Rapid Seek

Made by:

Abdul Aziz Zafar (469987)

Key Components

Lexicons

A comprehensive list of unique terms found in the indexed data.

Forward and Backward Indices

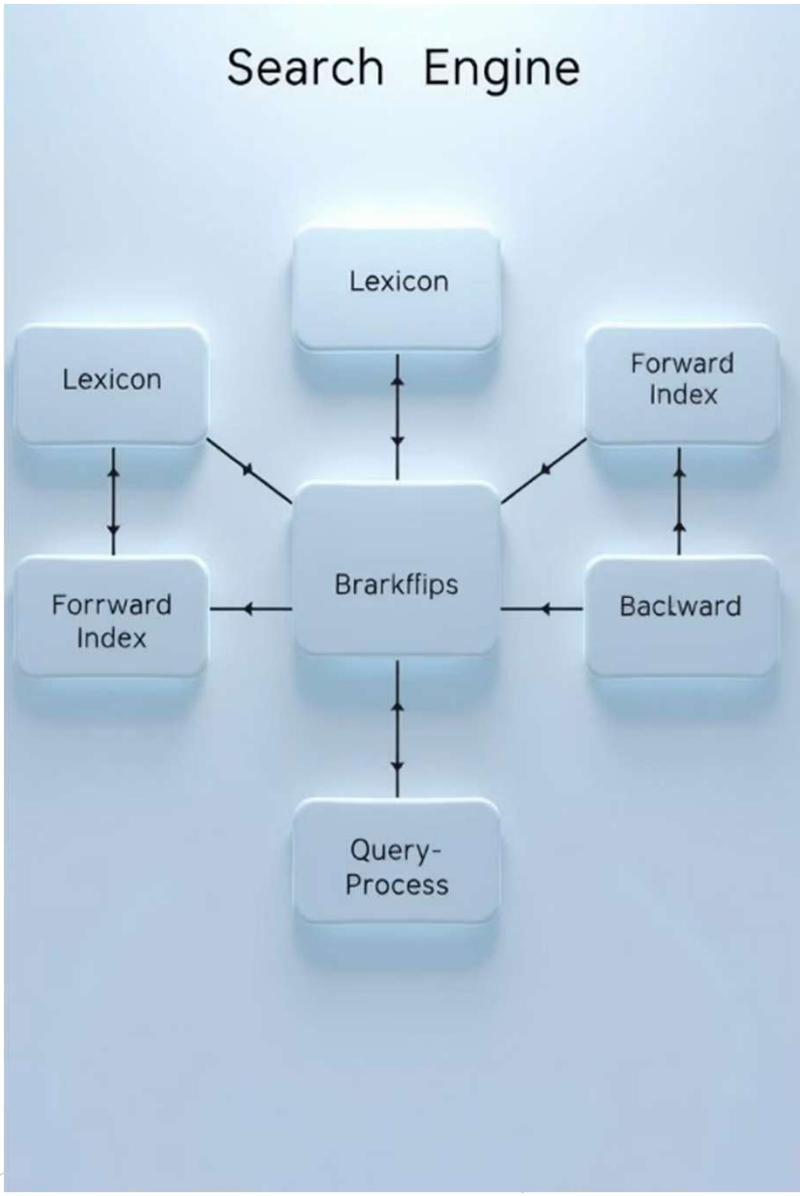
Data structures that link documents to terms and vice versa.

Query Processing

The process of analyzing and executing search queries to retrieve relevant results.

