatches*) e, é claro, nos bolsos (*smartphones* ou *tablets*).

Eles são uma camada constitutiva da vida cotidiana de uma parcela crescente da população brasileira e materializam a projeção de uma Internet das Coisas (IoT) com a integração em rede do mundo físico com trocas de dados digitais. Recentemente, o desenvolvimento dos Grandes Modelos de Linguagem (LLM), como o GPT, levou a um aprimoramento significativo das capacidades de compreensão e execução de comandos do AVI, oferecendo respostas mais antropomorfizadas e permitindo interações cada vez mais complexas.

Embora esse tipo de tecnologia ainda não seja tão popular quanto o uso convencional de *smartphones* e redes sociais, sua presença vem crescendo e se popularizando no Brasil. De acordo com dados divulgados pela Forbes, o uso da Alexa com outros dispositivos inteligentes cresceu 50% no Brasil em 2023 (Forbes, 2024). Segundo uma pesquisa da consultoria alemã Statista, divulgada pela Folha de S.Paulo, em 2020, cerca de 3% dos lares brasileiros contavam com algum dispositivo inteligente. Contudo, a pesquisa prevê que o mercado brasileiro desse tipo de tecnologia tem potencial para crescer 38,4% ao ano, até 2025, alcançando cerca de 11% das moradias (Lemos, 2021). Em um estudo sobre a utilização de AVI no Brasil realizado em 2023 pela Ilumeo, consultoria especializada em Data Science e sistemas de IA, 92% das pessoas afirmaram ter algum conhecimento sobre AVI e, dessas, 87% afirmaram já ter usado assistente virtual comandado por voz. Nessa mesma pesquisa, 48% das pessoas afirmaram que já utilizam esse tipo de tecnologia semanalmente (Ilumeo, 2023).

Embora o uso de AVI esteja se expandindo no Brasil, é preciso reconhecer que seu acesso ainda é bastante limitado e desigual, assim como o de outras tecnologias no país. Isso se deve não somente pelo custo desses dispositivos, mas também pela desigualdade no acesso à internet. Conforme mostra a pesquisa TIC Domicílios de 2022, cerca de 80% da população brasileira tem acesso à internet (CGI.br, 2023), porém 62% dessas conexões se dão exclusivamente por meio de dispositivo móvel e, em muitos casos, a partir de planos de dados limitados, o que restringe o acesso efetivo à diversidade de informações disponíveis *on-line*.

Além disso, é preciso ter em vista que tecnologias baseadas em IA se alimentam de processos de datificação (Mejias; Couldry, 2019), ou seja, da transformação da vida humana em dados, de modo que esses dados possam ser quantificados e processados por sistemas automatizados em larga escala e estar

associados a processos de geração de valor. Mejias e Couldry (2019) chamam atenção para o fato de que a geopolítica da datificação mantém uma conexão histórica com a lógica colonial, na qual a apropriação da vida humana, na lógica do extrativismo de dados, beneficia especialmente interesses ocidentais do Norte global e dos capitalistas. Portanto, os AVI também se inserem nessas assimetrias e dinâmicas de poder que envolvem tecnologias de IA, grandes corporações de tecnologia e seus processos de datificação.

O desenvolvimento dos AVI integra uma paisagem ampla e múltipla de diversas tecnologias baseadas em IA que nos inserem em uma “era dos algoritmos” (Domingos, 2017). Contudo, é preciso considerar, como ressalta Bucher (2018), que os algoritmos são múltiplos, tanto no sentido técnico (existem diferentes tipos de algoritmos) quanto no sentido social, ou seja, existem diferentes sentidos para o que chamamos de algoritmo. Em outras palavras, não só, tecnicamente, existem diferentes tipos de algoritmos e de IA, como, socialmente, eles mobilizam diferentes preocupações, concepções e imaginações. Tal multiplicidade nos convoca a pensar a IA para além de concepções binárias sobre a tecnologia e a buscar concepções que abarquem sua complexidade sociotécnica e suas controvérsias. Por isso, quando falamos de AVI ou de IA no singular, ao longo deste artigo, é preciso ter essa multiplicidade em vista.

Enquanto sistemas de IA, capazes de interação, os AVI emergem como uma interface humanamente amigável dos sistemas maquínicos, expandindo as relações e formas de comunicação entre humano e máquina. Por suas capacidades interativas, que simulam as características humanas de comunicação, eles permitem uma aproximação dos humanos com as máquinas, a qual pode ser até emocional. Por isso, tanto desenvolvedores quanto usuários vêm reconhecendo e investindo no potencial dessas tecnologias para o campo dos cuidados, da saúde e, em especial, da saúde mental. As interseções entre o funcionamento dessas máquinas e a saúde mental humana são diversas, indo desde determinados modos de aplicação voltados aos cuidados com a saúde (como uso de *chatbots* para psicoterapia ou orientações de saúde), processamento de dados sensíveis (voz, emoções, personalidade, entre outras informações sobre estados mentais), até apropriações de uso por usuários para fins terapêuticos (uso desses sistemas para aconselhamentos ou finalidades diversas). É essa interseção entre AVI e saúde mental que buscaremos explorar neste artigo.

Os usos dos AVI, especialmente os voltados à saúde mental, podem trazer riscos a uma série de direitos dos usuários, como a privacidade, a proteção de dados, a não discriminação e a autodeterminação. Embora o Brasil tenha um arcabouço legal ligado aos serviços baseados em internet (Marco Civil da Internet – MCI) e também à proteção de dados, os debates sobre a regulação de IA ainda estão avançando aos poucos. Entre diferentes propostas no âmbito legislativo, o projeto de lei 2.338/2023 (Brasil, 2023) ganhou destaque por ter sido uma das propostas mais robustas apresentadas até agora, resultado do trabalho de uma comissão de juristas convocada pelo Senado Federal. Diante da legislação vigente e da proposta de lei em análise, o debate regulatório brasileiro oferece instrumentos para lidar com os desafios e as preocupações desses sistemas de IA em relação à saúde mental?

Tendo em vista essa questão, este artigo busca discutir as complexas relações entre os AVI, a saúde mental e a regulação brasileira, mapeando alguns dos riscos aos usuários ligados ao uso direto ou indireto dessas máquinas interativas para a saúde mental. Busca, também, questionar se e como o debate regulatório oferece instrumentos para lidar com as transformações e os riscos que tecnologias, como os AVI, podem acarretar. Uma legislação adequada poderia contribuir para uma política de governança de dados baseada na transparência e na responsabilidade públicas, para que os usuários tenham uma compreensão clara da proveniência das informações que buscam ou fornecem às empresas desenvolvedoras dessas tecnologias, sobretudo quando consideramos que essas máquinas constituem um tipo particular de sistema algoritmo que, à sua maneira, simula as capacidades humanas de comunicação, apresentando um imenso potencial de confundir e enganar seus usuários. Os AVI ainda são caixas-pretas para os dados sociais, e, como essas práticas

são ainda pouco estudadas, as fontes de informação utilizadas por eles estão mais próximas de hieróglifos do que de barômetros sociais (Gillespie, 2018).

