
RSS Feeds and its Implementation at INFLIBNET

Rajeev Kumar

Abstract

RSS is a Web 2.0 technology that is being used by millions of web users around the world to keep track of their favourite websites. In the 'old days' of the web to keep track of updates on a website visitors had to 'bookmark' websites in their browser and manually return to them on a regular basis to see what had been added. RSS has become a valuable technology for everything from casual web users to webmasters. According to a recent Yahoo survey only 12% of internet users are aware of RSS and a mere 4% have knowingly used RSS. The paper discusses about the different aspects of RSS such as meaning of RSS, evolution of RSS, how to create RSS and how to get / read RSS Feeds including basics of XML, which is the basic language for writing programs for RSS Feeds. At the end of this paper, details of the project about the INFLIBNET's RSS Reader is also discussed.

Keywords : RSS, Feed, RSS Reader, Aggregator, XML.

1. Introduction

In the present age of Internet and rapid changing web technology, it is almost impossible to visit the website these days without seeing the mention of RSS or icons pointing to RSS resources. RSS is most commonly used on news sites and blogs. It is one of important advertising tools about the new products and informing latest research papers in a journal. Visitor can be updated as soon as new content is added. RSS (Really Simple Syndication) is a method of delivering web content to users that does not necessarily require a web browser. In technical words, RSS is an XML-based format for content distribution. Webmasters create an RSS file containing headlines and descriptions of specific information. While the majority of RSS feeds currently contain news headlines or breaking information, the scope of long term uses of RSS are broad. RSS is a defined standard based on XML with the specific purpose of delivering updates to web-based content. Using this standard, webmasters provide headlines and fresh content in a succinct manner. RSS Reader, a program to see the result of feeds, is very important. Visitors use RSS Readers or RSS Aggregators to collect and monitor their desired / favourite feeds in one centralized program or location. Content viewed in the RSS Reader or RSS Aggregator is known as an RSS feed.

2. What is RSS?

RSS stands for more than one thing. Usually, RSS stands for "Really Simple Syndication". But it can also mean "Rich Site Summary" or "RDF Site Summary", where RDF stands for "Resource Description Framework." RDF Site Summary is the first version of RSS created in March 1999. This version became known as RSS 0.9. In July 1999, a new version RSS 0.91 was produced and renamed RSS as "Rich Site Summary". A new version RSS 2.0 was released in September 2002 and now RSS became Really Simple Syndication.

The technical acronym for RSS is “Really Simple Syndication”, an XML format that is created to syndicate news, and be a means to share content on the web. Wikipedia defines RSS as follows: “RSS is a family of Web feed formats used to publish frequently updated works such as blog entries, news headlines, audio, and video in a standardized format. An RSS document (which is called a “feed,” “web feed,” or “channel”) includes full or summarized text plus metadata such as publishing dates and authorship. Web feeds benefit publishers by letting them syndicate works quickly and automatically. They benefit readers who want to subscribe to timely updates from favoured websites or to aggregate feeds from many sites into one place.”

3. How to Read RSS Feeds

RSS Feeds are read using RSS Reader. RSS Reader, which is also termed as Feed Reader or RSS Aggregator, is a software which can be web based or desktop based. The user subscribes to a feed by entering the feed’s link into the Reader by clicking an RSS icon in a browser that initiates the subscription process. The RSS Reader checks the user’s subscribed feeds regularly for new work, downloads any updates that it finds and provides a user interface to monitor and read the feeds. There are many feed readers going around with a variety of approaches and features. Some the them are freely available, some are purchasable. Some work as web services and some are limited to Windows or Linux platform. However a good place to start is with a couple of free and easy to use web based one like Google Reader and Bloglines. In the table -1 and 2 given below, list of some RSS Reader for Windows / Linux platforms are given. Now-a-days some browsers have also built-in RSS Reading facility such as Mozilla, Internet Explorer7. In these browsers, after typing the URL of the RSS Link, one can get the feed.

