



Andrea Müller, Noud van Krusbergen

# Slimme vergaarbak

## Nieuws verzamelen en filteren met Tattler

**Zelfs als je maar van een paar onderwerpen op de hoogte wilt blijven, heb je al snel tientallen RSS-feeds die je in de gaten moet houden. Tattler neemt je een deel van het werk uit handen door alleen het echt interessante nieuws eruit te filteren.**

In de Hannibal-serie van schrijver Thomas Harris komt de tabloid National Tattler voor, waarvan de reporter Freddy Lounds aan een vreselijk einde komt in een brandende rolstoel. De webapplicatie Tattler is minder opwindend, maar wel informatiever dan het fictieve roddelblad. Als je Tattler eenmaal geïnstalleerd hebt, doorzoekt de software het web en sociale diensten op berichten over de opgegeven onderwerpen en komen alleen relevante berichten door het filter heen.

De nieuwsaggregator is in eerste instantie bedoeld als hulpje voor de mediabranche om er voor te zorgen dat die alleen de belangrijkste berichten onder ogen krijgen. Zodoende hoeven ze geen bergen RSS-feeds bij te houden waartussen maar sporadisch iets interessants zit. Het enige wat je hoeft te doen, is Tattler de zoekwoorden mee te geven waarin je geïnteresseerd bent. Vervolgens kijkt Tattler met behulp van een cron-job regelmatig op internet en bij web 2.0-diensten of het bijbehorende berichten kan vinden. Die krijg je niet allemaal te zien, maar alleen die sites met de grootste dichtheid aan zoekwoorden.

Tattler is gebaseerd op de code van het open-source contentmanagementsysteem Drupal, waar ook de installer van geleend is. Om Tattler te installeren moet je een Linux- of Windows-systeem met een Lamp-stack hebben. PHP moet minstens versie 5.2 zijn en van MySQL moet ook een 5-versie aanwezig zijn. Onder Windows kun je dat mooi doen met het installeren van de Xampp-distributie, die alles bevat wat je nodig hebt (zie de softlink aan het eind van dit artikel). De ontwikkelaars adviseren om te installeren als subdomein, omdat de subfolder-installatie nog niet voldoende getest is. In de praktijk hebben we daar echter geen problemen mee ondervonden. Met PHP 5.3.5 deden zich wel wat bijverschijnselen voor, die redelijk eenvoudig op te lossen waren, maar helemaal vertrouwen deden we het niet. Vandaar dat we teruggingen naar PHP5.2.9, waarna het systeem vlekkeloos werkte.

### Voorbereidingen

Er zijn geen kant-en-klare pakketten voor Tattler, zodat je een gecomprimeerd tar- of zip-

archief van de website moet downloaden. De inhoud daarvan – de map tattler – pak je uit in de folder htdocs. Bij Linux staat die bij bijna alle distributies in de map /var/www. OpenSuse en Suse Linux Enterprise gebruiken in plaats daarvan de map /srv/www. Bij Windows staat de folder htdocs in de Xampp-map.

Voordat je het installatiescript kunt aanroepen, moet je een database voor Tattler aanmaken alsook een gebruiker die daar volledige toegang toe heeft. Als je het pakket op een server bij een provider installeert, heb je meestal een grafisch front-end als phpMyAdmin om dat te kunnen doen. Je kunt phpMyAdmin natuurlijk ook lokaal installeren, maar aanmaken van de database en gebruiker gaat eigenlijk net zo snel via de console. Begin met het volgende commando:

```
mysql -u root -p
```

Hiermee start je het databasebeheer met de gebruiker root en maak je vervolgens de database tattler aan:

```
create database tattler;
```

Met het volgende commando maak je de databasegebruiker 'tattleruser' aan met het wachtwoord 'geheim' en geef je hem volledige toegang tot de zojuist aangemaakte database:

```
grant ALL ON tattler.* to tattleruser@localhost
identified by 'geheim';
flush privileges;
```

Bij het uitvoeren van het installatiescript moet het bestand settings.php in de map

Bij de tweede stap van de installatie geef je aan welke database je voor Tattler aangemaakt hebt.

tattler/sites/default voor iedereen beschrijfbaar zijn, omdat de installer anders afbreekt met een foutmelding. Geef in de tattler-map dan ook het volgende commando:

```
chmod 666 sites/default/settings.php
```

Maak dan nog de map 'files' aan in tattler/profiles en geef daar ook iedereen schrijfrechten voor:

```
chmod 777 tattler/profiles/files
```

De modus 777 is nodig omdat het execute-bit gezet moet zijn om überhaupt in de map te kunnen komen.

