

Contents

1. Reference:BlueSpiceExtendedSearch	2
2. Manual:Extension/BlueSpiceExtendedSearch	7

Reference:BlueSpiceExtendedSearch

Extension\ BlueSpiceExtendedSearch

Overview			
Description:	Elasticsearch search backend		
State:	stable	Dependency:	BlueSpice
Developer:	HalloWelt	License:	GPL-3.0-onlyProperty "BSExtensionInfoLicense" (as page type) with input value "</br>GPL-3.0-only" contains invalid characters or is incomplete and therefore can cause unexpected results during a query or annotation process.
Type:	BlueSpice	Category:	Search and Navigation
Edition:	BlueSpice free, BlueSpice pro, BlueSpice Farm, BlueSpice Cloud		
View help page			

Features

ExtendedSearch replaces the default MediaWiki search engine. It is based on ElasticSearch engine, and provides many improvements over the standard MediaWiki search both in terms of quality of indexed content and user interface.

- Search titles.
- Search the full content.
- Search uploaded or linked files (Office documents and PDFs).
- Search image data.
- Search-as-you-type and auto-complete.
- Ignore upper and lower case (case-insensitive).
- Search with the operators AND, OR, NOT.
- Search with wildcards.
- Search for phrases.
- Fuzzy search.
- Search for sentence fragments.

Some aspects of this extension can be configured on [Special:BlueSpiceConfigManager](#), under section "ExtendedSearch". Here wiki administrators can configure:

- external file paths
- layout of the autocomplete box
- language filter
- exclude patterns

Configuration (server)

For boosting by match percent, the following settings are available in the extension itself:

```
"ESMatchPercentBoostFactor": {
  "description": "How much to boost the result based on the percent of its
title taken up by the search term. Set to 0 to disable",
  "public": false,
  "value": "0.5"
},
"ESMatchPercentTitleField": {
  "description": "Field on which to base the match percent boosting. If
empty, default title field of the source will be used",
  "public": false,
  "value": ""
},
```

Description:

- **ESMatchPercentBoostFactor:** The value here is multiplied by the the match percent and by the score of the document. Example: The search term is "Test". The search finds for example "Testpage" and - let's say - assigns a score of 20 to this result. Your search term matches 50 percent of the found name, so its match factor is 0.5. This factor is then multiplied by the value in this setting (by default 0.5), which gives the total boost factor of 0.25. The final score is calculated as "original score + (original score * boost factor)" => $20 + (20 * 0.25) = 20 + 5 = 25$. This means that because your search term matches 50% of the result, it got boosted by the factor of 5.
- **ESMatchPercentTitleField:** Field on which the match analysis is performed. Defaults to nothing, so whatever the search lookup sets. It can be set to "prefixed_text" to calculate match percent on the title with namespace prefix, for example, or to something like "basename" to calculate the match percent only on the page name, without the namespace prefix.

Technical Information

This information applies to BlueSpice 3 . Technical details for BlueSpice cloud can differ in some cases.

Requirements

MediaWiki: 1.31
BlueSpiceFoundation: 3.2

Integrates into

- BlueSpiceArticleInfo
- BlueSpiceExtendedSearch
- BlueSpiceExtendedStatistics
- BlueSpicePrivacy
- BlueSpiceSimpleFarmer
- BlueSpiceTagCloud
- BlueSpiceVisualEditorConnector
- VisualEditor

Special pages

- BSSearchAdmin
- BSSearchCenter

Permissions

Name	Description	Role
extendedsearch-search-externalfile	Search for external files	accountmanager, admin, author, bot, commenter, editor, maintenanceadmin, reader, reviewer, structuremanager
extendedsearch-search-repofile	Search for files	accountmanager, admin, author, bot, commenter, editor, maintenanceadmin, reader, reviewer, structuremanager
extendedsearch-search-specialpage	Search for special pages	accountmanager, admin, author, bot, commenter, editor, maintenanceadmin, reader, reviewer, structuremanager
extendedsearch-search-wikipage	Search for pages	accountmanager, admin, author, bot, commenter, editor, maintenanceadmin, reader, reviewer, structuremanager

Configuration

Name	Value
ESAllowIndexingDocumentsWithoutContent	true
ESAutoRecognizeSubpages	true
ESAutoSetLangFilter	false
ESBackendClass	'\\BS\\ExtendedSearch\\Backend'
ESBackendHost	'127.0.0.1'