Table – 1 : RSS Reader for Windows Platform

Sl. No.	RSS Reader	Source
1.	Awasu	http://www.awasu.com
2.	EffNews	http://www.iffbot.org/zone/effnews-exe.htm
3.	FeedReader	http://www.feedreader.com
4.	InfoSnorkel	http://www.blueelephantsoftware.com/rss_aggregator.htm
5.	NewzSpider	http://www.newzspider.com/
6.	RSS Captor Pro	http://www.rsscaptor.com
7.	RSSReader	http://www.rssreader.com
8.	SharpReader	http://sharpreader.com
9.	SurfPack	http://www.surfpack.com
10.	Syndirella	http://www.yole.ru/projects/syndirella/

Table – 2 : RSS Reader for Linux Platform

Sl. No.	RSS Reader	Source
1.	Bottom Feeder	http://www.cincomsmalltalk.com/BottomFeeder/
2.	Composite	http://www.jelovic.com/rssaggregator/
3.	Eclipse RSS Reader	http://morphine.sourceforge.net/presence/projects/rss/index.html
4.	K.R.S.S.	http://krss.sourceforge.net/
5.	Liferea	http://liferea.sourceforge.net
6.	Mozilla Thunderbird	http://www.mozilla.com/thunderbird/
7.	Straw	http://www.nongnu.org/straw/
8.	Syndicator	http://syndicator.sourceforge.net

4. Benefits of RSS

4.1 Benefit to Webmasters

As the web has become more crowded, webmasters have been striving to provide fresh and up to date content for their website visitors. Many webmasters have discovered they can easily utilize the information in RSS feeds to provide fresh web content. As discussed earlier, RSS feeds are composed in XML, which is a very simple markup language. Similar to HTML, XML uses tags to identify fields. Webmasters can easily parse the RSS feed and dynamically create web pages that contain headlines and summaries. The feeds will continuously update, supplying a steady stream of automatically generated fresh content. RSS allows webmasters to:

- 1.) Provide fresh and relevant content on their website, which encourages users to return.
- 2.) Constantly changing content means that search engine spiders will visit more frequently.
- 3.) Automate content delivery.

The benefits of RSS feeds are not limited to webmasters, surfers too benefit from the technology as well.

4.2 Benefit to Web Surfers

The beauty of RSS is that readers can quickly scan headlines (titles) and read articles of interest. Because the information is condensed and provided in a single location users can generally review more information in a shorter time frame. Additional information is only a click away. Best of all readers choose the feeds they wish to see, there is no spam with RSS. If one is not completely thrilled with the content appearing in a feed, simply it can be removed from the newsreader. The

technology is a pull technology rather than push technology, meaning the content is not forced on the consumers, who pull the content they want to see.

RSS allows for users to:

- 1.) Easily locate information.
- 2.) Read condensed information or 'soundbytes' with clearly marked and dated topic material.
- 3.) Classify and categorize information in an easy to navigate manner.
- 4.) Maximize their time without having to deal with spam.

RSS feeds can be viewed in a news aggregator or reader, which constantly updates and shows unread feeds. It is found the functionality of the newsreaders to be similar to a simple email client. Consumers generally enter the URL of any RSS feeds that interest them. Topics with a common theme can be segregated into related groups.

4.3 Benefit to Content Developer

While the benefits to users and webmasters are clear the distribution opportunities made available to content developers should not be overlooked. Information contained in the RSS feed can be easily syndicated, increasing content distribution and reach. RSS allows for content developers to:

- 1.) Increase exposure in niche markets.
- 2.) Communicate with user bases and reach potential customers via an alternate communication method.
- 3.) Disseminate relevant information.
- 4.) Define themselves as an industry expert.
- 5.) Automate content delivery.

RSS has effectively standardized the format for content delivery and has effectively defined the accepted standard for content distribution and syndication. RSS will likely rival email as a means of content distribution in another few years. The sheer simplicity makes the technology very appealing. The distribution potential, while albeit difficult to measure, is still attractive to all parties making the likelihood that RSS popularity will only continue to grow.

5. Creating Rss Feeds

RSS Feeds are created using XML format. Hence to create one's own RSS Feeds to be delivered through website, knowledge of XML is must. XML is a markup language much like HTML. In the next subsections, we will try to understand how to write XML codes for RSS Feeds. However there are several RSS Feed creation software also available, for which knowledge XML programming is not needed.