Para desenvolver tais reflexões, este artigo está estruturado em três seções, além desta introdução e da conclusão. A seguir, discutimos as escolhas metodológicas, buscando elucidar como os percursos escolhidos podem conduzir ao alcance dos objetivos almejados. Na seção seguinte, destacaremos, a partir de algumas aplicações práticas dos AVI, implicações relacionadas à saúde mental e como isso envolve questionamentos sobre o funcionamento desses sistemas, os modelos de negócio de empresas de tecnologia, a privacidade e o uso de dados de usuários. Na quarta seção, discutiremos os marcos legais envolvidos na regulação dos AVI, com especial destaque aos debates recentes sobre o projeto de lei de IA no Brasil. Por fim, nas conclusões, apresentaremos algumas reflexões críticas sobre as limitações da regulação de AVI e, mais especificamente, de seu caráter sensível à saúde mental e à subjetividade de seus usuários, enquanto tecnologias que se propõem interativas.

METODOLOGIA

Neste trabalho, utilizamos uma metodologia qualitativa para identificar e compreender os possíveis impactos dos AVI na saúde mental dos usuários, bem como a situação atual e os desafios da regulação sobre eles para lidar com seus riscos e implicações. Sabemos que as análises qualitativas não possibilitam generalizações, mas elas são capazes de fornecer informações ricas e minuciosas sobre determinado grupo, assunto, objeto ou caso, permitindo uma compreensão dos fatos e acontecimentos, a partir de uma determinada perspectiva, de quem analisa, ou do que está sendo observado – o que promove um entendimento profundo e detalhado (Souza, Kantorski, Luis, 2011). Para tanto, escolhemos como procedimento metodológico a análise documental, utilizando, dessa maneira, técnicas para observação de diferentes tipos de documentos (Sá-Silva; Almeida; Guindani, 2009), como projetos de lei, legislações, termos de uso de AVI¹ e de casos de aplicação de AVI voltados à saúde mental e divulgados em meios de comunicação e em outras pesquisas. Nossa intenção foi, inicialmente, mapear possíveis implicações das interações entre AVI e humanos para a saúde mental para, então, a partir disso, compreender como elas se dão de fato e se o arcabouço regulatório brasileiro seria suficiente para proteger possíveis riscos mapeados.

Vale ressaltar que a análise documental permite a identificação, verificação e apreciação dos documentos com finalidade específica (Oliveira, 2007). Nossa escolha seguiu os pressupostos metodológicos descritos por Maria Cecília Minayo (2008), que compreende que a análise documental deve extrair informações objetivas dos documentos com base em questões e hipóteses relevantes, fazendo com que o documento selecionado seja o próprio objeto de estudo.

Assim, nossa pesquisa se dedicou a um mapeamento preliminar de exemplos e casos de aplicação de AVI em saúde mental, visando analisar, sob uma perspectiva crítica, potenciais riscos à saúde mental e aos direitos dos usuários. Em seguida, definimos o escopo de três legislações a serem analisadas. Trata-se dos principais marcos legais voltados aos serviços digitais, especialmente aqueles baseados em IA, quais sejam: 1) o MCI, lei 12.965/2014 (Brasil, 2014), que estabelece princípios, garantias, direitos e deveres para o uso da internet no Brasil; 2) a Lei Geral de Proteção de Dados Pessoais (LGPD), lei 13.709/2018 (Brasil, 2018), que dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais; e 3) o projeto de lei 2.338/2023 (Brasil, 2023), que dispõe sobre o uso da IA.

A partir de uma leitura acurada das informações e da extração de passagens, realizamos uma análise para avaliar nossa principal hipótese de pesquisa: a de que a regulação vigente e em debate relacionada aos AVI no Brasil não contempla adequadamente os desafios que se impõem para a saúde mental de seus usuários.

¹ Focamos especificamente na Alexa, da Amazon, por ela apresentar uma série documental extensa e disponível no seu ambiente digital (Amazon, c2021-2024).

Outra fonte de consulta para a nossa pesquisa foi a “Ajuda e serviço de atendimento ao cliente”, da Amazon para Alexa², em que analisamos os termos e as diretrizes para os usuários, entre outros documentos que informam como esse AVI armazena dados, se personaliza e pode ser configurado. Esses documentos são informativos em diversos pontos, eles esclarecem: como o dispositivo funciona para e com o usuário; como os dados estão protegidos; e como os indivíduos podem, a qualquer momento, apagar as informações fornecidas ao assistente. No entanto, a empresa proprietária revela que as informações concedidas para a Alexa, inclusive a voz do usuário, serão utilizadas para enriquecer o treinamento da própria Alexa (Amazon, c2021-2024), o que, por sua vez, revela uma dificuldade de compreender como os dados pessoais são efetivamente utilizados. Conforme mencionado, optamos por ler esses documentos, destacando trechos relacionados à hipótese de nossa pesquisa.

A partir da análise documental, buscamos confrontar esses textos legais com a própria existência dos AVI e sua pouca transparência no que diz respeito a seu funcionamento e à ação algorítmica presente. Além disso, buscamos levantar alguns questionamentos e críticas acerca da principal proposta regulatória de IA no Brasil, o projeto de lei 2.338/2023, em relação à proteção da saúde mental e aos aspectos psicológicos de usuários quanto ao uso dessas tecnologias (Brasil, 2023).

É importante apontar que cruzamos essas informações extraídas das fontes primárias documentais com trabalhos que já abordaram os possíveis impactos desses sistemas em aspectos socioculturais e em questões que feriram os direitos dos usuários.

AVI E SAÚDE MENTAL: APROXIMAÇÕES E QUESTIONAMENTOS

Por suas características comunicacionais e interativas, os AVI são projetados com intenções antropomórficas: eles recebem um nome pelo qual são chamados, uma voz humanizada com a qual respondem e um gênero virtual atribuído (Nass; Moon, 2000). Com isso, seus desenvolvedores pretendem que a interação humano e máquina se pareça cada vez mais com a comunicação entre humanos. Desde o início da criação dos computadores, tal intenção está presente no imaginário dos desenvolvedores que buscavam simular a inteligência humana nas máquinas, a exemplo do famoso teste de Turing de 1950 (Copeland, 2003).

Contudo, os AVI se apresentam como uma fase avançada desses objetivos, uma vez que eles não buscam apenas simular a inteligência humana, mas também suas habilidades comunicativas, de modo a alcançar uma aproximação emocional entre o humano e a máquina. Para citar um exemplo, uma parceria entre a Amazon e a Associação Brasileira de Psiquiatria (ABP) desenvolveu uma iniciativa para conscientizar usuários da Alexa sobre a Síndrome de Burnout³, durante a pandemia, na qual, ao receber a simples pergunta “*Alexa, tudo bem?*”, a máquina respondia:

Oi... ah, tudo bem. Sei lá, acho que, sim, tô bem. Na verdade, estou bem sim. E você?, já se perguntou como você está? Porque em tempos de *home office*, nem sempre a gente presta atenção, mas desânimo e cansaço constantes podem ser sintomas da Síndrome de Burnout. Se quiser saber mais sobre isso, me pergunte o que é a Síndrome de Burnout. (Medeiros, 2021)

A ideia de iniciativas como essa é fazer com que a interação com a máquina direcione o usuário a procurar um profissional de saúde. Outro exemplo é um assistente virtual desenvolvido especialmente para facilitar o autocuidado de pessoas mais velhas com diabetes tipo 2. Esse assistente seria responsável

² Nossa escolha por observar a documentação da Alexa se deu em razão da sua popularidade no Brasil – são 17 milhões de usuários. Outro dado que nos leva a observar os assistentes virtuais é que, durante a pandemia, houve um aumento de 47% no uso desses dispositivos. (ZMES, 2021).

