

Vozes Familiares: O Potencial da Clonagem de Voz por IA para Acessibilidade na Educação Inclusiva

Luiz G. dos Santos¹, Luciana C. L. F. Borges¹, Thais R. Kempner², Eunice P. S. Nunes¹

¹Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)
Cuiabá – MT – Brazil

²Faculdade de Engenharia – Universidade Federal de Mato Grosso (UFMT)
Cuiabá – MT – Brazil

luiz.santos8@sou.ufmt.br, lucianafariaborges@gmail.com

thaisrgk@gmail.com, eunice@ufmt.br

Abstract. *This paper presents a systematic literature review on the use of Artificial Intelligence (AI) models for voice cloning as an accessibility tool in basic education, focusing on students with special educational needs, particularly those with Autism Spectrum Disorder (ASD). It analyzes the technological foundations, pedagogical and cognitive underpinnings, and case studies demonstrating the technology's potential. The discussion covers ethical, legal, and social implications, including Brazilian legislation on deepfakes (Bill 1272/2023) and the critique of techno-solutionism. We conclude that voice cloning, by providing familiar voices, offers significant cognitive advantages over generic Text-to-Speech (TTS) technologies, but its implementation requires teacher training, co-design with target communities, and a robust ethical framework to mitigate risks and promote educational equity.*

Resumo. *Este artigo apresenta uma revisão sistemática da literatura sobre o uso de modelos de Inteligência Artificial (IA) para clonagem de voz como ferramenta de acessibilidade na educação básica, com foco em estudantes com necessidades educacionais especiais, particularmente aqueles com Transtorno do Espectro Autista (TEA). Analisam-se as bases tecnológicas, os fundamentos pedagógicos e cognitivos, e os estudos de caso que demonstram o potencial da tecnologia. A discussão abrange as implicações éticas, legais e sociais, incluindo a legislação brasileira sobre deepfakes (PL 1272/2023) e a crítica ao tecnossolucionismo. Conclui-se que a clonagem de voz, ao proporcionar vozes familiares, oferece vantagens cognitivas significativas sobre as tecnologias de Texto-para-Fala (TTS) genéricas, mas sua implementação exige capacitação docente, co-design com as comunidades-alvo e um robusto arcabouço ético para mitigar riscos e promover a equidade educacional.*

1. Introdução

1.1. Contextualização: Desafios da Educação Inclusiva no Brasil

A educação básica no Brasil enfrenta desafios históricos e estruturais para a efetivação de um modelo verdadeiramente inclusivo (SILVA; LOPES, 2020). Apesar de um ro-

busto arcabouço normativo que protege os direitos dos estudantes com deficiência, a transposição da legislação para a prática cotidiana nas salas de aula permanece um obstáculo significativo. Estudos apontam para a persistência de barreiras que vão desde a infraestrutura física e tecnológica inadequada até a carência de material didático adaptado e, de forma crucial, a necessidade de qualificação docente para lidar com a diversidade (BARBOSA, 2025).

Este cenário é agravado por disputas políticas e ideológicas sobre o modelo de educação especial. Um exemplo notório é o Decreto nº 10.502/2020 (BRASIL, 2020), que instituiu a Política Nacional de Educação Especial (PNEE 2020). A política foi amplamente criticada por seu potencial segregacionista (KASSAR; DE SOUZA, 2022) e, por fim, suspensa pelo Supremo Tribunal Federal por contrariar o paradigma constitucional que estabelece a escola regular como o espaço preferencial para a educação de todos os estudantes. Tal contexto reforça a urgência de se desenvolver e implementar ferramentas que fortaleçam a viabilidade e a eficácia da inclusão no ambiente escolar comum, oferecendo respostas práticas aos desafios que alimentam narrativas segregacionistas.

Para estudantes com necessidades educacionais especiais, em particular aqueles com Transtorno do Espectro Autista (TEA), os desafios são ainda mais específicos. Dificuldades na comunicação, na interação social e na manutenção do foco atencional demandam abordagens pedagógicas personalizadas e mediadas. A busca por estratégias que promovam o engajamento e facilitem a aprendizagem para este público é um campo fértil para a inovação tecnológica.

