

Challenges and Innovations in Accessing and Researching Oral Memories: The Case of the AMOPC Archive

Desafios e Inovações no Acesso e Pesquisa de Memórias Orais: O Caso do Arquivo AMOPC

Paulo Barbosa

Escola Superior de Comunicação Social - IP Lisboa, Portugal

Abstract

In the digital era, oral history audiovisual archives are crucial primary sources for understanding the past through the voices of its protagonists. However, a fundamental challenge persists: unlike text-based documents, verbal content within audiovisual recordings remains largely inaccessible to direct and precise keyword searching. This is particularly critical for extensive oral history interviews, where locating specific information can require hours of viewing. This article presents the case of the Archive of Oral Memory of Communication Professions (AMOPC), an initiative by the School of Communication and Media Studies (ESCS/IPL) in Portugal, which addresses this challenge by innovating in the processing, publication, and, notably, the searchability of its audiovisual collection focused on the transformations in Portuguese communication professions. The paper details the evolution of AMOPC's documentary processing strategies, from manual methods to the integration of Artificial Intelligence (AI) tools. A key innovation discussed is the development and functionality of the AMOPC search engine, a tool capable of performing textual searches within audiovisual content, searching multiple archives with a single search, and providing direct links to the exact timestamps where terms appear. Furthermore, the article contextualizes the relevance of oral history, analyzes the current landscape of oral archives in Portugal, and discusses the implications of these technological advancements for archives, museums, and the future of oral memory preservation.

Keywords: Oral history, Digital archives, Memory preservation, Artificial intelligence, Communication professions

Introdução

Na era digital, os arquivos audiovisuais de história oral representam importantes fontes primárias para a compreensão do passado através das vozes de seus protagonistas.

É neste contexto que surge o Arquivo de Memória Oral das Profissões da Comunicação (AMOPC), uma iniciativa da Escola Superior de Comunicação Social (IPL) e que foi criado com o objectivo de preservar a em discurso directo à memória dos profissionais da comunicação social através dos registos orais das suas histórias de vida pessoais e profissionais.

É crucial diversificar as fontes por isso a expressão adaptada "Dos que não escrevem, não reza a história"

evidencia a necessidade de dar voz a uma diversidade de atores, dos mais notáveis aos anónimos, cujas trajetórias enriquecem o registo histórico coletivo.

O AMOPC não se limita a recolher e preservar testemunhos sobre o campo da comunicação em Portugal; assume também o desafio de inovar nas formas de tratamento, publicação e pesquisa do seu acervo audiovisual.

No entanto, um desafio fundamental persiste. Ao contrário de um livro em formato digital, onde é possível localizar rapidamente uma palavra específica, o conteúdo verbal em vídeos e áudios permanece amplamente inacessível à pesquisa direta e precisa. Este problema torna-se ainda mais crítico quando se trata de extensas entrevistas de história oral, cujo visionamento integral pode exigir horas para localizar informações concretas.

Este artigo apresenta o caso do AMOPC como um estudo sobre os desafios e as soluções encontradas para tornar verdadeiramente acessíveis os arquivos de memória oral na era digital. Os objetivos principais são:

- Contextualizar a relevância da história oral e analisar o atual panorama dos arquivos orais em Portugal, identificando lacunas comuns no acesso e pesquisa.
- Detalhar a evolução das estratégias de processamento documental do AMOPC, desde métodos manuais até a integração de ferramentas de Inteligência Artificial (IA).
- Apresentar o desenvolvimento e funcionamento do motor de busca AMOPC, uma ferramenta inovadora capaz de pesquisar textualmente dentro de conteúdos audiovisuais e fornecer links diretos para os momentos exatos em que os termos aparecem.
- Discutir as implicações destas inovações para arquivos e museus, bem como as perspectivas futuras para o campo da preservação da memória oral.

Estruturalmente, no capítulo 2 o artigo revê sucintamente a trajetória da história oral e o seu estado atual em Portugal.

Seguidamente no 3º capítulo é detalhada a criação e o tratamento documental do arquivo AMOPC, com ênfase na integração tecnológica no desenvolvimento e funcionamento do motor de busca AMOPC.

O capítulo 4 explora as implicações mais vastas e o futuro dos arquivos orais digitais.

Finalmente no capítulo 5 sintetiza as conclusões principais do trabalho.

2. A RELEVÂNCIA, A ORIGEM E EVOLUÇÃO DO USO DOS RELATOS ORAIS NA INVESTIGAÇÃO CIENTÍFICA

2.1 A relevância da memória oral

A expressão “Dos fracos não reza a história” sugere que apenas os feitos dos poderosos foram e serão registrados para a posteridade. Durante séculos, a história foi documentada por detentores de poder e acesso à escrita, enquanto as experiências dos analfabetos ou de quem não publicava eram frequentemente ignoradas.

Se formos rigorosos, o que realmente garantiu um lugar nos arquivos históricos não foi o poder em si, mas sim ter acesso à publicação em forma de livro ou periódicos, por isso a expressão deveria ser: “Dos que não escrevem, não reza a história.”

Felizmente, metodologias contemporâneas como a história oral democratizaram o acesso à memória coletiva, permitindo que vozes anteriormente silenciadas possam contribuir para a construção do conhecimento histórico. O registro oral da memória oral dá a oportunidade de preservar o ponto de vista do cidadão comum que viveu experiências relevantes que passam a poder ser estudadas para além do ponto de vista da classe alfabetizada dominante.

2.2 Antecedentes históricos da utilização de registros orais

O uso de fontes orais remonta aos primórdios da historiografia. Heródoto (c. 484-425 a.C.), frequentemente designado como o “pai da História”, destacou-se pela utilização deliberada de testemunhos orais na sua obra *Histórias*. Sua metodologia baseava-se em três procedimentos que prefiguram práticas contemporâneas: a recolha sistemática de informações através de entrevistas com pessoas de diversas origens; o registro meticuloso dos relatos; e a identificação das fontes, ainda que genericamente (Momigliano, 1990).

