

RSS 3.0 “Really Simple Syndication”

Publisher’s Note

It is difficult to say whether Aaron’s Swartz RSS 3.0 specification is or will become a “standard” or even a draft standard. Because it is cited and because some software has been developed to author or display RSS 3.0, it is necessary to include the “specification” so authors and software developers can have an authoritative source of information.

To put the “specification” in context, Swartz’ “The Road to RSS 3.0”, an introduction to RSS3.0, has been included. Some comments about this work are also included to provide background and opinions.

Swartz describes himself writing: “Aaron Swartz is a teenage writer, coder, and hacker. He was a finalist for the ArsDigita Prize for excellence in building non-commercial web sites at the age of 13. At 14 he co-authored the RSS 1.0 specification, now used by thousands of sites to notify their readers of updates. He's a member of the W3C's RDF Core Working Group which is developing the format for the Semantic Web and Metadata Advisor to the Creative Commons. He's also the author of rss2email, xmltramp, HTML diff, and html2text.” He was awarded the Apple World Wide Developers Conference Student Scholarship (an waiver to attend underage), the ArsDigita prize, and in 2000 the four-year Scholarship For Excellence at North Shore Country Day School.¹ He expects to attend Stanford University Fall 2004.

¹ For additional information see Swartz’ Webpage at <http://www.aaronsw.com/about> viewed 4 July 2004.

Background and Context

The Road to RSS 3.0

Edd Dumbill: RSS Moves Forward. “The goal of RSS 1.0 has been to fix some problems, provide an extensible framework for the future, and bring RSS into community ownership. [...] RSS 1.0 provides a solid framework on which to build.”

Dave Winer: The Road to RSS 2.0. “And further, there’s talk all over the place about RSS 2.0, a belief that now’s the time to really get RSS on a strong foundation, one that’s solid and frozen, and at the same time extensible.”

There’s been a lot of talk in the community about how RSS 2.0 is too complicated. I haven’t heard any objections, so I’m going to move ahead with the following changes that will result in RSS 3.0.

1. Remove XML. XML is just too complicated and is against the spirit of RSS, which is Really Simple Syndication. I don’t want people to have to buy one of these 200 page XML books to understand RSS. And XML sucks up bandwidth like nobody’s business. Instead, we’ll go back to RFC822-style fields. There are lots of available parsers for those.
2. Remove namespaces. Namespaces are just a waste of time. If people want to add an element to RSS, then just send it to me and I’ll add it to my list of all elements in use. This system is easy to use and doesn’t result in any wasteful URIs all over the place.
3. HTML forbidden. No one needs HTML. Email has been just fine for years before Microsoft introduce their stupid rich HTML extensions. HTML is for those loser newbies. Any intelligent Internet user deals in plain text.

I’ll leave some time for comment and then put up a spec. Then we’ll deploy.

Aaron Swartz, www.aaronsw.com/weblog/000574, September 6, 2002.

Aaron is Awesome

There’s been a lot of talk in the community about how RSS 2.0 is too complicated. I haven’t heard any objections, so I’m going to move ahead with the following changes that will result in RSS 3.0.

Andrew Wooster, nextthing.org/archive.php?date=2002-09-07#8-AaronIsAwesome, September 7, 2002

RSS 3.0?

Aarom Swartz proposed RSS 3.0, a complete waste of time and a lame application of inverse extremism to RSS 2.0.²

Dave Park, docuverse.com/blog/donpark/default.aspx?date=2002-09-09, September 9, 2002

“I'm pretty sure RSS 3.0 is mostly an exasperated joke from the midst of the Great RSS War.

“On the plus side, it certainly has a good claim to the Really Simple Syndication name.”

Todd Larason, molelog.molehill.org/blox/Computers/Internet/Web/RSS/RSS30.writeback,
4 February 2003.

A letter to Harvard about Syndication

A Note to Harvard University:

Dave Winer handed you a major gift when he turned RSS 2.0 over to Harvard. One that, I'm sure, you might not yet fully appreciate.

I've seen the light. Syndication will clearly be a major part of what happens next in the computer world. Already my ability to read Web sites has increased ten fold (I now read about 640 RSS feeds in the time it used to take me to read less than 60 HTML-based Web sites).

You think RSS isn't changing things? Heck, just look at politics. Here's a new RSS news aggregator that one of the top presidential candidates, Howard Dean, is using to push news out to his followers.

The fact that Harvard now owns the RSS specification will let Harvard play in a whole new realm of technology that our society will use. That is if Harvard doesn't blow it between now and 2005.

That's what this letter is all about.

Today Harvard's spec, RSS 2.0, is the leader in the syndication race. But, if everything remains the way it is today, RSS won't be on top for long.

Why not?

