Network Working Group
Request for Comments: 4685
Category: Standards Track

J. Snell

September 2006

Atom Threading Extensions

Status of This Memo

Copyright Notice

Abstract

This memo presents a mechanism that allows feeds publishers to express threaded discussions within the Atom Syndication Format.

Table of Contents

1. Introduction

This document defines an extension for expressing threaded discussions within the Atom Syndication Format [RFC4287].

2. Notational Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, [RFC2119], as scoped to those conformance targets.

The XML Namespaces URI [W3C.REC-xml-names-19990114] for the XML elements and attributes described in this specification is: http://purl.org/syndication/thread/1.0

In this document, the namespace prefix "thr:" is used for the above Namespace URI. Note that the choice of namespace prefix is arbitrary and not semantically significant.

This specification uses a shorthand form of terms from the XML Infoset [W3C.REC-xml-infoset-20040204]. The phrase "Information Item" is omitted when naming Element and Attribute Information Items. Therefore, when this specification uses the term "element," it is referring to an Element Information Item in Infoset terms. Likewise, when this specification uses the term "attribute," it is referring to an Attribute Information Item.

This specification allows the use of IRIs [RFC3987]. Every URI [RFC3986] is also an IRI, so a URI may be used wherever an IRI is named. When an IRI that is not also a URI is given for dereferencing, it MUST be mapped to a URI using the steps in Section 3.1 of [RFC3987]. When an IRI is serving as an identifier, it MUST NOT be so mapped.

Some sections of this specification are illustrated with a non-normative RELAX NG Compact schema [RELAXNG]. In those sections, this specification uses the atomCommonAttributes, atomMediaType, and atomURI patterns, defined in [RFC4287].

However, the text of this specification provides the sole definition of conformance.

3. The ’in-reply-to’ Extension Element

The "in-reply-to" element is used to indicate that an entry is a response to another resource. The element MUST contain a "ref" attribute identifying the resource that is being responded to.

The element is not unlike the references and in-reply-to email message headers, defined by [RFC2822]. However, unlike the in-reply-to header, the "in-reply-to" element is required to identify the unique identifier of only a single parent resource. If the entry

is a response to multiple resources, additional "in-reply-to" elements MAY be used. There is no direct equivalent to the references header, which lists the unique identifiers of each preceding message in a thread.

```
in-reply-to =
  element thr:in-reply-to {
    atomCommonAttributes,
    ref,
    href?,
    source?,
    type?,
    ( undefinedContent )
  }

ref = attribute ref { atomURI }
href = attribute href { atomURI }
type = attribute type { atomMediaType }
source = attribute source { atomURI }
```

The "ref" attribute specifies the persistent, universally unique identifier of the resource being responded to. The value MUST conform to the same construction and comparison rules as the value of the atom:id element, as defined in Section 4.2.6 of [RFC4287]. Though the IRI might use a dereferenceable scheme, processors MUST NOT assume that it can be dereferenced.

If the resource being responded to does not have a persistent, universally unique identifier, the publisher MUST assign an identifier that satisfies all the considerations in Section 4.2.6 of [RFC4287] for use as the value of the "ref" attribute. In that case, if a representation of the resource can be retrieved from an IRI that can be used as a valid atom:id value, then this IRI SHOULD be used as the value of both the "ref" and "href" attributes.

The "source" attribute MAY be used to specify the IRI [RFC3987] of an Atom Feed or Entry Document containing an atom:entry with an atom:id value equal to the value of the "ref" attribute. The IRI specified, once appropriately mapped to a corresponding URI, MUST be dereferenceable.

The "href" attribute specifies an IRI that may be used to retrieve a representation of the resource being responded to. The IRI specified, once appropriately mapped to a corresponding URI, MUST be dereferenceable.

The "type" attribute MAY be used to provide a hint to the client about the media type [RFC4288] of the resource identified by the "href" attribute. The "type" attribute is only meaningful if a corresponding "href" attribute is also provided.

This specification assigns no significance to the order in which multiple "in-reply-to" elements appear within an entry.

An example of an entry with a response follows:

```
<feed xmlns="http://www.w3.org/2005/Atom"
      xmlns:thr="http://purl.org/syndication/thread/1.0">
  <id>http://www.example.org/myfeed</id>
  <title>My Example Feed</title>
  <updated>2005-07-28T12:00:00Z</updated>
  <link href="http://www.example.org/myfeed" />
  <author><name>James</name></author>
  <entry>
    <id>tag:example.org,2005:1</id>
    <title>My original entry</title>
    <updated>2006-03-01T12:12:12Z</updated>
    <link
      type="application/xhtml+xml"
      href="http://www.example.org/entries/1" />
    <summary>This is my original entry</summary>
  </entry>
  <entry>
    <id>tag:example.org,2005:1,1</id>
    <title>A response to the original</title>
    <updated>2006-03-01T12:12:12Z</updated>
    <link href="http://www.example.org/entries/1/1" />
    <thr:in-reply-to
      ref="tag:example.org,2005:1"
      type="application/xhtml+xml"
      href="http://www.example.org/entries/1"/>
    <summary>This is a response to the original entry</summary>
  </entry>
</feed>
```

To allow Atom processors that are not familiar with the in-reply-to extension to know that a relationship exists between the entry and the resource being responded to, publishers are advised to consider including a "related" link referencing a representation of the resource identified by the in-reply-to element. Although such links are unlikely to be processed as a reference to a predecessor in a threaded conversation, they are helpful in at least establishing a semantically meaningful relationship between the linked resources.

