

MIT Technology Review

Publicado por TEC



Edição
Digital
Voz

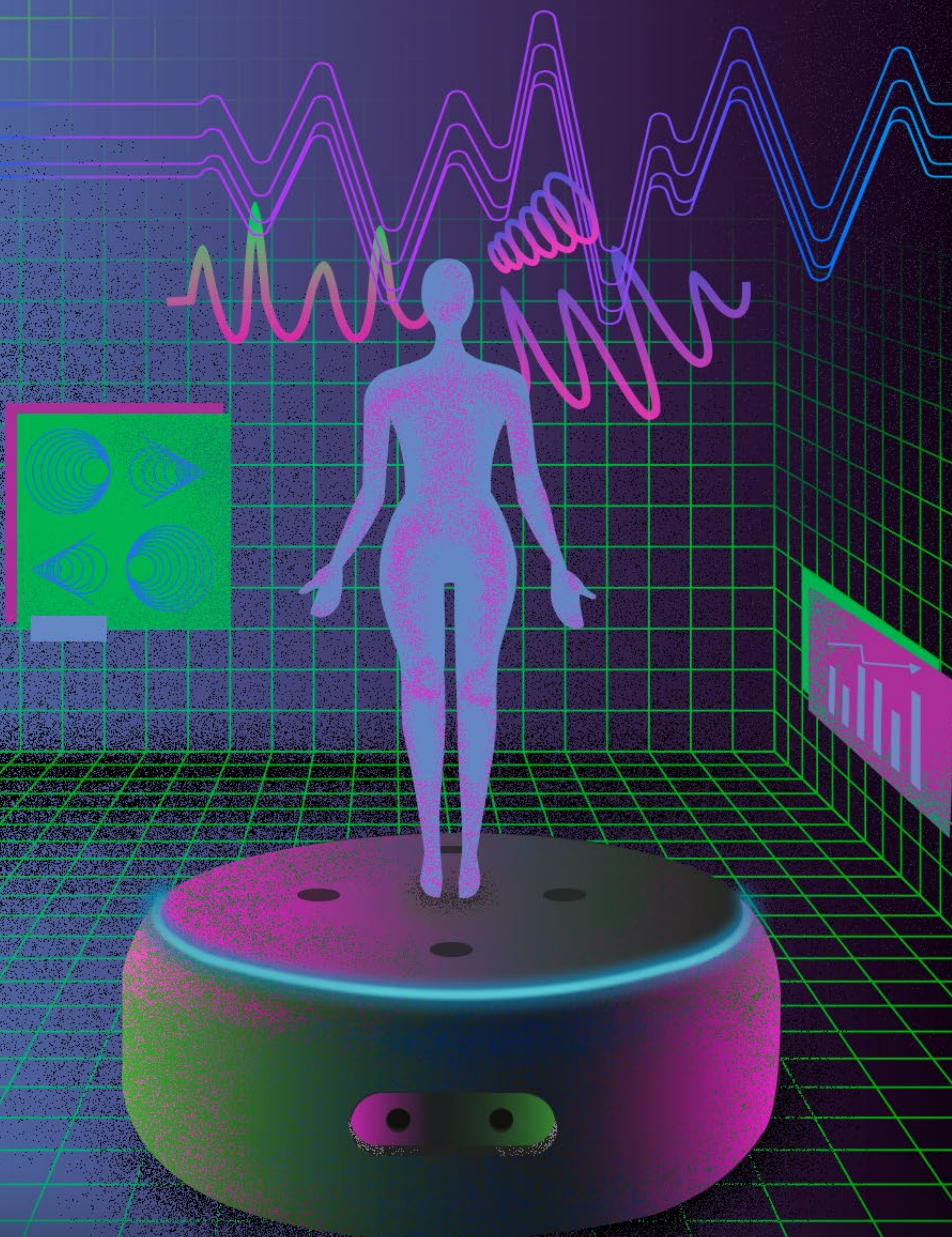
Ano 01

Número 02

Outubro

2021

A voz no comando



**O boom dos jogos por voz
está dando às crianças um
descanso das telas**

pág. 20

**TR Q+A Talita Bruzzi Taliberti:
a simplicidade da voz**

pág. 39

**Voz: humana vs sintética.
Qual é a melhor aplicação
de cada uma?**

pág. 45

Editorial

Editor-chefe
Andre Miceli

Editor-executivo
Carlos Aros

Editor-executivo
Rafael Coimbra

Editora, biotecnologia
Roberta Arinelli

Revisora
Camila Pepe

Revisora
Laura Murta

Gerente, operações
Rafael Coelho

Coordenador, conteúdo
Fernando Arbex

Coordenadora, tradução
Paula Arbex

Editorial EUA

Editor in chief
Mat Honan

Executive editor
Michael Reilly

Senior editor, biomedicine
Antonio Regalado

Senior editor, climate and energy
James Temple

Senior editor, digital culture
Abby Ohlheiser

Senior editor, cybersecurity
Patrick Howell O’Neill

Senior editor, AI
Will Douglas Heaven

Editor, Pandemic Technology Project
Lindsay Muscato

Conselho editorial

Lindalia Sofia

Luciana Bazanella

Marcos Facó

Pyr Marcondes

Ricardo Cappra

Sergio Gordilho

Yann Duzert

Design

Chief Creative Officer
Iago Ribeiro

Coordenador, design
Pedro Brito

Designer
Amanda Bordignon

Designer
Suelen Rapello

Ilustradora
Alessandra Benigno

Corporativo

Presidente do conselho de administração
Roberto Carneiro

Chief Executive Officer e publisher
Andre Miceli

Associado
Rodrigo Ferreira

Associado
Allan Fonseca

Gerente, recursos-humanos
Adriana Oliveira

Gerente, tecnologia
Mariana Aragão

Client Service Manager
Ted Hu

Global head, custom content and international
Nico Crepaldi

Marketing e vendas

Gerente, vendas
Daniel Salvador

Gerente, vendas
Rogério Oliveira

Gerente, vendas
Marcelo Augusto

Email
anuncie@mittechreview.com.br

Web
www.mittechreview.com.br/anuncie

Clientes e assinaturas

Email
assinaturas@mittechreview.com.br

Reimpressões e licenciamento
licencas@mittechreview.com.br

Assine
mittechreview.com.br/assine

Web
mittechreview.com.br/minha-conta



MIT Technology Review

Avenida Rio Branco, 108
10º andar - Centro
CEP: 20040-001
+55 21 2224-4525

Operação de vendas avulsas
Distribuição em Bancas
BrancaLeone

Distribuidora
EDICASE GESTÃO DE NEGÓCIOS
www.edicase.com.br

Nossa missão é inspirar a inovação e a aquisição de conhecimento, bem como aumentar a conscientização sobre o poder da tecnologia na sociedade, das ciências humanas e negócios, a fim de construir um futuro melhor para os amantes e líderes de tecnologia de língua portuguesa.

Technology Review, Inc., is an independent nonprofi t 501(c)(3) corporation wholly owned by MIT; the views expressed in our publications and at our events are not always shared by the Institute.



Capa

25

O SOM AO REDOR

Por Natanael Damasceno

38

TR Q+A TALITA BRUZZI TALIBERTI: A SIMPLICIDADE DA VOZ

Por Carlos Aros

04

Campeão de audiência e faturamento

O mercado de podcasts cresceu a passos largos durante a pandemia e a perspectiva é de que seja ampliado ainda mais, muito além dos U\$S 11,4 bilhões que vale hoje.

Por Carlos Aros

12

Os dubladores de Inteligência Artificial soam mais humanos do que nunca — e estão prontos para contratação

Uma nova onda de startups está usando o *deep learning* para construir dubladores sintéticos para assistentes digitais, personagens de videogame e vídeos corporativos.

Por Karen Hao

20

O boom dos jogos por voz está dando às crianças um descanso das telas

Eles estão surgindo como uma opção atraente para os pais preocupados com os efeitos do tempo de exposição às telas na saúde das crianças.

Por Tanya Basu

45

Voz: humana vs sintética. Qual é a melhor aplicação de cada uma?

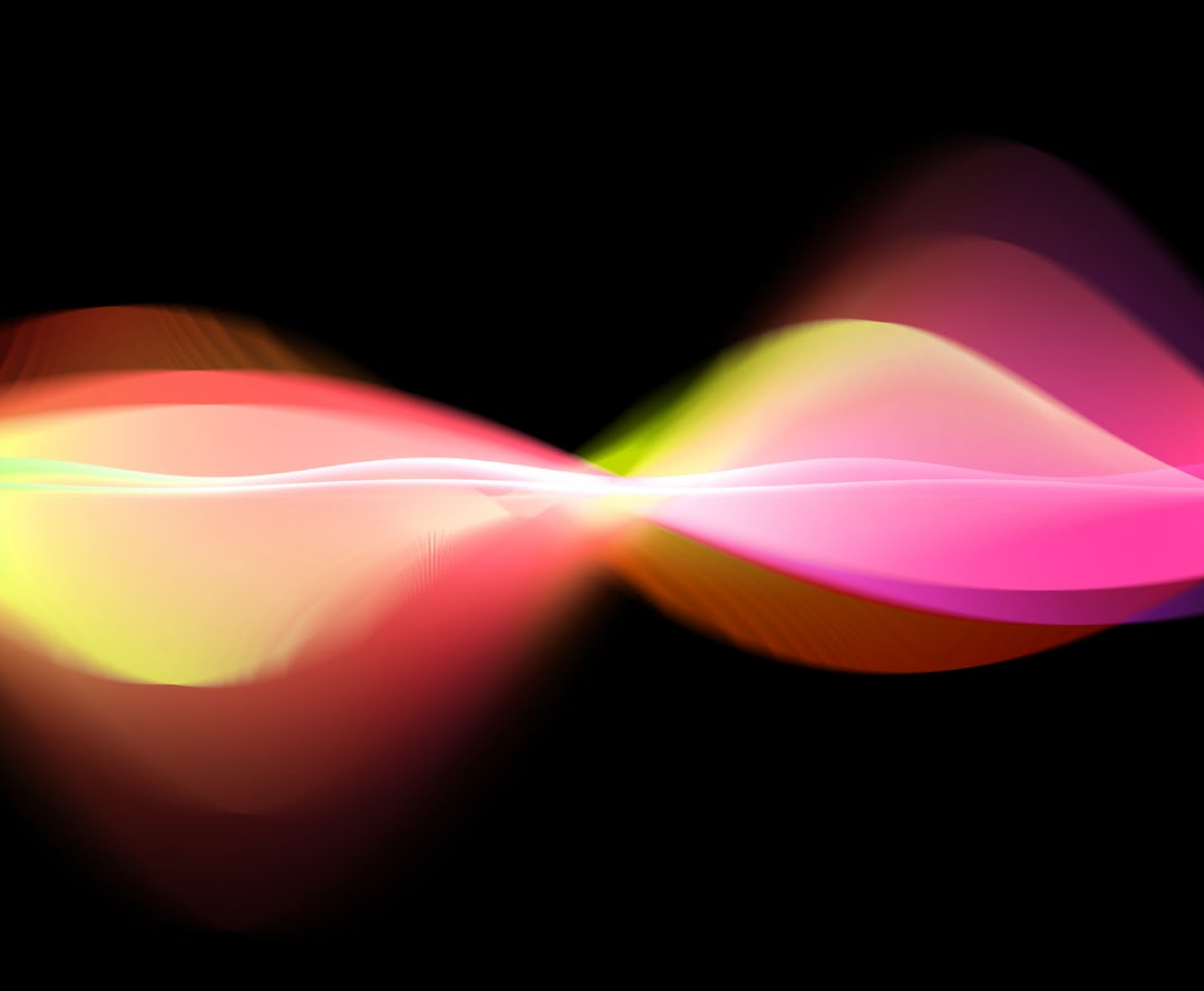
Por Simone Kliass e Jason Bermingham

51

As mulheres continuam sendo apagadas da indústria da tecnologia

A decisão de TikTok de usar a voz de uma mulher sem sua permissão é apenas um exemplo recente de um problema que alguns pensam erroneamente que já está superado.

Por Mar Hicks



Campeão de audiência e faturamento

O mercado de podcasts cresceu a passos largos durante a pandemia e a perspectiva é de que seja ampliado ainda mais, muito além dos U\$S 11,4 bilhões que vale hoje.



Por **Carlos Aros**
Editor-executivo da MIT Technology Review Brasil e comentarista de tecnologia no Grupo Jovem Pan

Foto por **Koy**

Qual podcast você está ouvindo atualmente? Qual a sua plataforma de podcasts preferida? Que formato você mais gosta: entrevista ou documentário? Talvez fossem dúvidas difíceis de responder há três ou quatro anos, mas é inevitável que, hoje em dia, você consiga abordar pelo menos um dos três questionamentos acima. Isso porque os podcasts — vou aqui, propositalmente, evitar definições e conceitos — consolidaram-se como um meio importante para informação, entretenimento e, acredite, redução do estresse provocado pela

enxurrada de informação e pela multiplicação das telas ao nosso redor.

Embora não sejam novos, com uma história que começa no início dos anos 2000, os podcasts ganharam uma nova perspectiva com o surgimento e consolidação de plataformas como Spotify e Deezer, que abriram caminho para a criação de hábito de consumo a partir da música — logo em seguida, inseriram podcasts no cotidiano das pessoas. E esse foi o empurrão necessário para a criação de um mercado que deve crescer cerca de 31% ao ano até 2028, sendo que, em 2020, o setor foi avaliado em U\$S 11,4 bilhões, segundo a empresa de pesquisas Grand View Research.

Essa indústria tem players como o próprio Spotify, que reuniu um grupo diverso de criadores de conteúdo para assinar produções originais da plataforma, além de produtoras que entregam

conteúdos originais e também branded content para marcas dispostas a entrar nesse canal de contato estreito com o consumidor, já que os podcasts têm audiências fiéis. Segundo a pesquisa Culture Next, 2/3 dos millennials e da Geração Z no mundo todo ouvem podcasts semanalmente.

A PANDEMIA E AS OPORTUNIDADES PARA OS PODCASTS

Quando, em março de 2020, o mundo alterou o seu compasso para se proteger de uma ameaça até então desconhecida, ninguém imaginava que as mudanças seriam tão profundas e que o enfrentamento ao desafio demoraria tanto tempo. Entre as adaptações, um aumento considerável do tempo gasto em frente aos mais diferentes tipos de tela. Prova disso é que a



Microsoft registrou, apenas um mês após o início da pandemia, aumento de 70% no número de usuários do Skype, enquanto o Zoom experimentou crescimento de 169% nos três primeiros meses de 2020.

