

VOZES SINTÉTICAS: EXPLORANDO OS AVATARES DA IA NA COMUNICAÇÃO

SYNTHETIC VOICES: EXPLORING AI AVATARS IN COMMUNICATION

CELESTINO JOANGUETE¹

RESUMO

O artigo apresenta uma pesquisa sobre o fenômeno das vozes sintéticas, desenvolvidas pela Inteligência Artificial (IA), e explora como essas vozes estão transformando diversos setores da comunicação, desde assistentes virtuais e redes sociais até educação, marketing, documentários e publicidade. O estudo reflete sobre os avanços que aproximam as vozes sintéticas cada vez mais da fala humana, proporcionando interações mais naturais e personalizadas. No entanto, a ascensão dessas tecnologias suscita questões sobre autenticidade e o potencial impacto na substituição de vozes humanas em determinadas profissões. Este estudo visa analisar como as vozes sintéticas influenciam a percepção humana e identificar os desafios e oportunidades associados ao seu uso em diversos contextos, inclusive da educação. Para operacionalizar a pesquisa, foi adotado um modelo metodológico fundamentado na análise qualitativa e de conteúdo das ferramentas de IA e os seus avatares voltados para o campo da comunicação. Assim, o objetivo é fornecer uma compreensão abrangente das dinâmicas na interação entre humanos e vozes sintéticas, contribuindo para o desenvolvimento de diretrizes éticas e práticas recomendadas para a implementação dessa tecnologia. Ficou evidente no estudo que as vozes sintéticas da IA estão se aproximando cada vez mais da naturalidade da fala humana. Essa aproximação não só melhora a funcionalidade de diversas aplicações tecnológicas, como também levanta questões do potencial impacto na substituição de vozes humanas em diferentes campos profissionais.

Palavras-chave: Vozes sintéticas; Inteligência Artificial; Avatares; Vozes

Evolução das vozes sintéticas

A evolução das interfaces de comunicação e da IA alcançou um marco impressionante ao aprender a imitar a linguagem humana com uma fidelidade surpreendente, sinalizando que o futuro da IA está agora ao nosso alcance. Nos últimos anos, a IA tem revolucionado a síntese de voz, desenvolvendo vozes artificiais que se aproximam cada vez mais da naturalidade humana. Essas vozes são criadas por sistemas avançados de aprendizado de máquina, nos quais algoritmos processam grandes volumes de dados de áudio para aprender e reproduzir com precisão os padrões da fala humana.

A tecnologia de Text-To-Speech, TTS, tem sido fundamental nesse avanço, transformando texto escrito em discurso audível por meio de regras linguísticas e modelos de síntese de voz. Além disso, a síntese de voz baseada em deep learning² utiliza redes neurais treinadas com

1 Docente e pesquisador na Escola de Comunicação e Artes da Universidade Eduardo Mondlane, Moçambique, com dedicação à pesquisa nas áreas de Mídias Digitais e Inteligência Artificial aplicada à comunicação. Possui Doutorado em Ciências da Comunicação pela Universidade do Minho, Portugal, e Pós-Doutorado em Ciências de Educação Ambiental, Peru. Também é pós-graduado em Inteligência Artificial e Machine Learning pela UNOPAR, Brasil, e em Jornalismo Político pela Universidade do Porto, além de ser especializado em Mídias Digitais pela Universidade de Ohio, EUA.

2 *Deep learning* é um subcampo do aprendizado de máquina que utiliza redes neurais artificiais com múltiplas camadas para processar grandes volumes de dados e aprender padrões complexos.

extensos conjuntos de dados de áudio, permitindo capturar nuances como entonação, ritmo e emoção na produção de fala sintética.

A solução TTS da *ReadSpeaker*³ demonstra de forma clara e precisa como o avanço da IA está sendo estrategicamente direcionados para melhorar tanto a estética quanto a funcionalidade dos produtos e serviços digitais. Ao agregar um elemento humano às vozes sintéticas, essa inovação tecnológica não apenas promove, mas efetivamente realiza um significativo avanço qualitativo na maneira como interagimos e nos relacionamos com as interfaces digitais.

