

Drupal in der Praxis: Google News importieren mit Feeds und Feedspammers

Auf der Webseite <http://www.berggenuss.de> werden im Zuge der Umstellung auf Drupal 7 mehrere Google Alerts zu verschiedenen Themen der Berglandwirtschaft eingerichtet.

Wir legen dazu einen eigenen Inhaltstyp an, der es erlaubt, diese News zu Themen bzw. Alpen-Regionen und Bergprodukten zuzuordnen.
Um die Zuordnung nicht komplett händisch vornehmen zu müssen, richten wir eine Teilautomatische Zuordnung beim Import ein.

Folgende Grundeinstellungen wurden gemacht:

Abstand Pixel

30

Grundeinstellungen

Attached to: [keine] [Einstellungen](#)
 Periodic import: every 30 Minuten
 Import on submission

Abrufer [Ändern](#)

HTTP Fetcher [Einstellungen](#)
 Download content from a URL.

Parser [Ändern](#)

Common syndication parser [Einstellungen](#)
 Parse RSS and Atom feeds.

Prozessor [Ändern](#)

Node processor [Einstellungen](#) [Zuordnung](#)
 Create and update nodes.

Grundeinstellungen

Name *

 A human readable name of this importer.

Beschreibung

 A description of this importer.

Attach to content type

 If "Use standalone form" is selected a source is imported by usir

Periodic import

 Choose how often a source should be imported periodically. [Re](#)

Import on submission
 Check if import should be started at the moment a standalor

Process in background
 For very large imports. If checked, import and delete tasks s
 cron task in any case. [Requires cron to be configured.](#)

Abstand Pixel

Als Parser verwenden wir den Common syndication parser.

Einstellungen des Node-Parsers.

Zum Testen stellen wir 'Update existing nodes' ein und 'Skip hash check'. Erfahrungsgemäß werden wir die gleichen News mehrmals einlesen, bis die Feineinstellungen stimmen.

Als Inhaltstyp wählen wir den, der vorher angelegt wurde.

Abstand Pixel

30

Settings for Node processor

Update existing beiträge

- Do not update existing nodes
- Replace existing nodes
- Update existing nodes (slower than replacing them)

Existing beiträge will be determined using mappings that are a "unique target".

Skip hash check

Force update of items even if item source data did not change.

Textformat *

Filtered HTML ▼

Select the input format for the body field of the nodes to be created.

Inhaltstyp

News Artikel ▼

Select the content type for the nodes to be created. **Note:** Users with "import google_rss_feed_berglawirtsch" level permissions. Further, users with "clear google_rss_feed_berglawirtschaft permissions" will be able to d

Autor

Montviso ○

Select the author of the nodes to be created – leave empty to assign "anonymous".

Autorisieren

Check that the author has permission to create the node.

Expire nodes

Nie ▼

Select after how much time nodes should be deleted. The node's published date will be used for determining

Speichern

Abstand Pixel

Nun kommen wir zum Mapping.

Hier werden vom Feeds Importer bereits die richtigen Vorschläge gemacht. Einzig die Zuordnung von Beschreibung zu Body fehlt.

Diese fügen wir händisch hinzu.

Die Einstellung unique bei der URL sorgt dafür, daß News mit der gleichen URL aktualisiert werden, statt ein zweites Mal eingelesen.

Das ist sehr wichtig, um unerwünschte Duplikate zu vermeiden.

Abstand Pixel

30

QUELLE	TARGET	TARGET CONFIGURATION	Zeilenreihenfolge anzeigen
+ Titel	Titel	Not used as unique. 	<input type="checkbox"/> Entfernen
+ Erschienen am	Erschienen am		<input type="checkbox"/> Entfernen
+ Item URL (link)	URL	Used as unique. 	<input type="checkbox"/> Entfernen
+ Item GUID	GUID	Used as unique. 	<input type="checkbox"/> Entfernen
+ Beschreibung	Fehlt		<input type="checkbox"/> Entfernen
+ Beschreibung	Body		<input type="checkbox"/> Entfernen
Select a source ▾	Select a target ▾		<input type="button" value="Hinzufügen"/>

LEGENDE

Abstand Pixel

Jetzt kommt der erste Test:

Dazu rufen wir die Seite

domainname/import auf und wählen unseren zuvor angelegten Feed.

Wenn wir ihn anklicken wird nach der URL gefragt, von der wir die News beziehen möchten.

In unserem Fall ist es

<http://news.google.de/news?pz=1&cf=all&ned=de&hl=de&q=Berglandwirtschaft...>

Natürlich kann man dem Parameter 'q' auch Begriffe in Kombination übergeben, wie z.B. Bergkäse+Alpen.

Wenn Begriffe mit Umlauten oder anderen Sonderzeichen verwendet werden, dann müssen diese umgewandelt sein.

Das geht ganz fix wenn man die URL mit den gewünschten Begriffen einmal im Browser-URL-Fenster einfütert.

Danach steht oben die korrekte URL die man nun dem Importer mit teilt.

