

PESQUISA/INICIAÇÃO CIENTÍFICA - PESQUISA/ INICIAÇÃO CIENTÍFICA

**DESENVOLVIMENTO DE UM BANCO DE DADOS EM XML DE ARTIGOS
CIENTIFICOS SOBRE O MULTILINGUISMO**

Milena Novais De Oliveira (milenanovaes567@gmail.com)

Jorge Luiz Nunes Dos Santos Junior (jorgesantosjunior@gmail.com)

Este trabalho foi desenvolvido no âmbito do Grupo de Estudos em Linguagem e Transculturalidade (GELT/CNPq), da Universidade Federal da Grande Dourados (UFGD) e situa-se no campo da Linguística Aplicada, do Multilinguismo e da Linguística Computacional. O objetivo da pesquisa foi desenvolver um banco de dados em Extensible Markup Language (XML) capaz de permitir o processamento digital de informações relacionadas à artigos científicos sobre o Multilinguismo, disponibilizados na plataforma Scielo. Inicialmente, criou-se um quadro com duas colunas para organizar o título e o DOI dos artigos. O critério para a seleção dos artigos científicos foi a presença das seguintes palavras-chave: multilinguismo, linguística, educação e ensino. Após essa etapa de coleta, que totalizou 30 textos, procedeu-se com a inserção das informações relacionadas aos dados e aos metadados de cada artigo em um arquivo XML, que foi estruturado de modo a permitir o processamento digital das informações. Desse modo, autor, nome do artigo, ano de publicação, palavras-chave, resumo, DOI e referência foram alocadas no interior de tags em XML. Posteriormente, a partir do uso do software BaseX, foi possível realizar dois tipos de buscas ao banco de dados, a saber: i) pesquisa simples, em que informações são recuperadas por meio de ferramentas prontas e disponíveis na interface do BaseX; ii) busca complexa,

em que linhas de código foram escritas no editor do BaseX com a finalidade de recuperar informações de maneira específica. Desse modo, foi possível executar buscas a partir de palavras presentes nos resumos e, ainda, filtrar informações por meio dos metadados de cada artigo como, por exemplo, autor, título, ano de publicação e palavras-chave. Conclui-se, por fim, que a estruturação de dados em XML possibilita o processamento de informações de maneira digital e semiautomática no campo da pesquisa científica, conferindo rapidez e exatidão na busca de informações mediante a seleção de filtros. Destaca-se, ainda, a perspectiva futura de integrar o banco de dados a uma aplicação web, que permitirá exibir resultados sobre o multilinguismo em um website dinâmico. Além disso, a quantidade de artigos catalogados pode ser ampliada continuamente, para que a base de dados possa abranger mais trabalhos relacionados ao multilinguismo, bem como suas interfaces com áreas correlatas. Gostaria de expressar meu sincero agradecimento ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro essencial à realização deste trabalho.

Palavras-chave: multilinguismo; pesquisa científica; basex; xml.