² On August 28, 2002 Dave Park defined “inverse extremism” saying: “Inverse Extremism: Extremism can be useful when used appropriately, but is not normally effective in its usual form which attempts to pull opinions from the middle toward one of the two edges or extremes. More effective form of extremism is Inverse Extremism which attempts to push opinions away from the opposite extreme. Inverse Extremism is effective because it relies on [negative] emotions to repulse subjects instead of logic or inference to attract subjects. “For example, instead of lobbying against abortion, one could form an Inverse Lobbying organization that propagates extreme abortionist views. In religious terms, this is equivalent to becoming the demon instead of demonizing others. “disclaimer: Inverse Extremism, as I have described it here, is entirely of my imagination. If this concept has already been described somewhere else before, please let me know.”

Because the market is changing. Just over the weekend there was a corner turn in the Atom camp. Atom is a format (and an API) that competes with RSS. Why is that? Because Atom started with the RSS spec and improved on it. What was the corner turn? Over the weekend Sam Ruby shipped a set of slides that spelled out quite clearly just how it is better.

That alone didn't mean much. But, today, my favorite news aggregator (NewsGator) supports both Atom and RSS. NewsGator is built on Microsoft's .NET platform. Why is that important? Well, today it might not seem to be. But, we're building our next version of Windows (code-named Longhorn) and Longhorn gives tons of new capabilities to .NET developers that haven't existed before.

Why is that a problem? Because Microsoft's developers are starting to compare RSS 2.0 and Atom and I'm seeing more and more of them switch to Atom because of the advantages laid out in Sam Ruby's slides.

What does that trend mean? Well, the value of the gift that Dave Winer gave you is going down every day. It might not look important today. Very few people are supporting Atom today. Well, except for Google, Six Apart, and IBM. Do they matter to this industry? Will the products they ship have an impact on the weblogging and syndication markets? To the Internet itself? You betcha!

Which is why I'm writing this letter. It's a roadmap of how Harvard will end up being the syndication leader in 2006, instead of Atom, er Google and IBM.

Here's what I'd do if I were at Harvard and in charge of the RSS spec:

- 1) Announce there will be an RSS 3.0 and that it will be the most thought-out syndication specification ever.
- 2) Announce that RSS 3.0 will ship on July 1, 2005. That date is important. For one, 18 months is long enough to really do some serious work. For two, RSS 3.0 should be positioned as "the best way to do syndication on Microsoft's Longhorn." The betas for Longhorn should really be rocking by that date, so you'll have tons of new developers trying to build innovative things for Longhorn. More on that later. For three, it would freeze the market for 18 months because "Mr. Safe" will not want to move away from RSS before he sees what the future of RSS will be. Also, "Mr. Safe" will want to stick on a platform that will be compatible with RSS 3.0. Today that platform is RSS 2.0.
- 3) Open up a mailing list, a wiki, and a weblog to track progress on RSS 3.0 and encourage community inclusion.
- 4) Work with Microsoft to ensure that RSS 3.0 will be able to take advantage of Longhorn's new capabilities (in specific, focus on learning Indigo and WinFS). Build a prototype (er, have MSN build one) that would demonstrate some of the features of RSS 3.0 -- make this prototype so killer that it gets used on stage at the Longhorn launch (in fact, make it even better than that, so it gets included with every copy of Longhorn that's shipped).

- 5) Make sure RSS 3.0 is simply the best-of-breed syndication protocol. Translation: don't let Microsoft or Google come up with a better spec that has more features.

Why would you do all of this?

Well, imagine what'll happen to Harvard's name recognition if your syndication format gets demonstrated on stage by Bill Gates? Imagine where future software engineering students will want to attend. Harvard or Stanford? Hmmm. Stanford generated Google. You do the math. How much does a single student pay nowadays? \$150,000+ to attend Harvard for four years? How many students decide to attend Stanford because that's where Google and Yahoo were started?

But, it'd take some vision. It'd take some chutzpah.

Of course, if you don't have the vision, that's OK. Atom is there to take over if you fumble the football.

Robert Soble, radio.weblogs.com/0001011/2003/12/a5764, 15 December 2003

Harvard active in RSS 3.0? Not a chance.

Scoble is way off in his suggestion that Harvard take an active role in the development of RSS 3.0.³ Giving a piece of intellectual property to Harvard is much like giving it to the pope -- it is a pair of safe hands without armies to wield the power it has been granted. Harvard is a highly decentralized organization, carefully divided into various academic fiefs. The central administration has little control over the other 95%, other than the funding stream from surprisingly small portions of the endowment. And appointment of tenured faculty. There's no one "in charge" of the RSS spec. It didn't get magically added to someone's job description, the way it would if it had been given to a Fortune 500 corporation. Harvard is a passive partner that would likely allow you to hire lawyers and sue in its name over RSS if you really wanted to and if there was an academically justifiable reason.

