

**Nome:** Fernando Paladini

**Matrícula:** 13200642

## **Leitura de “Introdução à RI e busca na Web”**

O texto para leitura (*Introdução à recuperação de informações e busca na Web*) é o capítulo 27 do livro de Elmasri/Navathe 6a edição. Ele está disponível na xerox do CTC e o livro todo na Biblioteca Central, tanto em cópias impressas, quanto em formato eletrônico, para leitura via Internet.

Após a leitura, responder as seguintes questões, em no máximo uma folha (4 mil caracteres) e fazer upload de suas respostas, em formato texto, nesta pagina.

1. Defina recuperação de informação.
2. Descreva sucintamente os modelos de espaço vetorial, probabilístico e semântico para recuperação de informação.
3. Defina cobertura e precisão em sistemas de recuperação de informação.

### **Respostas**

1) A recuperação de informação estuda e aplica técnicas de indexação, pesquisa e recuperação de dados desestruturados. Ela difere das consultas comuns de banco de dados relacionais pois estes trabalham com dados estruturados, de forma que uma linguagem de consulta bem definida e restrita pode ser criada. Já dados desestruturados não possuem tal facilidade, pois devido à sua diversidade de conteúdo e natureza, torna-se impossível representar os dados de forma estruturada.

2) Não consegui compreender bem os conceitos de espaço vetorial, probabilístico e semântico.

3) A precisão é uma definição bem simples que tem como objetivo avaliar a relevância de uma pesquisa. Ela é definida como o número de documentos relevantes recuperados por uma pesquisa dividido pelo número total de documentos recuperados por essa pesquisa. Não encontrei em nenhuma parte do capítulo o conceito de “cobertura”. Seria o conceito relacionado à “precisão”, chamado de “revocação”?