A evolução das vozes sintéticas é um testemunho do progresso notável da tecnologia ao longo das últimas décadas. De acordo com Story (2019), as tentativas iniciais de criar vozes artificiais remontam ao século XVIII, com dispositivos mecânicos como o “Euphonia” de Joseph Faber, que tentou replicar a fala humana usando um sistema complexo de alavancas e tubos. No entanto, foi apenas com o advento da computação moderna, no século XX, que os avanços significativos começaram a ser feitos.

Nos anos 1950 e 1960, os primeiros esforços computacionais para síntese de voz foram realizados. Um dos marcos mais notáveis dessa era foi o “Vocoder”, uma tecnologia inicialmente desenvolvida para fins militares durante a Segunda Guerra Mundial, que mais tarde foi adaptada para experimentos de síntese de voz. Em 1961, o Laboratório Bell desenvolveu o “IBM Shoebox”, um dos primeiros sistemas de reconhecimento de fala, que podia reconhecer 16 palavras faladas (Bell Labs, 1961).

Durante os anos 1970 e 1980, o foco da pesquisa mudou para o desenvolvimento de sistemas de síntese baseados em regras, que utilizavam um conjunto predefinido de regras linguísticas e fonéticas para gerar fala artificial. O sistema DECTalk, lançado na década de 1980, tornou-se um dos sintetizadores de voz mais conhecidos dessa época, usado por indivíduos como Stephen Hawking (Dennis Klatt, 1980s).

Nos anos 1990, a síntese de voz baseada em concatenação tornou-se popular. Essa técnica envolve a colagem de pequenos fragmentos de fala humana gravada para formar palavras e frases. Esta abordagem proporcionou uma qualidade de som mais natural em comparação com os métodos baseados em regras.

Dito isto, sumariamente assume-se que as tecnologias de síntese de voz evoluíram significativamente ao longo do tempo, passando por várias gerações, cada uma com suas próprias características e avanços. A primeira geração foi dominada pela síntese baseada em regras, seguida pela síntese por concatenação, que trouxe maior naturalidade. A terceira geração, com síntese por parametrização, melhorou a expressividade emocional. A chegada das redes neurais na quarta geração permitiu uma reprodução ainda mais fiel da fala humana, culminando na atual geração de síntese de voz baseada em IA generativa, capaz de personalização avançada e interações naturais em tempo real.

3 *ReadSpeaker* é uma tecnologia de conversão de texto em fala (text-to-speech) que permite a leitura automática de conteúdo digital em voz alta, facilitando o acesso à informação para diferentes públicos, como pessoas com deficiência visual ou dificuldades de leitura.

Aplicações das vozes sintéticas

As vozes sintéticas geradas pela Inteligência Artificial (IA) têm encontrado aplicações variadas e inovadoras em diferentes setores, transformando a forma como interagimos com a tecnologia e consumimos informação. No entendimento de Klatt (1987) e Orynbay *et al.* (1987), as tecnologias de IA têm permitido que essas vozes se tornem mais naturais e capazes de entender contextos complexos, melhorando significativamente a aceitação junto dos usuários.

No campo da dublagem, a interseção entre dubladores e IA desperta um interesse considerável. À medida que a IA continua a se aprimorar, surgem questionamentos profundos sobre o futuro das profissões tradicionais, como a dublagem. (Johnson, 2023). Desconstruindo a afirmação de Johnson, a integração da IA na dublagem redefine não apenas os processos criativos, mas também levanta questões sobre autenticidade e inovação. Historicamente marcado pela habilidade humana de transmitir emoções e sutilezas através da voz, esse campo agora se encontra à beira de uma revolução. A integração da inteligência artificial na dublagem não apenas redefine os processos criativos, mas também levanta debates sobre autenticidade, empregabilidade e inovação.

Além disso, os avanços recentes na tecnologia de síntese de voz e no reconhecimento emocional têm impulsionado significativamente a integração da IA na dublagem. Sistemas baseados em Tecnologia de Processamento de Linguagem Natural (PLN) estão cada vez mais aptos a analisar textos e gerar falas naturais que reproduzem com precisão as inflexões e emoções humanas. Esses desenvolvimentos estão abrindo novas possibilidades para dublagens automatizadas em jogos, filmes e séries, onde a consistência vocal e a expressão emocional desempenham papéis cruciais.