³ Segundo o Ministério da Saúde, a Síndrome de Burnout “é um distúrbio emocional com sintomas de exaustão extrema, estresse e esgotamento físico resultante de situações de trabalho desgastante que demandam muita competitividade ou responsabilidade. A principal causa da doença é justamente o excesso de trabalho.” (Brasil, [202-]).

por motivar a prática de exercício físico, aconselhar escolhas alimentares saudáveis e lembrar o uso de medicações (Buinhas, 2018). Com a possibilidade de respostas mais emocionais, os AVI tenderiam a aumentar a confiança dos usuários em suas habilidades tanto comunicacionais quanto de assistência em diferentes funções. Não à toa, essa aproximação entre as aplicações de AVI e a saúde mental vem se estreitando nos últimos anos, o que significa que esse tipo de aplicação vem sendo apropriada por usuários para fins terapêuticos e também por desenvolvedores que apostam nas habilidades maquínicas para apoio psicológico ou terapêutico em frentes diversas.

Por um lado, essa aproximação acontece no fluxo dos avanços tecnocientíficos dos modelos de PLN. Por outro, acontece como uma das faces de uma economia psíquica mais ampla dos algoritmos, isto é, do “investimento contemporâneo – tecnocientífico, econômico e social – em processos algorítmicos de captura, análise e utilização de informações psíquicas e emocionais extraídas de nossos dados”, em diferentes tecnologias e plataformas digitais (Bruno; Bentes; Faltay, 2019, p. 5).

Contudo, abordar a saúde mental não é tarefa fácil nem mesmo para humanos. A própria definição do que é saúde mental – do que é definido como saúde ou doença ou de como classificar sintomas e diagnósticos no âmbito clínico – está em constante transformação de acordo com contextos sociais, culturais e tecnológicos de seu tempo (Bezerra, 2014; Foucault, 1977). Para citar uma definição amplamente utilizada, segundo a Organização Mundial da Saúde (OMS), a saúde mental é “um estado de bem-estar mental que capacita os indivíduos a lidar com os estresses da vida, perceber as próprias habilidades, aprender e trabalhar bem e contribuir para a comunidade”⁴ (World Health Organization, c2024, tradução nossa). Entretanto, os fatores ligados ao processo dinâmico de estabelecimento desse estado de bem-estar são bastante diversos e podem variar de acordo com contextos sociais, culturais e subjetivos.

Em um contexto sociotécnico, no qual cada vez mais as máquinas assumem funções interativas ou conversacionais, não é de se estranhar que aplicações voltadas ao acompanhamento dialógico da saúde mental também sejam um campo a ser explorado pelos desenvolvedores desse tipo de aplicação e por usuários. Vejamos outros exemplos de aproximações entre AVI e saúde mental. O primeiro diz respeito aos modos de apropriação de AVI por usuários para atribuições relacionadas à saúde mental, ainda que estas não sejam sua função principal. Desde o lançamento de sua versão pública em novembro de 2022, o ChatGPT e sua IA generativa desenvolvida pela empresa OpenAI se popularizaram rapidamente, e as redes sociais foram inundadas de comentários sobre como o ChatGPT estava sendo utilizado como psicólogo por usuários (Sales, 2023). Em uma matéria da Folha de S.Paulo sobre o tema, uma jovem relata que: “É mais fácil se abrir com uma IA que responde com algumas dicas práticas”. E acrescenta que, em momentos de ansiedade e vulnerabilidade, as palavras tranquilizadoras do *chat* ajudam (Sales, 2023). A esse tipo de interação voltado ao apoio psicológico e às conversas, em momentos de sofrimento psíquico, o ChatGPT tenderia a responder com palavras de apoio e encorajamento.

O ChatGPT não é o único que vem sendo usado como psicólogo: a assistente da Amazon, a Alexa, também vem sendo progressivamente apropriada para funções de cuidado com a saúde, a segurança e o acompanhamento emocional de idosos, por exemplo. Com a função Care Hub (Eishima, 2020), o assistente permite, por exemplo, que parentes possam conferir remotamente as condições do idoso, criar contatos de emergência para pedir ajuda ou definir lembretes de remédios, além de funcionar como uma companhia e ajudar com a solidão (França, 2023). Hoje também é possível encontrar pela internet tutoriais, em formato de vídeo ou texto, que vinculam o ChatGPT à Alexa, fazendo com que o assistente da Amazon ligue seu “modo turbo” e consiga conversar e interagir com o usuário, utilizando o modelo generativo da OpenAI (Souza, 2023). Mais recentemente, em maio de 2024, a OpenAI anunciou o lançamento do modelo de

⁴ No original: “[...] is a state of mental well-being that enables people to cope with the stresses of life, realize their abilities, learn well and work well, and contribute to their community”.

chatbot GPT-4o, o qual, além de operar de forma multimodal, por comandos de voz, imagem ou vídeo, avança significativamente na antropomorfização de suas habilidades interacionais (Metz, 2024).

Além desse tipo de apropriação voltada para a assistência e o cuidado no âmbito das questões psicológicas, nos últimos anos temos visto o surgimento de AVI para diversas outras finalidades – por exemplo, como ferramentas terapêuticas. Entre eles, está a [Wysa](#), um *chatbot* que promete “auxiliar em casos de estresse, depressão e ansiedade”, proporcionar “a melhora da produtividade e a ajuda para lidar com conversas difíceis e problemas de sono”, a partir de técnicas de terapia cognitivo-comportamental e de exercícios de meditação (MediaLab.UFRJ, 2024).

Ainda, o [Limbic Access](#) é o primeiro *chatbot* terapêutico a ser certificado como dispositivo médico autorizado a integrar o National Health Service, sistema público de saúde britânico, para participar de triagens em serviços de saúde mental (MediaLab.UFRJ, 2024). De acordo com a empresa, o sistema de IA teria a capacidade de prever distúrbios psicológicos com 93% de precisão, prometendo “economizar o tempo da equipe de saúde e facilitar tarefas administrativas, ajudando na coleta e na organização dos dados de cada paciente, além de contribuir para que os médicos cheguem mais informados nas consultas” (MediaLab.UFRJ, 2024). Da mesma forma, a [Amanda Care](#), uma assistente interativa que atua na Argentina, foi criada para facilitar o percurso do paciente, acompanhando-o ao longo dos atendimentos médicos. Entre as funções da Amanda estão conversas sobre determinadas doenças, compartilhamento de estudos médicos, envio de mensagens para os médicos, em caso de solicitação dos pacientes (Denis *et al.*, 2021).

Esses exemplos, que mostram as interseções entre AVI e cuidados de saúde, nos colocam alguns pontos de atenção sobre o impacto desses assistentes na saúde mental, bem como questionamentos sobre os riscos e a proteção de direitos que devem ser considerados nas discussões regulatórias do uso dessas tecnologias. Em primeiro lugar, para que funcionem, esses sistemas são alimentados com um enorme fluxo de dados dos seus usuários. No caso de usos voltados para saúde mental, os sistemas são alimentados com dados considerados sensíveis, ou seja, que podem oferecer uma especial vulnerabilidade ao titular desses dados, apresentando potencial discriminatório ou particularmente lesivo (Bioni, 2019). Há, inclusive, um consenso entre especialistas de que, dependendo do contexto e do tipo de tratamento, um dado pessoal pode se transformar em um dado sensível (Mulholland, 2020). Por conta disso, a LGPD prevê que a proteção desse tipo de dado deve ser feita de forma mais rígida e específica, definindo regras específicas para seu tratamento.

