

292

RECUPERAÇÃO PELO CONTEÚDO DE DOCUMENTOS COM DIREITO AUTORAL.*Wanderley Augusto Radaelli Junior, José Valdeni de Lima (orient.) (UFRGS).*

Na área da Recuperação de Informação geralmente trabalha-se com documentos não estruturados. O conteúdo de um documento pode ser tratado sem estruturação para efeitos de indexação textual, mas na recuperação pode ser fundamental saber em qual parte da estrutura a informação se encontra. Adicionalmente, no armazenamento, indexação e recuperação de documentos também é importante respeitar exigências dos autores e/ou editores, tais como: recuperação de conteúdo sem permitir a impressão e/ou a publicação do mesmo, verificação de autenticidade e de integridade do texto, etc. Dentre os principais objetivos estabelecidos destaca-se a necessidade de, dado um arquivo qualquer definido segundo o padrão XML (Extensible Markup Language) processá-lo em busca de sua estrutura lógica, ou seja, marcadores, delimitadores, os quais definem capítulos e subseções de um livro digital, por exemplo. Após identificar a estrutura lógica, o documento é então indexado, tendo como objetivo criar um índice completo, o qual possa ser consultado pelo conteúdo contextualizado por sua estrutura, utilizando palavras-chave fornecidas pelo usuário da aplicação. Esta interação é feita com auxílio da interface gráfica oferecida pelo protótipo. Os métodos e materiais utilizados na elaboração do programa constituem-se na linguagem de programação Java, utilizada conjuntamente ao DOM Parser para documentos XML, o qual é um padrão W3C (World Wide Web Consortium) integrado a API (Application Programming Interface) de Java. Finalmente, é importante destacar que o protótipo implementado permite a recuperação pelo conteúdo dos livros digitais sem permitir a publicação e/ou a visualização do referido conteúdo. Atualmente, é empregado o algoritmo de criptografia AES (Advanced Encryption Standard), o qual é um padrão reconhecido, entretanto, futuramente pretende-se utilizar uma alternativa baseada no método das assinaturas. (PIBIC).