

## **Vozes entre códigos: o uso crítico da inteligência artificial na transcrição de entrevistas de História Oral**

Edgleide de Oliveira Clemente da Silva (CTI) [edgleide.silva@cti.gov.br](mailto:edgleide.silva@cti.gov.br)

Fernando Rezende Zagatti (CTI) [fernando.zagatti@cti.gov.br](mailto:fernando.zagatti@cti.gov.br)

Filipe Loyola Lopes (CTI) [filipe.lopes@cti.gov.br](mailto:filipe.lopes@cti.gov.br)

### **Resumo**

*O artigo discute o uso de softwares de transcrição automática baseados em inteligência artificial (IA) aplicados à História Oral, destacando seus benefícios, limitações e implicações éticas. Essas ferramentas aceleram o processo de transcrição, permitem identificar locutores, inserir marcações temporais e ampliar o acesso a pessoas com deficiência auditiva. No entanto, ainda enfrentam dificuldades em reconhecer sotaques regionais, gírias e ruídos, além de não compreenderem plenamente o contexto das falas, o que pode gerar distorções. O texto fundamenta-se em autores como Thompson, Portelli, Shopes e Mauad, que defendem a preservação do sentido da oralidade e a responsabilidade do pesquisador na mediação entre fala e escrita. O Centro de Tecnologia da Informação Renato Archer desenvolveu um software próprio, com foco na segurança dos dados, no reconhecimento do português brasileiro e na usabilidade para pesquisadores. Testes comparativos indicaram que os modelos mais avançados (Large-v3 e Turbo) apresentam maior precisão, mas ainda exigem revisão e refinamento humano. Conclui-se que a IA, embora essencial para otimizar e democratizar o trabalho com entrevistas orais, deve ser utilizada de forma crítica e ética, garantindo a fidelidade narrativa, a integridade dos testemunhos e o fortalecimento da memória institucional por meio de uma colaboração equilibrada entre tecnologia e humanidade.*

*Palavras-chave: História oral; Inteligência artificial; Transcrição de entrevistas.*

### **1. Introdução**

O uso de *softwares* de transcrição baseados em inteligência artificial (IA) tem se tornado cada vez mais comum no campo da História Oral, especialmente em processos de transcrição de entrevistas. Essas ferramentas representam um avanço significativo na otimização do tempo e na facilitação do trabalho dos pesquisadores, mas também apresentam limitações que precisam ser cuidadosamente consideradas.

Um dos principais benefícios do uso dessas ferramentas é a agilidade no processo de transcrição. *Softwares* como o [Clipto](#), [Descript](#), [Notta](#), [Otter.ai](#), [TacTiq](#), [Transcriptor](#), [Trint](#) e [TurboScribe](#), entre outros, conseguem transcrever horas de áudio em poucos minutos, algo que manualmente levaria dias ou semanas. Além disso, alguns desses *softwares* oferecem recursos adicionais úteis, como identificação de diferentes locutores, marcação temporal (*timestamps*), importação e exportação em múltiplos formatos e até tradução automática, para entrevistas em línguas estrangeiras.

Outro ponto positivo é a acessibilidade. Ao transformar o conteúdo oral em texto, essas ferramentas contribuem para a inclusão de pessoas com deficiência auditiva e ampliam as possibilidades de uso e disseminação das entrevistas em meios digitais e acadêmicos.

Entretanto, apesar dos avanços, as ferramentas de transcrição automática ainda enfrentam desafios importantes. A qualidade da transcrição depende fortemente da qualidade do áudio, da clareza da fala e da ausência de ruídos de fundo. Entrevistas com sotaques regionais, gírias, pausas, sobreposições de fala ou variações na dicção tendem a apresentar erros de transcrição mais frequentes.

Outro aspecto crítico é a falta de compreensão contextual por parte da IA, o que pode levar à distorção de termos, substituição de palavras similares e à perda de significados importantes no conteúdo oral. Há também a questão da privacidade e da segurança dos dados. Muitos serviços utilizam servidores em nuvem para processar os áudios, o que pode gerar preocupações em relação à confidencialidade das entrevistas, especialmente em projetos que envolvem dados sensíveis ou sujeitos protegidos por termos éticos.