Tucídides (c. 460-400 a.C.), autor da *História da Guerra do Peloponeso*, também recorreu a testemunhos orais, demonstrando uma abordagem mais crítica à verificação das informações, particularmente ao avaliar relatos de testemunhas oculares (Finley, 1986).

Estas contribuições representam marcos fundamentais no desenvolvimento do pensamento investigativo, apesar de não se enquadrarem na concepção contemporânea de ciência (Lloyd, 1979)

2.3 A emergência da história oral moderna como metodologia científica

A formalização da história oral como metodologia científica é um fenómeno mais recente, que se desenvolveu gradualmente a partir do século XIX, atingindo maturidade em meados do século XX.

No século XIX, apesar da hegemonia da metodologia histórica positivista que privilegiava documentos escritos oficiais, surgiram precursores importantes como Jules Michelet (1798-1874) que incorporou testemunhos diretos na sua análise da

Revolução Francesa, procurando captar a experiência das classes populares (Gossman, 2009). Henry Mayhew (1812-1887) documentou metulosamente as condições de vida do proletariado londrino em *London Labour and the London Poor* (1851), utilizando entrevistas sistemáticas (Thompson, 2000).

A institucionalização formal da história oral como disciplina académica ocorre apenas em meados do século XX, impulsionada pela invenção do gravador de fita magnética. Allan Nevins (1890-1971) fundou em 1948 o *Columbia Oral History Research Office*, que foi o primeiro programa académico dedicado especificamente à recolha e preservação sistemática de entrevistas gravadas para fins historiográficos (Grele, 1991).

Anteriormente, o *Federal Writers' Project*, desenvolvido durante a década de 1930 nos Estados Unidos, resultou nas *Slave Narratives* - um extenso corpus de entrevistas com ex-escravos que, embora não formalmente classificado como “história oral”, constitui um exemplo significativo de recolha sistemática de testemunhos de grupos marginalizados (Perks & Thomson, 2016).

No Reino Unido, George Ewart Evans iniciou na década de 1950 entrevistas com trabalhadores rurais, documentando práticas e tradições em vias de desaparecimento. Sua obra *Ask the Fellows Who Cut the Hay* (1956) demonstra a aplicação da metodologia para preservação da cultura popular e memória social das classes trabalhadoras (Thompson, 2000).

2.4 O desenvolvimento da história oral em Portugal

O desenvolvimento da história oral em Portugal apresenta características peculiares que refletem o contexto político-social do país. Sua emergência tardia, quando comparada ao contexto anglo-saxónico, deve-se principalmente às limitações impostas pelo regime do Estado Novo, que restringia abordagens historiográficas mais participativas (Godinho, 2001). Os primeiros esforços significativos neste campo surgiram no âmbito da antropologia e etnografia portuguesas através do trabalho pioneiro de Jorge Dias (1907-1973) e Ernesto Veiga de Oliveira (1910-1990), que a partir da década de 1940 recolheram testemunhos orais em comunidades rurais para o *Atlas Etnográfico* (Leal, 2000).

A consolidação da história oral como metodologia reconhecida ocorreu fundamentalmente após a Revolução dos Cravos em 1974, impulsionada pela necessidade de documentar experiências anteriormente silenciadas, relacionadas ao Estado Novo, à resistência antifascista e à guerra colonial (Godinho, 2001). Este período de democratização criou condições favoráveis para a valorização de vozes e narrativas alternativas à historiografia oficial.

Na década de 1980, observou-se uma viragem significativa quando historiadores como Fernando Rosas começaram a incorporar sistematicamente entrevistas como fontes complementares, especialmente em instituições como o Instituto de História Contemporânea da NOVA FCSH e o Instituto de

Ciências Sociais da Universidade de Lisboa (Pimentel, 2015). Manuel Villaverde Cabral destacou-se como pioneiro na aplicação de metodologias próximas da história oral, combinando análise estatística com trajetórias individuais nos seus estudos sobre o mundo rural e migrações (Almeida, 2007).

Pesquisadores como Paula Godinho (2001), Luísa Tiago de Oliveira (2004) e Dalila Cabrita Mateus (1999, 2004) exemplificam a prática e utilização da história oral em Portugal, ilustrando tanto suas potencialidades quanto seus desafios metodológicos. Godinho utilizou entrevistas não diretivas no seu trabalho sobre resistência rural, destacando questões de confidencialidade ao abordar temas sensíveis como prisão e tortura. Oliveira recorreu à memória oral para analisar o Serviço Cívico Estudantil, revelando complexidades ausentes nos documentos escritos da época. Mateus, nos seus estudos sobre elites fundadoras e a PIDE/DGS, reconhece a riqueza informativa dos testemunhos orais, ressaltando simultaneamente a necessidade de crítica rigorosa por tratar-se de “reconstrução da memória”.

Apesar destes avanços, a institucionalização da história oral em Portugal manteve-se fragmentada, surgindo simultaneamente em diferentes centros de investigação e disciplinas das ciências sociais (Freire, 2006). Nos anos 1990, verificou-se uma maior estruturação de projetos académicos e aumento de teses que utilizaram esta metodologia, embora sua aceitação plena como disciplina de ensino superior tenha ocorrido apenas a partir de 2001.

Um problema central identificado na prática da história oral em Portugal é que, mesmo quando utilizados, os testemunhos orais frequentemente não são citados ou explicitamente nomeados como fontes de informação, impedindo a discussão sobre sua importância metodológica. Adicionalmente, o arquivamento de documentos provenientes de recolhas orais é quase inexistente no país, não por obstáculos legais, mas por falta de vontade organizacional, recursos e priorização institucional.

As especificidades dos testemunhos orais exigem crítica rigorosa e cruzamento com outras fontes. Contudo, apesar dos desafios, a utilização da história oral permite investigar novas questões e construir narrativas historiográficas mais inclusivas e diversificadas, contribuindo para uma cultura de pesquisa mais aberta à multiplicidade de vozes e experiências.

2.5 O estado da arte dos arquivos de registos orais em Portugal (abril 2025)

Além do resumo das origens e passado recente da investigação feita com recurso a registos orais em Portugal, achamos que o estado da arte atual dos arquivos de história oral devia ser completado com o levantamento dos arquivos disponíveis. Identificámos 36 arquivos, cuja listagem dinâmica está disponível para consulta pública em <https://sites.google.com/escs.ipl.pt/amopc-pesquisa>.

