

COMING, READY OR NOT: RSS

by Belinda Weaver

For people who are not sure what RSS is, and may not even care that they do not know, the recent claim by Lockergnome's Chris Pirillo, author of *Poor Richard's E-mail Publishing* (the first book on e-mail newsletters), that RSS may be an answer to the growing problem of junk e-mail, or spam, should alert them to the fact that this technology is worth knowing about.

Not that the spam angle is necessarily the main point about RSS – it simply underlines the fact that this technology still has ways of surprising us. If you need to keep abreast of news and announcements, using RSS can shave hours off that job for you. If you want the news from a favourite site, but not the ads or the pop-ups, RSS can probably do that for you as well. Want to use targeted news headlines from elsewhere to add value to your site? RSS again. Want to get your own information out to people as soon as it's written? Use RSS.

According to Barry Parr, 'RSS makes it possible for smart, busy people to browse dozens of news sites in the time it would have taken them to review a couple.' The technology can be used in a seemingly inexhaustible number of ways.

The definition on PC Webopedia outlines some of these. 'Syndicated content includes such data as news feeds, events listings, news stories, headlines, project updates, excerpts from discussion forums or even corporate information.'

According to Jonathan Eisenzopf,

'RSS is becoming a vital "What's New" mechanism that serves a variety of purposes while helping to attract traffic from many different locations on the Web ... RSS is a better way to share data than more common approaches, such as fetching and parsing HTML, or using proprietary APIs, database dumps, and cobranding ... Under the RSS model, each site publishes a file describing the contents of its "channel". Other sites can subscribe to that channel and grab its contents. The RSS file could be converted to HTML and displayed directly on a subscriber site, or it might be edited first to select only those items that are appropriate for the site's audience. The nice thing about RSS, of course, is that once you've built the system to subscribe to one RSS channel, you can subscribe to thousands of them.'

According to Danny Sullivan, compiler of SearchEngineWatch,

'There are a wide-range of Web sites that "syndicate" their content in this way. Among the top 100 most subscribed feeds reported by Radio UserLand ... are technology headlines from the New York Times, the daily Dilbert cartoon, PDABuzz.com and former MTV

VJ Adam Curry's weblog. The beauty of these feeds means that you can effectively create your own custom newspaper or magazine of recent content ... In short, RSS is a way for Web site owners to let you know what new content they have available within their Web sites.... These feeds can be a great way for anyone to receive customized news information from a growing number of sources.'

And that's a key issue. It's about getting your information out, so that people stumble across it, and then come to your site to see even more of what you have and what you do. RSS is consistently recommended by enthusiasts as a way of driving traffic to Web sites.

What is it?

The acronym initially stood for RDF Site Summary (or Rich Site Summary). For many people now, it is simply shorthand for Really Simple Syndication. When you think of syndication, you may think about cartoonists or columnists whose work appears in a number of sources. RSS-style syndication is not that. According to WebReference.com, RSS is

'a lightweight XML format designed for sharing headlines and other Web content... RSS has evolved into a popular means of sharing content between sites (including the BBC, CNET, CNN, Disney, Forbes, Motley Fool, Wired, Red Herring, Salon, Slashdot, ZDNet, and more). RSS solved a myriad of problems webmasters commonly face, such as increasing traffic, and gathering and distributing news. RSS can be the basis for additional content distribution services.'

The technology was developed by Userland (another popular blog software service) in 1997 and then further developed by Netscape. It is simply a way of creating an XML 'feed' that news aggregators such as Google News and others can pick up. The feed may be news, or announcements, or lists of new books or Web sites - anything, really. Increasingly, bloggers, especially those using RSS-friendly blogging software, such as MoveableType, are creating feeds. The main point is to make the information available to people who are interested in what you do, or who *would* be interested, if only they knew about it. Having an RSS feed is certainly a good way to get material 'out there', as each word in the feed can be searched for, and found by, RSS search software, which is different from the more hit-or-miss approach of Web search engines to indexing information on Web pages.

