Link Relations and HTTP Header Linking

draft-nottingham-http-link-header-03

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on June 4, 2009.

Abstract

This document specifies relation types for Web links, and defines a registry for them. It also defines how to send such links in HTTP headers with the Link header-field.
Table of Contents

1. Introduction .................................................. 3
2. Notational Conventions ........................................ 3
3. Links .......................................................... 3
4. Link Relation Types ........................................... 4
5. The Link Header Field ........................................ 5
6. IANA Considerations .......................................... 6
7. Security Considerations ........................................ 10
8. References ..................................................... 10
   8.1. Normative References .................................... 10
   8.2. Informative References ................................... 11
Appendix A. Notes on Using the Link Header with HTML .......... 11
Appendix B. Notes on Using the Link Header with Atom .......... 12
Appendix C. Acknowledgements ..................................... 13
Appendix D. Document history ..................................... 13
Author’s Address .................................................. 14
Intellectual Property and Copyright Statements .................. 15
1. Introduction

A means of indicating the relationships between documents on the Web, as well as indicating the type of those relationships, has been available for some time in HTML [W3C.REC-html401-19991224], and more recently in Atom [RFC4287]. These mechanisms, although conceptually similar, are separate. However, links between resources need not be format-specific; it can be useful to have typed links that are independent of the format, especially when a resource has representations in multiple formats.

This document defines typed link relations, independent of the context they occur in. It does so by clarifying the status of the link relation registry established by Atom, and registering in it the relations that are defined by HTML.

Furthermore, an HTTP header-field for conveying typed links was defined in [RFC2068], but removed from [RFC2616], due to a lack of implementation experience. Since then, several use cases for doing so have surfaced. However, because it was removed, the status of the Link header is unclear, leading some to consider minting new application-specific HTTP headers instead of reusing it. This document addresses this by re-specifying the Link header with updated but backwards-compatible syntax.

[[ Feedback is welcome on the ietf-http-wg@w3.org mailing list, although this is NOT a work item of the HTTPBIS WG. ]]

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.

This document uses the Augmented Backus-Naur Form (ABNF) notation of [RFC2616], and explicitly includes the following rules from it: quoted-string, token, SP (space). Additionally, the following rules are included from [RFC3986]: URI and URI-Reference, and from [RFC4288]: type-name.

3. Links

In the context of this specification, a link is comprised of:
A link can be viewed as a statement of the form "(subject) has a (relation type) at (object)", where for an outbound link the subject is the context of use and the object is the resource identified by the target URI, and for an inbound link the subject is the resource identified by the target URI and the object is the context of use.

This specification does not define a general syntax for expressing links, nor the specific context for a given link; it is expected that applications of link relations will specify both aspects. One such application is communication of links through HTTP headers, specified in Section 5.

Such applications may further constrain or extend links (e.g., associating a media type hint, only allowing links in one direction).

4. Link Relation Types

A link relation type identifies the semantics of a link. For example, an outbound link with the relation type "copyright" indicates that the resource identified is a statement of the copyright terms applying to the current context of the link.

Relation types are not to be confused with media types [RFC4288]; they do not identify the format of the representation that results when the link is dereferenced. Rather, they only describe how the current context is related to another resource.

As such, relation types are not format-specific, and MUST NOT specify a particular format or media type that they are to be used with. Likewise, a relation type SHOULD NOT specify what its context of its use is.

Relation types are URIs. Although specific applications of links may specify the use of URI-References, they must also indicate how to resolve them to absolute URIs.

Although anyone may mint a URI to be used as a relation type, it is expected that a few standard types will predominate. To facilitate this, Section 6.2 establishes an IANA registry of relation types that share a common base URI.
5. The Link Header Field

The Link entity-header field provides a means for conveying one or more links in HTTP headers. It is semantically equivalent to the <LINK> element in HTML, as well as the atom:link feed-level element in Atom [RFC4287].

```
Link = "Link" ":" #link-value
link-value = "<" URI-Reference ">" *( ";" link-param ) )
link-param = ( ( "rel" "=" relation-type )
              | ( "rev" "=" relation-type )
              | ( "type" "=" type-name )
              | ( "title" "=" quoted-string )
              | ( link-extension ) )
link-extension = token [ "=" ( token | quoted-string ) ]
relation-type = URI-Reference |
               "<" URI-Reference *( SP URI-Reference) ">
```

For example:

```
Link: <http://example.com/TheBook/chapter2>; rel="previous";
     title="previous chapter"
```

indicates that chapter2 is previous to this resource in a logical navigation path.

Each link-value conveys one target URI inside angle brackets ("<>"). If it is relative, it MUST be resolved as per [RFC3986]. Note that because it is conveyed in a header, base URIs from content are not applied to it.

The context of links conveyed in the Link header field is the representation that the header is part of.

