

Biblioteca digital de peças teatrais

Angela Maria Silva

Especialista em biblioteconomia.
E-mail: angelams@dirbi.ufu.br

Ilmério Reis Silva

Doutor em ciência da computação (2000) pela Universidade Federal de Minas Gerais
E-mail: ilmerio@facom.ufu.br

Luiz Humberto Martins Arantes

Especialista em filosofia
E-mail: lharantes@yahoo.com.br

Resumo

Biblioteca Digital de Peças Teatrais - BDTeatro é um projeto para preservação, formatação, armazenamento e disseminação de Peças Teatrais. Este projeto foi desenvolvido pela Faculdade de Computação, Faculdade de Artes, Filosofia e Ciências Sociais e Biblioteca Central da Universidade Federal de Uberlândia - UFU. A Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) forneceu parte do suporte financeiro para o PROJETO. BDTeatro foi proposto para resolver os problemas de preservação e disseminação de uma coleção com aproximadamente 800 peças teatrais em papel. Esta coleção apresenta sinais de deterioração e sua digitalização foi a motivação deste projeto. Adicionalmente, a ausência de um catálogo on-line e de um processo automatizado para pesquisa de informações na coleção também motivou o projeto. Este artigo descreve o desenvolvimento da Biblioteca Digital de Peças Teatrais, onde foi usada uma perspectiva interdisciplinar de trabalho, garantindo desta maneira, padrões de qualidade no processo de construção do catálogo, na entrada dos dados e nos mecanismos de pesquisa de informações na coleção.

Palavras-chave

Indexação automática; sistemas de recuperação da informação – Teatro; Teatro – Indexação e resumos; Teatro – pesquisa; Teatro – Pesquisa – Indexação e resumos; Bibliotecas digitais.

Digital library of theater plays

Abstract

Digital Library of Theater Plays - BDTeatro is a project for preservation, formatting, storage, and dissemination of theater plays. This project was developed by the School of Computation, School of Philosophy, Arts, and Social Science, and Central Library of Federal University of Uberlândia(UFU). The Foundation to Research Support of the Minas Gerais State (FAPEMIG) provided partial financial support for the project. BDTeatro was proposed to deal with problems of preservation and dissemination of a collection with approximately eight hundred paper-based theater plays. This collection presents signs of deterioration and its digitalization has motivated this project. Besides, the absence of an on-line catalogue and an automated process to search information in the collection has also motivated the project. This paper describes the development of the Digital Library of Theater Plays where an interdisciplinary perspective of work was used, assuring this way standards of quality in the process of building up the catalog, in text data entrance and in mechanisms of searching by information in the collection.

Keywords

Automatic Indexing; Information retrieval systems - theater; Theater – Indexing and abstracts; Theater – search; Theater – Search – Indexing and abstracts; Digital Libraries.

INTRODUÇÃO

O século XX preparou o terreno para que o presente milênio se consolide como a era digital. Para alguns, a relação com as tecnologias trará perdas no tocante às relações humanas, tornando-as artificiais e instituindo indivíduos pouco sociáveis, mas, por outra perspectiva, acredita-se que a vida digital potencializará a comunicação humana e, por conseqüência, não acarretará o desaparecimento de antigos suportes comunicativos. Desse modo, ao contrário, o uso cada vez maior de tecnologias no dia-a-dia tem também ajudado a constituir novas bases para a difusão do conhecimento e da cultura (Levy, 1996).

Nas artes, campo por excelência da experimentação, as novidades sempre foram criadas e recriadas. No teatro, por exemplo, desde a chegada da luz elétrica, os mecanismos tecnológicos têm sendo bem recebidos, tornaram-se um meio importante, mas não o fim último para a especificidade da arte teatral.

A difusão de tecnologias que acelerou a imprensa escrita também muito contribuiu para maior divulgação de textos teatrais e na ativação de novos leitores. Os espectadores da cena passaram a ser também apreciadores de peças, incentivando maior circulação dos clássicos da dramaturgia, os quais, assim como os romances, ganharam espaços em bibliotecas domésticas e públicas.