Nach Durchführung des Imports sieht das dann so aus:

Abstand Pixel

30



- Neuer Alias [artikel/wo-das-edelweiss-aus-fichtenholz-blueht-alm-herbst-besuch-beim-finanznachric](#)
- Neuer Alias [artikel/aelperleben-auf-luen-vorarlberg-online-25879.html](#) für [node/25879](#) erstellt.
- Neuer Alias [artikel/rettung-migros-tochter-pachtet-sennaria-surselva-zu-schweizerbauerch-25880.ht](#)
- Neuer Alias [artikel/ein-langer-almsommer-geht-zu-ende-aktiv-verzeichnisde-pressemittellung-2588](#)
- Neuer Alias [artikel/offizielles-label-fuer-bergkaese-umstritten-st-galler-tagblatt-25882.html](#) für [node/](#)
- Neuer Alias [artikel/ein-himmlicher-anblick-vorarlberger-nachrichten-25883.html](#) für [node/25883](#) ers
- Neuer Alias [artikel/so-laesst-sich-der-sommer-aushalten-fotos-patrick-schluetter-neue-westfaelische](#)
- Neuer Alias [artikel/cool-idee-aus-allgaeuer-milch-gemeinde-betreibt-eisdiele-aktiv-verzeichnisde-](#)
- Created 8 beiträge.

Google RSS Feed Berglandwirtschaft

Status

Last import: 0 Sek. ago.
18 imported items total.

Importieren

URL *

Enter a feed URL.

IMPORTIEREN

Abstand Pixel

30
 In Log kann man auch noch mal sehen, ob alles geklappt hat:
[domainname/admin/reports/feeds](#)

Und natürlich schauen wir gleich ins Frontend, wo die News ausgegeben werden sollen:

Abstand Pixel

30

Abstand Pixel

30
 sieht schon mal gut aus.

Besucher der Webseite sollen die News filtern können und wie oben bereits beschrieben, wurde der Inhaltstyp mit entsprechenden Feldern versehen.

Qualitätssicherung:

Die News sollen vor dem Einstellen gesichtet und nur freigeschaltet werden, wenn sie thematisch auf die Webseite passen.

Deshalb wurde im Inhaltstyp News Artikel der Haken bei 'Veröffentlicht' entfernt.

Um nicht alle Tags händisch pflegen zu müssen, werden einige Vorbelegungen gemacht und für bestimmte Themen Standardwerte im Feeds-Import gemacht.

Dazu benötigt man das Modul Feeds Tamper.

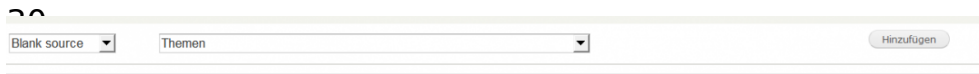
Nehmen wir als Beispiel den Alert für 'Genusswandern + Alpen'.

Wir unterstellen, daß die News zum Thema 'Sanfter Tourismus' passt und entsprechend getaggt werden soll.

Die Taxonomy für dieses Thema hat die ID 368.

Im Mapping des Moduls Feed für diesen Import fügen wir ein Feld hinzu, das so aussieht:

Abstand Pixel



The screenshot shows a form with a dropdown menu set to 'Blank source', a text input field containing '30', and a dropdown menu set to 'Themen'. A 'Hinzufügen' button is visible to the right of the input field.

Abstand Pixel

Nun gehen wir zurück in die Feeds Übersicht und wählen bei diesem Import zum Thema 'Genusswandern + Alpen' den Link auf Tamper.

In der Tamper-Konfiguration sehen wir alle Felder wieder, die wir gemappt haben. Uns interessiert davon nur das Feld für 'Thema', wo wir auf 'Plugin hinzufügen' klicken.

In der folgenden Maske suchen wir oben das Plugin 'Default Value' und verpflegen den Rest wie hier im Bild ersichtlich:

Abstand Pixel

30

The plugin to add

Set default value ▼

Beschreibung *

Default für Thema M

A useful description of what this plugin is doing.

CONFIGURE SET DEFAULT VALUE

Standardwert

368

This field will be set to the value specified.

Hinzufügen

Abstand Pixel

Nachdem wir die vier neu eingetragenen Alerts kurz gesichtet und danach freigegeben haben, erscheinen sie im Frontend mit dem Zusatz: 'Thema: Sanfter Tourismus.'

Abstand Pixel

??

Verwöhnhotel Berghof in den Salzburger Bergen - X-ACES

Verwöhnhotel Berghof in den Salzburger Bergen

X-ACES

Egal, ob jemand nur ein paar Schritte bis zur nächsten Alm **wandern** möchte oder eine ausgedehnt alpine Touren für Bergerfahrene beginnen ebenso vor der Hoteltür wie **Genusswanderungen** für c
...

Themen: [Sanfter Tourismus](#)

Abstand Pixel

Benso können wir den Alert von Berglandwirtschaft auf selbiges Thema legen und Bergkäse zum Thema 'Bergprodukte' Taggen.

Damit ist ein Teil der Thematischen Zuordnung bereits erledigt.

Das Tamper-Modul kann mit einem Plugin ergänzt werden, welches erlaubt, jeglichen PHP-Code beim Import auszuführen.

Es ist angedacht, die Inhalte der News zu durchsuchen und auf Vorkommen bestimmter Länder- oder Regionen- Begriffe zu prüfen, die dann automatisch an der News gepflegt werden können.

Dazu brauchts dann aber etwas mehr Zeit....;-)

Der Import der RSS Feeds kann zwar händisch erfolgen, er passiert aber auch mit dem regelmäßigen Cronlauf automatisiert.