EXT: TEQneers SEO Enhancements

Extension Key: tq_seo

Language: en

Keywords: SEO, Sitemap, Pagetitle, Metatags, Google, Yahoo, Bing, Google Analytics, Piwik, P3P

Copyright 2000-2011, Markus Blaschke, <blaschke@teqneers.de> (TEQneers GmbH & Co. KG)

This document is published under the Open Content License

available from http://www.opencontent.org/opl.shtml

The content of this document is related to TYPO3

- a GNU/GPL CMS/Framework available from www.typo3.org

Table of Contents

Introduction

What does it do?

This extensions will handle multiple page metatags and pagetitle manipulations in TYPO3. It's a replacement and major improvement for the "metatags" extension.

The following features are supported:

– Indexed XML sitemap for Google and other search engines (with support for ALL extensions, pibase and extbase, not only tt_news)

– Real sitemap builder via TYPO3 scheduler

– Metatags

– Manipulation of Metatags via stdWrap (Beta)

– Automatic fetching metatags from tt_news (single-display module)

– Pagetitle manipulations

– Manipulation of pagetitle via stdWrap (Beta)

– Multi-language and/or multi-tree TYPO3 installations

The following metatags are supported:

– Language (auto-detected)

– Description

– Keyword

– Copyright information

– Reply-To E-Mail-Address

– Author and publisher

– Rating and distribution

– Revisit

– Geo-location information

– Searchengine Crawler orders

– Google Analytics (with anonymize IP and download link tracking, asynchronous implementation since 4.0)

– Piwik integration (asynchronous implementation since 4.0)

– Google, MSN, Yahoo and WebOfTrust Verification

– Canonical URL (Custom URL or autogenerated to current page)

– IE Compatibility-Mode

– Google Expire Date

– Index, up, prev und next link-metatags

– P3P-HTTP-Headers

The following pagetitle transformations are supported:

– Prefix and suffix (inheritable)