Name	Value
ESBackendPort	'9200'
ESBackendTransport	'http'
ESCompactAutocomplete	false
ESDefaultSearchOperator	'AND'
ESEnableSearchHistoryTracking	true
ESEnableTypeFilter	true
ESExternalFilePaths	array ()
ESLookupModifierRegExPatterns	array (0 => '[0-9]{2}\\-[0-9]{2}\\-[0-9]{4}', 1 => '[0-9]{4}\\-[0-9]{2}\\-[0-9]{2}', 2 => '[0-9]{2}\\-[0-9]{4}\\-[0-9]{2}', 3 => '[0-9]{2}\\W[0-9]{2}\\W[0-9]{4}', 4 => '[0-9]{4}\\W[0-9]{2}\\W[0-9]{2}', 5 => '[0-9]{2}\\W[0-9]{4}\\W[0-9]{2}', 6 => '[0-9]{2}\\.[0-9]{2}\\.[0-9]{4}', 7 => '[0-9]{4}\\.[0-9]{2}\\.[0-9]{2}', 8 => '[0-9]{2}\\.[0-9]{4}\\.[0-9]{2}', 9 => '[0-9]{2}\\W[0-9]{2}\\W[0-9]{4}', 10 => '[0-9]{4}\\W[0-9]{2}\\W[0-9]{2}', 11 => '[0-9]{2}\\W[0-9]{4}\\W[0-9]{2}', 12 => '[0-9]{1}\\-[0-9]{2}\\-[0-9]{4}', 13 => '[0-9]{4}\\-[0-9]{2}\\-[0-9]{1}', 14 => '[0-9]{1}\\-[0-9]{4}\\-[0-9]{2}', 15 => '[0-9]{4}\\-[0-9]{1}\\-[0-9]{2}', 16 => '[0-9]{2}\\-[0-9]{4}\\-[0-9]{1}', 17 => '[0-9]{2}\\-[0-9]{1}\\-[0-9]{4}', 18 => '[0-9]{1}\\W[0-9]{2}\\W[0-9]{4}', 19 => '[0-9]{4}\\W[0-9]{2}\\W[0-9]{1}', 20 => '[0-9]{1}\\W[0-9]{4}\\W[0-9]{2}', 21 => '[0-9]{4}\\W[0-9]{1}\\W[0-9]{2}', 22 => '[0-9]{2}\\W[0-9]{4}\\W[0-9]{1}', 23 => '[0-9]{2}\\W[0-9]{1}\\W[0-9]{4}', 24 => '[0-9]{1}\\.[0-9]{2}\\.[0-9]{4}', 25 => '[0-9]{4}\\.[0-9]{2}\\.[0-9]{1}', 26 => '[0-9]{1}\\.[0-9]{4}\\.[0-9]{2}', 27 => '[0-9]{4}\\.[0-9]{1}\\.[0-9]{2}', 28 => '[0-9]{2}\\.[0-9]{4}\\.[0-9]{1}', 29 => '[0-9]{2}\\.[0-9]{1}\\.[0-9]{4}', 30 => '[0-9]{1}\\W[0-9]{2}\\W[0-9]{4}', 31 => '[0-9]{4}\\W[0-9]{2}\\W[0-9]{1}', 32 => '[0-9]{1}\\W[0-9]{4}\\W[0-9]{2}', 33 => '[0-9]{4}\\W[0-9]{1}\\W[0-9]{2}', 34 => '[0-9]{2}\\W[0-9]{4}\\W[0-9]{1}', 35 => '[0-9]{2}\\W[0-9]{1}\\W[0-9]{4}', 36 => '[0-9]{2}\\-[0-9]{2}\\-[0-9]{2}', 37 => '[0-9]{2}\\W[0-9]{2}\\W[0-9]{2}', 38 => '[0-9]{2}\\.[0-9]{2}\\.[0-9]{2}', 39 => '[0-9]{2}\\W[0-9]{2}\\W[0-9]{2}', 40 => '[0-9]{1}\\-[0-9]{1}\\-[0-9]{4}', 41 => '[0-9]{4}\\-[0-9]{1}\\-[0-9]{1}', 42 => '[0-9]{1}\\-[0-9]{4}\\-[0-9]{1}', 43 => '[0-9]{1}\\W[0-9]{1}\\W[0-9]{4}', 44 => '[0-9]{4}\\W[0-9]{1}\\W[0-9]{1}', 45 => '[0-9]{1}\\W[0-9]{4}\\W[0-9]{1}', 46 => '[0-9]{1}\\.[0-9]{1}\\.[0-9]{4}', 47 => '[0-9]

Name	Value
	<pre>{4}\\.[0-9]{1}\\.[0-9]{1}', 48 => '[0-9]{1}\\.[0-9]{4}\\.[0-9]{1}', 49 => '[0-9]{1}\\\\[0-9]{1}\\\\[0-9]{4}', 50 => '[0-9]{4}\\\\[0-9]{1}\\\\[0-9]{1}', 51 => '[0-9]{1}\\\\[0-9]{4}\\\\[0-9]{1}', 52 => '[0-9]{1}\\.[0-9]{1}\\.[0-9]{2}', 53 => '[0-9]{2}\\.[0-9]{1}\\.[0-9]{1}', 54 => '[0-9]{1}\\.[0-9]{2}\\.[0-9]{1}', 55 => '[0-9]{1}\\[0-9]{1}\\[0-9]{2}', 56 => '[0-9]{2}\\[0-9]{1}\\[0-9]{1}', 57 => '[0-9]{1}\\[0-9]{2}\\[0-9]{1}', 58 => '[0-9]{1}\\.[0-9]{1}\\.[0-9]{2}', 59 => '[0-9]{2}\\.[0-9]{1}\\.[0-9]{1}', 60 => '[0-9]{1}\\.[0-9]{2}\\.[0-9]{1}', 61 => '[0-9]{1}\\\\[0-9]{1}\\\\[0-9]{2}', 62 => '[0-9]{2}\\\\[0-9]{1}\\\\[0-9]{1}', 63 => '[0-9]{1}\\\\[0-9]{2}\\\\[0-9]{1}', 64 => '[0-9]{2}\\.[0-9]{4}', 65 => '[0-9]{2}\\[0-9]{4}', 66 => '[0-9]{2}\\.[0-9]{4}', 67 => '[0-9]{2}\\\\[0-9]{4}', 68 => '[0-9]{4}\\.[0-9]{2}', 69 => '[0-9]{4}\\[0-9]{2}', 70 => '[0-9]{4}\\.[0-9]{2}', 71 => '[0-9]{4}\\\\[0-9]{2}', 72 => '[0-9]{2}\\.[0-9]{2}', 73 => '[0-9]{2}\\[0-9]{2}', 74 => '[0-9]{2}\\.[0-9]{2}', 75 => '[0-9]{2}\\\\[0-9]{2}',)</pre>
ESMatchPercentBoostFactor	'0.5'
ESMatchPercentTitleField	"
ESOfferOperatorSuggestion	true
ESRecentBoostFactor	'0.5'
ESSearchCenterDefaultFilters	array (0 => 'namespace_text', 1 => 'categories',)
ESSearchInRawWikitext	true
ESSourceConfig	<pre>array ('wikiage' => array ('skip_namespaces' => array (0 => 8, 1 => 9,),), 'repofile' => array ('extension_blacklist' => array (0 => 'mp4',), 'max_size' => 20000000,), 'externalfile' => array ('extension_blacklist' => array (0 => 'mp4',), 'max_size' => 20000000,),)</pre>
ESWildcardingOperators	array (0 => '+', 1 => ' ', 2 => '*', 3 => '(', 4 => ')', 5 => '~',)
ESWildcardingSeparators	array (0 => ',', 1 => '.', 2 => ';', 3 => '-', 4 => '_',)
ExtendedSearchExternalFilePathsExcludes	array ()
TagSearchSearchFieldTemplatePath	'/resources/templates'

