

RSS

Web feed formats used to publish frequently updated content such as blog entries, news headlines, and podcasts in a standardized format. An RSS document (which is called a "feed" or "web feed" or "channel") contains either a summary of content from an associated web site or the full text. RSS makes it possible for people to keep up with web sites in an automated manner that can be piped into special programs or filtered displays.

The benefit of RSS is the aggregation of content from multiple Web sources in one place. RSS content can be read using software called an "RSS reader", "feed reader" or an "aggregator", which can be web-based or desktop-based. A standardized XML file format allows the information to be published once and viewed by many different programs. The user subscribes to a feed by entering the feed's link into the reader or 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 content, downloading any updates that it finds and provides a user interface to monitor and read the feeds.

The initials "RSS" are used to refer to the following formats:

- a) Really Simple Syndication (RSS 2.0).
- b) RDF Site Summary (RSS 1.0 and RSS 0.90).
- c) Rich Site Summary (RSS 0.91).

RSS formats are specified using XML, a generic specification for the creation of data formats although RSS formats have evolved since March 1999, the RSS icon ("") first gained widespread use in 2005/2006.

History

Screenshot of an RSS feed in Safari 2.0.4
Screenshot of an RSS feed in Safari 2.0.4

The RSS formats were preceded by several attempts at syndication that did not achieve widespread popularity. The basic idea of restructuring information about web sites goes back to as early as 1995, when Ramanathan V. Guha and others in Apple Computer's Advanced Technology Group developed the Meta Content Framework (MCF). For a more detailed discussion of these early developments, see the history of web syndication technology.

RDF Site Summary, the first version of RSS, was created by Guha at Netscape in March 1999 for use on the My.Netscape.Com portal. This version became known as RSS 0.9. In July 1999, Dan Libby of Netscape produced a new version, RSS 0.91 that simplified the format by removing RDF elements and incorporating elements from Dave Winer's scriptingNews syndication format. Libby also renamed RSS to Rich Site Summary and outlined further development of the format in a "futures document".

This would be Netscape's last participation in RSS development for eight years. As RSS was being embraced by web publishers who wanted their feeds to be used on My.Netscape.Com and other early RSS portals, Netscape dropped RSS support from My.Netscape.Com in April 2001 during new owner AOL's restructuring of the company, also removing documentation and tools that supported the format.

Two entities emerged to fill the void, with neither Netscape's help nor approval: The RSS-DEV Working Group and Winer, whose UserLand Software had published some of the first publishing tools outside of Netscape that could read and write RSS.

Winer published a modified version of the RSS 0.91 specification on the UserLand web site, covering how it was being used in his company's products, and claimed copyright to the document. A few months later, UserLand filed a U.S.

trademark registration for RSS, but failed to respond to a USPTO trademark examiner's request and the request was rejected in December 2001.

The RSS-DEV Working Group, a project whose members included Guha and representatives of O'Reilly Media and moreover, produced RSS 1.0 in December 2000. This new version, which reclaimed the name RDF Site Summary from RSS 0.9, reintroduced support for RDF and added XML namespaces support, adopting elements from standard metadata vocabularies such as Dublin Core.

In December 2000, Winer released RSS 0.92 a minor set of changes aside from the introduction of the enclosure element, which permitted audio files to be carried in RSS feeds and helped spark podcasting. He also released drafts of RSS 0.93 and RSS 0.94 that were subsequently withdrawn.

In September 2002, Winer released a major new version of the format, RSS 2.0 that redubbed its initials Really Simple Syndication. RSS 2.0 removed the type attribute added in the RSS 0.94 draft and added support for namespaces.

Because neither Winer nor the RSS-DEV Working Group had Netscape's involvement, they could not make an official claim on the RSS name or format. This has fueled ongoing controversy in the syndication development community as to which entity was the proper publisher of RSS.

One product of that contentious debate was the creation of an alternative syndication format, Atom that began in June 2003. The Atom syndication format, whose creation was in part motivated by a desire to get a clean start free of the issues surrounding RSS, has been adopted as IETF Proposed Standard RFC 4287.

In July 2003, Winer and UserLand Software assigned the copyright of the RSS 2.0 specification to Harvard's Berkman Center for the Internet & Society, where he had just begun a term as a visiting fellow. At the same time, Winer launched the RSS Advisory Board with Brent Simmons and Jon Udell, a group whose purpose was to maintain and publish the specification and answer questions about the format.

