

Network Working Group	E. Hammer-Lahav	
Internet-Draft	Yahoo!	
Intended status: Informational	October 23, 2009	
Expires: April 26, 2010		

[TOC](#)

Host-Meta: Web Host Metadata draft-hammer-hostmeta-01

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 April 26, 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 in effect on the date of publication of this document (<http://trustee.ietf.org/license-info>). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

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

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 'host-meta:Scope' Element](#)
 - [3.2. The 'Link' Element](#)
- [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](#)
- [8. 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 collection of resources identified by URIs with a common scheme and authority 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 sharing its metadata store with 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 usage of the root HTTP resource as a placeholder for host metadata that is not specific to the root resource, and 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. The name 'host-meta' was chosen based on the HTTP 'Host' header defined by [\[RFC2616\]](#) (Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1," June 1999.).

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' authorities and 'http' scheme with a link providing scope-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:host-meta='http://host-meta.net/ns/1.0'>

  <host-meta:Scope scheme='http' authority='example.com' />
  <host-meta:Scope scheme='http' authority='www.example.com' />

  <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 "host-meta:" 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 and unreserved.

2. Metadata Scope

[TOC](#)

The scope of each host-meta document is defined by one or more scheme-authority pairs, each applies to the collection of resources identified by URIs with the same scheme and authority as defined by [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#). The scope MUST be expressed explicitly within the document by using the one or more ['host-meta:Scope' elements \(The 'host-meta:Scope' Element\)](#).

The host-meta scope does not apply to a subset, nor does it apply to other combinations of scheme and authority (e.g., using another scheme, port, or a different hostname in the same domain) not explicitly declared. For example, resources with URIs beginning with 'http://example.com/', 'http://example.com:8080/', 'https://example.com/', and 'http://www.example.com/' all have different and non-overlapping scopes.

The scope declared within the host-meta document MUST match the scheme and authority used to [obtain \(Obtaining host-meta Documents\)](#) the document. However, protocols MAY use a different scheme to obtain the document than the one they are interested in, as long as both scheme and authority combinations are declared. For example, a protocol can use HTTP to obtain metadata about a scheme other than 'http' for which a host-meta document cannot be obtained using the method described in [Section 4 \(Obtaining host-meta Documents\)](#).

Protocols MAY place additional requirements on protocol-specific metadata retrieved from multiple host-meta documents. For example,

protocols can require that protocol-specific metadata obtained for the 'example.com:80' authority to be the same in both host-meta documents for the 'http' and 'https' schemes.

Any changes in scope, the retrieval of metadata for one scheme using another, and other requirements should only occur after a careful consideration of their security implications and authoritativeness.

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 format for describing resources. Since the host-meta XRD document does not describe a single resource but a collection of resources, this memo defines additional XRD elements and processing rules. The XRD elements not mentioned in this memo are allowed and used as defined in [\[OASIS.XRD-1.0\] \(Hammer-Lahav, E. and W. Norris, "Extensible Resource Descriptor \(XRD\) Version 1.0 \(work in progress\)," .\)](#).

host-meta document SHOULD NOT include the 'Subject' or 'Alias' XRD elements since these elements require a valid URI to identify the resource being described, which is not available for the host-meta scope. The use of these elements in host-meta is undefined. Instead, host-meta defines the ['host-meta:Scope' element \(The 'host-meta:Scope' Element\)](#) 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 is the entire collection of resources included in the document scope, with the exception of 'Link' elements with a 'URITemplate' child element as defined in [Section 3.2 \(The 'Link' Element\)](#).

3.1. The 'host-meta:Scope' Element

[TOC](#)

The 'host-meta:Scope' 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 'host-meta:Scope' elements (order does not matter). The element includes two REQUIRED attributes: 'scheme' and 'authority' and MUST NOT include a content.

The 'scheme' and 'authority' attributes use the same syntax and semantics defined by [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#), and MUST NOT contain empty values. If the 'authority'

attribute does not include a port number, the port is implicitly set to the default port number for the given scheme, and does not include any other port number.

