

Network Working Group	E. Hammer-Lahav	
Internet-Draft	Yahoo!	
Intended status: Informational	November 10, 2009	
Expires: May 14, 2010		

[TOC](#)

host-meta: Web Host Metadata draft-hammer-hostmeta-04

Abstract

This memo describes a method for locating host metadata for Web-based protocols.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and 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 May 14, 2010.

Copyright Notice

Copyright (c) 2009 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 BSD License.

Table of Contents

- [1. Introduction](#)
 - [1.1. Example](#)
 - [1.2. Namespace and Version](#)
 - [1.3. Notational Conventions](#)
- [2. Metadata Scope](#)
- [3. The host-meta Document Format](#)
 - [3.1. The 'hm:Host' Element](#)
 - [3.2. The 'Link' Element](#)
 - [3.2.1. Template Syntax](#)
- [4. Obtaining host-meta Documents](#)
- [5. Security Considerations](#)
- [6. IANA Considerations](#)
 - [6.1. The host-meta Well-Known URI](#)
- [Appendix A. Acknowledgments](#)
- [Appendix B. Document History](#)
- [7. Normative References](#)
- [§ Author's Address](#)

1. Introduction

[TOC](#)

Web-based protocols often require the discovery of host policy or metadata, where host is not a single resource but the entity controlling the collection of resources identified by URIs with a common host as defined by [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#). While these protocols have a wide range of metadata needs, they often define metadata that is concise, has simple syntax requirements, and can benefit from storing its metadata in a common location used by other related protocols.

Because there is no URI or a resource available to describe a host, many of the methods used for associating per-resource metadata (such as HTTP headers) are not available. This often leads to the overloading of the root HTTP resource (e.g. 'http://example.com/') with host metadata that is not specific to the root resource (e.g. a home page or web application), and which often has nothing to do it.

This memo registers the "well-known" URI suffix 'host-meta' in the Well-Known URI Registry established by [\[I-D.nottingham-site-meta\] \(Nottingham, M. and E. Hammer-Lahav, "Defining Well-Known URIs," September 2009.\)](#), and specifies a simple, general-purpose metadata document for hosts, to be used by multiple Web-based protocols.

Please discuss this draft on the apps-discuss@ietf.org mailing list.

1.1. Example

[TOC](#)

A simple host-meta document for the 'example.com' and 'www.example.com' hosts with a link providing host-wide copyright information and a link template providing a URI for obtaining resource-specific metadata for each resource within the host-meta document scope:

```
<?xml version='1.0' encoding='UTF-8'?>
<XRD xmlns='http://docs.oasis-open.org/ns/xri/xrd-1.0'
      xmlns:hm='http://host-meta.net/ns/1.0'>

  <hm:Host>example.com</hm:Host>
  <hm:Host>www.example.com</hm:Host>

  <Link>
    <Title xml:lang='en-us'>Site License Policy</Title>
    <Rel>license</Rel>
    <URI>http://example.com/license</URI>
  </Link>
  <Link>
    <Title xml:lang='en-us'>Resource Descriptor</Title>
    <Rel>describedby</Rel>
    <URITemplate>http://meta.example.com?uri={uri}</URITemplate>
  </Link>
</XRD>
```

1.2. Namespace and Version

[TOC](#)

The host-meta document uses the XRD 1.0 XML namespace URI [\[W3C.REC-xml-names-19990114\]](#) (Hollander, D., Layman, A., and T. Bray, "Namespaces in XML," January 1999.):

`http://docs.oasis-open.org/ns/xri/xrd-1.0`

The XML namespace URI for the host-meta specific extension elements defined in this specification is:

`http://host-meta.net/ns/1.0`

[TOC](#)

1.3. 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 [\[RFC2119\] \(Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," March 1997.\)](#).

This specification uses the namespace prefix "hm:" for the extension Namespace URI identified in [Section 1.2 \(Namespace and Version\)](#). Note that the choice of namespace prefix is arbitrary and not semantically significant. Element names without a namespace prefix belong to the XRD 1.0 XML namespace identified in [Section 1.2 \(Namespace and Version\)](#). This document uses the Augmented Backus-Naur Form (ABNF) notation of [\[RFC5234\] \(Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF," January 2008.\)](#). Additionally, the following rules are included from [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#): reserved, unreserved, and host.

2. Metadata Scope

[TOC](#)

Each host-meta document describes one or more hosts. The scope MUST be expressed explicitly within the document using the ['hm:Host' elements \(The 'hm:Host' Element\)](#). The host-meta scope does not apply to any other hostname (or sub-domain) not explicitly declared. For example, 'example.net', 'example.com', and 'www.example.com' all have different and non-overlapping scopes.

