

PERSPECTIVAS DE UTILIZAÇÃO DA IA EM MUSEUS: O CASO DO MUSEU DE COMUNICAÇÃO HIPÓLITO JOSÉ DA COSTA

Vinicius Bard¹

Estela Galmarino²

Carlos Barcellos³

RESUMO: Este trabalho explora a aplicação de inteligência artificial (IA) no acervo do Museu de Comunicação Social Hipólito José da Costa (Musecom). O objetivo é apresentar e analisar iniciativas e possibilidades de uso de IA em um contexto museológico, destacando suas potencialidades e limites a partir da experiência prática da instituição. A pesquisa adota uma metodologia exploratória e aplicada, dividida em três frentes. A primeira, já em fase de implementação, é a transcrição e diarização de entrevistas da coleção “Galeria de Vozes” usando a biblioteca Whisper para reconhecimento de fala, que apresentou uma taxa de erro de palavras (WER) de 1,06%. A segunda frente utiliza grandes modelos de linguagem (LLMs) como GPT e Copilot para a geração de resumos descritivos, facilitando a catalogação e pesquisa. A terceira e última frente é uma proposta conceitual de um *chatbot* localmente configurado com a arquitetura de Recuperação Aumentada de Geração (RAG). Treinado com dados do acervo do museu, o chatbot visa auxiliar pesquisadores e a equipe na recuperação de informações de forma eficiente e segura. A combinação dessas etapas demonstra não apenas resultados práticos já alcançados, mas também propõe um caminho futuro para a integração de IA na preservação, documentação e acesso a acervos museológicos.

¹ Graduado em Análise e Desenvolvimento de Sistemas (Unisinos). Graduado em Museologia (UFRGS). Mestrando no Programa de Pós Graduação em Museologia e Patrimônio da Universidade Federal do Rio Grande do Sul e colaborador do Museu de Comunicação Social Hipólito José da Costa.

² Graduada em História pela Universidade Federal do Rio Grande do Sul. Analista em Assuntos Culturais lotada no Museu de Comunicação Social Hipólito José da Costa.

³ Graduado em Produção Fonográfica pela Universidade do Vale do Rio dos Sinos e colaborador do Museu de Comunicação Social Hipólito José da Costa.

Palavras-chave: inteligência artificial. acervo museológico. museologia digital.

1 INTRODUÇÃO

O *Museu de Comunicação Hipólito José da Costa* (MuseCom) está localizado no centro histórico de Porto Alegre, capital do Estado do Rio Grande do Sul. Foi criado pelo Decreto Estadual nº 24.366, de 30 de dezembro de 1975, e tem como missão: “Preservar e difundir os suportes e a memória das distintas formas de comunicação presentes na sociedade gaúcha, tendo como foco nas produções originárias e/ou referentes ao Estado do Rio Grande do Sul” (Museu de Comunicação, 2024, p.4).

O acervo institucional possui as mais variadas tipologias de objetos associados às diferentes áreas da comunicação, isso inclui, mas não restringe-se, jornais, material publicitário, fotografias, filmes em película e digitais, registros de rádio e fonografia em geral. Esta diversidade de coleções implica em uma variedade de suportes da informação e objetos, deste modo a instituição conta com acervos analógicos e digitais em formatos como: películas, quadruplex, betamax, VHS, MiniDV, CDs, DVDs, fitas cassete, disquetes, além de um servidor de armazenamento em rede (*Network Attached Storage – NAS*) que contém objetos digitalizados e nato-digitais.

Esta diversidade não é apenas uma questão de armazenamento, pois ela impõe desafios significativos de processamento técnico, preservação e, sobretudo, de acesso e difusão. A natureza híbrida do acervo, mesclando o analógico e o digital, exige que a instituição desenvolva e aprimore constantemente suas práticas de museografia digital, seja através da digitalização de suportes físicos ou da gestão de coleções que já nascem digitais.

