

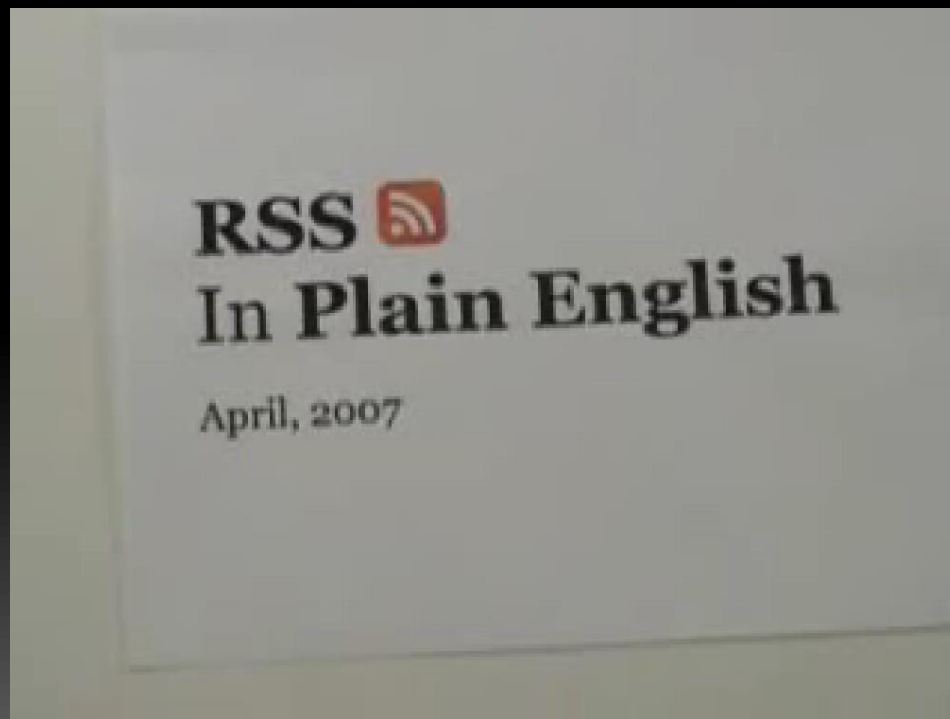


RSS

Service Oriented Architecture

Let Show It by Video...

- <http://video.google.com/videoplay?docid=1881080775770831914>





Really Simple Study

ANY PROBLEMS?!

What does RSS stand for?

- It depends!
- Some call it Rich Site Summary
- Others call it Really Simple Syndication
- It also be called RDF Site Summary

So, what is it?

- RSS is a **text-based** format, a type of **XML**. RSS files are often labeled as XML.
- RSS version **1.0 is also RDF** (Resource Discovery Framework), which is important only because an RSS file may be labeled as RDF.
- By using **XSLT** (Extensible Stylesheet Language), these XML files can be converted to a much more **friendly view**.

[\[日記\]不願面對的真相](#)



Brings Content to YOU

- Pull technology – “feeds” external information to your Web site
- Information from different sources can be displayed in a single view
- Distributes content management

Example – Univ. of Texas

The screenshot shows a Mozilla Firefox browser window titled "Colleges & Academic Units - Mozilla Firefox" with the address bar displaying "http://www.utexas.edu/dept/". The page content is organized into three main sections, each with a list of items on the left and a summary box on the right. The first section is "COLLEGE OF NATURAL SCIENCES", the second is "SCHOOL OF NURSING", and the third is "COLLEGE OF PHARMACY". Three yellow boxes with black borders are overlaid on the page, each containing the text "Source: [Unit Name]".