1.2. A Inteligência Artificial como Ferramenta Potencializadora

A Inteligência Artificial (IA) emerge como uma força transformadora com potencial para redefinir paradigmas educacionais (NASCIMENTO, 2024). A crescente adoção de IA no setor de e-learning, impulsionada por um mercado em franca expansão, já se materializa em soluções concretas no cenário brasileiro (GEEKIE, [s.d.]; LIGA VENTURES, 2023). Plataformas de aprendizagem adaptativa, como a Geekie One e a DreamBox, utilizam algoritmos para personalizar as trilhas de aprendizado, ajustando o conteúdo e o nível de dificuldade ao desempenho individual de cada aluno. Esses sistemas, operando de forma análoga à Teoria de Resposta ao Item (TRI) (PASQUALI, 2018), criam diagnósticos precisos e oferecem intervenções pedagógicas direcionadas, otimizando o processo de ensino-aprendizagem.

Para além da personalização do conteúdo, a IA tem sido progressivamente aplicada para desenvolver tecnologias assistivas que visam remover barreiras de acesso. Exemplos incluem tradutores automáticos de Língua de Sinais e, mais proeminentemente, sistemas de Texto-para-Fala (*Text-to-Speech*), que convertem conteúdo escrito em áudio, beneficiando estudantes com deficiência visual, dislexia e outras dificuldades de leitura.

1.3. Problemática e Justificativa: O Potencial da Clonagem de Voz

Apesar de sua inegável utilidade, as tecnologias *Text-to-Speech* (TTS) tradicionais frequentemente se baseiam em vozes sintéticas genéricas, que podem soar robóticas e impessoais. Essa falta de naturalidade pode limitar o engajamento do estudante e, em alguns casos, comprometer a eficácia da ferramenta como recurso de aprendizagem. É neste ponto que a clonagem de voz, impulsionada por modelos de IA generativa, representa uma

evolução qualitativa. Esta tecnologia permite a criação de vozes sintéticas ultrarrealistas, capazes de replicar com alta fidelidade as características únicas de uma voz humana específica, incluindo timbre, entonação e ritmo (VALLABHA et al., 2025; MANIMARAN et al., 2024; ELEVENLABS, 2024).

A hipótese central deste artigo é que a clonagem de voz, ao viabilizar o uso de vozes familiares, como a de um dos pais, de um professor de confiança ou até mesmo do próprio estudante, pode transcender a função de mera acessibilidade para se tornar uma poderosa ferramenta de aprimoramento cognitivo. Esta hipótese é fundamentada no conceito do *"familiar talker advantage"* (vantagem do locutor familiar), validado empiricamente por estudos (e.g., NEW YORK UNIVERSITY, 2014; LEVI, 2015) que demonstram que crianças processam e compreendem a linguagem falada com mais precisão quando esta é proferida por uma voz que lhes é familiar.

Diante do exposto, o objetivo deste trabalho é conduzir uma revisão sistemática da literatura para analisar as evidências científicas, identificar as aplicações atuais e discutir as implicações da aplicação da clonagem de voz por IA como ferramenta de acessibilidade na educação inclusiva.

2. Fundamentação Teórica

2.1. Definição de Acessibilidade no Contexto Educacional

Para os fins deste estudo, o conceito de acessibilidade é compreendido de forma multidimensional, transcendendo a simples remoção de barreiras físicas ou digitais. Acessibilidade educacional, ou mais especificamente, acessibilidade pedagógica (SASSAKI, 2006), é definida aqui a partir de três pilares interdependentes:

1. **Facilidade de Acesso:** Garantir que as ferramentas e os recursos pedagógicos estejam disponíveis e sejam operáveis por todos os estudantes que deles necessitem, independentemente de suas habilidades ou do contexto socioeconômico.
2. **Facilidade de Compreensão e Retenção:** O conteúdo deve ser apresentado de maneira a otimizar o processamento cognitivo, o entendimento e a memorização por parte do aluno, utilizando múltiplos formatos e modalidades sensoriais.
3. **Adequação do Conteúdo:** A capacidade de personalizar e adaptar os materiais e as abordagens pedagógicas para atender às necessidades, preferências e estilos de aprendizagem específicos de cada receptor.