– Absolute pagetitle without any prefix and suffix

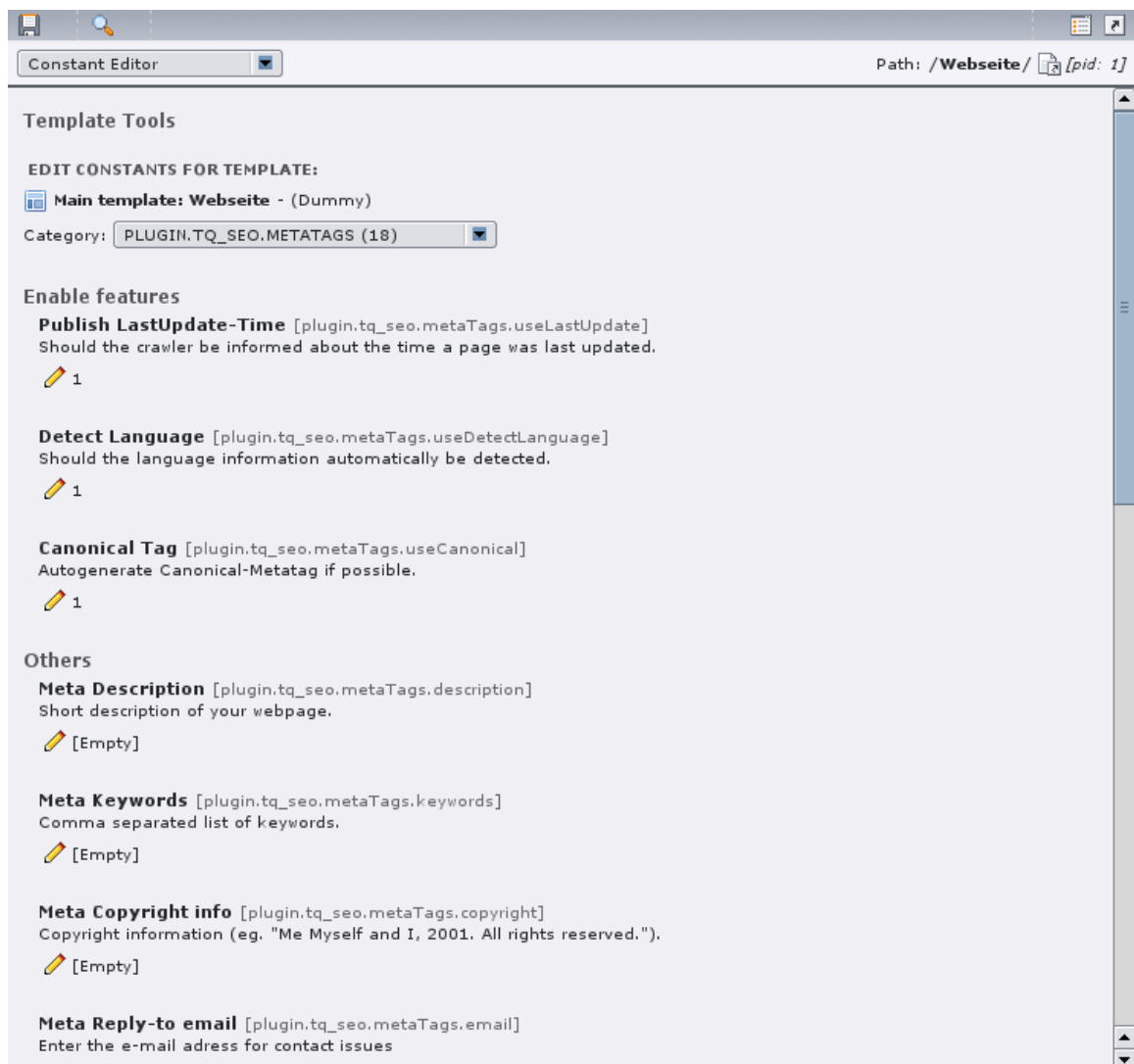– Sitetitle glue options

– Sitetile positon

Screenshots

Illustration 1: TYPO3 Metatag-Constants

Constant Editor ▼   Path: **/Webseite/** [pid: 1]

Template tools

EDIT CONSTANTS FOR TEMPLATE:

Main template: Webseite - (MUST BE REPLACED WITH PROJECT NAME)

Category: PLUGIN.TQ_SEO.CRAWLER (6) ▼

Crawler Index [plugin.tq_seo.metaTags.robotsIndex]
Should the crawler (eg. Google) index the page and subpages
🖉 1

Crawler Follow [plugin.tq_seo.metaTags.robotsFollow]
Should the crawler (eg. Google) follow links
🖉 1

Crawler Archive [plugin.tq_seo.metaTags.robotsArchive]
Should the crawler (eg. Google) put the content of the page into the archive (eg. Google Cache)
🖉 1

Crawler Snippet [plugin.tq_seo.metaTags.robotsSnippet]
Should the crawler (eg. Google) show the snippet in the search result
🖉 1

Crawler ODP [plugin.tq_seo.metaTags.robotsOdp]
Should the crawler (eg. Google) use the description from the OpenDirectoryProject
🖉 1

