Pseudo Wildcard Search   
(GSA 7.0, Connector 3.x)

At the moment Google Enterprise Search cannot do wildcard searches. (searching for part of a filename).

As a workaround for a full wildcard search we can offer a search feature that can search for parts of a filename that has pre-defined separation characters (e.g.: underscore ‘\_’ ).

Let’s assume a document with the name “BYT\_BSC\_APP\_PROC” exists. Underscores are defined separation characters.

All words between separation characters can be used for as a search term but not parts of the words.

Also the full filename or parts of the full filename, as long as all characters between separation characters are used can be used for a search. This is true as long as the search term starts at the beginning or the end of the full filename but not somewhere in the middle

Best explained with an example.

The following search terms would find the document in this example :

* **From the beginning of the filename going forward in blocks**
  + BYT
  + BYT\_BSC
  + BYT\_BSC\_APP
* **From the end of the filename going backwards**
  + PROC
  + APP\_PROC
  + BSC\_APP\_PROC
* **The full filename**
  + BYT\_BSC\_APP\_PROC
* **Terms between separation characters**
  + BSC
  + BYT
  + APP
  + PROC
* **A combination of above**
  + e.g.: BYT\_BSC PROC 🡪   
    a filename that starts with BYT\_BSC and has the term PROC in the filename or in the content of a document.
  + The search only for filenames of documents (and ignoring the content) the following syntax can be used : “BYT\_BSC intitle:PROC”
  + or alternative “BYT\_BSC inurl:PROC” – which would also include the whole URL in the search (that includes names of site collection and lists etc…)
* **Search terms in the middle of filename cannot be found :**
  + BSC\_APP only
  + Or any combination like : SC\_APP or BSC\_ or \_APP or BYT\_B, …

Again here in a table to visualize it better :

|  |
| --- |
| **BYT\_BSC\_APP\_PROC** |
| 🡪 vom Start ... |
| BYT |
| BYT\_BSC |
| BYT\_BSC\_APP |
| 🡨 von Ende |
| PROC |
| APP\_PROC |
| BSC\_APP\_PROC |

# Edge Cases

This should list special cases of combinations that can occur and need to be defined.

|  |  |  |
| --- | --- | --- |
| Case | Proposal | Notes |
| Multiple separator in a row.  BYT\_\_BSC\_APP\_\_\_LAST | Reduce multiple separators to only one and then build combinations. BYT\_BSC\_APP\_LAST … | If they are not reduced it cannot be determined what token gets how many separator characters.  Example : aa\_\_bb. Is it aa\_ bb or aa \_bb ? |
| Combination of separation characters. Assuming \_ and : are such characters. BYT\_BSC:APP\_:\_Last | Threat all separation characters the same. Maybe |  |
| Separation char is the last character Aa\_bb\_\_c\_ | * Forward parsing remove it * Backward parsing add it; users need to seach for c\_ | Not sure what google does with filenames that end with underscore.   * 1 underscore - Removes it? * With 2 underscores ??   d |