COLLEGE OF NATURAL SCIENCES

- Astronomy
- Jackson School of Geosciences
 - Bureau of Economic Geology
 - Geological Sciences
 - Institute for Geophysics
- Human Ecology
- Marine Science
- Mathematics
- Physics

College of Natural Sciences

- CNS Graduation Information for Undergraduates
- Women in Natural Sciences: Researching and Obtaining Undergraduate Research Opportunities
- UTeach, the Science and Mathematics Secondary Teacher Preparation Program
- College of Natural Sciences Outreach

SCHOOL OF NURSING

Source: School of Nursing

The School of Nursing

- Luci Baines Johnson Funds Endowment
- National Institutes of Health Funds \$800,000 Grant
- Children's Wellness Center-Nursing in the Community

COLLEGE OF PHARMACY

- Medicinal Chemistry

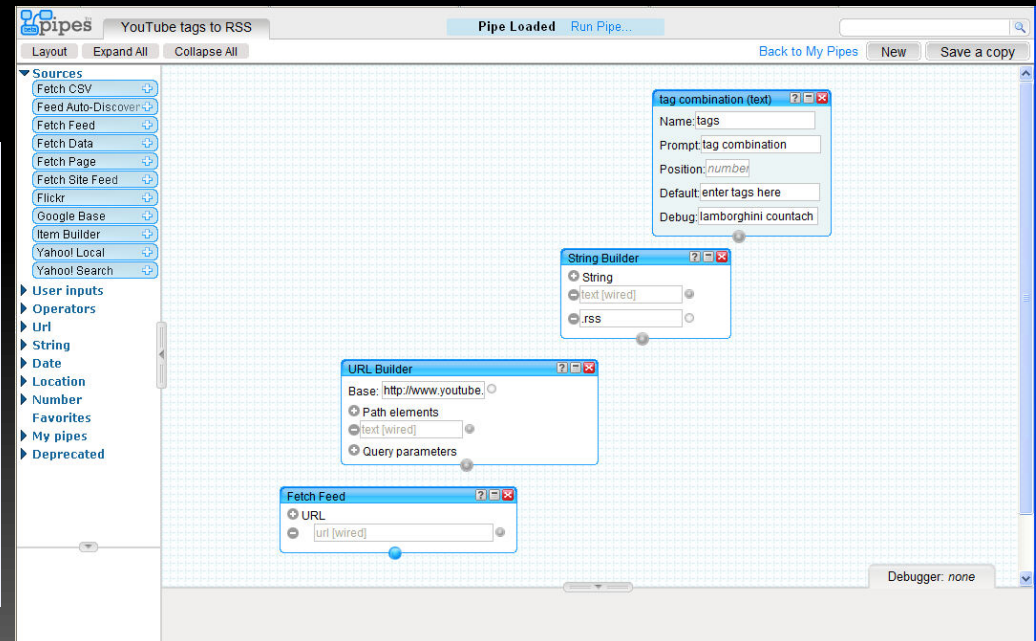
Source: College of Pharmacy

The College of Pharmacy

- Longhorn Pharmacy Focus Online, Alumni Magazine
- Dean's Welcome

Yahoo! Pipes

- <http://pipes.yahoo.com/>
- <http://mr6.cc/?p=709>
- <http://www.jumpcut.com/fullscreen?id=594F555C568011DC9D24000423CEF5B0&type=movie>



What about Atom?

- **Atom** is a format quite **similar to RSS**. It was created by people who felt that RSS could be **improved** upon.
- The basic difference is that while **Atom** is somewhat **more complex** (for *producers* of Atom feeds), it is also able to carry more complex information, and it is consistent across the syndication, storage, and editing of information.
- Almost everything regarding RSS applies equally well to Atom.
- You can learn more about Atom at the official website, AtomEnabled.org.

RSS – Two Parts

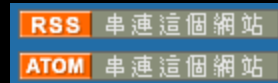
- **Feeds** – These are the actual content items that are produced by web authors. These could be taken from blogs, journals, news sites, etc.
- **Aggregators** – An aggregator is a program or site that collects the feeds for reading. Aggregators can be installed as stand-alone programs, or they can be web-based.

Advantages of RSS

- By using an aggregator, you can read feeds from multiple websites in one location.
- No SPAM!
- No Pop-Ups!
- Most blogs now have feeds, and more journals and news sites are adding them.
- Organization!!

How to find RSS feeds


- Look for a button like these within a site:



Web-Based Aggregators

- [Bloglines](#) – Currently the most popular in this category.
- [Newsgator](#) – Allows for delivery of content to Outlook.
- [Litefeeds](#) – good for PDAs & Mobile Phones
- [Google Reader](#)
- [Netvibes](#)
- Personalized Google Page
- My Yahoo! Page
- ...

