Unified Index

This applies to the search and capabilities in Tiki to filter (search filters, permissions, etc) content. It is important if not critical to keep a fresh Unified-Index as Tiki relies on the freshness of the Unified-Index to display and filter content.

Tiki can support multiple search engines internally. Each of those will have different capabilities and limitations. The default engine should provide capabilities good enough for small and medium sites. Larger sites may need additional infrastructure to get the most performance. Please see: Unified Index Comparison

Fields

Below is a matrix between the fields and the object types.
Legend:

X - Available
/ - Static value
? - Depends on the data

Field	Type	Tokenized*	Sortable	wiki page	forum post	blog post	article	file	trackeritem	sheet	comment	user	Available in Tiki version
object_type	Generic		X	X	X	X	X	X	X	X	X	X	7
object_id	Generic		X	X	X	X	X	X	X	X	X	X	7
title	Generic	X	X	X	X	X	X	X	X	X	?	X	7
title_initial	Generic			X	X	X	X	X	X	?	?		?
title_firstword	Generic			X	X	X	X	X	X	?	?		?
language	Generic		X	X			X		?				7
creation_date	Generic		X	X	X	X	X	X	X		X	X	7/15
modification_date	Generic		X	X	X	X	X	X	X	X	X		7
contributors	Generic			X	X	X	X	X	X	X	X		7
description	Generic	X		X			X	X		X			7
contents	Generic	X		X	X	X	X	X	X	X	X	X	7
wiki_content	Specific			X								X	7
wiki_uptodateness	Specific		X	X									7
wiki_approval_state	Specific			X									11
post_content	Specific				X								7
post_snippet	Specific				X								14

Field	Type	Tokenized*	Sortable	wiki page	forum post	blog post	article	file	trackeritem	sheet	comment	user	Available in Tiki version
parent_thread_id *(not to be confused with parent_object_id)*	Specific				X								8
root_thread_id	Specific				X								14
parent_contributors	Specific				X								14
blog_id	Specific		X			X							7
blog_excerpt	Specific					X							7
blog_content	Specific					X							7
topic_id	Specific		X				X						7
article_content	Specific						X						7
article_topline	Specific						X						7
article_subtitle	Specific						X						7
article_author	Specific						X						9
article_type	Specific						X						9
article_heading *available as description*	Specific						X						9
published	Specific						X						13
sitetitle	Specific						X						13
siteurl	Specific						X						13
gallery_id	Specific		X					X					7
filename_id	Specific		X					X					7
filetype	Specific	X	X					X					7
filesize	Specific	X	X					X					15
file_comment	Specific							X					7
file_content	Specific							X					7
tracker_id	Specific		X						X				7
tracker_status	Specific		X						X				7
tracker_field_*PERMNAME/ID* *(see below for more details)*	Specific	X	?						X				7

Field	Type	Tokenized*	Sortable	wiki page	forum post	blog post	article	file	trackeritem	sheet	comment	user	Available in Tiki version
sheet_content	Specific									X			7
comment_content	Specific	X									X		7
user_country	Specific	X	X									X	10
groups	Specific	X	X									X	?
hits	Specific				X								15
lastpost_title	Specific				X								15
lastpost_modification_date	Specific				X								15
lastpost_contributors	Specific				X								15
lastpost_post_content	Specific				X								15
lastpost_post_snippet	Specific				X								15
lastpost_hits	Specific				X								15
lastpost_thread_id	Specific				X								15
view_permission	Internal			X			X			X			7
parent_object_type	Internal		X		X	X	X	X	X		X		7
parent_object_id	Internal		X		X	X	X	X	X		X		7
parent_view_permission	Internal				X	X	X	X	X		X		7
global_view_permission	Internal										X		7
hash	Internal			X									7
url	Internal			X							X		7
categories	Global			X	X	X	X	X	X	X	X		7
deep_categories	Global			X	X	X	X	X	X	X	X		7
allowed_groups	Global			X	X	X	X	X	X	X	X		7
freetags	Global			X	X	X	X	X	X	X	X		7
freetags_text	Global			X	X	X	X	X	X	X	X		7
adv_rating_*ID*	Global		X	X	X	X	X	X	X	X	X	X	7
comment_count	Global		X	X	X	X	X	X	X	X	X		8
relations	Global			X	X	X	X	X	X	X	X	X	8
attachments	Global			X	X	X	X	X	X	X	X		7

Field	Type	Tokenized*	Sortable	wiki page	forum post	blog post	article	file	trackeritem	sheet	comment	user	Available in Tiki version
attachment_contents	Global	X		X	X	X	X	X	X	X	X		7
geo_located	Global			X	X	X	X	X	X	X	X	X	9
geo_location	Global			X	X	X	X	X	X	X	X	X	9
visits	Global		X	X				X					9.2

** Tokenized - as in decomposed in words for full text search*

Searchable and not searchable

Some fields are indexed and are searchable and other are not searchable (more explanation is required). It will influence results that will be outputted when using the plugins List, CustomSearch and ListExecute (probably).
or exemple, deepcategories (child from a parent category) are indexed but not searchable. When you set one of thos plugins you need to specify the `searchable_only` parameter as follow, `searchable_only="0"`
Note that searching for not searchable objects in the Tiki "Experiment with plugin LIST tool" will output "no result found"
The plugin should start as follow:

```
{list(searchable_only="0")}
```

Tracker Fields

In general, tracker fields are indexed as **tracker_field_*PERMNAME/ID***. *PERMNAME/ID* is your tracker field permanent name or ID. However, many tracker field types have additional useful variants of the main field (see below) that are indexed for each field.