The RSS feed itself is generally an XML file containing a number of items, each marked up in a certain manner. According to WebReference, each 'RSS text file contains both static information about your site, plus dynamic information about your new stories, all surrounded by matching start and end tags. Each story is defined by an <item> tag, which contains a headline TITLE, URL and DESCRIPTION'.

An item in an RSS feed would look something like this:

```
...  
<item>  
  <title>Christmas Opening Hours</title>  
  <link>http://www.yourlibrarysite.com/</link>  
  <description>The Shoal Bay branch will be  
closed from 5 pm on Christmas Eve until 9am on 2  
January, 2004. All other branches will reopen on  
December 27 and will operate for their normal  
hours.</description>  
</item>  
...
```

To be able to read an RSS feed, you need to 'parse' it in a reader. This can be done from desktop RSS readers, or you can use any of a number of free RSS reader programs on the Web. Parsing simply means outputting the file, not in plain XML, which would not look very attractive on screen, but in such a way that you can read and make sense of it. Instead of a mass of jumbled-looking text, you will see headlines and brief descriptions, all of which will include a link for the full story. Most feeds are reasonably plain-looking – rather like a list of headlines on a Web site, with each headline being a live link to the source of the information. Some

feeds include logos and images, but on the whole it's the information that is vital.

Chris Pirillo's RSS Quickstart Guide is as good a place as any to start. Pirillo lists the three steps necessary to use RSS. Getting an RSS reader is step one. Populating the reader

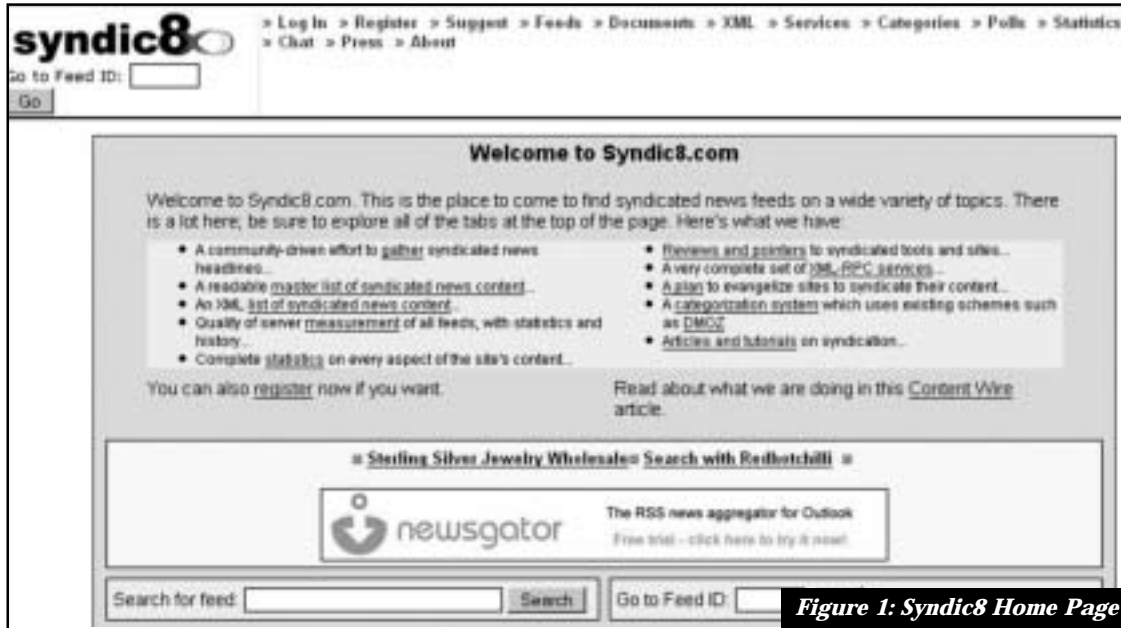


Figure 1: Syndic8 Home Page

If it all looks very reminiscent of HTML, that is not surprising: XML files have start and end tags like HTML and are designed to trigger a display just as HTML does. Also like HTML, XML files do not look very attractive or readable in their raw form. But just as HTML is transformed once a browser 'reads' the tags and triggers a correct display, so too are RSS files made comprehensible by RSS readers.

software with the feeds you want to receive is step two. This is also called 'subscribing'. Step three is the easy part – just sit back and read the content you have pulled in.

