

Resenha SDA

Aluno: Rafael Tenfen	
DATA: 13/11/2020	
ARTIGO Titulo: Reprint of: The anatomy of a large-scale hypertextual web search engine	
Autores: Sergey Brin, Lawrence Page	

Tópico 1 - RESUMO

Um parágrafo curto (3-4 frases) resumindo o artigo e discutindo um ou mais dos seguintes pontos:

O que você achou interessante sobre o trabalho?
Qual é o princípio essencial que o artigo explora?

O google por muito tempo foi uma caixa preta, algo que desperta curiosidade sobre o seu funcionamento e arquitetura para realizar o feito de apresentar aproximadamente 222.000.000 de resultados em incríveis 0,59 segundos ao pesquisar "bill clinton". A essência do artigo é explicar o funcionamento e arquitetura necessária para chegar nesse resultado.

Tópico 2 - Qual é a principal contribuição do artigo?

Descrever o funcionamento e arquitetura do motor de busca, ranqueamento e indexação de páginas do Google.

Tópico 3 - Pontos positivos (forças)

- Especificações do ranqueamento de páginas, o que sempre foi uma dúvida para inúmeras pessoas.

- A importância de o link ser referenciado por outras páginas, com as definições nos textos âncoras para o ranqueamento de páginas.

- Possibilidade, mesmo que pequena de um link do Google levar a lugar nenhum, ou sejam, uma página inexistente, pois ele não realiza a verificação de todas as páginas que ele corresponde.

- Apresentação da arquitetura do Google, com explicações de sua indexação.

- Definições do crawler que possui cerca de 300 conexões abertas simultaneamente e que cada crawler mantém o seu próprio cache de DNS.

Tópico 4 - Pontos negativos (limitações) do artigo

- Definição de indexação de SPA (*Single Page Application*) para páginas em que suas página são renderizadas utilizando JavaScript.
- Especificar o funcionamento do programa DumpLexicon.
- Demonstrar a evolução do algoritmo e arquitetura do motor de busca com o passar do tempo.
- Apresentar o funcionamento do Google Acadêmico.
- Como é a idexação das páginas anunciadas pelo Google.

Tópico 5 - Referências complementares (lidas ou a serem lidas):

- Ron Weiss, Bienvenido Velez, Mark A. Sheldon, ChanathipManprempre, Peter Szilagyi, Andrzej Duda, David K. Gifford, HyPursuit: a hierarchical network search engine that exploits content-link hypertext clustering, in: Proceedings of the 7th ACM Conference on Hypertext, New York, 1996.
- Junghoo Cho, Hector Garcia-Molina, Lawrence Page, Efficient crawling through URL ordering, in: Seventh International WebConference (WWW 98), Brisbane, Australia, April 14–18, 1998.
- NORUZI, Alireza. Google Scholar: The new generation of citation indexes. Libri, v. 55, n. 4, p. 170-180, 2005.
- LEE, Micky. Google ads and the blindspot debate. Media, Culture & Society, v. 33, n. 3, p. 433-447, 2011.