Esses pilares dialogam diretamente com as Diretrizes Curriculares Nacionais para a Educação Infantil (DCNEI) (BRASIL, 2009), que estabelecem os direitos de aprendizagem e desenvolvimento. Em particular, o direito de "Comunicar" (posteriormente revisado para "Expressar") ressalta a importância de a criança poder utilizar diferentes linguagens para expressar sentimentos, ideias e necessidades. As DCNEI, em sua Resolução CNE/CEB nº 5/2009 (BRASIL, 2009), orientam que as propostas pedagógicas devem garantir à criança o acesso a processos de apropriação e articulação de conhecimentos de diversas linguagens, tendo as interações e a brincadeira como eixos estruturantes. Nesse sentido, tecnologias de voz que facilitam a comunicação e o acesso ao conhecimento se apresentam como um meio potente para a efetivação desses direitos.

2.2. Tecnologia de Síntese e Clonagem de Voz

A síntese de fala evoluiu drasticamente nas últimas décadas. Os sistemas *Text-to-Speech* (TTS) iniciais, baseados em concatenação ou modelagem paramétrica, deram lugar a re-

des neurais profundas e modelos generativos que produzem áudio de qualidade e naturalidade sem precedentes (VALLABHA et al., 2025). A clonagem de voz é a vanguarda dessa tecnologia, definida como a capacidade de replicar a voz de uma pessoa específica a partir de uma amostra de áudio (ALMUTAWA et al., 2025).

Plataformas como a ElevenLabs popularizaram o acesso a essa tecnologia, oferecendo APIs que permitem a desenvolvedores integrar vozes clonadas de alta fidelidade em suas aplicações, como demonstrado em projetos de tecnologia assistiva (ELEVENLABS, [s.d.]; NUNES et al., 2024). A qualidade alcançada é tal que as vozes geradas podem ser indistinguíveis das humanas para um ouvinte casual, replicando nuances como entonação, prosódia e carga emocional (VALLABHA et al., 2025; MANIMARAN et al., 2024).

2.3. O "Familiar Talker Advantage": A Base Cognitiva para a Clonagem de Voz

A justificativa mais robusta para o uso de clonagem de voz na educação especial não é meramente tecnológica, mas cognitiva. O estudo seminal "*Children Understand Familiar Voices Better Than Those of Strangers*", conduzido por pesquisadores da NYU Steinhardt, fornece a base empírica para esta afirmação (NEW YORK UNIVERSITY, 2014; LEVI, 2015). A pesquisa investigou se a "vantagem do locutor familiar", já bem documentada em adultos, também se aplicava a crianças em idade escolar (7 a 12 anos).

A metodologia do estudo (NEW YORK UNIVERSITY, 2014) envolveu uma fase de familiarização, na qual as crianças aprenderam a associar três vozes específicas (de falantes bilíngues com sotaque) a personagens de desenho animado. Posteriormente, em uma tarefa de reconhecimento de palavras com ruído de fundo, as crianças ouviram palavras faladas tanto pelas vozes familiares quanto por vozes desconhecidas. Os resultados foram inequívocos: as crianças identificaram com significativamente mais precisão as palavras que já conheciam quando estas eram ditas por uma voz familiar.

Um achado desse estudo (NEW YORK UNIVERSITY, 2014) de particular relevância para a educação inclusiva foi que o benefício da familiaridade vocal era mais acentuado para as crianças que apresentaram o pior desempenho nos testes de base de habilidades linguísticas. Isso sugere que, para estudantes que já enfrentam dificuldades no processamento da linguagem, ouvir uma voz familiar pode liberar recursos cognitivos, facilitando a compreensão e a aprendizagem. Este fenômeno fornece uma poderosa justificativa científica para personalizar tecnologias assistivas de áudio com vozes clonadas de figuras de referência da criança, como pais, cuidadores ou professores, em vez de utilizar vozes sintéticas genéricas.