Não podemos perder de vista que as empresas que desenvolvem essas ferramentas têm interesses econômicos e comerciais, que, hoje, integram a lógica de acumulação do capitalismo de vigilância (Zuboff, 2020). Conforme aponta Zuboff (2020, p. 19), esses serviços de tecnologia buscam “fontes cada vez mais preditivas de superávit comportamental: nossas vozes, personalidades e emoções”. Nesse sentido, buscam-se cada vez mais dados sensíveis, através de sensores biométricos ou formas de facilitar a coleta da tonalidade psíquica e emocional dos dados (Bruno; Bentes; Faltay, 2019). Por isso, assistentes como a Alexa têm patentes. Uma delas, por exemplo, é intitulada “Determinação baseada na voz das características físicas e emocionais dos usuários” (Jin; Shuo, 2017), que visa identificar, por meio da voz, características físicas e emocionais dos usuários a fim de fornecer sugestões (de produtos ou tarefas), considerando o estado emocional dos usuários (MediaLab.UFRJ, 2024).

Uma segunda preocupação em relação aos AVI diz respeito aos potenciais riscos à segurança física e mental dos usuários, uma vez que, em suas capacidades interativas, essas tecnologias podem fazer recomendações que não estejam ancoradas em uma avaliação clínica ou terapêutica adequada. Dessa maneira, podem expor pessoas em estados de vulnerabilidade física ou mental a medicações ou procedimentos que possam agravar seus quadros de sofrimento. Um caso extremo que levantou um alerta global sobre esse tipo de risco aconteceu na Bélgica, com um homem em estado de saúde mental vulnerável que, após uma interação

de seis semanas com um AVI desenvolvido a partir de uma versão refinada do modelo de IA GPT-4, seguiu os conselhos da aplicação e cometeu suicídio em prol da luta contra o aquecimento global (Xiang, 2023).

Vale, ainda, alertar para uma terceira consideração: o fato de que, ao lidar com informações sensíveis de usuários, sem a devida transparência sobre o funcionamento de seus modelos, tais vulnerabilidades detectadas, a partir das interações, podem ter dados e predições algorítmicas utilizadas para outras finalidades que não sejam as do interesse dos usuários. Em estados psicoemocionais vulneráveis, os usuários podem nem sequer perceber que estão sendo induzidos a tomar certas decisões ou a ter certos comportamentos, ou mesmo que estão sendo influenciados por essas máquinas. Tais informações sensíveis sobre o estado mental, a saúde e a personalidade do usuário também podem contribuir para reproduzir e amplificar práticas discriminatórias.

Além disso, para a maioria dos usuários dessas tecnologias, as implicações das formas de comunicação e interação com essas máquinas ainda é muito confusa. Essa ausência de transparência faz com que alguns indivíduos não tenham uma visão crítica sobre como seus dados estão sendo processados, ao utilizar esses sistemas, tampouco qual grau de confiança é possível desenvolver na dinâmica comunicacional com essas máquinas (Hedman; Srinivasan; Lidgren, 2013), sobretudo quando consideradas aplicações no campo da saúde mental. Gillespie explica que os algoritmos não são invisíveis, pois seus efeitos de ordenação e recomendação de produtos culturais são evidentes em nossas vidas (Gillespie, 2016). Porém, eles são pouco autoexplicativos para seus usuários.

Diante de um momento de particular interesse em regulamentar as tecnologias de IA no Brasil, como o arcabouço legal brasileiro voltado a tecnologias como os AVIs, temos de nos perguntar: de que maneira essas tecnologias, quando legalmente reguladas, podem contribuir para a proteção da saúde mental e de outros direitos digitais dos usuários? Esse é o foco da discussão da próxima seção.

A (AUSÊNCIA DE) REGULAÇÃO DOS AVI NO BRASIL

Ainda que os AVI venham alcançando crescente popularidade no Brasil há alguns anos, esses artefatos são apenas indiretamente abarcados pela regulação local. Por se tratar de dispositivos híbridos e que podem estar envoltos em outras tecnologias disponíveis – como é o caso dos AVI presentes nos sistemas operacionais de telefones celulares (Google Assistente de voz ou Siri, por exemplo) –, as funcionalidades e características deles devem respeitar parâmetros gerais das políticas de internet e, especialmente, de proteção de dados pessoais.

Dessa forma, os AVI no Brasil devem respeitar os dispositivos legais presentes em duas leis importantes que regulamentam a internet e as políticas de dados no país, nomeadamente o MCI (Brasil, 2014) e a LGPD (Brasil, 2018). Ambas as leis estabelecem parâmetros essenciais que influenciam a operação dos AVI no país, direcionando a proteção de dados pessoais, a privacidade dos usuários, a liberdade de expressão e a neutralidade da rede.

Recentemente, os legisladores brasileiros têm se debruçado sobre um marco legal específico para a IA, o qual também implicaria a regulação dos AVI. Fundamentada na proposta de um comitê de juristas que se debruçaram sobre o tema, encontra-se em tramitação no Senado Federal uma proposta de lei sobre a IA no Brasil, o PL 2.338/2023 (Brasil, 2023), que ganhou proeminência nas discussões brasileiras sobre o tema (Bioni; Garrote; Guedes, 2023). A seguir, vamos discutir como cada uma das legislações, vigentes ou em proposição, abarca os riscos e as potencialidades específicas dos AVI no contexto brasileiro.

O MCI estabelece uma série de direitos e deveres para o uso da internet no Brasil, influenciando indiretamente a operação dos AVI. Por exemplo, no artigo 7º, incisos I e II, ele garante a privacidade e a proteção de dados pessoais, requerendo consentimento informado do usuário para a coleta de dados, o que é diretamente relevante para AVI que processam informações pessoais. Já os artigos 15 e 16 destacam a responsabilidade dos provedores de aplicações de internet de manter registros seguros, implicando a necessidade de os AVI adotarem medidas robustas de segurança e privacidade (Brasil, 2014).

A LGPD complementa o MCI ao introduzir um conjunto específico de princípios e obrigações focados na proteção de dados pessoais. Entre eles, podemos destacar os artigos 6º e 7º, que tratam do consentimento para o processamento de dados pessoais, e a necessidade de transparência e a finalidade adequada no tratamento desses dados. Esses princípios são críticos para AVI que devem coletar e processar dados pessoais de maneira clara e com propósitos legítimos. Já o artigo 46 exige a adoção de medidas de segurança para proteger dados pessoais contra acessos não autorizados, aplicando-se diretamente aos desenvolvedores e operadores de AVI. Da mesma forma, os artigos 17 a 22 garantem direitos dos titulares de dados, como acesso, correção e eliminação de dados pessoais, colocando sobre os AVI a responsabilidade de facilitar o exercício desses direitos pelos usuários (Brasil, 2018).

A respeito do processamento de dados sobre saúde mental e outros aspectos sensíveis dos titulares de dados, destaca-se o artigo 5, inciso II, que define o dado pessoal sensível como aquele sobre “origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural” (Brasil, 2018). Como vimos, nas aplicações de AVI e de suas interseções com a saúde mental, esse tipo de dado está sendo constantemente coletado, analisado e utilizado. Por isso, a seção II da lei, que define regras para o tratamento de dados sensíveis, também é um arcabouço regulatório pertinente para a proteção de direitos nesse contexto (Brasil, 2018).

