Table of Contents

FilterPlugin **1**
Description 1
Syntax Rules 1
SUBST 1
STARTSUBST, STOPSUBST 1
EXTRACT 2
STARTEXTRACT, STOPEXTRACT 2
FORMATLIST 2
MAKEINDEX 3
Examples 3
EXTRACT Example 1: convert table into text 3
EXTRACT Example 2: convert text into table 4
MAKEINDEX example 1: creating an index from a chunk of text 5
A 7
B 7
C 7
D 8
E 8
F 8
H 8
I 7
I 7
K 8
L 8
M 8
N
O
P
P
R
S
T
U
V
W
m
p
t
MAKEINDEX example 2: creating an index for a search result
C
I
L
O
R
S
S
T
W
Plugin Installation Instructions
Plugin Info

FilterPlugin

Powered by
WikiRing Consultants

Description

This plugin allows to substitute and extract information from content by using regular expressions. There are three different types of new functions:

1. FORMATLIST: maniplulate a list of items; it is highly configurable to define what constitutes a list and how to extract items from it
2. SUBST, STARTSUBST/STOPSUBST: substiture a pattern in a chunk of text
3. EXTRACT, STARTEXTRACT/STOPEXTRACT: extract a pattern from a text

While the START-STOP versions of SUBST and EXTRACT work on inline text, the normal versions process a source topic before including it into the current one.

Syntax Rules

SUBST

Syntax: `%SUBST{topic="..." ...}%`

insert a topic by processing its content.

- `topic="..."`: name of the topic text to be processed
- `text="..."`: text to be processed (has got higher precedence than 'topic')
- `pattern="..."`: pattern to be extracted or substituted
- `format="..."`: format expression or pattern substitute
- `header="..."`: header string prepended to output
- `footer="..."`: footer string appended to output
- `limit="<n>"` maximum number of occurences to extract or substitute counted from the start of the text (defaults to `100000` aka all hits)
- `skip="<n>"` skip the first n occurences
- `exclude="..."`: skip occurences that match this regular expression
- `include="..."`: skip occurences that don't match this regular expression
- sort="on,off,alpha,num" order of the formatted items (default "off")
- `expand="on,off"`: toggle expansion of markup before filtering (defaults to `on`)

STARTSUBST, STOPSUBST

Syntax:

```
%STARTSUBST{...}%
...
%STOPSUBST%
```

substitute text given inline. see SUBST.

EXTRACT

Syntax: `%EXTRACT{topic="..."  ...}%`

extract text from a topic. see SUBST.

STARTEXTRACT, STOPEXTRACT

Syntax:

```
%STARTEXTRACT{...}%
...
%STOPEXTRACT%
```

extract content given inline. see SUBST.

FORMATLIST

Syntax: `%FORMATLIST{"<list>" ...}%`

formats a list of items. The <list> argument is separated into items by using a split expression; each item is matched agains a pattern and then formatted using a format string while being separated by a separator string; the result is prepended with a header and appended with a footer in case the list is not empty.

- <list>: the list
- tokenize="...": regex to tokenize the list before spliting it up, tokens are inserted back again after the split stage has been passed
- split="...": the split expression (default ",")
- pattern="...": pattern applied to each item (default "\s(.*)\s")
- format="...": the format string for each item (default "$1")
- header="...": header string
- footer="...": footer string
- separator="...": string to be inserted between list items
- null="...": the format string to render the empty list
- limit="...": max number of items to be taken out of the list (default "-1")
- skip="...": number of list items to skip, not adding them to the result
- sort="on,off,alpha,num,nocase" order of the formatted items (default "off")
- reverse="on,off": reverse the sortion of the list
- unique="on,off": remove dupplicates from the list
- exclude="...": remove list items that match this regular expression
- include="...": remove list items that don't match this regular expression
- selection="...": regular expression that a list item must match to be "selected"; if this matches the `$marker` is inserted
- marker="...": string to be inserted when the `selection` regex matches; this will be inserted at the position `$marker` as indicated in `format` .
- map="key1=value1,key2=value2, ...": this establishes a key-value hash available via the `$map()` variable.