API Modules

- [bs-extendedsearch-autocomplete](#)
- [bs-extendedsearch-query](#)
- [bs-extendedsearch-resultrelevance](#)
- [bs-extendedsearch-stats](#)
- [bs-extendedsearch-triggerupdate](#)
- [bs-extendedsearch-type-store](#)

Hooks

- [BSInsertMagicAjaxGetData](#)
- [BSUEModulePDFBeforeCreatePDF](#)
- [BeforePageDisplay](#)
- [GetPreferences](#)
- [LoadExtensionSchemaUpdates](#)
- [NamespaceManager::editNamespace](#)
- [ResourceLoaderTestModules](#)
- [SkinTemplateOutputPageBeforeExec](#)

Extended search

A [quality version](#) of this page, [approved](#) on *26 April 2021*, was based off this revision.

ExtendedSearch replaces the default MediaWiki search engine. It is based on the [Elasticsearch](#) engine and provides many improvements over standard MediaWiki search both in terms of quality of indexed content and user interface. Pages and files are indexed immediately and available for searching in real-time.

Contents

1 Usage	8
2 Autocomplete	8
2.1 Namespace and subpage "pills"	9
2.2 Configuration	9
2.3 Search operators	9
2.4 Dates as search terms	10
3 Search Center (extended search)	10
3.1 Search operators	11
3.2 Elements	11
3.2.1 Results relevance	13
3.2.2 Exporting search results	13
4 Configuration	14
5 SearchAdmin	15
6 Related info	16

Usage

Extended search offers two ways of searching for content:

1. **Title search** using the quick search
2. **Fulltext search** using the search center

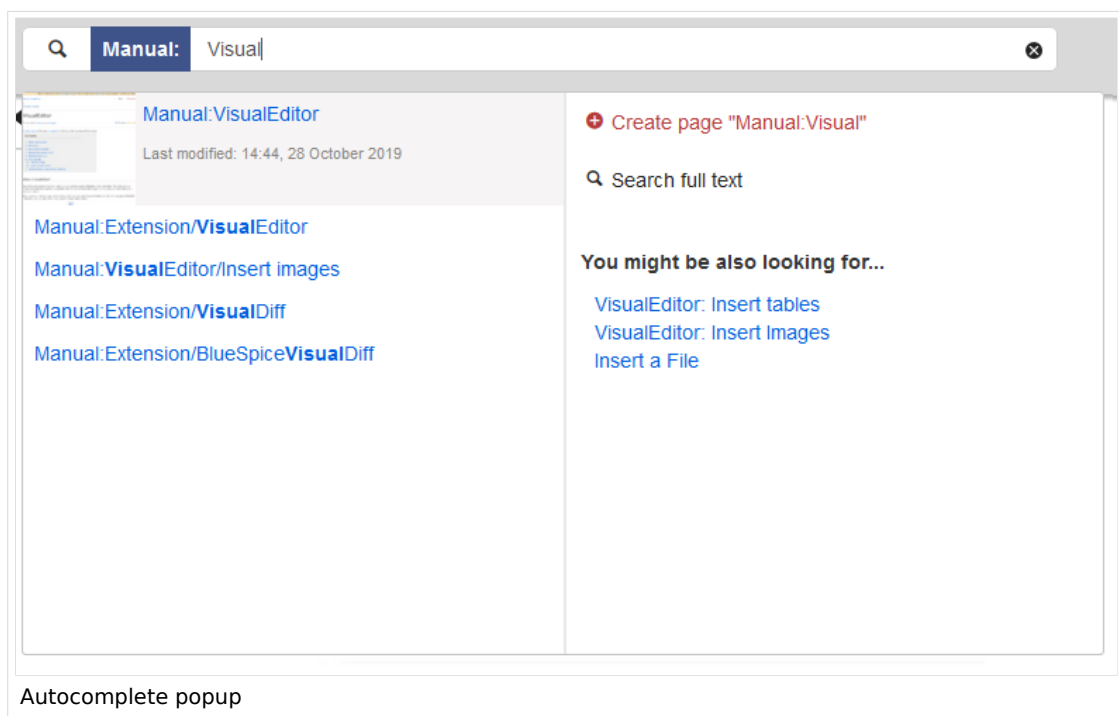
By default, the search indexes:

- wiki pages: all content pages, except for pages in the MediaWiki namespace
- special pages: all pages in the namespace *Special*
- repo files: files uploaded to the wiki - for text files like Word documents and PDFs, this includes content and metadata
- external files: files from arbitrary locations from accessible storage paths (e.g., mounted drives)
- social entities: entities (pages) created by the [BlueSpiceSocial](#) extension, like blog posts and user profiles

Autocomplete

The search box in the main toolbar offers with autocomplete functionality is available from any page on the wiki. Autocomplete matches only page and file titles, but does not provide a full-text search.