Diante deste cenário, a instituição tem explorado ativamente a implementação de novos sistemas e tecnologias. Contudo, processos museográficos tradicionais, como a transcrição de acervos sonoros ou a descrição e análise de grandes volumes de conteúdo, permanecem como "gargalos" operacionais, sendo tarefas de alta demanda e baixa automação. É precisamente nesta lacuna que a emergência de ferramentas baseadas em Inteligência Artificial (I.A.) se apresenta como uma oportunidade estratégica para otimizar e qualificar o tratamento técnico dos acervos.

Assim, neste trabalho analisaremos três iniciativas do MuseCom que envolvem a aplicação de conceitos de I.A em acervos e coleções museológicas. Duas destas iniciativas que serão apresentadas já foram testadas e possuem considerações favoráveis para sua implementação como parte do rol de processos institucionais, enquanto a terceira e última iniciativa está em estágio de idealização, e sua descrição neste trabalho serve como um exercício de proposta conceitual.

Desta forma este trabalho apresenta a experiência inicial do Museu na verificação de utilização de tecnologias de inteligência artificial, em especial o modelo *Whisper*, desenvolvido pela *OpenAI*, para transcrição automática dos áudios da coleção, também foram testados os modelos *GPT-4* e *DeepSeek* como apoio à análise e descrição dos conteúdos, visando comparativos de desempenho e qualidade. Este é um processo inicial e que serviu para verificar a viabilização de implementação de tecnologias baseadas em inteligência artificial de maneira concomitante aos processos museográficos realizados na instituição.

2 A GALERIA DE VOZES

Ao iniciar os estudos de viabilidade para testes com I.A., identificamos que a coleção *Galeria de Vozes* destacou-se como um terreno propício para o desenvolvimento de aplicações e a experimentação de novas metodologias. Isso ocorre pois esta coleção é composta quase que unicamente por documentos sonoros, sobretudo

cerca de 1.390 fitas magnéticas, cujo conteúdo foi majoritariamente gravado entre as décadas de 1970 e 1990, sendo esta iniciativa retomada recentemente em 2025.

Estes registros documentam produções originárias do Museu de Comunicação mas também de instituições parceiras, além das entrevistas, depoimentos, palestras e outros registros sonoros que documentam aspectos relevantes da trajetória da comunicação no estado. Seu conteúdo abrange entrevistas, depoimentos, palestras e outros registros sonoros, e integra também a coleção gravações de eventos culturais de importância histórica, como seminários e palestras promovidos ou acompanhados pela instituição. Um exemplo notável é o conjunto de registros do 1º Congresso Estadual de Cultura do Rio Grande do Sul, realizado em 1989, evento que antecedeu a criação da Secretaria de Estado da Cultura no Estado.

Este conjunto foi identificado como prioritário durante o inventário geral dos acervos da instituição, realizado a partir de 2019. Esta priorização está relacionada principalmente a dois fatores críticos: seu alto valor histórico e o risco de perda das informações contidas, inerente à obsolescência dos suportes analógicos baseados em fitas magnéticas, que tendem a decair naturalmente com o passar do tempo.

Diante desse contexto, tornou-se prioritária a adoção de estratégias de digitalização, tratamento e descrição desses conteúdos, visando garantir sua preservação. E deste modo, como parte das ações de preservação e difusão, 138 dessas entrevistas já foram digitalizadas. Estes materiais estão acessíveis sob demanda e, para fomentar a descoberta, amostras de 3 minutos de áudio estão disponíveis publicamente no repositório digital da instituição, via Tainacan disponível no endereço: <<https://acervos.musecom.rs.gov.br/>> (Musecom, 2025, doc. eletrônico).

O tratamento técnico da coleção, contudo, envolve desafios significativos enfrentados por uma equipe multidisciplinar. As características intrínsecas do acervo, sobretudo o seu suporte analógico, qualidade de áudio variável, múltiplos falantes e vocabulário técnico ou regionalizado exigem cuidados metodológicos que vão além dos protocolos tradicionais de documentação. Foram justamente essas limitações

operacionais, especialmente o tempo necessário para transcrição e descrição manual, que motivaram a exploração das tecnologias de inteligência artificial descritas neste trabalho.