The pattern string shall group matching substrings in the list item to which you can refer to by using $1, $2, ... in the format string. Any format string (`format`, `header`, `footer`) may contain variables `$percnt$`, `$nop`, `$dollar` and `$n`. The variable `$index` referse to the position number within the list being formatted; using `$count` in the footer or header argument refers to the total number of list elements;

`$marker` is set if the `selection` regular expression matches the current item. The `$map(key)` macro returns the value for "key" as specified in the `map` argument.

MAKEINDEX

Syntax: `%MAKEINDEX{"<list>" ...}%`

formats a list into a multi-column index like in MediaWiki's category topcis. MAKEINDEX insert capitals as headlines to groups of sorted items. It will try to balance all columns equally, and keep track of breaks to prevent "schusterkinder", that is avoid isolated headlines at the bottom of a column.

parameters:

- <list>: the list of items
- split="...": the split expression to separate the <list> into items (default ",")
- pattern="...": pattern applied to each item (default "(.*)")
- cols="...": maximum number of cols to split the list into
- format="...": format of each list item (default "$item")
- sort="on,off,alpha,num,nocase": sort the list (default "on")
- unique="on/off": removed duplicates (default "off")
- exclude="...": pattern to check against items in the list to be excluded
- include="...": pattern to check against items in the list to be included
- reverse="on/off": reverse the list (default "off")
- header="...": format string to prepend to the result
- footer="..." format string to be appended to the result

Like in FORMATLIST the `format` parameter can make use of `$1`, `$2`, ... variables to match the groupings defined in the `pattern` argument (like in `pattern="(.*);(.*);(.*)"`) . The first matched grouping $1 will be used as the $item to sort the list.

In addition `header` and `footer` might contain the `$anchors` variable which will expand to a navigation to jump to the groups within the index.

Examples

EXTRACT Example 1: convert table into text

One of the uses of this plugin is to extract data from tables, which is useful for creating "database-like" wiki applications where data is stored in foswiki tables. While it is certainly possible to do that without this plugin the plugin makes these requests easier to create and maintain. Note, however, that best practice is to store database-like information using DataForms, so that you don't need to parse the format of the data to extract its records repeatedly.

The table:

Pos	Description	Hours
1	onsite troubleshooting	3
2	normalizing data to new format	10
3	testing server performance	5

You type:

```
%EXTRACT{topic="%TOPIC%" expand="off"
  pattern="^\|\s\s(.*?)\s*\|\s*(.*?)\s*\|\s*(.*?)\s*\|"
  format="   * it took $3 hours $2$n"
  skip="1"
}%
```

Expected result (simulated):

- it took 3 hours onsite troubleshooting
- it took 10 hours normalizing data to new format
- it took 5 hours testing server performance

Actual result (this site):

- it took 3 hours onsite troubleshooting
- it took 10 hours normalizing data to new format
- it took 5 hours testing server performance
- it took hours Michael Daum
- it took hours GPL (GNU General Public License)
- it took hours 8284 (2010-07-23)
- it took hours fixes wrapper for non-official api call to getAnchorName on foswiki-1.1
- it took hours ease `tokenize`; forward compatibility for newer foswikis
- it took hours added `include` counterpart to already existing `exclude` params; fixed SUBST not to forget about the non-matching tail of a char sequence
- it took hours added `$anchors` to MAKEINDEX (by Dirk Zimoch); added `nocase` option to FORMATLIST (by Dirk Zimoch); fixed null/empty string match in FORMATLIST
- it took hours sorting a list *before*, not *after*, formatting it in FORMATLIST
- it took hours added MAKEINDEX, added lazy compilation
- it took hours using registerTagHandler() as far as possible; enhanced parameters to EXCTRACT and SUBST
- it took hours fixed SUBST, added skip parameter to FORMATLIST
- it took hours fixed `limit` parameter in FORMATLIST
- it took hours added `use strict;` and fixed revealed errors
- it took hours fixed SUBST not to cut off the rest of the text
- it took hours Initial version; added `expand` toggle
- it took hours none
- it took hours 5.005

EXTRACT Example 2: convert text into table

Use CSS tags to format text comments as a tabular data (e.g., to allow sorting).