There are a growing number of RSS reader software programs available for download from the Web. It is up to you whether you opt to download RSS reader

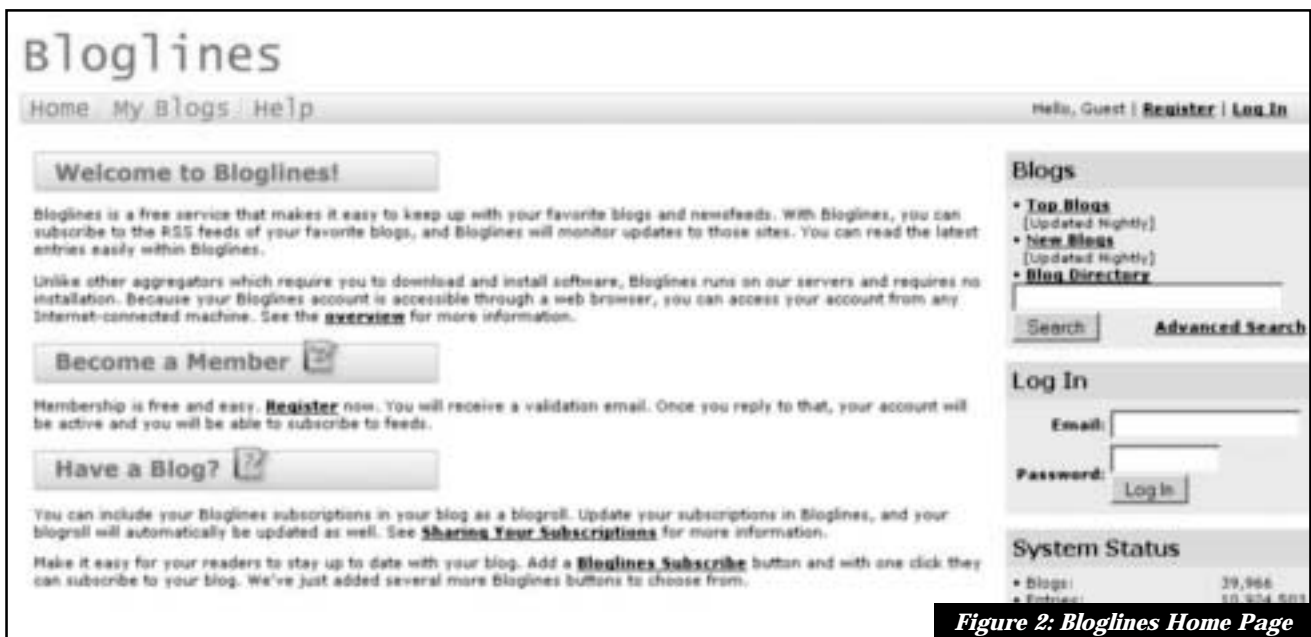


Figure 2: Bloglines Home Page

software to your PC so you can read feeds, or whether you use some of the Web-based aggregator services that show you the feeds you want to see without the need for any extra software. Certainly, until you figure out whether this technology is right for you, it might be easier to use a Web-based reader.

The two best known services for feeds are probably Newsisfree (<http://www.newsisfree.com/>) and Syndic8 (<http://www.syndic8.com/>). Newsisfree launched its service in June 2001 and can hook you up with literally thousands of feeds. In addition, they are always adding new

ones. Confusingly, these feeds are referred to as 'channels', but they are simply RSS feeds from sites as diverse as news, Web logs, and other sites that want to announce something to the world. At Newsisfree, you can browse categories or search by keyword or personal name. If you register, you will be able to log in to a more personalised page. As a guest, click on the NewsCenter section to see what's available.