Os assistentes virtuais e chatbots representam uma das aplicações mais visíveis dessas vozes. Desde a introdução da Siri da Apple em 2011, assistentes como Amazon Alexa, Google Assistant e Microsoft Cortana utilizam vozes sintéticas para interagir de forma natural e eficiente com os usuários (Klatt, 1987; Orynbay *et al.*, 1987; Wang *et al.*, 2022). Eles realizam desde tarefas triviais como responder perguntas e fornecer informações sobre o clima até controlar dispositivos domésticos inteligentes e gerenciar agendas pessoais. Paralelamente, os chatbots são amplamente adotados no atendimento ao cliente, oferecendo suporte e resolvendo questões básicas sem intervenção humana, o que reduz custos e aumenta a eficiência empresarial (Orynbay *et al.*, 1987; Wang *et al.*, 2022).

Outro campo da massificação das vozes sintéticas assumidas por avatares são redes sociais. Os influencers digitais têm explorado as vozes sintéticas de maneiras inovadoras. Personalidades virtuais como Lil Miquela e Shudu Gram⁴ utilizam essas tecnologias para se comunicar com seguidores, promovendo produtos e causas de forma personalizada e envolvente. Notas claramente nas redes sociais a incorporação de vozes sintéticas para melhorar a acessibilidade, incluindo funcionalidades como leitura de texto em voz alta para ajudar usuários com deficiências visuais a navegar mais facilmente (Klatt, 1987; Orynbay *et al.*, 1987; Wang *et al.*, 2022). Nestes casos, percebe-se que as redes sociais estão a expandir as possibilidades de interação e criando conteúdos mais dinâmicos e inclusivos.

4 Lil Miquela: How a virtual avatar became one of Time's 25 most influential people. Texto escrito por Thea Sokolowski e publicado na revista Outside Insight. Disponível em <https://outsideinsight.com/insights/lil-miquela-how-a-virtual-avatar-became-one-of-times-25-most-influential-people/.05/07/2024>

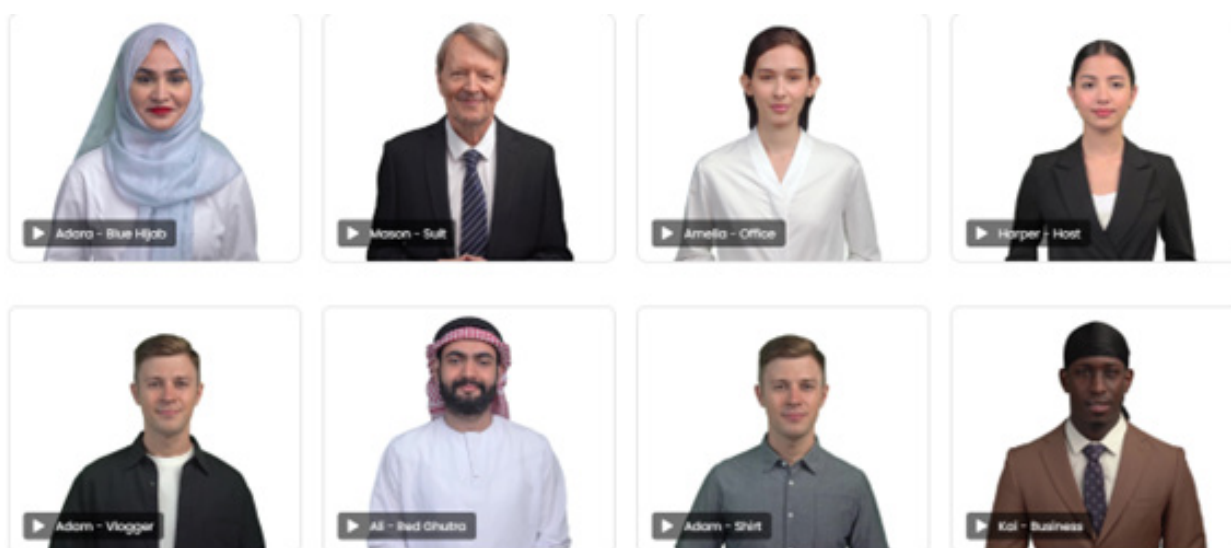
No campo da educação e treinamento, as vozes sintéticas estão a tornar o aprendizado mais acessível e adaptável. Em plataformas de *e-learning*, essas tecnologias narram lições, oferecem *feedback* e respondem perguntas dos alunos, sendo particularmente úteis em cursos *online* que dependem menos da interação humana direta (Klatt, 1987; Orynbay et al., 1987; Wang et al., 2022). A personalização é outra vantagem, permitindo ajustar a voz conforme estilos de ensino e necessidades individuais, inclusive em aplicativos de idiomas para prática de pronúncia correta.