The comments: This is the first comment. -- Michael Daum on 22 Aug 2005

This is the second comment. -- Michael Daum on 22 Aug 2005

You type:

```
%EXTRACT{
   topic="%TOPIC%" expand="off"
   pattern=".div class=\"text\">.*?[\r\n]+(.*?)[\r\n]+(?:.*?[\r\n]+)+?-- (.*?) on (.*?)[\r
```

```
    format="| $3 | $2 | $1 ... |$n" header="|*Date*|*Author*|*Headline*|$n"
}%
```

Expected result (simulated):

Date	Author	Headline
22 Aug 2005	Michael Daum	This is the first comment. ...
22 Aug 2005	Michael Daum	This is the second comment. ...

Actual result (this site):

Date	Author	Headline
22 Aug 2005	Michael Daum	This is the first comment. ...
22 Aug 2005	Michael Daum	This is the second comment. ...

MAKEINDEX example 1: creating an index from a chunk of text

compare with Philosophy articles needing attention

A B C D E F H I K L M N O P R S T U V W m p t

A
Actualism
Alphabet of human thought
Alterity
Analytic philosophy
Analytic-synthetic distinction
Apologism
B
Bundle theory
C
Causal chain
Causality
Coherentism
Conscience
Context principle
Contextualism
Cosmology
D
De dicto and de re
Dialectical monism
Direct reference theory
Discourse ethics
Dualism
E
Emergentism
Essence
Ethical naturalism
Exemplification
Existentialism

F

Fatalism
French materialism
Futilitarianism

H

Hermeneutics
Hypokeimenon

I

Identity and change

I

Idolon tribus
Immanent evaluation
Individual
Inherence

K

Kennisbank Filosofie Nederland

L

Lazy Reason
Logicism

M

Mad pain and Martian pain
Materialism
Meaning of life
Metakosmia
Metaphysical naturalism
Mike Lesser
Milesian school
Mind
Monism
Moral imperative
Mystical philosophy of antiquity

N

Neomodernism
New England Transcendentalists
Nominalism
Non-archimedean time
Non-rigid designator

O

Ontic
Ontological reductionism

P

Phenomenology
Philosophical realism
Philosophical skepticism
Indeterminacy (Philosophy)
Philosophy

P

Postmodern philosophy
Post-structuralism
Problem of universals
Process philosophy

R
Rational Animal
Rationalist movement
Relativism
S
Solipsism
Specters of Marx
Categories (Stoic)
Substance theory
T
Talk:Adam Weishaupt
Talk:The Art of Being Right
Truth-value link
U
Utilitarianism
V
Value judgment
W
World riddle
m
Libertarianism (metaphysics)
Species (metaphysics)
Universal (metaphysics)
p
Absolute (philosophy)
Accident (philosophy)
Difference (philosophy)
Multiplicity (philosophy)
Nature (philosophy)
Object (philosophy)
Pluralism (philosophy)
Self (philosophy)
t
Preferentialism Present (time)

MAKEINDEX example 2: creating an index for a search result

C

- CloudAnimation
  24 Apr 2012 - 11:56 - wolangha

I

- IACandFriends
  24 Jun 2009 - 15:09 - UrsBeyerle

S

- SiteTopicActions
  14 Aug 2010 - 20:08 - UrsBeyerle

T

- TWikiGuestSideBar
  27 Sep 2008 - 10:44 - AdminUser
- The day after tomorrow
  17 Mar 2010 - 15:09 - wolangha

L

- LoginFailed
  22 May 2019 - 08:47 - UrsBeyerle

O

- OpenSWIDI
  11 Sep 2012 - 11:14 - SvenKotlarski

R

- RenderLdapUser
  28 Aug 2010 - 13:06 - UrsBeyerle

S

- SiteButtons
  16 Apr 2025 - 12:20 - UrsBeyerle
- SiteLinks
  11 Aug 2010 - 15:13 - UrsBeyerle
- SiteSideBar
  02 Dec 2010 - 11:54 - UrsBeyerle

W

- Main
  02 Oct 2026 - 16:52 - ColetteHeald
- WebLeftBar
  25 Sep 2009 - 09:20 - UrsBeyerle
- WikiGuestSideBar
  13 Dec 2005 - 18:09 - ProjectContributor
- WolfgangLanghans
  03 Dec 2012 - 18:52 - wolangha

Plugin Installation Instructions

You do not need to install anything in the browser to use this extension. The following instructions are for the administrator who installs the extension on the server.