Syndic8 accepts submissions from readers and bloggers as well as offering feeds added by volunteers who are helping to build the Syndic8 database. With luck, this service will eventually become what Danny Sullivan calls 'an Open Directory for RSS feeds'. You can search for feeds here.

Bloglines (<http://www.bloglines.com/>) is another free service for reading blogs and newsfeeds. It looks less busy than Newsisfree or Syndic8, and it is easy to read the latest entries within Bloglines. You will need to register, then choose the feeds you want.

If these sites don't seem to provide anything that would interest you, a better starting point might be LISfeeds (http://www.lisfeeds.com), which aggregates a lot of different library news and blog sites. Billing itself as 'one place for library news headlines', the site allows you to read a number of feeds on library-related issues. If you click on a feed in the left hand column, the content of it will be displayed in the right. It is rather a 'plain vanilla' presentation, but for speed, it can't be beaten. Rather than visiting all these separate sites individually, which would be exceptionally time-consuming, one can quickly scan the content of interesting ones and move on.

Since LISfeeds includes library-related keeping-up-to-date content, such as Tara Calishain's

researchbuzz, Gary Price's ResourceShelf, The Virtual Chase and BeSpacific, it is an easy way to keep on top of these sorts of services. ResourceShelf is probably one of the most useful weblogs for librarians, focusing as it



Figure 3 : amphetaDesk Headlines

does on new databases and invisible Web resources, as well as advances in search and information technology for information retrieval. Rita Vine's Sitelines (<http://www.workingfaster.com/sitelines>) blog, which covers Internet searching, is not available at LISfeeds, but can be added to an RSS reader, since its output is available as XML.

If you would rather go the software download road, AmphetaDesk (<http://www.disobey.com/amphetaDesk/>) is a free piece of aggregator software that translates your chosen news feeds into a readable page of headlines which you can browse at your leisure. AmphetaDesk resides on your computer's desktop, downloads the news feeds that you specify, and displays them as a Web page. You can customise the display of that page. There are other reader programs available, such as NewsGator, a news aggregator that runs within Microsoft Outlook (<http://www.newsgator.com/default.aspx>), and Newzcrawler (<http://www.newzcrawler.com/>), which will allow you to read not only XML/RSS feeds, but also postings to Usenet newsgroups and ordinary Web pages.

If you want to look beyond the world of library blogs and feeds, then you need to know about other tools. Feedster (<http://www.feedster.com/>) is a new tool for



Figure 4 : Feedster Search Tool

RSS feeds and blogs. Enter a keyword into Feedster, and it will return matches, with the very latest feeds displayed first. All entries are date stamped. Each entry comes with two links – one for the specific item, the second for the feed or blog itself. This is a handy way to discover new feeds or blogs. Should you enter the name of an actual service e.g. SiteLines, or ResourceShelf, the Feedster database will display all entries in date order, making it easy to get an idea of the scope of a feed. Daypop (<http://www.daypop.com/>) is another useful tool for blogs and feeds. Searches there can be limited to news, blogs or RSS feeds.

Increasingly, sites that output their information in RSS-friendly formats are making it obvious. Many blogs especially include small icons that say 'XML' or 'RSS'. Clicking on those icons will display the feed. Cut and paste the URL of interesting feeds into your reader and away you go.

Keeping current is obviously one of the most useful things about RSS feeds. Reading them in one spot, as LISfeeds allows you to do, is obviously faster and simpler than visiting a whole bunch of different sites. RSS feeds are rather like the benefits of e-mail over newsgroups. With a reader, the feeds come to you, rather than you needing to go out and hunt them down. However, unlike e-mail, RSS feeds contain no spam, which is why Chris Pirillo predicted they would eventually replace e-mail publishing.

