



door Egon Willighagen
<egonw(at)linuxfocus.org>

Over de auteur:

Hij heeft een meestersgraad in chemie aan de Universiteit van Nijmegen, en hij verricht zijn PhD onderzoek over moleculaire representatie aan dezelfde Universiteit. Hij speelt basketbal en schrijft java-programma's.

Vertaald naar het Nederlands door:
CyberProphet
<cyberprophet/at/winux.be>

BibTeXML in DocBook XML gebruiken om wetenschappelijke artikelen te schrijven



Kort:

In dit artikel zal ik je laten zien hoe je een soort BibTeX referentiesysteem met DocBook XML kunt gebruiken. Om dit mogelijk te maken, heb ik gereedschappen ontwikkeld om dit proces te vergemakkelijken - dit zijn pakketten die meegeleverd worden met de JReferences distributie.

Inleiding

LaTEX gebruikers weten hoe bruikbaar BibTeX is. Dit is een heel handig gereedschapje die referenties aan andere wetenschappelijke literatuur toevoegt zonder dat je veel naar de eigenlijke afdruk moet omkijken, daar die toch automatisch correct wordt aangemaakt - i.e. de correcte weergave doorgeven, maar zonder dat je manueel de type moet ingeven. Net zoals Latex zelf. Meer zelfs, in wetenschappelijke literatuur is het gebruikelijk om referenties te nummeren met superscript nummers, op deze manier¹. Deze nummers moeten in opeenvolgende volgorde staan. Ook dat regelt BibTeX.

Docbook is steeds meer mijn favorietere tekstgereedschap aan het worden, door zijn ordelijke XML-gebaseerde syntax, de grote ondersteuning voor het maken van webpagina's (e.g. CDK website <http://cdk.sf.net/>, is volledig in DocBook gesschreven) en de man pagina's. De volgende stap voor mij was om DocBook te gebruiken voor het schrijven van wetenschappelijke artikelen. Daarom had ik BibTeX nodig voor DocBook. En daarom schreef ik JReferences.

JReferences doet meer dan BibTeX. Net als BibTeX heeft het gereedschappen om referenties uit gewone tekstdatabases automatisch te nummeren, maar het kan meer. Het ondersteunt meerdere formaten (zowel om in te geven als om weer te geven), het heeft een MySQL backend die gebruikt kan worden met een PHP frontend. Ook probeert het een referentie-database te zijn zoals EndNote. Alhoewel het een open broncode project betreft (onder de GPL licentie) heeft het buiten mezelf nog maar bitter weinig ontwikkelaars aangetrokken, de ontwikkeling gebeurt dan ook moeizaam. Dat wil echter niet zeggen dat het daarom niet bruikbaar is, en van het tegendeel zal ik je in dit artikel overtuigen.

Wanneer dit artikel gepubliceerd werd was JReferences ontwikkeld tot versie 0.7.2. Dit artikel werd dan ook met deze versie in het achterhoofd geschreven.

Een DocBook artikel

Bekijk eens onderstaand voorbeeld dat terug te vinden is in het JReferences pakket.

```
<?xml version="1.0"?>
<!DOCTYPE article PUBLIC
"-//JReferences//DTD DocBook JReferences Module //EN"
"../dtd/jreferences.dtd" []>
<article>
  <jref:mode>Number</jref:mode>
  <articleinfo>
    <title>Test Article</title>
    <author><firstname>Egon</firstname>
      <surname>Willighagen</surname></author>
    <date> 3 May 2000</date>
  </articleinfo>
  <section>
    <title>Some section</title>
    <para>This is a text with a reference
<jref:cite id="Steinbeck99"/>.</para>
    <para>And now for some more serious tests, we
    add a second reference <jref:cite id="Bachrach99"/>.
    And again the first reference <jref:cite id="Steinbeck99"/>.
    </para>
  </section>
  <jref:reflist/>
</article>
```

Dit voorbeeld zal ik regel voor regel doorlopen. De eerste regel is een gewone (optionele) regel die je de syntax van het bestand vertelt, wat dus XML is. De tweede tot de vierde regel vertelt je dat de XML taal die gebruikt werd DocBook is, maar dat het de JReferences modules gebruikt heeft in plaats van de normale DTD. Normale DocBook XML kent JReferences niet, en het gebruik van de standaard DTD zou onbruikbare documenten kunnen opleveren. De JReferences module kent echter zowel DocBook als JReferences (voor DocBook kenners: nog geen SVG of MathML). Dus, door deze module te gebruiken kun je het bestand terug valideren zoals het hoort - het voorbeeld hierboven is een correct DocBook document.

De vijfde regel bevat de begin-tag van het artikelelement. Tot zo ver gaat alles goed. Op de zevende regel begint de pret: het eerste jref element. Het <jref=mode> element wordt gebruikt om aan

JReferences te vertellen welk soort van genummerde referentie er gebruikt dient te worden. In de inleiding heb ik al vermeld dat referentienummers vaak in de superscriptstijl gebruikt worden. Maar er zijn vele alternatieven. JReferences ondersteunt [1], ², en [Steinbeck99]. Deze laatste geeft de code die in de referentie gebruikt werd. Het voorbeeld gebruikt de eerste optie.