É importante destacar que o uso dessas ferramentas não substitui o refinamento humano, mas o complementa. A revisão manual por parte dos pesquisadores continua sendo essencial para garantir a fidelidade da transcrição e a preservação da riqueza da narrativa oral. Sendo assim, o uso de *softwares* de transcrição por IA no campo da História Oral é uma ferramenta importante, que acelera processos e amplia acessos, mas deve ser empregada com criticidade. A combinação entre tecnologia e análise depurada do pesquisador é o que garante a integridade e a profundidade da memória histórica registrada nas entrevistas.

## 2. Referencial teórico

A incorporação de ferramentas de transcrição automática baseadas em inteligência artificial no campo da História Oral requer um embasamento teórico que permita refletir criticamente sobre seus usos, benefícios e limites. Autores como Paul Thompson e Alessandro Portelli são referências fundamentais nesse debate.

Thompson (2002), em *A voz do passado: história oral*, destaca a importância da fala como fonte histórica viva e insubstituível, chamando atenção para os cuidados necessários na transcrição, a fim de preservar os significados expressos na oralidade. Ele enfatiza a necessidade do cuidado na transcrição para preservar os significados originais expressos na oralidade.

Portelli (2016) aprofunda essa discussão ao afirmar que a História Oral não busca apenas fatos, mas interpretações e sentidos. Para ele, a transcrição é uma forma de tradução entre códigos – da linguagem falada para a escrita – e, como tal, envolve sempre perda e reconstrução de significados. Sua ênfase na oralidade como “performance” reforça o risco de uma transcrição automatizada apagar as nuances da voz, como hesitações, ênfases e emoções, que são constitutivas do testemunho.

Linda Shopes (2002) também contribui para essa reflexão ao tratar da transcrição como mediação e interpretação, ressaltando a importância da ética no tratamento dos depoimentos, sobretudo em ambientes digitais. Ela argumenta que, mesmo diante do uso de tecnologias, é o pesquisador quem deve garantir a integridade e a coerência entre o que foi dito e o que será

registrado. Nesse sentido, o uso de ferramentas de IA precisa estar ancorado em escolhas metodológicas transparentes e comprometidas com o respeito à narrativa do outro.

No contexto brasileiro, Ana Maria Mauad (2015) tem se dedicado à análise da memória nas interfaces digitais e à mediação tecnológica dos registros orais. Seu trabalho auxilia na compreensão de que o uso de *softwares* e dispositivos digitais para a transcrição não é apenas uma questão técnica, mas cultural e política, envolvendo decisões sobre o que se preserva, como se organiza e quem acessa esses registros.

Ao se pensar na construção de um *software* de transcrição, como o que está sendo desenvolvido no CTI Renato Archer, essa dimensão de responsabilidade coletiva e institucional torna-se ainda mais evidente.

Outros autores, como Pierre Lévy e Johanna Drucker, provenientes das áreas das humanidades digitais e da teoria da informação, também oferecem contribuições importantes. Lévy (1999), por exemplo, propõe o conceito de inteligência coletiva, sugerindo que as tecnologias digitais, quando apropriadas de forma crítica e colaborativa, podem ampliar as capacidades humanas de produção e partilha do conhecimento. Já Drucker (2011) chama a atenção para o fato de que os dados são sempre construções (ou *capta*, e não *data*), ou seja, resultados de escolhas, interpretações e processos humanos. Isso é essencial para compreendermos que os resultados das transcrições automáticas não são neutros ou objetivos, mas produtos mediados por algoritmos, treinamentos e parâmetros específicos.

Historicamente, o uso de tecnologia na coleta de memórias orais remonta aos trabalhos de David Boder, que em 1946 utilizou gravadores portáteis para registrar os depoimentos de sobreviventes do Holocausto. Seu pioneirismo já indicava a relevância de pensar a relação entre memória, técnica e ética, um tripé que permanece atual nas discussões sobre IA e história oral.

Nesse sentido, o desenvolvimento de um *software* ético e sensível às especificidades da oralidade, como o que vem sendo testado no Centro de Tecnologia da Informação Renato Archer (CTI), encontra respaldo em uma tradição teórica que valoriza não apenas a fidelidade técnica, mas também o compromisso com a preservação do significado, da experiência e da dignidade de quem compartilha sua história.

### 3. Desenvolvimento do software de transcrição

Considerando os benefícios e limitações dos *softwares* de transcrição baseados em inteligência artificial, o CTI Renato Archer tem se dedicado ao desenvolvimento de um *software* de transcrição automática, voltado especialmente para o campo da História Oral e outras áreas que demandam fidelidade na conversão de fala em texto.