The indexing for tracker fields will vary depending on the field type. As a general rule, tracker_field_*PERMNAME/ID* will be used as the field and will be sortable. However, there are a few exceptions:

- Image and File fields are not indexed
- TextArea is not sortable

Multilingual fields are indexed as multiple fields

- The main one (tracker_field_*PERMNAME/ID*) contains all languages
- tracker_field_*PERMNAME/ID_lang* contains one language only (tracker_field_12_fr for example)

Rating and related fields store as multiple fields

- tracker_field_*PERMNAME/ID* contains the average
- tracker_field_*PERMNAME/ID*_sum contains the vote totals
- tracker_field_*PERMNAME/ID*_count contains the number of votes

Items List and Item Link fields

- tracker_field_*PERMNAME/ID*_text contains the text instead of the IDs of the linked/listed items

Language of the tracker item

- If a language field is set for the tracker item, that language is indexed as the item language, i.e. the **language** field.

Some used in buildQuery/tiki-searchindex.php (need explanation on whether these are real fields or just helpers):

type: refers to **object_type**

deep: if this is set, **categories** will be considered **deep-categories**

autocomplete: Will search for items with title starting with this

Rebuild search index

Please see Rebuild Unified Index

MySQL limits for very big numbers of tracker fields

How to detect

If *Unified Search* is configured with the *MySQL Full Text Search* engine, you might encounter a failure to reindex with the following symptoms:

- `php console.php index:rebuild` stops without displaying the list of indexing statistics:

php console.php index:rebuild [14 mai 2020 09:26 EDT] Started rebuilding index... Unified search engine: MySQL, version 10.1.45-MariaDB (it takes a very long time and in the end nothing shows up)

If you investigate with producing a log it ends this way:

php console.php index:rebuild --log [14 mai 2020 16:10 EDT] Started rebuilding index... logging to file: temp/Search_Indexer_mysql_database_name_console.log Unified search engine: MySQL, version 10.1.45-MariaDB (i takes a very long time and in the end nothing shows up) tail temp/Search_Indexer_mysql_database_name_console.log ... 2020-05-14T16:21:47-04:00 ERR (3): Indexing failed while processing "2512" (type trackeritem) with the error "Could not perform index modification: Too many columns 2020-05-14T16:21:47-04:00 ERR (3): WARNING: PDO::query(): SQLSTATE[HY000]: General error: 1117 Too many columns {"code":null,"file":"/path/to/tikiroot/lib/core/TikiDb/Pdo.php","line":104} 2020-05-14T16:21:47-04:00 INFO (6): addDocument trackeritem 2513 {"memoryUsage":"46.5 MiB"} 2020-05-14T16:21:47-04:00 ERR (3): Indexing failed while processing "2513" (type trackeritem) with the error "Could not perform index modification: Too many columns" 2020-05-14T16:21:47-04:00 ERR (3): WARNING: PDO::query(): SQLSTATE[HY000]: General error: 1117 Too many columns {"code":null,"file":"/path/to/tikiroot/lib/core/TikiDb/Pdo.php","line":104} 2020-05-14T16:21:47-04:00 INFO (6): addDocument trackeritem 2514 {"memoryUsage":"46.5 MiB"}

Technical explanation (thanks Victor)

[+]

Solutions

The following options in `Control panels → Search` help for this situation:

- **MySQL use short field names**
  Due to frm file constraints, number of search fields that one index can hold is usually limited to about 1500. This can be exceeded if you have numerous tracker fields. Enabling this option will try to shorten the field names internally that should allow you to use 300-500 more fields. Switching this option requires full index rebuild.
  Summary: This makes no difference in how you use Tiki.
- **Don't index non searchable fields**
  Indexing will skip adding all tracker fields that are not marked as "searchable". This wil free index space but also make it impossible to use those fields in search index queries
  Summary: You need to review the *Searchable* property of you tracker fields. A full index rebuild will be necessary after changes in the tracker fields *Searchable* properties.

Hint: If the first option is enough for your site, you may ignore the second one.
Indexing log common error troubleshooting

When rebuilding your index use the following command to create a log file.
Note you need to point to the PHP for the PHP version your Tiki uses.

```
php console.php i:r -p --log
```

In the log you may see some error, this is an attempt to help and diagnose the possible issues.

Error	Possible fix

addDocument article 275
ERR (3): NOTICE: Trying to access array offset on value of type
bool.../...lib/core/Search/ContentSource/ArticleSource.php","line":53

Edit article275, Advanced tab check
the URL validity in the "Source" field.

Error: Index is being rebuilt at the moment and cannot start another rebuild process.
This error happens when rebuilding dies with fatal error, our PHP shutdown functions are not executed and thus doesn't clear the preference flag that index is being rebuilt. So you try to rebuild and it is blocked because it thinks it's still rebuilding. So make it's not (reboot or kill the process) and then

php console.php preferences:delete unified_manticore_index_rebuilding

Unified Index storage in the database (MySQL full-text search)

By design the unified-index store data in the MySQL database using MyISAM no matter if your Tiki uses innoDB or MyISAM (not all versions of InnoDB offer FULLTEXT).

Duplicate unified-index table

For some reasons you may found that your database has several index (tables like index_...alphanumeric...). This happen when they were issues rebuilding the index or when moving and upgrading the database (Tiki upgrade). You can delete those tables (always consider saving a backup of your data prior any work) and rebuild the unified-index. Tiki will recreate and relink everything properly. You can find information about the unified-index in use at the Search control panel, General settings under Unified search index.

Related

- Email in Unified Index
- Cron Job to Rebuild Search Index
- PluginList
- PluginCustomSearch
- Unified Search

Developer Notes

See https://dev.tiki.org/Unified-Index#Developer_Notes.

alias names for this page

Unified Search | UnifiedSearch | Enterprise search | Search Index | SearchIndex | UnifiedIndex