Defining Well-Known Uniform Resource Identifiers (URIs)

draft-nottingham-rfc5785bis-04

Abstract

This memo defines a path prefix for "well-known locations", "/.well-known/", in selected Uniform Resource Identifier (URI) schemes.

Note to Readers

_RFC EDITOR: please remove this section before publication_

This draft is a proposed revision of RFC5875.

The issues list for this draft can be found at
https://github.com/mnot/I-D/labels/rfc5785bis [1].

The most recent (often, unpublished) draft is at

Recent changes are listed at https://github.com/mnot/I-D/commits/gh-pages/rfc5785bis [3].

See also the draft’s current status in the IETF datatracker, at

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 https://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."
1. Introduction

Some applications on the Web require the discovery of information about an origin [RFC6454] (sometimes called "site-wide metadata") before making a request. For example, the Robots Exclusion Protocol (http://www.robotstxt.org/ [5]) specifies a way for automated processes to obtain permission to access resources; likewise, the Platform for Privacy Preferences [W3C.REC-P3P-20020416] tells user-agents how to discover privacy policy before interacting with an origin server.

While there are several ways to access per-resource metadata (e.g., HTTP headers, WebDAV’s PROPFIND [RFC4918]), the perceived overhead
(either in terms of client-perceived latency and/or deployment difficulties) associated with them often precludes their use in these scenarios.

When this happens, one solution is designating a "well-known location" for data or services related to the origin overall, so that it can be easily located. However, this approach has the drawback of risking collisions, both with other such designated "well-known locations" and with resources that the origin has created (or wishes to create).

To address this, this memo defines a path prefix in HTTP(S) URIs for these "well-known locations", "/.well-known/". Future specifications that need to define a resource for such metadata can register their use to avoid collisions and minimise impingement upon origins’ URI space.

Well-known URIs can also be used with other URI schemes, but only when those schemes’ definitions explicitly allow it.

1.1. Appropriate Use of Well-Known URIs

As per [RFC7320], "publishing independent standards that mandate particular forms of URI substructure is inappropriate, because that essentially usurps ownership." Well-known URIs are not an escape hatch from the requirements therein; they are a very limited carve-out of the path name space owned by the authority, ceded to standard use for a designated purpose.

That purpose is to facilitate discovery of information about an origin when it isn’t practical to use other mechanisms; for example, when discovering policy that needs to be evaluated before a resource is accessed, or when the information applies to many (or all) of the origin’s resources.

Typically, the resource(s) identified by a well-known URI will make information about the origin (e.g., policy) available directly, or provide references to other URIs that provide it. In general, that information should be applicable to most origins (i.e., Web sites - while acknowledging that some origins might not use a particular well-known location, for various reasons).

In keeping with the Architecture of the World-Wide Web [W3C.REC-webarch-20041215], well-known URIs are not intended for general information retrieval or establishment of large URI namespaces.
Specifically, well-known URIs are not a "protocol registry" for applications and protocols that wish to use HTTP as a substrate. Instead, such applications and protocols are encouraged to use an absolute URI to bootstrap their operation, rather than using a hostname and a well-known URI.

Exceptionally, the registry expert(s) may approve such a registration for documents in the IETF Stream [RFC5741], in consultation with the IESG, provided that the protocol in question cannot be bootstrapped with a URI (e.g., the discovery mechanism can only carry a hostname). However, merely making it easier to locate it is not a sufficient reason. Likewise, future use unsupported by the specification in question is not sufficient reason to register a well-known location.

Well-known locations are also not suited for information on topics other than the origin that they are located upon; for example, creating a well-known resource about a business entity or organisational structure presumes that Internet hosts and organisations share structure, and are likely to have significant deployment issues in environments where this is not true.

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 [RFC2119].

3. Well-Known URIs

A well-known URI is a URI [RFC3986] whose path component begins with the characters "/.well-known/", and whose scheme is "HTTP", "HTTPS", or another scheme that has explicitly been specified to use well-known URIs.

Applications that wish to mint new well-known URIs MUST register them, following the procedures in Section 5.1.

For example, if an application registers the name ‘example’, the corresponding well-known URI on ‘http://www.example.com/’ would be ‘http://www.example.com/.well-known/example’.