E quando se tornou insuportável encarar tantas telas — até cunhamos um termo para isso, a fadiga do Zoom —, corremos para os fones de ouvido. No período livre, mas ainda em isolamento, as músicas e os podcasts foram companheiros importantes para aplacar o sentimento de solidão de 66% das pessoas da Geração Z entrevistadas pelo Spotify e, ao mesmo tempo, permitiram mergulho profundo em um universo de cores e imagens criadas pela nossa mente.

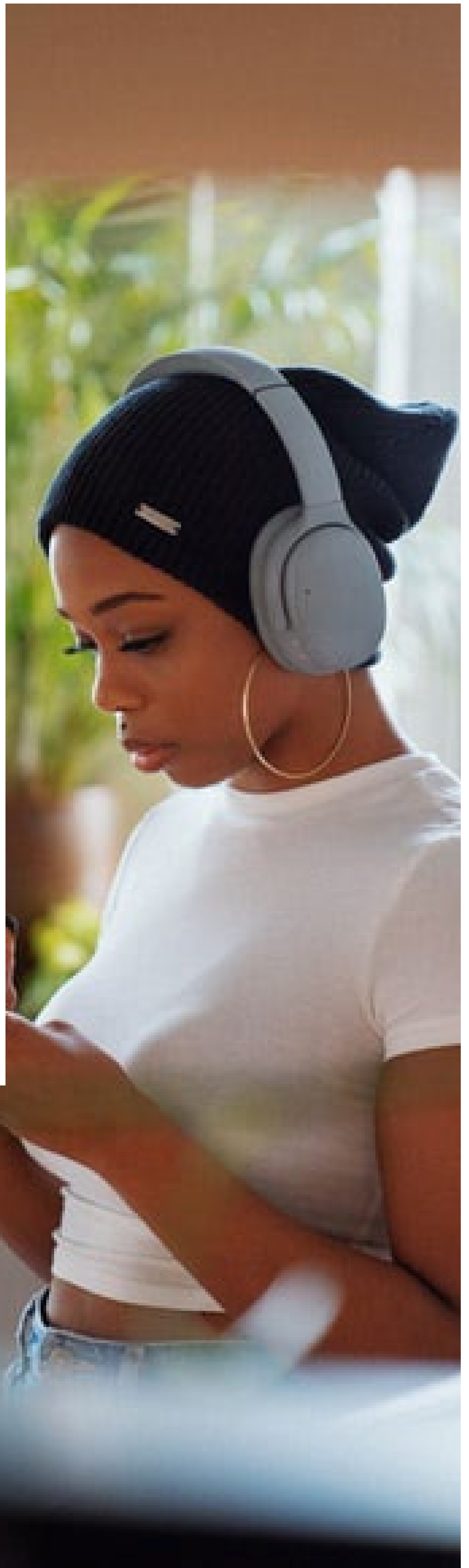
Para além do escapismo e do combate à solidão pandêmica dos adultos, os conteúdos em

áudio ajudaram a diminuir a exposição das crianças às telas. Como contamos recentemente aqui na MIT Technology Review Brasil, plataformas interativas por voz ganharam destaque justamente pela possibilidade de exercitar a criatividade e estimular o desenvolvimento de outras habilidades, e ainda por cima reduzem a exposição dos mais jovens às telas. Mas esse é um tópico para outra conversa...

Um estudo realizado pelo Ibope mostra que 57% dos entrevistados começaram a ouvir programas em áudio digital durante a pandemia. Em um universo de 100 milhões de brasileiros consumindo conteúdos nesse formato, 28 milhões se assumem ouvintes de podcasts. Os números provam que há um grande mercado consumidor que ainda será impactado pelo formato e por suas variações.

E uma das razões possíveis para o sucesso dos podcasts

é a proximidade com a audiência. Assim como o rádio, que também se baseia no áudio como recurso para informar, entreter, emocionar e engajar, os podcasts não criam barreiras para que a comunicação seja efetiva. É tudo simples, direto e ao pé do ouvido. No Brasil, um dos formatos mais bem aceitos pelo público é de “entrevista” ou “mesa redonda”, em que duas ou mais pessoas se reúnem para uma conversa. Mas o tempo é ainda o definidor do formato. Em geral, os episódios com duração média de trinta minutos ganham a preferência da maioria, ou 31% dos entrevistados pelo Ibope.





AUDIÊNCIA, ENGAJAMENTO E A PRESENÇA DAS MARCAS

Entre as características mais valorizadas pela audiência de podcats, considerando os millennials e a Geração Z pesquisados pelo Spotify, estão a naturalidade, a espontaneidade e a simplicidade adotadas pelos criadores de conteúdo. Outro ponto é a confiança depositada no que os criadores e seus convidados apresentam.

“Entre os millennials e a Geração Z do Brasil, 34% disseram que confiavam mais em podcasts do que em meios de comunicação tradicionais, incluindo noticiários de TV, jornais e rádios nacionais”, diz o estudo Culture Next.

Essa percepção de maior

confiança pode estar amparada, entre outros aspectos, pelo fato de que os criadores “falam a língua” do público. São, como já descrito aqui, “figuras reais”, que se posicionam de forma horizontal e dividem dúvidas, angústias e aspirações com essas gerações que compõem parte considerável da audiência. E essa conexão genuína que surge entre os fãs e os criadores tem um valor enorme, em especial numa época em que a “influência digital” se tornou ativo valioso e amplamente explorado pelas marcas.

O formato tradicional de publicidade pode até não ser bem aceito pela audiência, em especial nas plataformas digitais. A interrupção do conteúdo para a exibição de anúncios é incômoda, mas os spots dentro dos podcasts respondem por 76% das receitas, segundo pesquisa da PwC para o Interactive Advertising Bureau (IAB). Além do spot comercial, outro aspecto



que o podcast incorpora do rádio é o merchandising, quando o apresentador se envolve na apresentação do comercial.

“Os podcasts demonstraram que são um ambiente particularmente eficaz para os anúncios: 41% de todos os ouvintes dizem que confiam mais nos anúncios se os escutarem durante um podcast e 81% afirmam que realizaram alguma ação após escutar um anúncio em um podcast”, salienta o Spotify no estudo Culture Next.

A audiência cresceu e o interesse das marcas também, tanto que o Spotify lançou uma plataforma para que os anunciantes tenham mais opções para inserção de publicidade em conteúdos originais ou exclusivos, com métricas e informações que ajudam no direcionamento estratégico da campanha.

Sinal de maturidade do mercado de podcasts e do interesse das empresas.

Mas, de uns tempos para cá, com a evolução do marketing de conteúdo, as marcas resolveram criar produtos próprios para garantir mais do que simplesmente alguns minutos de visibilidade. Como geradoras de conteúdo — e não só anunciantes de um produto ou serviço —, as marcas têm o desafio de entregar relevância à audiência e estabelecer uma conexão tentando imprimir as características valorizadas pelo público (naturalidade, simplicidade, etc). A forma buscada, na maior parte dos casos, é a associação com personalidades que já têm boa aceitação e engajamento com o público. Os influenciadores hoje também “emprestam influência” por meio de podcasts que apresentam, em projetos de longa ou curta duração.



ASSISTENTES VIRTUAIS E DIVERSIDADE NO CONTEÚDO

Temos vivenciado, ano após ano, grandes transformações no mercado de podcasts. Seja pela melhora na qualidade das produções, pelo aumento da audiência ou ainda pelo crescimento da receita, a evolução é notável. Novos formatos foram inseridos no mercado, cada vez mais a segmentação de público tem sido valorizada pelas produções e há mais diversidade entre os temas e criadores.

Avanços notáveis, claro, mas ainda é preciso avançar mais. E a jornada iniciada pela mudança de hábitos provocada pela pandemia mostra o caminho a seguir. As pessoas, no mundo todo, descobriram os podcasts e também que as interfaces controladas por voz são presenças constantes em

casa, no carro, no trabalho e onde mais se puder imaginar. Esses dispositivos priorizam a experiência para o áudio, assim como os podcasts, e dessa sinergia podem derivar bons resultados.

Google e Amazon lideram essa frente porque têm as plataformas (Google Assistente e Alexa) e os dispositivos (Google Home e Echo). Correm por fora, seja pela preferência do público ou pela ampla oferta de conteúdo original e exclusivo, Spotify e Apple, que lideram o mercado global de podcasts. Otimizar conteúdos, para que tenham bom desempenho nos dispositivos inteligentes de áudio, não só vai colocar os podcasts em um habitat preparado para eles como vai potencializar a experiência do ouvinte. Quem já utiliza Alexa ou Google Assistente sabe o que isso representa.

Em outra frente, tão importante quanto a anterior, o investimento em conteúdo relevante e de qualidade



deve ser constante. E essa relevância está diretamente relacionada com o interesse que a audiência tem sobre os temas e as pessoas envolvidas. A diversidade é um caminho para isso, porque o olhar do criador de conteúdo precisa estar na mesma direção que o da audiência.

Por isso, é fundamental que os criadores continuem surgindo do mesmo contexto que o nicho de público deles. Isso cria conexão, aumenta o engajamento e a relevância. É bom para o criador, para a audiência e, por consequência, para as marcas. ■





Os dubladores de Inteligência Artificial soam mais humanos do que nunca — e estão prontos para contratação

Uma nova onda de startups está usando o *deep learning* para construir dubladores sintéticos para assistentes digitais, personagens de videogame e vídeos corporativos.

Ilustração por **Alessandra Benigno**

A postagem no blog da empresa transborda entusiasmo, como aquele dos infomerciais dos anos 90. WellSaid Labs descreve o que os clientes podem esperar de seus “oito novos atores de voz digital!”. Tobin é “dinâmico e perspicaz” Paige é “equilibrada e expressiva”. Ava é “elegante, autoconfiante e profissional”.

Cada um é baseado em um dublador real, cuja semelhança (com consentimento) foi preservada usando



Inteligência Artificial.

As empresas agora podem licenciar essas vozes para dizer tudo aquilo que for precisarem. Elas simplesmente inserem o texto no mecanismo de voz, e logo obtêm um trecho de áudio nítido com uma performance de som natural.

WellSaid Labs, uma startup sediada em Seattle, Estados Unidos, que se originou da pesquisa sem fins lucrativos da Allen Institute of Artificial Intelligence, é a mais recente empresa que oferece vozes de IA aos clientes. Por enquanto, é especializada em vozes para vídeos corporativos de *e-learning*. Outras startups fazem vozes para assistentes digitais, operadores de call center, e até mesmo personagens de videogame.

Não muito tempo atrás,

essas vozes *deepfake* tinham uma péssima reputação por serem usadas em fraudes telefônicas e truques na internet. Desde então, a melhora da qualidade, contudo, despertou o interesse de um número crescente de empresas. Avanços recentes no *deep learning* tornaram possível reproduzir muitas das sutilezas da fala humana. Essas vozes pausam e respiram nos momentos certos. Elas podem mudar de estilo ou de acordo com emoções. Você pode identificá-las se elas falarem por muito tempo, mas, em trechos de áudio curtos, algumas se tornaram indistinguíveis das vozes humanas.

As vozes de IA também são baratas, podem ser reproduzidas em larga escala e são fáceis de trabalhar. Ao contrário da gravação de um dublador humano, as vozes sintéticas podem modificar seu roteiro em tempo real,



abrindo novas oportunidades para a personalização da publicidade.

A ascensão de vozes hiper-realistas não é, contudo, isenta de consequências. Os dubladores reais, em particular, estão se perguntando o que isso representa para seus meios de subsistência.

COMO EMULAR UMA VOZ

Vozes sintéticas já existem há algum tempo. As mais antigas, incluindo as vozes da Siri e Alexa originais, simplesmente colavam palavras e sons, obtendo um efeito robótico e desajustado. Fazer com que soassem mais naturais era uma tarefa manual trabalhosa.

O *deep learning* modificou isso. Os desenvolvedores de voz não precisavam mais ditar o ritmo, a pronúncia ou a entonação exata da

fala gerada. Em vez disso, eles podiam alimentar algumas horas de áudio em um algoritmo e fazer com que o algoritmo aprendesse esses padrões por conta própria.

Ao longo dos anos, os pesquisadores usaram esse princípio básico para construir mecanismos de voz cada vez mais sofisticados. O que a WellSaid Labs construiu, por exemplo, faz uso de dois modelos principais de *deep learning*. O primeiro prevê, a partir de uma passagem de texto, os pontos gerais de como o locutor soará, incluindo sotaque, tom e timbre da voz. A segunda completa os detalhes, como a respiração e a maneira como a voz ressoa no ambiente.

No entanto, fazer uma voz sintética convincente exige mais do que um simples apertar de botão. Parte do que torna a voz humana tão humana é sua inconsistência,



expressividade e capacidade de entregar as mesmas falas em estilos completamente diferentes, dependendo do contexto.

Capturar essas nuances envolve encontrar os dubladores certos para fornecer os dados de treinamento apropriados e ajustar os modelos de *deep learning*. WellSaid diz que o processo requer pelo menos uma ou duas horas de áudio e algumas semanas de trabalho para desenvolver uma réplica sintética que pareça realista.

As vozes de IA tornaram-se populares entre as marcas que buscam manter um som consistente nas milhões de interações com os seus clientes. Hoje, com a onipresença de alto-falantes inteligentes e o surgimento de agentes de atendimento ao cliente automatizados, bem como assistentes digitais embutidos em carros e dispositivos inteligentes,

as marcas podem precisar produzir mais de cem horas de áudio por mês. Elas também não querem mais usar as vozes genéricas oferecidas pela tecnologia tradicional de conversão de texto em voz — uma tendência que se acelerou durante a pandemia à medida que mais e mais clientes abandonaram as interações na loja para interagir virtualmente com as empresas.

“Se eu sou a voz da Pizza Hut, certamente não posso soar como a da Domino’s, e certamente não posso soar como a da Papa John’s”, diz Rupal Patel, professor da Northeastern University e fundador e CEO da VocaliD, que promete construir vozes personalizadas que correspondem à identidade da marca de uma empresa.

“Essas marcas pensaram nas cores. Eles pensaram nas fontes. Agora eles precisam começar



a pensar sobre a maneira como a voz delas soa também”.

Enquanto as empresas costumavam contratar diferentes locutores para diferentes mercados — Nordeste versus Sul dos EUA ou França versus México — algumas empresas de IA de voz podem manipular o sotaque ou mudar a linguagem de uma única voz de maneiras diferentes. Isso abre a possibilidade de adaptar anúncios em plataformas de *streaming* dependendo de quem está ouvindo, mudando não apenas as características da voz, mas também das palavras que estão sendo faladas. Um anúncio de cerveja poderia dizer a um ouvinte para dirigir-se a um bar diferente, dependendo se está sendo transmitido em Nova York ou em Toronto, por exemplo. A Resemble.ai, que projeta vozes para publicidades e assistentes inteligentes,

diz que já está trabalhando com clientes para lançar essas propagandas de áudio personalizados no Spotify e no Pandora.

