Web Linking
draft-nottingham-rfc5988bis-01

Abstract

This specification defines a way to indicate the relationships between resources on the Web ("links") and the type of those relationships ("link relation types").

It also defines the use of such links in HTTP headers with the Link header field.

Note to Readers

This is a work-in-progress to revise RFC5988.

The issues list can be found at https://github.com/mnot/I-D/labels/rfc5988bis.

The most recent (often, unpublished) draft is at https://mnot.github.io/I-D/rfc5988bis/.

Recent changes are listed at https://github.com/mnot/I-D/commits/gh-pages/rfc5988bis.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on November 4, 2016.
Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction ........................................... 3
2. Notational Conventions .................................. 3
3. Links ................................................... 3
4. Link Relation Types .................................... 5
   4.1. Registered Relation Types .......................... 5
   4.1.1. Registering Link Relation Types ............... 6
   4.2. Extension Relation Types ........................... 7
5. Link Serialisation in the Link HTTP Header Field .... 7
   5.1. Link Target ....................................... 8
   5.2. Link Context ...................................... 8
   5.3. Relation Type ..................................... 9
   5.4. Target Attributes .................................. 9
   5.5. Examples .......................................... 10
6. IANA Considerations ..................................... 11
   6.1. Link HTTP Header Field Registration ............. 11
   6.2. Link Relation Type Registry ....................... 11
7. Security Considerations ................................ 12
8. Internationalisation Considerations ........................ 12
9. References .............................................. 13
1. Introduction

This specification defines a way to indicate the relationships between resources on the Web ("links") and the type of those relationships ("link relation types").

HTML [W3C.REC-html5-20141028] and Atom [RFC4287] both have well-defined concepts of linking; this specification generalises this into a framework that encompasses linking in these formats and (potentially) elsewhere.

Furthermore, this specification formalises an HTTP header field for conveying such links, having been originally defined in Section 19.6.2.4 of [RFC2068], but removed from [RFC2616].

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 [RFC7230], including the #rule, and explicitly includes the following rules from it: quoted-string, token, SP (space), LOALPHA, DIGIT.

Additionally, the following rules are included from [RFC3986]: URI and URI-Reference; from [RFC6838]: type-name and subtype-name; from [W3C.CR-css3-mediaqueries-20090915]: media_query_list; from [RFC5646]: Language-Tag; and from [I-D.ietf-httpbis-rfc5987bis], ext-value and parmname.

3. Links

In this specification, a link is a typed connection between two resources, and is comprised of:

- A _link context_,
- a _link relation type_ (Section 4),
o a _link target_, and

o optionally, _target attributes_.

A link can be viewed as a statement of the form "{link context} has a {link relation type} resource at {link target}, which has {target attributes}".

Link contexts and link targets are both IRIs [RFC3987]. However, in the common case, the link context will also be a URI [RFC3986], because many protocols (such as HTTP) do not support dereferencing IRIs. Likewise, the link target will be sometimes be converted to a URI (see [RFC3987], Section 3.1) in places that do not support IRIs (such as the Link header field defined in Section 5).

This specification does not place restrictions on the cardinality of links; there can be multiple links to and from a particular target, and multiple links of the same or different types between a given context and target. Likewise, the relative ordering of links in any particular serialisation, or between serialisations (e.g., the Link header field and in-content links) is not specified or significant in this specification; applications that wish to consider ordering significant can do so.

_Target attributes_ are a set of key/value pairs that describe the link or its target; for example, a media type hint. This specification does not attempt to coordinate their names, cardinality or use, but individual link relations, link serialisations and link applications can do so. This specification does provide common target attributes for use in the Link HTTP header field.

Links are conveyed in _link serialisations_; they are the "bytes on the wire", and can occur in various forms. This specification does not define a general syntax for links, nor does it mandate a specific context for any given link; it is expected that serialisations of links will specify both aspects. One such serialisation is communication of links through HTTP headers, specified in Section 5.

Finally, links are consumed by _link applications_. Generally, an application will define the link relation types it uses, along with the serialisations that they might occur within. For example, the application "Web browsing" looks for the "stylesheet" link relation type in the HTML link serialisation.
4. Link Relation Types

In the simplest case, a link relation type identifies the semantics of a link. For example, a link with the relation type "copyright" indicates that the resource identified by the link target is a statement of the copyright terms applying to the current link context.

Link relation types can also be used to indicate that the target resource has particular attributes, or exhibits particular behaviours; for example, a "service" link implies that the identified resource is part of a defined protocol (in this case, a service description).