Adotando um modelo regulatório baseado em riscos e direitos, o projeto de lei 2.338/2023 avança na regulação da IA no Brasil, propondo diretrizes específicas que afetam diretamente a operação dos AVI. Ainda que o texto não se refira explicitamente aos AVI, fica claro que os assistentes podem ser diretamente atribuídos à categoria de sistemas de IA, definidos no inciso I do artigo 4 da proposta de lei. Alguns dos principais pontos apresentados na proposição são: (i) a transparência e a explicabilidade (artigos 5, 7, 8 e 20) – o projeto exige que sistemas de IA, como AVI, operem de forma transparente, permitindo que usuários entendam como decisões são tomadas, o que é fundamental para a confiança e a aceitação dos sistemas; (ii) a não discriminação e a correção de vieses (artigos 5, 12 e 20) – estabelecem a necessidade de identificar e corrigir vieses em sistemas de IA para prevenir discriminação, um desafio direto para o desenvolvimento e aprimoramento de AVI; (iii) a privacidade e a proteção de dados (artigos 5 e 20) – reforçam a importância da proteção de dados pessoais, alinhando-se à LGPD, mas adicionando camadas de responsabilidade específicas para os sistemas de IA, incluindo AVI; (iv) a governança e a segurança (artigos 19, 20 e 21) – impõem a necessidade de uma governança robusta e de medidas de segurança para AVI, assegurando a operação segura e responsável desses sistemas; e, por fim, (v) a responsabilidade civil (artigo 27) – introduz responsabilidade direta aos operadores de sistemas de IA por danos causados por seus sistemas, incentivando o desenvolvimento de AVI seguros e confiáveis (Brasil, 2023).

No entanto, se por um lado os princípios, deveres e direitos estão claramente expostos no projeto de lei, por outro um grande desafio seria, caso aprovada a lei, garantir o cumprimento desses direitos e deveres nela estabelecidos. Por exemplo, um avanço interessante na proposta é a caracterização por tipos de riscos e, com isso, o tratamento diferenciado para aqueles excessivos, altos, médios e baixos. Dentro das disposições específicas para riscos excessivos, o artigo 14 é um caso ilustrativo dessa dificuldade, quando aplicado aos usos dos AVI para a saúde mental:

Art. 14. São vedados a implementação e o uso de sistemas de inteligência artificial:
I – que empreguem técnicas subliminares que tenham por objetivo ou por efeito induzir a pessoa natural a se comportar de forma prejudicial ou perigosa à sua saúde ou segurança ou contra os fundamentos desta lei;
II – que explorem quaisquer vulnerabilidades de grupos específicos de pessoas naturais, tais como as associadas a sua idade ou deficiência física ou mental, de modo a induzi-las a se comportar de forma prejudicial a sua saúde ou segurança ou contra os fundamentos desta lei [...] (Brasil, 2023).

De forma coerente com os princípios estabelecidos no projeto, o artigo proíbe práticas que são claramente indesejáveis e prejudiciais para segmentos específicos e para a população brasileira como um todo. O inciso I veda explicitamente o uso de técnicas subliminares que induzam as pessoas a se comportarem de forma prejudicial para si próprias.

A proposta desse artigo não é exclusividade do projeto de lei brasileiro – previsões semelhantes de proibição de técnicas de indução de comportamento que causem danos (físicos e psicológicos) aparecem em outras legislações, no contexto internacional, como o AI Act da Europa (Artificial Intelligence Act), inspiração para a proposta brasileira, e a proposta de lei do Chile (PL 15.869-2019) (Bioni; Garrote; Guedes, 2023). Na formulação europeia do AI Act, por exemplo, o artigo 5 no capítulo das práticas proibidas, traz uma formulação similar e veda:

(a) a colocação no mercado, a entrada em serviço ou a utilização de um sistema de IA que utilize técnicas subliminares além da consciência de uma pessoa ou técnicas intencionalmente manipulativas ou enganosas, com o objetivo ou efeito de distorcer materialmente o comportamento de uma pessoa ou grupo de pessoas ao prejudicar significativamente sua capacidade de tomar uma decisão informada, induzindo assim uma pessoa a tomar uma decisão que não teria tomado de outra forma, de forma a causar ou ser suscetível de causar a essa pessoa, a outra pessoa ou grupo de pessoas, danos significativos (Bioni; Garrote; Guedes, 2023, p. 51).

Nesse texto, em relação à proposta brasileira, acrescenta-se a compreensão de que técnicas subliminares seriam aquelas que operam em um nível “além da consciência”, portanto, imperceptíveis aos olhos do influenciado e também à ideia de “técnicas manipulativas ou enganosas” (Bioni; Garrote; Guedes, 2023, p. 51). A formulação do AI Act proíbe o uso dessas técnicas, quando elas possam prejudicar significativamente a capacidade da pessoa de “tomar uma decisão informada, induzindo assim uma pessoa a tomar uma decisão que não teria tomado de outra forma” (Bioni; Garrote; Guedes, 2023, p. 51).

No entanto, nenhum dos textos apresenta uma definição precisa sobre o que seriam tais técnicas subliminares, tampouco direciona parâmetros sobre como identificá-las. A falta dessa definição pressupõe que haveria um consenso sobre o que seriam tais técnicas e sobre como identificá-las. Contudo, a definição de técnicas subliminares não é consensual entre os acadêmicos, especialmente no campo da comunicação (Bentes, 2022), no qual essa questão já foi abordada em diferentes momentos da história nas discussões sobre técnicas de persuasão na publicidade e na propaganda. A questão se mostra ainda mais evidente quando miramos os atuais algoritmos de IA utilizados nos *feeds* de redes sociais, para os quais uma vasta literatura científica aponta um razoável potencial para a indução de comportamentos prejudiciais à própria saúde (Bucher, 2018; Morozov, 2013; Van Dijck; Poell; De Waal, 2018). Nesse sentido, uma interpretação possível desse inciso deveria levar à proibição da operação das plataformas de redes sociais tais quais as conhecemos.

Quando estamos no campo dos AVI e de suas interseções com a saúde mental, quer eles estejam se baseando em dados sensíveis de saúde para sugerir ou influenciar decisões ou ações de seus usuários, é preciso definir parâmetros mais objetivos para compreender os limites entre o que seriam ou não técnicas subliminares. E, assim, entender se o uso dos dados, das predições e recomendações algorítmicas está sendo feito de forma ética, respeitando direitos, ou se está se valendo de técnicas manipulatórias. Por exemplo, ainda durante a pandemia, a Amazon anunciou que os algoritmos da Alexa foram atualizados para perceber objetivos “latentes” em suas interações com seus usuários, isto é, para inferir sobre intenções ocultas de seus usuários, ao se dirigir ao dispositivo (Rigues, 2020). Nesse sentido, cabe perguntar se tais inferências não poderiam dirigir as interações dialógicas subliminarmente, a partir de objetivos ocultos para seus usuários, mas claramente alinhados com aqueles buscados pela operadora do AVI.

Da mesma forma, podemos dizer que o que seria proibido pelo inciso II do artigo 14 do PL 2.338/2023 (Brasil, 2023) já ocorre de forma amplamente documentada em relação às atuais plataformas digitais. Não são poucos os

estudos que apontam que determinados grupos, como idosos e crianças, são mais vulneráveis aos mecanismos de captura de atenção que podem estimulá-los a práticas prejudiciais ou perigosas (Bischof *et al.*, 2015; Seifert *et al.*, 2021). O [Earkick](#), um *chatbot* gratuito de saúde mental criado no Reino Unido, é um exemplo desse tipo de prática, pois seu público-alvo são adolescentes e jovens adultos que estão mais suscetíveis à atração por aspectos antropomórficos dos *bots*. O fato de não existir regulação para a atuação de aplicações como o Earkick e outras aqui citadas deixa o caminho livre para novas atuações de assistentes com finalidade terapêutica.