After typing in the search bar, an autocomplete popup appears:



This popup is divided in two columns.

Left column:

- **Primary results:** These are the results that are direct matches for the search term. The exact matches are shown in bold letters.

- When looking for the term "quality management", it is also possible to type in "qual man" and the search starts looking for matches
- The best match is often shown with a thumbnail preview as a top search result

Right column:

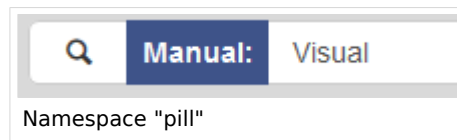
- **Tools and secondary results:** Tools are displayed in the top portion of the right column and contain:
 - a button to create the page with the current search term as the title (only if user has permission to create new pages)
 - a button to execute full-text search for the current term. Selecting this option leads to the Search Center.

In the bottom portion of the right column, secondary results are displayed. Secondary results are matches that do not directly match the given term, but might be relevant for the user. These show up only in certain contexts:

- If there are significantly more matches for a similar term than the term actually searched for (in a situation of a typo for example)
- If the namespace or subpage pill is present (see next section). In this case, secondary results will show direct matches in other namespaces/pages.

Namespace and subpage "pills"

Autocomplete search bar supports pills for namespaces and subpages.



If a user types `Manual:`, a pill for the namespace "Manual" appears and the search for the following term is limited to this namespace. Subpage pills work in the same way. If the user types in `PageName/`, any term after that is matched in the context of that page. If you search for "TestPage/Te", for example, the search will match only subpages of "TestPage" that contain "Te" in the title, like "TestPage/Test" and "TestPage/TestSubpage".

Configuration

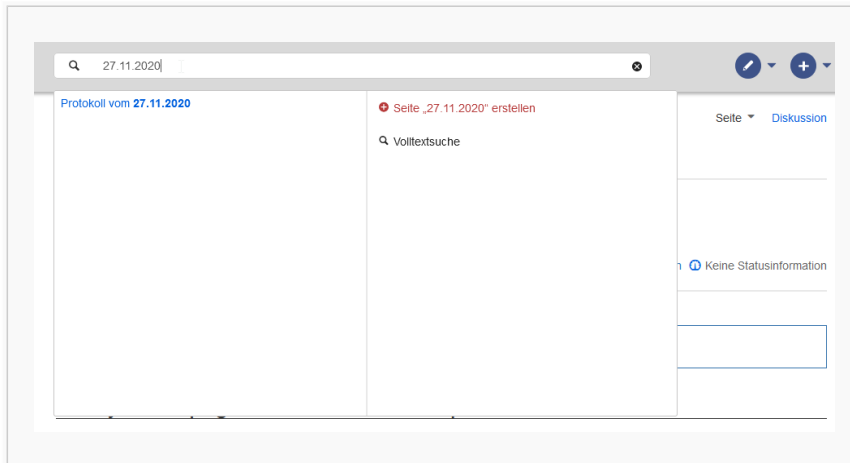
Previously explained is the full version of the autocomplete popup. Wiki admins can also choose to use "Simple autocomplete", which contains only one column with primary results and tools, without secondary results. This type of autocomplete is also more compact. This setting is available from [BlueSpiceConfigManager](#), under *Extension > BlueSpiceExtendedSearch > Use simple autocomplete*.

Search operators

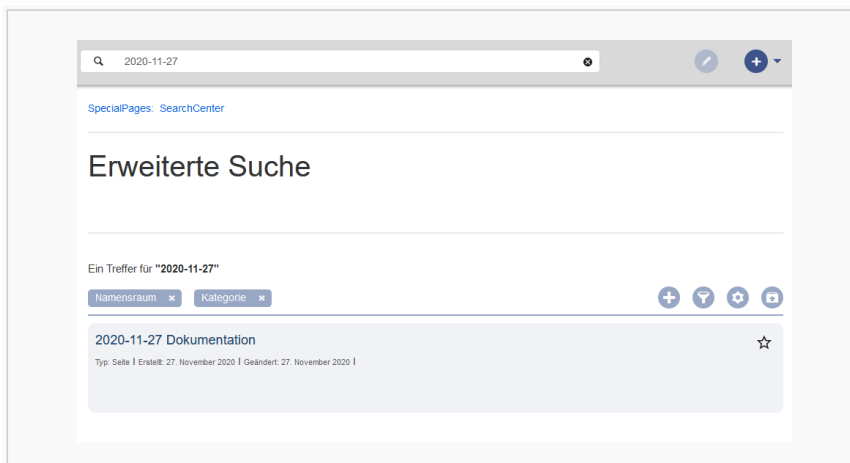
Since the autocomplete search only searches for titles, search operators such as AND or OR are not supported. However, they are supported using the Search Center.

Dates as search terms

The search recognizes certain date formats. In the search center, these date formats are treated with surrounding quotation marks.