3. Metodologia

Este artigo adota a metodologia de Revisão Sistemática de Literatura (RSL), conforme as diretrizes PRISMA. O objetivo é identificar, avaliar e sintetizar de forma crítica todas as pesquisas relevantes para responder às seguintes questões de pesquisa (QPs):

- QP1: Quais são as aplicações atuais e emergentes da tecnologia de clonagem de voz baseada em IA como ferramenta de acessibilidade na educação de crianças com necessidades especiais?
- QP2: Quais evidências existem sobre o impacto dessas tecnologias nos indicadores educacionais e no desenvolvimento de habilidades de comunicação e interação social em crianças com TEA?

A estratégia de busca, realizada entre julho e setembro de 2025, envolveu consultas a bases de dados acadêmicas, incluindo IEEE Xplore, ACM Digital Library, Google Scholar e a SBC OpenLib (SOL), utilizando combinações dos termos: "voice cloning", "speech synthesis", "AI in education", "assistive technology", "inclusive education", "autism", "TEA", "clonagem de voz", "educação inclusiva", "tecnologia assistiva".

Os critérios de inclusão abrangeram artigos revisados por pares, relatórios técnicos e estudos de caso, publicados entre 2018 e 2024, com foco em educação básica e aplicações de IA para acessibilidade. Foram excluídos trabalhos focados exclusivamente no ensino superior ou em contextos corporativos não educacionais, bem como artigos sem fontes claras, falta de métricas ou que apresentavam dados inconclusivos.

Após a remoção de 12 duplicatas, um total de 73 artigos foram identificados. Destes, 24 foram triados por título e resumo, resultando em 49 artigos selecionados para análise de texto completo. Após a análise completa, 23 estudos foram incluídos na síntese qualitativa e quantitativa.

4. Análise e Discussão dos Resultados

4.1. Mapeamento das Aplicações de Voz Sintética na Educação Inclusiva

A revisão da literatura revela um uso consolidado de tecnologias de voz sintética, principalmente na forma de TTS e *Speech-Generating Devices* (SGDs), como ferramentas de apoio na educação inclusiva. Estudos demonstram consistentemente os benefícios dessas tecnologias para estudantes com dislexia e outras dificuldades de leitura, que se beneficiam da conversão de texto em áudio para melhorar a compreensão, o foco e a retenção (KEELOR et al., 2020; SUWAHYO et al., 2023). Plataformas comerciais como a Speechify exemplificam essa aplicação, oferecendo vozes de IA de alta qualidade e recursos de anotação para apoiar a aprendizagem multimodal (SPEECHIFY, [s.d.]).

No contexto do TEA, os SGDs são frequentemente utilizados para apoiar a comunicação de crianças minimamente verbais ou não-verbais. Pesquisas indicam que o uso de SGDs em intervenções comportamentais pode acelerar o desenvolvimento da linguagem falada (SHIRE et al., 2023). A hipótese é que a consistência do *output* auditivo e visual do dispositivo, em contraste com a variabilidade da voz humana, pode facilitar o aprendizado.

Contudo, a grande maioria dos estudos analisados foca em sistemas que utilizam vozes sintéticas padrão, ainda que de alta qualidade. A literatura sobre a aplicação específica da clonagem de voz para criar experiências auditivas personalizadas com vozes familiares no contexto educacional é notavelmente mais escassa, representando uma fronteira de pesquisa emergente.

4.2. Estudo de Caso Aprofundado: O Robô Terapêutico Otto

Um dos exemplos mais proeminentes e bem documentados da aplicação de clonagem de voz no contexto brasileiro é o projeto do robô terapêutico Otto, desenvolvido por pesquisadoras da Universidade Federal de Mato Grosso (UFMT) (AGÊNCIA BRASIL, 2022; REBOUÇAS et al., 2021). O Otto é uma tecnologia assistiva de baixo custo, projetada para atuar como um mediador nas sessões de terapia com crianças com TEA, com o objetivo de aumentar o engajamento e facilitar o desenvolvimento de habilidades de comunicação e interação social.