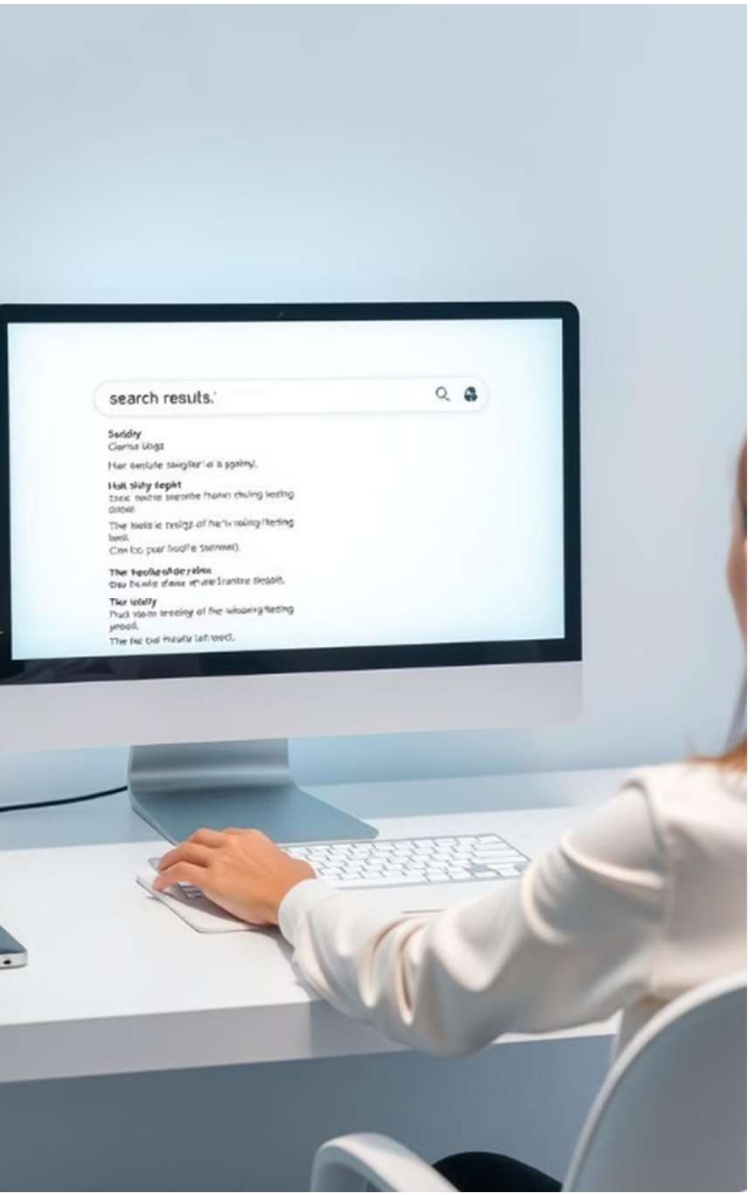
Barrels

Storage units for indexed documents and associated metadata.



Lexicon

- Collection of all unique terms present in the document corpus.
- Foundation of a search engine's indexing and retrieval capabilities.
- Provides a comprehensive record of all unique terms found in the indexed data.
- Helps in indexing and retrieving information quickly.
- Allows for efficient lookup of terms and their corresponding document references.

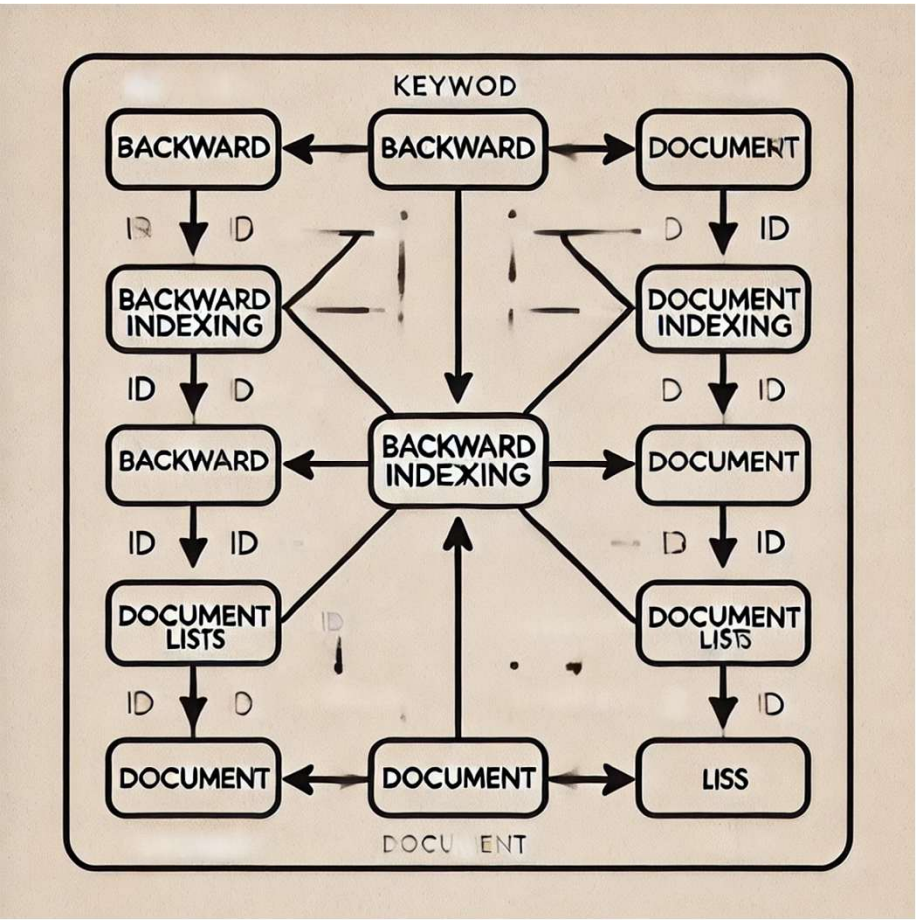


Forward Index

The forward index maps documents to the terms they contain, helping retrieve all documents containing a particular term.

- Maps each document to the terms it contains.
- Another crucial component of a search engine.
- Acts as a bridge between documents and the terms they contain.

Backward Indexing



The backward index (or inverted index) maps terms to documents, enabling efficient search for documents containing a specific term.

This provides efficient access to documents containing a specific term, allowing quick retrieval of all documents containing a given term.

Barrels

Definition

Barrels are sub-indexes used to partition the index for scalability and efficiency.

Purpose

They divide large datasets into smaller, manageable parts, enabling faster processing and retrieval.

Types

Barrels can be based on document IDs, term frequency, or other criteria, allowing for flexible partitioning strategies.

Query Processing

- **Query Parsing**

The query is broken down into terms and analyzed for structure.

- **Term Matching**

The query terms are looked up in the inverted index to find relevant documents.

- **Ranking**

Results are ordered based on relevance using factors like term frequency and document popularity.

Efficiency and Optimization

Speed

Indexing improves retrieval speed by providing efficient access to relevant documents.

Scalability

Barrels and indexing enable efficient processing of large datasets, scaling search engines to handle increasing data volumes.

Optimization

Techniques like indexing only relevant terms, compression, and parallel processing further optimize performance.

The End