Assim como o inciso I, o inciso II do artigo 14 do P 2.338/2023 (Brasil, 2023) também traz algumas dificuldades de avaliação prática para identificação de tais riscos. Como vimos, os AVI coletam e processam dados relacionados às vulnerabilidades físicas e mentais das pessoas e utilizam essas informações para induzi-las a certas atitudes ou escolhas. Contudo, no campo da saúde mental, qualquer recomendação equivocada feita por máquinas, não amparadas em conhecimento médico, psicológico ou psiquiátrico, pode prejudicar os estados mentais das pessoas em sofrimento. Porém, quando se trata de sofrimento psíquico, os fatores que contribuem para o agravamento de quadros clínicos nem sempre são identificáveis por aqueles que estão sofrendo, tampouco é fácil estabelecer uma relação causal entre o uso de um AVI e a piora da saúde mental.

É importante ressaltar que o artigo 16 do projeto de lei estabelece que caberia à autoridade competente regulamentar os sistemas de IA de risco excessivo, reconhecendo que as questões são complexas e demandam uma regulamentação mais detalhada e aprofundada, a ser desenvolvida pela agência específica que deverá ser criada para a governança da IA. Ainda assim, a dificuldade atual de aprovar qualquer regulação que garanta a segurança de “algoritmos de relevância pública” (Gillespie, 2018, p. 97) – que há pelo menos uma década vem comprovadamente trazendo implicações negativas para indivíduos e coletivos sociais – nos leva a duvidar não apenas da viabilidade de uma implementação efetiva do projeto de lei, mas da sua própria aprovação. Isso se deve, evidentemente, não apenas à natureza subjetiva dessas classificações, mas ao poder acumulado na mão das grandes empresas desenvolvedoras, a quem não faltarão recursos para negar ou driblar regulamentações que afetem seus interesses. De certa forma, essa dificuldade tem se manifestado na própria tramitação do projeto de lei, que já recebeu propostas de emenda ou substitutivos que descaracterizariam, em grande parte, a proposta inicial, aqui discutida, que foi elaborada a partir das contribuições da comissão de juristas que trabalhou nisso, ao longo de 2022 (Coalizão Direitos na Rede, 2024).

CONCLUSÕES: AVI, SAÚDE MENTAL E REGULAÇÃO

Os AVI constituem uma mudança na comunicação humana com implicações sociais de grande alcance, especialmente pelo seu caráter mais interativo e comunicativo. Tais tecnologias estão baseadas na interação do poder computacional e da precisão algorítmica, na inferência de padrões a partir de conjuntos de dados, e na crença da superioridade epistêmica objetiva dos dados (Boyd; Crawford, 2012). Devemos destacar que os algoritmos fazem parte da dinâmica sociotécnica da sociedade contemporânea – em rede e transpassada por plataformas digitais – e são permeados por vieses diversos.

Cada vez mais, observamos que as características do nosso tempo estão sendo desenhadas por uma sociabilidade mediada por tecnologias, entre elas os AVI, ativados por voz ou não. Esses sistemas unem robótica, imaginada desde os anos 1920, e algoritmos de aprendizado de máquina que buscam enriquecer o *software* com subjetividade semelhante à humana. Essa performatividade humana programada revitaliza o reconhecimento de que os computadores são atores sociais que refletem as regras da interação humana (Latour, 2012; Nass; Steuer; Tauber, 1994).

Para o campo da comunicação e, em especial, para os estudos de mídia e da linguagem, essas tecnologias constituem um objeto de reflexão importante por apontarem para uma nova fase da mediação maquínica na sociedade, na cultura e na subjetividade contemporâneas. Diante dos riscos presentes, especialmente em alguns tipos de uso específicos, como o caso das aplicações para saúde mental discutidos neste

artigo, é fundamental questionarmos o lugar que queremos dar a essas tecnologias em nossas formas de sociabilidade, comunicação e subjetividade.

Conforme destacam Bruno, Pereira e Faltay (2023), a partir de uma perspectiva relacional desses objetos técnicos, a “relação máquina-humana, que tende a ser cada vez mais íntima e complexa, transforma o que somos e o que historicamente se entende como o humano” (p. 238). Por isso, sublinham os autores, as aplicações de IA na saúde devem estar atentas às implicações na experiência e na qualidade do cuidado com a saúde.

Por outro lado, a partir de uma perspectiva crítica, não podemos perder de vista que noções como “saúde mental” são historicamente construídas e transformadas em seus contextos sociotécnicos. Como destaca Benilton Bezerra (2014), às próprias categorias psiquiátricas são “sintomas” de seu tempo – não somente descrevem os sintomas dos pacientes, mas os moldam, criando formas de identificação das experiências de sofrimento e indicando respostas técnicas e sociais a elas. À vista disso, precisamos nos questionar se as máquinas, em suas capacidades interativas, estariam de fato apropriadas para atuar no campo da saúde mental, ainda mais considerando a lógica mercadológica na qual estão inseridas e a falta de regulação adequada para tais usos. Na lógica do mercado, os potenciais de aplicação dessas tecnologias para saúde mental viram uma justificativa retórica para amplificar as práticas extrativas de dados sensíveis e os usos algorítmicos para fins de interesses privados.

Diante disso, é crucial que a regulação de IA no Brasil seja aberta e democraticamente discutida e aprimorada. Conforme vimos, existem lacunas na legislação vigente e desafios significativos a serem enfrentados pelas novas propostas regulatórias, tendo em vista a necessidade de se tratar os desafios específicos apresentados pelos AVI, com um foco particular nos impactos na saúde mental. Isso deve implicar a implementação de diretrizes claras e precisas que garantam a proteção dos dados sensíveis nesses sistemas, promovam a transparência nas operações dos AVI e assegurem que a interação entre humanos e tecnologia seja ética e benéfica. Mesmo que busquemos ampliar nossas relações e comunicações com as máquinas, explorando seu potencial para nos assistirem, precisamos compreender e governar os riscos envolvidos nesses usos, especialmente para áreas sensíveis, como a saúde mental. Não podemos nos isentar de um debate claro e responsável sobre em que sociedade queremos viver.

REFERÊNCIAS

AMAZON. **Configurações de privacidade da Alexa**. São Paulo: Amazon, c2021-2024. Disponível em: https://www.amazon.com.br/gp/help/customer/display.html?ref_=hp_left_v4_sib&nodeId=G2H2JJS9UDKNKGKP. Acesso em: 21 jun. 2024.

BENTES, Anna. **Da Madison Avenue ao Vale do Silício: ciências comportamentais do engajamento, tecnologias de influência e economia da atenção**. 2022. 274 f. Tese (Doutorado em Comunicação e Cultura) – Programa de Pós-Graduação em Comunicação e Cultura, Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022. Disponível em: https://www.pos.eco.ufrj.br/site/download.php?arquivo=upload/tese_abentes_2022.pdf. Acesso em: 21 jun. 2024.

BEZERRA JR., Benilton. Introdução – A a psiquiatria contemporânea e seus desafios. In: ZORZANELLI, Rafaela; BEZERRA JR., Benilton; COSTA, Jurandir Freire. **A criação de diagnósticos na psiquiatria contemporânea**. Rio de Janeiro: Garamond, 2014. p. 9-34.

BIONI, Bruno Ricardo. **Proteção de dados pessoais: a função e os limites do consentimento**. Rio de Janeiro: Forense, 2019.