In December 2005, the Microsoft Internet Explorer team and Outlook team announced on their blogs that they were adopting the feed icon first used in the Mozilla Firefox browser. A few months later, Opera Software followed suit. This effectively made the orange square with white radio waves the industry standard for RSS and Atom feeds, replacing the large variety of icons and text that had been used previously to identify syndication data.

In January 2006, Rogers Cadenhead relaunched the RSS Advisory Board without Dave Winer's participation, with a stated desire to continue the development of the RSS format and resolve ambiguities. In June 2007, the board revised their version of the specification to confirm that namespaces may extend core elements with namespace attributes, as Microsoft has done in Internet Explorer 7. In their view, a difference of interpretation left publishers unsure of whether this was permitted or forbidden.

Incompatibilities

As noted above, there are several different versions of RSS, falling into two major branches (RDF and 2.*). The RDF or RSS 1.* branch includes the following versions:

1. RSS 0.90 was the original Netscape RSS version. This RSS was called RDF Site Summary, but was based on an early working draft of the RDF standard, and was not compatible with the final RDF Recommendation.
2. RSS 1.0 is an open format by the RSS-DEV Working Group, again standing for RDF Site Summary. RSS 1.0 is an RDF format like RSS 0.90, but not fully compatible with it, since 1.0 is based on the final RDF 1.0 Recommendation.
3. RSS 1.1 is also an open format and is intended to update and replace RSS 1.0. The specification is an independent draft not supported or endorsed in any way by the RSS-Dev Working Group or any other organization.

The RSS 2.*branch (initially UserLand, now Harvard) includes the following versions:

- RSS 0.91 is the simplified RSS version released by Netscape, and also the version number of the simplified version originally championed by Dave Winer from Userland Software. The Netscape version was now called Rich Site Summary; this was no longer an RDF format, but was relatively easy to use. It remains the most common RSS variant.
- RSS 0.92 through 0.94 are expansions of the RSS 0.91 format, which are mostly compatible with each other and with Winer's version of RSS 0.91, but are not compatible with RSS 0.90. In all Userland RSS 0.9x specifications, RSS was no longer an acronym.
- RSS 2.0.1 has the internal version number 2.0. RSS 2.0.1 was proclaimed to be "frozen", but still updated shortly after release without changing the version number. RSS now stood for Really Simple Syndication. The major change in this version is an explicit extension mechanism using XML namespaces.

For the most part, later versions in each branch are backward-compatible with earlier versions (aside from non-conformant RDF syntax in 0.90), and both versions include properly documented extension mechanisms using XML Namespaces, either directly (in the 2.* branch) or through RDF (in the 1.* branch). Most syndication software supports both branches. Mark Pilgrim's article "The Myth of RSS Compatibility" discusses RSS version compatibility in more detail.

The extension mechanisms make it possible for each branch to track innovations in the other. For example, the RSS 2.* branch was the first to support enclosures, making it the current leading choice for podcasting, and as of mid-2005 is the format supported for that use by iTunes and other podcasting software; however, an enclosure extension is now available for the RSS 1.* branch, `mod_enclosure`. Likewise, the RSS 2.* core specification does not support providing full-text in addition to a synopsis, but the RSS 1.* markup can be (and often is) used as an extension. There are also several common outside extension packages available, including a new proposal from Microsoft for use in Internet Explorer 7.

The most serious compatibility problem is with HTML markup. Userland's RSS reader—generally considered as the reference implementation—did not originally filter out HTML markup from feeds. As a result, publishers began placing HTML markup into the titles and descriptions of items in their RSS feeds. This behavior has become expected of readers, to the point of becoming a de facto standard, though there is still some inconsistency in how software handles this markup, particularly in titles. The RSS 2.0 specification was later updated to include examples of entity-encoded HTML; however, all prior plain text usages remain valid.

Modules

The primary objective of all RSS modules is to extend the basic XML schema established for more robust syndication of content. This inherently allows for more diverse, yet standardized, transactions without modifying the core RSS specification.

To accomplish this extension, a tightly controlled vocabulary (in the RSS world, "module"; in the XML world, "schema") is declared through an XML namespace to give names to concepts and relationships between those concepts.