Each link-value MUST have at least one "rel" or "rev" parameter whose value indicates the relation type. If the "rel" parameter is used, it indicates that the link’s direction for that relation type is outbound; if the "rev" parameter is used, the given relation type’s direction is inbound.

If the relation-type is a relative URI, its base URI MUST be considered to be "http://www.iana.org/assignments/relation/", and the corresponding value MUST be present in the link relation registry.

Relation-types that include a semicolon (";") or comma (",")) MUST be quoted.

The title parameter MAY be used to label the destination of a link.
such that it can be used as identification within a human-readable menu.

Note that link-values may contain multiple relations; for example

```
Link: <http://example.org/>; rel="index start";
     rel="http://example.net/relation/other";
     rev=copyright
```

Here, the link "http://example.org/" has outbound links of the types "http://www.iana.org/assignments/relation/index",
"http://www.iana.org/assignments/relation/start",
and
"http://example.net/relation/other", as well as an inbound link of type "http://www.iana.org/assignments/relation/copyright".

6. IANA Considerations

6.1. Link Header Registration

This specification updates the Message Header Registry entry for "Link" in HTTP [RFC3864] to refer to this document.

Header field: Link
Applicable protocol: http
Status: standard
Author/change controller:
  IETF (iesg@ietf.org)
  Internet Engineering Task Force
Specification document(s):
  [ this document ]

6.2. Link Relation Type Registry

This specification establishes the Link Relation Type Registry, located at <http://www.iana.org/assignments/relation/>, and updates Atom [RFC4287] to refer to it in place of the "Registry of Link Relations".

The semantics of relation types is described in Section 4. This registry is intended to contain widely-used, standard relation types so that they may be used in "short form" (i.e., as a relative URI) in applications that allow this.

Registered relation types have an implicit base URI of <http://www.iana.org/assignments/relation/>, and their name SHOULD be limited to the sgml-name rule for simplicity and backwards-compatibility.
Names that differ only in case (e.g., "Foo" and "foo") MUST NOT be registered.

New relation types can be registered by IETF Review, as outlined in [RFC5226]. Specifications of new values should use the following template:

- Relation Name:
- Description:
- Reference:

The Link Relation Type registry's initial contents are:

- Relation Name: alternate
  - Description: Designates a substitute for the link's context.
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: appendix
  - Description: Refers to an appendix.
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: bookmark
  - Description: Refers to a bookmark or entry point.
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: chapter
  - Description: Refers to a chapter in a collection of resources.
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: contents
  - Description: Refers to a table of contents.
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: copyright
  - Description: Refers to a copyright statement that applies to the link's context.
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: current
  - Description: Refers to a resource containing the most recent item(s) in a collection of resources.
  - Reference: [RFC5005]

- Relation Name: edit
- Description: Refers to a resource that can be used to edit the link’s context.
  - Reference: [RFC5023]

- Relation Name: edit-media
  - Description: Refers to a resource that can be used to edit media associated with the link’s context.
  - Reference: [RFC5023]

- Relation Name: enclosure
  - Description: Identifies a related resource that is potentially large and might require special handling.
  - Reference: [RFC4287]

- Relation Name: first
  - Description: A URI that refers to the furthest preceding resource in a series of resources.
  - Reference: <http://www.iana.org/assignments/link-relations/first>

- Relation Name: glossary
  - Description: Refers to a glossary of terms.
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: help
  - Description: Refers to a resource offering help (more information, links to other sources information, etc.)
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: index
  - Description: Refers to an index.
  - Reference: [W3C.REC-html401-19991224]

- Relation Name: last
  - Description: A URI that refers to the furthest following resource in a series of resources.
  - Reference: <http://www.iana.org/assignments/link-relations/last>

- Relation Name: license
  - Description: Refers to a license associated with the link’s context.
  - Reference: [RFC4946]

- Relation Name: next
  - Description: Refers to the next resource in a ordered series of resources.
  - Reference: [W3C.REC-html401-19991224]
o Relation Name: next-archive
  o Description: Refers to the immediately following archive resource.
  o Reference: [RFC5005]

o Relation Name: payment
  o Description: indicates a resource where payment is accepted.
  o Reference: <http://www.iana.org/assignments/link-relations/payment>

o Relation Name: prev
  o Description: Refers to the previous resource in an ordered series of resources. Synonym for "previous".
  o Reference: [W3C.REC-html401-19991224]

o Relation Name: previous
  o Description: Refers to the previous resource in an ordered series of resources. Synonym for "prev".
  o Reference: [W3C.REC-html401-19991224]

o Relation Name: prev-archive
  o Description: Refers to the immediately preceding archive resource.
  o Reference: [RFC5005]

o Relation Name: related
  o Description: Identifies a related resource.
  o Reference: [RFC4287]

o Relation Name: replies
  o Description: Identifies a resource that is a reply to the context of the link.
  o Reference: [RFC4685]

o Relation Name: section
  o Description: Refers to a section in a collection of resources.
  o Reference: [W3C.REC-html401-19991224]

o Relation Name: self
  o Description: Conveys an identifier for the link’s context.
  o Reference: [RFC4287]

o Relation Name: start
  o Description: Refers to the first resource in a collection of resources.
  o Reference: [W3C.REC-html401-19991224]

o Relation Name: stylesheet
7. Security Considerations

The content of the Link header-field is not secure, private or integrity-guaranteed, and due caution should be exercised when using it.