For example,

```
<feed xmlns="http://www.w3.org/2005/Atom"
      xmlns:thr="http://purl.org/syndication/thread/1.0">
  <id>http://www.example.org/myfeed</id>
  <title>My Example Feed</title>
  <updated>2005-07-28T12:00:00Z</updated>
  <link href="http://www.example.org/myfeed" />
  <author><name>James</name></author>
  <entry>
    <id>tag:example.org,2005:1,1</id>
    <title>A response to the original</title>
    <updated>2006-03-01T12:12:12Z</updated>
    <link href="http://www.example.org/entries/1/1" />
    <thr:in-reply-to
      ref="tag:example.org,2005:1,0"
      type="application/xhtml+xml"
      href="http://www.example.org/entries/1"
      source="http://www.example.org/myfeed" />
    <link
      rel="related"
      type="application/xhtml+xml"
      href="http://www.example.org/entries/1" />
    <summary>This is a response to the original entry</summary>
  </entry>
</feed>
```

4. The 'replies' Link Relation

An Atom link element with a rel attribute value of "replies" may be used to reference a resource where responses to an entry may be found. If the type attribute of the atom:link is omitted, its value is assumed to be "application/atom+xml".

A "replies" link appearing as a child of the Atom feed or source element indicates that the referenced resource likely contains responses to any of that feed's entries. A "replies" link appearing as a child of an Atom entry element indicates that the linked resource likely contains responses specific to that entry.

An atom:link element using the "replies" rel attribute value MAY contain a "thr:count" attribute whose value is an unsigned, non-negative integer, conforming to the canonical representation of the XML Schema nonNegativeInteger data type [W3C.REC-xmlschema-2-20041028], that provides a hint to clients as to the total number of replies contained by the linked resource. The value is advisory and may not accurately reflect the actual number of replies.

The link MAY also contain a "thr:updated" attribute, whose value is a [RFC3339] date-time stamp conforming to the same construction rules as the Atom Date Construct defined in [RFC4287], and is used to provide a hint to clients as to the date and time of the most recently updated reply contained by the linked resource. The value is advisory and may not accurately reflect the actual date and time of the most recent reply.

For example,

```
<feed xmlns="http://www.w3.org/2005/Atom"
      xmlns:thr="http://purl.org/syndication/thread/1.0">
  <id>http://www.example.org/myfeed</id>
  <title>My Example Feed</title>
  <updated>2005-07-28T12:00:00Z</updated>
  <link href="http://www.example.org/myfeed" />
  <author><name>James</name></author>
  <entry>
    <id>tag:entries.com,2005:1</id>
    <title>My original entry</title>
    <updated>2006-03-01T12:12:12Z</updated>
    <link href="http://www.example.org/entries/1" />
    <link rel="replies"
          type="application/atom+xml"
          href="http://www.example.org/mycommentsfeed.xml"
          thr:count="10" thr:updated="2005-07-28T12:10:00Z" />
    <summary>This is my original entry</summary>
  </entry>
</feed>
```

Although Atom feed, entry, and source elements MAY each contain any number of atom:link elements using the "replies" link relation, this specification assigns no significance to the presence or order of such links. Multiple replies links appearing within an atom:entry may reference alternative representations of the same set of responses or may reference entirely distinct resources containing distinct sets of responses. Processors MUST NOT assume that multiple replies links are referencing different representations of the same resource and MUST process each replies link independently of any others.

5. The 'total' Extension Element

The "total" element is used to indicate the total number of unique responses to an entry known to the publisher. Its content MUST be an unsigned non-negative integer value conforming to the canonical representation of the XML Schema nonNegativeInteger data type [W3C.REC-xmlschema-2-20041028].

```
   total = element thr:total { xsd:nonNegativeInteger }
```

Atom entries MAY contain a "total" element but MUST NOT contain more than one.

There is no implied relationship between the value of the "total" element of an Atom entry and any individual or aggregate values of the "thr:count" attributes of its Atom link elements having a "replies" relation.

6. Considerations for Using thr:count, thr:updated, and total

The thr:count, thr:updated, and total extensions provide additional metadata about the thread of discussion associated with an entry. The values are intended to make it easier for feed consumers to display basic contextual information about the thread without requiring that those consumers dereference, parse, and analyze linked resources. That said, there are a number of considerations implementors need to be aware of.

First, these extensions MUST NOT be assumed to provide completely accurate information about the thread of discussion. For instance, the actual total number of responses contained by a linked resource MAY differ from the number specified in the thr:count attribute. Feed publishers SHOULD make an effort to ensure that the values are accurate. The non-authoritative nature of "external reference metadata", like the replies link attributes, is discussed in detail in Section 3.3 of the W3C document "Tag Finding 12: Authoritative Metadata" [TAG12].