Stand-alone readers

- [SharpReader](#) – for Windows
- [FeedDemon](#) – Also for Windows
- [NetNewsWire](#) – Good choice for Macs
- [Straw](#) – Great for Linux
- [RSS Toolbar](#) – Internet Explorer & Firefox (Windows only)
- Outlook 2007 rz



RSS Via E-Mail

- <http://www.rssfwd.com/>



Check the Complete List

- <http://www.newsonfeeds.com/faq/aggregators>

Are people talking about RSS?

- [35 ways to use RSS](#)
- [25 New Ways To Use RSS](#)
- [About.com](#) has some ideas
- [Businesses](#) are using it!
- Microsoft has an [RSS blog!](#)

Creating Feeds

- Most blog sites create the feeds for you.
 - <http://blogger.com>
 - <http://livejournal.com>
 - <http://typepad.com>
 - <http://wordpress.com>
- Create them in Dreamweaver
 - <http://www.rnsoft.com/products/rssdreamfeeder/>

Creating Feeds

- FeedForAll
 - <http://www.feedforall.com/>
- Moveable Type
 - <http://www.sixapart.com/movabletype/>
- Create your own with XML.
 - <http://www.rss-specifications.com/creating-rss-feeds.htm>
 - <http://www.make-rss-feeds.com/making-an-rss-feed.htm>

How to Make RSS Feed

- RSS feeds contain what are referred to as "items"
 - title
 - description
 - link
- The title and description should be written to describe the content and the link should reference the webpage that contains that actual content.

An Item May Look Like...

```
<title>The Title Goes Here</title>  
  <description>The description goes  
  here</description>  
  <link>http://www.linkgoeshere.com</link>
```

Items May Look Like...

```
<item>
```

```
  <title>The Title Goes Here</title>
```

```
  <description>The description goes here</description>
```

```
  <link>http://www.linkgoeshere.com</link>
```

```
</item>
```

```
<item>
```

```
  <title>Another Title Goes Here</title>
```

```
  <description>Another description goes  
  here</description>
```

```
  <link>http://www.anotherlinkgoeshere.com</link>
```

```
</item>
```

A Channel Contains Items...

```
<channel>
```

```
  <title>The Channel Title Goes Here</title>
```

```
  <description>The explanation of how the items are  
  related goes here</description>
```

```
  <link>http://www.directoryoflinksgohere</link>
```

```
<item> </item>
```

```
<item> </item>
```

```
<item> </item>
```

```
</channel>
```


RSS Designation

```
<?xml version="1.0"?>
```

```
<rss version="2.0">
```

```
<channel>
```

```
  <title>The Channel Title Goes Here</title>
```

```
  <description>The explanation of how the items are related goes  
  here</description>
```

```
  <link>http://www.directoryoflinksgohere</link>
```

```
<item> </item>
```

```
<item> </item>
```

```
<item> </item>
```

```
</channel>
```

```
</rss>
```



You Got a RSS Feed

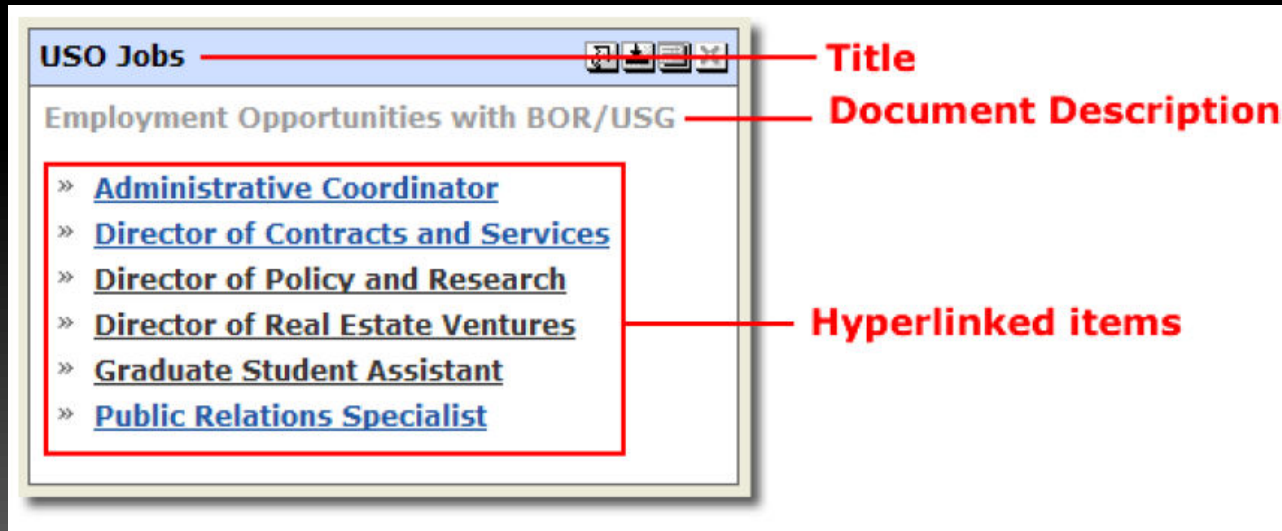
- Make sure to save it as an xml file
- Check your work
 - <http://feedvalidator.org>

