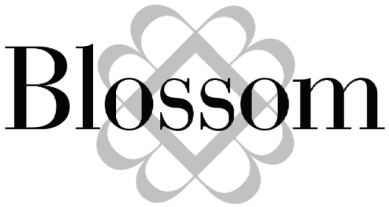

Fine tuning with page weights and meta-tags

Generating highly relevant results for a query is a hard problem. Blossom, like other modern search vendors, regularly tweaks its ranking algorithm to improve the results. For some queries, you may prefer specific pages appear near the top of the results regardless of how they are ranked by the search engine.

This note presents the primary ways you can influence a page's ranking. Basically, there are three things you can do:

1. Alter the text on the page.
2. Alter the meta-tags on the page.
3. Manually adjust the page weight.

We discuss each of these below.

Automatic ranking

The goal of automatic ranking is to match the words in a query against the words in each document. The better the match, the higher the score and the higher the page appears in the result list.

By default, Blossom Search uses proximity search, which means that, in order for a document to be considered relevant to a query, all of the words in the query must appear near one another in the document; the closer together the words appear, the higher the relevancy score. In addition, Blossom uses document structure like sentences and paragraphs to form boundaries for the terms in a query. For example, if the query terms appear close together in a document but span a paragraph boundary, the document will not be considered relevant.

If the query words appear in a document more than once, that also boosts the score. To keep long documents from always appearing at the top of the rankings simply because they have more words, document length also contributes to the score. The same is true for old documents, under the assumption that newer are more relevant than older.

Where in a document the query words appear is also important. The title of a document is usually a good indicator of its content, so matching words in a page title count the most.

A document can be given a description using the “Description” meta-tag. While not as important as a title, matching words that appear in the description are also counted heavily. Next in importance are headlines in the document, such as text enclosed within an H1 tag.

Keywords and keyphrases meta-tags

The rules for using meta-tags are certainly one of the least understood aspects of Web pages. While they were intended to make searching more effective, the major Web-wide search engines either ignore them or don’t explain how they are interpreted.[1] So, first we must say that Blossom *does* read and use meta-tags to provide an alternate way of influencing the relevance of a page for particular query words.

Very often there are words not in the page title that are descriptive of a page. Synonyms are a good example; there are usually several ways to say the same thing. Rather than putting all the synonyms into the title text, you can use the *keyphrases* meta-tag. Matches in keyphrases are treated just like those in page titles; they give a big boost the page. The only difference is that title matches are shown in the page snippet section of the search results, while keyphrase matches are not shown.

Analogously, matches in a *keywords* meta-tag are treated like matches in page descriptions. Again, while matches in descriptions appear in the search results, matches in keywords do not. In the case of both keywords and keyphrases, all of the words are grouped together without regard to the use of commas. For example,

```
<META NAME="Keywords" CONTENT="price schedule, list of fees">
```

will match queries like “price list” and “fee schedule” in addition to the more obvious “price schedule” and “list of fees”.

Manual page weights

There may be certain pages you always want to be rated highly if they match at all, for example, a FAQ page or a directory page might be your first choice if they match the query. Rather than creating long lists of keywords and keyphrases, you can instead just boost (or lower) the score for the page no matter what the query. Using the Blossom weight command you can set the starting score for a page to be something other than zero.

The weight command is placed inside a comment, for example placing this comment on a page sets the starting score to 30:

```
<!--Blossom:Weight=30-->
```

By default, pages start with a score of zero. Using a positive score pushes a page towards the top of the result list; a negative weight drops the page lower in the list.

[1] The opaque policy has arisen because, in the battle to appear higher in search results, website designers have used meta-tags in unscrupulous ways.