Relation types are not to be confused with media types [RFC6838]; 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.

Relation types SHOULD NOT infer any additional semantics based upon the presence or absence of another link relation type, or its own cardinality of occurrence. An exception to this is the combination of the "alternate" and "stylesheet" registered relation types, which has special meaning in HTML for historical reasons.

There are two kinds of relation types: registered and extension.

4.1. Registered Relation Types

Well-defined relation types can be registered as tokens for convenience and/or to promote reuse by other applications, using the procedure in Section 4.1.1.

Registered relation type names MUST conform to the reg-rel-type rule, and MUST be compared character-by-character in a case-insensitive fashion. They SHOULD be appropriate to the specificity of the relation type; i.e., if the semantics are highly specific to a particular application, the name should reflect that, so that more general names are available for less specific use.

Registered relation types MUST NOT constrain the media type of the link context, and MUST NOT constrain the available representation media types of the link target. However, they can specify the behaviours and properties of the target resource (e.g., allowable HTTP methods, request and response media types that must be supported).
Historically, applications have sometimes referred to registered relation types with a URI [RFC3986] (e.g., Appendix B) by prefixing their names with an application-defined base URI. This practice is NOT RECOMMENDED, because the resulting strings will not be considered equivalent to the registered relation types by other processors. Applications that do use such URIs internally MUST NOT use them in link serialisations that do not explicitly accommodate them.

### 4.1.1. Registering Link Relation Types

Relation types are registered on the advice of a Designated Expert (appointed by the IESG or their delegate), with a Specification Required (using terminology from [RFC5226]).

The Expert(s) will establish procedures for requesting registrations, and make them available from the registry page.

Registration requests consist of at least the following information:

- Relation Name:
- Description:
- Reference:

The Expert(s) MAY define additional fields to be collected in the registry.

General requirements for registered relation types are described in Section 4.1.

See the registry for examples of the description field; generally, it SHOULD identify the semantics in terms of the link's context and target.

Registrations MUST reference a freely available, stable specification.

Note that relation types can be registered by third parties, if the Expert(s) determine that an unregistered relation type is widely deployed and not likely to be registered in a timely manner.

Decisions (or lack thereof) made by the Expert(s) can be first appealed to Application Area Directors (contactable using app-ads@tools.ietf.org email address or directly by looking up their email addresses on [http://www.iesg.org/](http://www.iesg.org/) website) and, if the appellant is not satisfied with the response, to the full IESG (using the iesg@iesg.org mailing list).
4.2. Extension Relation Types

Applications that don't wish to register a relation type can use an extension relation type, which is a URI [RFC3986] that uniquely identifies the relation type. Although the URI can point to a resource that contains a definition of the semantics of the relation type, clients SHOULD NOT automatically access that resource to avoid overburdening its server.

The URI used for an extension relation type SHOULD be under the control of the person or party defining it, or be delegated to them. These URIs also SHOULD NOT use the base URI defined by an application for registered relation types (as per Section 4.1).

When extension relation types are compared, they MUST be compared as strings (after converting to URIs if serialised in a different format, such as a XML QNames [W3C.REC-xml-names-20091208]) in a case-insensitive fashion, character-by-character. Because of this, all-lowercase URIs SHOULD be used for extension relations.

Note that while extension relation types are required to be URIs, a serialisation of links can specify that they are expressed in another form, as long as they can be converted to URIs.

5. Link Serialisation in the Link HTTP Header Field

The Link header field provides a means for serialising one or more links into HTTP headers.
5.1. Link Target

Each link-value conveys one target IRI as a URI-Reference (after conversion to one, if necessary; see [RFC3987], Section 3.1) inside angle brackets ("<>"). If the URI-Reference is relative, parsers MUST resolve it as per [RFC3986], Section 5. Note that any base IRI from the message’s content is not applied.

5.2. Link Context

By default, the context of a link conveyed in the Link header field is identity of the representation it is associated with, as defined in [RFC7231], Section 3.1.4.1, serialised as a URI.

When present, the anchor parameter overrides this with another URI, such as a fragment of this resource, or a third resource (i.e., when
the anchor value is an absolute URI). If the anchor parameter's value is a relative URI, parsers MUST resolve it as per [RFC3986], Section 5. Note that any base URI from the body's content is not applied.

Consuming implementations can choose to ignore links with an anchor parameter. For example, the application in use might not allow the link context to be assigned to a different resource. In such cases, the entire link is to be ignored; consuming implementations MUST NOT process the link without applying the anchor.