As indústrias de jogos e entretenimento também estão sendo beneficiadas. Sonantic, uma empresa especializada em vozes emotivas que podem rir, chorar, sussurrar ou gritar, trabalha com fabricantes de videogames e estúdios de animação para fornecer as vozes de seus personagens. Muitos de seus clientes usam as vozes sintetizadas apenas na pré-produção e mudam para atores de voz reais na produção final. A Sonantic afirma que alguns começaram a usá-las ao longo de todo o processo, possivelmente em personagens com um número menor de falas. A Resemble.ai e outros também trabalharam com filmes e programas de TV para consertar a performance dos atores quando as palavras são

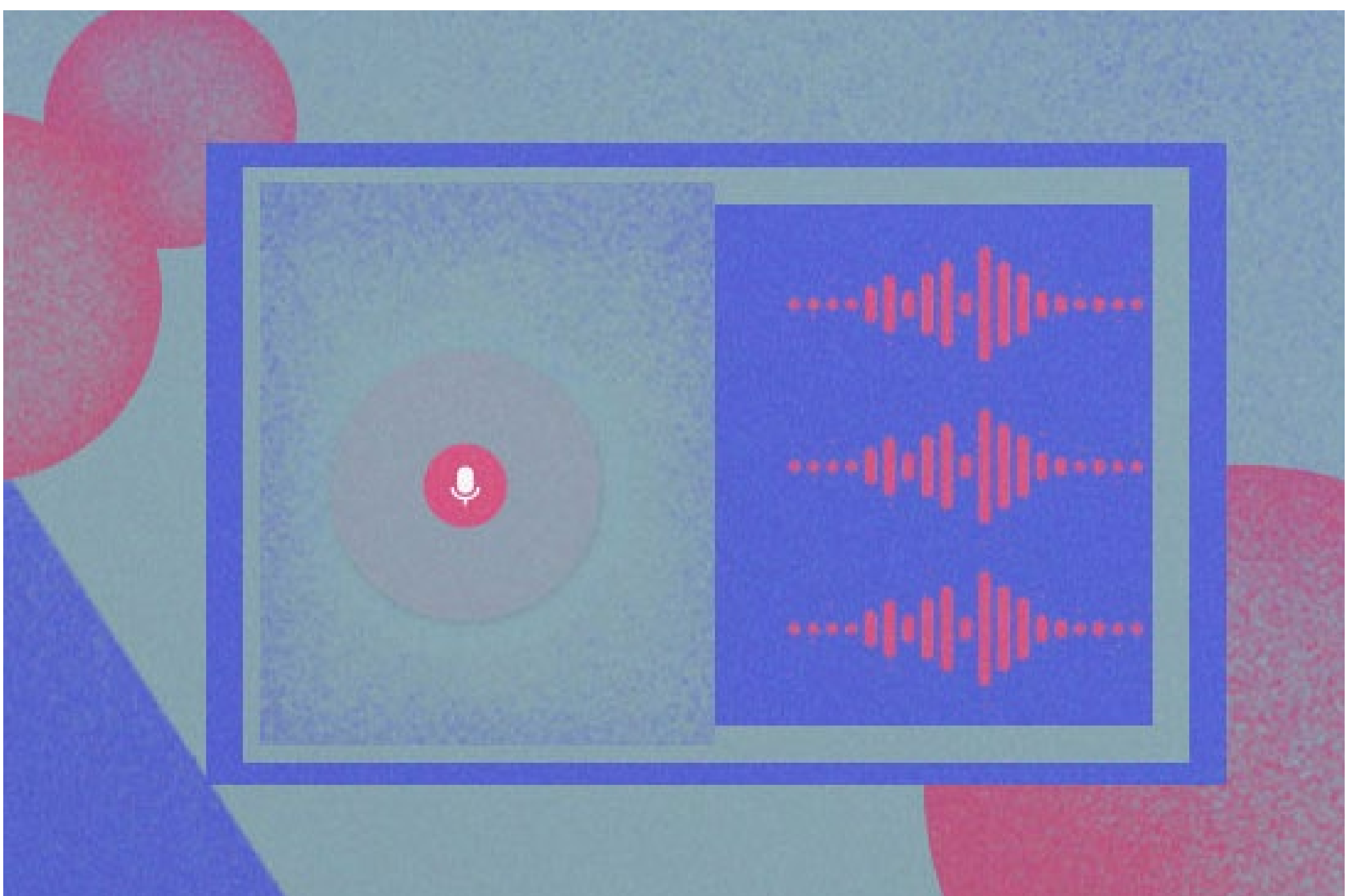
distorcidas ou pronunciadas incorretamente.

Existem, todavia, limitações para o quão longe a IA pode chegar. Ainda é difícil manter o realismo de uma voz por longos períodos de tempo, o que pode ser necessário para um *audiobook* ou um podcast. Há, também, pouca capacidade de controlar o desempenho de uma voz de IA tal qual um diretor faria com um artista. “Ainda estamos nos primórdios da fala sintética”, diz Zohaib Ahmed, o fundador e CEO da Resemble.ai,

comparando-os aos dias em que a tecnologia CGI era usada principalmente para retoques, em vez de criar mundos inteiramente novos a partir de telas verdes.

UM TOQUE HUMANO

Em outras palavras, os dubladores humanos não se tornarão obsoletos ainda. Projetos expressivos, criativos e longos ainda são melhor executados por humanos. Além disso, para cada voz sintética feita





por essas empresas, um dublador também precisa fornecer os dados de treinamento originais.

Alguns atores, contudo, estão cada dia mais preocupados com seus meios de subsistência, disse um porta-voz do SAG-AFTRA, o sindicato que representa os dubladores nos Estados Unidos. Se eles não têm medo de serem automatizados pela IA, eles estão preocupados em serem pouco recompensados ou em perder o controle sobre suas vozes, que constituem sua marca e reputação.

Esta é a temática de uma ação judicial contra o TikTok movida pelo dublador canadense Bev Standing, que alega que o recurso de narração embutido no aplicativo usa uma cópia sintética de sua voz sem sua permissão. A experiência de Standing também se assemelha a de Susan

Bennett, a voz original da americana Siri, que foi paga por suas gravações iniciais, mas não pelo uso contínuo de sua semelhança vocal em milhões de dispositivos Apple.

Algumas empresas estão buscando ser mais responsáveis em como se relacionam com a indústria de dublagem. As melhores, diz o representante da SAG-AFTRA, procuraram o sindicato para descobrir a melhor maneira de recompensar e respeitar os dubladores por seu trabalho.

Muitas agora usam um modelo de participação nos lucros para pagar aos atores cada vez que um cliente licencia sua voz sintética específica, o que abriu um novo fluxo de receita passiva. Outros envolvem os atores no processo de criação de sua imagem de IA e lhes dão poder de veto sobre os projetos em que serão usados.



O SAG-AFTRA também está insistindo por uma legislação que proteja os atores de réplicas ilegítimas de sua voz.

Mas, para Patel da VocaliD, o objetivo das vozes de IA não é, em última análise, replicar o desempenho humano ou automatizar o trabalho de locução existente. Em vez disso, a promessa é que elas poderiam abrir possibilidades inteiramente novas. E se no futuro, diz ela, vozes sintéticas pudessem ser usadas para adaptar rapidamente materiais educacionais on-line para públicos diferentes? “Se você está tentando alcançar, digamos, um grupo de crianças do centro da cidade, não seria ótimo se essa voz realmente soasse como se fosse de sua comunidade?”. ■





O boom dos jogos por voz está dando às crianças um descanso das telas

Eles estão surgindo como uma opção atraente para os pais preocupados com os efeitos do tempo de exposição às telas na saúde das crianças.

Em uma casa em Massachusetts, nos Estados Unidos, uma família se reúne ao redor da mesa da cozinha. Os pais ligam o aparelho que fica na cabeceira da mesa. Segundos depois, uma voz alegre

silencia a conversa e toda a família se senta para ouvir. Não, esta não é uma cena da era dourada do *wireless*. Esta é a casa dos Daniels durante o confinamento da Covid-19, e o foco da atenção da família é o Yoto, um dispositivo de áudio projetado para competir com os aplicativos e vídeos que as crianças modernas costumam usar. Não é diferente de um tocafitas: os usuários podem clicar em um cartão pré-gravado com uma história ou jogo. O Yoto também pode



reproduzir novos episódios diários, como o noticiário infantil que a família Daniels ouve todas as manhãs. “Não ter uma tela foi o grande atrativo”, diz Brian Daniels, o pai.

“Promover o pensamento criativo e a diversão era importante para nós. Queríamos algo que pudesse prender a atenção deles e ainda ser produtivo”.

Os estudos científicos que pesquisam sobre se o tempo na tela é realmente ruim para as crianças ainda estão incertos, e muitos especialistas acreditam que os perigos foram exagerados. Apesar disso, muitos pais ainda estão preocupados com o fato de seus filhos passarem muito tempo olhando para tablets e televisores. Para essas famílias, os jogos por voz e áudio parecem uma alternativa mais saudável. Como resultado, os jogos de voz para toda a família

dispararam nas listas de download durante este ano pandêmico.

Muitos desses jogos são executados em assistentes de voz como Alexa, da Amazon, ou Google Assistant. Quando noticiamos pela primeira vez os jogos por voz em 2019, a maioria deles tomou a forma de “escolha-sua-própria-história-de-aventura” e competições de trivialidades. Agora eles estão ficando mais sofisticados. O jogo Lemonade Stand, por exemplo, permite que as crianças pratiquem a administração de uma empresa, enquanto no Kids Court, um assistente de voz julga as discussões entre as crianças fazendo-as falar sobre suas diferenças. A maioria dos jogos de voz pode ser jogada de forma assíncrona, o que significa que os jogadores não precisam estar no mesmo lugar ou jogar ao mesmo tempo. Max Child, o fundador da Volley, uma



empresa de jogos de voz, diz que os avós costumam jogar em seu tempo livre e enviar mensagens de texto aos netos sobre seu progresso, e as crianças avançam quando podem.

“Uma pessoa poderia brincar com os filhos ou pode reunir os amigos para um anoite de jogos”, diz Child. Um dos jogos de voz mais populares do Volley é Yes Sire, um conto envolvente que faz com que o jogador se imagine como o governante implacável de um feudo.

Os pais também gostam do fato de que muitos jogos de voz oferecem uma tela em branco para a criatividade. Parker, o filho de seis anos de Kate e Brian Daniels, é um artista ávido. “Ele é nosso filho criativo”, diz Kate, descrevendo como ele ouve as histórias e depois as representa na praia em frente à casa da família com sua irmã de quatro anos, Charlotte. “Gosto de pensar em como

posso mudar a história”, interrompe Parker, descrevendo a vez na qual ouviu um conto sobre um castelo e imaginou como seria.

À noite, as crianças clamam pela continuação da história de Brian para dormir, que ele grava como um MP4 e carrega em cartões em branco para elas ouvirem se tiverem se comportado.

Mas, não são apenas as crianças que estão evitando a tela; os adultos também são um mercado crescente para jogos de voz. O Volley, por exemplo, tem alguns jogos “maduros”, como Love Taps, Sherlock e Infected, que incluem linguagem explícita e conteúdo voltado para um público mais velho. E tanto Child quanto o criador do Yoto, Ben Drury, dizem que, além de aproveitar a oportunidade de participar com os netos, os mais velhos também descobriram que é empoderador poder jogar com suas vozes em vez de



aprender a navegar em um *console* ou usar um controle.

Os jogos de voz não são apenas uma questão de competição. No início deste ano, Nina Meehan e Jonathan Shmidt Chapman, ambos profissionais do teatro juvenil, criaram o “K’ilu Kit: Passover Adventure” para o feriado judaico que ocorre em abril. Eles perceberam que, pelo segundo ano consecutivo, a pandemia interromperia a reunião usual de familiares e amigos para o Séder de Pessach, o jantar ritual no qual recontam a história do êxodo judeu do Egito.

“E não é a mesma coisa no Zoom”, diz Shmidt Chapman. “Muitas dessas histórias bíblicas são difíceis de explicar para uma criança de três a oito anos. Como podemos transmitir essa história de uma maneira apropriada para a idade?”. O K’ilu Kit tenta tornar a história do êxodo significativa,

compreensível e divertida para as crianças com a ajuda de elementos interativos: uma chama de papel enrolada em uma lanterna torna-se a sarça ardente através da qual Deus diz a Moisés para conduzir os israelitas para fora do Egito, por exemplo.

“A experiência de áudio orienta as crianças a fazerem as coisas fisicamente com instruções, em vez de apenas ouvir”, diz Meehan. “A história da Páscoa judia é a história do reconhecimento de tópicos complexos sobre a liberdade da escravidão e da opressão. É assim que as crianças podem aprender a história da Páscoa. Não se trata apenas de olhar para uma tela ou ouvir a história e sim dos níveis de importância e de compreensão”.

O entretenimento conduzido por voz é o único capaz de fornecer esse tipo de compreensão, de acordo com Naomi Baron, professora emérita de Linguística



da American University e autora de *How We Read Now: Strategic Choices for Print, Screen, and Audio*.

“A preocupação com o tempo de exposição a telas não tem sido apenas com as horas em que nossos olhos ficam grudados nela, mas a superficialidade da maior parte da interação”, diz Baron.

“Você não está fazendo esforço mental”.

