Guidelines and Registration Procedures for New URI/IRI Schemes
draft-ietf-iri-4395bis-irireg-04

Abstract

This document updates the guidelines and recommendations for the definition of Uniform Resource Identifier (URI) schemes, and extends the registry and guidelines to apply when the schemes are used with Internationalized Resource Identifiers (IRIs). It also updates the process and IANA registry for URI/IRI schemes. It obsoletes RFC 4395.

Status of this Memo

Copyright Notice

Table of Contents

TOC

1. Introduction

The Uniform Resource Identifier (URI) protocol element and generic syntax is defined by **[RFC3986]**. Each URI begins with a scheme name, as defined by Section 3.1 of RFC 3986, that refers to a specification for identifiers within that scheme. The URI syntax provides a federated and extensible naming system, where each scheme's specification may further restrict the syntax and define the semantics of identifiers using that scheme. As originally defined, URIs only allowed a limited repertoire of characters chosen from US-ASCII.

An Interationalized Resource Identifier (IRI), as defined by **[RFC3987bis]**, extends the URI syntax to allow characters from a much greater repertoire, to accomodate resource identifiers from the world's languages. The same schemes used in URIs are used in IRIs. The term Resource Identifier (RI) is used as a shorthand for both URIs and IRIs. **[RFC3987]** introduced IRIs by defining a mapping between URIs and IRIs; **[RFC3987bis]** updates that definition, allowing an IRI to be interpreted directly without translating into a URI.

This document obsoletes **[RFC4395]**, which in turn obsoleted **[RFC2717]** and **[RFC2718]**. Recent documents have used the terms "URI"/"IRI" for all resource identifiers, avoiding the term "URL" and reserving the term "URN" explicitly for those URIs/IRIs using the "urn" scheme name (**[RFC2141]**). URN "namespaces" (**[RFC3406]**) are specific to the "urn" scheme and are not covered explicitly by this specification.

This document extends the URI scheme registry to be a registry of URI/IRI schemes (i.e., applicable to both URIs and IRIs). This document also provides updated guidelines for the definition of new schemes, for consideration by those who are defining, registering, or evaluating those definitions, as well as a process and mechanism for registering URI/IRI schemes within the IANA URI scheme registry. There is a single namespace for registered schemes. Within that namespace, there are values that are approved as meeting a set of criteria for permanent URI/IRI schemes. Other scheme names may also be registered provisionally or historically, without necessarily meeting those criteria. The intent of the registry is to:

- provide a central point of discovery for established URI/IRI scheme names, and easy location of their defining documents;
- discourage use of the same scheme name for different purposes;
- help those proposing new scheme names to discern established trends and conventions, and avoid names that might be confused with existing ones;
- encourage registration by setting a low barrier for provisional registrations.

There is no separate, independent registry or registration process for IRIs: the URI Scheme Registry is to be used for both URIs and IRIs. Previously, those who wish to describe resource identifiers that are useful as IRIs were encouraged to define the corresponding URI syntax, and note that the IRI usage follows the rules and transformations defined in **[RFC3987]**. This document changes that advice to encourage explicit definition of the scheme and allowable syntax elements within the larger character repertoire of IRIs, as defined by **[RFC3987bis]**.

A scheme definition cannot override the overall syntax for IRIs. For example, this means that fragment identifiers (#) cannot be re-used outside the generic syntax restrictions, and in particular scheme-specific syntax cannot override the fragment identifier syntax because it is generic.

2. Conformance Guidelines

Within this document, the key words MUST, MAY, SHOULD, REQUIRED, RECOMMENDED, and so forth are used within the general meanings established in **[RFC2119]**, within the context that they are requirements on future registration specifications.

3. Guidelines for Permanent URI/IRI Scheme Definitions

This section gives considerations for new URI/IRI schemes. Meeting these guidelines is REQUIRED for permanent scheme registration. Meeting these guidelines is also RECOMMENDED for provisional registration, as described in **Section 4**.

3.1. Demonstratable, New, Long-Lived Utility

The use and deployment of new URI/IRI schemes in the Internet infrastructure may be costly; some parts of URI/IRI processing may be scheme-dependent, and deployed software already processes URIs and IRIs of well-known schemes. Introducing a new scheme may require additional software, not only for client software and user agents but also in additional parts of the network infrastructure (gateways, proxies, caches) **[W3CWebArch]**. URI/IRI schemes constitute a single, global namespace; it is desirable to avoid contention over use of short, mnemonic scheme names. For these reasons, the unbounded registration of new schemes is harmful. New schemes should have utility to the Internet community beyond that available with already registered schemes. The registration document SHOULD discuss the utility of the scheme being registered.

3.2. Syntactic Compatibility

[RFC3986] defines the generic syntax for all URI schemes, along with the syntax of common URI components that are used by many URI schemes to define hierarchical identifiers. **[RFC3987]** and subsequently **[RFC3987bis]** extended this generic syntax to cover IRIs. All URI/IRI scheme specifications MUST define their own syntax such that all strings matching their scheme-specific syntax will also match the <absolute-URI> grammar described in **[RFC3987bis]**.

New schemes SHOULD reuse the common components of **[RFC3987bis]** for the definition of hierarchical naming schemes. However, if there is a strong reason for a scheme not to use the hierarchical syntax, then the new scheme definition SHOULD follow the syntax of previously registered schemes.