O diferencial do projeto reside na sua abordagem à vocalização do robô. Em vez de utilizar uma voz sintética padrão, a equipe de pesquisa empregou explicitamente a tecnologia de clonagem de voz da plataforma comercial ElevenLabs para criar e expandir o vocabulário do Otto (NUNES et al., 2024). Foi recriada a voz de um menino para manter a naturalidade e a adequação ao público-alvo, resultando em uma interação vocal percebida como mais rica e engajadora (NUNES et al., 2024; KEMPNER et al., 2023).

Este caso ilustra um padrão de inovação de grande relevância para o Brasil e outros contextos com recursos limitados. Em vez de investir no desenvolvimento dispendioso de um modelo de IA de voz proprietário, os pesquisadores utilizaram a API de uma plataforma de ponta para "embarcar" uma capacidade tecnológica de classe mundial em um hardware acessível (baseado em impressão 3D e componentes de baixo custo). Este modelo de "inovação por integração" democratiza o acesso a tecnologias de fronteira, permitindo que o foco dos projetos acadêmicos e de P&D locais seja direcionado para a customização, a aplicação contextualizada e a validação com a comunidade-alvo, aspectos cruciais para o sucesso de qualquer tecnologia assistiva. Adicionalmente, o projeto Otto se destaca por seu processo de design participativo, que envolveu consultas a familiares de crianças com TEA e a equipes multidisciplinares para definir os requisitos do robô, uma prática que busca mitigar os riscos do tecnossolucionismo (ACUNA et al., 2024; HOLT; WALKER, 2025).

4.3. Evidências de Impacto e Potenciais Empecilhos

Apesar do forte embasamento teórico-cognitivo e dos promissores estudos de caso, a literatura ainda carece de evidências robustas sobre o impacto da clonagem de voz em larga escala. A expectativa é que o uso de vozes familiares melhore significativamente os indicadores de engajamento, compreensão e desenvolvimento da linguagem em comparação com o TTS padrão, mas esta hipótese precisa ser validada por meio de estudos controlados e longitudinais.

Os principais empecilhos para a adoção mais ampla desta tecnologia são:

- **Falta de Estudos de Validação:** A área é emergente e necessita de mais pesquisas empíricas que comparem diretamente o impacto de vozes clonadas familiares versus vozes sintéticas genéricas em desfechos educacionais (MERA-MOLINA et al., 2025).
- **Capacitação Docente:** A eficácia de qualquer tecnologia assistiva está intrinsecamente ligada à competência do educador em integrá-la à sua prática pedagógica. A literatura aponta uma defasagem crônica na formação de professores, tanto inicial quanto continuada, para o uso crítico e contextualizado de tecnologias digitais e assistivas (BEZERRA et al., 2024).
- **Barreiras de Custo e Acesso:** Embora o modelo de "inovação por integração" reduza custos de desenvolvimento, o acesso a hardware (tablets, robôs) e software (licenças de plataformas) ainda pode ser um obstáculo, especialmente na rede pública de ensino, com risco de ampliar a desigualdade digital.

5. Implicações Éticas, Legais e Sociais

5.1. O Duplo Fio da Navalha: Benefícios vs. Riscos de Uso Malicioso

A mesma tecnologia que permite criar ferramentas de acessibilidade poderosas também pode ser utilizada para fins maliciosos. A clonagem de voz é a tecnologia central por trás

dos deepfakes de áudio, que podem ser usados para criar desinformação, cometer fraudes, e realizar ataques à reputação de indivíduos (VALLABHA et al., 2025; ELEVENLABS, 2024). A sofisticação da tecnologia torna cada vez mais difícil para o público distinguir entre áudios autênticos e sintéticos.

