

289

ARQUIVAMENTO E GERENCIAMENTO DE VERSÕES DE DOCUMENTOS ELETRÔNICOS. *Andrea Raymundo Balle, Renata de Matos Galate (orient.)* (UFRGS).

Este trabalho está sendo desenvolvido como parte do projeto PERXML, que tem como objetivo o desenvolvimento de uma plataforma de armazenamento e a recuperação de dados históricos, para a especificação do perfil do usuário. Este gerenciamento servirá de base para o processo de adaptação de conteúdos de EAD e de portais Web baseada no histórico do perfil do usuário. O objetivo deste trabalho é desenvolver uma ferramenta para a recuperação de dados contidos no currículo Lattes do Cnpq. Os dados a serem extraídos são as informações pessoais, a área de pesquisa e as publicações dos usuários no contexto de ciência da computação. Quando o usuário faz seu cadastro, ele deve fornecer um arquivo XML de seu currículo Lattes. A partir deste, há a recuperação dos dados nele contidos com ajuda do parser XML DOM. Para classificar a área de pesquisa, inicialmente extrai-se a área contida no Currículo Lattes e as palavras-chave de cada artigo, utilizando o parser DOM. A partir dos títulos dos artigos publicados, extrai-se as palavras relevantes. Para isso, há uma pesquisa de palavras a serem ignoradas, a criação de uma stop-list e a programação de um método para a recuperação dos dados. Pode-se, classificar a área de pesquisa do usuário com essas palavras, cruzando-as com os parâmetros estabelecidos pela ACM que serão implementados como um dicionário. Todos os dados recuperados são armazenados no perfil do usuário. Além dos resultados obtidos para o projeto PERXML, o trabalho desenvolvido alimentará outros projetos do mesmo grupo de pesquisa. Os títulos dos artigos serão utilizados no projeto DIGITEX, para classificar o usuário como um pesquisador experiente ou não, o que determinará se seus artigos passarão por uma banca examinadora. A área de pesquisa será utilizada no projeto REC-Semântica, que recomenda para o usuário outros artigos de sua área. (PIBIC).