Note that depending on HTTP status code and response headers, the link context might be "anonymous" (i.e., no link context is available). For instance, this is the case on a 404 response to a GET request.

5.3. Relation Type

The relation type of a link is conveyed in the "rel" parameter's value. The "rel" parameter MUST NOT appear more than once in a given link-value; occurrences after the first MUST be ignored by parsers.

The "rev" parameter has been used in the past to indicate that the semantics of the relationship are in the reverse direction. That is, a link from A to B with REL="X" expresses the same relationship as a link from B to A with REV="X". "rev" is deprecated by this specification because it often confuses authors and readers; in most cases, using a separate relation type is preferable.

Note that extension relation types are REQUIRED to be absolute URIs in Link headers, and MUST be quoted if they contain a semicolon (";") or comma (",") (as these characters are used as delimiters in the header field itself).

5.4. Target Attributes

The "hreflang", "media", "title", "title*", "type", and any link-extension link-params are considered to be target attributes for the link.

The "hreflang" parameter, when present, is a hint indicating what the language of the result of dereferencing the link should be. Note that this is only a hint; for example, it does not override the Content-Language header field of a HTTP response obtained by actually following the link. Multiple "hreflang" parameters on a single link-value indicate that multiple languages are available from the indicated resource.
The "media" parameter, when present, is used to indicate intended destination medium or media for style information (see \[W3C.REC-html5-20141028\], Section 4.2.4). Its value MUST be quoted if it contains a semicolon (";") or comma (","), and there MUST NOT be more than one "media" parameter in a link-value.

The "title" parameter, when present, is used to label the destination of a link such that it can be used as a human-readable identifier (e.g., a menu entry) in the language indicated by the Content-Language header field (if present). The "title" parameter MUST NOT appear more than once in a given link-value; occurrences after the first MUST be ignored by parsers.

The "title*" parameter can be used to encode this label in a different character set, and/or contain language information as per \[I-D.ietf-httpbis-rfc5987bis\]. The "title*" parameter MUST NOT appear more than once in a given link-value; occurrences after the first MUST be ignored by parsers. If the parameter does not contain language information, its language is indicated by the Content-Language header field (when present).

If both the "title" and "title*" parameters appear in a link-value, processors SHOULD use the "title*" parameter's value.

The "type" parameter, when present, is a hint indicating what the media type of the result of dereferencing the link should be. Note that this is only a hint; for example, it does not override the Content-Type header field of a HTTP response obtained by actually following the link. The "type" parameter MUST NOT appear more than once in a given link-value; occurrences after the first MUST be ignored by parsers.

### 5.5. Examples

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.

Similarly,

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

indicates that the root resource ("/" ) is related to this resource with the extension relation type "http://example.net/foo".
The example below shows an instance of the Link header field encoding multiple links, and also the use of RFC 5987 encoding to encode both non-ASCII characters and language information.

Link: </TheBook/chapter2>; rel="previous"; title*=UTF-8'de'letztes%20Kapitel,
</TheBook/chapter4>; rel="next"; title*=UTF-8'de'n%c3%a4chstes%20Kapitel

Here, both links have titles encoded in UTF-8, use the German language ("de"), and the second link contains the Unicode code point U+00E4 ("LATIN SMALL LETTER A WITH DIAERESIS").

Note that link-values can convey multiple links between the same link target and link context; for example:

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

Here, the link to "http://example.org/" has the registered relation type "start" and the extension relation type "http://example.net/relation/other".

6. IANA Considerations

In addition to the actions below, IANA should terminate the Link Relation Application Data Registry, as it has not been used, and future use is not anticipated.

6.1. Link HTTP Header Field 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):
    [RFC&rfc.number;]

6.2. Link Relation Type Registry

This specification updates the registration procedures for the Link Relation Type registry; see Section 4.1.1. The Expert(s) and IANA will interact as outlined below.
IANA will direct any incoming requests regarding the registry to the processes established by the Expert(s); typically, this will mean referring them to the registry HTML page.

The Expert(s) will provide registry data to IANA in an agreed form (e.g., a specific XML format). IANA will publish: * The raw registry data * The registry data, transformed into HTML * The registry data in any alternative formats provided by the Expert(s)

Each published document will be at a URL agreed to by IANA and the Expert(s), and IANA will set HTTP response headers on them as (reasonably) requested by the Expert(s).

Additionally, the HTML generated by IANA will: * Take directions from the Expert(s) as to the content of the HTML page's introductory text and markup * Include a stable HTML fragment identifier for each registered link relation

All registry data documents MUST include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions (<http://trustee.ietf.org/license-info>).