Crawler YDir [plugin.tq_seo.metaTags.robotsYdir]
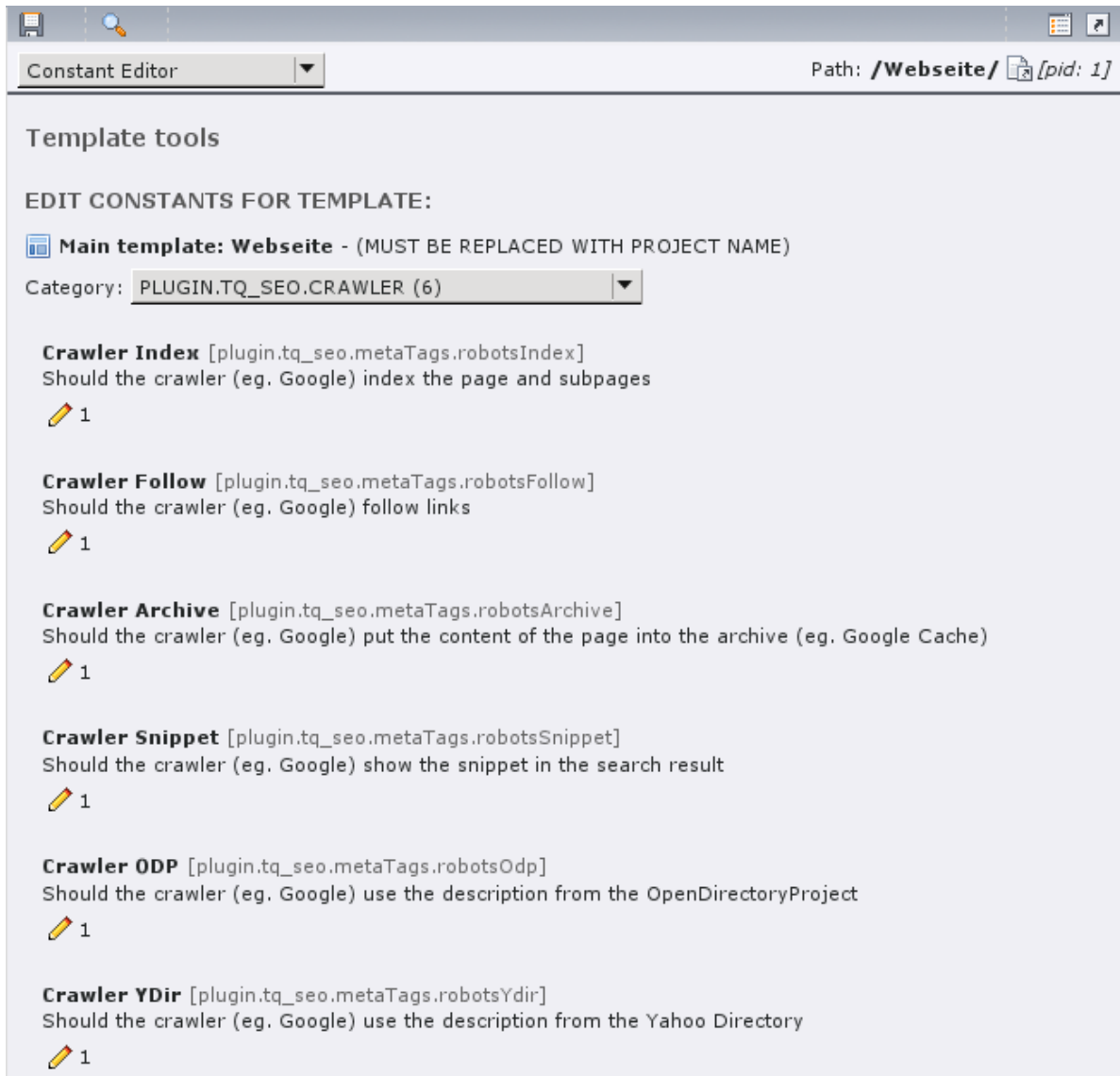Should the crawler (eg. Google) use the description from the Yahoo Directory
🖉 1

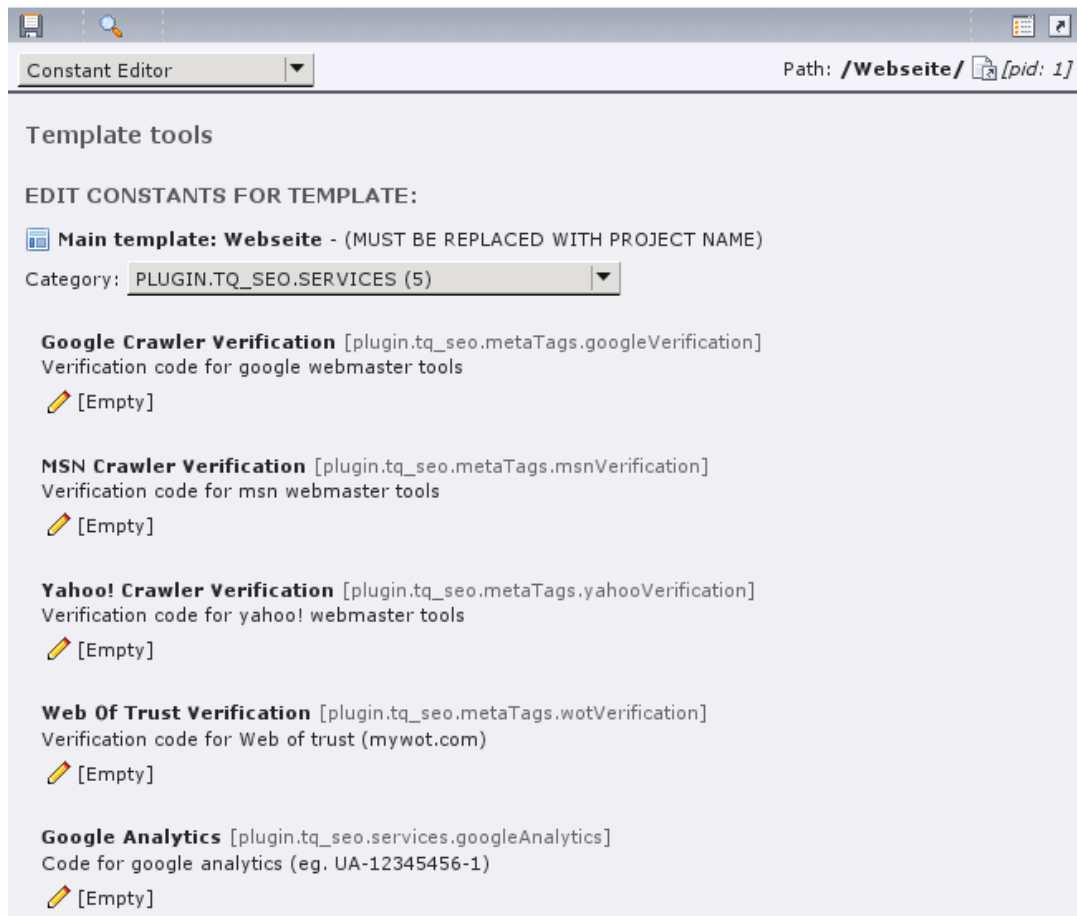*Illustration 2: TYPO3 Crawler-Constants*