Com histórias de áudio e jogos, no entanto, as informações não são apresentadas a você em uma bandeja. A imaginação é necessária e requer mais foco e atenção do que olhar para uma tela. Baron afirma que a pesquisa mostrou que, com esse tipo de aprendizado, a compreensão e a recordação são muito maiores para o desenvolvimento de leitores. Ela acrescenta que os ouvintes mais velhos também podem se beneficiar,

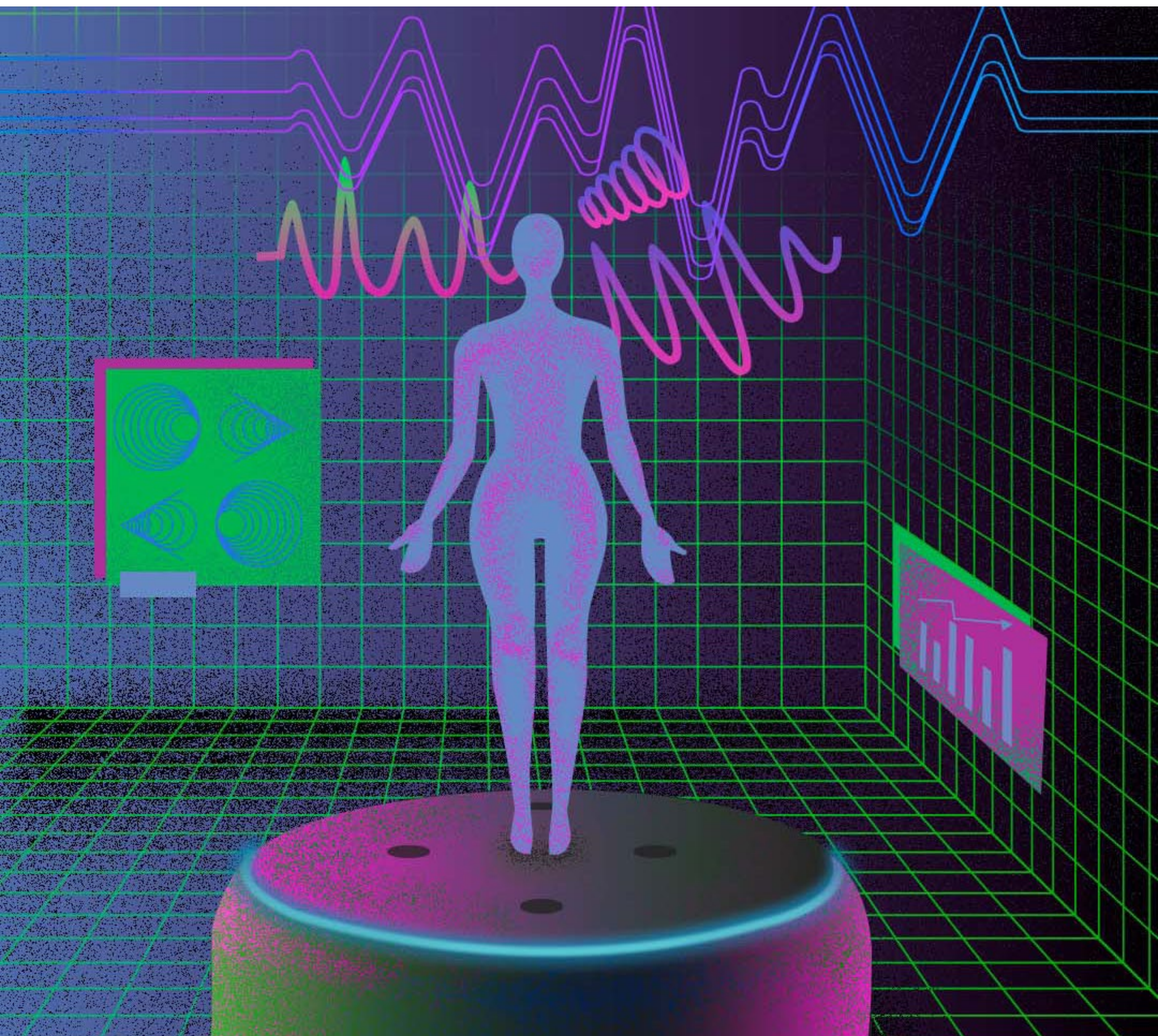
principalmente se o inglês não for sua primeira língua, seu estilo de aprendizagem for menos visual ou se eles tiverem alguma deficiência visual.

Se o tempo de exposição a telas é “bom” ou “ruim” ainda é discutível, e é muito cedo para dizer se a popularidade dos jogos de áudio e voz durante a pandemia vai acabar, pois as vacinas tornam possível mais uma vez os encontros presenciais. Eles não são perfeitos. Os jogos de voz costumam os usuários, principalmente crianças que estão aprendendo a pronunciar palavras e a interagir com a tecnologia. Os Daniels, no entanto, dobraram seu consumo por áudios. A família comprou recentemente seu segundo Yoto, que Price, um bebê de 21 meses, descobriu como usar. “Ele canta junto com o aparelho. Ele adora”, diz Kate. Charlotte concorda: “Eu adoro porque toca música e histórias”. ■

O som ao redor

O refinamento das tecnologias de voz e as inúmeras possibilidades de aplicação no dia a dia apontam para caminhos promissores para empresas e reforçam o debate sobre a necessidade de regulação neste universo.

Por Natanael Damasceno





O encontro da voz com as tecnologias digitais, embora não seja uma novidade, tem apontado caminhos inovadores para esta que é uma das formas mais primordiais de comunicação. O último ano, aliás, tão marcado pelas restrições impostas pela pandemia do novo coronavírus, pode até ser considerado um ponto de virada para estas tecnologias e para a própria voz. Se por um lado, o confinamento forçado jogou luz sobre as possibilidades do seu uso nas interações e nas dinâmicas sociais; por outro, a urgência em tornar os dias mais leves diante da nova rotina chamou a atenção para como a voz aliada à Inteligência Artificial pode entreter e facilitar tarefas rotineiras.

Isso pode ser notado, por exemplo, no número de usuários da Internet que usam a voz para acionar mecanismos de busca ou que utilizam comandos de voz em seus dispositivos,

sejam eles smartphones ou desktops: são 45,3% em todo o mundo e 40,2% no Brasil, de acordo com uma pesquisa conduzida em 2020 pela GWI com usuários da Internet com idade entre 16 e 64 anos, e publicada na edição de 2021 do relatório Global Digital, produzido anualmente pela agência We Are Social em parceria com a Hootsuite. São números que, comparados às edições anteriores, apontam para uma tendência de crescimento do uso da voz como interface de comunicação (foram 43% e 36%, respectivamente, registradas no relatório de 2020, ante 39% e 34%, registrados na pesquisa do ano anterior). O mesmo documento mostra que, no Brasil, mais de 3/4 dos entrevistados (78,4%) afirmaram que ouvem música por meio de serviços de *streaming* e pouco menos da metade (44,4%) dizem que ouvem podcasts.

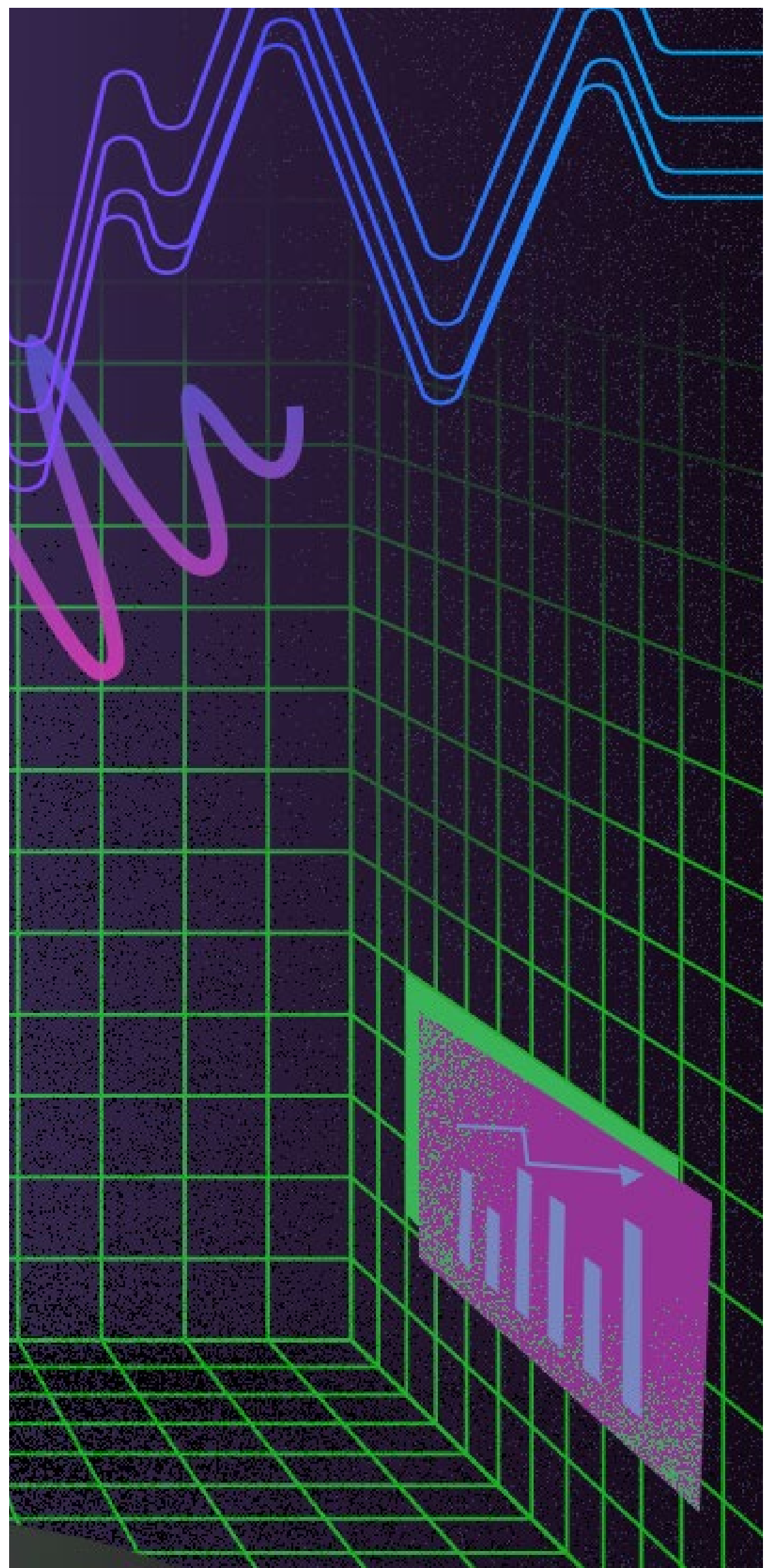
Segundo o Google, o Brasil é o terceiro país com mais usuários ativos do Google Assistant. E, de acordo com

os dados que a empresa reúne a partir das atividades dos usuários, mais de 60% dos brasileiros já usaram a voz para controlar celulares, para obter orientações no trânsito, fazer buscas na internet e reproduzir música. Mais de 90% já enviaram alguma mensagem de voz usando o smartphone. E mais de um terço dos entrevistados usam o Google Assistant pelo menos três vezes na semana, sendo que 21% usam diariamente. Daí se depreende que os brasileiros gostam de usar a voz para se comunicar.

Essa percepção de que a voz tem ganhado espaço nos corações e mentes dos consumidores não fica restrita às agências e institutos de pesquisa. Empresas há muito estão atentas ao potencial dessas tecnologias como produto, ferramenta de inovação, meio indutor de vendas e instrumento para fortalecer suas marcas. Na verdade, são muitas as possibilidades, segundo Conrado Caon, entusiasta do tema e CTO da holding AdVentures Inc, que tem desenvolvido diferentes

estratégias com o uso de tecnologias de voz para empresas como Domino's, Mondelez, Disney e outras.

Ele afirma que o potencial de exploração dessas tecnologias pelas empresas é enorme e que, apesar de o Brasil estar, segundo ele, entre os países que mais as utilizam, ainda há um vasto campo a ser desbravado. “Temos um potencial muito grande. Dá para ver que a curva que





retrata o interesse nessas tecnologias é exponencial e está em franco crescimento. E há aí uma relação entre o conteúdo e o interesse das empresas. Quando começamos, era preciso mostrar o que a gente tinha e convencer da importância de investir na inovação. Hoje, os clientes vêm em busca do que viram, do que ouviram falar. E se antes as empresas inovadoras estavam muito mais presentes, hoje as mais tradicionais nos procuram manifestando interesse nessas estratégias”, explica.

Conrado destaca pelo menos seis campos em que as estratégias baseadas em tecnologias de voz têm sido utilizadas com mais sucesso. O *audio interactive*, ou conteúdo baseado em interação; o *voice commerce*, com o uso de recursos de voz como estratégia em *e-commerces*; o *flash briefing* ou *sonic branding*, com o uso do som para reforçar ou compor a identidade das marcas; o *voice search*, ou a otimização para serviços de busca acionados pela voz; as assistentes de voz e

os *voice apps*, que utilizam a tecnologia para criar aplicativos que educam, informam, controlam dispositivos, entretêm e fazem transações financeiras e comerciais. O CTO conta que tem desenvolvido projetos em cada uma delas.

“A ideia de gerar facilidade, uma resposta mais rápida e uma economia de tempo é sempre bem-vinda. Numa economia tão ágil, com todo mundo correndo para fazer mais em menos tempo, a voz traz o benefício da agilidade para a solução de problemas”, diz Caon.

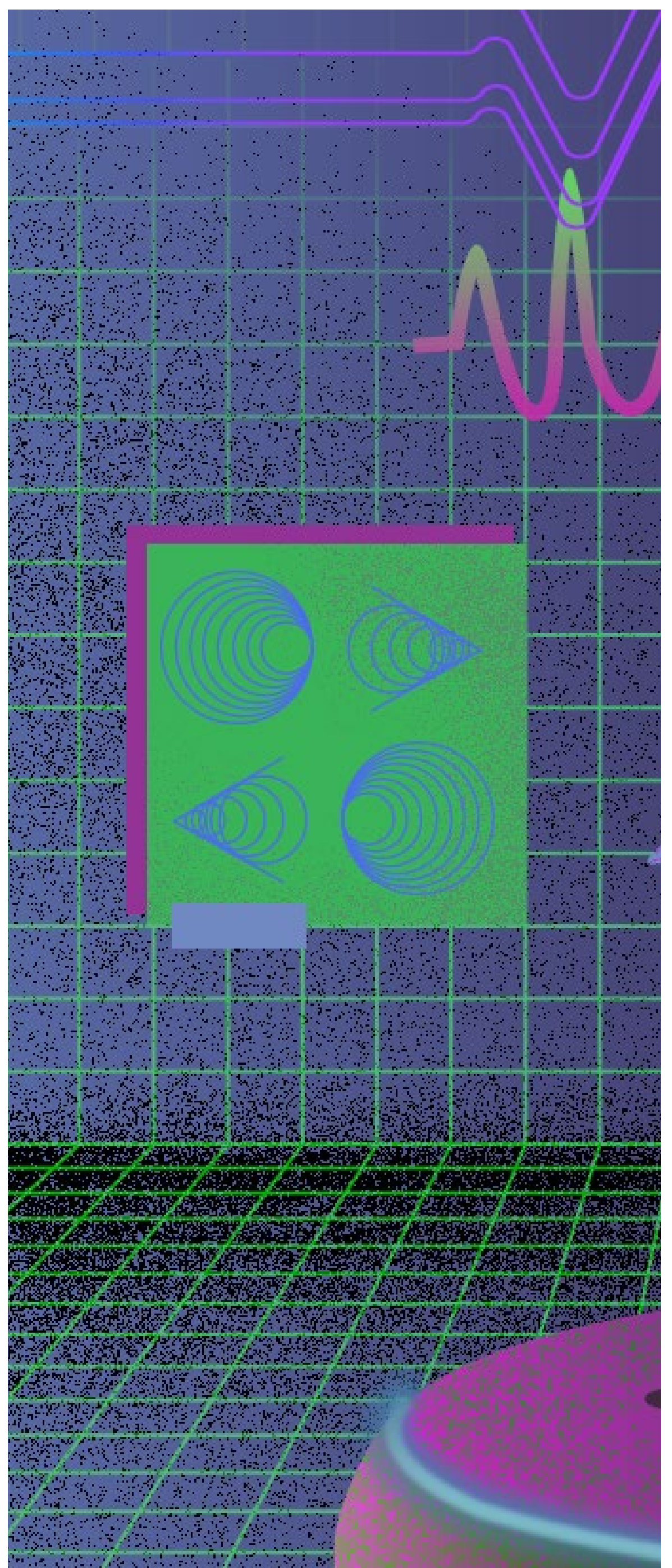
De fato, os brasileiros, cada vez mais familiarizados com as tecnologias de voz, estão mais inclinados ao uso dos assistentes pessoais e, assim, conversam mais com as máquinas, mesmo que seja apenas dentro de casa. A constatação é de uma pesquisa conduzida pela consultoria Ilumeo, especializada em *data*

science, divulgada em março. O levantamento foi feito com base nas respostas de 1.100 entrevistados de todo o Brasil.