Date format 1

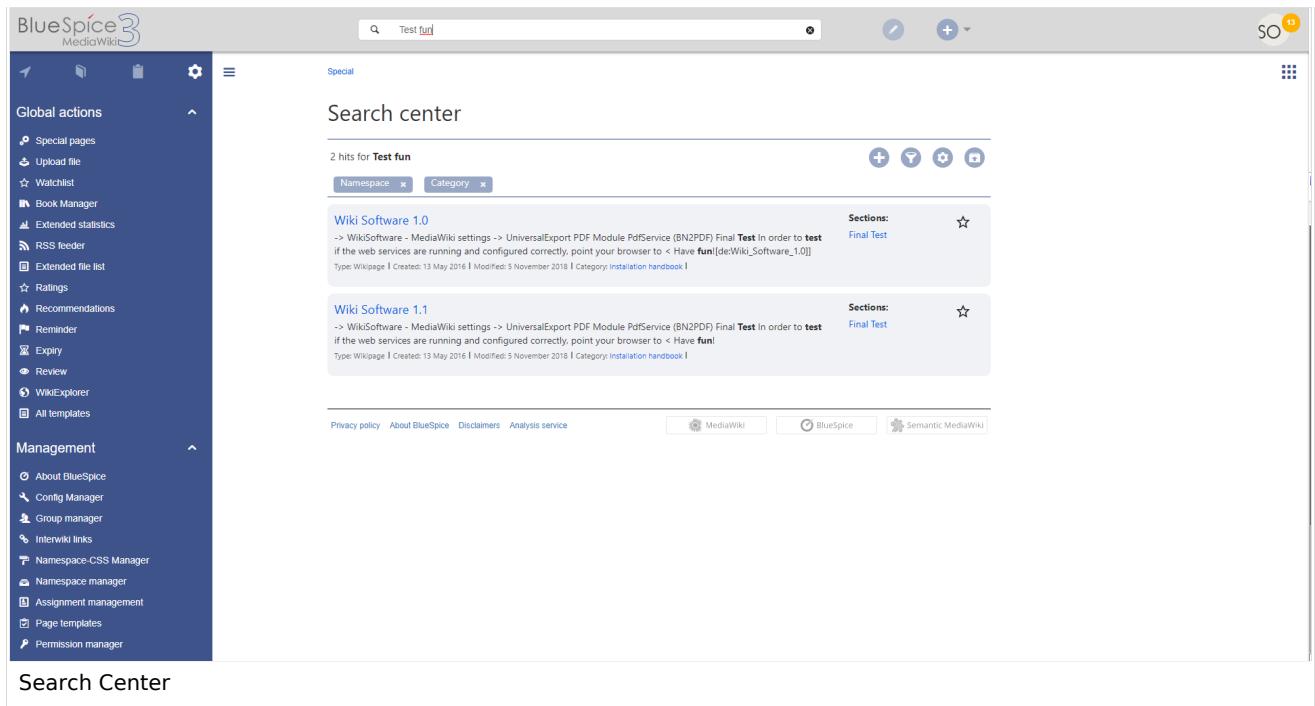


Date format 2

Search Center (extended search)

As soon as you trigger the full-search with the magnifying glass symbol or the *Enter* key, the wiki switches to the Search Center. This is a special page that offers the full functionality of the search. The search in the Search Center is carried out using the same search bar that is also used for autocomplete.

Here, the search engine carries out a full text search, which also includes the metadata of documents. It does not index the wikitext of a wiki page, but rather the rendered content, that is the output by the browser.



Search operators

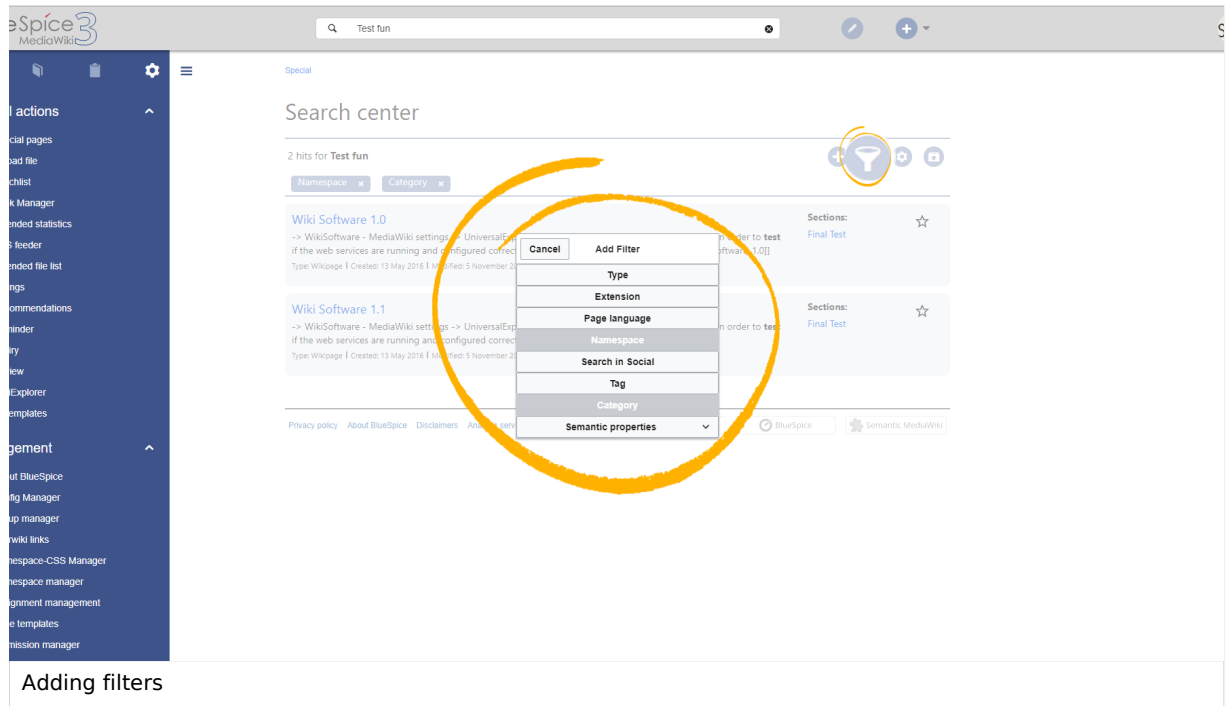
To use search operators in the full-text search, the following syntax can be used:

- Sandbox **AND** Widgets - returns pages that contain both the term sandbox and the term widgets
- Sandbox **OR** Widgets - returns pages that contain either the term sandbox or the term widgets.
- (Sandbox) **NOT** (Widgets) - returns pages that contain the term sandbox but not the term widgets..

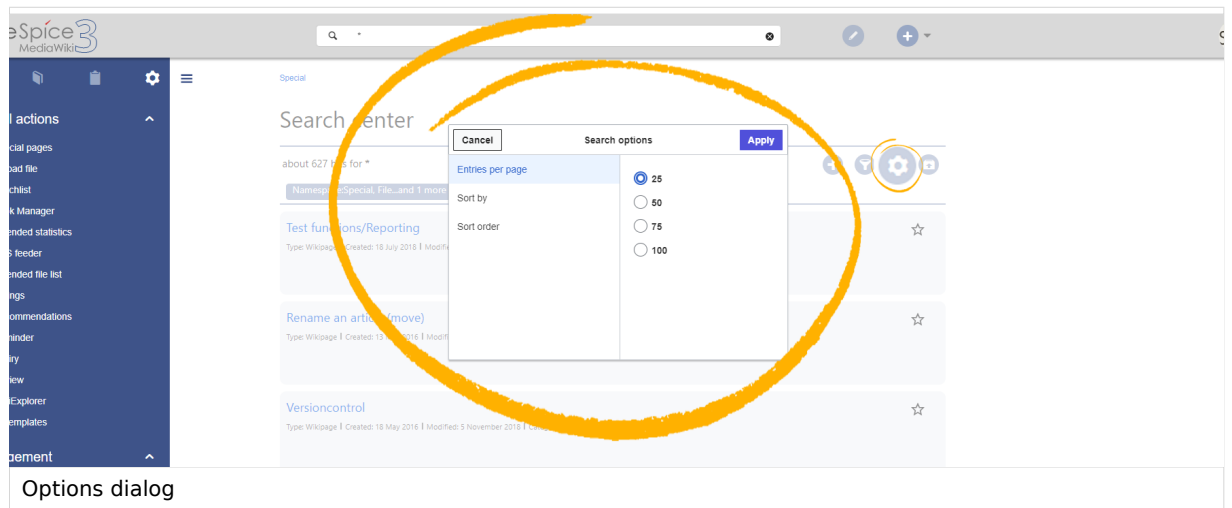
The search operators have to be written in uppercase letters.

Elements

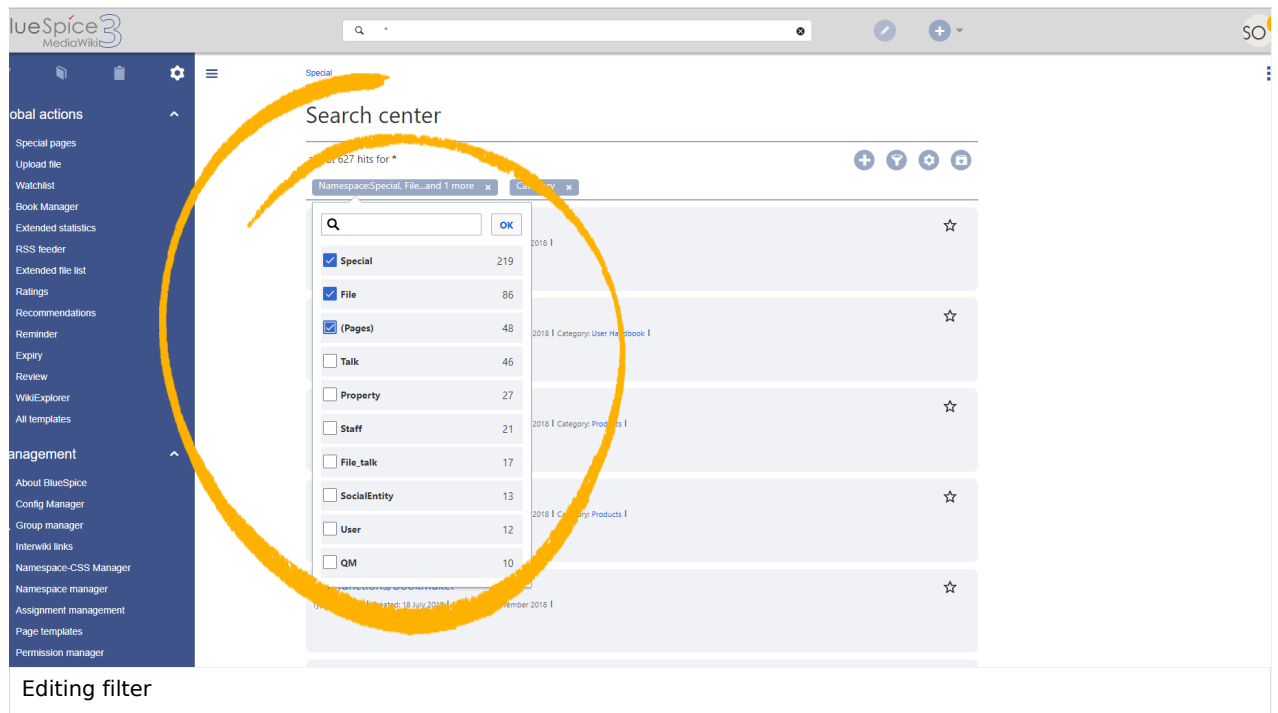
- **Hit count** - shows how many matches a search has returned. It may be exact or approximate. Approximation is due to dynamic check for permissions to view each page in the results, some results may not be shown to the user with no adequate permissions event if it matches the search term.
- **Tools** - these are located in the top right as a set of buttons:
 - Create a page - represented by a "plus" button - will create a page with current search term as a title (only visible if user has permission to create pages)
 - Add filters - represented by a "funnel" icon - opens a dialog where users can select the type of filter to apply to the search.