Paul Eng, of the US-based ABCnews.com, also stresses the pluses of RSS feeds: 'Since the reader software can pick up multiple news feeds from many Web sites simultaneously, an RSS user can have all the information from their favorite Web sites display on one computer screen.' Amy Gahran, publisher of the *Contentious* Web log (<http://blog.contentious.com/>), says 'RSS technology offers Web publishers a number of advantages in keeping visitors up-to-date with new information. A lot of sites use e-mail announcements to keep visitors coming back to their sites. However, spam is killing e-mail.' Instead, with an RSS feed, "A [subscriber] is getting stuff directly from the publisher, and the only thing that comes into the feed reader is stuff [they] asked for," says Gahran.'

Keeping up with ALIA news, media releases, and with vacancies in library and information science can be done via RSS at ALIA's Web site (<http://alia.org.au/rss/>). Just paste the link for each XML feed into an RSS reader to see the contents of the feed.

You can certainly customise what you get as feeds – you only see the content you want. Some sites are even starting to provide feeds than can be read by PDAs, as well as by PC or browser-based readers.

Gahran goes on to say:

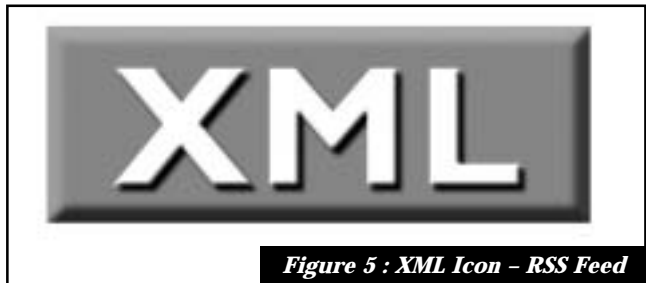
'Right now, it's used for news, but sports teams can use it to publish statistics, music groups can use it to

publish tour dates, government agencies can use it to publish regulatory updates. There are a lot of potential uses for RSS that have yet to be tapped.'

According to Paul Eng, 'The big draw of RSS for Web site publishers is that such technology helps to ensure its content is seen by readers – and helps keep the amount of traffic up on their Web sites.'

Gahran lists a number of potential new uses for RSS feeds:

- Press releases and other media alerts/materials
- Announcements of new products, beta releases, and so on
- Sales, specials, packages, and promotions
- New reports, studies, surveys, papers
- Industry, legislative, regulatory, and research progress and developments
- Incident and outage reports



Danny Sullivan has written a useful article on how to make an RSS feed. It covers the different versions of RSS, how to create a file from scratch, ways of validating it, then getting it syndicated. For people who don't want to grapple with XML, but still want to be able to start syndicating content, the RSS Headline Creator (<http://www.webdevtips.com/webdevtips/codegen/rss.shtml>) may be able to assist.

Probably the biggest problem for RSS at this point is that it is not fully integrated into the normal workings of the Web. You still need special software to read feeds, and creating them is not a simple process. However, this will presumably change as the format gathers popularity.

Eng quotes Randy Cassingham, a Colorado editor and publisher, who feels bullish about the growth and spread of RSS: 'When you look at the Web,' says Cassingham, 'it wasn't designed for fill-in forms and multimedia content a few years ago. But it grew and adapted. I think RSS will adapt too.'

He's probably right.