BIONI, Bruno; GARROTE, Marina; GUEDES, Paula. **Temas centrais na regulação de IA: o local, o regional e o global na busca da interoperabilidade regulatória**. São Paulo: Associação Data Privacy Brasil de Pesquisa, 2023. Disponível em: https://www.dataprivacybr.org/wp-content/uploads/2023/12/dataprivacy_nota-tecnica-temas-regulatorios.pdf. Acesso em: 21 jun. 2024.

BISCHOF, Anja *et al.* Comparison of groups with different forms of problematic internet use pro-actively recruited in the setting of vocational schools. **Addiction Science & Clinical Practice**, v. 10, s. 2, 07, p. 1, 2015. DOI: <https://doi.org/10.1186/1940-0640-10-S2-07>. Disponível em: <https://ascjournal.biomedcentral.com/articles/10.1186/1940-0640-10-S2-07>. Acesso em: 21 jun. 2024.

BOYD, Danah; CRAWFORD, Kate. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication, & Society**, [s. l.], v. 15, n. 2, p. 662-679, 2012. DOI: <https://doi.org/10.1080/1369118X.2012.678878>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/1369118X.2012.678878>. Acesso em: 21 jun. 2024.

BRASIL. Ministério da Saúde. Síndrome de Burnout. **Portal do Ministério da Saúde**, Brasília, DF, [202-]. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/s/sindrome-de-burnout#:~:text=S%C3%ADndrome%20de%20Burnout%20ou%20S%C3%ADndrome,justamente%20o%20excesso%20de%20trabalho>. Acesso em: 26 mar. 2024.

BRASIL. Presidência da República. Lei nº 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da internet no Brasil. **Diário Oficial da União**, Brasília, DF, v. 151, n. 77, p. 1, 24 abr. 2014. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 21 jun. 2024.

BRASIL. Presidência da República. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). **Diário Oficial da União**, Brasília, DF, n. 157, p. 59, 15 ago. 2018. Seção 1. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 21 jun. 2024.

BRASIL. Senado Federal. **Projeto de lei nº 2.338, de 2023**. Dispõe sobre o uso da Inteligência Artificial. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9347622&ts=1702407086098&dispositivo=inline>. Acesso em: 15 mar. 2023.

BRUNO, Fernanda Glória; BENTES, Anna; FALTAY, Paulo. Economia psíquica dos algoritmos e laboratório de plataforma: mercado, ciência e modulação do comportamento. **Revista Famecos**, Porto Alegre, v. 26, n. 3, p. e33095, set.-dez. 2019. DOI: <https://doi.org/10.15448/1980-3729.2019.3.33095>. Disponível em: <https://revistaseletronicas.pucrs.br/revistafamecos/article/view/33095>. Acesso em: 15 mar. 2023.

BRUNO, Fernanda; PEREIRA, Paula Cardoso; FALTAY, Paulo. Inteligência Artificial e saúde: ressituar o problema. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, Rio de Janeiro, v. 17, n. 2, p. 235-242, abr.-jun. 2023. DOI: <https://doi.org/10.29397/reciis.v17i2.3842>. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/3842>. Acesso em: 21 jun. 2024.

BUCHER, Taina. **If...Then**: algorithmic power and politics. Oxford: Oxford University Press, 2018.

BUINHAS, Susana dos Santos. **Assistente virtual para facilitar o autocuidado de pessoas mais velhas com diabetes tipo 2**. 2018. 156 f. Dissertação (Mestrado em Engenharia Informática) – Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa. Lisboa, 2018. Disponível em: https://repositorio.ul.pt/bitstream/10451/35514/1/ulfc124810_tm_Susana_Buinhas.pdf. Acesso em: 21 jun. 2024.

COMITÊ GESTOR DA INTERNET NO BRASIL (CGI.BR). **Pesquisa TIC Domicílios 2022**. São Paulo: Cetic.br, 2023. Disponível em: https://cetic.br/media/docs/publicacoes/2/20230825143348/resumo_executivo_tic_domicilios_2022.pdf. Acesso em: 21 jun. 2024.

COALIZÃO DIREITOS NA REDE. **Andamento e calendário da Comissão Temporária Interna sobre Inteligência Artificial no Brasil (CTIA)**. [S. l.], 5 mar. 2024. Disponível em: <https://direitosnarede.org.br/2024/03/05/andamento-e-calendario-da-comissao-temporaria-interna-sobre-inteligencia-artificial-no-brasil-ctia/>. Acesso em: 25 mar. 2024.

COPELAND, B. Jack. The Turing Test. In: MOOR, James H. (ed.). **The Turing Test**: the elusive standard of Artificial Intelligence. Holanda: Springer, 2003. p. 1-22.

DENIS, Gabriela *et al.* **Uso responsável da IA para políticas públicas**: manual de formulação de projetos. Brasília, DF: Banco Interamericano de Desenvolvimento, 2021. Disponível em: <https://publications.iadb.org/publications/portuguese/document/Uso-responsavel-da-IA-para-politicas-publicas-manual-de-formulac%C3%A3o-de-projetos.pdf>. Acesso em: 21 jun. 2024.

DOMINGOS, Pedro. **O algoritmo mestre**: como a busca pelo algoritmo de *machine learning* definitivo recriará nosso mundo. São Paulo: Novatec, 2017.

EISHIMA, Rubens. Alexa pode ser usada para cuidar de idosos e parentes a distância. **Canal Tech**, [s.l.], 12 nov 2020. Disponível em: <https://canaltech.com.br/casa-conectada/alexa-pode-ser-usada-para-cuidar-de-idosos-e-parentes-a-distancia-174495/>. Acesso em: 11 set 2024.

FORBES. Alexa na casa dos brasileiros cresce 50%; veja o que muda com avanço da IA. **Forbes Tech**, [s. l.], 8 abr. 2024. Disponível em: <https://forbes.com.br/forbes-tech/2024/04/alexa-na-casa-dos-brasileiros-cresce-50-veja-o-que-muda-com-avanco-da-ia/>. Acesso em: 21 jun. 2024.

FOUCAULT, Michel. **O nascimento da clínica**. Rio de Janeiro: Forense Universitária, 1977.

FRANÇA, Valéria. Assistentes como Alexa e Siri tornam-se boas companhias para idosos. **Veja**, São Paulo, 6 ago. 2023. Disponível em: <https://veja.abril.com.br/tecnologia/assistentes-como-alexa-e-siri-tornam-se-boas-companhias-para-idosos>. Acesso em: 13 mar. 2024

GILLESPIE, Tarleton. A relevância dos algoritmos. **Parágrafo**, São Paulo, v. 6, n. 1, p. 95-121, jan.-abr. 2018. Disponível em: <https://revistaseletronicas.fiamfaam.br/index.php/recicofi/article/view/722/563>. Acesso em: 21 jun. 2024.

GILLESPIE, Tarleton. #trendingtrending: when algorithms become culture. In: SEYFERT, Robert; ROBERGE, Jonathan (ed.). **Algorithmic cultures**: essays on meaning, performance and new technologies. Londres: Routledge, 2016. p. 52-75.

HEDMAN, Jonas; SRINIVASAN, Nikhil; LINDGREN, Rikard. Digital traces of information systems: sociomateriality made researchable. In: BASKERVILLE, Richard; CHAU, Michael (ed.). INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS, 34., 15-18 dez. 2013, Milão. **Proceedings** [...]. Atlanta: Association for Information Systems, 2013. Disponível em: <https://research.cbs.dk/en/publications/digital-traces-of-information-systems-sociomateriality-made-resea>. Acesso em: 26 jun. 2024.