Illustration 3: SEO Services (eg. verification codes)

Path: /Webseite/ Products [22]

Edit Page "Products"

| General | Access | Metadata | SEO | Appearance | Behaviour | Resources | Extended |

Pagetitle

<title>-TAG (absolute, overwrites prefix/suffix)

Prefix (inheritable)   Suffix (inheritable)

Inheritance mode

Normal

Searchengines

Exclude page from SearchEngines

Canonical URL (if your content is a copy from another webpage, link here to it)

Sitemap

Priority (0-100)

0

Page change frequency

Default (see constants)

Page [22]

Illustration 4: SEO-pagesettings in TYPO3

7

Users manual

Installation

– Install Extension via Extension Manager.

– Include "static extension template" (Template → Info/Modify → Edit the whole template record → "Include static (from extensions):" and select "TEQneers SEO")

– Modify your metatags via constants editor

Indexed Sitemap

The sitemap will automatically collect all cacheable sites and provides a xml- and plaintext-output – that's why it is "indexed".

The XML-Sitemap (eg. for Google) is available with: index.php?type=841132

The TXT-Sitemap is available with: index.php?type=841131

If you have more than one tree in your TYPO3 you will have to add the root-PID to your Sitemap, eg:

Tree #1 with PID 123: index.php?id=123&type=841132

Tree #2 with PID 234: index.php?id=234&type=841132

If you have also enabled "sitemap_ObeySysLanguage" in the extension configuration you also have to add the language-id for your Sitemap - eg. for seperated language-domain eg. example.com (only english sites) and example.de (only german sites).

The sitemap will index ALL cacheable pages with full extension support (like tt_news and all other "clean" extensions).

If your extension doesn't use cHash or use no_cache the outwill WILL NOT included in the sitemap (and also will not be indexed by index_search).

Also the sitemap indexes all generated "typolink" (BETA).

Warning:

The TQ Seo Sitemap relies on the TYPO3 caching system. If any extension (or configuration – eg. RealURL configuration) break the caching system and makes TSFE non-cacheable (TSFE->no_cache) the sites will NOT INDEXED!

Make sure no extension will set no_cache and the cHash of your link is valid. This is the only way to get only valid URLs into your sitemap.

This sitemap supports both, pibase- and extbase-Extensions without problems. However the developer must take care of the cHash-handling.

Robots.txt

The robots.txt can be gerated with type 841133, eg.:

index.php?type=841133

If possible and enabled the robots.txt buidler will automatically add the link to the sitemap generator or the static sitemap files (will require TYPO3 Scheduler task to generate the static sitemap).

Metatags

Constant	Description:	Default:
Last Update time	Publish the update time of the current page	enabled
Detect Language	Publish the current TYPO3-FE-language as metatag	enabled
Canonical Tag	Publish canonical link if possible (TYPO3-Cache is enabled and cHash is valid) or if user entered canonical tag into page options.	enabled
Canonical Tag (Strict mode)	Enable strict mode (all wrong GET-parameters will generate a canonical-tag to the self without GET-params). eg. /index.php?id=123&foo=bar would generate a Canonical-Tag poiting to /index.php?id=123 if the cHash is wrong or caching is disabled	disabled
Publish Page Expire Time	Publish expire date from the "End Date" of the page. Currently only used for Google.	enabled
Link generation	Automatic generate index and up-links. Next/Prev wil follow in the next version.	enabled
Enable Dublin Core (DC.) metatags	Enable/Disable output of dublin core (DC) metatags	enabled
Description	Default description of your pages. Overwritten by description of page	
Keywords	Default list of keywords Overwritten by keywords of page	
Copyright info	Copyright information of your page	
Reply-to email	E-Mail address for contact	
Author	Default author Overwritten by author of page	
Publisher	Publisher of the website	
Language	Overwrite language detection	
Rating	Rating of the website	
Distribution	Distribution of your website	
Revisit after	Number of days between search engine visits	
Geo Location	Geo-Location of your webpage with latitude, longitude, region and placename	
PICS-label	Platform for Internet Content Selection see http://www.w3.org/PICS/labels.html	
P3P Compact Policy	Your P3P Compact Policy. More informations about P3P: http://www.w3.org/P3P/ http://www.w3.org/TR/P3P/ http://en.wikipedia.org/wiki/P3P http://www.p3pwriter.com/LRN_111.asp	
P3P Policy Url	Link (full URL) to your P3P Policy File	

