

Apresentação do Arquivo Digital



Comissão de Planeamento e Informática
I Colóquio Guerra da Cal / III Seminário de Lexicologia
Compostela, Galiza, 12 de outubro de 2011

The screenshot shows the homepage of the digital archive. At the top, there is a navigation bar with 'Entrar' and 'Página Inicial' followed by a dropdown menu showing 'Lista de Categorias'. The main header features the title 'Arquivo Digital' in large blue letters and the logo of the Academia Galega da Língua Portuguesa. Below the header, the main content area is divided into several sections: 'Arquivo Digital da Academia Galega da Língua Portuguesa' with a brief description and a PDF download link; 'Categorias no repositório' with a list of six categories; 'Todo o repositório' with a search input field and a 'Ir' button; 'Visualizar' with a list of filters; and 'Minha conta' with 'Entrar' and 'Cadastrar' links.

<http://arquivo.academiagalega.org>

Introdução

O **Arquivo Digital da Academia Galega da Língua Portuguesa** (em diante, o Arquivo ou AD ou ADAGLP) é um repositório de materiais digitalizados acerca da língua portuguesa na Galiza e da cultura galega em geral. O Arquivo contém materiais de texto, áudio, vídeo e imagem em formatos padronizados comuns. O material está protegido pelas licenças *Creative Commons*, permitindo-se o seu livre acesso, utilização, distribuição e cópia sem alterações e com atribuição da fonte original. Nesta altura trabalham no projeto três membros da Academia Galega da Língua Portuguesa, um colaborador da Associação Pró-AGLP, e um técnico em informática.

O AD está albergado num servidor da Internet dedicado, a cargo da AGLP e organizado com a ferramenta *DSpace* de catalogação e acesso. Encontra-se no URL

<http://arquivo.academiagalega.org>

O Arquivo utiliza o extenso protocolo *Dublin Core* de metadados para a catalogação das características contextuais e de formato dos materiais: título, data, autor, género, extensão ou duração, etc. etc. Na altura, consta de mais de 50 Conjuntos Documentais textuais ou audiovisuais, que compreendem a uns 200 itens ou documentos de livre acesso.

Prevê-se a incorporação ao Arquivo de um grande número de materiais, nesta altura ora disponíveis em diversos lugares da Internet (particularmente de membros da AGLP ou de algumas associações), ora arquivados ainda privadamente. A parte mais cumprida do Arquivo neste seu lançamento é o material audiovisual de eventos comunicativos públicos de diversos géneros (palestras ou relatórios, mesas-redondas, entrevistas, debates, lançamentos editoriais, recitais) em torno da situação da língua da Galiza, na sua relação com o resto da Lusofonia e no seu conflito sociolinguístico atual com o espanhol, ou em torno da cultura, literatura ou política. Estão também disponíveis todos os números publicados do *Boletim da Academia Galega da Língua Portuguesa* e o seu Anexo 1. O Arquivo assegurará em todos os casos o carácter público e isento de restrições de distribuição dos materiais. Com a sua apresentação organizada, pretende-se oferecer ao público, aos pesquisadores e às instituições dados de primeira mão para o conhecimento não só da situação social da língua portuguesa na Galiza mas também dos aspetos formais e estruturais das suas diversas realizações orais e/ou escritas.

Estrutura do Arquivo

Os materiais do Arquivo Digital estão organizados hierarquicamente. Na sua estrutura inicial, a página de entrada do Arquivo apresenta as **Categorias** básicas dos materiais (chamadas “Comunidades” pela plataforma *DSpace*) e um motor de Busca. O motor permite pesquisas gerais ou avançadas por qualquer campo dos metadados.

A descrição das Categorias básicas é a seguinte:

1. **Publicações da Academia.** Monografias ou publicações independentes, incluindo aquelas da própria Academia Galega da Língua Portuguesa, tais como o seu *Boletim*.
2. **Trabalhos académicos.** Escritos de carácter técnico-científico, em revistas ou outras publicações especializadas.
3. **Outros textos.** Textos jornalísticos e doutra natureza sobre a língua ou sobre o debate científico, em revistas ou publicações gerais.
4. **Eventos comunicativos.** Palestras, debates, recitais literários, lançamentos editoriais, etc. de âmbito público, recolhidos *in praesentia* ou dos meios de comunicação.
5. **Informativos.** Materiais audiovisuais de natureza informativa (documentários, reportagens, etc.) elaborados profissionalmente com fins comerciais, dos quais se tenham assegurada permissão para a sua distribuição. Inclui a subcategoria **Programas de rádio.**
6. **Vários.**