## Tattler instellen

Om het installatiescript van Tattler te starten, ga je in de browser naar de pagina <http://tattler.mijndomein.nl>. Het script test eerst of PHP en MySQL in de juiste versie geïnstalleerd zijn en vraagt bij de volgende stap om de gegevens van de database. Geef daar de naam van de eerder aangemaakte database op, evenals de naam van de gebruiker en diens wachtwoord.

Dan volgen de instellingen voor de Tattler-pagina. Geef het adres op waarmee Tattler bereikbaar zal zijn, bijvoorbeeld [tattler.mijndomein.nl](http://tattler.mijndomein.nl), evenals het e-mailadres van de websitebeheerder. Hij is de enige die zich na het installeren bij het systeem kan aanmelden.

De volgende stap is het opgeven van API-sleutels. Dit is optioneel. Hier kun je API-sleutels invullen van webdiensten als Flickr en Compete, die Tattler later dan kan gebruiken. Tattler haalt foto's van Flickr die bij de onderwerpen horen en gebruikt de Compete-key voor het grafisch weergeven van trends. Deze API-keys zijn gratis en aan te vragen via de links op deze webpagina. De link naar de



Technorati-website werkt echter niet omdat deze dienst op dit moment geen sleutels uitdeelt. Als je op 'Save and continue' klikt, wordt de installatie afgerond. Vervolgens krijg je een link naar je site.

## Onderwerpen

Als je de eerste keer contact maakt met de webdienst, krijg je een bijna lege pagina te zien. Dat is te verhelpen door een nieuw thema aan te maken. Dat kan het snelst door op de tekstballon naast het Tattler-pictogram linksboven te klikken. Dat opent een formulier waarmee je via 'Add a New Topic' invoervelden krijgt voor de titel van het thema, de zoekwoorden (topic keywords) en de zogeheten user tags. Dat zijn tags die Tattler automatisch aan berichten toevoegt die bij het onderwerp passen.

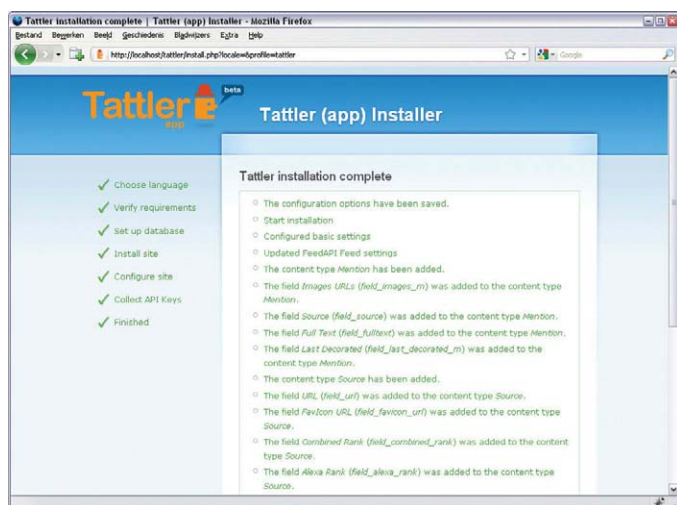
Zo kun je met de tag `lang:dutch` bijvoorbeeld een tag voor voornamelijk Nederlandstalige berichten instellen, als dat voor de trefwoordenlijst van toepassing is tenminste. Als je bijvoorbeeld berichten over het thema 3D-tv wilt hebben, stel je met `track:3d-tv` de bijbehorende tag in. De user tags kun je later als zoekterm

voor de ingebouwde zoekfunctie rechtsboven gebruiken om de resultaten te filteren.

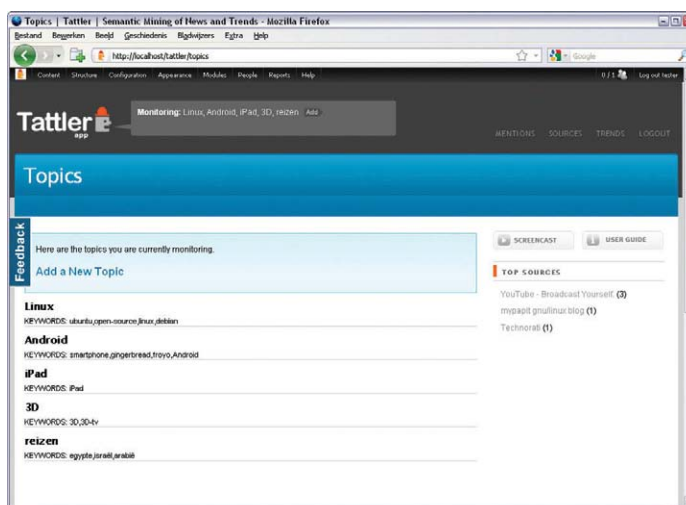
Als je een thema hebt aangemaakt, moet je de cron-job van Tattler starten om de eerste berichten op te halen. Het script voor het bijwerken heet `cron.php` en staat in de `tattler-map`. Dat kan ook via het administratormenu linksboven met 'Run cron'. Via de commandline kun je de cron-job bijvoorbeeld starten met

```
curl --silent http://tattler.mijndomein.nl/cron.php
```