Some metatags also have markers which could be build in, following metatags supports markers:

– Title

– Description

– Keywords

– Copyright

– Publisher

Following Markers are available:

Marker	Description:	Example:
%YEAR%	Current year	2011

UserAgent

Constant	Description:	Default:
IE Compatibility Mode	Compatibility mode for Microsoft Internet Explorer	

Crawler

Constant	Description:	Default:
Index	Should the crawler index your website?	enabled
Follow	Should the crawler follow links on your website?	enabled
Archive	Is the crawler allowed to archive the page (eg. google cache)	enabled
Snippet	Should the crawler use the snippet/description in search results	enabled
ODP	Should the crawler use the OpenDirectoryProject to display the description in search results	enabled
YDir	Should the crawler use the YahooDirectory to display the description in search results	enabled

Services

Metatag	Description:	Default:
Crawler Verification	Verification code for Google, MSN and Yahoo webmaster tools and Web of trust	
Google Analytics	The google analytics code for using on your site (Will not be shown in frontend if BE-user is logged in, can be re-enabled in BE-Login-Mode: plugin.tq_seo.services.googleAnalytics.showIfBeLogin = 1)	
Google Analytics Cookie Domain Name	If you want to limit the current google analytics to one domain (or subdomain) set the domain name here, eg.: - "auto" (default in google analytics) - "none" - single domain (eg. "example.com") - subdomain (eg. ".example.com") Fore more informations visit: http://code.google.com/intl/de-DE/apis/analytics/docs/gaJS/gaJSApiDomainDirectory.html#_gat.GA_Tracker_._setDomainName	
Google Analytics Anonymize IP	Anonymize the last part of the IP (may be required in some countries)	disabled
Google Analytics Track Downloads (BETA)	Try to track downloads with google analytics. See res/ga-track-download.js for more details Currently supported files: doc,docx,xls,ppt,odt,ods,pdf,zip,tar,gz,txt,vsd,vxd,rar,exe,wma,mov,avi,ogg,ogm,mkv,wmv,mp3,webm	disabled
Piwik URL	Url to your Piwik installation (without http:// and https://)	
Piwik ID	Tracking id of your website in your piwik	

Metatag	Description:	Default:
Piwik Download & Click Domain Name	Specifies which domains are internal domains: - single domain (eg. "example.com") - subdomain (eg. ".example.com") For more informations visit: http://piwik.org/docs/javascript-tracking/	
Piwik Cookie Domain Name	Specifies the domain name for the tracking cookie: - single domain (eg. "example.com") - subdomain (eg. ".example.com") For more informations visit: http://piwik.org/docs/javascript-tracking/	
Piwik DoNotTrack	Opt Out users with Mozilla's DoNotTrack browser setting	enabled (since 4.0)

Pagetitle

Metatag	Description:	Default:
Apply tmpl-sitetitle to absolute <title>	There is a prefix/suffix for your pagetitle defined in your root template settings. If you use the SEO-Absolute-Pagetitle settings you can disable this suffix/prefix here.	disabled (since 1.1.0)
Apply tmpl-sitetitle to prefix/suffix <title>	There is a prefix/suffix for your pagetitle defined in your root template settings. If you use the SEO-Pagetitle-Suffix/Prefix settings you can disable this suffix/prefix here.	enabled
Sitetitle glue	Glue between Pagetitle (from Page) and Sitetitle (from template)	:
Sitetitle glue spacer (before)	Add spacer before glue string	disabled
Sitetitle glue spacer (after)	Add spacer after glue string	enabled
Sitetitle position	Position of Sitetitle (from template) Possible options: Sitetitle-Pagetitle (eg. Example Company: About us) Pagetitle-Sitle (eg. About us: Example Company)	Sitetitle-Pagetitle (0)