O panorama dos arquivos de memória oral em Portugal caracteriza-se por notável diversidade de

iniciativas. Instituições como o Arquivo da Memória (2019) e o Museu da Pessoa (inspirado no projeto brasileiro de 1991, com reflexo em Portugal desde 2007) representam esforços significativos na recolha e disponibilização online de histórias de vida. Museus como o Museu do Aljube - Resistência e Liberdade (2015) e o Museu Municipal de Palmela também integram testemunhos orais nas suas atividades.

Observa-se forte presença académica na promoção destes arquivos. Centros como o Instituto de História Contemporânea (IHC-NOVA FCSH), o Centro de Estudos Sociais da Universidade de Coimbra, o CIDEHUS da Universidade de Évora e a Escola Superior de Comunicação Social (ESCS/IPL) lideram projetos relevantes.

Existem também arquivos temáticos importantes, como o Arquivo de Memória Oral das Profissões da Comunicação (AMOPC, ESCS/IPL, 2017). O arquivo da “Memória Oral da Diplomacia Portuguesa” (2012), o “Arquivo Digital do Cante” (CIDEHUS/UÉ) e as coleções do MEMORIAMEDIA e-Museu do Património Cultural Imaterial (IELT/NOVA FCSH). Iniciativas como “Palavras em Movimento” sobre o cinema português e o projeto “MEMnews” (2025) sobre a memória do “consumo” de notícias complementam este panorama.

O formato predominante é o audiovisual, frequentemente disponibilizado online através de websites próprios ou plataformas como YouTube. Contudo, alguns arquivos importantes, como o Arquivo Sonoro de Michel Giacometti, mantêm acesso primordialmente presencial.

Apesar desta vitalidade, os arquivos de memória oral enfrentam desafios significativos. A verificação realizada em abril de 2025 identificou URLs inválidos e problemas técnicos em diversas plataformas, sugerindo dificuldades na manutenção e sustentabilidade a longo prazo dos recursos digitais. A natureza dos arquivos estarem frequentemente ligados a projetos de investigação com curta duração temporal e dependentes de financiamentos específicos, contribui para alguma descontinuidade no registo, tratamento e publicação.

A dispersão das iniciativas e a ausência de um portal agregador centralizado dificultam a descoberta e o acesso unificado aos diversos acervos. Embora existam portais setoriais como o das Instituições de Memória da Defesa Nacional, falta uma visão integrada do panorama nacional da memória oral.

Not obstante os desafios, estes arquivos constituem fontes primárias insubstituíveis para investigação, recursos pedagógicos valiosos e promovem o diálogo intergeracional ao dar voz a experiências frequentemente marginalizadas pela história oficial. A crescente digitalização democratiza o acesso a estas memórias.

Em síntese, o estado da arte dos arquivos de memória oral em Portugal apresenta-se como um ecossistema rico e diversificado, marcado pela proliferação de iniciativas digitais e forte envolvimento institucional. A predominância do acesso online facilita a disseminação, mas a sustentabilidade a longo prazo e a fragmentação dos acervos emergem como

desafios prementes que precisam ser abordados para assegurar a preservação deste patrimônio cultural.

3. O Arquivo AMOPC: Da Construção do Acervo à Pesquisa Potenciada por Tecnologia

3.1. Gênese, Objetivos e Metodologia de Recolha do Arquivo AMOPC

O Arquivo de Memória Oral das Profissões da Comunicação (AMOPC), iniciado em 2017 por investigadores da Escola Superior de Comunicação Social (IPL), responde à necessidade premente de documentar e preservar a história recente das profissões da comunicação em Portugal, um setor marcado por rápidas transformações tecnológicas e laborais. O objetivo primordial é capturar as experiências e perspectivas de profissionais de diversas áreas (jornalismo, audiovisual, publicidade, relações públicas), criando um recurso primário para futuras investigações (Subtil & Garcia, 2017). O projeto adota uma abordagem interdisciplinar, para compreender tanto trajetórias individuais como dinâmicas estruturais da comunicação social.

A construção do acervo AMOPC assenta numa metodologia rigorosa. A seleção dos entrevistados segue um critério de diversificação máxima (Patton, 2002), abrangendo diferentes setores e níveis de notoriedade profissional, para garantir uma visão plural do campo comunicacional. Os registos são realizados em vídeo de alta definição (HD) e áudio de alta qualidade, seguindo padrões internacionais (IASA, 2009), com foco no entrevistado e registo áudio off-screen das perguntas do entrevistador. O protocolo de entrevista é semi estruturado, partindo de um núcleo comum de questões biográficas e profissionais, mas privilegiando a narrativa aberta e cronológica do entrevistado para captar a profundidade da sua experiência. A integridade do testemunho é um pilar, com as entrevistas disponibilizadas na íntegra e oferecendo ao entrevistado a possibilidade de revisão pré-publicação, em respeito pelo consentimento informado.

3.2. Desafios no Acesso e Pesquisa em Arquivos Audiovisuais de História Oral

A disponibilização online de grandes volumes de registos audiovisuais, como os do AMOPC, levanta desafios significativos no que concerne à acessibilidade e à pesquisa eficaz da informação contida. Métodos tradicionais de catalogação, baseados em metadados essenciais ou mesmo indexação temática manual com timecodes associados a tópicos específicos (cf. Pugh, 2005) - implementados nas fases iniciais do AMOPC -, permitem uma navegação geral, mas revelam-se insuficientes para responder a necessidades de pesquisa mais precisas e rápidas.

Identificamos duas limitações principais que dificultam a exploração eficiente destes acervos por parte dos investigadores:

Falta de Precisão na Pesquisa: A pesquisa por termo num arquivo audiovisual extenso raramente conduz o

utilizador diretamente ao ponto relevante da gravação. Idealmente, tal como a busca num livro indica a página exata, a pesquisa num arquivo de memória oral deveria direcionar para o timestamp preciso onde a palavra ou frase foi proferida. No entanto, funcionalidades padrão frequentemente apontam apenas para o ficheiro como um todo, obrigando a um moroso visionamento para localizar a referência pretendida.