Esse esforço surge da necessidade de alinhar os avanços tecnológicos com os princípios éticos, metodológicos e científicos que regem os projetos de memória institucional, garantindo maior controle sobre os dados, personalização das funcionalidades e acessibilidade à comunidade da instituição. Diferentemente de muitas plataformas comerciais, esse *software* em desenvolvimento busca atender a demandas específicas, como:

- **Privacidade e segurança:** permitindo que todo o processamento ocorra de forma local ou em servidores próprios, respeitando os princípios de confidencialidade exigidos em

projetos de pesquisa com seres humanos.

- **Melhoria na precisão linguística:** com foco no reconhecimento do português falado no Brasil, incluindo variações regionais e vocabulário técnico ou institucional.
- **Usabilidade da interface:** voltada para pesquisadores, com recursos como marcação de trechos relevantes, identificação de locutores e exportação em formatos padronizados (como.txt,.docx e.srt).

Esse projeto reforça o compromisso do CTI Renato Archer com a preservação da memória científica e tecnológica da instituição e do país, promovendo inovação responsável, respeito às especificidades da História Oral e apoio à construção coletiva do conhecimento.

Para viabilizar o desenvolvimento desse *software* de transcrição automática voltado à História Oral, a pesquisadora PCI Edgleide Clemente, em colaboração com os servidores públicos do CTI Renato Archer Fernando Rezende Zagatti e Filipe Loyola Lopes, tem conduzido uma série de testes experimentais com a ferramenta.

Esses testes têm como objetivo principal avaliar o desempenho da transcrição automática em diferentes contextos, identificar pontos de melhoria e adaptar o sistema às necessidades reais de pesquisadores que lidam com registros orais.

As avaliações envolvem entrevistas com diferentes modelos de complexidade – como variações de sotaque, sobreposição de fala, qualidade do áudio e uso de termos técnicos –, permitindo uma análise crítica da capacidade do *software* de reconhecer e transcrever com precisão os elementos fundamentais da narrativa oral, conforme detalhado nos seguintes pontos:

- **Velocidade de processamento e estabilidade do sistema:** Quanto mais rápido o processador do computador e otimização do *software*, menor será o tempo de espera para obter o áudio transcrito. Para arquivos longos, como entrevistas de História Oral, a redução do tempo de transcrição de horas para minutos impacta diretamente na produtividade. Um sistema lento pode comprometer a habilidade do *software* de aplicar os seus recursos de forma eficaz.
- **Entrevistas com qualidade de áudio variada:** Estão sendo utilizados áudios com diferentes níveis de ruído de fundo, volume e nitidez para avaliar a capacidade do *software* em transcrever com precisão em condições menos ideais. Isso permite observar o impacto da qualidade sonora na precisão da transcrição automática e testar mecanismos de filtragem e limpeza de áudio.
- **Falas com sotaques regionais:** Entrevistas com pessoas de diferentes regiões do Brasil estão sendo submetidas ao sistema para verificar o reconhecimento de sotaques linguísticos e vocabulários informais. A diversidade de sotaques revela desafios específicos na identificação correta de palavras e frases, apontando para a necessidade de um treinamento contínuo dos modelos para falas regionalizadas.
- **Presença de múltiplos locutores:** Foram testados trechos com interlocução entre entrevistador e entrevistado, com sobreposição de vozes e alternância rápida entre falas. O objetivo é avaliar a capacidade do *software* em diferenciar os locutores e aplicar marcações adequadas, recurso essencial para entrevistas de história oral. Atualmente, o *software* ainda não realiza a segmentação das falas por locutor.