Sitemap

Metatag	Description:	Default:
Enable	Enables output (if set on root-pid of tree) and indexing for the whole subtree	enabled
Page limit	Limit pages on sitemap-xml-pages	10000
Limit to current language	Limit output of the sitemap to the current language. This will enable multi-language-domain sitemaps. eg: www.example.com (FE-Language is english) will output only english pages www.example.de (FE-Language is german) will output only german pages This option was ported from the extension configuration and will replace this configuration.	disabled
Default change frequency	Default change frequency for sitemap cache (will be cached!)	
Page priority	Default page priority if the page have no own priority set Page priority will be calculated by: ([page priority] – [priority modificator]) * (1/[page depth] * [page multiplier])	1
Page priority depth multiplier	Page depth multiplier, see formula in page priority	1
Page priority depth modificator	Page depth modificator, see formula in page priority	1

Advanced manipulations (stdWrap support)

If you want to modify some things you can use stdWraps

Metatags

TypoScript Node	Description:	Type
plugin.tq_seo.metaTags.stdWrap.title	Manipulation for title	stdWrap
plugin.tq_seo.metaTags.stdWrap.description	Manipulation for description	stdWrap
plugin.tq_seo.metaTags.stdWrap.keywords	Manipulation for keywords	stdWrap
plugin.tq_seo.metaTags.stdWrap.copyright	Manipulation for copyright	stdWrap
plugin.tq_seo.metaTags.stdWrap.language	Manipulation for language	stdWrap
plugin.tq_seo.metaTags.stdWrap.email	Manipulation for email	stdWrap
plugin.tq_seo.metaTags.stdWrap.author	Manipulation for author	stdWrap
plugin.tq_seo.metaTags.stdWrap.publisher	Manipulation for publisher	stdWrap
plugin.tq_seo.metaTags.stdWrap.distribution	Manipulation for distribution	stdWrap
plugin.tq_seo.metaTags.stdWrap.rating	Manipulation for rating	stdWrap
plugin.tq_seo.metaTags.stdWrap.lastUpdate	Manipulation for last update (date)	stdWrap

Pagetitle

TypoScript Node	Description:	Type
plugin.tq_seo.pageTitle.stdWrap.before	Manipulation of the raw page title (before TQ-Seo processing)	stdWrap
plugin.tq_seo.pageTitle.stdWrap.after	Manipulation of the processed page title (after TQ-Seo processing)	stdWrap

Google Analyitics / Piwik customizations

Customization codes (fast/simple)

You can add additional javascript code to the default Google Analytics and/or Piwik integration.

TypoScript Node	Description:	Type
plugin.tq_seo.services.googleAnalytics.customizationCode	Customization Code for Google Analytics	TS-Content Object (TEXT, COA, ...)
plugin.tq_seo.services.piwik.customizationCode	Customization Code for Piwik	TS-Content Object (TEXT, COA, ...)

Example for Google Analytics in TypoScript-Setup:

```
plugin.tq_seo.services.googleAnalytics.customizationCode = COA
plugin.tq_seo.services.googleAnalytics.customizationCode {
  10 = TEXT
  10.value (
_gaq.push(['_setClientInfo', false]);
_gaq.push(['_setAllowHash', false]);
_gaq.push(['_setDetectFlash', false]);
_gaq.push(['_setDetectTitle', false]);
)
}
```

Template customization (advanced)

The Google Analytics and Piwik integration is done by using a FLUIDTEMPLATE object in TypoScript. If you don't like the integration and want to modify the integration feel free to use your own templates and pass your own variables to FLUIDTEMPLATE.

TypoScript Node	Description:	Type
plugin.tq_seo.services.googleAnalytics.template	Template rendering object for Google Analytics (contains a FLUIDTEMPLATE)	COA
plugin.tq_seo.services.piwik.template	Template rendering object for Piwik (contains a FLUIDTEMPLATE)	COA

It's quite easy, for more information read:

– http://forge.typo3.org/projects/typo3v4-mvc/wiki/FLUIDTEMPLATE%20Content%20Object

– http://typo3.org/development/articles/the-fluidtemplate-cobject/

Example for your own Google Analytics Template:

```
## Google Analytics template
plugin.tq_seo.services.googleAnalytics.template.10.file = fileadmin/templates/service-ga.html

## if you need some variables you also can set these:
plugin.tq_seo.services.googleAnalytics.template.10.variables {
  myOwnStuff = TEXT
  myOwnStuff.value = foobar
}
```