Cada categoria pode conter **subcategorias**, chamadas pela plataforma DSpace “subcomunidades”. Em Eventos Comunicativos, por exemplo, a Comissão de trabalho do Arquivo Digital estão amplamente implementadas as subcategorias **Conferências e debates, Entrevistas, Apresentações e Recitais.**

Por sua vez, as subcategorias podem conter outras subcategorias, de maneira recorrente, ou podem ligar diretamente para os **Conjuntos Documentais** de materiais relacionados.

Cada Conjunto Documental contém fichas individuais de cada arquivo para a pesquisa. Para uma conferência, por exemplo, podem incluir-se materiais disponíveis tais como a gravação de áudio e/ou de vídeo com áudio, um cartaz de anúncio, uma ou várias fotografias, e o resumo ou o texto completo da palestra.

Os formatos dos documentos são os de uso mais comum, procurando-se a máxima qualidade possível e a máxima fidelidade possível aos originais. Por exemplo, as gravações de áudio estão codificadas em MP3 de alta qualidade (320 Kbps estéreo, ou 160 Kbps mono). Para os textos escolheu-se exclusivamente o formato PDF, com uma resolução mínima de 300 pontos por polegada, com segurança para permitir a visualização, cópia e impressão do documento.

Tanto cada Conjunto Documental quanto cada documento que os compõe são catalogados conforme uma notação que indica a data completa e outras informações.

A página do ficheiro de uma gravação, por exemplo, apresenta o registo básico, com informações de algumas das suas propriedades contextuais e de formato. A opção “Apresentar o registo completo” oferece informação mais detalhada conforme os metadados utilizados.

Os metadados são configuráveis e adaptáveis às necessidades do repositório. A Comissão de trabalho estabeleceu o conjunto de metadados adequados para os materiais de que dispõe a Academia.

Cada ficheiro de áudio, vídeo, texto ou imagem pode ser descarregado para o computador do utilizador. A equipa de trabalho está a estudar as possibilidades técnicas para a plataforma DSpace permitir também a reprodução do documento em tempo real a meio de uma interface.

Além da descrição em metadados, cada ficheiro contém também informações inseridas na forma de etiquetas ou comentários incrustados, acerca de algumas das suas propriedades contextuais (participantes, lugar, data, etc.), da proveniência (“Arquivo Digital da Academia Galega da Língua Portuguesa”) e da licença *Creative Commons* para a sua utilização e distribuição.

A equipa de trabalho do Arquivo tem o máximo interesse por esta contextualização de cada documento, de modo a um utilizador poder traçar a sua origem e características em qualquer momento da sua circulação previsivelmente acelerada a meio das novas tecnologias. Este aspeto da digitalização e preservação dos documentos, infelizmente ausente em muitos repositórios e bibliotecas digitais, pretende facilitar a sua organização posterior pelos utilizadores.

Conclusão: Prazos e previsões

O Arquivo Digital projeta-se como um instrumento com continuidade e com longo e estável percurso. Por sinal, na altura dispõe-se já do registo de mais de 300 eventos comunicativos das últimas décadas para serem digitalizados, catalogados e disponibilizados publicamente, além dos que continuarão a ser recolhidos e dos numerosos documentos textuais por organizar. Como ficou descrito, a finalidade do repositório não é a de um simples armazém de dados sobre a língua e a cultura, mas a de uma ferramenta sistematizada com o máximo rigor possível. Deseja-se assim maximizar a sua utilidade pública para a sociedade galega, para as culturas lusófonas e para a comunidade internacional em geral, objetivo último da constituição e do trabalho da própria Academia Galega da Língua Portuguesa.

O projeto de **Arquivo Digital da Academia Galega da Língua Portuguesa** é uma tarefa laboriosa, que requer notáveis recursos em termos de esforço e de tempo, e que precisaria da máxima colaboração pelo coletivo da AGLP, além de apoio profissional dedicado. A Comissão do AD está a verificar a solidez das fases iniciais, para garantir a sua consolidação sem mais problemas ulteriores dos que puderem surgir e surgem logicamente em empresas desta natureza. Um prazo realista para a disponibilização do grosso do material recolhido é dez anos.