Let's Check a Real RSS

```
<?xml version="1.0" encoding="UTF-8" ?>
- <rss version="2.0">
- <channel>
  <title>NotePage Software News - SMS and Paging Software News</title>
  <description>News relating to SMS messaging, text messaging and paging software.</description>
  <link>http://www.notepage.net/press.htm</link>
  <language>en-us</language>
  <lastBuildDate>Sun, 8 Jul 2007 08:37:13 -0400</lastBuildDate>
  <pubDate>Tue, 10 Jul 2007 09:00:00 -0400</pubDate>
  <webMaster>webmaster@notepage.net</webMaster>
  <generator>FeedForAll v2.0 (2.0.1.0) http://www.feedforall.com</generator>
- <image>
  <url>http://www.notepage.net/notepage-logo-tweak.gif</url>
  <title>NotePage Press Releases</title>
  <link>http://www.notepage.net/software.htm</link>
  <description>press and announcements related to NotePage's software</description>
  <width>134</width>
  <height>138</height>
</image>
- <item>
  <title>Multifunction Messaging Server Updated</title>
  - <description>
    - <![CDATA[
      NotePage, Inc. is pleased to announce the release of <a href="http://www.notepage.net/pagegate.htm" target="_blank" >PageGate version 5</a>
      solution that incorporates a wide variety of messaging protocols for sending messages to mobile phones, wireless devices, fax machines and p
      <br />
      Wireless messaging communication is important in todays fast paced environment. PageGate allows critical alerts to be sent instantly to wirel
      <br />
      To ensure delivery even in the event of connectivity loss, PageGate version 5 supports failover options and redundant connectors. Messages t
      <br />
      Messages can be sent into PageGate via: a Windows client, web page, email, commandline, text file, serial port, modem, wireless modem, wirel
      <br />
      PageGate's modular design allows users to select from eight different interfaces, allowing users to scale PageGate to meet their business ne
      <br />
      PageGate supports multiple telecom carriers, multiple connectors, direct connections to paging terminals, group paging, scheduled messages,
      <br />
      PageGate supports a variety of delivery methods including the following protocols: Fax, GSM-AT, SMTP, SNPP, TAP, Tone, UCP and WCTP. Message
    ]]>
  </description>
  <link>http://www.notepage.net/press-release-pagegate5.htm</link>
  <guid isPermaLink="true">http://www.notepage.net/press-release-pagegate5.htm</guid>
  <pubDate>Tue, 10 Jul 2007 09:00:00 -0400</pubDate>
</item>
```