O segundo problema é o âmbito limitado da pesquisa uma vez que esta fica geralmente circunscrita aos registos do arquivo específico consultado. Contudo, a necessidade investigativa pode abranger testemunhos sobre um determinado tema (e.g., o papel dos jornalistas no 25 de Abril) disponíveis em múltiplas fontes online (outros arquivos, conferências, entrevistas de rádio/TV, podcasts), exigindo consultas manuais dispersas e morosas. Verifica-se a falta de interoperabilidade e agregação na pesquisa.

O objetivo inicial da criação de uma nova ferramenta de pesquisa visou superar estas lacunas. A pesquisa de palavras por timestamp e a pesquisa agregada em múltiplas fontes online constitui um objetivo central para potenciar a utilidade da ferramenta de pesquisa motor de busca AMOPC para aceder aos arquivos de história oral na era digital.

3.3 Definição do Âmbito: Que Registos Oraís Incluir no Arquivo e na Pesquisa?

Um aspeto crucial considerado na conceção do motor de busca do AMOPC prende-se com a definição do âmbito dos registos oraís a incluir e disponibilizar para pesquisa. Colocou-se a seguinte questão: deverá o acesso para fins de investigação limitar-se aos registos académicos, sistematicamente recolhidos no âmbito de projetos de história oral, ou deverá o arquivo assumir uma abordagem mais abrangente, à semelhança do historiador que utiliza uma variedade de fontes primárias (periódicos, correspondência pessoal, etc.), incluindo registos que não foram originalmente produzidos com o objetivo de documentar histórias de vida?

Para decidir se devemos ou não incluir no motor de busca registos oraís produzidos com objetivos não académicos provenientes de fontes não estritamente académicas ou de projetos de história oral tradicionais (como rádio, televisão, conferências, podcasts, etc.) decidimos ponderar argumentos a favor e contra a inclusão de registos oraís.

3.3.1 Argumentos a Favor da Inclusão de Registos Não Académicos

Encontrámos tres argumentos que justificam a inclusão de registos oraís com origens diversas (rádio, TV, conferências, etc.), para além das entrevistas académicas de história oral, que são os seguintes:

- O primeiro argumento a favor prende-se com a ampliação do Âmbito e Contextualização da Pesquisa. Queremos dizer com isto que a integração de registos provenientes de diferentes meios e contextos de produção permite ao investigador aceder a um leque muito mais vasto de vozes, perspectivas e situações

em que os temas ou indivíduos relevantes para as profissões da comunicação social foram abordados publicamente. Enquanto as entrevistas de história oral tendem a captar a memória e a reflexão a posteriori, estes outros registos captam o discurso e a interação no momento, em variados formatos (debate, notícia, intervenção pública, etc.). Esta diversidade fornece um contexto essencial para as narrativas de história oral e facilita a análise da evolução do discurso público sobre determinados temas ou a participação de indivíduos em eventos específicos. Consequentemente, a pesquisa torna-se mais completa, rica e multifacetada.

- O segundo argumento consiste na potencial descoberta de Fontes Dispersas e Aumento da Eficiência do Investigador na medida em que numerosos testemunhos orais de valor histórico e investigativo (como discursos, painéis de discussão, intervenções públicas) encontram-se dispersos por diferentes plataformas e arquivos digitais, frequentemente com pouca ou nenhuma indexação temática que permita a pesquisa no seu conteúdo verbal. Ao agregar e disponibilizar estas fontes num único ponto de pesquisa, um arquivo ou motor de busca centralizado oferece um serviço de valor inestimável ao investigador, otimizando a identificação e localização destas fontes e poupando tempo e esforço consideráveis. Esta funcionalidade responde diretamente à necessidade prática identificada, por exemplo, no contexto da produção de um documentário que requer a recolha de testemunhos de diversas proveniências.

- O terceiro argumento a favor podemos sintetizar como o reconhecimento de Diferentes Formas de Oralidade e Memória Pública, isto é, embora a história oral académica, com a sua metodologia e foco na experiência de vida, seja uma fonte fundamental, não é a única forma de expressão oral com relevância histórica. Discursos proferidos em conferências, participações em debates televisivos ou radiofónicos, e conversas em podcasts são também registos importantes do pensamento, da argumentação, da ação e das interações de indivíduos em momentos-chave da história social e profissional. Incluir estes formatos reconhece o valor dessas outras "vozes" e discursos públicos como fontes primárias para a história das profissões da comunicação e da sociedade, complementando as narrativas baseadas na memória privada e refletida captada em entrevistas de história oral mais estruturadas.

3.3.2 Argumentos Contra a Inclusão de Registos Não Académicos

Apesar das vantagens, a inclusão de registos orais de origens diversas e não estritamente académicas também apresenta desvantagens que devem ser consideradas:

- O primeiro é a potencial Diluição da Especificidade e Rigor Metodológico da História Oral uma vez que um arquivo de memória oral de origem académica assenta numa metodologia específica, que envolve protocolos éticos rigorosos (nomeadamente, consentimento

informado para a gravação, transcrição e depósito permanente em arquivo, com controlo sobre o acesso e a utilização) e um foco particular na memória subjetiva, na experiência de vida e na construção narrativa individual. Incluir indiscriminadamente registos produzidos para outros fins (como entretenimento, informação noticiosa, difusão pontual em eventos) pode potenciar a diluição do propósito central do arquivo enquanto coleção de história oral, levantar questões sobre a homogeneidade, comparabilidade e fiabilidade das fontes para determinados tipos de análise, e gerar confusão junto dos utilizadores quanto à natureza e proveniência do material disponibilizado.

- Outro argumento contra a inclusão de registos orais feitos fora do âmbito académico tem a ver com desafios éticos, legais e de proveniência uma vez que a integração de material proveniente de rádios, televisões, conferências ou plataformas online levanta questões éticas e legais complexas. O consentimento concedido para uma emissão de rádio ou TV ou para uma apresentação pública em conferência difere substancialmente do consentimento explícito para o depósito e acesso público e permanente num arquivo para fins de investigação histórica.