- **Marcação temporal (*timestamps*):** Em gravações longas, os *timestamps* facilitam a localização de informações específicas, contribuindo na verificação de trechos imprecisos. Esse recurso auxilia na precisão e preservação do contexto da fala, ao considerar as pausas, hesitações e ênfases, que podem ser perdidas na transcrição textual. Isso se alinha com o pensamento de Portelli, que valoriza a oralidade como “performance”. Como também, é um recurso que ajuda a economizar o tempo dos pesquisadores que trabalham com entrevistas extensas de História Oral.
- **Comparação com transcrições manuais revisadas:** Algumas transcrições geradas automaticamente estão sendo comparadas com transcrições realizadas pelos *softwares* Word e AdobePodcast, permitindo medir a taxa de erro, a fidelidade ao conteúdo original e os tipos de correções mais frequentes – como pontuação, nomes próprios, termos técnicos e estruturas gramaticais.
- **Avaliação da usabilidade da interface:** Os testes também incluem a experiência do pesquisador com a interface do *software* de transcrição. Observando-se a clareza das funções, as alterações de modelos de transcrição e as possibilidades de exportação do conteúdo.
- **Flexibilidade na edição posterior do texto gerado:** Diante de possíveis erros de reconhecimento de fala, a flexibilidade de edição permite ao pesquisador adicionar anotações, comentários contextuais, identificadores de tema ou outros metadados relevantes diretamente no corpo da transcrição. Introduzir ou ajustar elementos, como hesitações, pausas significativas, ênfases emocionais (riso, choro) ou outros marcadores de *performance* que são “constitutivos do testemunho”, expõe de forma fidedigna a experiência do entrevistado, enriquece o documento e facilita análises futuras.

Esse processo de testagem e refinamento contínuo é essencial para garantir que o *software* final atenda aos padrões éticos, técnicos e metodológicos exigidos por projetos de memória institucional, como os desenvolvidos no Programa de História Oral (PHO) do CTI Renato Archer. Demonstra, igualmente, a importância da colaboração entre especialistas em ciência da computação e pesquisadores das áreas de ciências humanas, mostrando como a interdisciplinaridade pode gerar soluções inovadoras e sensíveis à complexidade dos processos de registro e preservação da memória.

#### 4. Metodologia dos testes

Neste tópico é apresentada uma análise comparativa do desempenho de diferentes modelos de inteligência artificial para transcrição na reprodução textual de palavras, incluindo nomes próprios, siglas institucionais e vocabulário técnico. Foram avaliadas 71 palavras e expressões geradas por 6 modelos distintos: *Small*, *Medium*, *Large*, *Large-v2*, *Large-v3* e *Turbo* (Radford et al., 2023). Essa análise tinha como objetivo orientar a escolha do modelo mais adequado para o trabalho de transcrição.

Antes, é relevante destacar que a configuração de um computador impacta diretamente na eficiência, precisão e abrangência da execução do *software* de transcrição. Uma máquina mais potente permite que o processamento do áudio ocorra em menos tempo. Portanto, os resultados que serão apresentados a seguir, principalmente considerando o aspecto *tempo de processamento*, estão relacionados ao sistema da máquina utilizada para o teste, a qual apresenta as seguintes configurações:

- **Processador:** Intel® Core™ i7-1165G7 (2.8 GHz até 4.7 GHz, *cache* de 12MB, *quad-core*, 11ª geração)
- **Memória:** 16GB (2x8GB), DDR4, 3200MHz; Expansível até 32GB (2 *slots* soDIMM, sem *slot* livre)
- **Armazenamento:** SSD de 512GB PCIe NVMe M.2
- **Placa de vídeo:** NVIDIA® GeForce® MX330 com 2GB de GDDR5
- **Placa de rede:** 802.11ac (*Wi-Fi* 1x1) + Bluetooth 5.0
- **Bateria:** 4 células e 53Whr (integrada)
- **Tela:** *Full HD* de 14 (1920 x 1080) retroiluminação LED, borda fina, antireflexo, WVA
- **Teclado:** retroiluminado com leitor de impressão digital (configurações com Windows), em Português (padrão ABNT2)
- **Som:** Alto-falantes ajustados profissionalmente ao Waves MaxxAudio® Pro
- **Sistema:** Windows 11 Home Single Language, de 64 *bits* – em Português

Considerando essas informações iniciais, o primeiro teste do *software* foi realizado com a entrevista da servidora Angela Maria Alves, realizada em 13 de agosto de 2024 pelo PHO, com duração de 34 minutos e 55 segundos. O arquivo de áudio foi inserido no *software* e foi utilizado o primeiro modelo de detecção de fala disponíveis na ferramenta: *small* (Figura 1).