No Brasil, muitos desses acervos estão sendo disponibilizados por meio da rede internet, proporcionando maior visibilidade às coleções. Entre as instituições que já disponibilizam catálogos *on-line* com acervos específicos na área de teatro, podemos citar o Centro de Documentação e Informação em Arte, da Funarte

< <http://www.funarte.gov.br>>; O Centro Cultural de São Paulo

< <http://sampa3.prodam.sp.gov.br/ccsp/ccsp>>; a Escola de Comunicação e Artes da USP

< <http://www.rebeca.eca.usp.br>>; a Sociedade Brasileira de Autores (SBAT)

<<http://www.sbat.com.br>>, além de outras instituições que disponibilizam o acesso a peças teatrais, nos catálogos globais das bibliotecas, a exemplo da Unirio <www.unirio.br> e UERJ <www.uerj.br> .

O advento dos meios tecnológicos, posteriormente à Internet, tem provocado o surgimento de novos tipos de leitores. Em apenas alguns cliques, uma criança ou um adulto acessa inúmeras páginas e conteúdos. Em apenas um Cd-Rom já é possível encontrar, por baixo custo, clássicos da literatura brasileira e, inclusive, peças teatrais.

Muitas iniciativas importantes têm sido desenvolvidas tanto no sentido de reunir quanto de disponibilizar conteúdos digitais de acesso público, a exemplo do Programa Prossiga <www.prossiga.cnpq.br>, que disponibiliza bibliotecas virtuais temáticas, como a “Biblioteca Virtual de Literatura”, na qual é disponibilizado um *link* na área de artes cênicas, com conteúdos nacionais e internacionais na área. Também merecem destaque a Biblioteca Virtual do Estudante de língua Portuguesa <<http://www.bibvirt.futuro.usp.br/index.html?principal.html&2>> e a Biblioteca Nacional <www.bn.br>, que disponibilizam o acesso a textos integrais de peças teatrais de domínio público.

Para os pesquisadores da cultura brasileira, é um sinal de que seus objetos de pesquisa, em breve, estarão mais acessíveis do que nunca, principalmente quando os documentos dos principais acervos estiverem digitalizados e acessíveis ao público em geral.

A perspectiva das novas tecnologias de informação, aliada à atuação das bibliotecas, abre novas possibilidades de acesso a esses conteúdos, por meio da construção de bibliotecas digitais. De acordo com Baeza-Yates (1999), “Biblioteca digital tem sido um tema cada vez mais discutido nas literaturas de ciência da informação e de informática. Ao mesmo tempo, dezenas de projetos de implementação estão em andamento em diversos países, notadamente nos Estados Unidos, Reino Unido e Austrália”.

Os documentos disponibilizados via Internet constituem uma grande biblioteca digital de conteúdo genérico. O mecanismo mais utilizado pelos usuários da Internet para encontrar documentos de seu interesse são máquinas de busca. Mas, em geral, as máquinas de busca não conseguem satisfazer as necessidades do usuário. Um dos motivos é a diversidade de temas e perfis de usuários. Uma tendência para disponibilizar dados na internet são Bibliotecas Digitais de conteúdo específico específico (Ribeiro-Neto; Assunção, 2001; Vale *et alii*, 2001), que possibilitam maior precisão nas pesquisas.

Com base em contextos relacionados a assuntos, é possível desenvolver mecanismos especializados que facilitem ao usuário encontrar e visualizar informações de seu interesse.

A recuperação de informação em uma biblioteca digital é feita por meio de um sistema de recuperação de informação (SRI) que busca informações contidas em documentos da biblioteca digital que possam ser úteis para o usuário.

O desenvolvimento de um SRI especializado em peças teatrais, descrito neste artigo, possibilitará um avanço tecnológico na área de recuperação de informação na área de teatro, a partir da criação da Biblioteca Digital de Peças Teatrais (Bdteatro) do acervo de textos teatrais Sandro Polloni, projeto desenvolvido em parceria entre a Faculdade de Computação, a Faculdade de Artes, Filosofia e Ciências Sociais e a Biblioteca Central.

