

1.1

Strict format of url (according to rules)

More general urls (anything with http:// or for example c://)

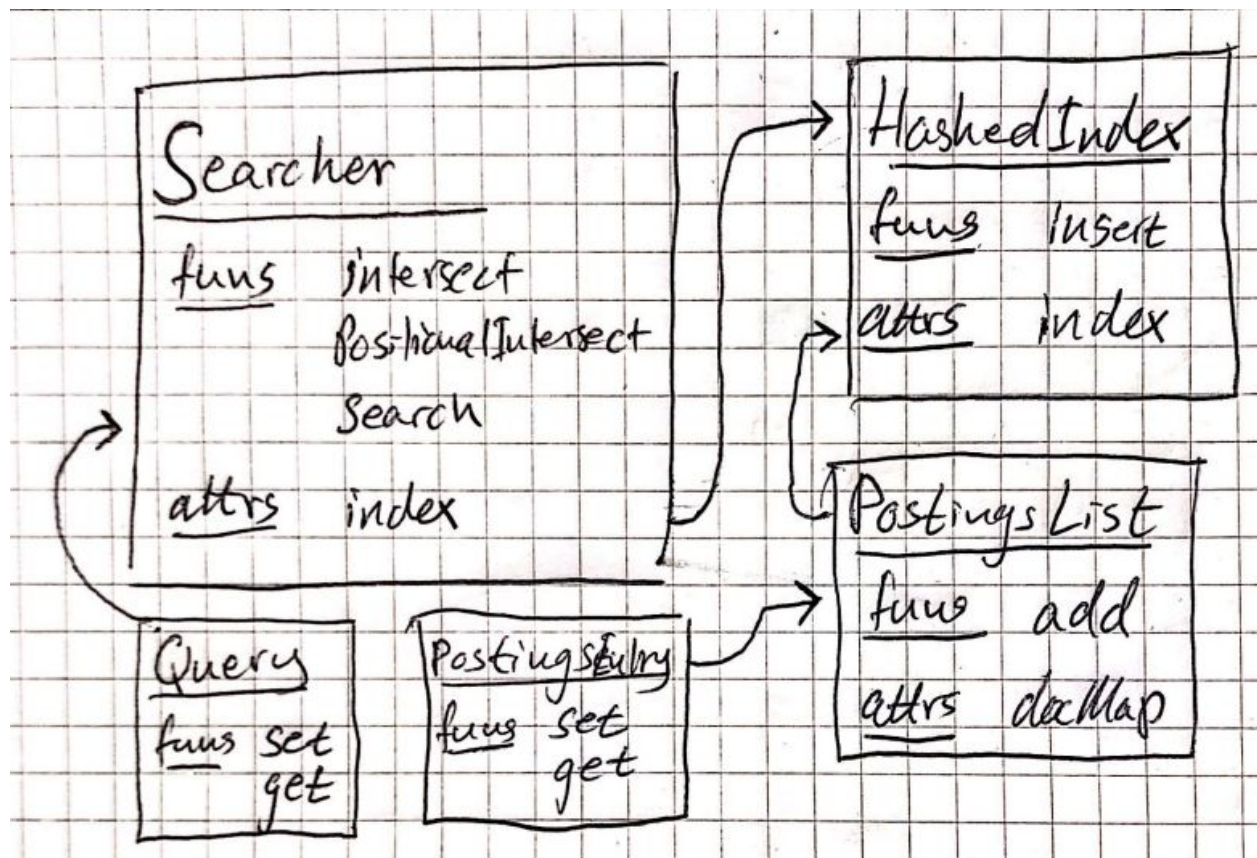
Cleanup (file:here, this:, alpha=)

1.2

No examination

1.3

Code changes and data structure:



1.4

Why are fewer documents generally returned in phrase query mode than in intersection query mode?

In a phrase query like “term1 term2 term3”, the searcher tries to find term1 AND term2 AND term3 in a document, while searching for an intersection query of the same time tries to find term1 OR term2 OR term3, which will clearly give more results.

1.5

Recall and precision

1.6

Why can we not simply set the query to be the entire information need description?

Different authors can phrase the same information differently, which is why one should start with a broad search and refine it iteratively based on the search results to find the information they are looking for.