Segundo o estudo, 87% dos entrevistados já usaram assistentes virtuais e 92% tinham algum conhecimento da tecnologia. Cerca de 48% utilizam com frequência um assistente virtual comandado por voz e mais da metade dos entrevistados afirmou que tem usado mais as funções de assistente virtual por voz do celular, tem visto mais valor para o uso da voz para controlar os aparelhos e tem visto mais soluções tecnológicas com base em reconhecimento de voz. Além disso, 61% dos entrevistados manifestaram o interesse em usar mais aparelhos eletrônicos comandados por voz.

Muito além de registrar a familiaridade dos entrevistados com as assistentes virtuais, o estudo listou os anseios dos consumidores e mapeou as oportunidades para o desenvolvimento de marcas,

produtos e serviços. Felipe Senise, Head of Strategy da Ilumeo e um dos responsáveis pela condução da pesquisa, chama atenção para algumas das conclusões. A primeira é que o assistente virtual por voz não é apenas uma tecnologia de produto, mas algo que agrega valor às marcas. Ele constatou ainda que o uso geral do





recurso ainda é muito básico, embora as pessoas queiram usá-la em mais situações e com mais frequência, o que criaria oportunidades para além do segmento de produtos eletrônicos. “As pessoas querem pedir em vez de fazer”, explica.

Segundo o estudo, as pessoas gostariam de usar a voz não apenas para interagir com as assistentes virtuais, mas também para ajudar nas atividades rotineiras da casa, permitindo o controle de aparelhos de TV e de eletrodomésticos como o ar-condicionado, o ventilador, chuveiros e torneiras e também do sistema de iluminação. O curioso é que, para alguns desses desejos, já existe tecnologia disponível — como produtos, softwares e plataformas que permitem a automação do controle dos sistemas de água e luz de uma residência. Ou mesmo nos seus carros, com funcionalidades que vão das interações mais simples ao cruzamento de tecnologias como a dos carros autônomos. Além disso, querem poder solicitar

serviços de delivery, bancários, e serviços educacionais e de entretenimento por meio de comandos de voz. Segundo Felipe, outra percepção importante é a de que, apesar da privacidade levantar questionamentos, ela não se apresenta como uma grande barreira de uso como o preço.

Os entrevistados afirmaram ainda que gostariam de ver a tecnologia associada aos aplicativos de transporte, embora a maioria dos aplicativos, como a Uber e a 99, já possua funcionalidade integrada aos assistentes dos smartphones Apple ou Android. Um sinal de que não basta desenvolver a tecnologia. É preciso um esforço maior de comunicação.

Ainda no campo das possibilidades, a consultoria britânica Juniper Research estima que o *voice commerce* movimentará US\$ 80 bilhões em 2023. O uso das tecnologias de voz pelos *e-commerces*, a propósito, é uma das tendências apontadas por um relatório



publicado este ano pela Rain Agency que, a partir do acompanhamento do que foi desenvolvido nesse sentido em 2020, faz uma projeção do que deve ganhar relevância este ano em termos de uso da voz nas tecnologias digitais.

O relatório diz que, no campo voltado para o consumidor, as empresas já têm uma série de opções para aumentar as vendas, sensibilizar os clientes e fazer um trabalho consistente de fortalecimento de suas marcas. No varejo, por exemplo, a tecnologia de voz deixou sua marca durante a pandemia, uma vez que a Covid-19 mudou o paradigma tradicional de compras.

A voz se mostrou uma tecnologia valiosa para as empresas, criando experiências sem atritos que limitam a interação direta entre as pessoas.

A projeção é que, num futuro próximo, estas tecnologias possam incrementar a

eficiência dos funcionários e melhorar ainda mais a experiência dos clientes nas lojas.

Outro exemplo é o amadurecimento e o avanço dos anúncios de voz interativos. Embora a tecnologia não seja exatamente nova, ela tem sido utilizada com mais frequência por grandes empresas, incluindo Amazon e Spotify. A previsão é que, à medida que os serviços de *streaming* de áudio ganhem espaço no mercado, o uso de anúncios de voz interativos cresça junto com ele.

O estudo cita ainda, entre outras tendências, a incorporação da voz em *wearables* avançados e uma consequente mudança de comportamento possibilitada pelo aumento da capacidade dos usuários de personalização e interação. Além disso, aponta para a contribuição da voz na construção de uma nova maneira de trabalhar. Combinadas, a assistência de voz e a tecnologia de áudio preenchem uma variedade

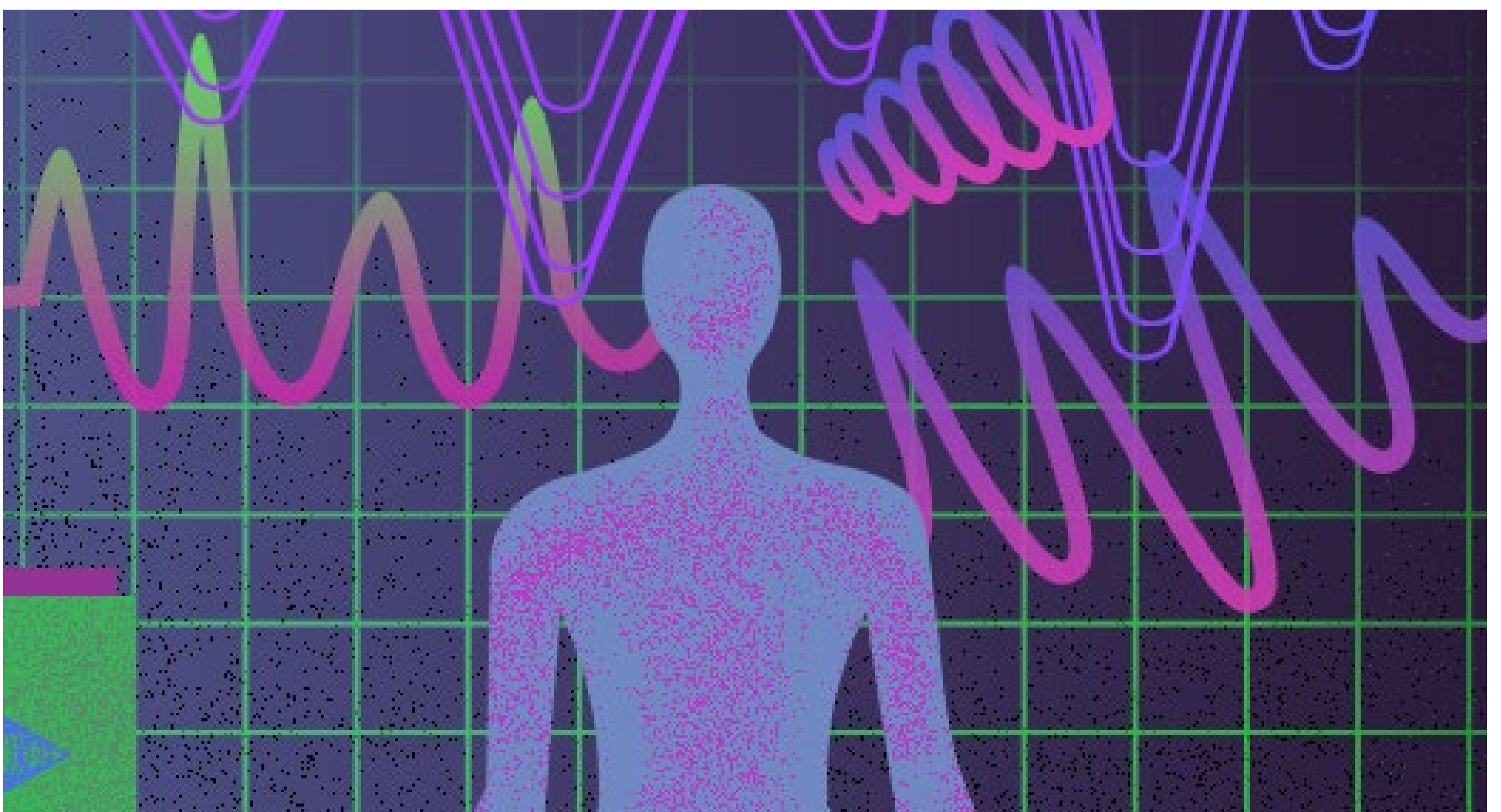
de funções importantes para a força de trabalho, incluindo o esforço para melhorar a comunicação e a produtividade.

A pesquisa também aponta como tendência a combinação das tecnologias de voz e a Internet das Coisas para o aprimoramento dos conceitos e aplicações da casa inteligente, bem como o uso da voz na criação e popularização de jogos e aplicativos de entretenimento.

Por fim, o estudo constata que há uma diversificação dos lugares em que a voz apareceu nos ecossistemas da marca. O estudo afirma que 52% das 50 maiores

empresas dos Estado Unidos adotaram o Alexa e 42% adotaram o Google Assistant criando aplicativos de voz em suas plataformas. Ainda assim, segundo o estudo, poucas foram as empresas que viram a voz como um ativo que merece ser tratado como um verdadeiro produto digital.

Ana Laura Gomes, Customer Success Manager da Adobe Inc., concorda que há uma demanda crescente para este tipo de aplicação. E dá como exemplo o aprimoramento do software Adobe XD, voltado para o desenvolvimento de interfaces. Diante da demanda dos desenvolvedores, a empresa tratou de incorporar as





tecnologias de áudio. “A gente não está fazendo de fato o produto final, mas as pessoas que estão criando coisas para lidar com estes produtos finais precisam que a ferramentas aceitem e trabalhem com áudio. Outra demanda que está aparecendo é o uso da IA para, da mesma forma que nossas ferramentas identificam similaridades entre imagens, algo semelhante aconteça nas ferramentas de áudio.”

Ela conta ainda que a Adobe, que há mais de dez anos desenvolve tecnologias de voz para seus aplicativos, trabalha continuamente para aprimorá-las, e que novidades serão apresentadas durante o próximo Adobe Max, conferência anual organizada pela empresa. A iniciativa reúne, durante três dias, especialistas e o público leigo para a apresentação dos produtos, novidades e tendências no setor. A edição deste ano, que será realizada entre os dias 26 e 28 de outubro, trará, segundo a CSM, uma novidade que dá um passo

além na aplicação das tecnologias de voz aplicadas à Inteligência Artificial.

Já uma outra pesquisa, conduzida pelo Instituto de Transformação Digital da Capgemini, uma multinacional francesa de consultoria, e divulgada em 2018, mostrava a maneira como os consumidores já estavam usando dispositivos de assistentes de voz e, ainda, as oportunidades que este comportamento tem trazido para as empresas se conectarem com seus clientes. O estudo, para o qual foram entrevistados mais de 5 mil consumidores nos Estados Unidos, Reino Unido, França e Alemanha, descobriu que os consumidores preferem cada vez mais interagir com *bots* do que com seres humanos, especialmente quando se trata de pesquisar produtos, aprender sobre novos serviços ou acompanhar consultas de atendimento ao cliente pós-compra. As pessoas estão desenvolvendo uma forte preferência pela interação com as empresas usando assistentes de voz. A

conclusão é que a tecnologia vai render benefícios concretos para empresas e marcas que forem capazes de garantir boas experiências, seja induzindo a um aumento de vendas, seja aumentando o valor da marca.

POTENCIAL DE SEGREGAÇÃO

Naquele momento, os entrevistados que utilizavam assistentes de voz viam a experiência como extremamente positiva, com 71% deles satisfeitos com a utilização da tecnologia. Outros aspectos favoráveis à tecnologia seriam a conveniência (48%), a capacidade de ter as mãos





livres para fazer outras coisas (48%) e a automação das tarefas de compras rotineiras (41%). No entanto, apontou também dois aspectos críticos, que têm potencial para limitar o acesso amplo à tecnologia. O aspecto cultural — com 81% dos entrevistados registrando uma expectativa de que a tecnologia compreenda mais claramente sua dicção e sotaque — e o aspecto financeiro, pois o relatório mostra que as assistentes são mais populares entre pessoas com idade entre 33 e 45 anos e que quase 1 em cada 5 (algo como 17% deles) dispõe de um rendimento familiar anual de mais de US\$ 100 mil.

O potencial de segregação das novas tecnologias de voz é um dos pontos de atenção com relação aos riscos e problemas inerentes a elas listados pela Dra. Adriana Braga, Professora Associada do Departamento de Comunicação da PUC-Rio, Pesquisadora do CNPq e Vice-Presidente da Media Ecology Association, entidade norte americana que se debruça

sobre a pesquisa, a crítica e a aplicação da ecologia das mídias em contextos educacionais, industriais, políticos, cívicos, sociais, culturais e artísticos. Ela chama a atenção para a necessidade de discussão a respeito dos aspectos negativos que sempre acompanham a introdução de novas tecnologias e que acabam relegados a um segundo plano. “Essas novas tecnologias digitais geralmente vêm acompanhadas de uma euforia que gira em torno das possibilidades econômicas relacionadas a elas. Mas, aliada à falta de regulamentação, essa euforia abre a guarda para riscos reais com relação à segurança dos usuários e mesmo às implicações éticas das aplicações. E quanto mais se desenvolvem, quanto mais são refinadas, há mais vulnerabilidade.”

Da avalanche de desinformação distribuída pela internet às possibilidades expandidas de crimes virtuais, ela lista exemplos em várias



aplicações das tecnologias digitais. E no campo das tecnologias de voz, que tem se alargado rapidamente, ela destaca, além da tendência à segregação evidenciada pelo grau de letramento e até pelo sotaque dos usuários, questões relacionadas à segurança e à privacidade, com aplicações que facilitam a fraude e com o risco iminente de captação de informações pelos microfones que se multiplicam exponencialmente; e até questões de gênero na definição da identidade das assistentes pessoais.