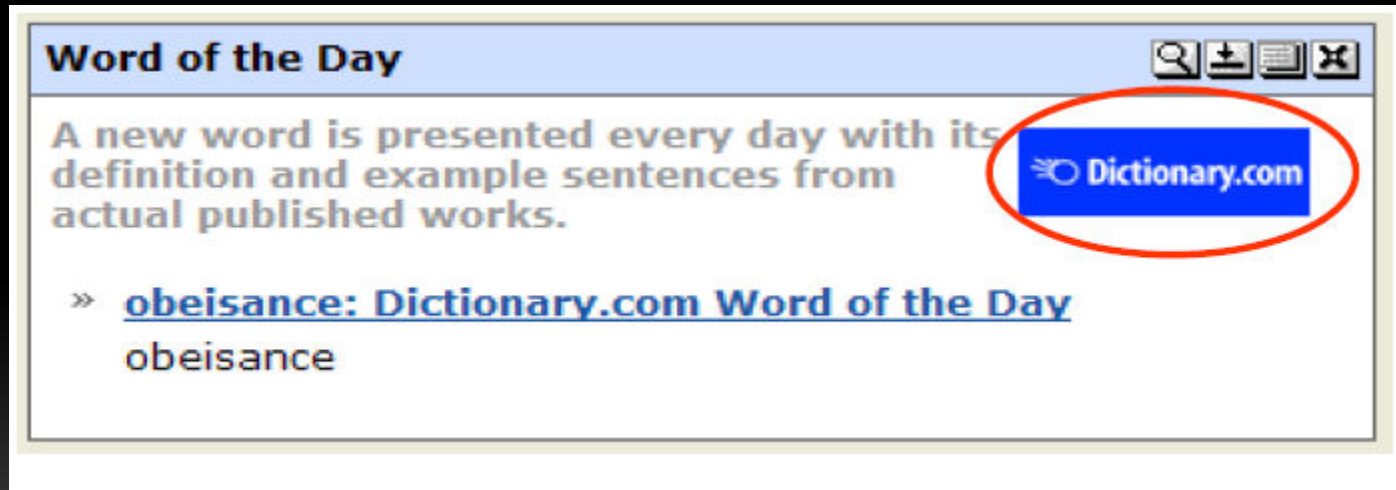
RSS Core Elements

- All RSS feeds have:
 - RSS Title
 - RSS Description
 - Hyperlinked list of items



