Elasticsearch provide different analyzer under string field, mostly deal with Natural Language Processing such as tokenization, removing stopwords, normalization of the data.

\_source: control the return output, include and exclude are the often use contribute

Multi Get: allow you to get document base on there index, type, id

Fuzzy Query:

Tried to match string as close as possible by using Damerau-Levenshtein distance formula(number of insertions, deletions, substitutions, and transpositions)

Fuzziness: represent the maximum edit distance

prefix\_length: The number of initial characters which will not be “fuzzified”. This helps to reduce the number of terms which must be examined. Defaults to 0

transpositions: action such as olivre in instead of delete r and add e behind the word which mean the edit distance is 2, it just change the position of r and e

max\_expansion: The maximum number of terms that the fuzzy query will expand to. Defaults to 50

When using fuzzy query knowing that the system use implement analyzer automatically there is a chance that the result after the search would not make sense base on the analyzer query used

Usually fuzzy search mix with match query, the exact fuzzy query usually used as term query(which mean find the exact term in certain field)

Best example of Fuzzy query could be use would be text correction, text suggestion

Using prefix\_length, max\_expansion would significantly increase performance

More like this Query:

Find documents that are similar to the text user provide, the mlt selects few representative terms of these input document and form the query using these terms. Using Lucene scoring formula

Like: require filed to put in the documents or text user tent to search for

Unlike: conjunction of like, its for user to exclude the term that doesn’t want to search for

Field: list of fields to fetch and search

Docs: use to search a list of documents

Max\_query\_term: maximum term that the query could select out of the text user provide

Min\_term\_freq: minimum term frequency the term will be ignore If it is less than the number user provide, default 2