- Por fim o último factor a considerar tem a ver com o elevado custo e complexidade de processamento técnico e preservação na medida em que o tratamento técnico, a indexação e a preservação a longo prazo de grandes volumes de material audiovisual proveniente de fontes altamente heterogêneas constitui um empreendimento logisticamente exigente, oneroso e tecnicamente complexo. A qualidade técnica dos registos varia consideravelmente, os formatos digitais são múltiplos e a criação de metadados consistentes e suficientemente detalhados para permitir uma pesquisa eficaz tornasse um desafio significativo.

Ponderados os argumentos apresentados, e reconhecendo o potencial significativo de enriquecimento e contextualização da pesquisa para os investigadores, optou-se por integrar no motor de busca do AMOPC a funcionalidade que permite a pesquisa para além dos registos de história oral feitos em âmbito académico. Assim, a ferramenta foi concebida para incluir e processar registos provenientes de diversos contextos, como conferências, programas de rádio, televisão, podcasts, e outros testemunhos orais publicamente acessíveis online, sempre que tecnicamente viável e eticamente permissível. Esta decisão visa maximizar o potencial de descoberta e acesso à informação para a comunidade académica e outros utilizadores interessados na história e nos temas das profissões da comunicação social.

3.4. A Solução Tecnológica: O Motor de Busca AMOPCI

Para colmatar as limitações atuais da ferramenta de pesquisa no arquivo AMOPC e possibilitar a pesquisa noutros arquivos com registos relevantes sobre a comunicação social tivemos que criar uma nova ferramenta a que demos o nome de motor de busca AMOPC com as seguintes características:

- Permitir uma pesquisa precisa no interior de todo o discurso transcrito e não só no índice do vídeo;
- O resultado da pesquisa apresentar contexto através da transcrição do depoimento antes e depois da palavra pesquisada ter sido encontrada;
- executar a pesquisa também e em simultâneo em registos orais para além dos que foram registados e publicados pelos investigadores do AMOPC agregando ainda fontes externas relevantes. Esta funcionalidade era tradicionalmente feita nas bibliotecas através do clipping e permitia ter acesso a recortes de imprensa organizados tematicamente por nome de pessoa, filme, empresa etc

3.4.1. Processamento Documental e Arquitetura do Sistema

A base tecnológica do motor de busca reside num fluxo de trabalho complexo de processamento documental. As entrevistas são transcritas integralmente utilizando software de edição de vídeo (DaVinci Resolve), que não só realiza a transcrição automática com elevada precisão para o português, mas também identifica e diferencia a voz de cada interveniente (Blackmagic Design, 2023). Esta ferramenta gera arquivos de legendas (formato .srt ou similar) conteúdo o texto, a identificação do locutor e os timestamps de início e fim de cada fala.

Subsequentemente, estas legendas são processadas por um script (desenvolvido em Python) que agrupa as falas de cada interveniente, e associa o URL público do vídeo correspondente.

A transcrição completa deve ser corrigida manualmente para identificar e corrigir os erros do processo informático e depois o documento resultante é submetido como fonte à ferramenta de IA notebookLM. Nesta plataforma é solicitado que seja construído o índice da transcrição da entrevista de modo a que seja feita referência a cada tema abordado com a obrigatoriedade de assinalar no índice os nomes de pessoas, empresas, instituições antecidos do timestamp em que foi detectado o termo pesquisado. Deste modo passamos também a ter um índice pormenorizado da transcrição do vídeo / áudio que serve para ser inserido na página do AMOPC e também na descrição do vídeo no youtube.

Este documento estrutura a informação numa base de dados otimizada para pesquisa. Cada entrada na base de dados contém, no mínimo: título/identificação do vídeo, timestamp de início da fala, transcrição do segmento de fala, identificação do locutor (quando disponível) e o URL direto para aceder a esse ponto específico do vídeo.

Para a disponibilização pública, foi criada uma interface web dinâmica (acessível em <https://parabarbosa.github.io/AMOPC-pesquisa/>) que interage com esta base de dados online. O utilizador insere termos de busca, e o sistema retorna os registos encontrados onde o termo foi usado.



Figura 1 – motor de busca AMOPC disponível em <https://parabarbosa.github.io/AMOPC-pesquisa/>

A primeira versão do motor de busca AMOPC ficou operacional em março de 2025 e a respectiva base de dados foi carregada com 260 vídeos sendo que estão catalogados todos os vídeos do AMOPC e para testar a pesquisa em multiplataforma carregamos nesta base de dados quatro vídeos do arquivo audiovisual da ESCS com o registo de conferências, cinco vídeos do podcast Maluco Beleza, um vídeo do arquivo Memória para Todos e seis episódios do podcast Artigo 38 da rádio observador conduzidos por João Miguel Tavares. Ao todo estão catalogadas mais de 125 horas de registos em que alguns têm apenas os índices, outros têm a transcrição completa e bastantes têm os índices e a transcrição.

3.4.2. Funcionalidades e Eficiência na Pesquisa

O motor de busca AMOPC oferece as seguintes funcionalidades:

Pesquisa textual integral quer nos índices do vídeo quer nas transcrições completas das entrevistas indexadas, superando a limitação da pesquisa por metadados ou índices manuais.

Resultados Contextualizados e Precisos: cada resultado apresenta um excerto da transcrição (contexto imediato do termo pesquisado), a identificação da fonte (entrevista/arquivo) e o nome de quem fala.

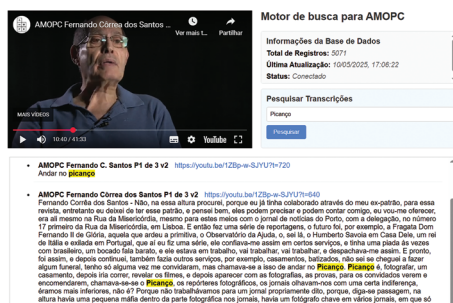


Figura 2 - contexto com a transcrição da entrevista antes e depois do termo "pianço" ter sido encontrado

Acesso Direto ao Segmento Audiovisual: Um hiperlink com timestamp direciona o utilizador para o ponto exato do vídeo (e.g., no YouTube) onde o termo foi proferido, permitindo verificação auditiva e visual imediata (cf. Hedstrom, 2010).