For example, these two scope declarations are equivalent:

```
<host-meta:Scope scheme='http' authority='example.com' />
<host-meta:Scope scheme='http' authority='example.com:80' />
```

3.2. The 'Link' Element

[TOC](#)

The XRD 'Link' element, when used with the 'URI' child element, conveys a link relation between the collection of resources included in the host-meta document scope 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>
```

However, there are cases in which individual resources within the same scope do not share the same target URI, but follow a common pattern in how the target URI is constructed. 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 appending a suffix (such as ';by') 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>{+uri};by</URITemplate>
</Link>
```

3.2.1. Template Syntax

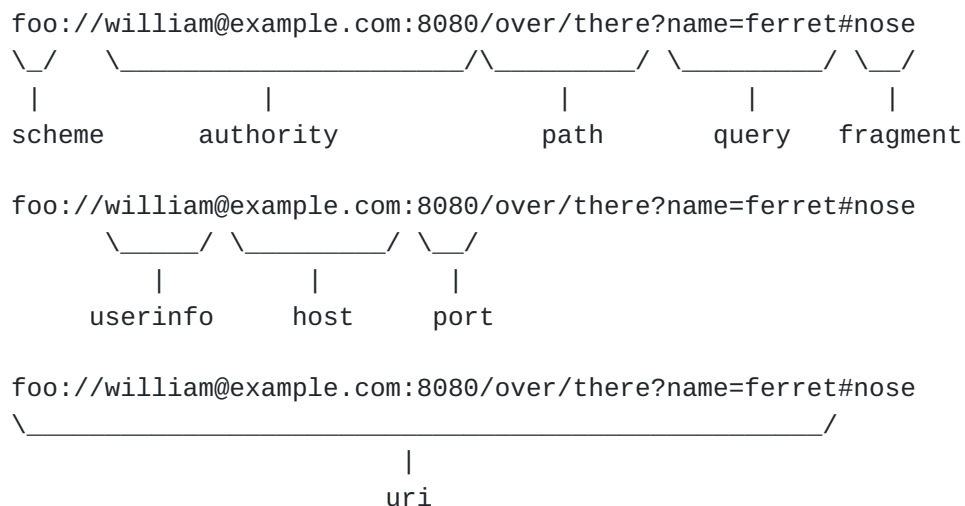
[TOC](#)

host-meta defines a simple template syntax and a vocabulary 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 reserved and 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.). If a variable name is not prefixed by a '+' character, any reserved character SHALL also be percent-encoded.

To construct a URI using a template, the context URI is parsed into its URI components and each component value assigned to a variable name. The variable set is based on the URI vocabulary defined by [\[RFC3986\]](#) (Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax," January 2005.) section 3 and includes: 'scheme', 'authority', 'path', 'query', 'fragment', 'userinfo', 'host', and 'port'. In addition, this memo defines the 'uri' variable as the entire context URI excluding the fragment component and the '#' fragment separator.



Protocols MAY define additional variables and syntax rules, but SHOULD only do so for protocol-specific relation types, and MUST NOT change the meaning of the variables defined above. If a template uses an unknown syntax or contains unknown variable names, 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-var | ( 1*var-char )
uri-var      = "scheme" | "authority" | "path"      |
               "query"   | "fragment"   | "userinfo" |
               "host"    | "port"       | "uri"
var-char     = ALPHA / DIGIT / "." / "_"

```

For example, given the input URI 'http://example.com/r/1?f=xml#top', each of the following templates will produce the associated output URI:

```

{+uri}&test -->
http://example.com/r/1?f=xml&test

http://example.org?q={uri} -->
http://example.org?q=http%3A%2F%2Fexample.com%2Fr%2F1%3Ff%3Dxml

http://meta.{host}:8080{+path}?{+query} -->
http://meta.example.com:8080/r/1?f=xml

```

4. Obtaining host-meta Documents

[TOC](#)

The host-meta document for a given scheme-authority pair is identified by a URI constructed using the desired scheme, authority, and '/.well-known/host-meta' as the URI path. For example, given the 'http' scheme and 'www.example.com:80' authority, the host-meta URI is:

```
http://www.example.com/.well-known/host-meta
```

The host-meta document is obtained by dereferencing the request host-meta URI. The semantics of the protocol used to obtain the host-meta document apply. Therefore, if the server indicates that the host-meta resource is located elsewhere (in HTTP, 3xx response status codes), the client **MUST** try to obtain the resource from the location provided. This means that the host-meta document for one authority **MAY** be retrieved from a different authority. Likewise, if the resource is not available or exists (in HTTP, 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 authority is using this URI 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.

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.

Scoping metadata to a single scheme-authority pair is the default host-meta scope. Protocols that change the scope without careful consideration can incur security risks.

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, 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]]
-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).
[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