References

Cohen, Steven M. (2002) 'RSS For Non-Techie Librarians', [Online] <http://www.llrx.com/features/rssforlibrarians.htm/>

Eisenzopf, Jonathan 'Making Headlines with RSS: Using Rich Site Summaries To Draw New Visitors', [Online]

<http://www.newarchitectmag.com/archives/2000/02/eisenzopf/>

EEVL 'RSS Primer for Publishers and Content Providers', [Online] http://www.eevl.ac.uk/rss_primer/
Eng, Paul (2003) 'Feeding the Need for News: Software Keeps Surfers on Top of the Web's Latest Info, [Online] <http://abcnews.go.com/sections/SciTech/FutureTech/RSSWeb031029-1.html>

Gahrn, Amy (2003) 'The Challenge of RSS Evangelism', [Online] <http://blog.contentious.com/archives/000034.html>

Gahrn, Amy (2003) 'Overlooked RSS Opportunities,' [Online] <http://blog.contentious.com/archives/000035.html>

Parr, Barry (2003) 'RSS, E-mail, and the Web: Massaging the Media' [Online], <http://mediasavvy.com/archives/000417.shtml#000417>

PC Webopedia (2003) 'RSS', [Online], <http://www.webopedia.com/TERM/R/RSS.html>

Pirillo, Chris (2003) 'RSS Quickstart Guide', [Online], <http://rss.lockergnome.com/resources/articles/quickstart.phtml>

Sullivan, Danny (2003) 'RSS: Your Gateway To News & Blog Content', [Online]. <http://www.searchenginewatch.com/sereport/article.php/2175281>

Sullivan, Danny (2003) 'Making an RSS feed'. [Online]. <http://www.searchenginewatch.com/sereport/article.php/2175271>

WebReference.com (2003) 'Introduction to RSS', [Online], <http://www.webreference.com/authoring/languages/xml/rss>

Belinda Weaver is the Coordinator of ePrints@UQ, the University of Queensland Library.

NET NOTE

ELECTRONIC VERSION OF GEOSCIENCE THESAURUS AVAILABLE

Charted Information has recently released an electronic version of the Geoscience, Minerals and Petroleum Thesaurus (GeMPeT). GeMPeT includes over 9,000 terms relating to all facets of the geoscience industry; it aims to be the indexing tool of choice for geoscience professionals.

The electronic version is available on CD-ROM, and licensed for either a single site or for multiple sites, allowing it to be loaded onto corporate intranets and integrated into information management applications.

The CD-ROM provides the thesaurus in a number of formats including: comma delimited data; a Term

Tree file, suitable for loading into Term Tree thesaurus management software; and a DB/TextWorks format. Additional specialist formatting may also be available.

GeMPeT had its genesis in the *Australian Geoscience, Minerals and Petroleum Thesaurus*, previously published by the Australian Mineral Foundation - the last edition was published in 1994.

GeMPeT is priced from \$350 (plus GST). Further information, including sample pages, is available at <http://www.chartedinfo.com.au/thesaurus.html>. Discounts are available for organisations wishing to purchase both the print and electronic versions of the thesaurus.

NET NOTE

AUSTLIT UPDATE

AustLit (<http://www.austlit.edu.au/>) provides online access to selected full text creative and critical works. In conjunction with the National Library of Australia's PANDORA Archive, and the University of Sydney Library's digitising project, SETIS (Scholarly Electronic Text and Image Service), AustLit's coverage of accessible full text is increasing. AustLit currently has records relating to

over 7,000 works. Over half the available works are poems and nearly 1,000 are short stories. There are also more than 70 novels and numerous reviews and critical works.

The Full Text Help Page provides access to these and other full text resources, such as freely available e-journals and Web sites relating to authors, publishers and other relevant literary organisations.

NET NOTE

CROSSREF DROPS DOI RETRIEVAL FEES

CrossRef (<http://www.crossref.org>) was established by scholarly publishers as an independent, not-for-profit body in 2000. Its membership now represents 250 publishers. CrossRef software represented the first full scale implementation of the DOI (Digital Object Identifier) system. Now CrossRef has announced

that it will drop its DOI retrieval fees for all members and affiliates from January 2004, thereby giving all CrossRef users unlimited access to DOIs. DOI links from citations and bibliographic databases to full text are expected to increase greatly as a result. The system holds several hundred thousand book and conference proceeding DOIs, in addition to 700 books from Oxford University Press, and 8,700 journals.