Als je nu een foutmelding krijgt omdat de commandline-downloader `curl` niet op je systeem aanwezig is, gebruik je in plaats daarvan `wget`, dat standaard bij de meeste Linux-distributies wordt meegeleverd. Klik na het aanroepen van de Tattler-site op de link 'All' om de eerste berichten te zien. Je kunt ook alleen naar de berichten van de laatste week, maand of van het hele jaar kijken. Via de tabbladen direct boven de berichten kun je filteren op berichtentype, waarbij je de keuze hebt uit News, Tweets en Multimediacontent. Bij de laatste mogelijkheid staan ook foto's van Flickr die bij het thema passen, als je bij het installeren een API-sleutel ingevoerd hebt tenminste.



Alles op groen: als de installatie achter de rug is, kun je naar de startpagina van Tattler gaan.



Met een klik in de tekstballon ga je naar het themaoverzicht, waar je nieuwe onderwerpen kunt aanmaken.



**Via het menu Content kom je bij een lijst van alle nieuwsbronnen.**

Als alles werkt en je het eerste nieuws ziet, kun je de cron-job vast instellen. Geef als gebruiker root het commando `crontab -e` en maak een cron-job met het volgende commando:

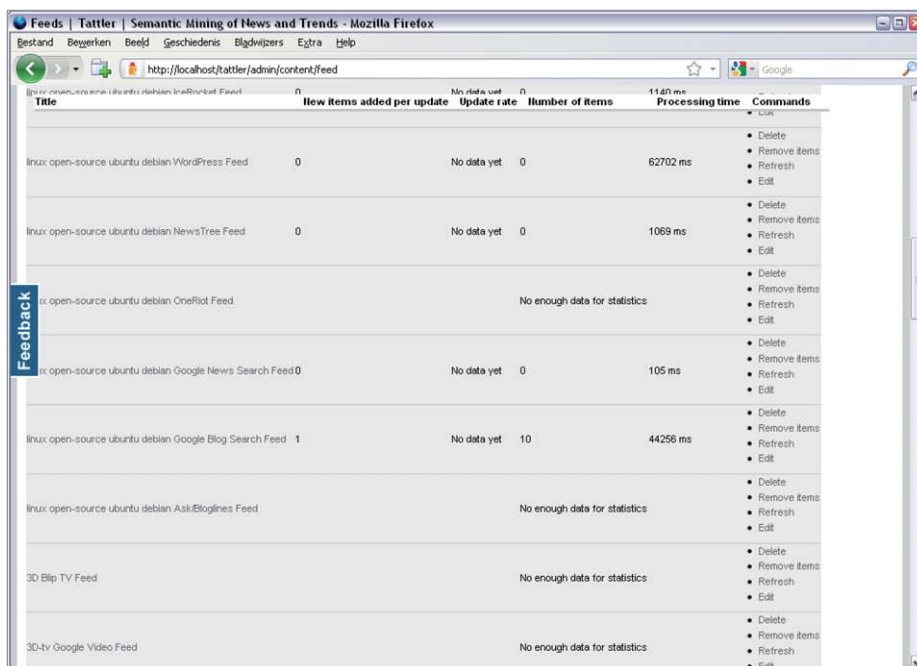
```
* /10 * * * * curl --silent --compressed http://tattler.mijndomein.nl/cron.php
```

Dit commando start de cron-job iedere 10 minuten. In de vorige c't stond een artikel over het aanpassen van cron-jobs [1]. Als je Tattler op een Windows-server gebruikt, die geen cron-service heeft, raden de ontwikkelaars aan om de Drupal-module Poormanscron te installeren. Nadat je die via de softlink hebt gedownload, zet je hem in de map 'modules'.

**Feeds en trends**

Via de optie 'Content / Edit / List' in de beheer-menubalk linksboven krijg je te zien welke nieuwsbronnen Tattler gevonden heeft. Die kun je bewerken door op 'edit' te klikken of er meerdere tegelijk selecteren door een vinkje voor de betreffende berichten te zetten om een bepaalde actie bij 'Update options' op uit te voeren. Op die manier zijn bronnen te verwijderen, als bookmark in te stellen of op een zwarte lijst te zetten. Dan zie je die alleen als je op de hoofdpagina expliciet voor de optie 'Blacklist' kiest.