“Essa é uma questão que para mim é emblemática, primordial, que é o fato de todas as assistentes virtuais serem mulheres, o que acaba reproduzindo um estereótipo feminino”.

Ela cita a campanha criada pelo Bradesco em resposta às mensagens ofensivas enviadas à sua assistente, a Bia, baseada no movimento

“Hey, update my Voice”. capitaneado pela Unesco, para que a IA que rege a personagem reaja de forma firme contra o assédio, sem submissão. “Logo que a Bia foi lançada, o Bradesco identificou milhares de mensagens ofensivas direcionadas à assistente. E inspirados na iniciativa da Unesco, fizeram uma grande campanha para combater o assédio a partir da mudança do tom de voz. Mas, a questão que se coloca, tanto aqui quanto lá, é que há uma questão de gênero muito profunda, que tem origem nos valores de quem está por trás da tecnologia. Há uma naturalização do papel feminino como aquela que ajuda, a que organiza.”

Outros dois pontos sublinhados diante do refinamento das tecnologias de voz associadas à Inteligência Artificial é o potencial para o crime e para a desinformação, bem como os riscos éticos associados ao uso das tecnologias na antropomorfização das máquinas. “O risco aqui é colocar mais confusão onde



já está confuso. Quando se fala em pós-verdade e em *fake news*, isso pode potencializar o problema. Bem como a questão do transhumanismo, que é a ideia de ressuscitar os mortos por meio da máquina. Esse é um anseio antigo, o de produzir esse simulacro, mas que na verdade abastarda a capacidade humana porque um sistema operacional pode simular, mas nunca vai chegar próximo das características humanas. E isso vai sendo feito de modo sutil e muito perigoso. O pensamento que sai daí é: “se a máquina é igual a mim, a inteligência da máquina é igual a minha”. Mas, não é.

A pesquisadora ressalta, no entanto, que suas considerações não representam um fechamento ao uso dessas novas tecnologias. Ao contrário, mostram que é fundamental uma discussão permanente que jogue luz sobre todos os aspectos relacionados à adoção delas no dia a dia. Em vez de uma postura condenatória, ela defende uma postura mais vigilante

associada a um esforço de regulação e de transparência. “Neil Postman e Marshall McLuhan (dois acadêmicos fundadores dos estudos de Ecologia das Mídias) defendem uma agenda de educação para as mídias desde o ensino médio, como forma de preparar as pessoas para lidar com estas questões. O que temos hoje é um descompasso muito grande entre o que se discute na academia e a realidade da maioria das pessoas. É preciso achar modos mais civilizados de lidar com a questão. Do jeito que está, não é uma coisa que satisfaça a sociedade. Claro que você quer o seu smartphone, você precisa da internet. Mas vejo que há uma certa impotência diante de determinados aspectos relacionados a essas novas tecnologias. É preciso pensar a partir da nova realidade digital da comunicação.” ■



A SIMPLICIDADE DA VOZ

**T R :
Q + A**

**Talita Bruzzi
Taliberti**

Por **Carlos Aros**

A assistente virtual Alexa está completando dois anos no Brasil, e, para o futuro, a Amazon quer romper barreiras e tornar as tecnologias acessíveis para cada vez mais pessoas.

Foto por **Mariana Pekin**

Há alguns anos talvez fosse estranho conversar com uma caixa de som e fazer pedidos para que o aparelho traga respostas ou execute tarefas envolvendo outros dispositivos da casa, como acender uma lâmpada, por exemplo. Mas, de dois anos para cá, não é estranho para o brasileiro dizer “Alexa, bom dia”, e esperar que uma voz suave, vinda da caixinha de som, lhe informe a previsão do tempo, os próximos compromissos da agenda e ainda, de quebra, apresente um fato curioso relacionado ao dia que está começando.



É com o objetivo de tornar essa e outras interações com a assistente virtual Alexa cada vez mais rotineiras e naturais que **Talita Bruzzi Taliberti** lidera a operação brasileira de Alexa na Amazon. A executiva é entusiasta das tecnologias de voz e defensora do aprimoramento constante das aplicações e das experiências baseadas nessa interação, que tem potencial único de aproximar das tecnologias pessoas que, por outros meios, talvez não fossem digitalizadas.

“Por ser natural, [a voz] é muito fácil de ser adotada. Então, você tira a barreira de quem tem dificuldade de aprendizado com uma nova tecnologia. Seja para pessoas com deficiência ou com idade muito avançada e, no extremo, podemos falar sobre pessoas iletradas”, argumenta Talita sobre o potencial inclusivo das tecnologias de voz.

Nesta entrevista exclusiva para a **MIT Technology Review Brasil**, a Country Manager de Alexa Brasil na Amazon, Talita Bruzzi Taliberti, faz um balanço sobre os dois anos de operação no país, explica os desafios e a estratégia para o desenvolvimento de aplicações baseadas em voz e indica os próximos passos para a Alexa no Brasil.

Confira abaixo os principais trechos da entrevista, disponível na íntegra em nosso canal no YouTube (acesse pelo QR Code da página 45).

MIT Technology Review: Como você avalia essa jornada das tecnologias de voz, dos assistentes virtuais, em especial durante esses dois anos de Alexa no Brasil?

Talita Bruzzi Taliberti: No início era tudo uma grande incógnita, uma tecnologia muito nova, disruptiva. Mas eu posso dizer que, nesses dois anos que estamos completando no Brasil, foi uma surpresa pra gente e também para os brasileiros o quanto essa tecnologia é de

fácil uso, de muita utilidade e de muita diversão.

A gente vem passando, infelizmente, por um momento difícil com essa pandemia, e as pessoas acabam ficando muito mais tempo dentro de casa. Ter uma tecnologia como a Alexa no centro da nossa casa trazendo facilidades e até diversão foi uma grande companhia. E nesse período, quase dois anos de Brasil, a Alexa evoluiu muito, se torna

cada vez mais inteligente, entende bem a linguagem e a cultura brasileira. A Alexa foi se adaptando, trazendo mais informações e mais conhecimentos.

As *skills*, que são as aplicações de voz para Alexa, são um dos exemplos de inovação. Quando a gente lançou, eram 300 *skills*, hoje já são mais de mil e quinhentas. Os dispositivos também mostram uma clara evolução, já lançamos mais de 15 dispositivos desde que apresentamos o primeiro Echo aqui no Brasil.

Outro segmento que trouxe oportunidades e que a gente viu evolução é o de casas inteligentes. Hoje já são mais de 500 dispositivos que se conectam com Alexa facilitando e tornando mais democrática a automação residencial, algo antes difícil de ser implementado.

TR: Como é o trabalho para fortalecer o ecossistema de Alexa com as *skills* e o envolvimento das marcas para criação de aplicações e soluções para os usuários?

Talita: Do ponto de vista de *skills*, a gente tem um

trabalho de preparar mais o segmento, principalmente na linha de desenvolvimento de aplicações de voz. A gente disponibiliza uma série de materiais, o Alexa Skill Kit, com a nossa tecnologia de integração e APIs para os desenvolvedores, uma série de treinamentos, realizamos eventos de tempos em tempos, exatamente para fornecer uma série de materiais para que desenvolvedores se capacitem no desenvolvimento de aplicações de voz. Esses desenvolvedores podem pensar em *skills* para si próprios e para outras empresas que estão percebendo essa tendência de crescimento do uso de voz na interação com o usuário para trazer novos serviços, novas formas de transacionar e se relacionar com o cliente.

Além disso, temos algumas atividades voltadas para esses desenvolvedores em capacitação. Uma delas é o Alexa Live, evento global no qual compartilhamos casos de uso e novas funcionalidades. No ano passado fizemos o Prêmio Alexa de Acessibilidade e estimulamos que os desenvolvedores



produzissem novas aplicações voltadas para pessoas com deficiência para facilitar que essas pessoas tenham mais independência e mais autonomia. Foram mais de 90 *skills* desenvolvidas.

A gente está trazendo novas oportunidades de negócios para quem está embarcando nesse movimento de aplicações de voz. Esse ano, por exemplo, tivemos um brasileiro como primeiro sul-americano reconhecido em um programa global da Amazon, o Alexa Champion, que reconhece os principais desenvolvedores de tecnologia de voz do mundo.

Ou seja, são oportunidades para os desenvolvedores e também para as marcas que querem entrar nesse segmento de voz para ter mais um ponto de contato, uma forma de se relacionar com o cliente.

TR: Nesse contexto de desenvolvimento de aplicações e soluções, como atender demandas que surgem pelos usuários?

Talita: Quando a gente fala sobre experiência do usuário,

[a voz] trouxe uma nova perspectiva. A gente precisa pensar no desenvolvimento dessas aplicações e experiências com outro viés. Estamos muito acostumados a pensar no *Graphic User Interface*, mas quando a gente vai para o ambiente de voz, é preciso pensar como se não enxergássemos, precisamos pensar na experiência com voz, e aí surge o *Voice User Interface*. Então, por exemplo, enquanto no GUI é muito comum pensar num menu *drop-down* com uma série de opções, se a gente leva isso para voz fica extremamente maçante. É um outro nível de *mindset*. Claro que, mesmo com a voz, é possível ter uma experiência complementar com uma tela. Mas, o principal é imaginar que aquela experiência não teria um recurso visual.

TR: A experiência de Alexa mostra que é possível explorar o desenvolvimento de soluções baseadas em voz como um caminho viável para fazer com que as tecnologias cheguem a mais pessoas?

Talita: Esse é um assunto extremamente importante. A voz — ou a fala — é muito

natural, seja porque é uma das primeiras linguagens que a gente aprende, uma criança que tá começando a aprender alguns símbolos muitas vezes aprende a falar antes. E por ser natural, é muito fácil de ser adotada. Então, você tira a barreira de quem tem dificuldade de aprendizado com uma nova tecnologia. Seja para pessoas com deficiência ou com idade muito avançada e, no extremo, podemos falar sobre pessoas iletradas. A gente vê o WhatsApp, por exemplo, que é cada vez mais usado, e as pessoas mandando cada vez mais áudio, e isso facilita muito. Então, imagine poder usar a voz para comandos dentro de casa, que você precisa fazer de maneira rotineira.

Outro fator que a tecnologia de voz traz e que é muito interessante é que a voz, por ser natural, requer uma carga cognitiva muito mais baixa do que outras atividades, e permite que seja muito mais fácil executar ações concomitantemente. Enquanto estou cozinhando, posso incluir um produto na lista de compras com um



simples “Alexa, adicione leite à minha lista de compras”, por exemplo.

Por causa dessa simplicidade que as tecnologias de voz têm sido adotadas com muita velocidade.

TR: E o mercado brasileiro? Nesses dois anos, é possível dizer que a Amazon atingiu as expectativas com a Alexa falando português?

Talita: A gente sabia que as pessoas iam gostar e que daria certo, por isso que a gente se dedicou para desenvolver a tecnologia da inteligência artificial da Alexa, um trabalho monstruoso, para desenvolver as tecnologias



de reconhecimento da fala, gerar respostas e entender o que o cliente está pedindo. Mas, foi uma grande surpresa o nível de adoção e como as pessoas gostam e interagem com a Alexa.

O Brasil, hoje, é um dos maiores do mundo onde as pessoas, por exemplo, dão bom dia para a Alexa, e a gente rapidamente percebeu essa peculiaridade na cultura brasileira, que é de ser próximo e amigável. Então, as pessoas gostam não só de dizer bom dia, mas de pedir por favor e de falar obrigado para a Alexa. E a partir do momento que agente começou a perceber isso, a gente viu a importância de trazer um conteúdo relevante naquele bom dia. As pessoas gostam porque voltam no dia seguinte para saber o conteúdo que virá com o bom dia da Alexa.

E outra questão são as diferenças de sotaque, que trouxe a necessidade para a Alexa de aprender muito rápido as muitas formas de falar no Brasil. Ela não muda o sotaque dela, mas consegue reconhecer os diferentes tipos de entonação e jeitos de falar. A Alexa aprendeu até

mesmo o jeito do brasileiro falar inglês. Foi importante também que a Alexa se mostrasse verdadeiramente brasileira, as características que ela tem na personalidade e que se conectam com as pessoas.

TR: O isolamento social forçado pela pandemia estimulou ou acelerou a interação e a adoção da assistente virtual e dos dispositivos inteligentes pelos brasileiros?

Talita: A Alexa foi lançada no Brasil na virada de outubro para novembro de 2019, então a gente tinha poucos meses de Brasil quando começou a pandemia que colocou todo mundo dentro de casa. Mas acredito que, sim, as pessoas passaram a ficar muito tempo em casa e com isso tiveram mais oportunidades de conexão e interação com os dispositivos. Além disso, as pessoas viram muito valor ao trazer mais conveniência e praticidade para o dia a dia dentro de casa, na hora de gerenciar as tarefas. Tudo o que foi possível utilizar como aliado para facilitar a vida foi adotado. Isso vale muito para o segmento da

casa conectada, por exemplo.

TR: Quais desafios precisam ser superados para ampliar a adoção de tecnologias de voz como a Alexa? O que precisa evoluir?

Talita: Mais do que um desafio, é como agente se coloca e o que a gente acredita que precisa melhorar. Muito se fala que a Amazon acredita sempre que todo dia é o primeiro dia, esse é o nosso *mindset*, de que estamos sempre começando e precisamos continuamente evoluir e oferecer uma experiência cada vez melhor para o cliente.