O Banco de Textos Teatrais Sandro Polloni

O Banco de Textos Teatrais Sandro Polloni, hoje sob a guarda da Biblioteca da Universidade Federal de Uberlândia, guarda uma das relíquias da memória do teatro brasileiro: um acervo de peças de teatro datilografadas e mimeografadas que estão sofrendo a ação do tempo. Em seu conjunto, encontram-se obras que foram escritas em Minas Gerais e encenadas neste e em outros estados.

A coleção consta de aproximadamente 800 textos de peças teatrais e foi doada em 1993 pelos artistas Sandro Polloni e Maria D’ella Costa, ambos figuras de destaque na história teatral brasileira. A coleção possui importância ímpar na área de teatro, pois congrega um acervo de peças teatrais nacionais e algumas estrangeiras, além de traduções de obras da dramaturgia brasileira para outros idiomas, não disponíveis em outros acervos.

Os textos são muito consultados tanto pela comunidade interna da universidade, quanto pela comunidade artística de Uberlândia e região. Porém, o acesso à coleção não é eficiente, pelo fato de a maioria das obras apresentar problemas de preservação (rasuras, desgaste na impressão, folhas soltas), dificultando o acesso à leitura e oferecendo riscos de perda futura, além da falta de um processo automatizado de catalogação que possua uma ferramenta eficiente de busca das informações da coleção.

Essas dificuldades motivaram o projeto de criação da biblioteca digital, cujos principais objetivos foram possibilitar a preservação da coleção, para garantir a

continuidade do acesso e contribuir com o incremento da pesquisa, com o desenvolvimento de um sistema de qualidade na busca de informações e da disponibilização do acesso via Internet.

O SISTEMA DESENVOLVIDO PARA IMPLEMENTAÇÃO DA BIBLIOTECA DIGITAL DE PEÇAS TEATRAIS

O sistema desenvolvido para a BDTeatro consiste em um subsistema de digitalização das peças teatrais e um subsistema de busca das peças teatrais, descritos a seguir.

O subsistema de digitalização das peças teatrais

O subsistema de digitalização das peças teatrais consiste de um subsistema de digitação das peças e um subsistema de digitalização semi-automática, utilizando *scanner*, OCR e um formatador das peças.

O sistema de digitação é utilizado para entrada de dados catalográficos das peças teatrais e dos textos de peças cujo estado de conservação não possibilitam entrada via *scanner*. Os textos e os dados catalográficos são digitados e armazenados em formato XML. No sistema de digitalização semi-automática, os textos são obtidos automaticamente de um *scanner* e OCR e formatados de forma manual em XML. Todo o sistema de digitalização foi desenvolvido em Java (SUN, 2004). Os dados catalográficos e a estrutura dos arquivos XML gerados serão descritos a seguir, destacando a contribuição do curso de artes cênicas na concepção do formato do texto teatral, bem como do processo de catalogação em geral.

O estudo das estruturas que fazem funcionar o texto teatral

A participação do curso de artes cênicas neste processo tem contribuído para a mobilização de vários saberes, mais especificamente, aqueles referentes aos procedimentos e às estruturas que fazem funcionar o texto teatral, visando à definição do formulário de digitação das peças.

Sobre o cuidado que a obra teatral merece, o crítico teatral Décio de Almeida Prado apresenta uma observação lapidar para o pesquisador que deseja conhecer suas singularidades:

Como caracterizar, em teatro, a personagem? Os manuais de *playwriting* indicam três vias principais: o que a personagem revela sobre si mesma, o que faz, e o que os outros dizem a seu respeito (Prado, 1972).

Além das personagens, o olhar também deve ser lançado sobre a rubrica. Um exemplo disso pode ser observar a rubrica do dramaturgo. Dessa maneira, como Luiz Fernando Ramos, entender que a rubrica, e em termos gerais também o texto, são:

[...] o registro literário de uma certa poética cênica, o vestígio ou a marca de um método. O estilo de cada encenador e/ou, dramaturgo, quando exerce essa condição de montador de um espetáculo imaginário, estará estampado nas didascálias. Será lá, nesta cena desejada, que se encontrará a referência mais próxima, literária, do formato que já assumiu ou ainda assumirá a cena na leitura de seu autor. O método, ou a poética de encenação de cada um vai repercutir neste texto e ter sua expressão literária (Ramos, 1999).