De meeste gebruikers zullen bij het instellen van Tattler al een assortiment aan RSS-feeds hebben die ze bijhouden. Tattler zal die waarschijnlijk niet allemaal automatisch gevonden hebben, dus die kun je nog handmatig toevoegen bij een bepaald thema. Daar zijn twee mogelijkheden voor. Als je maar een paar bronnen hoeft toe te voegen, kan dat via het menu-item 'Content / Add'. Als je dan op 'Feed' klikt, opent een formulier om een newsfeed toe te voegen.



Als je daarentegen meteen meerdere feeds als nieuwsbron wilt toevoegen, is er een snellere manier. Tattler kan namelijk bronnen in het OPML-formaat importeren. Bijna iedere RSS-reader kan zijn feeds in dat formaat exporteren. Om een OPML-bestand te importeren, ga je naar het menu-item 'Content / Feeds / Import OPML'.

Om er achter te komen welke ingestelde thema's op een bepaald moment veel nieuws opleveren, klik je op het item 'Trends'. Dan krijg je een overzicht te zien waarin Tattler het verkeer van de afzonderlijke thema's grafisch weergeeft.

**Tattler voor teams**

Standaard kan alleen de administrator die bij het installeren van Tattler aangemaakt is deze dienst gebruiken. Als je Tattler openbaar wilt maken of voor een bepaalde groep gebruikers wilt ontsluiten, moet je de instellingen bij

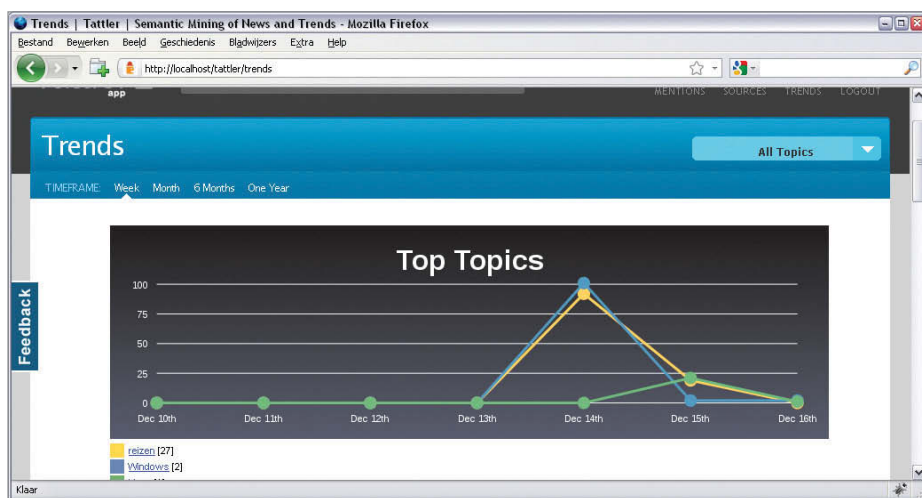
'People / User settings' aanpassen. Daar kun je bepalen of alleen gebruikers die door de administrator zijn aangemaakt de site mogen bekijken of dat ook gebruikers die zichzelf via het inlogformulier hebben aangemeld dat mogen. In het laatste geval heb je ook de mogelijkheid de geregistreerde gebruikers eerst door de administrator te laten accorderen.

Als je alleen bepaalde gebruikers toegang tot je Tattler wilt verlenen, moet je hen als administrator toevoegen via 'People / Users / Add user'. Daarna kun je nauwkeurig instellen wat ze wel en niet mogen door hen een bepaalde rol toe te kennen. Via 'People / Roles' kun je zien welke rollen er zijn. Dat zijn de anonieme gebruikers, die standaard niets mogen, tot administrators aan toe. Via 'People / Permissions' kun je aanpassen wat de rechten van de afzonderlijke rollen zijn. Daar kun je bijvoorbeeld instellen of aangemelde gebruikers nieuwe thema's mogen aanmaken.

Zoals gezegd is Tattler gebaseerd op Drupal. Drupal wordt door een levendige community continu bijgewerkt, vandaar dat je niet moet schrikken als je bij 'Reports / Available updates / List' een aantal rode (update required) en gele (update available) modules ziet. Op de website van Tattler staat verder niets over het updaten ervan, daarvoor moet je bij Drupal zelf zijn. Tattler is bij Release Candidate 1 aanbeland, maar die is al van oktober 2009. Producent phase2tech heeft de ontwikkeling een tijdje op een zacht pitje gezet. Het is dan ook nog onbekend of en wanneer de definitieve versie zal verschijnen en of die met Drupal 6 of de nieuwe versie 7 zal werken. (nkr)

**Literatuur**

[1] Mirko Döller, Tijdmachines, Routineklusjes onder Linux met Cron, Anacron en At automatiseren, c't 3/2011, p.120



Bij de Trend-weergave geeft Tattler een grafisch overzicht van de drukte bij de verschillende onderwerpen.