Scheduler tasks

Task	Description:	Frequency
TQ SEO Cleanup	This task cleans up old database entries in the tx_tqseo_sitemap table.	One run per day
TQ SEO sitemap.txt builder	This task builds a real sitemap.txt file in the upload directory. Directory: uploads/tx_tqseo/sitemap_txt/ Sitemap: sitemap-r{ROOTPID}.txt.gz If language domain support is active: Sitemap: sitemap-r{ROOTPID}-l{LANG}.txt.gz {ROOTPID} is the Page-UID from the root pages in your TYPO3 installations. {LANG} is the language id (only active if language domains are active). Hint: These files are already gziped.	One run per day
TQ SEO sitemap.xml builder	This task builds a real sitemap.xml files in the upload directory. Directory: uploads/tx_tqseo/sitemap_xml/ Sitemap-Index: index-r{ROOTPID}.xml.gz Sitemap-Page: sitemap-r{ROOTPID}-p{PAGE}.xml.gz If language domain support is active: Sitemap-Index: index-r{ROOTPID}-l{LANG}.xml.gz Sitemap-Page: sitemap-r{ROOTPID}-l{LANG}-p{PAGE}.xml.gz {ROOTPID} is the Page-UID from the root pages in your TYPO3 installations. {PAGE} is the current page of the sitemap. {LANG} is the language id (only active if language domains are active). The index will refer to all page sitemaps so you only have to reference to the sitemap index. Hint: These files are already gziped.	One run per day

FAQ

General

Problem: Is this extension mutli-language and/or multi-tree ready?

Solution: Check it out, there should be no problems at all with multi-language multi-tree TYPO3-installations.

Problem: In my TYPO3 the feature XYZ of this extension doesn't work. What's wrong?

Solution: The tq_seo-Extension was carefully written. We're at TEQneers what to deliver correct pages and extensions will the power of the TYPO3 caching system (and those extensions are the fastest). For some features (eg. sitemap, canonical-tag) we have to trust in the caching-system – it's the only way to make sure that we don't deliver wrong informations to our search engines. If our extension doesn't work correct in your TYPO3 installation maybe there is something wrong – maybe the configuration, your setupTS or one of your extension.

Indexed Sitemap

Problem: The sitemap is empy.

Solution: Check if you created the "tx_tqseo_sitemap_pages" database-table! We're using InnoDB as MySQL-Engine so you have to make sure that you have InnoDB-Support in your Hosting-Account (if not you should switch your hoster, InnoDB is a Standard-Feature in MySQL). As with 2.0.1 there is an error-message if the table doesn't exist.

Problem: My sitemap is still empty, but the database is ok.

Solution: Maybe you disabled the whole TYPO3 cache (config.no_cache=1 or somehthing else)? Enable the cache – this will also speed up your TYPO3 installation.

Problem: The generated pages from my extension are not included in the sitemap!

Solution: You have to make sure that all generated pages are cacheable. The extension has to pass a valid cHash-token!

Problem: There is only one site in my XML-Sitemap, what's wrong? In the sitemap-database are thousands of pages.

Solution: That's correct. TQ_Seo always uses "Sitemap Groups" (as defined in http://www.sitemaps.org/protocol.php#index), each group can contain about 50.000 pages/urls so if we have more than 50.000 urls we have to use sitemap groups. Visit the url defined in the LOC-tag and you will see that the sub-sitemap will contain all your stored URLs.

Problem: The generated pages from my extension still are not included in the sitemap!

Solution: You've enabled RealURL? Please check the RealURL configuration if you have specified values that are not passed with your url. All variables that are not passed with the url will result in a NO-CACHE.

Problem: The sitemap is still not working! No page is indexed and the table tx_tqseo_sitemap_pages is empty!

Solution: Double check your installation and disable all third-party extensions, make sure that no extension disables the TYPO3-cache! RealUrl (if properly configured) and TemplaVoila are working wonderfull with TQ Seo Sitemap but some old

extensions might break the TYPO3 caching system and you will not notice it. Our sitemap indexer relies on the indexing system to make sure that only valid urls are stored and delivered to the search engines like google.

Problem: I want to limit each sitemap to it's domain (eg. example.com for english pages, example.de for german pages). Is this possible?

Solution: Yes, just enable the "Enable language-domain support" in seo control center in your website/rootpage settings. (replaces the old extension configuration sitemap_ObeySysLanguage and TypoScript constants setting).

Problem: My tt_news entries are not indexed, what's wrong?