Tudo isto para entender que o texto teatral não é um texto qualquer, que ele possui uma especificidade, portanto, uma estrutura e uma apresentação que o diferencia dos demais. Assim, foi necessário considerar tais características no momento de construção do formulário que iria receber a digitalização dos dados. A constatação resultou na necessidade de dissecar os textos teatrais do referido acervo e, assim, reconhecer as suas divisões em atos, cenas, partes, episódios, quadros, dentre outras possíveis.

A implementação do estudo

As divisões da peça teatral foram implementadas por meio de um documento XML que pode ser representado pela árvore mostrada na figura 1, a seguir. Observe na figura 1 que o nó “registro_marc” identifica a subárvore onde são registrados os dados de catalogação. Os dados são identificados por códigos específicos para campo e subcampo, por exemplo, o campo 100 identifica o autor. Este campo pode ser subdividido em outros subcampos, como o nome pessoal (subcampo a).

Na subárvore “texto-peça” são registrados os dados de estrutura e texto da peça. Destacam-se as divisões da peça em ato, episódio, parte, cena ou outro tipo de divisão. Essas divisões são então subdivididas em textos classificados em monólogo, diálogo, letra-música, poesia ou outro. Cabe destacar que em cada texto pode aparecer a rubrica, conforme descrição anterior. Esta é a estrutura da peça teatral na biblioteca digital. A digitação do texto das peças é feita pelo do programa digitar, desenvolvido durante o projeto, utilizando linguagem Java. A tela de entrada de dados é mostrada na figura 2, a seguir.

