Cheatography

elastic search Cheat Sheet

by hajuko via cheatography.com/20601/cs/3329/

Analyzer

A Combination of **Character Filters**, **Tokenizers**, **Token Filters**

Processing pipeline

Input -- *String*--> **Character Filters** -- *String*--> **Tokenizer** -- *Tokens*---> **Token Filters** -- *Tokens*--> **Output**

Character Filters

The Character filter removes words from the input Strings.

e.g. **html_strip char filter**:

Input: *" An example
 *

Output: *"An example*"

Tokenizer

The way the input is divided into parts (tokens)

e.g. **whitespace tokenizer**

Input: *The quick brown fox*

Tokens: [*The, quick, brown, fox*].

Token Filters

A Tokenfilter returns a subset of given tokens.

e.g. **stop word filter** that removes all the **stop words** from the tokens, such as *"the"*, *"and"*, ...

By **hajuko**
cheatography.com/hajuko/

Not published yet.
Last updated 10th May, 2016.

Sponsored by **Readable.com**
Measure your website readability!
https://readable.com