Pesquisa Multi-Arquivo Agregada: A arquitetura permite indexar e pesquisar simultaneamente em diferentes coleções e fontes online (desde que disponíveis na base de dados). Testes preliminares integram conteúdos do AMOPC, do arquivo “Memória Para Todos” (ambos no YouTube) do podcasts Artigo 38 de João Miguel Tavares alojado na plataforma Spotify (com limitações porque não permite utilizar o timestamp), e o podcast em vídeo Maluco Beleza de Rui Unas de forma a mimetizar a lógica de agregação de fontes como se faz no clipping.

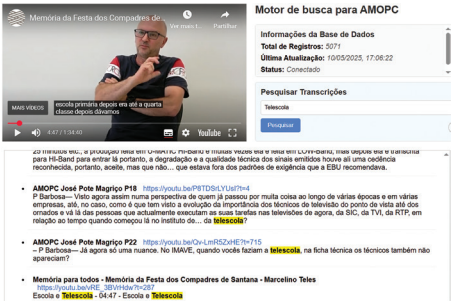


Figura 3 - Na pesquisa por “telescola” foram encontradas todas as referências no interior do arquivo AMOPC e além disso em vídeos do arquivo Memória para todos

A eficácia foi testada com a pesquisa de termos como “Telescola”, gírias profissionais como “Picanço”, ou nomes próprios como “Cicciolina”. Nos três casos os termos foram localizados rapidamente em múltiplas entrevistas dentro e fora do arquivo AMOPC, com acesso direto ao momento exato e transcrição do que foi dito antes e depois para dar um contexto ao resultado da pesquisa. Foi possível transformar horas de pesquisa manual em segundos de busca direcionada.

As deficiências que foram observadas durante a fase de testes foram duas. A primeira foi não existir uma opção para o operador poder limitar a pesquisa apenas dentro de um arquivo específico. A segunda deficiência é na verdade um erro e tem a ver com os registos feitos em áudio e alojados na plataforma spotify. Neste caso o motor de busca AMOPC consegue encontrar a gravação mas o operador ao clicar na hiperligação para abrir e escutar o podcast não consegue abrir na janela o podcast. Além disso, como a plataforma Spotify não tem a função de timestamp a gravação não pode ser ouvida automaticamente no instante onde foi usada a palavra pesquisada.

3.5. Implicações para a Investigação e o Futuro dos Arquivos de Memória Oral

A implementação deste sistema, alicerçada na aplicação de tecnologias de Inteligência Artificial (IA) como o Reconhecimento Automático de Fala (ASR), representa um avanço significativo. Transforma o conteúdo verbal dos arquivos audiovisuais em texto pesquisável, abrindo caminho para métodos avançados de análise de conteúdo ou discurso (cf. Bardin, 2011) e pesquisa exploratória.

Esta abordagem alinha-se com conceitos emergentes como a “arquivística aumentada” (augmented archival science), onde a tecnologia potencia as tarefas de curadoria, análise e acesso, complementando a interpretação humana crítica (cf. Abramowitz et al., 2024). O percurso do AMOPC, partindo de práticas metodológicas consolidadas em história oral (cf. Thompson, 2000; Ritchie, 2015), ilustra como a inovação tecnológica pode superar desafios inerentes ao tratamento e disseminação de grandes volumes de dados audiovisuais.

A combinação da recolha rigorosa de narrativas orais com o processamento inteligente e a pesquisa agregada resulta num recurso dinâmico, mais acessível e funcional. Este motor de busca não é apenas uma ferramenta de acesso, mas um instrumento que potencia novas formas de investigação, facilitando análises comparativas, estudos de linguagem, mapeamento de redes e a descoberta de conexões inesperadas entre testemunhos. Aponta, assim, novos caminhos para o futuro os vídeos e áudios passarem a ser pesquisados como um livro permitindo deste modo maximizar o seu valor para a comunidade académica, criadores de conteúdo audiovisual e para os utilizadores em geral (cf. Frisch, 2016; Huvila, 2008).

4. Inovação e o Futuro dos Arquivos de Memória Oral

O caso do AMOPC, com a sua integração de ferramentas digitais para tratamento e pesquisa, não só otimiza o acesso ao seu próprio acervo, mas também aponta caminhos promissores para o futuro da preservação e exploração da memória oral. As inovações desenvolvidas têm implicações que transcendem o projeto, sugerindo novas abordagens para estações de televisão, museus, exposições, arquivos, além de ser possível dar uma “segunda vida” aos registos orais feitos no âmbito de investigações concretas.

No âmbito dos arquivos das estações de televisão a possibilidade de pesquisar o uso de certas palavras no interior dos programas em vídeo torna mais fácil e rápida a pesquisa de imagens de arquivo.

A capacidade de integrar testemunhos orais pesquisáveis em exposições e acervos digitais oferece uma oportunidade única para adicionar profundidade e dimensão humana a objetos e narrativas históricas. Museus e arquivos podem adotar modelos como o do AMOPC para:

- Ligar artefactos ou documentos a testemunhos orais que revelem o seu uso, significado ou contexto de produção. Deste modo além de um texto a acompanhar um objeto exposto o registo oral de quem usou o objecto extravasa a descrição factual e criando conexões mais fortes com o público (cf. Sandell, 2007).
- Utilizar a pesquisa precisa em arquivos orais para destacar perspectivas diversas e frequentemente sub-representadas sobre eventos ou temas históricos, enriquecendo a curadoria e promovendo uma “autoridade partilhada” (Frisch, 1990).

- Integrar excertos de vídeo/áudio relevantes, facilmente localizáveis através de motores de busca de registos orais tornam as exposições mais interativas e ricas.

A última consequência que podemos identificar que pode ter o registo público e pesquisável de registos orais é o de no âmbito de um estudo académico o investigador depois de ter feito entrevistas com algumas fontes; poder arquivar as entrevistas numa base de dados pública com os registos completos e desta forma outros investigadores passarem a ter acesso à entrevista oral para obter informação para novos objetivos de estudo.