De volgende paar regels bevatten gewone DocBook inhoud, en de volgende echt interessante regel is de 17e. In deze regel wordt de eerste referentie geciteerd. Latex gebruikers zouden hiervoor `\cite{ }` gebruiken; De JReferences code hiervoor is `<jref:cite id= "enigeID"/>` De ID correspondeert met een referentie in de database, die later uitgelegd zal worden. De volgende paragraaf van de sectie bevat nog twee citaten, waarvan één nogmaals de eerste referentie is.

Om de referenties ook daadwerkelijk in te voegen wordt in regel 24 `<jref:reflijst/>` gebruikt. Dit JReferences commando zal geconverteerd worden naar een DocBook geformateerde lijst van referenties, in de volgorde waarin ze werden geciteerd.

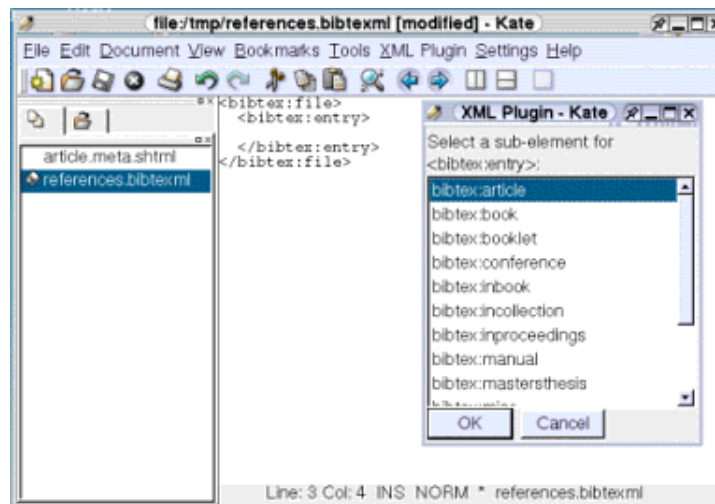
De BibTeXML database

Het JReferences systeem heeft een database nodig die heel wat weg heeft van de *.bib bestanden in Latex/BibTeX. JReferences kan een BibTeXML backend gebruiken, maar kan ook overweg met andere systemen (zoals MySQL). BibTeXML werd ontwikkeld door Vidar Gundersen en Zeger Hendrikse. Het voorbeeld in de JReferences distributie (0.7.2) maakt nog geen gebruik van BibTeXML, maar het voorbeeldartikel zou een BibTeXML-bestand kunnen bevatten zoals hier:

```
<?xml version="1.0" encoding="UTF-8"?>
<bibtex:file xmlns:bibtex="http://www.bitjungle.com/~bibtex/">
<bibtex:entry bibtex:id="Steinbeck99">
  <bibtex:article>
    <bibtex:title>JChemPaint - Using
      the Collaborative Forces of the Internet to
      Develop a Free Editor for 2D Chemical
      Structures</bibtex:title>
    <bibtex:author>Steinbeck, C. and
      Krause, S. and
      Willighagen, E.</bibtex:author>
    <bibtex:year>2000</bibtex:year>
    <bibtex:volume>5</bibtex:volume>
    <bibtex:pages>93-98</bibtex:pages>
  </bibtex:article>
</bibtex:entry>
<bibtex:entry bibtex:id="Bachrach99">
  <bibtex:article>
    <bibtex:title>End-User Customized Chemistry Journal
      Articles</bibtex:title>
    <bibtex:author>Bachrach, S. and
      Krassavine, A. and
      Burleigh, D.</bibtex:author>
    <bibtex:journal>J.Chem.Inf.Comput.Sci.</bibtex:journal>
    <bibtex:year>1999</bibtex:year>
    <bibtex:volume>39</bibtex:volume>
    <bibtex:pages>81-85</bibtex:pages>
  </bibtex:article>
</bibtex:entry>
```

```
</bibtex:file>
```

De tweede regel bevat de begin-tag voor het hoofdelement <bibtex:file>. Zo'n file (bestand) bevat één of meerdere <bibtex:entry> elementen. En elke ingang bestaat uit een BibTeX XML referentie met type: article (artikel), book (boek), incollection (in collectie), unpublished (niet gepubliceerd), misc (allerhande) en others (andere). Elke dergelijke referentie bevat specifieke elementen voor dat type, maar enkele zijn algemeen, zoals <bibtex:titel> en <bibtex:jaar>. De JReferences distributie levert de BibTeX XML DTD mee zodat elke editor die DTDs ondersteunt gemakkelijk BibTeX XML documenten kan aanpassen. Meer nog, JReferences bevat zelfs Meta DTD's voor Kate onder KDE 3.x (zie het artikel DocBook XML documenten bewerken), die automatisch geïnstalleerd wordt in de map \$HOME/.kde/share/apps/katexmlplugin.