3 MUSEUS, COMUNICAÇÃO E TECNOLOGIA

A incorporação de tecnologias digitais em museus e espaços museológicos é um processo impulsionado pela digitalização geral da informação, que transformou o ciberespaço no principal canal de comunicação e suporte de memória da humanidade (Lévy, 2010). Uma das características deste processo é a busca por novas formas de preservação, interpretação e mediação dos acervos. No contexto da museologia digital, a adoção de ferramentas baseadas em inteligência artificial se insere como uma continuidade dos esforços de informatização e automação que, desde o final do século XX, vêm transformando as práticas museográficas.

No caso do Museu de Comunicação Hipólito José da Costa, a convergência entre comunicação, tecnologia e museologia representa não apenas uma oportunidade de inovação institucional, mas também um desafio conceitual. Isso porque as ferramentas de IA não apenas otimizam fluxos de trabalho, mas introduzem novas dimensões interpretativas e epistemológicas no processo de mediação museal.

A seguir, são apresentadas as três frentes de experimentação que constituem o núcleo deste estudo: a transcrição e diarização de áudios com o Whisper, o uso de modelos de linguagem para descrição e apoio à catalogação, e o desenvolvimento conceitual de um chatbot baseado em RAG (Recuperação Aumentada de Geração), voltado à mediação entre acervo e público.

3.1 Transcrição e Diarização

A etapa de transcrição constitui um dos maiores gargalos em acervos sonoros e audiovisuais. O processo manual é exaustivo, requer alto nível de concentração e, frequentemente, conhecimento técnico sobre o contexto das gravações. Desta forma, a etapa de transcrição dos conteúdos sonoros da coleção *Galeria de Vozes* foi realizada com o uso da biblioteca Whisper, desenvolvida pela OpenAI (2023), empregada como solução de reconhecimento automático de fala (ASR – *Automatic Speech Recognition*). A escolha do Whisper deve-se à sua alta taxa de acerto, especialmente em contextos de áudio com diferentes níveis de ruído, qualidade e inteligibilidade.

Para aprimorar a separação das falas e a clareza dos conteúdos transcritos, especialmente em segmentos com múltiplos interlocutores ou sobreposição de vozes, o processo foi complementado por técnicas de detecção de atividade de voz (VAD – *Voice Activity Detection*) e diarização automática. Para isso, utilizou-se o repositório *whisper-diarization* (Ashraf, 2025.), que integra o modelo Whisper a ferramentas adicionais de segmentação e identificação de turnos de fala. Essa abordagem permitiu distinguir interlocutores de forma automática, estruturando o conteúdo em blocos organizados de discurso, o que contribuiu para a legibilidade e a análise posterior das entrevistas.

O fluxo de trabalho estabelecido teve início com a importação das matrizes secundárias de áudio, produzidas em conformidade com os padrões da *International Association of Sound and Audiovisual Archives* (IASA, 2017). Os arquivos, no formato .WAV, apresentam resolução mínima de 48 kHz de taxa de amostragem e 24 bits de profundidade de palavra. O processamento foi realizado em uma estação de trabalho equipada com processador AMD Ryzen 5 5600X, 16 GB de memória RAM DDR4 (2666 MHz) e GPU NVIDIA GeForce RTX 3060 (12 GB de VRAM).

A biblioteca foi testada a partir de um conjunto de 5 documentos provenientes da *Galeria de Vozes*, totalizando cerca de 3,5 horas de gravação. As transcrições automáticas geradas foram comparadas com versões manuais previamente revisadas pela equipe institucional, resultando em uma taxa média de erro de palavras (*Word*

Error Rate – WER) de 1,06%, desempenho considerado de excelência mesmo em contextos laboratoriais.