- Search options - represented by a "gear" icon - opens a dialog where users can change settings for the current search. It provides options like the number of results per page, sorting field and order.



- Export search - represented by the "export" icon - opens a dialog to export current search results to a dedicated page (only visible to authorized users, by default admin users).
- **Filter:** Filters are displayed in form of pills below the results counter. Each filter pill represents a different attribute. Filter values can be changed by clicking on the filter pill and selecting the desired options from the popup that will open. Filter can be removed by clicking on the "x" button in each pill.



- **Results:** If current search criteria yield any results, a list of results is displayed. The information available in the results is depending on the result type):
 - Main title of the result (page title, file name...)
 - Result type
 - Size and extension (for files)
 - Thumb image (for images) or file type icon (for other file types)
 - Redirect information
 - Original name (for wikipages, if the page display title is different than the actual page title)
 - Secondary information like page categories, modification and creation date, author, etc.
 - Snippet of the matches in the content of the page (if there are matches in the page content),
 - Sections with matches in the content (for wikipages)

In case there is an exact match for a given search term, the result that matches exactly is displayed as a "featured" result. Such a result displays a preview image of the page and always displays the content snippet, regardless of whether there are matches in the content or not.

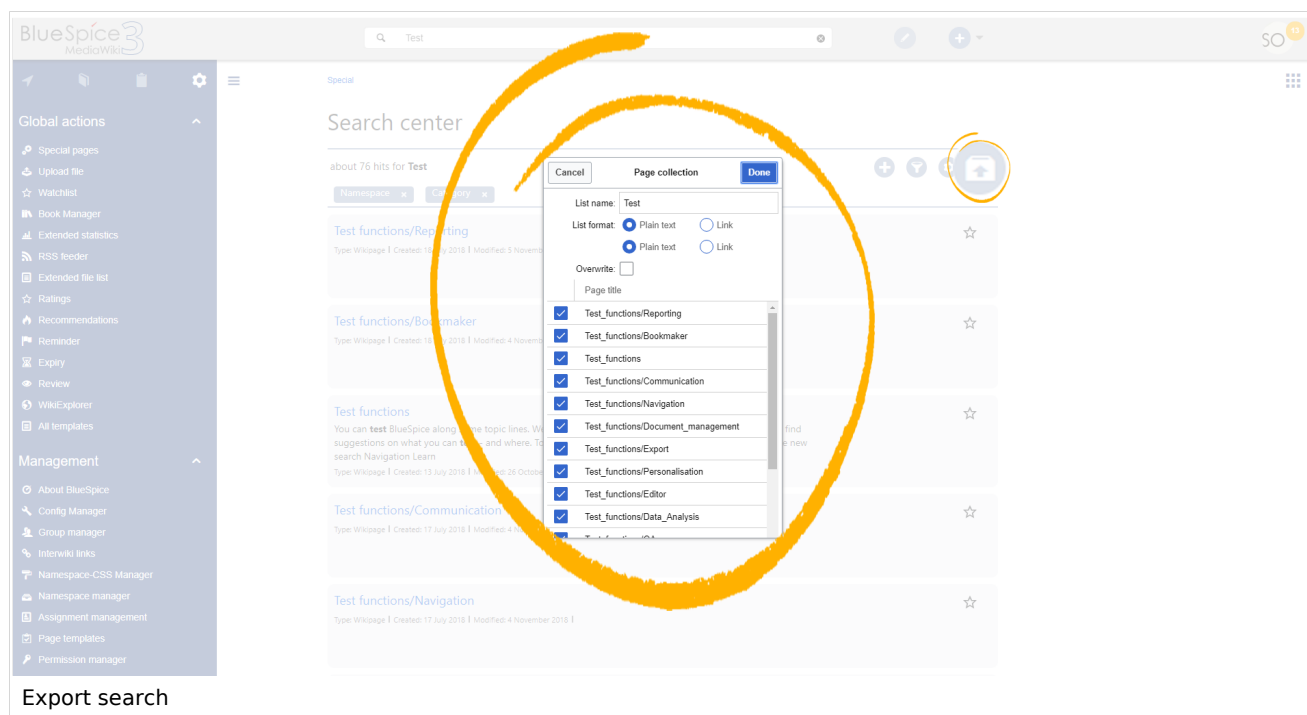
Results relevance

All logged-in users can mark certain results as "favorites" by clicking on the star symbol in the right corner of each result. Files marked as "favorite" will score higher in future searches - this is not visible immediately, however, but happens over time. One user's relevance settings won't affect other user's searches. Results can be unmarked as "favorite" by clicking on the "star" symbol again.