5.1 Using XML Programming

- ◆ **Basics of XML** : XML stands for eXtensible Markup Language, which was designed to transport and store. XML is very easy to learn. Like HTML (Hyper Text Markup Language), tags are basic thing in XML. But XML tags are not predefined. One may define his / her own tags. XML is a W3C Recommendation. It is just plain text. Software that can handle plain text can also handle XML.

Structure of XML documents: -

XML documents form a tree structure given as below: -

```
<root>
  <child>
    <subchild> ..... </subchild>
  </child>
</root>
```

The elements in an XML document form a tree type of structure. The tree starts at root and branches to the lowest level of the tree. All elements can have sub elements terms as child elements. The relationship between various elements of an XML document are described using the terms parent, child and siblings. Parent elements have children and children on the same level are known as siblings. All elements can have text content and attributes.

Let us take an example of XML document which contains a note regarding meeting from Director to All Scientist: -

```
<?xml version="1.0"?>
<note>
  <to>All Scientists</to>
  <from>Director</from>
  <heading>Meeting</heading>
  <body>Please make your presence in the meeting to be held tomorrow with
  Google representatives regarding sharing of our union database.</body>
</note>
```

In the above XML program, first line is the XML declaration. It defines the XML version (1.0). The next line describes the root element, called note, of the. The next 4 lines describe 4 child elements of the root (to, from, heading and body) and finally the last line defines the end of the root element.

Following are basic syntax rules of XML: -

- ◆ The first element must be a root element, which is the parent of all other elements.
- ◆ Name of an XML element can contain letters, numbers and other characters, but it must not start with a number or punctuation character. Also it must not start with letters xml, XML, Xml etc. and it can not contain spaces.

- ◆ All XML elements must have a closing tags.
- ◆ XML tags are case sensitive and hence opening and closing tags must be written with the same case.
- ◆ XML elements must be properly nested.
- ◆ Tags may contain some special property i. e. additional information about elements with associated values, which is called tag's attribute and their value. XML attributes must be quoted.

5.2 Making RSS Using XML

After learning basics for XML, now let us try to understand how to write RSS using above concepts.

Here is the structure an RSS Feed using XML: -

```
<?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> A Website address is given here as link </link>
<item>
<title> The Title of First Item Goes Here </title>
<description> The description of first item goes here </description>
<link> The web link consisting the detailed of first item goes here </link>
</item>
<item>
<title> The Title of Second Item Goes Here </title>
<description> The description of second item goes here </description>
<link> The web link consisting the detailed of second item goes here </link>
</item>
</channel>
</rss>
```

Description of the code: -

The first line in the document is the XML declaration which defines the XML version used in the document.

The next line is the RSS declaration which identifies that this is an RSS document (in this case, RSS version 2.0).

The next line contains the <channel> element. This element is used to describe the RSS feed. In fact, RSS Feed is a series of items, these items are chained together to create what is called a "Channel". The Channel appears at the top of the file and tells the use how the items relate to each other.

The <channel> element has three required child elements:

- ◆ <title> - Defines the title of the channel
- ◆ <description> - Describes the channel (e.g. Free web building tutorials)
- ◆ <link> - Defines the hyperlink to the channel
- ◆ Each <channel> element can have one or more <item> elements. The open channel tag <channel> occurs before the first item and the close tag </channel> appears after the last item. Each <item> element defines an article or news or detailed information or story in the RSS feed.
- ◆ The <item> element has three required child elements:
 - ◆ <title> - Defines the title of the item
 - ◆ <description> - Describes the item
 - ◆ <link> - Defines the hyperlink to the item
- ◆ At the end of each item, there must be closed tag of item element.
- ◆ Finally, the two last lines close the channel and rss elements.
- ◆ Besides above, following elements may be used under <channel>, which are optional: -
 - ◆ <category> - it defines one or more categories for the feed.
 - ◆ <copyright> - it notifies about copyrighted material.
 - ◆ <generator> - it specifies the program used to generate the feed (about this more details are in next subsection).
 - ◆ <image> it allows an image to be displayed when RSS Reader present a feed.
 - ◆ <language> it specifies the language the feed is written in.
 - ◆ <pubDate> - it defines the last publication date for the content of the feed.
 - ◆ <webmaster> - it defines the e-mail address to the webmaster of the feed.