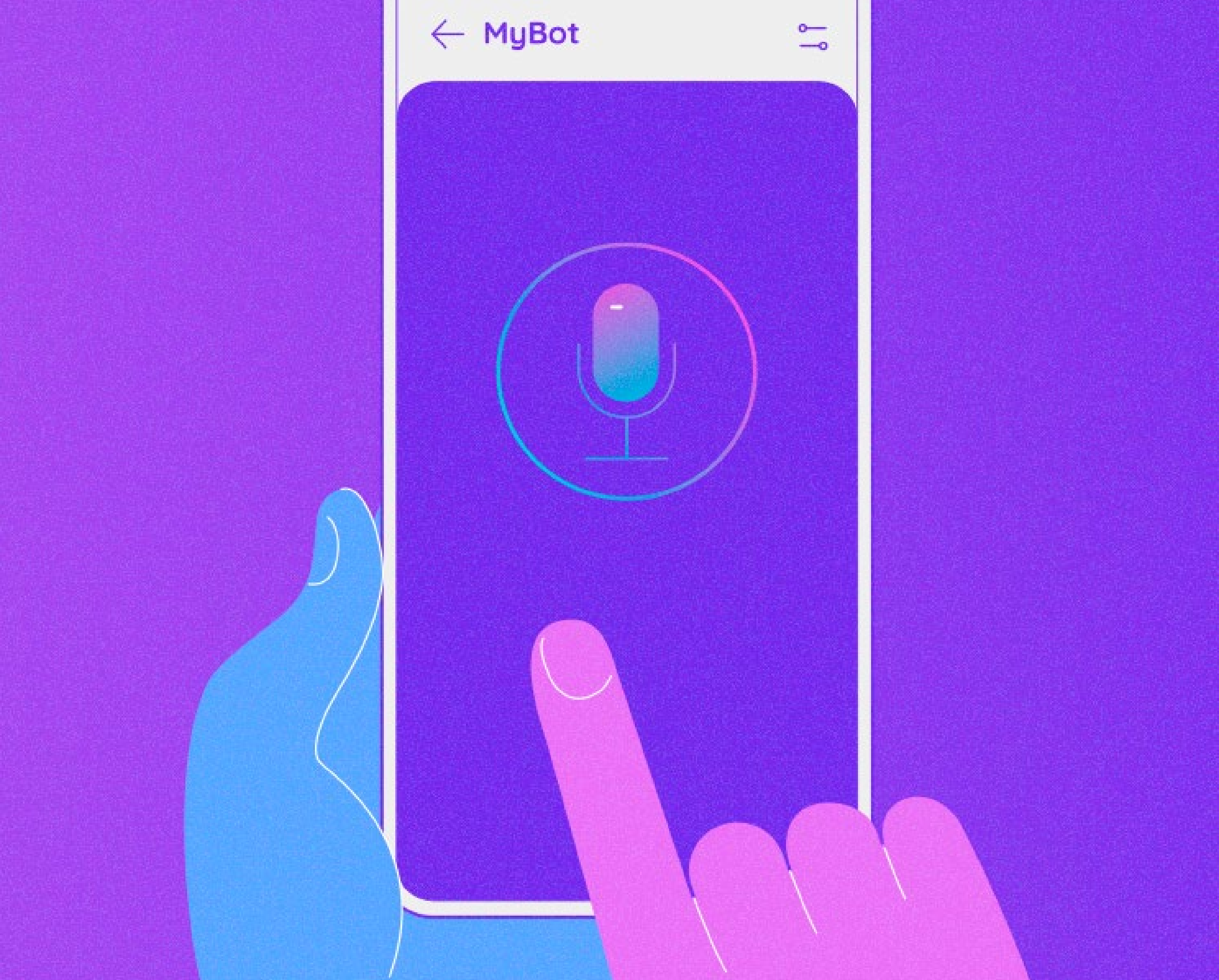
Nesse sentido, existem alguns pilares que precisamos evoluir, como, por exemplo, aumentar a capacidade de aprendizado da Alexa e fazer com que ela seja cada vez mais inteligente. Precisamos que ela cada vez mais entenda o contexto para que traga respostas aos comandos ou pedidos de forma mais acertada. Outro ponto é como trazer cada vez mais informações, que a base de dados seja cada vez mais capaz de responder todas as perguntas do cliente.

Outro aspecto é a naturalidade da conversação. A gente tem um desafio para que essa interação com a Alexa seja cada vez mais natural, mais humana e mais simples. O objetivo não é fazer as pessoas esquecerem que é um robô, mas tornar essa interação ainda mais fluída faz com que seja ainda mais fácil de [o recurso] ser usado.

E, do outro lado, trazer cada vez mais dispositivos, seja como Amazon ou por meio de parcerias com empresas. Quanto mais opções a gente oferece para o cliente, mais a gente tem a opção de ele encontrar o que quer. Essa é a nossa bandeira.

A íntegra da entrevista pode ser acessada no canal da MIT Technology Review Brasil no YouTube. Aponte a câmera ou clique no código para acessar. 





Voz: humana vs sintética. Qual a melhor aplicação de cada uma?

Por **Simone Kliass**
e **Jason Bermingham**

Você é impactado por gravações de voz do momento em que acorda até a hora em que vai dormir. Talvez você ouça um podcast no Spotify ao preparar o café da manhã. Pouco tempo depois, no caminho para o trabalho, um comercial toque no rádio enquanto o Waze pede para

you pegar a próxima direita. De repente o trânsito para. Atrasado, você pede para a Siri fazer uma chamada — só que a ligação cai direto na URA da empresa. De quem são todas essas vozes? São de profissionais que ganham a vida em frente a um microfone e que fazem parte da sua rotina diária, muitas vezes sem você perceber.

A fala sempre foi a ferramenta mais valiosa do ser humano. É o que nos separa dos outros seres do nosso planeta. Desde os primórdios da evolução do Homo Sapiens, a nossa



prevalência como espécie foi ligada a comunicação verbal. Até hoje, apesar de todos os avanços tecnológicos alcançados pela sociedade, a voz continua sendo uma peça central da nossa identidade individual e coletiva. Por meio dela contamos as nossas histórias, compartilhamos o nosso conhecimento e impomos as nossas convicções.

Mas, mesmo que não tenha substituído a nossa principal ferramenta de comunicação, a tecnologia também evoluiu, e hoje somos capazes de reproduzir a fala humana com alta precisão e em grande escala, possibilitando que muitas gravações antes feitas por profissionais de voz possam ser programadas utilizando dados digitais. Uma “voz sintética” se refere a conteúdo sonoro criado digitalmente para replicar a voz humana e transmitir informações a ouvintes humanos. Por meio de conteúdo pré-gravado e posteriormente modificado

por meio de algoritmos, hoje podemos recriar a voz humana digitalmente com fidelidade surpreendente.

Entre nós, profissionais de voz, o debate sobre as vantagens e desvantagens das vozes sintéticas tem sido acalorado. Por um lado, elas representam uma nova fronteira e, portanto, prometem novas oportunidades. Afinal, alguém precisa gravar essas vozes, certo? Por trás de cada voz virtual há um profissional que foi pago pelo seu tempo no estúdio e, idealmente, também pelo direito de uso da sua voz. Por outro lado, por mais realistas que as vozes sintéticas se tornam, mais trabalhos podemos perder para elas.

Uma coisa é certa: as vozes sintéticas vieram para ficar. Mas, a questão que não quer calar é: qual será a relação entre as vozes gravadas por profissionais humanos e as vozes geradas por computadores nos próximos



anos? Em busca de respostas objetivas, procuramos a Rupal Patel, CEO da VocaliD, uma empresa que cria vozes sintéticas através de AI, *machine learning* e *crowdsourcing*. Para Rupal, uma voz sintética pode ser útil para textos muito grandes, que mudam constantemente e que precisam ser gravados em prazos curtíssimos. Chatbots, notícias em tempo real e anúncios personalizados são alguns dos formatos que frequentemente utilizam vozes sintéticas.

“Para nós, o importante é oferecer consistência e diferenciação às empresas que já utilizam serviços genéricos de text-to-speech, como Amazon Polly e Google TTS,” explica Rupal.

Para criar o banco de vozes da VocaliD, Rupal trabalha com diversos profissionais

da nossa área para gerar uma espécie de dublê digital. Por sua vez, essas vozes são licenciadas para empresas que precisam de uma solução rápida e consistente e que seja disponível *on-demand*. Além disso, cada contratação é negociada com o talento. Apenas usuários registrados têm acesso à plataforma para gravar, editar e produzir o áudio de roteiros. Ou seja, o modelo de negócios da VocaliD é trabalhar junto com os profissionais de voz para que possamos monetizar em cima do licenciamento do nosso trabalho.

Esta ressalva é importante, porque o uso da voz sintética levanta um questionamento: quem é o dono da gravação da voz? Ela deixa de pertencer ao talento? Já vimos casos em que o profissional perdeu o rastreamento do uso da sua voz, como o da locutora Bev Standing, que se surpreendeu ao se ouvir como a voz de vídeos do TikTok. O aplicativo, por seu lado, obteve a voz sintética

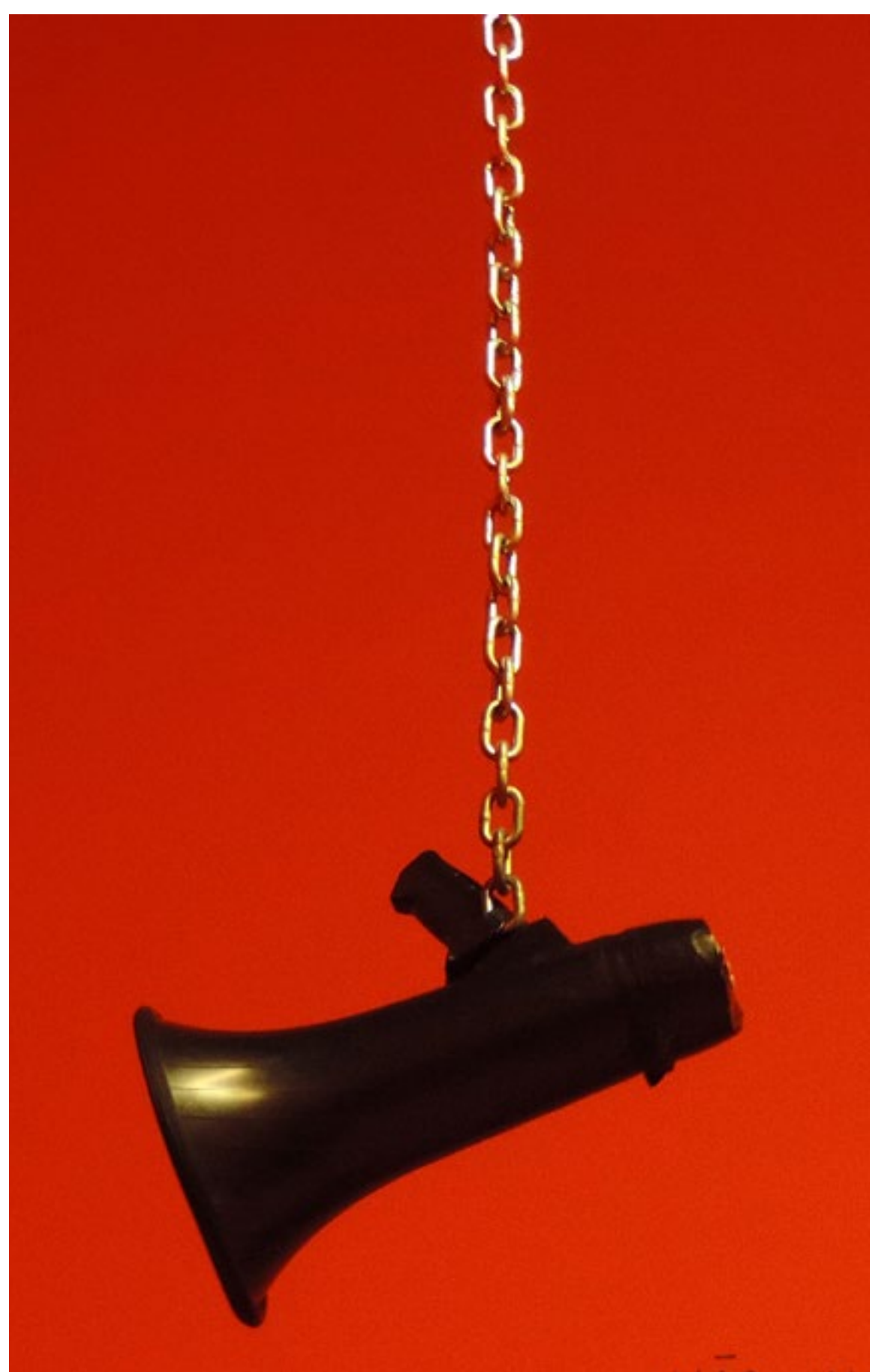


por meio de uma empresa chinesa para a qual a Bevtinha gravado anteriormente.

Mais do que nunca, a regulamentação desse mercado precisa caminhar de forma paralela ao desenvolvimento da tecnologia. Devemos debater questões de transparência, de diversidade de gênero, geográfica, de raça, diversidade cultural, passando pelos direitos de uso da voz de alguém, da atenção a possíveis fraudes com emulação de vozes, até o uso sem aviso de uma voz recriada, como aconteceu no documentário “Roadrunner” sobre o chef Anthony Bourdain.

A empresa Soundthinkers participou recentemente no Brasil de uma ação que recriou a voz do pai do Zico com muito sucesso e transparência. O senhor José Antunes Coimbra nunca tinha visto o filho fazer um gol no Maracanã. Com sua voz reconstituída, ele pede esse gol em uma surpresa

para o jogador. Para Paulo Dytz, CEO da Soundthinkers, a questão ética no uso de assistentes de voz é densa. De acordo com ele, “podemos estar tratando de uma voz sintetizada de maneira 100% artificial ou com a voz de uma pessoa, um profissional que ‘emprestou’ sua própria voz para o assistente. Além disso, há uma questão comportamental vinculada à conversação com este tipo de tecnologia. Podemos desenvolver hábitos prejudiciais humanizando demais ou de menos. Como diversos outros aspectos da





vida, a chave é o equilíbrio”. Essa discussão não seria completa sem incluir também o olhar de alguém que entende de tecnologia, mas também ganha a vida com a voz. Anne Ganguzza é locutora, produtora e apresentadora do podcast “Voice and AI: Future of VO”. Durante a nossa conversa, ela usou o exemplo de *e-learning* para mostrar porque uma voz sintética não pode competir com uma narração humana. “A minha experiência de mais de 20 anos como professora em sala de aula me serviu bem, e reforçou o meu compromisso para me envolver com o ouvinte quando gravo um roteiro. O meu objetivo é ser a professora favorita do estudante. Quero deixar claro que estou apaixonada tanto pelo assunto, como pela jornada educacional do aluno. Para isso, utilizo técnicas de atuação que focam no compromisso individualizado com o estudante em um tom de conversação que

destaca de forma audível os momentos de ensino, bem como expressam o encorajamento e a intenção para o público específico.” Ou seja, como Anne deixa claro, a nuance emocional de um profissional de voz que tem um compromisso com o roteiro que grava não é algo que se recrie com uma voz sintética.

Em conclusão, acreditamos que as vozes sintéticas têm o seu lugar, sim. Empresas como a Obsidian estão usando essas vozes durante a concepção de vídeo games com dezenas de personagens e milhares de linhas de diálogo. Porém, antes de lançar a versão final, vozes humanas são utilizadas para substituir as vozes sintéticas. Também há o caso da Amazon, que tem apenas uma pequena porcentagem dos seus livros listados disponíveis no Audible. Como o mercado de audiolivros é bem aquecido nos EUA, algumas empresas estão usando vozes



sintéticas para a narração de títulos menos conhecidos e contratando locutores para aqueles com mais audiência.