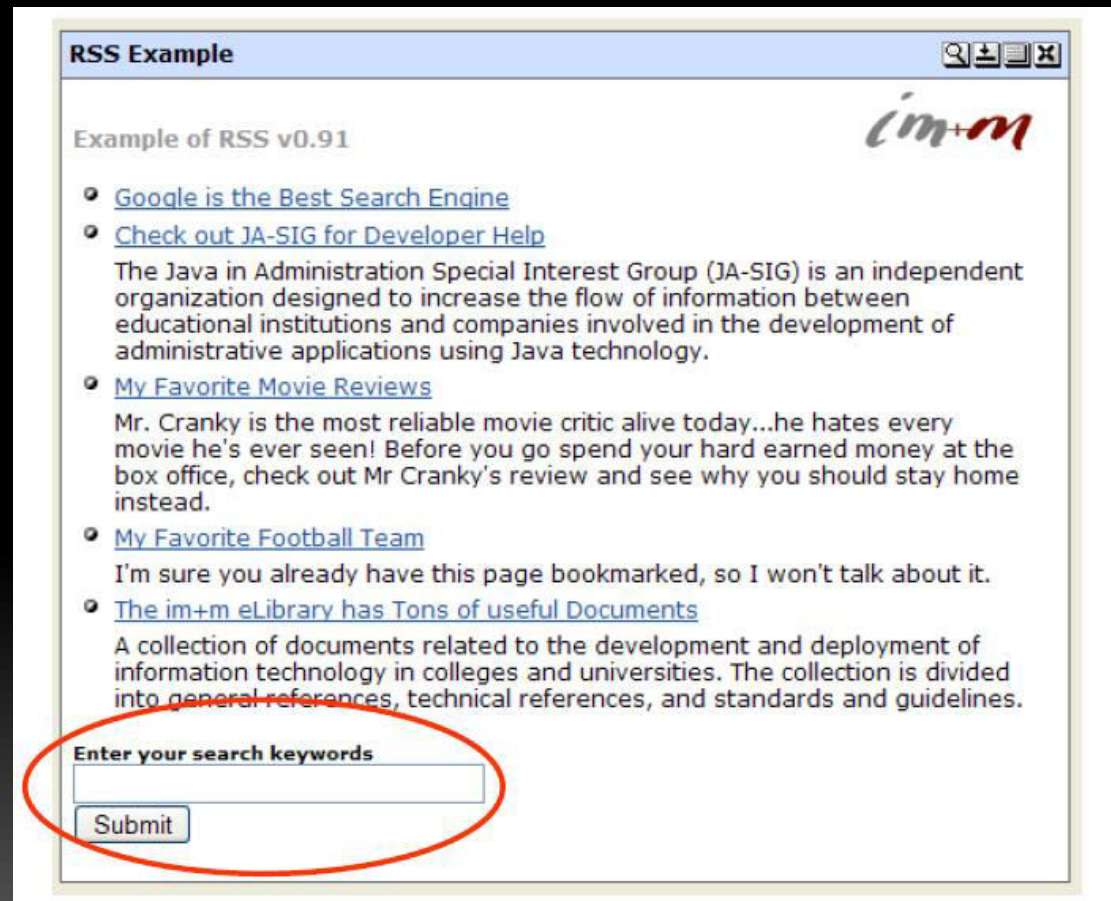
RSS Enhanced Elements

- RSS Image



RSS Enhanced Elements

- Text Input



The screenshot shows a web browser window titled "RSS Example". The page content includes the text "Example of RSS v0.91" and the "im+m" logo. There are four list items, each with a title and a description:

- [Google is the Best Search Engine](#)
- [Check out JA-SIG for Developer Help](#)
The Java in Administration Special Interest Group (JA-SIG) is an independent organization designed to increase the flow of information between educational institutions and companies involved in the development of administrative applications using Java technology.
- [My Favorite Movie Reviews](#)
Mr. Cranky is the most reliable movie critic alive today...he hates every movie he's ever seen! Before you go spend your hard earned money at the box office, check out Mr Cranky's review and see why you should stay home instead.
- [My Favorite Football Team](#)
I'm sure you already have this page bookmarked, so I won't talk about it.

Below the list items is a search section with the text "Enter your search keywords" above a text input field. A "Submit" button is located below the input field. A red circle highlights the input field and the "Submit" button.

At the bottom of the screenshot, there is a text input field and a "Submit" button, which are circled in red.

Enhanced Elements

- Document and item-level metadata
- Information about the feed: publisher, copyright, contact information, last updated date
- Modules
 - Dublin Core
<http://web.resource.org/rss/1.0/modules/dc/>
 - Streaming Media
- Embedded HTML**

RSS Versions

- At least four distinct, incompatible flavors of RSS (some say there are as many as seven)
- All RSS versions are XML
 - Must be well-formed
 - Must start with `<?xml version="1.0"?>`

RSS Version

- RSS 0.90 was the original Netscape RSS version.
- RSS 1.0 is an open format by the RSS-DEV Working Group
- RSS 1.1 is also an open format and is intended to update and replace RSS 1.0

RSS Version

- The RSS 2.* branch (initially UserLand, now Harvard)
 - RSS 0.91 is the simplified RSS version released by Netscape
 - RSS 0.92 through 0.94 are expansions of the RSS 0.91 format
 - RSS 2.0.1 has the internal version number 2.0.

RSS 0.9

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns="http://my.netscape.com/rdf/simple/0.9/">
  <channel>
    <title>Mozilla Dot Org</title>
    <link>http://www.mozilla.org</link>
    <description>the Mozilla Organization web site</description>
  </channel>
  <item>
    <title>New Status Updates</title>
    <link>http://www.mozilla.org/status/</link>
  </item>
</rdf:RDF>
```

RSS 0.90

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns="http://my.netscape.com/rdf/simple/0.9/">

  <channel>
    <title>Mozilla Dot Org</title>
    <link>http://www.mozilla.org</link>
    <description>the Mozilla Organization web site</description>
  </channel>

  <item>
    <title>New Status Updates</title>
    <link>http://www.mozilla.org/status/</link>
  </item>
</rdf:RDF>
```



RSS 0.9x

```
<?xml version="1.0"?>
```

```
<rss version="0.91">
```

```
  <channel>
```

```
    <title>XML.com</title>
```

```
    <link>http://www.xml.com/</link>
```

```
    <description>XML.com features a rich mix of information and services for the XML  
community.</description>
```

```
    <language>en-us</language>
```

```
    <item>
```

```
      <title>Normalizing XML, Part 2</title>
```

```
      <link>http://www.xml.com/pub/a/2002/12/04/normalizing.html</link>
```

```
      <description>In this second and final look at applying relational normalization techniques  
to W3C
```

```
XML Schema data modeling, Will Provost discusses when not to normalize, the scope of uniqueness  
and
```

```
the fourth and fifth normal forms.</description>
```

```
    </item>
```

```
  </channel>
```

```
</rss>
```

