

Laboration 3 –

RDF och RSS med PHP och Databaser

I denna laboration kommer ni att generera XML-innehåll från data i en databas, och därefter transformera detta XML-innehåll med hjälp av PHP. Mer specifikt ska ni skapa en RSS-baserad news-feed från databasen, och sedan skapa ett stylesheet som transformerar denna feed till HTML.

Laborationen är obligatorisk och ger upp till 7 poäng. Om den redovisas innan deadline (se kursinformationen) så ges ytterligare 2 bonuspoäng. Laborationen genomförs antingen i grupper om två, eller individuellt.

Teoribakgrund

- Föreläsningssanteckningar om Namespaces, RDF och RSS.
- Sid 61-65 i *XML in a Nutshell* handlar primärt om Namespaces, men exemplen är just RDF och Dublin Core.
- Sid 270-271 i *XML in a Nutshell* beskriver kort RDF och RSS.
- W3C har tutorials för RDF: http://www.w3schools.com/xml/xml_rdf.asp och för RSS: http://www.w3schools.com/xml/xml_rss.asp
- Den officiella platsen för Dublin Core finns på <http://dublincore.org/>
- På <http://www.php.net/> finns dokumentation och syntax för php-funktioner.

Förberedelse

I din webbmapp "TNM065" skapar du undermappen "lab-rss" och placerar där filerna för laborationen, som alla finns tillgängliga på Lisam. Dessa består av filerna 'start.php', 'prefix.php', 'postfix.php', 'index-rss.xml', 'out.xml', samt 'out.html'.

Dessutom finns en dump av databasen som används 'rsslab.sql', för den som eventuellt vill köra på egen server. Notera dock att det krävs stöd för XSLT-transformationer på servern för att göra den sista delen av laborationen.

Öppna din webbläsare och gå in på adressen

`http://www.student.itn.liu.se/~ditt-användarnamn/TNM065/lab-rss/start.php`

Om allt gått rätt och ni ser en lien tabell med år, vecka och några rader med namn, är ni redo att påbörja laborationen.


Laborationsuppgift 1: RSS-feed

Filen 'start.php' är ett PHP-skript som hämtar data om en fiktiv kaffelista ur en databas och skriver ut posterna i en tabell i HTML-format.

Börja med att spara om filen 'start.php' som 'rss.php'. Filen 'rss.php' ska nu modifieras så att den skapar ett rss-dokument istället för att skapa ett html-dokument (som i fallet 'start.php'). Börja med att ändra mime-typen som sätts på första raden av 'rss.php' från text/html till text/xml. Detta krävs för att en webbläsare som läser filen ska förstå att det inte rör sig om en websida utan ett xml/RSS-dokument. Normalt görs en sådan mappning av filtyper till mime-typer genom den ändelse filen har (dvs. filer som slutar på .html ges mime-typen text/html, och filer som slutar på .xml ges mime-typen text/xml). I detta fall, när ett program genererar innehållet, måste man explicit sätta mime-typ, om man vill att resultatet ska vara något annat än text/html. Om ni

återigen öppnar 'rss.php' bör ni istället för en websida se en listning på xml-strukturen som sidan innehåller.

Er uppgift är nu att modifiera 'rss.php' till att generera en RSS-feed från en databas med ett antal böcker. Innehållet, som i slutänden ska återskapas, finns i filen 'out.xml'. Om ni öppnar ett korrekt RSS-dokument i Firefox, som har en inbyggd RSS-läsare, kommer den visuella representationen se ut som nedan, med böckerna sorterade efter senaste utgivningsdatum:

 Prenumerera på den här kanalen med

☐ Använd alltid Aktiva bokmärken för att prenumerera på kanaler.

Bokkatalog

En bokkatalog.

[Modern PHP: New Features and Good Practices](#)
19 februari 2015 17:00

PHP is experiencing a renaissance, though it may be difficult to tell with all of the outdated PHP tutorials online. With this practical guide, you will learn how PHP has become a full-featured, mature language with object-orientation, namespaces, and a growing collection of reusable component libraries. Author Josh Lockhart, creator of PHP The Right Way, reveals these new language features in action. You will learn best practices for application architecture and planning, databases, security, testing, debugging, and deployment. If you have a basic understanding of PHP and want to bolster your skills, this is your book.

[RDF Database Systems: Triples Storage and SPARQL Query Processing](#)
26 november 2014 22:00

RDF Database Systems is a cutting-edge guide that distills everything you need to know to effectively use or design an RDF database. This book starts with the basics of linked open data and covers the most recent research, practice, and technologies to help you leverage semantic technology. With an approach that combines technical detail with theoretical background, this book shows how to design and develop semantic web applications, data models, indexing and query processing solutions.

[Beginning XML](#)
11 juli 2012 00:00

A complete update covering the many advances to the XML language. The XML language has become the standard for writing documents on the Internet and is constantly improving & evolving

[The Personal Internet Address & Password Log Book](#)
4 juli 2010 21:00

Are you tired of losing track of those login, usernames and passwords you create every time you visit a new website? Now you can keep important website addresses, usernames, and passwords in one convenient place!