Some RSS 2.0 modules with established namespaces:

- Ecommerce RSS 2.0 Module.
- Media RSS 2.0 Module.
- OpenSearch RSS 2.0 Module.

BitTorrent and RSS

Several BitTorrent-based peer-to-peer applications also support RSS. Such feeds (also known as Torrent/RSS-es or Torrentcasts) allow client applications to download files automatically from the moment the RSS reader detects them (also known as Broadcatching).

AGGREGATOR

This article is about a web application. For websites that compiles reviews, see review aggregator.

In computing, a feed aggregator, also known as a feed reader, news reader or simply as an aggregator, is client software or a Web application which aggregates syndicated web content such as news headlines, blogs, podcasts, and vlogs in a single location for easy viewing.

Functions

Aggregators reduce the time and effort needed to regularly check websites for updates, creating a unique information space or "personal newspaper." Once subscribed to a feed, an aggregator is able to check for new content at user-determined intervals and retrieve the update. The content is sometimes described as being "pulled" to the subscriber, as opposed to "pushed" with email or IM. Unlike recipients of some "pushed" information, the aggregator user can easily unsubscribe from a feed.

Aggregator features are frequently built into portal sites (such as My Yahoo! and iGoogle), modern Web browsers and email programs.

The aggregator provides a consolidated view of the content in a single browser display or desktop application. Such applications are also referred to as RSS readers, feed readers, feed aggregators, news readers or search aggregators. Aggregators with podcasting capabilities can automatically download media files, such as MP3 recordings. In some cases, these can be automatically loaded onto portable media players (like iPods) when they are connected to the PC.

Recently, so-called RSS-narrators have appeared, which not only aggregate text-only news feeds, but also convert them into audio recordings for offline listening.

The syndicated content an aggregator will retrieve and interpret is usually supplied in the form of RSS or other XML-formatted data, such as RDF/XML or Atom.

Varieties

The variety of software applications and components that are available to collect, format, translate, and republish XML feeds is a testament to the flexibility of the format and has shown the usefulness of presentation-independent data.

Web-based

Web-based aggregators are applications that reside on remote servers and are typically available as Web applications such as Google Reader or Bloglines. Because the application is available via the Web, it can be accessed anywhere by a user with an Internet connection.

More advanced methods of aggregating feeds are provided via AJAX coding techniques and XML components known as Web widgets. Ranging from full-fledged applications to small fragments of code that can be integrated into larger programs, they allow users to aggregate OPML files, email services, documents, or feeds into a single interface. Many customizable homepage/portal implementations such as iGoogle, Live.com, My Yahoo!, and Pageflakes provide such functionality.

In addition to personal aggregators, planet sites are used by online communities to aggregate community blogs in a centralized location. Such sites are named after the Planet aggregator, an application designed for this purpose.

Client software

Client software aggregators are installed applications designed to collect Web feed subscriptions and group them together using a user-friendly interface. The graphical user interface of such applications often closely resembles that of popular e-mail clients, using a three-panel composition in which subscriptions are grouped in a frame on the left, and individual entries are browsed, selected, and read in frames on the right.

Software aggregators can also take the form of news tickers which scroll feeds like ticker tape, alerters that display updates in windows as they are refreshed, web browser macro tools or as smaller components (sometimes called plugins or extensions), which can integrate feeds into the Operating System or software applications such as a Web browser.

Client Libraries

Many programming languages have libraries that are able to download, process, generate and upload RSS feeds. Perl for example has several libraries in the XML::RSS name space of CPAN.

Media aggregators

Media aggregators are sometimes referred to as "Podcatchers" due to the popularity of the term "podcast" used to refer to a web feed containing audio or video. Media aggregators refer to applications, client software or Web based, which maintain subscriptions to feeds that contain audio or video media enclosures. They can be used to automatically download media, playback the media within the application interface, or synchronize media content with a portable media player.

Feed filtering

One of the problems with news aggregators is that the volume of articles can sometimes be overwhelming, especially when the user has many Web feed subscriptions. As a solution, many feed readers allow users to tag each feed with one or more keywords which can be used to sort and filter the available articles into easily navigable categories. Another option is to import the user's Attention Profile to filter items based on their relevance to the user's interests.