Network Working Group
Internet-Draft
Expires: June 9, 2006

J. Snell
December 6, 2005

Atom Link No Follow
draft-snell-atompub-feed-nofollow-04.txt

Status of this Memo

Copyright Notice

Abstract

This memo presents a mechanism that allows feed publishers to express preferences over how a consumer processes Atom links and Content-By-Reference.

Table of Contents

1. Introduction

This docoument specifies a mechanism by which feed publishers MAY express how applications consuming Atom documents should handle links and referenced content. For example, a publisher may include an enclosure link within a feed but may not wish for applications to automatically download the enclosed file when it processes the feed; or, the publisher may not wish to allow applications to archive or index the enclosure in any way. The 'follow', 'index' and 'archive' attributes introduced herein provide the means for publishers to express these preferences.

```
noFollowAttributes = {
  attribute follow { "yes" | "no" }?,
  attribute index { "yes" | "no" }?,
  attribute archive { "yes" | "no" }?
}
```

Section 6.3 of the Atom Format specification indicates that Atom processors that encounter unknown extensions MUST ignore those extensions without altering their behavior. Because of this requirement, there can be no assumption that a particular software implementation will support the extensions defined herein.

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.

In this specification, "entry" refers to an atom:entry element.

In this specification, "feed" refers to an Atom Feed Document.

In this specification, "head section" refers to the children of a feed's document-wide metadata container; e.g., the child elements of the atom:feed element in an Atom Feed Document.

In this specification, the term "extension attribute" refers to a namespace qualified element attribute.

In this specification, the term "link" refers to an atom:link element.

In this specification, the term "referenced content" refers to an atom:content element that contains a @src attribute.

This specification uses XML Namespaces [W3C.REC-xml-names-19990114] to uniquely identify XML element and attribute names. It uses the following namespace prefix for the indicated namespace URI;

{Ed. Note: The Namespace must be changed before publication to reflect a proper IETF namespace scheme}
"x": "http://purl.org/atompub/nofollow/1.0"

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

3. The 'x:follow' extension attribute

The 'x:follow" attribute indicates whether applications should automatically attempt to follow links and referenced content (e.g., whether or not enclosure links should be automatically downloaded, etc). The value of the attribute is either "yes" or "no". If missing, the value is considered to be indeterminate. A value of "no" indicates that applications SHOULD NOT attempt to automatically resolve the referenced resource -- rather, the application should wait until a user explicitly requests the linked resource to be resolved.

For example,

```
<entry xmlns="http://www.w3.org/2005/Atom"
       xmlns:x="http://purl.org/atompub/nofollow/1.0">
  <id>...</id>
  <title>...</title>
  <link rel="enclosure"
        href="http://example.com/mediafiles/mysong.mp3"
        x:follow="no" />
  ...
</entry>
```

The 'x:follow' attribute MAY be contained by atom:link elements and atom:content elements that contain a 'src' attribute.

4. The 'x:index' extension attribute

The 'x:index' attribute indicates whether applications should index links and referenced content. The value of the attribute is either "yes" or "no". If missing, the value is considered to be indeterminate. A value of "no" indicates that applications SHOULD

NOT index the referenced resource.

For the sake of this specification, 'indexing' refers to the practice of parsing and processing the content of a resource for the purpose of populating a database used for searching and other forms of analysis. The intended purpose of using x:index="no" would be for the publisher to indicate their preference that the associated resource not be processed for searching or analysis purposes.

```
  <entry xmlns="http://www.w3.org/2005/Atom"
         xmlns:x="http://purl.org/atompub/nofollow/1.0">
    <id>...</id>
    <title>...</title>
    <link rel="enclosure"
          href="http://example.com/mediafiles/mysong.mp3"
          x:index="no" />
    ...
  </entry>
```

The 'x:index' attribute MAY be contained by atom:link elements and atom:content elements containing a 'src' attribute.

5. The 'x:archive' extension attribute

The 'x:archive' attribute indicates whether applications should archive the targets of links and content references. The value of the attribute is either "yes" or "no". If missing, the value is considered to be "yes". A value of "no" indicate that applications SHOULD NOT archive the referenced resource.

For the sake of this specification, 'archiving' refers to the practice of maintain a local copy of a resource as part of a historical record. This is different than the practice of maintaining locally cached copies of a resource for the sake of improving transmission performance and reducing network bandwidth. The intended purpose of using x:archive="no" would be for a publisher to indicate their preference that local copies of the asociated resource not be maintained for archival/historical purposes.

```
  <entry xmlns="http://www.w3.org/2005/Atom"
         xmlns:x="http://purl.org/atompub/nofollow/1.0">
    <id>...</id>
    <title>...</title>
    <link rel="enclosure"
          href="http://example.com/mediafiles/mysong.mp3"
          x:archive="no" />
    ...
  </entry>
```

The 'x:archive' attribute MAY be contained by atom:link elements and atom:content elements containing a 'src' attribute.

6. Security Considerations

There are no security considerations introduced by this specification.

7. IANA Considerations

There are no IANA considerations introduced by this specification.

8. References

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

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

Appendix A. Acknowledgements

The author gratefully acknowledges the feedback from the members of the Atom Publishing Format and Protocol working group during the development of this specification.

Author's Address

James M Snell

Phone:
Email: jasnell@gmail.com
URI: http://snellspace.com

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.