[XML Developers Guide](#)
21 december 2000 18:00

This is a comprehensive guide to XML for intermediate to advanced professional developers

(Notera att detta inte fungerar för t.ex. Chrome eller Safari, då dessa webbläsare inte har någon inbyggd RSS-läsare.)

Det är samma uppgifter för att logga in i databasen som används i 'start.php', men tabellen med data heter 'bookcatalog' istället för 'calender'. Den är definierad på följande sätt:

```
CREATE TABLE `bookcatalog` (  
  `link` varchar(255) NOT NULL default '',  
  `title` varchar(255) NOT NULL default '',  
  `description` longtext NOT NULL,  
  `author` varchar(255) NOT NULL default '',  
  `publish_date` datetime NOT NULL default '0000-00-00 00:00:00',  
  PRIMARY KEY (`link`)  
) ENGINE=MyISAM DEFAULT CHARSET=latin1;
```

Uppgifter i RSS-feeden som ni inte får hårdkoda, utan ska genereras via PHP är:

- Datumen i dc:date.
Datumet direkt under 'channel' ska vara aktuellt datum/tid, de övriga ska hämtas från databasen. Det är inget krav att ni ändrar till rätt tidszon.
- Länkarna till böckerna i rdf:li-elementen, som är barn till rdf:Seq.
Dessa hämtar ni från 'link'-kolumnen i databasen.
- All information i item-elementen:
<title>, <link>, <description>, <dc:creator> (där ni använder 'author' från databasen) och <dc:date> (som ni hämtar från 'publish_date')

Några tips:

Ni kan stöta på problem då innehållet i databasen för både 'title', 'link' och 'description' innehåller &-tecken. Det beror på att &-tecken används av XML för att visa att en entitet ska användas, men i det här fallet vill vi faktiskt att just ett &-tecken ska genereras. Lösningen är att ersätta '&' med '&' vilket är entiteten för ett &-tecken. Använd förslagsvis php-funktionen preg_replace (på liknande sätt som 'ö' ersätts med 'ö' i 'start.php'). Vissa länkar kan också innehålla mellanslag, som då måste ersättas med '%20'. Att kontrollera att data man hämtar från exempelvis en databas inte innehåller otillåtna tecken kan vara en viktig lärdom inför projektet.

Datumformatet i databasen skiljer sig åt från hur datum ska anges i dc:date. För att göra om ett mySql-datum till det format som ska förekomma i ett dc:date-element kan ni t.ex. använda 'strtotime' för att omvandla ett mySql-datum till ett php-datum, och sedan använda 'strftime' för att få ut en textrepresentation av datumet på önskat format. Ni kan också snegla på parametern "c" i date-funktionen.

När ni är klar med uppgiften ska er RSS-feed gå att validera på: <http://feedvalidator.org/>

Laborationsuppgift 2: Transformerings med XSLT

Som avslutning ska ni skapa ett xsl-stylesheet som transformerar det RSS-dokument som genereras från databasen till ett html-dokument, dvs. något som fungerar i alla webbläsare. Slutresultatet ska efterlikna det i 'out.html'.

De XSLT-transformationer ni gjorde i Lab 2 ägde rum på klienten, men det är från och med PHP5 även möjligt att göra det på servern, vilket ni ska göra i denna uppgift. Spara om er fil 'rss.php' med namnet 'rss2.php' och ersätt den första php-raden med följande:

```
<?php include 'prefix.php';?>
```

Lägg även till följande rad absolut sist i dokumentet:

```
<?php include 'postfix.php';?>
```

Den första raden gör att XML-innehållet som skapas lagras i en buffert istället för att skrivas ut. Den sista raden tar buffertens innehåll och applicerar en XSLT-transformation med filen 'index-rss.xsl'

Om ni öppnar filen 'rss2.php' i en webbläsare bör ni först se alla textnoder. Uppgiften är nu att modifiera 'index-rss.xsl' så att resultatet blir identiskt med 'out.html'. Rubrikerna ska vara h2-element som är länkade till de URL:er som finns i link-elementet, för respektive bok.

Redovisning

Laborationen redovisas genom inlämningsfunktionen i Lisam, där att ni laddar upp filerna 'rss.php' och 'rss2.php', samt anger URL:er till era filer. Om ni genomfört laborationen i par så gör ni endast **en gemensam inlämning**, där ni lägger till båda era namn som deltagare vid inlämningen. Om ni samarbetat, utan att meddela detta, och gör varsin inlämning med identiska lösningar, så skapas en onödig risk för att misstankar om fusk ska uppstå.

Om ni missat något och får poängavdrag, så ges möjlighet till komplettering. Efter att laborationen rättats får ni besked om vad som behöver kompletteras för full poäng, samt när sista datum för kompletteringen är (normalt en vecka efter att laborationens rättats). Observera att **endast en** komplettering är möjlig. Efter att en komplettering lämnats in, eller om ingen komplettering lämnas in före deadline, spikas slutpoängen för laborationen.