RSS 1.0

```
<?xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns="http://purl.org/rss/1.0/">
  <channel rdf:about="http://www.usg.edu/admin/humres/cojobs/jobs.rss">
    <title>Employment Opportunities with University System of Georgia</title>
    <link>http://www.usg.edu/admin/humres/cojobs/</link>
    <description>Employment Opportunities with BOR/USG</description>
    <items>
      <rdf:Seq>
        <rdf:li resource="http://www.usg.edu/admin/humres/cojobs/ga_partnership.phtml" />
      </rdf:Seq>
    </items>
    <item rdf:about="http://www.usg.edu/admin/humres/cojobs/ga_partnership.phtml">
      <title>Director of Policy and Research</title>
      <link>http://www.usg.edu/admin/humres/cojobs/ga_partnership.phtml</link>
    </item>
  </channel>
</rdf:RDF>
```

RSS 1.0 with HTML

```
<item rdf:about="http://www.immagic.com">
  <title>5/2: Win-Win Negotiation: Practice and Principles</title>
  <link>http://www.immagic.com</link>
  <description>Guest Instructor, Snuffleupagus. Assignments Due: Read "Getting to 'Yes'"
  book, Read "Leadership Challenge" (Chapters 1-2), Individual Assignment #1 (submit by
  email)</description>
  <content:items>
    <rdf:Bag>
      <rdf:li>
        <content:item>
          <content:format rdf:resource="http://www.w3.org/1999/xhtml"/>
          <content:encoding rdf:resource="http://www.w3.org/TR/REC-xml#dt-wellformed"/>
          <rdf:value rdf:parseType="Literal" xmlns="http://www.w3.org/1999/xhtml">
            Guest Instructor, <strong>Snuffleupagus</strong>.<br/>
            Assignments Due:<ul>
              <li>Read: <em>Getting to 'Yes'</em> book</li>
              <li>Read: <em>Leadership Challenge</em> (Chapters 1-2)</li>
            </ul>
          </rdf:value>
        </content:item>
      </rdf:li>
    </rdf:Bag>
  </content:items>
</item>
```

RSS 2.0

```
<?xml version="1.0"?>
<rss version="2.0">
<channel>
  <title>Jim Flowers: Blogs and Education</title>
  <link>http://radio.weblogs.com/0113212/categories/blogsAndEducation/</link>
  <description>Efforts by K-20 education</description>
  <copyright>Copyright 2004 Jim Flowers</copyright>
  <lastBuildDate>Wed, 13 Oct 2004 21:42:16 GMT</lastBuildDate>
  <docs>http://backend.userland.com/rss</docs>
  <managingEditor>jimdf@bellsouth.net</managingEditor>
  <item>
    <link>http://radio.weblogs.com/0113212/...</link>
    <description>
      <P><STRONG>e-Development</STRONG></P>
      <P>Asia worries about <A href="http://www.washingtonpost.com/...">Kerry outsourcing
      ideas</A>.</P>
    </description>
    <guid>http://radio.weblogs.com/0113212/...</guid>
    <pubDate>Wed, 13 Oct 2004 21:42:14 GMT</pubDate>
  </item>
</channel>
</rss>
```


Summary

- RSS 0.90
 - Very simple, very restrictive
- RSS 0.9x
 - Also simple, still restrictive
 - More features than 0.90
- RSS 1.0
 - Complicated, tightly structured, very restrictive
 - Many features
- RSS 2.0
 - Somewhat complicated, very loose
 - Many features
 - Support??



Which Version?!

- Depends on:
 - What tools you will use to create your feed?
 - How and where you will display your feed?

RSS Spec.