Solution: You're using realurl? Then check your realurl_conf.php. Errors (or misconfiguration) in the realurl-configuration will produce uncacheable sites (and you will not notice it). The most common issue is the configuration of the parameter "tx_ttnews[swords]" in the postVarSets-area. Remove it, you don't need it.

Others

Problem: I want to customize my google Analytics and/or piwik integration.

Solution: You can modify the code of the google analytics and piwiki integration with typoscript. Feel free to use *plugin.tq_seo.services.googleAnalytics.customizationCode* (STDWRAP) or *plugin.tq_seo.services.piwik.customizationCode* (STDWRAP). Also you can modify the FLUIDTEMPLATE and assign custom variables.

RealURL configuration

If you want to activiate "real" sitemap.xml feature (eg. http://example.com/sitemap.xml), configure realurl like this:

```php
<?php
$TYPO3_CONF_VARS['EXTCONF']['realurl']['_DEFAULT'] = array(

'init' => array(
    // ...
),

'preVars' => array(
    // ...
),

'fixedPostVars' => array(
    // ...
),

'postVarSets' => array(
    '_DEFAULT' => array(

        // TT-NEWS (example configuration)
        'date' => array(
            array(
                'GETvar' => 'tx_ttnews[year]' ,
            ),
            array(
                'GETvar' => 'tx_ttnews[month]' ,
                'valueMap' => array(
                    'january' => '01',
                    'february' => '02',
                    'march' => '03',
                    'april' => '04',
                    'may' => '05',
                    'june' => '06',
                    'july' => '07',
                    'august' => '08',
                    'september' => '09',
                    'october' => '10',
                    'november' => '11',
                    'december' => '12',
                ),
            ),
            array(
                'GETvar' => 'tx_ttnews[day]',
            ),
        ),

        // news pagebrowser
        'browse' => array(
            array(
                'GETvar' => 'tx_ttnews[pointer]',
            ),
        ),

        // news categories
        'news-category' => array (
            array(
                'GETvar' => 'tx_ttnews[cat]',
                'lookUpTable' => array(
                    'table' => 'tt_news_cat',
                    'id_field' => 'uid',
                    'alias_field' => 'title',
                    'addWhereClause' => ' AND NOT deleted',
                    'useUniqueCache' => 1,
                    'useUniqueCache_conf' => array(
                        'strtolower' => 1,
                        'spaceCharacter' => '-',
                    ),
                ),
            ),
        ),

        // news articles
        'article' => array(
            array(
                'GETvar' => 'tx_ttnews[tt_news]',
                'lookUpTable' => array(
                    'table' => 'tt_news',
                    'id_field' => 'uid',
                    'alias_field' => 'title',
                    'addWhereClause' => ' AND NOT deleted',
                    'useUniqueCache' => 1,
```

17

```php
                'useUniqueCache_conf' => array(
                    'strtolower' => 1,
                    'spaceCharacter' => '-',
                ),
            ),
        ),

        // ... other extensions ...
    ),
),

'fileName' => array(
    'defaultToHTMLsuffixOnPrev' => 1,
    'index' => array(
        // ...

        'sitemap.xml' => array(
            'keyValues' => array(
                'type' => 841132,
            ),
        ),

        'sitemap.txt' => array(
            'keyValues' => array(
                'type' => 841132,
            ),
        ),

        'robots.txt' => array(
            'keyValues' => array(
                'type' => 841133,
            ),
        ),

        '_DEFAULT' => array(
            'keyValues' => array(
                'type' => 0,
            )
        ),

    ),
),

'pagePath' => array(
    'type' => 'user',
    'userFunc' => 'EXT:realurl/class.tx_realurl_advanced.php:&tx_realurl_advanced->main',
    'spaceCharacter' => '-',
    'segTitleFieldList' => 'tx_realurl_pathsegment,alias,nav_title,title',
    'languageGetVar' => 'L',
    'expireDays'  => 30,
    'rootpage_id' => 1,
),

);
?>
```

Hooks

All hooks are defined in docs/examples/hooks.php (available with short description).

Beta-Features

- TypoLink Sitemap parser

Suggested Extensions

- sourceopt: HTML sourcecode optimization
- realurl/cooluri: Link SEO optimization

Known problems

– Don't use another SEO-extension like metatags

– Sourceopt might break sitemap-txt output, xml is still fine

To-Do list

– Backend Interface for easy metatag editing

– RealURL autoconfiguration

Found a bug? Problems?

The bugtracker can be found on TYPO3-forge:

http://forge.typo3.org/projects/extension-tq_seo

.. or contact us at blaschke@teqneers.de or info@teqneers.de.

ChangeLog

– see Changelog-File