

Análise de monografias

Claudivan Ribeiro

4 de maio de 2011

1 Recuperação de Informações em Bancos de Dados Textuais

Por *Marcela Ortega Garcia, 2009, 8.0*

1.1 Resumo

Recuperação de informação em bancos de dados não estruturados é uma subárea da computação que trata da busca e ranqueamento de textos através de uma consulta em linguagem natural. Um exemplo típico é a procura de páginas na internet. Para isso utiliza-se um conjunto de índices, que são o subconjunto de palavras no texto que melhor lhe atribui significado. Algumas heurísticas são utilizadas para a construção destes: as preposições, os artigos e as conjunções assim como os números são desconsiderados pois são informações que em geral ocorrem em todos os documentos e, portanto, não o torna distinguível. Além disso, os verbos são armazenados no infinitivo (recuperar, ao invés de recuperado) e todas os termos são convertidos para o singular. Isso diminui a quantidade de índices em memória e aumenta a relevância dos resultados.

A principal técnica de indexação utilizada é chamada de arquivos invertidos e consiste em armazenar em memória RAM cada palavra e a lista de documentos em que ela aparece. Dentre as implementação de busca em dados textuais estão a biblioteca Apache Lucene e a biblioteca Ferret. Para o ranqueamento do resultado, destaca-se o PageRank do Google.

1.2 Parte Técnica

O texto está claro, foi bem escrito e ilustrado com figuras, diagramas e tabelas de modo que um ingressante em computação pode entendê-lo perfeitamente.

Os fundamentos teóricos e as alternativas tecnológicas foram bem explicadas na medida do possível para um assunto abrangente como este.

Apenas a motivação e os objetivos não ficaram explícitos. Também poderia ter realçado os resultados. A implementação foi realizada no Centro de Estudos do Genoma Humano da USP com a biblioteca *Ferret* mas não foi discutido se o resultado foi satisfatório ou insatisfatório. Os usuários encontravam apontamentos relevantes sobre pacientes para suas buscas? Precisaram de algum treinamento para utilizar a caixa de busca, com sua sintaxe particular? Mostrar a comparação entre os resultados de uma consulta estruturada e uma consulta baseada em índices, mostrando como isso afeta a relevância dos resultados poderia ter enriquecido o resultado.

1.3 Parte Subjetiva

A estudante demonstrou forte empenho para desenvolver o projeto. Ficou inteiramente responsável pelo sistema do CEGH onde fazia reuniões com os usuários para levantamento de suas necessidades. É provável que a busca não estruturada tenha surgido de relatos em alguma destas reuniões.

Também mostra a gratificação que é ter implementado uma funcionalidade tão útil naquele sistema de 15 mil usuários, na época.

A relação do aprendizado no curso com o trabalho desenvolvido foi relatado juntamente com a lista de disciplinas mais utilizadas.

Assim como a empolgação e a energia, a frustração também fica clara por não ter conseguido tempo suficiente para iniciar os estudos sobre buscas semânticas, que seria o passo seguinte do projeto e de um possível mestrado.

1.4 Críticas e Elogios

Em geral o texto está bem redigido, com seções bem definidas. O conteúdo foi bem apresentado e baseado em livros referências da área.

A orientação deve ter sido bem conduzida pelo especialista em Banco de Dados que é o Professor João Eduardo Ferreira.

Um acréscimo que enriqueceria ainda mais a monografia seria uma maior exploração dos resultados.

2 Construção de gadgets na plataforma EzWeb

Por *Rafael Sato Sugiura, 2009, 7.5*

2.1 Resumo

Ezweb é uma plataforma web escrita com o *framework* django em que pode-se publicar pequenos aplicativos provedores de serviços, os gadgets.

2.2 Parte Técnica

Está fácil entender como foi a junção de diversas tecnologias para montar um gadget pois cada uma delas (XML, CSS, JavaScript etc) recebeu uma breve explicação.

Porém, não está claro no início do texto qual será o objetivo, quantos gadgets criar e com quais funcionalidades. E depois de mostrada a implementação, o leitor espera que haja alguma relação entre os gadgets criados.

2.3 Parte Subjetiva

O estudante mostrou-se engajado, com boa visão para resolver problemas da época, como um gadget que provia informações da Bovespa.

A grande frustração foi de não poder publicar seu trabalho por problemas de incompatibilidade de softwares utilizados.

Realçou a contribuição das matérias do curso para o trabalho com surpreendente destaque para Conceitos de Linguagem de Programação, que o ajudou no entendimento de algumas características de JavaScript.

2.4 Críticas e Elogios

A monografia descreve pontos práticos no desenvolvimento do projeto. Talvez tenha cometido excesso neste quesito, carecendo o texto de mais teoria, de conceitos mais abstratos. Isso diminuiria o tom de tutorial da obra.

No mais, o texto está claro e o leitor tem a vantagem de poder reproduzir os passos na escrita do código.