3. The host-meta Document Format

[TOC](#)

The host-meta document uses the XRD 1.0 document format as defined by [\[OASIS.XRD-1.0\] \(Hammer-Lahav, E. and W. Norris, "Extensible Resource Descriptor \(XRD\) Version 1.0 \(work in progress\)," .\)](#), which provides a simple and extensible XML-based schema for describing resources. This memo defines additional elements and processing rules needed to describe hosts. XRD elements not explicitly mentioned in this memo are permitted and used as defined in [\[OASIS.XRD-1.0\] \(Hammer-Lahav, E. and W. Norris, "Extensible Resource Descriptor \(XRD\) Version 1.0 \(work in progress\)," .\)](#).

The host-meta document root MUST be an 'XRD' element. The document SHOULD NOT include a 'Subject' element, as at this time no URI is available to identify hosts. The use of the 'Alias' element in host-meta is undefined and NOT RECOMMENDED.

This memo defines the ['hm:Host' element \(The 'hm:Host' Element\)](#) for declaring document scope. The subject (or "context resource" as defined by [\[I-D.nottingham-http-link-header\] \(Nottingham, M., "Web Linking," July 2009.\)](#)) of the XRD 'Type' and 'Link' elements are the hosts included in the document scope, with the exception of 'Link' elements with a 'URITemplate' child element for which the subject are individual resources included in the document scope as defined in [Section 3.2 \(The 'Link' Element\)](#).

3.1. The 'hm:Host' Element

[TOC](#)

The 'hm:Host' element is used to declare the scope of the host-meta document and is defined as a child element of the root 'XRD' element. The parent 'XRD' element MUST include one but MAY include more 'hm:Host' elements (order does not matter). If a host-meta document includes more than one 'hm:Host' element, it does not signify any relationship between the individual hosts other than sharing the same metadata.

The element value syntax ABNF:

```
Host-Element-Value = host
```

3.2. The 'Link' Element

[TOC](#)

The XRD 'Link' element, when used with the 'URI' child element, conveys a link relation between the host (or hosts) described by the document and a common target URI.

For example, the following link declares a common author for the entire scope:

```
<Link>
  <Rel>author</Rel>
  <URI>http://example.com/author</URI>
</Link>
```

In addition, a 'Link' element with a 'URITemplate' child element conveys relations whose context are individual resources within the host-meta document scope, and whose target is constructed by applying the context URI to a template.

For example, a blog with multiple authors can provide information about each article's author by providing an endpoint with a parameter set to the URI of each article. Each article has a unique author, but all share the same pattern of where that information is located:

```
<Link>
  <Rel>author</Rel>
  <URITemplate>http://example.com?author={uri}</URITemplate>
</Link>
```

3.2.1. Template Syntax

[TOC](#)

This memo defines a simple template syntax for URI transformation. A template is a string containing brace-enclosed ("{}") variable names marking the parts of the string that are to be substituted by the corresponding variable values.

Before substituting template variables, any value character other than unreserved (as defined by [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#)) MUST be percent-encoded per [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#).

This memo defines a single variable, 'uri', as the entire context URI. Protocols MAY define additional relation-specific variables and syntax rules, but SHOULD only do so for protocol-specific relation types, and MUST NOT change the meaning of the 'uri' variable. If a client is unable to successfully process a template (due to unknown variable names, or unknown or incompatible syntax) the parent 'Link' element SHOULD be ignored.

The template syntax ABNF:

```
URI-Template = *( uri-char | variable )
variable     = "{" var-name "}"
uri-char     = ( reserved | unreserved )
var-name     = "uri" | ( 1*var-char )
var-char     = ALPHA / DIGIT / "." / "_"
```

For example:

```
Input:    http://example.com/r?f=1
Template: http://example.org?q={uri}
Output:   http://example.org?q=http%3A%2F%2Fexample.com%2Fr%3Ff%3D1
```

[TOC](#)

4. Obtaining host-meta Documents

The host-meta document is obtained by making an HTTP [\[RFC2616\]](#) (Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1," June 1999.) GET request to the host's port 80, or an HTTPS [\[RFC2818\]](#) (Rescorla, E., "HTTP Over TLS," May 2000.) GET request to the host's port 443 for the `'/.well-known/host-meta'` path.

Servers MUST support at least one but SHOULD support both ports unless restricted by other considerations. If both ports are supported, they MUST serve the same document. Clients MAY attempt to obtain the host-meta document from either port, and SHOULD attempt the other port if the first fails, unless restricted by other considerations.

For example, the following request is used to obtain the host-meta document for the `'example.com'` host:

```
GET /.well-known/host-meta HTTP/1.1
Host: example.com
```

If the server response indicates that the host-meta resource is located elsewhere (a 301, 302, or 307 response status codes), the client SHOULD try to obtain the resource from the location provided in the response. This means that the host-meta document for one host MAY be retrieved from a different host. Likewise, if the resource is not available or exists (the 404 or 410 response status codes), the client SHOULD infer that metadata is not available via this mechanism.

If a representation is successfully obtained, but is not in the format described above, clients SHOULD infer that the path is being used for other purposes, and not process it as a host-meta document. To aid in this process, authorities using this mechanism SHOULD correctly label host-meta responses with the `"application/xrd+xml"` internet media type. The scope declared within the host-meta document MUST match the desired host.