Desta forma o uso do *Whisper* trouxe ganhos significativos em tempo e consistência, reduzindo o tempo médio de transcrição para aproximadamente 25 minutos, incluindo revisão humana. O resultado final das transcrições está em processo de incorporação ao repositório institucional do museu, o sistema Tainacan, ampliando as possibilidades de busca textual e de análise semântica futura.

3.2 Descrição com LLMs

A segunda frente experimental envolveu o uso de modelos de linguagem de larga escala (LLMs), como GPT-4 e DeepSeek, para apoio às etapas de descrição e indexação dos conteúdos transcritos. A proposta consistiu em avaliar a capacidade desses modelos em gerar resumos descritivos e sugestões de palavras-chave a partir das transcrições, de modo a auxiliar o trabalho técnico de catalogação.

Foram elaborados *prompts* específicos que orientaram os modelos a respeitar a terminologia museológica e os padrões de descrição adotados pelo MuseCom. Os resultados demonstraram que as *LLMs* foram capazes de sintetizar com precisão o conteúdo principal das entrevistas, destacando temáticas, nomes de pessoas, lugares e instituições citadas.

De forma geral a utilização destes modelos encontra-se em fase de análise qualitativa e técnica dos resultados, com o intuito de definir qual modelo será adotado como referência principal para a descrição automatizada dos conteúdos da coleção. Entre os critérios considerados estão a precisão das informações geradas, o nível de intervenção humana necessário para revisão, o tempo de processamento, o custo computacional e a facilidade de integração com as ferramentas já utilizadas pela equipe técnica do museu, como o Tainacan e os sistemas internos de documentação.

Além desses aspectos, a equipe também manifesta interesse em experimentar modelos de linguagem locais, executados em infraestrutura própria, de modo a reduzir a dependência de serviços externos e garantir maior autonomia, privacidade e soberania de dados no processo de geração e tratamento das transcrições e descrições automatizadas.

3.3 Chatbot

A terceira frente do projeto consiste na proposição conceitual de um *chatbot* baseado em arquitetura de Recuperação Aumentada de Geração (RAG – *Retrieval-Augmented Generation*), concebido para atuar como interface de pesquisa e mediação do acervo do Museu de Comunicação Hipólito José da Costa. O sistema tem como finalidade ampliar a acessibilidade e a eficiência na recuperação de informações, oferecendo respostas contextualizadas e semanticamente precisas a partir do corpus documental do museu.

O desenvolvimento proposto adota primariamente ferramentas e componentes *open source* (*Free Open Source Software* - FOSS), assegurando autonomia institucional e independência tecnológica. A orquestração do fluxo é realizada por meio da biblioteca *LangChain*, responsável por coordenar as etapas de processamento. O carregamento dos dados é conduzido pelo *DirectoryLoader*, enquanto a segmentação textual, necessária para dividir documentos extensos em blocos menores e semanticamente coesos, é executada através do *RecursiveCharacterTextSplitter*.

Cada segmento resultante seria então vetorizado com o uso de *HuggingFaceEmbeddings*, que converte o texto em vetores numéricos capazes de representar o significado semântico dos conteúdos. Esses vetores são armazenados em um banco de dados vetorial *ChromaDB*, utilizado para recuperar os trechos mais relevantes de acordo com a consulta feita pelo usuário.

Por fim, o componente de geração de linguagem natural é realizado localmente por meio do *Ollama*, que executa modelos de linguagem (LLMs) em infraestrutura própria, preservando a confidencialidade e a soberania dos dados.

Nesse processo, o *LangChain* recebe a pergunta do usuário e coordena a execução das etapas subsequentes. O *ChromaDB* é consultado para localizar os vetores correspondentes aos trechos mais relevantes do acervo, que são então comparados semanticamente via *HuggingFaceEmbeddings*. Esses resultados são enviados ao modelo de linguagem executado pelo *Ollama*, que sintetiza uma resposta textual coerente, fundamentada exclusivamente nas informações recuperadas, e a retorna ao usuário. Essa abordagem assegura fidelidade ao acervo e controle total sobre o ciclo de dados, permitindo que o chatbot opere em ambiente fechado, sem dependência de APIs externas.