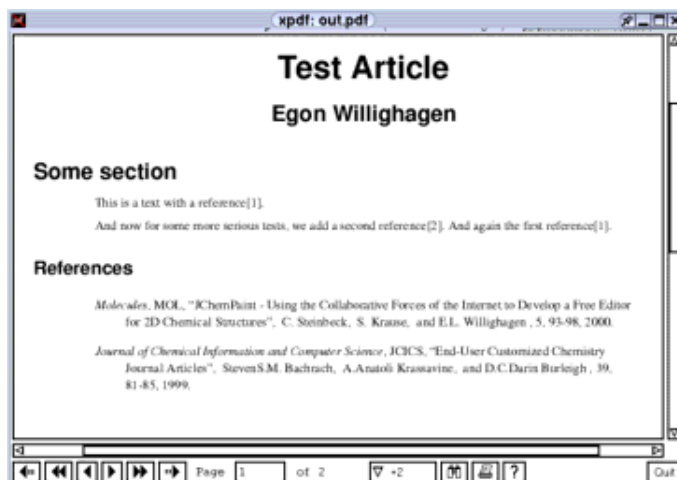
Het aanpassen van BibTeX XML bestanden met Kate, zijn XML plugin en de JReferences BibTeX XML Meta DTD.

Het genereren van een bibliografie

Beschouw nogmaals de twee bovenstaande voorbeelden. Het DocBook document is bewaard als article.docbookxml, en de referentie database is bewaard als references.bibtexml. JReferences bevat nog geen gereedschap zoals het BibTeX programma, maar hetzelfde kan bereikt worden met enkele commando's. De volgende commando's gaan ervan uit dat je JReferences onder een systeem van de Unix-familie hebt geïnstalleerd, zoals Linux (zie onder):

```
jref-clear --filedb  
jref-set --filedb --bibtexml references.bibtexml  
jref-number --filedb article.docbookxml > article-numbered.docbookxml
```

Het bestand dat verkregen wordt, article-numbered.docbookxml genaamd, is een waardig DocBook XML 4.1.2 document zonder enige <jref:*> elementen, en kan behandeld worden met elk ander gereedschap dat gebruikt wordt om DocBook XML documenten te converteren, b.v. naar PDF. (Zie b.v. eens PDF documenten maken met DocBook).



Het resultaat: een PDF document met genummerde referenties en een bibliografie.

Dat is alles wat je moet weten, of, toch niet....

Het formatteren van stijlen

Er is nog één interessant punt. BibTeX ondersteunt ook opmaakstijlen omdat de meeste publikaties specifieke eisen stellen aan hoe een bibliografie opgemaakt dient te worden. Op het ogenblik bevat JReferences slechts twee stijlen. De eerste is een soort van standaard DocBook XML formaat, dat eigenlijk niet echt een stijl genoemd kan worden. Maar er is ook de ACS stijl die door de Amerikaanse Chemische Societeit (ACS) vereist wordt.

Het `<jref:Referentielijst>` element heeft een `@style` attribuut, dat de stijl specificeert die in de plaats van de standaard stijl gebruikt dient te worden. Om de ACS stijl te gebruiken dien je regel 23 te vervangen door:

```
<jref:reflist style="ACS"/>
```

Het installeren van JReferences

JReferences vereist een java 1.3 (of hoger) installatie, Xerces, Log4J en DocBook XML DTD 4.1.2. Sommige werktuigen vereisen bijkomende programma's, zoals python (voor conversie van BibTeX naar BibTeXML) en Perl (om EndNote's Bibtex output op te ruimen).

Als dit allemaal geïnstalleerd is, kan JReferences geïnstalleerd worden met:

```
./configure --prefix=$HOME  
make  
make install
```

Als enkele onderdelen niet gevonden kunnen worden, probeer dan deze opties: `-with-xercesdir`,

-with-log4javadoc en -with-sgml. Om meer informatie te krijgen over deze opties typ je "./configure --help".

Het Project

JReferences is nu zo'n twee jaar oud, en terwijl het al vele malen is gedownload krijg ik niet veel respons. Dit is uiteraard zonder mijn persoonlijke ervaringen meegerekend. In de afgelopen maanden werd JReferences immers succesvol gebruikt voor het schrijven van een puur wetenschappelijk artikel. Maar zoals bij elk goed open-bron project zijn commentaar, foutrapporteringen, patches, ideeën en succesverhalen zeer welkom op de Project pagina van JReferences.

Referenties

- DocBook XML, <http://docbook.sf.net>
- BibTeX XML, <http://bibtexml.sf.net>
- JReferences, <http://sf.net/projects/jreferences>

Site onderhouden door het LinuxFocus editors team © Egon Willighagen "some rights reserved" see linuxfocus.org/license/ http://www.LinuxFocus.org	Vertaling info: en --> -- : Egon Willighagen < egonw@linuxfocus.org > en --> nl: CyberProphet < cyberprophet@winux.be >
--	--