- RSS V₂
 - <http://cyber.law.harvard.edu/rss/rss.html>
- RSS V₁
 - <http://web.resource.org/rss/1.0/spec>
- RSS V_{0.9x}
 - <http://backend.userland.com/rss093>
- RSS V_{0.9}
 - <http://www.purplepages.ie/RSS/netscape/rso.go.html>

RSS Writer Class

- **PHP Universal Feed Generator**
 - <http://www.ajaxray.com/blog/2008/03/08/php-universal-feed-generator-supports-rss-10-rss-20-and-atom>
 - RSS 1.0/ 2.0, ATOM
- **RSS Writer Class – PHP**
 - <http://snipplr.com/view/23/rss-writer-class/>
- **PHP RSS Writer Class**
 - <http://webscripts.softpedia.com/script/Snippets/PHP-RSS-Writer-Class-24923.html>

RSS Parser

- Magpie RSS
 - <http://magpierss.sourceforge.net/>
- RSS Parser and XML Parser for PHP 5+
 - <http://sourceforge.net/projects/rssphp/>
- Others
 - You can find them out in <http://sourceforge.net/>
 - Try RSS Parser, or XML Parser

Magpie RSS

- Very simple to use

```
require('rss_fetch.inc');  
$rss = fetch_rss($url);
```

Usage

- `$rss->channel`
 - Contains key-value pairs of all tags
 - Without nested tags
 - Found between the root tag (`<rdf:RDF>`, or `<rss>`) and the end of the document

Usage II

- `$rss->items`
 - An array of associative arrays
 - Each one describing a single item

```
<item rdf:about="http://protest.net/NorthEast/calendrome.cgi?span=event&ID=210257">
<title>Weekly Peace Vigil</title>
<link>http://protest.net/NorthEast/calendrome.cgi?span=event&ID=210257</link>
<description>Wear a white ribbon</description>
<dc:subject>Peace</dc:subject>
<ev:startdate>2002-06-01T11:00:00</ev:startdate>
<ev:location>Northampton, MA</ev:location>
<ev:enddate>2002-06-01T12:00:00</ev:enddate>
<ev:type>Protest</ev:type>
</item>
```

```
array(
  title => 'Weekly Peace Vigil',
  link => 'http://protest.net/NorthEast/calendrome.cgi?span=event&ID=210257',
  description => 'Wear a white ribbon',
  dc => array (
    subject => 'Peace'
  ),
  ev => array (
    startdate => '2002-06-01T11:00:00',
    enddate => '2002-06-01T12:00:00',
    type => 'Protest',
    location => 'Northampton, MA'
  )
);
```


Usage III

- image and textinput
 - `$rss->image`
 - `$rss-textinput`
 - Associative arrays including name-value pairs for anything found between the respective parent tags

Very Simple Example

```
require_once 'rss_fetch.inc';
$url = 'http://magpie.sf.net/samples/imc.1-0.rdf';
$rss = fetch_rss($url);
echo "Site: ", $rss->channel['title'], "<br>\n";
foreach ($rss->items as $item )
{
    $title = $item[title]; $url = $item[link];
    echo "<a href=$url>$title</a></li><br>\n";
}
```

Example - Display Top n Items

```
$num_items = 10;  
$rss = fetch_rss($url);  
$items = array_slice($rss->items, 0, $num_items);  
foreach ( $items as $item )  
{ ... }
```

Error Handling

```
# Magpie throws USER_WARNINGS only
# so you can cloak these, by only showing ERRORS
    error_reporting(E_ERROR);
# check the return value of fetch_rss()
$rss = fetch_rss($url);
if ( $rss )
    { ...display rss feed... }
else
    { echo "An error occurred! " . "Consider donating more $$$
    for restoration of services." . "<br>Error Message: " .
    magpie_error(); }
```

Only Display Today Posts

```
require_once('rss_utils.inc');
$today = getdate();
$date = mktime(0,0,0,$today['mon'], $today['mday'], $today['year']);
$rss = fetch_rss($url);
foreach ( $rss->items as $item )
{
    $published = parse_w3cdtf($item['dc']['date']);
    if ( $published >= $date )
    {
        echo "Title: " . $item['title']; echo "Published: " . date("h:i:s A", $published);
        echo "<p>";
    }
}
```

Parse Local File

```
require_once('rss_parse.inc');
$rss_file = 'some_rss_file.rdf';
$rss_string = read_file($rss_file);
$rss = new MagpieRSS( $rss_string );
if ( $rss and !$rss->ERROR)
    { ...display rss... }
else
    { echo "Error: " . $rss->ERROR; }
function read_file($filename)
{
    $fh = fopen($filename, 'r') or die($php_errormsg);
    $rss_string = fread($fh, filesize($filename) );
    fclose($fh);
return $rss_string;
}
```