Como profissionais de voz, vemos diariamente como um narrador humano consegue inspirar, motivar e criar laços com os seus ouvintes. Dentro do estúdio, usamos variações de tom, ritmo, projeção, entonação, cacos, respiros, pausas, gestos faciais e corporais, diferentes posturas corporais, entre outras técnicas, para criar uma conexão verdadeira com os nossos ouvintes. As vozes sintéticas mais elaboradas tentam “imitar” algumas dessas facetas, sempre dentro de personas digitais com grandes limitações. Às vezes são chamadas de “*AI voice actors*”. Mas, não são atores, não. Com uma voz virtual, você pode ser capaz de gravar as informações de forma mais rápida e talvez até mesmo mais barata (mas nem sempre). Porém, a que custo? Afinal, se o narrador parece não se importar com a

mensagem, por que o ouvinte deveria? A melhor resposta talvez esteja no meio termo, levando-se em consideração a praticidade, honrando sempre a transparência e a ética, estando totalmente consciente sobre o que você ganha e o que você perde quando escolhe uma voz para o seu projeto. Só não se esqueça que quando você precisa criar uma conexão verdadeira, não existe comparação entre a voz humana e a sintética. ■



As mulheres continuam sendo apagadas da indústria da tecnologia

A decisão de TikTok de usar a voz de uma mulher sem sua permissão é apenas um exemplo recente de um problema que alguns pensam erroneamente que já está superado.

Ilustração por **MSTECH**

No meio da noite de 24 de maio, o TikTok mudou sua voz. A voz onipresente feminina que podia ler o texto de qualquer vídeo em voz alta em uma cadência robótica e levemente afetada foi subitamente substituída por uma com um tom quase sorridente e otimista. Muitos usuários começaram a chamar a nova voz de “Garota do Vale da Estranheza” para expressar seu descontentamento. O cantor Lil Nas X até fez um TikTok sobre isso.

Mas, o que aconteceu com a velha voz? E quem era a



mulher por trás dela?

Quando pensamos no papel das mulheres na computação, geralmente pensamos em como, literal e figurativamente, elas foram silenciadas com mais frequência do que ouvidas.

As vozes e os corpos das mulheres podem ser encontrados em toda a história da computação — desde por meio de áudios em contagens regressivas de lançamento até serem visíveis em fotos — mas apenas recentemente os historiadores incluíram essas mulheres de volta na narrativa, explicando o que elas fizeram. Por muito tempo, as mulheres foram erroneamente consideradas periféricas para a história da computação, embora muitas vezes fossem elas as programadoras dos computadores.

E ainda há os casos em que,

quando ouvimos a voz de uma mulher como parte de um produto de tecnologia, podemos não saber quem ela é, se ela é mesmo real e, se for, se ela consentiu que sua voz fosse usada dessa forma. Muitos usuários do TikTok presumiram que a voz de conversão de texto em fala que ouviram no aplicativo não era uma pessoa real. Mas era: pertencia a uma dubladora canadense chamada Bev Standing, e Standing nunca deu permissão de uso à ByteDance, a empresa dona da TikTok.

Standing processou a empresa em maio deste ano, alegando que as maneiras como a voz dela estava sendo usada — principalmente a forma como os usuários podiam fazer com que ela falasse qualquer coisa, incluindo palavrões — estavam prejudicando sua marca e suas formas de sustento. Ao tornar-se conhecida como “aquela



voz no TikTok “, a qual você poderia usar para dizer o que quisesse, Standing teve reconhecimento sem remuneração e, ela alegou, prejudicou suas possibilidades de conseguir trabalho.

E, quando TikTok removeu abruptamente sua voz da plataforma, Standing descobriu da mesma maneira que o resto de nós: ouvindo a modificação e vendo notícias sobre ela. (TikTok não comentou com a imprensa sobre a mudança de voz.)

Aqueles familiarizados com a história da assistente de voz Siri, da Apple, podem estar experienciando um pouco de déjà vu: Susan Bennett, a mulher que dublou a Siri original, também não sabia que sua voz estava sendo usada para aquele produto até que ele foi lançado. Bennett acabou sendo substituída de seu papel como “voz feminina para o inglês americano”, e a Apple nunca a reconheceu

publicamente. Desde então, a Apple escreveu cláusulas de sigilo nos contratos dos dubladores e, mais recentemente, afirmou que sua nova voz é “inteiramente gerada por software”, eliminando a necessidade de dar crédito a qualquer pessoa.

Esses incidentes refletem um padrão problemático e comum na indústria de tecnologia. As formas como os feitos das pessoas são avaliadas, reconhecidas e pagas geralmente refletem sua posição na sociedade em geral, não suas contribuições reais. Uma das razões pelas quais os nomes de Bev Standing e Susan Bennett são amplamente conhecidos on-line é que eles são exemplos extremos de como o trabalho das mulheres é apagado mesmo quando está lá para que todos possam ver ou ouvir.

As formas como os feitos das pessoas são avaliados,



reconhecidos e pagos geralmente refletem sua posição na sociedade em geral, não suas contribuições reais.

Quando as mulheres do setor de tecnologia se manifestam, costumam ser orientadas a ficarem quietas — principalmente se forem negras. Timnit Gebru, que possui um doutorado em Ciência da Computação em Stanford, foi expulsa do Google, onde coliderou uma equipe de ética em Inteligência Artificial (IA), depois de falar sobre suas preocupações em relação aos grandes modelos de linguagem da empresa. Sua colíder, Margaret Mitchell (Doutora pela University of Aberdeen com foco na geração de linguagem natural), também foi demitida de seu cargo após falar sobre a demissão de Gebru. Em outras partes da indústria, denunciante como Sophie Zhang, no Facebook;

Susan Fowler, na Uber; e muitas outras mulheres se viram silenciadas e, muitas vezes, demitidas como resultado direto ou indireto de tentarem fazer seus trabalhos e mitigar os danos que viram nas empresas de tecnologia onde elas trabalharam.

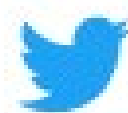
Mesmo as mulheres que fundaram startups podem ser apagadas em tempo real, e o problema é pior para as mulheres negras. Rumman Chowdhury, Doutora pela University of California, San Diego, fundadora e ex-CEO da Parity, uma empresa focada em ética em IA, viu seu papel na história de sua própria empresa ser minimizado pelo New York Times.

Em uma reportagem sobre a Parity, o jornal falhou em identificar Chowdhury como a CEO fundadora e, em vez disso, a descreveu apenas como “uma pesquisadora que fez uma ferramenta”



Rumman Chowdhury

@ruchowdh



Friends, I am tired. I work hard to build good things and bring in the right people. Parity is no different.

For the second time in two weeks, I have to fight a major media outlet for basic recognition of my work. The gaslighting is real. The erasure is real. 1/



Liz O'Sullivan @lizjosullivan

To set the record straight: Parity 2.0 wouldn't exist without @ruchowdh. In between Accenture and Twitter, Rumman worked tirelessly on Parity for a year. She got customers, proved the model, validated the need, and invented a new way to identify bias in AI beyond "just data". 1/2 twitter.com/lizjosullivan/...

4:56 PM · Jun 30, 2021



1.4K



34



Share this Tweet

[Tweet your reply](#)

na qual o negócio da Parity se baseia. Depois de uma reação pública significativa, o Times silenciosamente atualizou a história, sem emitir uma correção formal. Mas, ainda assim, não reconheceu Chowdhury como CEO fundadora da Parity, concentrando-se na jovem mulher branca que é sua sucessora.

Recentemente, milhares de criadores de conteúdo negros no TikTok, muitos

[Amigos, estou cansada. Trabalho muito para construir coisas boas e trabalhar com as pessoas certas. A Parity não é diferente. Pela segunda vez em duas semanas, tenho que lutar contra um grande veículo de comunicação pelo reconhecimento básico do meu trabalho. O gaslighting é real. O apagamento é real. 1 /].

[Para esclarecer as coisas: Parity 2.0 não existiria sem a @ruchowdh. Entre a Accenture e o Twitter, Rumman trabalhou incansavelmente na Parity por um ano. Ela conseguiu clientes, testou o modelo, validou necessidades e criou uma nova maneira de identificar preconceitos em IA além de “apenas dados”. ½].

deles mulheres, entraram em “greve”, recusando-se a coreografar novas danças para o *single* mais recente da cantora Megan Thee Stallion. As mulheres negras, em



particular, têm visto suas coreografias repetidamente copiadas e roubadas por criadores do TikTok que são mulheres brancas e que monetizam essas danças, e até mesmo passam a apresentá-las na televisão nacional, sem dar crédito aos criadores originais.

Quando olhamos para o impacto das vozes das mulheres na tecnologia hoje, podemos ver que elas lideraram exigências de responsabilização e também que foram literal e figurativamente subestimadas. As mulheres muitas vezes estão presentes na tecnologia sem serem ouvidas, desde seu trabalho de locução, que se torna a base das ferramentas de voz utilizadas por milhões de pessoas, sem serem remuneradas ou devidamente reconhecidas, até seu trabalho nos conceitos fundamentais da IA.

Embora as mulheres, e particularmente as mulheres negras, sejam frequentemente as primeiras pessoas a quem as empresas de tecnologia recorrem quando precisam mostrar sua diversidade ou se defender das críticas de que seus produtos agravam o sexismo e o racismo, essas mulheres lutam para que seus conhecimentos sejam levados ao máximo nível de gestão, e raramente estão em posição de definir a agenda para o desenvolvimento tecnológico.

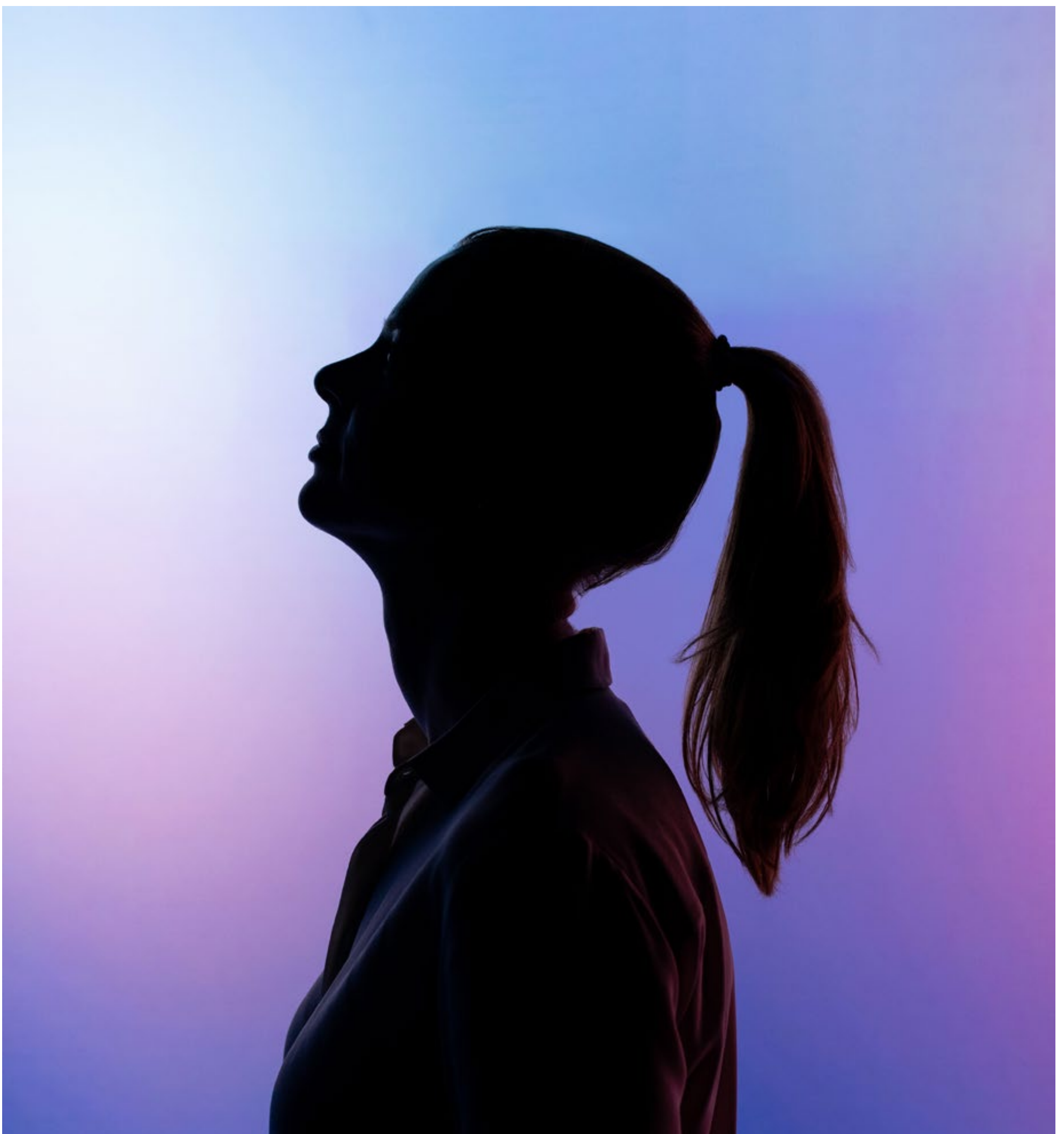
A boa notícia é que historiadores e jornalistas, assim como as próprias mulheres, têm trabalhado muito para reverter esse apagamento e estão tendo um sucesso significativo. Na última década, novos livros, artigos e filmes esclareceram o assunto e mudaram nossa compreensão sobre a importância das contribuições das mulheres



para a alta tecnologia. A má notícia é que essas contribuições ainda estão sendo apagadas em tempo real, incluindo o trabalho de mulheres que estão tentando resolver alguns dos problemas mais importantes da tecnologia hoje. Enquanto isso ainda ocorrer, não

importa a rapidez com que tentemos corrigir esses apagamentos, acabaremos no mesmo lugar. 

Mar Hicks é Professora Associada do Illinois Institute of Technology e autora do livro *Programmed Inequality* (MIT Press, 2017).



A MIT Technology Review Brasil, além da revista, produz conteúdo especializado diariamente.

Tome decisões mais embasadas
e mais conscientes sobre a tecnologia
por meio de jornalismo de autoridade,
influyente e confiável.



/MITTechReviewBR

Acesse nosso site:

mittechreview.com.br



TRANSFORMAÇÃO DIGITAL PARA SUA EMPRESA

[Digital]

[Aplicações]

[Infraestrutura]



**1º Lugar - Lista as
integradoras de TI mais
eficientes do Brasil.**

Informática Hoje



**PME's que mais crescem
no Brasil.**

Deloitte e Exame



Conheça mais sobre
nossos serviços.

Founding sponsor

MIT Technology Review Brasil

infobase.com.br

infobase 