Um caso concreto que pode tornar mais clara esta vantagem foi por exemplo uma entrevista feita com Teresa Fernandes para o AMPC com o objectivo imediato de fazer um documentário sobre auto-promoções televisivas. Uma vez que a conversa registrada abordou toda a sua carreira profissional ficou muita informação disponível na entrevista integral sobre publicidade, rádio e aspectos diversos do funcionamento da estação de televisão TVI desde a sua origem.

O AMOPC demonstra que a tecnologia pode tornar viável esta integração em larga escala, transformando arquivos orais de repositórios passivos em recursos ativos disponíveis para que seja fácil e rápida a pesquisa de informação através de fontes orais.

4.1. O Próximo Passo: Interação Inteligente com Arquivos Orais

O motor de busca do AMOPC representa um avanço significativo na localização de informação. Contudo, o futuro aponta para sistemas mais sofisticados: agentes de inteligência artificial (IA) capazes não apenas de encontrar, mas de analisar, sintetizar e responder a questões complexas com base no conteúdo dos arquivos. Imaginemos um investigador ou membro do público a poder “dialogar” com o arquivo, perguntando, por exemplo, “Quais foram os principais desafios éticos mencionados pelos jornalistas que cobriram a Revolução de 1974?” e receberem uma resposta sintetizada, fundamentada em múltiplos testemunhos.

Para alcançar esta funcionalidade será necessário:

- Ir além da transcrição, enriquecendo os dados com anotações semânticas que permitam à IA compreender contextos, temas e relações.
- Desenvolver ou adaptar modelos de IA especificamente treinados em grandes volumes de história oral para garantir respostas relevantes e fiáveis.
- Criar plataformas que permitam a interação em linguagem natural, indicando sempre claramente as fontes primárias utilizadas para gerar as respostas.

Esta evolução transformaria os arquivos orais em verdadeiras bases de conhecimento complementares às fontes escritas.

4.2. Requisitos para a criação de um motor de busca agregador de Arquivos Orais

A proliferação de arquivos de memória oral, embora positiva, resulta frequentemente numa paisagem fragmentada que dificulta a pesquisa abrangente. A funcionalidade do motor de busca AMOPC foi testada com registos pontuais mas pode ser expandida para criar portais agregadores mais vastos, a nível nacional ou temático. Tal iniciativa exigiria:

Um compromisso entre diferentes arquivos e museus para partilharem metadados e, idealmente, acesso (controlado) às transcrições ou conteúdos.

Adoção de protocolos e formatos de metadados padronizados para garantir a interoperabilidade entre sistemas (Benardou et al., 2010).

Desenvolvimento de uma plataforma tecnológica robusta e de um modelo de governação que assegure a atualização e manutenção contínuas (cf. Rumsey, 2016).

Um ecossistema agregado de memória oral democratiza radicalmente o acesso a estas fontes ricas, potenciando investigações comparativas e uma compreensão mais holística do passado através das vozes de quem o viveu.

Em suma, as tecnologias digitais e a IA oferecem oportunidades sem precedentes para potenciar o valor da história oral, movendo-a de um nicho especializado para uma fonte central e acessível na construção da memória coletiva.

5. Conclusão: Potenciar o acesso aos arquivos de Memória Oral

A preservação da memória oral constitui um pilar fundamental para a construção de uma compreensão histórica mais rica, plural e inclusiva. O Arquivo de Memória Oral das Profissões da Comunicação (AMOPC) insere-se neste esforço, focando-se num setor vital da sociedade portuguesa contemporânea. No entanto, como este artigo demonstrou, a simples recolha e armazenamento de testemunhos audiovisuais não garante a sua plena utilização. O verdadeiro desafio reside em tornar estes vastos e valiosos recursos efetivamente acessíveis e pesquisáveis.

O percurso do AMOPC ilustra uma resposta pragmática e inovadora a este desafio. Confrontado com os estrangulamentos do processamento manual de longas entrevistas, o projeto adotou progressivamente ferramentas tecnológicas, culminando na integração de software de edição com capacidades de IA (DaVinci Resolve) e plataformas de análise de texto (NotebookLM). Esta abordagem permitiu ultrapassar o obstáculo da transcrição morosa e abriu caminho para um tratamento documental mais profundo e eficiente, gerando transcrições cada vez mais precisas, identificando locutores e facilitando a criação semiautomática de índices temáticos e referenciais.

A principal contribuição apresentada pelo motor de busca AMOPC reside no desenvolvimento de uma ferramenta específica para o conteúdo do AMOPC e

potencialmente agregador de outros arquivos. Esta ferramenta supera as limitações da pesquisa baseada apenas em títulos e índice dos vídeos, permitindo aos utilizadores pesquisar textualmente todo o conteúdo verbal das entrevistas. A capacidade de devolver excertos contextuais e links diretos com timestamp para os segmentos de vídeo relevantes representa um salto qualitativo na eficiência da pesquisa em arquivos orais audiovisuais, aproximando a sua usabilidade à dos documentos textuais.

As implicações desta abordagem estendem-se para além do AMOPC. Demonstra um modelo viável para estações de televisão, arquivos e museus que procurem não só preservar testemunhos orais, mas também integrá-los ativamente nas suas narrativas e plataformas digitais, conferindo uma dimensão humana às suas coleções e facilitando novas formas de acesso a fontes primárias de investigação. O potencial futuro inclui o desenvolvimento de agentes de IA mais sofisticados para interação conversacional com os arquivos e a criação de ecossistemas agregados que combatam a fragmentação atual do panorama arquivístico.

Em suma, o caso AMOPC evidencia que a combinação criteriosa de metodologias consolidadas de história oral com as potencialidades das tecnologias digitais e da inteligência artificial é essencial para desbloquear o imenso valor contido nas memórias orais. Ao superar os desafios de acesso e pesquisa, não só se democratiza o acesso a estas fontes primárias, como também se potencia a sua contribuição para uma construção mais dinâmica, participada e completa da memória coletiva. Para isso é necessário agora e no futuro fazer a preservação da memória escrita e oral de modo a todos poderem contribuir para as narrativas que os historiadores do futuro vão fazer sobre a nossa época.