Em marketing e publicidade, as vozes sintéticas estão redefinindo como as marcas se comunicam com os consumidores. Anúncios de áudio personalizados são criados com síntese de voz para alcançar diferentes públicos-alvo, aumentando a relevância e o impacto das campanhas publicitárias (Klatt, 1987; Orynbay et al., 1987; Wang et al., 2022). Chatbots de voz oferecem uma interação mais interativa, personalizada e memorável com os consumidores, enquanto as vozes sintéticas em chamadas automáticas direcionam campanhas de marketing direto de forma eficiente. A capacidade de criar vozes únicas também fortalece o branding das marcas, construindo conexões emocionais mais profundas com os consumidores.

Na produção de documentários e narrativas, a síntese de voz oferece uma maneira flexível e econômica de narrar histórias de alta qualidade. Com a capacidade de ajustar entonação, emoção e estilo, criadores de conteúdo podem criar narrativas envolventes e multilíngues, aumentando a acessibilidade global (Klatt, 1987; Orynbay et al., 1987; Brown, 2020). Essas tecnologias também são usadas em audiolivros, oferecendo uma alternativa de leitura que pode ser apreciada em qualquer lugar, com opções de personalização para melhorar a experiência auditiva.

Os avatares, imagens únicas e antropomórficas geradas pela IA, juntamente com suas vozes e perfis personalizáveis, proporcionam uma narrativa mais adaptada, abrangendo conteúdos, idiomas e culturas variadas. Isso é exemplificado pela figura da IA da VINOD, que demonstra como a IA pode ser utilizada para criar experiências de usuário mais diversificadas e inclusivas. Esses avatares não apenas enriquecem a interação digital, mas também ampliam as possibilidades de personalização, permitindo que os usuários se conectem de maneira mais profunda e significativa com os conteúdos e contextos variados que encontram.

Figura 1: Imagens de Avatares da VINOD AI



Fonte: VINOD IA, 2024

No entanto, a aceitação pública das vozes sintéticas é um fenômeno complexo influenciado por diversos fatores. Segundo Klatt (1987), Orynbay *et al.* (1987) e Wang *et al.* (2022), estudos de caso e pesquisas de opinião revelam que a aceitação depende da naturalidade da voz, a clareza, entonação e a capacidade de expressar emoções. A adaptação ao contexto de uso, a confiança percebida na tecnologia e a personalização da interação também desempenham papéis fundamentais.

Aspectos culturais e demográficos, além da capacidade de interagir de forma empática, são igualmente importantes para determinar como as vozes sintéticas são recebidas e utilizadas pelos usuários em diferentes partes do mundo e em diferentes contextos de aplicação. No entanto, apesar da personalização e da aproximação aos seres humanos, há o risco de exclusão pela IA de vozes e idiomas de comunidades minoritárias ou idiomas não reconhecidos oficialmente pelos organismos internacionais. Essa questão levanta desafios significativos em termos de inclusão e representação, exigindo uma abordagem cuidadosa e sensível no desenvolvimento e na implementação de tecnologias de síntese de voz.

Implicações éticas

A substituição de vozes humanas por vozes sintéticas geradas por Inteligência Artificial (IA) levanta uma série de implicações éticas que precisam ser analisadas com profundidade. Acima de tudo, questões como autenticidade, o impacto nas profissões relacionadas à comunicação e o controle sobre o uso dessas vozes exigem uma abordagem cuidadosa.