FIGURA 1
Estrutura hierárquica de uma peça teatral

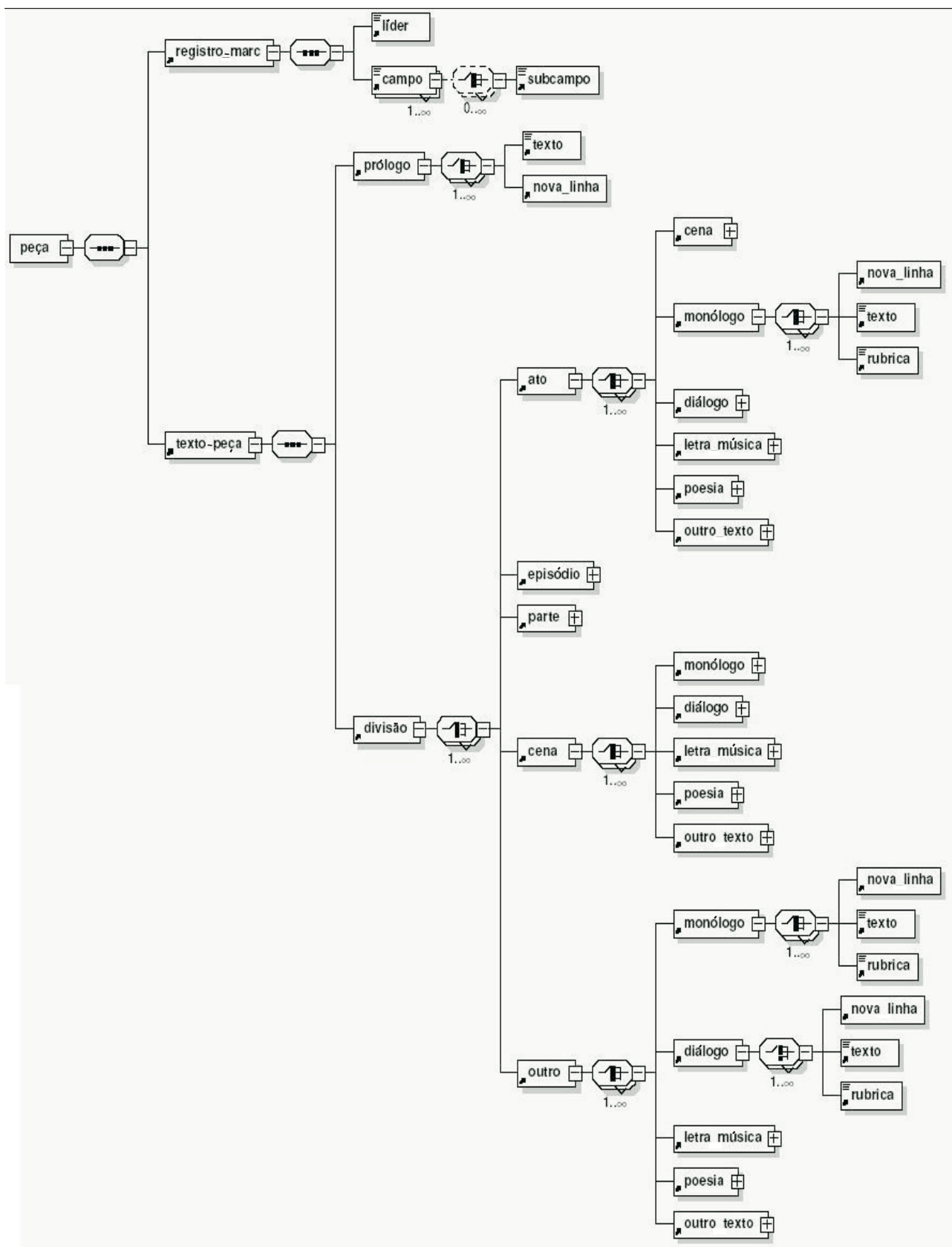
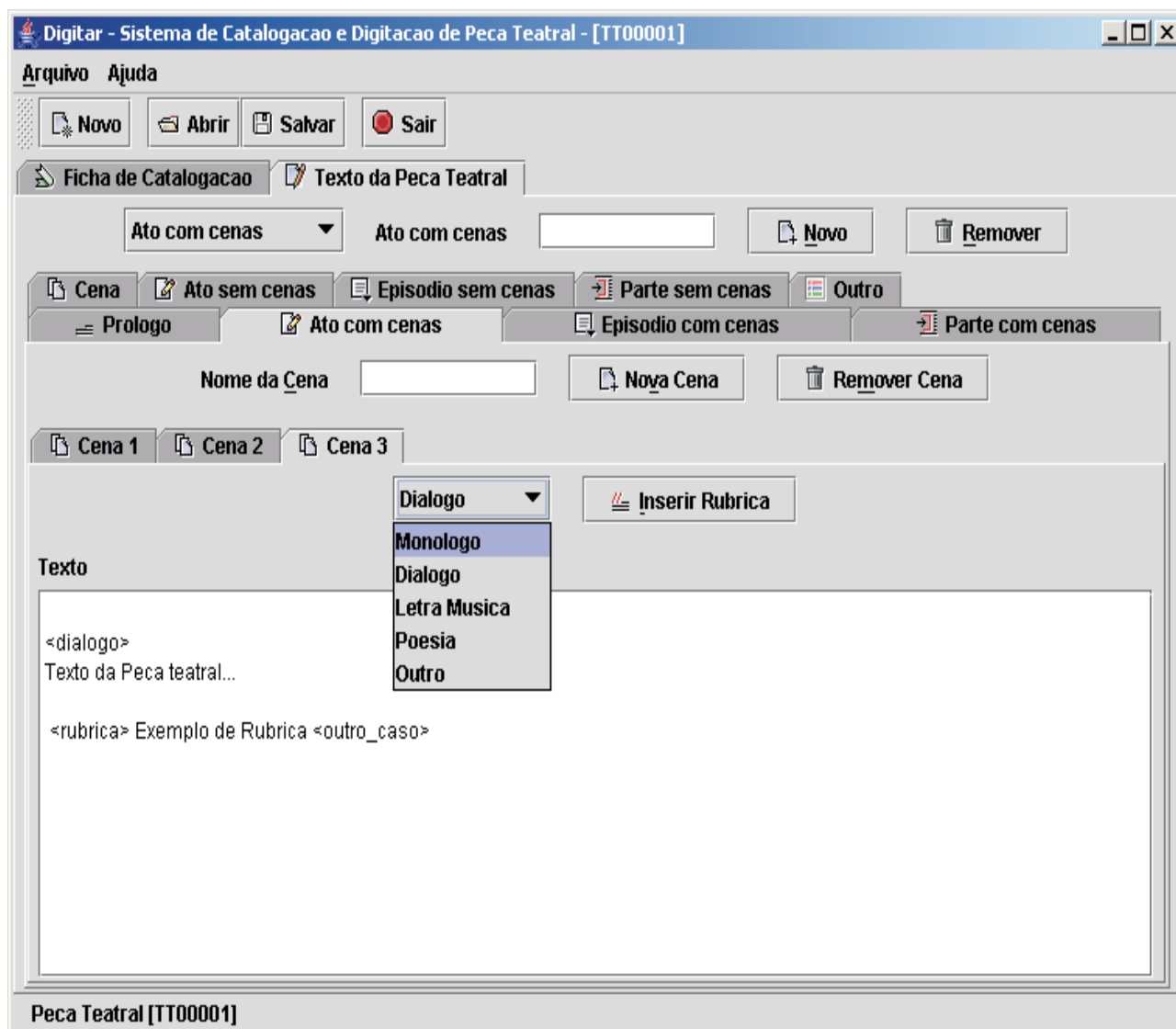