Atualmente, o chatbot encontra-se em fase de elaboração conceitual, com o desenvolvimento centrado na definição da arquitetura e na seleção das tecnologias de base. As próximas etapas preveem a realização de testes de aplicabilidade e usabilidade, tanto com a equipe técnica interna, visando avaliar a integração com os fluxos institucionais de documentação, quanto com o público visitante e pesquisadores, a fim de mensurar o potencial de mediação, acessibilidade e engajamento da ferramenta no contexto museológico.

4 CONSIDERAÇÕES FINAIS

As experiências apresentadas neste trabalho demonstram que a aplicação de tecnologias baseadas em inteligência artificial pode representar um avanço significativo na gestão, preservação e difusão de acervos museológicos. O uso do modelo Whisper para transcrição e diarização de conteúdos sonoros mostrou-se uma alternativa tecnicamente viável, eficiente e economicamente acessível, capaz de reduzir substancialmente o tempo de processamento e a carga de trabalho humano. A

automatização parcial dessas etapas não substitui o olhar técnico da museologia, mas o amplia, proporcionando novas possibilidades de análise e indexação de acervos sonoros que, historicamente, apresentavam desafios operacionais e de acesso.

A adoção experimental de modelos de linguagem de larga escala (LLMs) reforça o potencial da IA como ferramenta de apoio à documentação museológica. A geração automática de descrições, resumos e palavras-chave favorece a consistência e a padronização das informações, ao mesmo tempo em que libera a equipe técnica para atividades de maior valor interpretativo. Ainda assim, os resultados apontam para a necessidade de uma mediação humana contínua, responsável por validar, ajustar e contextualizar as informações produzidas.

Por fim, a proposta conceitual de um *chatbot* baseado em arquitetura RAG sinaliza um futuro promissor para a integração entre inteligência artificial e práticas museológicas. A adoção de ferramentas *FOSS* demonstra que é possível construir soluções de alta complexidade técnica com base em tecnologias abertas e acessíveis. Apesar de ainda se encontrar em fase de concepção, o projeto prevê etapas de testagem e validação junto à equipe interna e ao público, com o intuito de avaliar sua aplicabilidade e usabilidade no contexto museal.

Tais iniciativas reafirmam o papel do Museu de Comunicação Hipólito José da Costa como um espaço de experimentação tecnológica, comprometido com a preservação e a mediação do patrimônio cultural por meio das ferramentas digitais do presente.

REFERÊNCIAS

ASHRAF, Mahmoud. *Whisper-diarization*. Software. Disponível em: <<https://github.com/MahmoudAshraf97/whisper-diarization/>>. Acesso em nov. de 2025.

IASA. Comitê Técnico. **A salvaguarda do patrimônio audiovisual: ética, princípios e estratégia de preservação** (IASA-TC 03). 4th ed. Will Prentice e Lars Gaustad (ed.) [tradução de Ariane Gervásio e Marco Dreer; revisão de Carlos Roberto de Souza, Igor

Calado e Ines Aisengart Menezes] Londres: Associação Internacional de Arquivos Sonoros e Audiovisuais, 2017.

LÉVY, Pierre. **Cibercultura**. 3. ed. São Paulo: Editora 34, 2010.

MUSEU DE COMUNICAÇÃO HIPÓLITO JOSÉ DA COSTA. **Plano Museológico 2024–2028**. Porto Alegre, 2024.

MUSEU DE COMUNICAÇÃO HIPÓLITO JOSÉ DA COSTA. **Repositório Digital do Acervo MuseCom**. Porto Alegre, 2025. Disponível em:
<https://acervos.musecom.rs.gov.br/>. Acesso em: 31 out. 2025.

OPENAI. *Whisper: Speech Recognition Model*. 2023. Disponível em:
<https://openai.com/research/whisper>. Acesso em: 31 out. 2025.