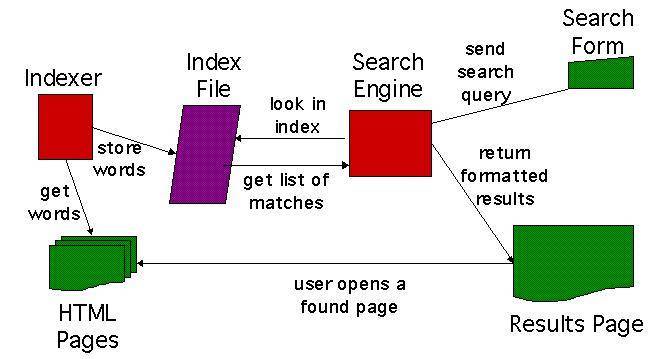
**ANATOMY AND WORKING OF SEARCH ENGINES**

Almost daily we use search engines and get the desired result but have you ever wondered what actually happens behind the scene? How is the data stored? Where is the data stored? Does it use file system/database servers/[cloud](https://krazytech.com/technical-papers/cloud-computing) for the storage? We will try to understand how a search engine works? Although it is a complex and tedious process, we are going to explain some of the important terminologies used in the search engines and the basic process involved.

Data is stored in a number of files on the internet, as ASCII files, binary files or in the databases. Search engines may vary on the way the data is stored. If the data is stored in the database the same can be queried easily to create search engines. For HTML files, graphics & PDFs search engine is an additional program.

[](https://krazytech.com/technical-papers/anatomy-and-working-of-search-engines/attachment/searching-meachanism)

A search engine that does not have a given content, searches it elsewhere. This data comes from a program that crawls many pages and reads the contents. Such a program is called a **ROBOT** or a **SPIDER**. It crawls the URLs specified by the search engine and marks when a new one is found. Google.com differentiates the pages that have been crawled and those that have not. The search engine crawls the data indexes the pages.  Whenever you search for a query, the pages that have been indexed are displayed on the results page. In the results, the search engine mostly displays the title of the related article and the excerpt.