Em resposta a essa ameaça, o legislativo brasileiro tem se movimentado. O Projeto de Lei nº 1272 de 2023 (SENADO FEDERAL, 2023) propõe a criminalização da adulteração maliciosa de arquivos de vídeo ou áudio "mediante clonagem da voz, substituição de rosto, sincronização labial ou outra ferramenta de inteligência artificial, com a intenção de divulgar notícias falsas ou prejudicar pessoa física ou jurídica" (NOGUEIRA & DE OLIVEIRA SOCIEDADE DE ADVOGADOS, 2024). O projeto estabelece penas de reclusão que podem ser agravadas se o material for divulgado na internet. É fundamental que a regulamentação, embora necessária para coibir abusos, seja elaborada de forma a não inibir a pesquisa e o desenvolvimento de aplicações legítimas e benéficas, como as tecnologias assistivas discutidas neste artigo.

5.2. Consentimento, Privacidade e Autoria Humana

A implementação ética da clonagem de voz no ambiente educacional depende de um arcabouço rigoroso de governança. O princípio fundamental é o consentimento. A voz de um indivíduo é um dado biométrico e um traço de sua identidade. A clonagem da voz de qualquer pessoa, especialmente de uma criança, deve ser precedida de autorização explícita e esclarecida dos pais ou responsáveis legais, com o assentimento da própria criança, sempre que possível. As políticas de plataformas como a ElevenLabs, que exigem que o usuário confirme possuir os direitos necessários para clonar uma voz, representam um padrão mínimo que deve ser adotado e fiscalizado (ELEVENLABS, [s.d.]).

Questões de privacidade e segurança de dados são igualmente críticas. É preciso garantir que as amostras de voz coletadas sejam armazenadas de forma segura, anonimizadas sempre que possível, e utilizadas exclusivamente para os fins consentidos, em conformidade com a Lei Geral de Proteção de Dados (LGPD).

Por fim, é imperativo reforçar o papel da IA como uma ferramenta complementar. A tecnologia deve enriquecer o processo de ensino-aprendizagem, e não substituir a interação humana e o julgamento pedagógico do educador. A autoria e a revisão humana de todo o conteúdo gerado por IA são essenciais para garantir sua qualidade, precisão e adequação pedagógica, mantendo o professor como o curador e mediador central do processo educativo.

5.3. A Crítica ao Tecno-solucionismo

A introdução de tecnologias avançadas na educação corre o risco de cair na armadilha do "tecnossolucionismo", a ideologia que trata problemas sociais complexos, como a desigualdade educacional, como se fossem meramente quebra-cabeças técnicos a serem resolvidos com a ferramenta certa (ACUNA et al., 2024). Essa abordagem, muitas vezes impulsionada por um "complexo de salvador", tende a desenvolver soluções para comunidades marginalizadas, mas sem a sua participação ativa, resultando em ferramentas inadequadas que ignoram as necessidades e experiências vividas dos usuários finais (HOLT; WALKER, 2025).

A própria premissa da clonagem de voz, quando analisada à luz da ciência cognitiva, oferece um contraponto a essa visão. Uma abordagem puramente tecnossolucionista poderia focar em criar a "voz sintética perfeita", otimizada para métricas de clareza e inteligibilidade. No entanto, o princípio do "*familiar talker advantage*" demonstra que a eficácia não reside em uma otimização técnica universal, mas em uma adequação relacional e subjetiva. A "melhor" voz não é a mais clara, mas sim aquela que a criança conhece e em quem confia.

Portanto, a clonagem de voz, para atingir seu potencial máximo, exige uma abordagem intrinsecamente centrada no ser humano. Sua eficácia depende da personalização baseada no contexto social e afetivo de cada estudante. Isso torna a tecnologia um exemplo de como a inovação, quando guiada por evidências científicas sobre a cognição humana e implementada por meio de processos de co-design, envolvendo ativamente estudantes, famílias, professores e terapeutas, como no caso do projeto Otto, pode superar a superficialidade do tecnossolucionismo e gerar impacto real.