Bibliografia

Abramowitz, Clara, Yoshua Bengio e Joelle Pineau. 2024. "Augmenting Archives: AI Tools for Enhanced Access to Oral History Collections." *Journal of Digital Archives* 15 (2): 112-129.

Almeida, Ana Nunes de. 2007. *"A sociologia e a descoberta da vida quotidiana"*. Lisboa: Instituto de Ciências Sociais.

Almeida, João Ferreira de. 2007. "Velhos e novos aspectos da epistemologia das ciências sociais." *Sociologia, Problemas e Práticas* 55: 11-24.

Andrade, Maria Ana Pires Cardoso de. 2008. *"Arquivos orais em Portugal - que testemunhos? uma reflexão sobre arquivos de oralidade em Portugal"*. Tese de Mestrado, Universidade não especificada.

AMOPC. 2024. *"Arquivo de Memória Oral das Profissões da Comunicação"*. Lisboa: Escola Superior de Comunicação Social, IPL. <https://www.amopc.org>. Acedido em abril de 2025.

Barbosa, Paulo. 2025. *"Arquivos e Museus que têm registos de memória oral em Portugal"*. <https://sites.google.com/escs.ipl.pt/amopc-pesquisa?usp=sharing>. Acedido em abril de 2025.

Bardin, Laurence. 2011. *"Análise de Conteúdo"*. Lisboa: Edições 70.

Benardou, Agiatis, Erik Champion e Costis Dallas. 2010. *"Digital Cultural Heritage: Critical Concepts in Media and Cultural Studies"*. Cambridge, MA: MIT Press.

Blackmagic Design. 2023. *"DaVinci Resolve 18: AI-Powered Audio Transcription White Paper"*. Port Melbourne: Blackmagic Design.

Cardina, Miguel. s.d. *"História Oral - Caminhos, Problemas e Potencialidades"* [Excerto].

Estanque, Elísio e José Manuel Mendes. 1997. *"Classes e desigualdades sociais em Portugal: um estudo comparativo"*. Porto: Edições Afrontamento.

Evans, George Ewart. 1956. *"Ask the fellows who cut the hay"*. Londres: Faber and Faber.

Finley, Moses I. 1986. *"Ancient History: Evidence and Models"*. Londres: Chatto & Windus.

Freire, João. 2006. *"Sociologia do Trabalho: Uma Introdução"*. Porto: Edições Afrontamento.

Frisch, Michael. 1990. *"A Shared Authority: Essays on the Craft and Meaning of Oral and Public History"*. Albany: State University of New York Press.

Godinho, Paula. 2001. *"Memórias da resistência rural no sul: Couço (1958-1962)"*. Oeiras: Celta Editora.

Gossman, Lionel. 2009. *"Between History and Literature"*. Cambridge, MA: Harvard University Press.

Grele, Ronald J. 1991. *"Envelopes of Sound: The Art of Oral History"*. Nova Iorque: Praeger.

Hedstrom, Margaret. 2010. "Archives, Memory, and Interfaces with the Past." *Archival Science* 10 (3): 321–333.

Huvila, Isto. 2008. "Participatory Archive: Towards Decentralized Curation, Radical User Orientation, and Broader Contextualization of Records Management." *Archival Science* 8 (1): 15-36.

IASA (International Association of Sound and Audiovisual Archives). Technical Committee. 2009. *"Guidelines on the Production and Preservation of Digital Audio Objects"*. 2ª ed. Auckland: IASA.

Leal, João. 2000. *"Etnografias Portuguesas (1870-1970): Cultura Popular e Identidade Nacional"*. Lisboa: Dom Quixote.

Lloyd, Geoffrey E. R. 1979. *"Magic, Reason and Experience: Studies in the Origin and Development of Greek Science"*. Cambridge: Cambridge University Press.

Mateus, Dalila Cabrita. 1999. *"Aluta pela independência: A formação das elites fundadoras da FRELIMO, MPLA e PAIGC"*. Lisboa: Editorial Inquérito.

Mateus, Dalila Cabrita. 2004. *"A PIDE/DGS na guerra colonial: 1961-1974"*. Lisboa: Terramar.

Momigliano, Arnaldo. 1990. *"The Classical Foundations of Modern Historiography"*. Berkeley: University of California Press.

Oliveira, Luísa Tiago de. 2004. *"Estudantes e povo na revolução: O Serviço Cívico Estudantil (1974-1977)"*. Oeiras: Celta Editora.

Oliveira, Luísa Tiago de. 2010. *"A história oral em Portugal"*. *SOCIOLOGIA, PROBLEMAS E PRÁTICAS* 63: 139-156.

Patton, Michael Quinn. 2002. *"Qualitative Research and Evaluation Methods"*. 3ª ed. Thousand Oaks, CA: Sage Publications.

Perks, Robert e Alistair Thomson, org. 2016. *"The Oral History Reader"*. 3ª ed. Londres: Routledge.

Pimentel, Irene Flunser. 2015. *"A História da PIDE"*. Lisboa: Temas e Debates.

Pimentel, Irene Flunser. 2015. *"História da oposição à ditadura portuguesa: 1926-1974"*. Lisboa: Figueirinhas.

Pugh, Mary Jo. 2005. **Providing Reference Services for Archives and Manuscripts**. Chicago: Society of American Archivists.

Ritchie, Donald A. 2015. **Doing Oral History**. 3ª ed. Oxford: Oxford University Press.

Rumsey, Abby Smith. 2016. **Digital Preservation and Access to Cultural Heritage**. Londres: Routledge.

Sandell, Richard. 2007. **Museums, Prejudice and the Reframing of Difference**. Londres: Routledge.

Subtil, Filipa e José Luís Garcia. 2017. "Dinâmicas comunicacionais e transformações sociais: proposta de criação de um arquivo de memória oral dos profissionais de comunicação." Comunicação apresentada no X Congresso Português de Sociologia, Covilhã, 10-12 julho.

Thompson, Paul. 2000. **The Voice of the Past: Oral History**. 3ª ed. Oxford: Oxford University Press.