Exporting search results

After executing the search, administrators can export (save) the results to a specific page using the "Export" command from the "Tools" menu. The list is exported to *MediaWiki:PageCollection /LISTENNAME*.

Only valid wikipages are exported. The pages contained in a list can be included in a book in the book function via *Mass add> Source: Page collection*.



Configuration

In the [BlueSpiceConfigManager](#), under settings for the extension *BlueSpiceExtendedSearch*, the following configuration options are available:

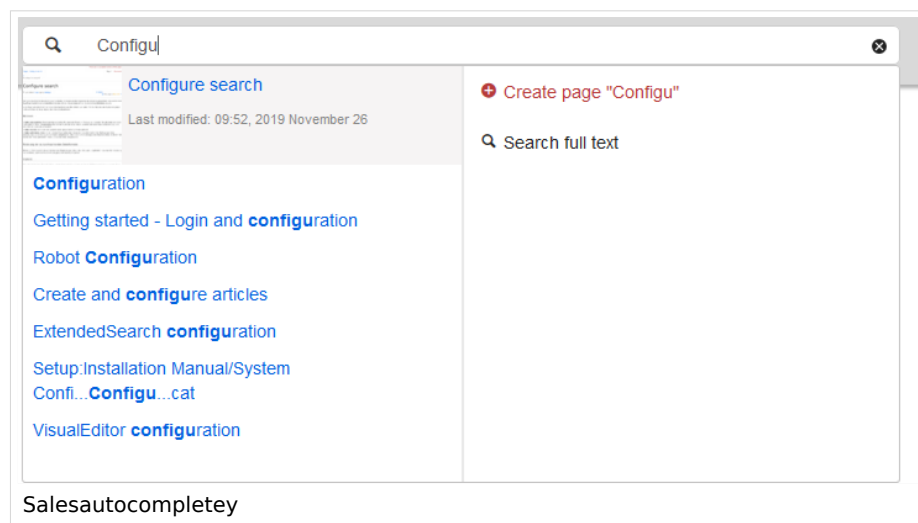
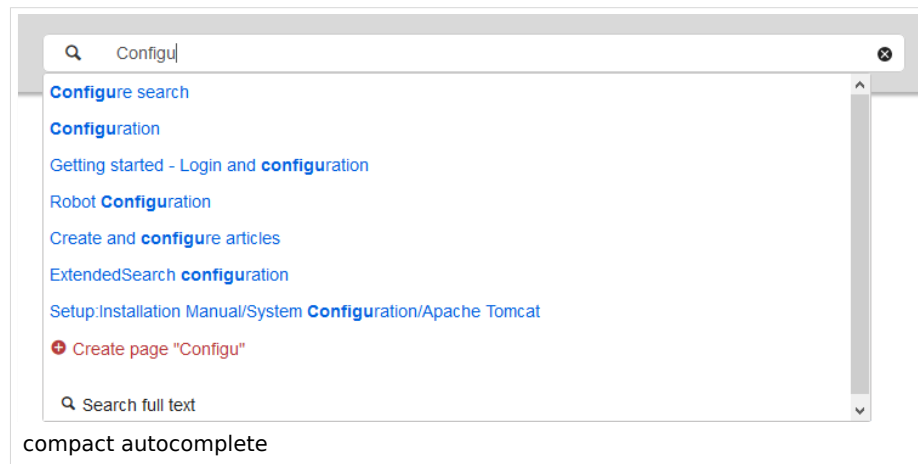
- External file paths:** External locations to scan for and index files from. These locations are entered in key/value pairs of Physical path and URL prefix.
 - Physical path:* Storage path. For mounted drives, for example: *X:\Documents*
 - URL prefix:* All files indexed from the indexed physical location are prefixed in the search results. It should be used if files are available from a web server. By default, the URL prefix is set to the same value as *Physical path*.

Examples for external file paths

Example file	Phys. Path	URL Prefix	Link in results list
S:\Sales\Info\glossaryy.pdf	S:\	file:///S:/	file:///S:/Sales/Info/glossary.pdf
		https://sharepoint/Sales/ documents/Sales/Info/glossary.pdf	https://sharepoint/Sales/ documents/Sales/Info/glossary.pdf
		https://intranet/showdocs	https://intranet/showdocsinline.aspx?
		inline.aspx?path=	path=/Sales/Info/glossary.pdf

/var/www/bluespice/w /mounted_ share/Sales/Info/glossary. pdf	/var/www/bluespice /w/ mounted_share/		https://wiki.company.local/w /mounted_ share/Sales/Info/glossary.pdf
--	---	--	--

- **Use simple autocomplete:** forces a more compact autocomplete (1-column Sales thumbnail preview).



- **Automatically set language filter:** This setting applies to multi-languageSalesIt returns seayarch results for a user only in the user's language. Users can set their language in their [preferences](#).
- **Exclude patterns:** Files at external file locations can be exluded from search based on regular expression patterns.

SearchAdmin

Special:SearchAdmin is a special page for wiki admins. This page provides stats on documents indexed in the search.

Related info

- [Reference:BlueSpiceExtendedSearch](#)