FIGURA 2
Tela para digitação das peças teatrais



A catalogação de peças

A catalogação requer o uso de padrões de qualidade, que refletem diretamente na qualidade da pesquisa e na comunicação científica, em que são considerados os padrões internacionais de qualidade. O sistema de catalogação para a biblioteca digital de peças teatrais compreende a definição dos metadados e a entrada de dados*.

* Os metadados são atributos de descrição dos documentos que permitem sua identificação sem ter que os executar; portanto, equivalentes aos atributos que descrevem as obras em uma biblioteca, porém realizados sempre de forma digital (Pavani, Ana. *Formação de diretores de projetos de teses e dissertações eletrônicas*. Curso ministrado em julho de 2003 na biblioteca universitária da UFMG.)

O subsistema para a entrada dos dados catalográficos foi criado em formato Marc 21*, que atende às necessidades de busca de informação nos textos teatrais. Foram inseridos apenas os parágrafos que dizem respeito ao texto teatral. A figura 3, a seguir, mostra a primeira tela de entrada dos dados de catalogação.

* O formato Machine Readable Cataloging (Marc) é um padrão para a representação e troca de informações bibliográficas, de forma legível por computador.

FIGURA 3
Tela de entrada de dados de catalogação

Como o resumo das peças é um dos campos a serem preenchidos, é necessária a leitura completa de peças. Esta atividade, competência de um dos bolsistas do projeto, passou a ser realizada também dentro de disciplinas ministradas no curso de artes cênicas.

Após a entrada dos dados catalográficos, a padronização dos assuntos e autores é feita por um bibliotecário, de acordo com o Código de Catalogação Anglo Americano - AACR2R (CÓDIGO, 1983-85) e as bases de dados de assuntos e autores mantidas pela biblioteca nacional <www.bn.br> e Library of Congress <www.Lcweb.loc.gov>.

O sistema de busca das peças teatrais

O Sistema de Busca das Peças Teatrais pode ser dividido em dois subsistemas: a busca utilizando dados catalográficos e a busca em texto completo.

A busca em dados catalográficos é realizada utilizando uma interface que coleta do usuário palavras-chaves associadas aos campos do registro Marc. Com base nas palavras informadas pelo usuário, o sistema gera uma consulta SQL, que é submetida ao gerenciador de banco de dados MySQL (MySQL AB., 2004). A resposta consiste de uma lista de peças teatrais que atenda à requisição do usuário. Veja na figura 4, a seguir, a interface de consulta e, na figura 5, a seguir, a lista de peças que atendem à consulta.