5. Security Considerations

[TOC](#)

The metadata returned by the host-meta resource is presumed to be under the control of the appropriate authority and representative of all the resources described by it. If this resource is compromised or otherwise under the control of another party, it may represent a risk to the security of the server and data served by it, depending on what protocols use it.

The host-meta scope is explicitly declared by the `'hm:Host'` elements listed in the document. Clients SHOULD evaluate the authority of a host-meta document obtained from one host to describe any other host. Protocols that change the scope from the one declared in the document without careful consideration can incur security risks.

Protocols using host-meta templates SHOULD evaluate the construction of their templates as well as any protocol-specific variables or syntax to ensure that the templates cannot be abused by an attacker. For example, a client can be tricked into following a malicious link due to a poorly constructed template which produces unexpected results when its variable values contain unexpected characters.

Protocols MAY restrict document retrieval to HTTPS based on their security needs. Protocols utilizing host-meta documents obtained via other methods not described in this memo SHOULD consider the security and authority risks associated with such methods.

6. IANA Considerations

[TOC](#)

6.1. The host-meta Well-Known URI

[TOC](#)

This memo registers the 'host-meta' well-known URI in the Well-Known URI Registry as defined by [\[I-D.nottingham-site-meta\] \(Nottingham, M. and E. Hammer-Lahav, "Defining Well-Known URIs," September 2009.\)](#).

URI suffix: host-meta

Change controller: IETF

Specification document(s): [[this document]]

Related information: None

Appendix A. Acknowledgments

[TOC](#)

This memo was initially based on [\[I-D.nottingham-site-meta\] \(Nottingham, M. and E. Hammer-Lahav, "Defining Well-Known URIs," September 2009.\)](#).

The author would like to acknowledge the contributions of everyone who provided feedback and use cases for this memo; in particular, Dirk Balfanz, DeWitt Clinton, Blaine Cook, Breno de Medeiros, Brad Fitzpatrick, James Manger, Will Norris, Mark Nottingham, John Panzer, and Drummond Reed.

Appendix B. Document History

[TOC](#)

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

-04

*Corrected the <hm:Host> example.

-03

*Changed scope to an entire host (per RFC 3986).

*Simplified template syntax to always percent-encode values and vocabulary to a single 'uri' variable.

*Changed document retrieval to always use HTTP(S).

*Added security consideration about the use of templates.

*Explicitly defined the root element to be 'XRD'.

-02

*Changed Scope element syntax from attributes to URI-like string value.

-01

*Editorial rewrite.

*Redefined scope as a scheme-authority pair.

*Added document structure section.

-00

*Initial draft.

7. Normative References

[TOC](#)

[I-D.nottingham-http-link-header]	Nottingham, M., " Web Linking ," draft-nottingham-http-link-header-06 (work in progress), July 2009 (TXT).
[I-D.nottingham-site-meta]	Nottingham, M. and E. Hammer-Lahav, " Defining Well-Known URIs ," draft-nottingham-site-meta-03 (work in progress), September 2009 (TXT).
[OASIS.XRD-1.0]	

	Hammer-Lahav, E. and W. Norris, " Extensible Resource Descriptor (XRD) Version 1.0 (work in progress) " (HTML).
[RFC2119]	Bradner, S. , " Key words for use in RFCs to Indicate Requirement Levels ," BCP 14, RFC 2119, March 1997 (TXT , HTML , XML).
[RFC2616]	Fielding, R. , Gettys, J. , Mogul, J. , Frystyk, H. , Masinter, L. , Leach, P. , and T. Berners-Lee , " Hypertext Transfer Protocol -- HTTP/1.1 ," RFC 2616, June 1999 (TXT , PS , PDF , HTML , XML).
[RFC2818]	Rescorla, E., " HTTP Over TLS ," RFC 2818, May 2000 (TXT).
[RFC3986]	Berners-Lee, T. , Fielding, R. , and L. Masinter , " Uniform Resource Identifier (URI): Generic Syntax ," STD 66, RFC 3986, January 2005 (TXT , HTML , XML).
[RFC5234]	Crocker, D. and P. Overell, " Augmented BNF for Syntax Specifications: ABNF ," STD 68, RFC 5234, January 2008 (TXT).
[W3C.REC-P3P-20020416]	Marchiori, M., " The Platform for Privacy Preferences 1.0 (P3P1.0) Specification ," World Wide Web Consortium Recommendation REC-P3P-20020416, April 2002 (HTML).
[W3C.REC-xml-names-19990114]	Hollander, D., Layman, A., and T. Bray, " Namespaces in XML ," World Wide Web Consortium FirstEdition REC-xml-names-19990114, January 1999 (HTML).

Author's Address

[TOC](#)

	Eran Hammer-Lahav
	Yahoo!
Email:	eran@hueniverse.com
URI:	http://hueniverse.com