Following are some sub-elements for <item> element, which are optional: -

- ◆ <author> - it specifies the e-mail ID to the author of the item.
- ◆ <category> - it defines one or more categories the item belong to.
- ◆ <comments> - it allows an item to link to comments about that item.
- ◆ <enclosure> - it allows a media file to be included with the item.

- ◆ <pubDate> - it defines the last publication date for the item.
- ◆ <source> - it specifies a third party source for the item.

After creation of the RSS, it may be published on the web using following steps: -

1. Name the RSS file with an .xml extension.
2. Upload the RSS file to the desired web directory on the web server.

Results may be seen using a RSS Reader or RSS Aggregation. Now-a-days browsers have built-in RSS reading facility.

5.3 Making RSS using available software

There are several RSS Feed creation software are available which crates RSS Feeds that comply with RSS specifications. Some are free, some may be get as free trial for some days and some are to be purchased. FeedForAll is one of the Feed creation software which is easy to use desktop application that allows users to create and edit RSS feeds. Some other software are MyRSSCreator, FeedFire, etc. For users who only need an RSS feed for their personal website, they can use services provided by some of the most popular blog managers such as Blogger, Radio, which offer built-in RSS Services.

6. INFLIBNET's RSS Reader

At INFLIBNET, a unique web-based RSS Reader is developed, which may be found through the link: <http://www.inflibnet.ac.in/feed/>

Uniqueness of the INFLIBNET's RSS Reader is that user, particularly research personals, may find RSS Feed either subject-wise or publisher-wise. They can see latest updates about their research area without searching related RSS links and then subscribing them to their RSS Reader. Beauty of INFLIBNET's RSS Reader lies in its "My RSS" feature. A user may register him/her by signing up followed by singing in and selection of his / her favourite RSS Title. Thereafter, any time after signing in he / she can directly see the feed and aware about the latest research in his / her area of research / interest without selecting RSS Title from large list of all titles.

Under this project, INFLIBNET maintains a database of RSS Links of all the journals subscribed under its e-resource consortia. However user may see the feed without registration also. For this, he / she has to select either "Subject wise" or "Publisher wise" option in home page of INFLIBNET's RSS Reader.

Details of Options available on the home page INFLIBNET's RSS Reader: -

1. INFLIBNET's News – This gives the output of latest news and updates of INFLIBNET Centre.
2. MY RSS – Here user can see their selected feeds, which is stored in INFLIBNET's database. This database includes tables related to registered users and their selected RSS Titles. Any time, user can edit (add new titles or delete old titles) their selection. First time, user has to register themselves.

3. Subject Wise RSS – In this, user may select the subject of their choice / interest and then related RSS titles will appear. After selection of RSS title related feed will appear.
4. Publisher Wise RSS – In this, user may select the publisher of their choice / interest and then related RSS titles will appear. After selection of RSS title related feed will appear.
5. All Available RSS – In this, list of all RSS Titles will appear in the drop down list. User may select the Title of their choice / interest and then feed will appear.

For this project, open source and ajax technologies are used. Database is developed using MYSQL, where as interfaces are based on PHP-AJAX.

7. Conclusion

RSS is one among Web 2.0 Technologies. RSS exists as a means to gather and display information quickly and easily. It's an easy way for users/visitors to keep up with news and information that's important to them, and helps them to avoid the conventional methods of browsing or searching for information on websites. Now the content one wants can be delivered directly to him/her without cluttering his/her inbox with e-mail messages. INFLIBNET's RSS Reader is one of the important and useful services provided by INFLIBNET Centre for research community.

References

1. Wikipedia
2. Wikiversity
3. Google Search
4. <http://savetimewithrss.blogspot.com>
5. <http://www.goarticles.com>
6. <http://news.bbc.co.uk/1/hi/help/3223484.stm>
7. <http://probblogger.net/what-is-rss/>
8. <http://www.xml.com>
9. <http://paulstamatiou.com/2005/11/13/how-to-getting-started-with-rss>

About Author

Mr. Rajeev Kumar, Scientist - B (CS), INFLIBNET Centre, Ahmedabad - 380 009, India. He recently developed the INFLIBNET's RSS Reader.