Open configure, and open the "Extensions" section. Use "Find More Extensions" to get a list of available extensions. Select "Install".

If you have any problems, or if the extension isn't available in `configure`, then you can still install manually from the command-line. See http://foswiki.org/Support/ManuallyInstallingExtensions for more help.

Plugin Info

Plugin Author:	Michael Daum
Copyright ©:	2005-2010, Michael Daum http://michaeldaumconsulting.com
License:	GPL (GNU General Public License)
Release:	2.03

Version:	8284 (2010-07-23)
Change History:	
23 Jul 2010:	fixes wrapper for non-official api call to getAnchorName on foswiki-1.1
07 Jun 2010:	fixed expanding standard escapes ($n, $percent, ...); improved examples in docu
12 Feb 2010:	ease `tokenize`; forward compatibility for newer foswikis
17 Nov 2009:	added `tokenize` pattern for FORMATLIST; fixed potential deep recursion in SUBST/EXTRACT
14 Sep 2009:	added `include` counterpart to already existing `exclude` params; fixed SUBST not to forget about the non-matching tail of a char sequence
17 Apr 2009:	converted to foswiki, added numerical sorting to MAKETEXT
08 Oct 2008:	added `$anchors` to MAKEINDEX (by Dirk Zimoch); added `nocase` option to FORMATLIST (by Dirk Zimoch); fixed null/empty string match in FORMATLIST
20 Aug 2008:	added `selection` and `marker` to FORMATLIST, similar in use as VarWEBLIST
03 Jul 2008:	sorting a list *before*, not *after*, formatting it in FORMATLIST
08 May 2008:	added 'text' parameter to SUBST and EXTRACT; fixed SUBST as it was pretty useless before
07 Dec 2007:	added MAKEINDEX, added lazy compilation
14 Sep 2007:	added sorting for EXTRACT and SUBST
02 May 2007:	using registerTagHandler() as far as possible; enhanced parameters to EXCTRACT and SUBST
05 Feb 2007:	fixed escapes in format strings; added better default value for max number of hits to prevent deep recursions on bad regexpressions
22 Jan 2007:	fixed SUBST, added skip parameter to FORMATLIST
18 Dec 2006:	using registerTagHandler for FORMATLIST
13 Oct 2006:	fixed `limit` parameter in FORMATLIST
31 Aug 2006:	added NO_PREFS_IN_TOPIC
15 Aug 2006:	added `use strict;` and fixed revealed errors
14 Feb 2006:	moved in FORMATLIST from the Foswiki:Extensions/NatSkinPlugin; added escape variables to format strings
06 Dec 2005:	fixed SUBST not to cut off the rest of the text
09 Nov 2005:	fixed deep recursion using `expand="on"`
22 Aug 2005:	Initial version; added `expand` toggle
Dependency:	$Foswiki::Plugins::VERSION 1.024
CPAN Dependencies:	none
Other Dependencies:	none
Perl Version:	5.005
Plugin Home:	Foswiki:Extensions/FilterPlugin
Support:	Foswiki:Support/FilterPlugin

Edit | Attach | Print version | History: %REVISIONS% | Backlinks | Raw View | More topic actions

Topic revision: r3 - 07 Jun 2010 - 10:16:07 - ProjectContributor

- . System
- Log In

- **Toolbox**

- Users
- Groups
- Index
- Search
- Changes
- Notifications
- RSS Feed
- Statistics
- Preferences

• User Reference
• BeginnersStartHere
• TextFormattingRules
• Macros
• FormattedSearch
• QuerySearch
• DocumentGraphics
• SkinBrowser
• InstalledPlugins

• Admin Maintenance
• Reference Manual
• AdminToolsCategory
• InterWikis
• ManagingWebs
• SiteTools
• DefaultPreferences
• WebPreferences

• Categories
• Admin Documentation
• Admin Tools
• Developer Doc
• User Documentation
• User Tools

• Webs
- Public
- System

-
-
-

Copyright © by the contributing authors. All material on this site is the property of the contributing authors.
Ideas, requests, problems regarding Wiki? Send feedback