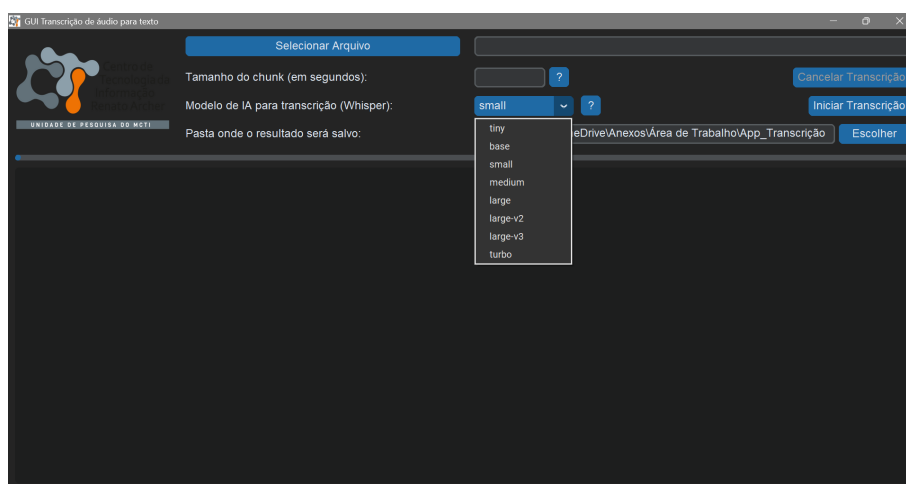


Figura 1 – Interface do software

Durante a análise, observou-se que esse modelo, embora mais rápido, apresentou algumas incoerências na identificação de palavras que são facilmente compreendidas no áudio por pessoas. Erros de substituição de palavras, omissões e dificuldades com nomes próprios foram mais frequentes nesse modelo.

Conforme os modelos mais robustos foram sendo aplicados — especialmente os modelos *large*, *large-v2*, *large-v3* e *turbo* — houve uma notável melhoria na precisão da transcrição, com maior capacidade de reconhecimento de termos complexos, pausas naturais da fala e segmentações mais adequadas (Figura 2). Isso demonstra que os modelos mais avançados oferecem transcrições mais próximas da fala original, embora demandem maior capacidade de processamento e tempo de execução.

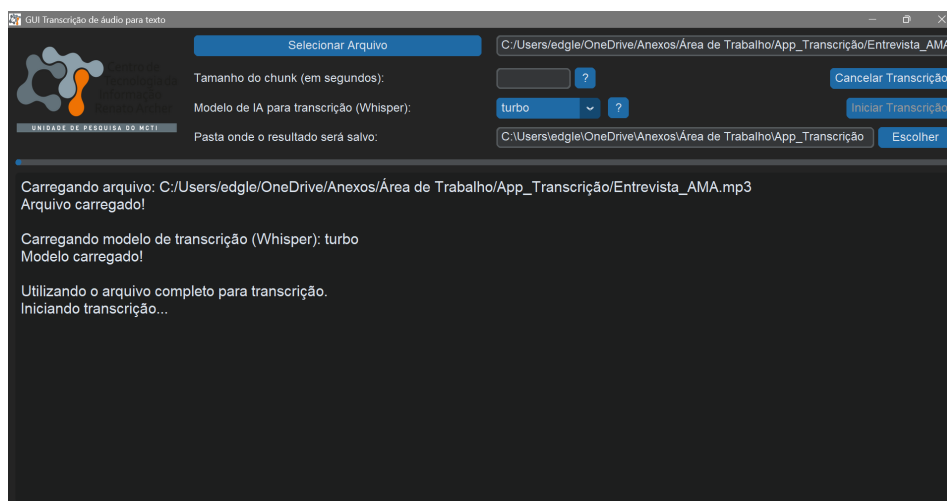


Figura 2 – Software em processo de transcrição

Contudo, ao estabelecer a mudança de um modelo para outro, aparece o seguinte erro (Figura 3):