Registered names MUST conform to the segment-nz production in [RFC3986]. This means they cannot contain the "/" character.

Note that this specification defines neither how to determine the authority to use for a particular context, nor the scope of the metadata discovered by dereferencing the well-known URI; both should be defined by the application itself.
Typically, a registration will reference a specification that defines the format and associated media type to be obtained by dereferencing the well-known URI.

It MAY also contain additional information, such as the syntax of additional path components, query strings and/or fragment identifiers to be appended to the well-known URI, or protocol-specific details (e.g., HTTP [RFC7231] method handling).

Note that this specification does not define a format or media-type for the resource located at "/.well-known/" and clients should not expect a resource to exist at that location.

Well-known URIs are only valid when rooted in the top of the path’s hierarchy; they MUST NOT be used in other parts of the path. For example, "/.well-known/example" is a valid use, but "/foo/.well-known/example" is not.

4. Security Considerations

This memo does not specify the scope of applicability of metadata or policy obtained from a well-known URI, and does not specify how to discover a well-known URI for a particular application. Individual applications using this mechanism must define both aspects.

Applications minting new well-known URIs, as well as administrators deploying them, will need to consider several security-related issues, including (but not limited to) exposure of sensitive data, denial-of-service attacks (in addition to normal load issues), server and client authentication, vulnerability to DNS rebinding attacks, and attacks where limited access to a server grants the ability to affect how well-known URIs are served.

Security-sensitive applications using well-known locations should consider that some server administrators might be unaware of its existence (especially on operating systems that hide directories whose names begin with "."). This means that if an attacker has write access to the .well-known directory, they would be able to control its contents, possibly without the administrator realising it.

5. IANA Considerations

5.1. The Well-Known URI Registry

This document specifies procedures for the well-known URI registry, first defined in [RFC5785].
Well-known URIs are registered on the advice of one or more experts
(appointed by the IESG or their delegate), with a Specification
 Required (using terminology from [RFC8126]).

To allow for the allocation of values prior to publication, the
expert(s) may approve registration once they are satisfied that such
a specification will be published.

Registration requests can be sent to the wellknown-uri-
review@ietf.org mailing list for review and comment, with an
appropriate subject (e.g., "Request for well-known URI: example").

5.1.1. Registration Template

URI suffix: The name requested for the well-known URI, relative to
"/well-known/"; e.g., "example".

Change controller: For Standards-Track RFCs, state "IETF". For
others, give the name of the responsible party. Other details
(e.g., postal address, e-mail address, home page URI) may also be
included.

Specification document(s): Reference to the document that specifies
the field, preferably including a URI that can be used to retrieve
a copy of the document. An indication of the relevant sections
may also be included, but is not required.

Related information: Optionally, citations to additional documents
containing further relevant information.

6. References

6.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119,
DOI 10.17487/RFC2119, March 1997,

Resource Identifier (URI): Generic Syntax", STD 66,
RFC 3986, DOI 10.17487/RFC3986, January 2005,

Headers, and Boilerplates", RFC 5741,
DOI 10.17487/RFC5741, December 2009,

Nottingham               Expires August 19, 2018                [Page 6]
6.2. Informative References


6.3. URIs


Appendix A. Frequently Asked Questions

1. Aren’t well-known locations bad for the Web?

   They are, but for various reasons – both technical and social – they are sometimes necessary. This memo defines a "sandbox" for them, to reduce the risks of collision and to minimise the impact upon pre-existing URIs on sites.

2. Why /.well-known?

   It’s short, descriptive, and according to search indices, not widely used.

3. What impact does this have on existing mechanisms, such as P3P and robots.txt?

   None, until they choose to use this mechanism.

4. Why aren’t per-directory well-known locations defined?

   Allowing every URI path segment to have a well-known location (e.g., "/images/.well-known") would increase the risks of colliding with a pre-existing URI on a site, and generally these solutions are found not to scale well, because they’re too “chatty”.

5. I want to use a well-known location to make it easy to configure my protocol that uses HTTP.

   This is not what well-known locations are for; see Section 1.1.

Appendix B. Changes from RFC5785

- Discuss appropriate and inappropriate uses more fully
- Adjust IANA instructions
- Update references
- Various other clarifications