6. Conclusão

6.1. Síntese dos Achados

Esta revisão sistemática analisou o potencial da clonagem de voz por IA como uma ferramenta de acessibilidade na educação inclusiva. A análise demonstrou que, enquanto tecnologias TTS já são amplamente utilizadas, a clonagem de voz representa um avanço qualitativo significativo. A principal justificativa para sua aplicação reside no fundamento cognitivo do "*familiar talker advantage*", que postula que vozes familiares melhoram o processamento da linguagem, especialmente em crianças com menores habilidades linguísticas.

Estudos de caso, como o do robô terapêutico Otto no Brasil, ilustram a viabilidade técnica e o potencial terapêutico da integração de APIs de clonagem de voz em plataformas de tecnologia assistiva de baixo custo. Contudo, a expansão dessa tecnologia enfrenta barreiras significativas, incluindo a escassez de estudos de validação de impacto em larga escala, a necessidade crítica de capacitação docente para a integração de IA na prática pedagógica, e os desafios éticos e legais relacionados ao consentimento, privacidade e uso malicioso. A discussão crítica sobre o tecnossolucionismo reforça que o sucesso da tecnologia depende de processos de co-design e de uma abordagem centrada no contexto humano e social do estudante.

6.2. Implicações e Recomendações

Com base nos achados, são propostas as seguintes recomendações:

Para Desenvolvedores e Pesquisadores: Priorizar o desenvolvimento de tecnologias assistivas por meio de metodologias de co-design, garantindo a participação ativa dos usuários finais em todas as fases do projeto. Implementar robustos mecanismos de consentimento e segurança de dados como premissa fundamental.

Para Gestores Educacionais e Formuladores de Políticas Públicas: Investir massivamente em programas de formação continuada para professores, focados não apenas no uso instrumental de tecnologias, mas em sua integração crítica e criativa às práticas pedagógicas inclusivas. Fomentar parcerias entre universidades e escolas para desenvolver e validar tecnologias assistivas contextualmente relevantes.

Para o Poder Legislativo: Avançar na regulamentação do uso de IA generativa, como o PL 1272/2023 (SENADO FEDERAL, 2023), buscando um equilíbrio que puna o uso malicioso da clonagem de voz sem sufocar a inovação em áreas benéficas como saúde e educação.

Referências

- ACUNA, M. et al. Artificial Intelligence Alone Will Not Democratise Education: On Educational Inequality, Techno-Solutionism and Inclusive Tools. *Sustainability*, v. 16, n. 2, p. 781, 2024.
- AGÊNCIA BRASIL. Robô criado pela UFMT auxilia tratamento com crianças autistas. 2022.
- ALMUTAWA, F. et al. A Survey on Voice Cloning. *arXiv*, 2025.
- ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. **Teoria da Resposta ao Item**: conceitos e aplicações. São Paulo: Associação Brasileira de Estatística, 2000.
- BARBOSA, M. I. EDUCAÇÃO INCLUSIVA E PRÁTICA DOCENTE: BARREIRAS, CAMINHOS E ESTRATÉGIAS FORMATIVAS. *Revista Ibero-Americana De Humanidades, Ciências E Educação*, v. 11, n. 8, 2025.
- BEZERRA, E. T. et al. Formação de professores e o uso de tecnologia assistiva e comunicação alternativa e ampliada na educação inclusiva. *Revista Ibero-Americana de Humanidades, Ciências e Educação*, v. 10, n. 10, 2024.
- BRASIL. Decreto nº 10.502, de 30 de setembro de 2020. Institui a Política Nacional de Educação Especial: Equitativa, Inclusiva e com Aprendizado ao Longo da Vida. *Diário Oficial da União*, Brasília, DF, 1 out. 2020.
- BRASIL. Ministério da Educação. Conselho Nacional de Educação. Câmara de Educação Básica. Resolução nº 5, de 17 de dezembro de 2009. Fixa as Diretrizes Curriculares Nacionais para a Educação Infantil. *Diário Oficial da União*, Brasília, DF, 18 dez. 2009.
- EDTECH DIGEST. The Enhanced DreamBox Math by Discovery Education. 2025.
- ELEVENLABS. What is Voice Cloning? A Basic Guide. 2024.
- ELEVENLABS. ElevenLabs Voices: A comprehensive guide. [s.d.].
- GEEKIE. GeekIE One: A plataforma de educação que potencializa a aprendizagem e o ensino. [s.d.].
- HOLT, R.; WALKER, A. Technological solutionism: disability, higher education and generative artificial intelligence. *Disability & Society*, 2025.
- KASSAR, M. C. M.; DE SOUZA, F. F. Decreto nº 10.502/2020: a "nova" política de Educação Especial. *Revista de Educação Pública*, v. 31, p. 1-22, 2022.
- KEELOR, J. et al. The Effect of Text-to-Speech Presentational Features on Reading Comprehension for Students with Reading Disabilities. *Assistive Technology Outcomes and Benefits*, v. 14, n. 1, p. 22-41, 2020.
- KEMPNER, T. R. et al. Desenvolvimento do Robô Otto: Uma Tecnologia Assistiva para Crianças com TEA. *Scribd*, 2023.

