

Graphical Method for Facet Refinement in Search Applications

An interaction method is disclosed to enable users to generate a complex query against an indexed collection of data without knowledge of query syntax through direct keyword entry and spontaneous discovery of frequent and highly correlated data entities by means of an intuitive graphical interaction.

In search applications where text analytics are done, you will often see facet data displayed in a tree structure on the left hand side which allows the user to drill down and narrow their search results. This is often seen on shopping sites like Amazon.com, where you will search and see categories of products on the left along with counts. When one of these is clicked, the search is refined to include only those items which are in that category. This is a very convenient method for narrowing results and is common throughout the web. The drawback of this approach is that the user is limited to only selecting a single category (facet) at a time. If a user searches for 'coffee' on Amazon.com, they can click to only show results in the Gourmet Food category or in the Grocery category, but not both. Also, they have no way to indicate that they do not want to see results in the Books category. We needed to expand this concept to give the user greater flexibility and control while still retaining the intuitive refinement that this paradigm provides.

Our idea was to expand the concept of faceted search refinement to give the user the ability to add multiple facet terms to the query. The common implementation would effectively add a boolean AND clause to the existing query to refine the result set. Using the example above, if the query were 'coffee' and the user clicked 'Gourmet Food', the query would be something like 'coffee' AND category:"Gourmet Food". We wanted to give the user the ability to create more complicated queries in a graphical way. Using our new mechanism, the user could generate a query like 'coffee' AND category:("Gourmet Food" OR "Grocery") AND NOT category:("Books").

Another issue with the existing implementation is that once the user refines by a particular category, such as "Gourmet Food", they cannot then click "Grocery" to get the same behavior because the result set will have changed and there may no longer be any results which are in the group "Grocery". Our method allows the user the greatest degree of flexibility for building complex queries without needing to edit the query syntax.

The preferred embodiment lists the facets with counts next to 2 icons: a plus and a minus. The user can click either of these icons next to the desired facet value, and that will generate either a boolean AND (for the plus) or a boolean NOT (for the minus) in the subsequent query. If the user clicks multiple plus icons before executing the search, the resulting query will have the form <original query> AND category:("a" OR "b" OR "c"), where "a" "b" and "c" are the facets which were selected. For multiple minus selections, the resulting query will be of the form <original query> AND NOT category:"a" AND NOT category:"b" AND NOT category:"c".

What makes this implementation significantly different and more useful for certain applications than prior art is the the function where the user can build a complex syntax query using multiple selection states for facets and this query is displayed as editable to the user where the user can not only edit the facet refinement but also include refinements not related to facet computation. Since the facet list is not hierarchical there is no chance of a breadcrumb or tree. Further, the user can be presented with the entirety of facet values and create a complex query of overlapping or disparate terms through multiple selections before commitment. The user also has the option to create a simple query syntax with the available one click operation.

Figure 1 is a screenshot of the preferred embodiment. As you can see, we show the facet value along with its count. If the user were to click 'file', then it would immediately be appended to the query much like on Amazon.com or other internet sites which included faceted navigation. The +/- approach gives greater flexibility for our users and can result in powerful queries without having to know any query syntax.

This method does not preclude the preferred embodiment, however. Alternate methods of display could easily be adapted to perform a similar functionality. For example, a tag cloud with context sensitive menus and/or graphical devices could be employed. Another example would be a drag and drop method of query refinement that utilizes the context sensitivity within the query field/builder. The basis of this disclosure is that there is a faceted categorization displayed to the user in a meaningful way that also gives them the ability to create a complex query based on that visualization that goes beyond a simple navigation structure.

Search

Keywords:

Sender:

Recipient:

Subject:

Dates between:
 5 : 5

▼ Phrases

- + = [file](#) (272)
- + = [attached file](#) (269)
- + = [van hooser](#) (231)
- + = [agreement](#) (198)
- + = [word ibm omnifind yahoo](#) (168)

[more...](#)

► Senders

▼ Recipients

- + = [gerald.nemec@enron.com](#) (4990)
- + = [eric.gillaspie@enron.com](#) (605)
- + = [shonnie.daniel@enron.com](#) (572)
- + = [barbara.gray@enron.com](#) (562)
- + = [steve.hooser@enron.com](#) (501)

[more...](#)

► Sender Domains

► Recipient Domains

► People

► Companies

► Locations

► Flags

New search Add to search

Figure 1.