Em primeiro lugar, uma das principais preocupações éticas é a questão da autenticidade. Embora as vozes sintéticas, por mais avançadas que sejam em termos de aproximação da fala humana, ainda careçam de uma dimensão essencial: a experiência subjetiva do falante. As vozes humanas carregam consigo uma série de elementos que vão além da simples pronúncia das palavras. Elas transmitem emoção, história, cultura e identidade pessoal. Assim, quando uma voz sintética substitui a de um ser humano em uma função específica, como na narração de um documentário ou em uma publicidade, há o risco de que se perca a autenticidade e o valor emocional que a voz humana transmite.

De acordo com Valenzuela, *et al.*(2024), as tecnologias de IA não conseguem replicar a profundidade da experiência humana que se reflete na fala e no contexto cultural. Desse modo, por mais que as vozes sintéticas possam ser eficientes em termos de clareza e precisão, elas não conseguem atingir a profundidade emocional e cultural que é uma característica inerente da comunicação humana. O uso generalizado de vozes sintéticas pode, portanto, desumanizar certos aspectos da comunicação, criando uma distância emocional entre os interlocutores.

Outra preocupação ética significativa refere-se ao impacto nas profissões relacionadas à voz, como locutores, dubladores, narradores e até mesmo professores e palestrantes. À medida que a IA avança na criação de vozes sintéticas que soam quase indistinguíveis da fala humana, não se pode ignorar que existe um risco real de que essas tecnologias substituam profissionais que dependem da voz para seu sustento. Por conseguinte, em setores como a publicidade e o entretenimento, já há casos de vozes sintéticas sendo utilizadas em lugar de locutores tradicionais, o que gera preocupação entre os profissionais do ramo sobre o futuro de suas carreiras. Um estudo

realizado por He *et al.* (2022) aponta que “a crescente adoção de vozes sintéticas em narrativas audiovisuais pode reduzir significativamente a demanda por locutores e dubladores humanos, sobretudo em setores que priorizam a automação e a redução de custos” (HE et al., 2022).

Além disso, essa transição para o uso de vozes sintéticas pode resultar em uma padronização e homogeneização das expressões vocais. Se, por um lado, a diversidade de sotaques, tons e timbres em vozes humanas reflete a riqueza cultural e social, por outro lado, as vozes sintéticas tendem a ser uniformizadas, eliminando particularidades regionais e individuais. Esse fenômeno pode não só empobrecer a comunicação, como também excluir grupos cujas vozes tradicionais não são representadas adequadamente por essas tecnologias.

No âmbito legal e regulatório, outra implicação ética envolve a questão do consentimento e da propriedade. Quando uma voz sintética é criada com base em gravações de uma voz humana, surgem dúvidas sobre quem detém os direitos de uso dessa voz. Por exemplo, o caso da cantora Elis Regina, que teve sua voz replicada em um anúncio publicitário da empresa Volkswagen no início de julho, em comemoração ao aniversário de 70 anos da empresa. Na campanha, foi utilizada a IA para recriar a imagem da cantora, falecida em 1982, que gerou grande polêmica e levantou a necessidade de regulamentações que protejam a propriedade vocal dos indivíduos.

Com efeito, no campo da IA, “a utilização de amostras vocais humanas para a criação de vozes sintéticas sem o devido consentimento constitui uma violação ética e legal dos direitos de personalidade” (Ouchchy; Coin e Dubljevi, 2020). Essa questão se torna ainda mais complexa em situações onde as vozes sintéticas são utilizadas após a morte de um indivíduo, como o caso da cantora Elis Regina, cuja voz foi recriada postumamente para novos projetos, sem que sua família estivesse ciente ou concordasse com o uso.

Esse exemplo indica a necessidade urgente de uma regulamentação específica para o uso de vozes sintéticas, de tal forma que garanta o respeito pelos direitos dos indivíduos cujas vozes são replicadas, bem como proteja os profissionais de setores impactados pela automação. Assim, é necessário que as empresas e desenvolvedores dessas tecnologias considerem diretrizes éticas que promovam o uso responsável e transparente dessas vozes, assegurando que o público seja sempre informado quando uma voz sintética estiver sendo utilizada em vez de uma voz humana.