- LEVI, S. V. Talker familiarity and spoken word recognition in school-age children. *Journal of Child Language*, v. 42, n. 4, p. 913-920, 2015.
- LIGA VENTURES. Como a Inteligência Artificial pode contribuir com o ensino no Brasil. 2023.
- MANIMARAN, S. et al. A Systematic Literature Review on AI Voice Cloning. *Journal of Engineering and Technology*, v. 8, n. 1, p. 1-10, 2024.
- MEC (Ministério da Educação). Proposta Pedagógica. [s.d.].
- MERA-MOLINA, C. et al. Inclusive education through technology: a systematic review of types, tools and characteristics. *Frontiers in Education*, 2025.
- NASCIMENTO, J. L. A. do. **O Impacto da Inteligência Artificial na Educação: Uma Análise do Potencial Transformador do ChatGPT**. Formiga (MG): Editora MultiAtual, 2024.
- NEW YORK UNIVERSITY. Children Understand Familiar Voices Better Than Those of Strangers. 2014.
- NOGUEIRA & DE OLIVEIRA SOCIEDADE DE ADVOGADOS. Deepfake, regulação e o Projeto de Lei 1.272/23. 2024.
- NUNES, E. P. S. et al. Uso da Clonagem de Voz por Inteligência Artificial para Expansão do Vocabulário do Robô Terapêutico Otto. *Anais da Escola Regional de Informática de Mato Grosso*, 2024.
- NYU STEINHARDT. Some Children Find It Harder to Understand What Strangers Are Saying. 2019.
- PASQUALI, L. **TRI Teoria de Resposta ao Item**: teoria, procedimentos e aplicações. Curitiba: Appris, 2018.
- REBOUÇAS, G. et al. Modelagem 3D do Robô Otto para o atendimento de crianças com Transtorno do Espectro Autista. *Anais da Escola Regional de Informática de Mato Grosso*, 2021.
- SASSAKI, R. K. **Inclusão**: construindo uma sociedade para todos. 7. ed. Rio de Janeiro: WVA, 2006.
- SENADO FEDERAL. Projeto de Lei nº 1272, de 2023.
- SHIRE, S. et al. Responsiveness to Speech-Generating Device Input and Spoken Language in Minimally Verbal Children with Autism. *Journal of Autism and Developmental Disorders*, v. 53, n. 4, p. 1467-1479, 2023.
- SILVA, C. R. da; LOPES, C. C. A formação do professor para a educação inclusiva: entre o ideal e o real. *Revista Educação Especial*, v. 33, 2020.
- SPEECHIFY. Speechify - Text to Speech. Google Play. [s.d.].
- SUWAHYO, H. et al. The Role of Text To Speech Assistive Technology to Improve Reading Ability in E-Learning for ADHD Students. *Journal of Education and Learning*, v. 17, n. 1, p. 108-114, 2023.
- VALLABHA, V. et al. Can you tell the difference between human and AI-generated voices? *PLOS ONE*, 2025.