Applications that take advantage of typed links should consider the attack vectors opened by automatically following, trusting, or otherwise using links gathered from HTTP headers.

8. References

8.1. Normative References


[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an

8.2. Informative References


Appendix A. Notes on Using the Link Header with HTML

HTML motivated the original syntax of the Link header, and many of the design decisions in this document are driven by a desire to stay compatible with these uses.

In HTML4, the link element can be mapped to links as specified here by using the "href" attribute for the target URI, and "rel" and rev" to convey both the relation type and its direction, as in the Link header. The context of the link is generally the entire HTML document.

All of the link relations defined by HTML4 have been included in the link relation registry, so they can be used without modification. However, extension link relations work differently in HTML4 and the Link header; the former uses a document-wide "profile" URI to scope the relations, while the latter allows the use of full URIs on individual relations.
Therefore, when using the profile mechanism in HTML4, it is necessary to map the profiled link relations to URIs when expressed in Link headers. For example, in HTML:

```html
<html>
<head profile="http://example.com/profile1/">
  <link rel="foo" href="/foo">
</head>
[...]
```

could be represented as a header like this;

```
Link: </foo>; rel="http://example.com/profile1/foo"
```

Profile authors should note this when creating profile URIs; it may be desirable to use URIs that end in a delimiter (e.g., "/" or "#"), to make extracting the specific relation in use easier.

HTML defines link relation values as case-insensitive, while the Link header's syntax does not. Therefore, it is important to case-normalise relation values in HTML before comparing or converting them to Link headers.

HTML also defines several attributes on links that are not explicitly defined by the Link header. Although most of these are believed to be defunct, they can be used as link-extensions.

**Appendix B. Notes on Using the Link Header with Atom**

Atom conveys links in the atom:link element, with the "href" attribute indicating the target URI and the "rel" attribute containing the relation type. The context of the link is either a feed or an entry, depending on where it appears; generally, feed-level links are candidates for transmission as a Link header. Since atom:link only specifies "rel", only outbound links are allowed by non-extended Atom syntax.

When serialising an atom:link into a Link header, it is necessary to convert IRIs (if used) to URIs. Additionally, since the base URI for link relations in Link headers is fixed, extension relation types (i.e., those not in the registry) must be represented as absolute URIs.

Note also that while the Link header allows multiple relations to be associated with a single link, atom:link does not. In this case, a
single link-value may map to several atom:link elements.

As with HTML, atom:link defines some attributes that are not explicitly mirrored in the Link header syntax, but they may also be used as link-extensions.

Appendix C.  Acknowledgements

This specification lifts the idea and definition for the Link header from RFC2068; credit for it belongs entirely to the authors of and contributors to that document. The link relation registrations themselves are sourced from several documents; see the applicable references.

The author would like to thank the many people who commented upon, encouraged and gave feedback to this draft, especially including Frank Ellermann and Julian Reschke.

Appendix D.  Document history

[[ to be removed by the RFC editor before publication as an RFC. ]]

-03

- Inverted focus from Link headers to link relations.
- Specified was a link relation type is.
- Based on discussion, re-added ‘rev’.
- Changed IESG Approval to IETF Consensus for relation registrations (i.e., require a document).
- Updated RFC2434 reference to RFC5226.
- Registered relations SHOULD conform to sgml-name.
- Cautioned against confusing relation types with media types.

-02

- Dropped XLink language.
- Removed ‘made’ example.
- Removed ‘rev’. Can still be used as an extension.
- Added HTML reference to introduction.
- Required relationship values that have a ; or , to be quoted.
- Changed base URI for relation values.
- Noted registry location.
- Added advisory text about HTML profile URIs.
- Disallowed registration of relations that only differ in case.
Clarified language about IRIs in Atom.

Added descriptions for 'first', 'last', and 'payment', referring to current IANA registry entries, as these were sourced from e-mail. Will this cause self-referential implosion?

Explicitly updates RFC4287.

Added 'type' parameter.

Removed unnecessary advice about non-HTML relations in HTML section.

Changed syntax of link-relation to one or more URI; dropped Profile.

Dropped anchor parameter; can still be an extension.

Removed Link-Template header; can be specified by templates spec or elsewhere.

Straw-man for link relation registry.

Initial draft; normative text lifted from RFC2068.

Author’s Address

Mark Nottingham

Email: mnot@mnot.net
URI: http://www.mnot.net/