Second, the values of the these extensions are volatile and may change at a faster rate than that of the containing entry. Frequent updates to these values, or to any part of the Atom document, could have a detrimental impact on the cacheability of the document using the attributes, leading to an increase in overall bandwidth consumption.

Feed publishers SHOULD consider a change to the values of the thr: count, thr:updated, and total extensions an "insignificant" update in terms of [RFC4287], meaning that the value of the containing feed, entry, or source element’s atom:updated element SHOULD NOT be affected by a change to the values of these extensions.

Lastly, implementors need to be aware that although the Atom specification [RFC4287] explicitly allows the link element to contain arbitrary extensions, the specification does not require that implementations support such extensions. Specifically, relating to the use of extensions, Atom does not define any level of mandatory

conformance on the part of feed consumers beyond a requirement that implementations ignore any extension the implementation does not understand. As a result, some implementations MAY NOT be capable of fully utilizing the extensions defined by this or any specification.

7. Security Considerations

As this specification defines an extension to the Atom Syndication Format, it is subject to the same security considerations defined in [RFC4287].

Feeds using the mechanisms described here could be crafted in such a way as to cause a consumer to initiate excessive (or even an unending sequence of) network requests, causing denial of service (to the consumer, the target server, and/or intervening networks). Consumers can mitigate this risk by requiring user intervention after a certain number of requests, or by limiting requests either according to a hard limit, or with heuristics.

The mechanisms described here can be used to construct threaded conversations spanning resources distributed across multiple domains. For example, an individual posting an entry to one weblog hosted on one Internet domain could mark that entry as a response to an entry from a different weblog hosted on a different domain. Implementors should note that such distributed responses can be leveraged by an attacker to attach inappropriate or unwanted content to a discussion. Such attacks can be prevented or mitigated by allowing users to explicitly configure the sources from which responses may be retrieved, or by applying heuristics to determine the legitimacy of a given response source.

Implementors should also note the potential for abuse that exists when malicious content publishers edit or change previously published content. In closed, centralized comment systems, after-the-fact editing of comments is typically not an issue, as such changes are easily prevented, detected, or tracked. With the form of distributed comments enabled through the use of the thr:in-reply-to extension, however, such changes become more difficult to detect, raising the possibility of serious attribution and repudiation concerns. XML Digital Signatures, as specified in Section 5.1 of [RFC4287], present one possible avenue for mitigating such concerns, although the presence of a valid XML Digital Signature within an entry is not, by itself, a reliable defense against repudiation issues.

8. IANA Considerations

This specification defines one new Atom link relation type that has been registered in the IANA Registry of Link Relation, as defined by [RFC4287].

Attribute Value: replies
Description: (see Section 4)
Expected display characteristics: (see Section 4)
Security considerations: (see Section 5)

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3339] Klyne, G. and C. Newman, "Date and Time on the Internet: Timestamps", RFC 3339, July 2002.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

[RFC3987] Duerst, M. and M. Suignard, "Internationalized Resource Identifiers (IRIs)", RFC 3987, January 2005.

[RFC4287] Nottingham, M. and R. Sayre, "The Atom Syndication Format", RFC 4287, December 2005.

[RFC4288] Freed, N. and J. Klensin, "Media Type Specifications and Registration Procedures", BCP 13, RFC 4288, December 2005.

[W3C.REC-xml-infoset-20040204] Tobin, R. and J. Cowan, "XML Information Set (Second Edition)", W3C REC REC-xml-infoset-20040204, February 2004.

[W3C.REC-xml-names-19990114] Hollander, D., Bray, T., and A. Layman, "Namespaces in XML", W3C REC REC-xml-names-19990114, January 1999.

[W3C.REC-xmlschema-2-20041028] Malhotra, A. and P. Biron, "XML Schema Part 2: Datatypes Second Edition", W3C REC REC-xmlschema-2-20041028, October 2004.

9.2. Informative References

[RELAXNG] Clark, J., "RELAX NG Compact Syntax", December 2001, <http://www.oasis-open.org/committees/relax-ng/compact-20021121.html>.

[RFC2822] Resnick, P., "Internet Message Format", RFC 2822, April 2001.

[TAG12] Fielding, R. and I. Jacobs, "Tag Finding 12: Authoritative Metadata", <http://www.w3.org/2001/tag/doc/mime-respect-20060412>.

Appendix A. Acknowledgements

The author gratefully acknowledges the feedback from Antone Roundy, Aristotle Pagaltzis, Byrne Reese, David Powell, Eric Scheid, James Holderness, John Panzer, Lisa Dusseault, M. David Peterson, Sam Ruby, Sylvain Hellegouarch, and the remaining members of the Atom Publishing Format and Protocol working group during the development of this specification. Any fault or weakness in the definition of this extension is solely the blame of the author.

Some portions of text in this document have been adapted from [RFC4287] in order to maintain a stylistic and technical alignment with that specification.

Author's Address

James M Snell

EMail: jasnell@gmail.com
URI: http://www.snellspace.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).