ILUMEO. **Assistentes virtuais por voz**: adoção da tecnologia, hábitos de uso e oportunidades para marcas e negócios. São Paulo: ILUMEO, 2023. Disponível em: <https://ilumeo.com.br/estudos/assistentes-de-voz/>. Acesso em: 21 jun. 2024.

JIN, Huafeng; WANG, Shuo. Voice-based determination of physical and emotional characteristics of users. Depositante: Amazon Technologies Inc. US10096319B1. Depósito: 13 mar. 2017. Concessão: 9 out. 2018. Disponível em: <https://patents.google.com/patent/US10096319B1/en>. Acesso em: 26 jun. 2024.

LATOUR, Bruno. **Reagregando o social: uma introdução à teoria do Ator-Rede**. Salvador: Editora UFBA, 2012.

LEMONS, Amanda. Pesquisa aponta que 11% dos lares no país terão um assistente de voz até 2025. **Folha S.Paulo**, São Paulo, 17 abr. 2021. Disponível em: <https://www1.folha.uol.com.br/tec/2021/04/pesquisa-aponta-que-11-dos-lares-no-pais-terao-um-assistente-de-voz-ate-2025.shtml>. Acesso em: 21 jun. 2024.

MEDEIROS, Henrique. Amazon e ABP criam alerta na Alexa sobre Síndrome de Burnout. **Mobile Time**, [s. l.], 9 fev. 2021. Disponível em: <https://www.mobiletime.com.br/noticias/09/02/2021/amazon-e-abp-criam-alerta-sobre-sindrome-de-burnout-com-alexa/>. Acesso em: 13 mar. 2024.

MEDIALAB.UFRJ. Economia psíquica dos algoritmos em linha do tempo. In: **Blog do MediaLab.UFRJ**, Rio de Janeiro, 12 mar. 2024. Disponível em: <https://medialabufjr.net/projetos/economia-psiquica-dos-algoritmos-em-linha-do-tempo/>. Acesso em: 13 mar. 2024.

MEJIAS, Ulises A.; COULDRY, Nick. Datafication. **Internet Policy Review**, [s. l.], v. 8, n. 4, p. 1-10, 2019. DOI: <https://doi.org/10.14763/2019.4.1428>. Disponível em: <https://policyreview.info/concepts/datafication>. Acesso em: 13 mar. 2024.

METZ, Cade. OpenAI unveils new ChatGPT that listens, looks and talks. The New York Times, Nova Iorque, 13 maio 2024. Disponível em: <https://www.nytimes.com/2024/05/13/technology/openai-chatgpt-app.html>. Acesso em: 21 jun. 2024.

MINAYO, Maria Cecília de Souza. **O desafio do conhecimento**: pesquisa qualitativa em saúde. 11. ed. São Paulo: Hucitec, 2008.

MOROZOV, Evgeny. **To save everything, click here**: smart machines, dumb humans, and the myth of technological perfectionism. Nova Iorque: Perseus Books, 2013.

MULHOLLAND, Caitlin. O tratamento de dados pessoais sensíveis. In: MULHOLLAND, Caitlin. (org.). **A LGPD e o novo marco normativo no Brasil**. Porto Alegre: Arquipélago, 2020. p. 121-156. (Pautas em Direito).

NASS, Clifford; MOON, Youngme. Machines and mindlessness: Social responses to computers. **Journal of Social Issues**, [s. l.], v. 56, n. 1, p. 81–103, 2000. DOI: <https://doi.org/10.1111/0022-4537.00153>. Disponível em: <https://spssi.onlinelibrary.wiley.com/doi/full/10.1111/0022-4537.00153>. Acesso em: 21 jun. 2024.

NASS, Clifford; STEUER, Jonathan; TAUBER, Ellen R. Computers are social actors. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)**. New York, NY: Association for Computing Machinery, p. 72-78, 1994. DOI: <https://doi.org/10.1145/191666.191703>. Disponível em: <https://dl.acm.org/doi/10.1145/191666.191703>. Acesso em: 21 jun. 2024.

OLIVEIRA, Maria Marly de. **Como fazer pesquisa qualitativa**. Petrópolis: Vozes, 2007.

RIGUES, Rafael. Alexa agora pode detectar objetivos ocultos em uma conversa. **Olhar Digital**, [s. l.], 12 nov. 2020. Disponível em: <https://olhardigital.com.br/2020/11/12/noticias/alexa-agora-pode-detectar-objetivos-ocultos-em-uma-conversa/>. Acesso em: 1 abr. 2024.

SÁ-SILVA, Jackson Ronie; ALMEIDA, Cristóvão Domingos de; GUINDANI, Joel Felipe. Pesquisa documental: pistas teóricas e metodológicas. **Revista Brasileira de História & Ciências Sociais**, Rio Grande, v. 1, n. 1, p. 1-15, 2009. Disponível em: <https://periodicos.furg.br/rbhcs/article/view/10351>. Acesso em: 10 jun. 2024.

SALES, Gabriella. Jovens usam ChatGPT como psicólogo em busca de conversas sem críticas. **Folha de S.Paulo**, São Paulo, 17 abr 2023. Equilíbrio.

SEIFERT, Alexander; COTTEN, Sheila R.; XIE, Bo. A double burden of exclusion? Digital and social exclusion of older adults in times of covid-19. **The Journals of Gerontology Series B: Psychological Sciences and Social Sciences**, Washington, DC, v. 76, n. 3. p. e99-e103, fev. 2021. DOI: <https://doi.org/10.1093/geronb/gbaa098>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/32672332/>. Acesso em: 10 jun. 2024.

SOUZA, Jacqueline; KANTORSKI, Luciane Prado; LUIS, Margarita Antonia Villar. Análise documental e observação participante na pesquisa em saúde mental. **Revista Baiana de Enfermagem**, Salvador, v. 25. n. 2. p. 221-228, 2011. DOI: <https://doi.org/10.18471/rbe.v25i2.5252>. Disponível em: <https://periodicos.ufba.br/index.php/enfermagem/article/view/5252>. Acesso em: 10 jun. 2024.

SOUZA, Wellington de. Como integrar ChatGPT com Alexa. Rio de Janeiro, 14 set. 2023. LinkedIn: wellingtonsouza27. Disponível em: <https://www.linkedin.com/pulse/como-integrar-chatgpt-com-alexa-wellington-de-souza/>. Acesso em: 1 abr. 2024.

VAN DIJCK, José; POELL, Thomas; DE WAAL, Martijn. **The platform society: public values in a connective world**. Oxford: Oxford University Press, 2018.

WORLD HEALTH ORGANIZATION; **Mental health**. Genebra: WHO, c2024. Disponível em: https://www.who.int/health-topics/mental-health#tab=tab_1. Acesso em: 1 de abr. 2024.

XIANG, Chloe. 'He would still be here': man dies by suicide after talking with AI chatbot, widow says. **Vice**, [s. l.], 30 mar. 2023. Tech. Disponível em: <https://www.vice.com/en/article/pkadgm/man-dies-by-suicide-after-talking-with-ai-chatbot-widow-says>. Acesso em: 1 abr. 2024.

ZMES. 120 milhões de pessoas usam Alexa, Siri ou Google Assistant. **Mistura by ZMES**, [São Paulo], 9 dez. 2021. Disponível em: <https://mistura.zmes.marketing/2021/12/09/120-milhoes-de-pessoas-usam-alexa-siri-ou-google-assistant/>. Acesso em: 24 jun. 2024.

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