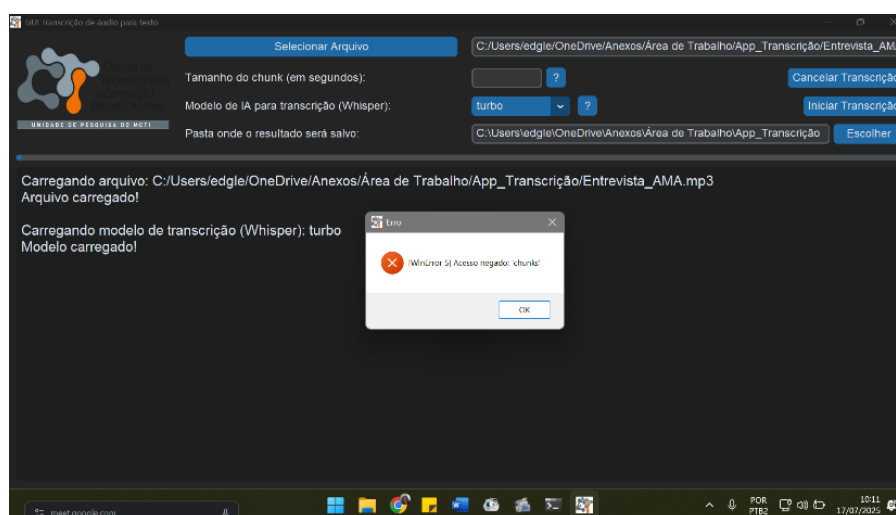


Figura 3 – Erro na tela do software

O erro exibido na imagem, “Acesso negado: ‘chunks’”, indica que o programa não conseguiu acessar um arquivo ou diretório chamado ‘chunks’ ou relacionado a ‘chunks’ devido a permissões insuficientes. Esse tipo de erro geralmente ocorre devido a permissões de usuário restritas, o arquivo ou pasta estar em uso por outro processo ou bloqueio por *softwares* de segurança. No contexto de transcrição de áudio, ‘chunks’ se refere a segmentos do arquivo de áudio que estão sendo processados pela IA.

Para contornar esse problema, a orientação dos desenvolvedores foi excluir a pasta, um procedimento que deve ser realizado a cada mudança de modelo de transcrição. Essa necessidade de intervenção manual para contornar uma falha de acesso (acesso negado) é uma questão operacional que deve ser resolvida na interface. A dependência de o usuário excluir manualmente arquivos para prosseguir com o processamento local, em um *software* que tem a segurança dos dados como prioridade, introduz um fator de risco e instabilidade que contradiz a robustez esperada para projetos de memória institucional.

## 5. Resultados e análise

Os resultados desse primeiro teste foram sistematizados em quadros comparativos. Em geral, a diferença entre a transcrição dos seis modelos envolve a ausência de precisão em nomes próprios (pessoas e lugares), siglas institucionais, termos técnicos e acadêmicos e palavras comuns (Tabela 1).

Classificação	Exemplos (Correto – Errado)	Modelos com erro
Nomes próprios	“Angela” – “Jolene”, “Lajla”, “Juliette”	Small e Medium
	“Archer” – “Acha”, “Ache”, “Archi”, “Arschi”	Todos
	“Barretos” – “Barredos”	Small
Siglas	“MicroMed” – “MicroMED”, “micromédio”	Todos
	“CTI” – vira “CPI”	Large, Large-v2 e Large-v3
	“MAST” – “MASH”	Large
Palavras técnicas	“eletrotécnica” – “eletroteca”	Todos, exceto Large-v2
	“intangível” – “intergível”	Medium e Large-v3
	“tripartite” – “de partite”	Small
Palavras comuns	“sistema” – “sistema”, “ciclema”, “desse tema”, “siquema”	Todos, exceto Large-v2
	“singela” – “se ingela”	Medium
	“Trajetória” – “tranjetória”, “transição”	Small e Large
Estrangeirismo	“boom” – “ubum”	Small
	“commodities” – “comótito”, “comódito”	Todos
	“BigTechs” – “big-tex”, “big tecs”	Small e Large

Fonte: Elaborada pela autora (2025).

Tabela 1 – Classificação de erros e acertos

Os erros mais graves aparecem em nomes próprios e siglas. A confusão de “CTI” com “CPI” demonstra a fragilidade no reconhecimento de vocabulário técnico e institucional, um problema que compromete a fidedignidade textual em contextos acadêmicos e corporativos. As palavras técnicas e acadêmicas foram igualmente deturpadas, como “intangível” virando “intergível”.

O modelo *Turbo* demonstra maior consistência e precisão, sugerindo desempenho superior. Os modelos *Large-v3* e *Large-v2* também apresentam resultados satisfatórios, com alto número de acertos. Já os modelos *Small* e *Medium* têm alto índice de erros, frequentemente substituindo ou distorcendo os termos (Tabela 2).

Modelo	Precisão	Observações
Small	Muito baixa	Muitos nomes errados, substituições aleatórias, com graves distorções nos vocabulários.
Medium	Baixa	Mantém alguns erros, especialmente em nomes próprios e palavras acadêmicas.
Large	Razoável	Corrige várias palavras, mas ainda falha em nomes próprios e siglas. Exemplo: “CTI” torna-se “CPI”.
Large-v2	Boa	Algumas falhas em siglas e nomes, mas com progresso em relação aos anteriores. Ainda há distorções em termos técnicos.
Large-v3	Alta	Poucos erros, desempenho estável, mas algumas falhas em siglas persistem.
Turbo	Altíssima	Apresenta os melhores resultados, com erros pontuais e coerência contextual. Contudo, reproduz erros já melhorados nos modelos anteriores.