FIGURA 4
Interface de consulta aos dados catalográficos

FIGURA 5
Lista em resposta a uma consulta utilizando dados catalográficos

Foram encontrados **1** peças contendo **coca-cola** Resultados 1 até 1 de 1 peças.

1 / 1

Localização: tt00098

Autor: Cardoso, Dejair.

Título: O dia em que a vovó explodiu de tanto beber **coca-cola** / Dejair Cardoso. -

Publicação: 1982.

Descrição física: 54p.

Assunto: Tatro brasileiro - Séc. XX.

Gênero: Comédia.

Resumo: Marilene prepara o seu casamento, junto à sua mãe. Recebe a notícia de que seu noivo, seu pai e seu irmão perderam o emprego na fábrica. Eles voltam a trabalhar como motorista de caminhão, mecânico e dono de bar, respectivamente. Marilene é muito orgulhosa e não aceitando esta nova condição, faz um plano de seqüestro do filho do dono da fábrica. O plano corria bem até o momento em que a vovó que sempre bebia **coca-cola**, morre e estraga o plano. Por fim Marilene fica atormentada mentalmente.

Histórico: Doação de Antonio Mercado.

Personagens: 03 femininos e 04 masculinos.

Num. de atos: 2

Num. de cenas: 4

Ambiente: Dois cômodos de uma casa pobre.




FIGURA 6
Interface de consulta em texto completo

The image shows a web interface for a digital theater library. At the top, there is a blue banner with the logo of two theater masks and the text 'BIBLIOTECA DIGITAL DE PEÇA TEATRAL'. To the right of the banner are two buttons for language selection: 'Español' and 'English'. Below the banner is a navigation bar with four buttons: 'Home', 'Pesquisa', 'Pesquisa Avançada', and 'Fale Conosco'. The main content area features a search bar with the text 'Pesquisar por:' followed by an input field containing 'floresta', a dropdown menu set to 'Todos os campos', and a 'Pesquisar' button. To the right of the search bar is a link for 'Pesquisa avançada'. Below this is a section for 'Pesquisar no Texto Completo:' with an input field containing 'floresta' and a 'Pesquisar' button. At the bottom of the page, there is a footer with navigation links: '| Home | Pesquisa | Pesquisa Avançada | Fale Conosco | Estatísticas |', copyright information: 'Copyright © 2003 Biblioteca Digital de Peças Teatrais Todos os direitos reservados', and a credit: '[powered by Lázaro Nunes]'.

A busca em texto completo consiste na coleta de palavras-chaves informadas pelo usuário e na montagem de um *ranking* de peças teatrais que atendem totalmente (ou parcialmente) à consulta do usuário. Essa busca segue os modelos tradicionais de RI e deve ser realizada quando o usuário trabalhar com incertezas em relação à sua consulta, ou seja, quando o usuário não tiver muita informação sobre o conteúdo da biblioteca digital. As figuras 6 e 7, a seguir, mostram um exemplo de consulta e resposta baseadas em texto completo.

A implementação da busca em texto completo consiste na indexação de todas as palavras encontradas no texto da peça e estatísticas de ocorrência das palavras nas peças. No momento da consulta, o sistema calcula a similaridade entre as palavras da consulta e as peças, gerando um *ranking* em ordem decrescente de similaridade da consulta com a peça. A equação de similaridade usada na BDTeatro segue a fórmula do cosseno do modelo vetorial clássico (Baeza-Yates, 1999).

FIGURA 7

Lista resposta (*ranking*) a uma consulta em texto completo

Foram encontrados 7 peças contendo floresta	Resultados 1 até 7 de 7 peças.
--	--------------------------------

1 / 7

Localização: tt00044**Autor:** Arco e Flexa, Jairo.**Título:** O país de Sir Ney / Jairo Arco e Flexa. -**Publicação:** [19--].**Descrição física:** 33p.**Assunto:** Teatro brasileiro - Séc. XX.**Gênero:** Comédia.