Outro aspecto é o impacto psicológico e social no consumidor. Quando os consumidores interagem com vozes sintéticas, especialmente em plataformas de atendimento ao cliente ou assistentes virtuais, podem desenvolver relações mais distantes e impessoais com as marcas e empresas. A confiança e o engajamento, muitas vezes baseados na interação humana, podem ser prejudicados quando essas vozes automatizadas são introduzidas de maneira inadequada. O relatório de Wang *et al.* (2022) sobre o impacto dos assistentes de voz revela que “a falta de interação emocional genuína com vozes sintéticas pode gerar frustração e alienação nos usuários, especialmente quando a tecnologia é incapaz de responder adequadamente a nuances emocionais ou situações complexas” (Wang *et al.*, 2022). Portanto, é fundamental que, além da eficiência técnica, o uso de vozes sintéticas leve em conta a dimensão afetiva e relacional da comunicação.

Por fim, há uma questão mais ampla de justiça social. Em um mundo cada vez mais automatizado, onde a IA ocupa funções que anteriormente eram desempenhadas por humanos, é necessário refletir sobre quem se beneficia dessas inovações. Se a substituição de vozes humanas por vozes sintéticas continuar em ascensão, há um risco de aumento da desigualdade, à medida

que certas profissões e grupos sociais podem ser abandonados. Assim sendo, o avanço das vozes sintéticas deve ser acompanhado por políticas públicas que garantam a proteção do trabalho humano e promovam a inclusão digital e tecnológica de todos os segmentos da sociedade.

Em síntese, as implicações éticas da substituição de vozes humanas por vozes sintéticas são amplas e complexas, abrangendo questões de autenticidade, impacto nas profissões relacionadas, consentimento, regulamentação, e efeitos psicológicos e sociais. Embora a tecnologia ofereça inovações valiosas, é crucial que seu desenvolvimento seja orientado por princípios éticos sólidos, visando proteger tanto os direitos dos indivíduos quanto o bem-estar coletivo. Como argumenta Suchman (2019), “a verdadeira inovação tecnológica só será alcançada quando a ética e a justiça social estiverem no centro de suas práticas” (Suchman, 2019).

Conclusões

As vozes sintéticas, impulsionadas pela IA, têm experimentado um avanço notável nas últimas décadas, aproximando-se cada vez mais da naturalidade humana. Este progresso é evidenciado pela transição das primeiras tentativas mecânicas no século XVIII para os sistemas modernos baseados em redes neurais, capazes de capturar nuances como entonação e emoção.

Johnson (2023), quanto ao avanço das vozes sintéticas representa não apenas um desenvolvimento tecnológico, mas também um ponto de reflexão, onde a distinção entre voz humana e artificial se torna cada vez mais tênue. Embora as aplicações em dublagem e assistentes virtuais sejam promissoras, há um desafio contínuo em garantir que essas vozes não apenas imitem, mas também compreendam a complexidade das interações humanas

No campo das aplicações, as vozes sintéticas estão amplamente integradas em diversas indústrias. Na dublagem, por exemplo, a IA está redefinindo processos criativos e levantando questões sobre autenticidade e inovação. Em assistentes virtuais e chatbots, essas vozes facilitam interações naturais e eficientes com os usuários, enquanto em redes sociais e marketing, elas melhoram a acessibilidade e o engajamento. Além disso, na educação e treinamento, as vozes sintéticas personalizáveis estão transformando o aprendizado online, oferecendo feedback e suporte adaptados às necessidades individuais dos alunos.

Story (2019) explica que a integração de vozes sintéticas em redes sociais e marketing não se limita apenas à eficiência e ao engajamento, mas também levanta questões sobre a autenticidade da comunicação digital. Ele sugere que o próximo passo na pesquisa seja explorar como essas vozes podem ser adaptadas cultural e emocionalmente para diferentes audiências globais, mantendo um equilíbrio entre personalização e aceitação pública.

Perspectivas futuras de estudo podem explorar ainda mais a evolução das vozes sintéticas, focando em avanços contínuos em tecnologias de IA para melhorar a naturalidade e a adaptabilidade dessas vozes em diferentes contextos linguísticos e culturais. Além disso, investigações sobre como a síntese de voz pode ser integrada de forma mais eficaz em áreas emergentes, como saúde, assistência social e interação humano-robô, são cruciais.