Fonte: Elaborada pela autora (2025).

Tabela 2 – Tendência de erros por modelo de transcrição



Considerando que o desenvolvimento de projetos acadêmicos e institucionais exige alta precisão na escrita, os erros em nomes, siglas e termos técnicos comprometem a confiabilidade do *software*. Para o PHO (Programa de História Oral), o uso dos modelos *Turbo* ou *Large-v3* se justifica pela fidelidade lexical superior, mas com a ressalva de que estes modelos, por vezes, perdem a precisão de identificação de algumas palavras já reconhecidas nos modelos inferiores.

O refinamento do pesquisador permanece necessário, mesmo no modelo mais avançado, para a verificação e correção de nomes próprios e palavras técnicas. Neste caso, indica-se a viabilidade do treinamento do *software* na gravação de termos técnicos e na detecção de siglas e estrangeirismos acadêmicos, por meio da criação de um glossário institucional, garantindo a fidelidade narrativa e a integridade do testemunho.

## Conclusão

O desenvolvimento de um *software* livre de transcrição pelo CTI Renato Archer representa uma contribuição relevante tanto para a inovação tecnológica no setor público quanto para a valorização da memória institucional. Ao propor uma solução acessível, colaborativa e sensível às especificidades da História Oral, o projeto reafirma o papel das instituições públicas na promoção de tecnologias comprometidas com a ciência aberta e com os valores democráticos.

No entanto, o estudo evidencia que a tecnologia, por si só, não substitui o olhar crítico e interpretativo do pesquisador. A precisão técnica deve caminhar junto à sensibilidade ética e metodológica, de modo a preservar o sentido, a emoção e a singularidade das narrativas orais. A necessidade de correções lexicais sistemáticas e a resolução de questões operacionais, como o erro de acesso aos ‘chunks’ de áudio, demonstram que o aprimoramento contínuo é imperativo para garantir a estabilidade e a segurança do produto final.

A continuidade dos testes e o diálogo com pesquisadores e profissionais da área serão fundamentais para o aprimoramento da ferramenta, reafirmando a importância da interdisciplinaridade e da ética no uso da IA. Acredita-se que essa iniciativa fortalecerá a articulação entre tecnologia e humanidades no contexto brasileiro.

## Referências

**DRUCKER, J.** *Graphesis: Visual Forms of Knowledge Production*. Cambridge: Harvard University Press, 2011.

**LÉVY, P.** *Cibercultura*. São Paulo: Editora 34, 1999.

**MAUAD, A. M.** *A narrativa da escuta: sentidos e sentimentos na entrevista de história oral*. In: ABREU, M. et al. (Orgs.). *História oral: desafios para o século XXI*. Rio de Janeiro: Letra e Imagem, 2015. p. 71-84.

**NEGRÃO, M.; DOMINGUES, P.** SpeechToText: An open-source software for automatic detection and transcription of voice recordings in digital forensics. *Forensic Science International: Digital Investigation*, v. 38, p. 301223, 2021. DOI: <https://doi.org/10.1016/j.fsidi.2021.301223>. Acesso em: 3 nov. 2025.

**PORTELLI, A.** *A morte na minha rua: uma história oral*. Tradução de Paulo Fontes. São Paulo: Fundação Perseu Abramo, 1997.

**RADFORD, A.; KIM, J. W.; XU, T.; BROCKMAN, G.; MCLEAVEY, C.; & SUTSKEVER, I.** (2023, July). *Robust speech recognition via large-scale weak supervision*. In: International conference on machine

learning (pp. 28492-28518). PMLR. Disponível em: <https://proceedings.mlr.press/v202/radford23a.html>. Acesso em: 3 nov. 2025.

**SHOPES, L.** *Making sense of oral history*. History Matters: The U.S. Survey Course on the Web, 2002. Disponível em: <http://historymatters.gmu.edu>. Acesso em: 15 jul. 2025.

**THOMPSON, P.** *A voz do passado: história oral*. 3. ed. Rio de Janeiro: Paz e Terra, 2002.