Resumo: A peça trata de uma história hilariante de um cavaleiro chamado Sir Ney, que foi coroado rei de um país que se chama Nova República. O então rei coroado tem a missão de salvar a pátria que se encontra em crise econômica, política, social e cultural. Retrata temas como as imposições do Fundo Monetário Internacional; o plano da reforma agrária; dívida externa, homossexualismo; desemprego; arbitrariedades governamentais. A peça faz alusão ao Brasil da época em questão, todas as suas dificuldades, e o inesquecível jeitinho brasileiro para contornar a situação.

Histórico: Doação de Sandro Polloni / Maria Della Costa em 27/02/1992.**Personagens:** 3**Num. de atos:** 1

2 / 7

Localização: tt00133**Personagens:****Num. de atos:**

CONSIDERAÇÕES FINAIS

A parceria entre a Faculdade de Computação, a Faculdade de Artes, Filosofia e Ciências Sociais e a Biblioteca Central para o desenvolvimento de uma biblioteca digital de peças teatrais tem sido uma experiência ímpar e muito enriquecedora para todos os participantes do projeto, promovendo um modelo de integração entre áreas distintas, gerando um produto de qualidade e abrindo perspectiva de pesquisa multidisciplinar.

O sistema está desenvolvido, encontra-se em fase de digitalização das peças, com expectativa de disponibilização da biblioteca digital ainda em 2004, na *home page* <<http://www.bdteatro.ufu.br>>. Será disponibilizada a pesquisa nos dados bibliográficos e nos textos completos. A disponibilização dos textos será

apenas para as peças de acesso público, ou devidamente autorizadas pelos autores.

Cabe destacar que o projeto, além de cumprir o objetivo de preservação da coleção e proporcionar grande qualidade à pesquisa, é uma aplicação inovadora de técnicas de digitalização, formatação e armazenamento de textos teatrais semi-estruturados, por meio exclusivo do uso de *softwares* livres. Esta aplicação também amplia o número de conteúdos digitais na área de artes disponíveis no Brasil, promovendo o incremento da pesquisa nesta área, o que contribui para outras iniciativas de preservação e disponibilização digital de acervos culturais.

Artigo recebido em 01-06-2004 e aceito para publicação de 18 a 21/10/2004.

REFERÊNCIAS

- ARANTES, L. H. M. *Teatro da memória: história e ficção na dramaturgia de Jorge Andrade*. São Paulo : Annablume : Fapesp, 2001.
- BAEZA-YATES. R.; RIBEIRO-NETO, B. *Modern information retrieval*. Essex, England : Addison Wesley, 1999. 513 p.
- CÓDIGO de catalogação anglo americano. 2. ed. São Paulo : FEBAB, 1983-1985. 2 v.
- LEVY, P. *O que é virtual*. São Paulo : Editora 34, 1996.
- MySQL AB. *MySQL open source database*. Disponível em: <<http://www.mysql.com>>. Acesso em: jan. 2004.
- PRADO, D. de A. A personagem no teatro. In: PRADO, D. de A . et al. (Org.). *A personagem de ficção*. São Paulo : Perpectiva, 1972.
- RAMOS, L. F. *O parto de Godot e outras encenações imaginárias: rubrica como poética da cena*. São Paulo : Hucitec, 1999.
- RIBEIRO NETO, B.; ASSUNÇÃO, R. T. Recuperação de documentos jurídicos baseados em um tesouro. In: SIMPÓSIO BRASILEIRO DE BANCO DE DADOS, 16., 2001, Rio de Janeiro. *Anais...* Rio de Janeiro : [s. n.], 2001.
- SUN: the source of java developers. Disponível em: <<http://java.sun.com>>. Acesso em: jan. 2004.
- VALE, R. et al. Recuperação de informação em coleções médicas utilizando categorização automática de documentos. In: SIMPÓSIO BRASILEIRO DE BANCO DE DADOS, 16. 2001, Rio de Janeiro. *Anais...* Rio de Janeiro, [s. n.], 2001.