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. Use of Transport Layer Security (TLS) with HTTP ([RFC2818] and [RFC2817]) is currently the only end-to-end way to provide such protection.

Link applications ought to consider the attack vectors opened by automatically following, trusting, or otherwise using links gathered from HTTP headers. In particular, Link header fields that use the "anchor" parameter to associate a link's context with another resource should be treated with due caution.

The Link header field makes extensive use of IRIs and URIs. See [RFC3987] for security considerations relating to IRIs. See [RFC3986] for security considerations relating to URIs. See [RFC7230] for security considerations relating to HTTP headers.

8. Internationalisation Considerations

Link targets may need to be converted to URIs in order to express them in serialisations that do not support IRIs. This includes the Link HTTP header field.
Similarly, the anchor parameter of the Link header field does not support IRIs, and therefore IRIs must be converted to URIs before inclusion there.

Relation types are defined as URIs, not IRIs, to aid in their comparison. It is not expected that they will be displayed to end users.

Note that registered Relation Names are required to be lower-case ASCII letters.

9. References

9.1. Normative References

[I-D.ietf-httpbis-rfc5987bis]
Reschke, J., "Indicating Character Encoding and Language for HTTP Header Field Parameters", draft-ietf-httpbis-rfc5987bis-01 (work in progress), March 2016.


9.2. Informative References


[W3C.REC-html-rdfa-20150317]

[W3C.REC-html5-20141028]

[W3C.REC-xml-names-20091208]
Appendix A. Link Serialisation in HTML

HTML [W3C.REC-html5-20141028] motivated the original syntax of the Link header field, and many of the design decisions in this document are driven by a desire to stay compatible with it.

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

All of the link relation types defined by HTML have been included in the Link Relation Type registry, so they can be used without modification. However, there are several potential ways to serialise extension relation types into HTML, including:

- As absolute URIs,
- using the RDFa [W3C.REC-html-rdfa-20150317] convention of mapping token prefixes to URIs (in a manner similar to XML name spaces).

Individual applications of linking will therefore need to define how their extension links should be serialised into HTML.

Surveys of existing HTML content have shown that unregistered link relation types that are not URIs are (perhaps inevitably) common. Consuming HTML implementations ought not consider such unregistered short links to be errors, but rather relation types with a local scope (i.e., their meaning is specific and perhaps private to that document).

HTML also defines several attributes on links that can be see as target attributes, including "media", "hreflang", "type" and "sizes".

Finally, the HTML specification gives a special meaning when the "alternate" and "stylesheet" relation types coincide in the same link. Such links ought to be serialised in the Link header field using a single list of relation-types (e.g., rel="alternate stylesheet") to preserve this relationship.

Appendix B. Link Serialisation in Atom

Atom [RFC4287] is a link serialisation that conveys links in the atom:link element, with the "href" attribute indicating the link target and the "rel" attribute containing the relation type. The context of the link is either a feed locator or an entry ID, depending on where it appears; generally, feed-level links are obvious candidates for transmission as a Link header field.
When serialising an atom:link into a Link header field, it is necessary to convert link targets (if used) to URIs.

Atom defines extension relation types in terms of IRIs. This specification re-defines them as URIs, to simplify and reduce errors in their comparison.

Atom allows registered link relation types to be serialised as absolute URIs using a prefix, "http://www.iana.org/assignments/relation/". This prefix is specific to the Atom serialisation.

Furthermore, link relation types are always compared in a case-sensitive fashion; therefore, registered link relation types SHOULD be converted to their registered form (usually, lowercase) when serialised in an Atom document.

Note also that while the Link header field allows multiple relations to be serialised in 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 field syntax, but they can also be used as link-extensions to maintain fidelity.

Appendix C. Changes from RFC5988

This specification has the following differences from its predecessor, RFC5988:

- The initial relation type registrations were removed, since they've already been registered by 5988.
- The introduction has been shortened.
- The Link Relation Application Data Registry has been removed.
- Incorporated errata.
- Updated references.
- Link cardinality was clarified.
- Terminology was changed from "target IRI" and "context IRI" to "link target" and "link context" respectively.
- Made assigning a URI to registered relation types application-specific.
- Removed misleading statement that the link header field is semantically equivalent to HTML and Atom links.
- More carefully defined how the Experts and IANA should interact.
- More carefully defined and used "link serialisations" and "link applications."
- Clarified the cardinality of target attributes (generically and for "type").
- Corrected the default link context for the Link header field, to be dependent upon the identity of the representation (as per RFC7231).

Author's Address

Mark Nottingham

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