John Stafford, stafford.typepad.com/the_next_america/2003/12/harvard_active_.html, 16 December 2003

RSS 3.0

RSS is, in short, a great idea. But in order to spread its adoption from blogs to large-scale publications, a few extra tags would help. All of these are parameters us BigPub media whores care about, as do our readers. This information, encoded as tags rather than random text strings in an item, would help human readers - or better yet, their software - decide whether or not to prioritize and read an item.

byline - the current author tag is specified as an email address only, in case the reader wants to reply. But we know that readers are drawn to specific authors, e.g. Bruce

³ According to scoble.weblogs.com, 4 July 2004, "Robert Scoble works at Microsoft as an evangelist on the Windows team."

Sterling vs Mickey Kaus vs Paul Boutin. If P.J. O'Rourke or Michael Crichton did a guest article on Slate, none of the RSS audience would know. It would be great if publications could identify the name (rather than the email) of the writer, so RSS fans can keep up on their favorite authors, subscribe to them, etc, and likewise avoid those they don't want to read.

publication - This seems obvious, especially for aggregated feeds like Channel Dean. I can't tell what's originally from the NY Times, Fox News, Mother Jones, or Dave's blog without clicking through. Relying on a human editor to include that info in the summary field seems like a bad idea.

wordcount - Less mission-critical than the above two, yet relevant. Is this a 300 word paragraph, a three thousand word magazine feature, or a 30,000-word book? That would help readers decide what to read and when. Granted, most casual readers don't know how long 300 words is, but I'm pretty sure they would develop a sense of that over time, given daily examples. Perhaps this could be length in bytes or something, but that gets confusing with enclosures and markup involved. Do more people (as opposed to more developers) grok bytes or wordcount? I don't know.

section - Front page? World? National? Local? Arts? Technology? Webhead? Gearbox? Slate and others use category for this. But unless I misunderstand the spec, a category is meant to be something larger than a publication, rather than a subset of a single pub.

I'm sure everyone has a tag they want added to the spec. But as a writer, rather than a publisher or a nameless employee of one, these are the important parameters to me and my readers. As metadata, they would help us decide rather or not an item in a feed is worth looking at, based on important traits. But they're not captured by the current set of RSS 2.0 tags. I hardly have time to get involved in it, but there's my two cents.

Posted on 3/8/04; 10:06:36 PM

Paul Boutin, paulboutin.weblogger.com/2004/03/08, March 8, 2004.
He writes for Slate magazine.

RSS 3.0 [The specification]⁴

Introduction (you should probably read [it](#) first). *This spec is not yet finalized. Feedback appreciated.*

Format

An **item** consists of a series of **lines** separated by "\n".

Each **line** is a series of letters, numbers, "-", ".", or "_" (called the **name**) followed by ": " followed by a series of characters (called the **value**). No two lines should start with the same name. If a line starts with a space or tab character, then it is a continuation of the value on the previous line. The newline in between is preserved. UTF-8 encoding is always used.

An item ends at the first blank line (that is, a line with no characters).

Document

An RSS 3.0 document consists of one head item followed by zero or more body items.

Head

The head is an **item**. Names for the lines are globally assigned. Names are case-insensitive. The assigned names are:

- title
- description
- link
- generator
- errorsto
- creator
- created
- last-modified
- language
- rights
- license
- guid
- uri
- subject

Most properties refer to the whole feed in addition to the item. i.e. last-modified is the last-modified date of the feed.

Body

The body is a series of zero or more items. Names for the lines are globally-assigned and case-insensitive. The assigned names are:

⁴ Posted 6 September 2002 at <http://www.aaronsw.com/2002/rss30>.

title
description
link
generator
creator
created
last-modified
language
rights
license
guid
uri
subject

Tokens

title

The title of the item.

description

A short description of the item.

link

A link to the item.

generator

The person or program that generated the item.

errorsto

An email address, optionally followed by a space and a name, of the person to send error reports about the feed to.

creator

An email address, optionally followed by a space and a name, of the person who created the item.

created

The date (in W3CDTF format) the item was created.

last-modified

The date (in W3CDTF format) the item was modified.

language

The language of the item, using the language tag format specified in RFC 3066.

rights

The copyright statement for the item.

license

A URI for the copyright license of the item.

guid

A globally unique identifier for the item.

uri

A globally unique identifier in the form of a URI for the item.

subject

The topic of the item.

Example

```
title: RSS 3.0 News
description: Latest updates on RSS 3.0.
link: http://www.aaronsw.com/2002/rss30
creator: me@aaronsw.com Aaron Swartz
errorsTo: me@aaronsw.com Aaron Swartz
language: en-US
```

```
title: Spec Introduced
created: 2002-09-06
guid: 00795648-C1E0-11D6-9AA6-003065F376B6
description:
  The spec was introduced to the world.
```

A few people noticed.

```
Title: Zooko Likes It
Created: 2002-09-06
GUID: 0894CB2F-C1E0-11D6-9649-003065F376B6
Description: Zooko says he likes the spec.
```