Wang *et al.* (2022), o futuro das vozes sintéticas reside na capacidade de oferecer não apenas uma replicação precisa da fala humana, mas também uma adaptação dinâmica às mudanças nos padrões linguísticos e culturais. Eles propõem que estudos futuros investiguem métodos avançados de aprendizado de máquina para melhorar a interpretação contextual e a expressão emocional das vozes sintéticas, ampliando suas aplicações em áreas como educação e assistência digital personalizada

Estudos futuros também podem abordar questões éticas relacionadas ao uso de vozes sintéticas, incluindo a privacidade, a manipulação de informações e o impacto psicológico e emocional dos usuários. Essas perspectivas ajudarão a orientar o desenvolvimento ético e sustentável das tecnologias de síntese de voz, promovendo benefícios sociais e culturais enquanto mitigam potenciais desafios e preocupações.

Referências

- BELL LABS. IBM Shoebox. 1961. Disponível em: <https://www.bell-labs.com/>. Acesso em: 10 set. 2024.
- HE, Xiwei; LIU, Yutong; WUJI, Siguleng; LUO, Ran. **The Impact of Artificial Intelligence Voice Assistants on Consumer Experience Based on the Cognitive-Affective-Conative Model**. 2022. Disponível em: <https://ssrn.com/abstract=4017405>. Acesso em: 10 set. 2024.
- KLATT, D. H. **Review of text-to-speech conversion for English**. The Journal of the Acoustical Society of America, v. 82, n. 3, p. 737-793, 1987.
- LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep learning**. Nature, v. 521, p. 436-444, 2015. Disponível em: <https://doi.org/10.1038/nature14539>. Acesso em: 10 set. 2024.
- ORYNBAY, L.; RAZAKHOVA, B.; PEER, P.; MEDEN, B.; EMERŠI, Ž. **Recent Advances in Synthesis and Interaction of Speech, Text, and Vision**. Electronics, v. 13, p. 1726, 2024. Disponível em: <https://doi.org/10.3390/electronics13091726>. Acesso em: 10 set. 2024.
- OUCHCHY, L.; COIN, A.; DUBLJEVI, V. **AI in the headlines: the portrayal of the ethical issues of artificial intelligence in the media**. AI & Society, v. 35, p. 927-936, 2020. Disponível em: <https://doi.org/10.1007/s00146-020-00965-5>. Acesso em: 10 set. 2024.
- SOKOLOWSKI, Thea. Lil Miquela: **How a virtual avatar became one of Time's 25 most influential people**. In: ***Outside Insight**. Disponível em: <https://outsideinsight.com/insights/lil-miquela-how-a-virtual-avatar-became-one-of-times-25-most-influential-people/>. Acesso em: 05 jul. 2024.
- STORY, B. H. **History of speech synthesis**. In: The Routledge Handbook of Phonetics*. Taylor and Francis, 2019. p. 9-33. Disponível em: <https://doi.org/10.4324/9780429056253-2>. Acesso em: 10 set. 2024.
- VALENZUELA, A.; PUNTONI, S.; HOFFMAN, D.; CASTELO, N.; DE FREITAS, J.; DIETVORST, B.; HILDEBRAND, C.; HUH, E.; MEYER, R.; SWEENEY, M. E.; TALAI FAR, S.; TOMAINO, G.; WERTENBROCH, K. **How Artificial Intelligence Constrains the Human Experience**. Journal of the Association for Consumer Research, v. 9, n. 3, 2024. Disponível em: <https://doi.org/10.1086/730709>. Acesso em: 11 set. 2024.
- WANG, X.; LIU, Y.; HE, W.; WUJI, S.; LUO, R. **The Impact of Artificial Intelligence Voice Assistants on Consumer Experience Based on the Cognitive-Affective-Conative Model**. Disponível em: <https://ssrn.com/abstract=4017405> ou <http://dx.doi.org/10.2139/ssrn.4017405>. Acesso em: 11 set. 2024. Acesso em: 10 set. 2024.