Schemes that are not intended for use with relative URIs/IRIs SHOULD avoid use of the forward slash "/" character, which is used for hierarchical delimiters, and the complete path segments "." and ".." (dot-segments).

Avoid improper use of "//". The use of double slashes in the first part of a URI/IRI is not an artistic indicator that what follows is a URI/IRI: Double slashes are used ONLY when the syntax of the <scheme-specific-part> contains a hierarchical structure. In URIs and IRIs from such schemes, the use of double slashes indicates that what follows is the top hierarchical element for a naming authority. (Section 3.2 of RFC 3986 has more details.) Schemes that do not contain a conformant hierarchical structure in their <scheme-specific-part> SHOULD NOT use double slashes following the "<scheme>:" string.

New schemes SHOULD clearly define the role of **[RFC3986]** reserved characters in URIs/IRIs

of the scheme being defined. The syntax of the new scheme should be clear about which of the "reserved" set of characters are used as delimiters within the URIs/IRIs of the new scheme, and when those characters must be escaped, versus when they may be used without escaping.

3.3. Well-Defined

While URIs/IRIs may or may not be defined as locators in practice, a scheme definition itself MUST be clear as to how it is expected to function. Schemes that are not intended to be used as locators SHOULD describe how the resource identified can be determined or accessed by software that obtains a URI/IRI of that scheme.

For schemes that function as locators, it is important that the mechanism of resource location be clearly defined. This might mean different things depending on the nature of the scheme.

In many cases, new schemes are defined as ways to translate between other namespaces or protocols and the general framework of URIs. For example, the "ftp" scheme translates into the FTP protocol, while the "mid" scheme translates into a Message-ID identifier of an email message. For such schemes, the description of the mapping must be complete, and in sufficient detail so that the mapping in both directions is clear: how to map from a URI/IRI into an identifier or set of protocol actions or name in the target namespace, and how legal values in the base namespace, or legal protocol interactions, might be represented in a valid URI or IRI. In particular, the mapping should describe the mechanisms for encoding binary or character strings within valid character sequences in a URI/IRI (See **Section 3.6** for guidelines). If not all legal values or protocol interactions of the base standard can be represented using the scheme, the definition should be clear about which subset are allowed, and why.

3.4. Definition of Operations

As part of the definition of how a URI/IRI identifies a resource, a scheme definition SHOULD define the applicable set of operations that may be performed on a resource using the RI as its identifier. A model for this is HTTP; an HTTP resource can be operated on by GET, POST, PUT, and a number of other operations available through the HTTP protocol. The scheme definition should describe all well-defined operations on the resource identifier, and what they are supposed to do.

Some schemes don't fit into the "information access" paradigm of URIs/IRIs. For example, "telnet" provides location information for initiating a bi-directional data stream to a remote host; the only operation defined is to initiate the connection. In any case, the operations appropriate for a scheme should be documented.

Note: It is perfectly valid to say that "no operation apart from GET is defined for this RI". It is also valid to say that "there's only one operation defined for this RI, and it's not very GET-like". The important point is that what is defined on this scheme is described.

3.5. Context of Use

In general, URIs/IRIs are used within a broad range of protocols and applications. Most commonly, URIs/IRIs are used as references to resources within directories or hypertext documents, as hyperlinks to other resources. In some cases, a scheme is intended for use within a different, specific set of protocols or applications. If so, the scheme definition SHOULD describe the intended use and include references to documentation that define the applications and/or protocols cited.

3.6. Internationalization and Character Encoding

When describing schemes in which (some of) the elements of the URI or IRI are actually representations of human-readable text, care should be taken not to introduce unnecessary variety in the ways in which characters are encoded into octets and then into characters; see **[RFC3987bis]** and Section 2.5 of **[RFC3986]** for guidelines. If URIs/IRIs of a scheme contain any text fields, the scheme definition MUST describe the ways in which characters are encoded and any compatibility issues with IRIs of the scheme.

Specifications for IRIs schemes MUST be described in terms of processing an IRI as a sequence of Unicode codepoints, without reference to the encoding of those code points as a sequence of bytes, using UTF-8 or UTF-16. The scheme specification SHOULD be as restrictive as possible regarding what characters are allowed in the URI/IRI, because some characters can create several different security considerations (see for example **[RFC4690]**).

All percent-encoded variants are automatically included by definition for any character given in an IRI production. This means that if you want to restrict the URI percent-encoded forms in some way, you must restrict the Unicode forms that would lead to them.

3.7. Clear Security Considerations

Definitions of schemes MUST be accompanied by a clear analysis of the security implications for systems that use the scheme; this follows the practice of Security Consideration sections within IANA registrations **[RFC5226]**.

In particular, Section 7 of RFC 3986 **[RFC3986]** describes general security considerations for URIs, while **[RFC3987bis]** gives those for IRIs. The definition of an individual URI/IRI scheme should note which of these apply to the specified scheme.

3.8. Scheme Name Considerations

Section 3.1 of RFC 3986 defines the syntax of a URI scheme name; this sytax remains the same for IRIs. New registered schemes registrations MUST follow this syntax, which only allows a limited repertoire of characters (taken from US-ASCII). Although the syntax for the scheme name in URI/IRIs is case insensitive, the scheme names itself MUST be registered using lowercase letters.

URI/IRI scheme names should be short, but also sufficiently descriptive and distinguished to avoid problems.

Avoid names or other symbols that might cause problems with rights to use the name in IETF specifications and Internet protocols. For example, be careful with trademark and service mark names. (See Section 7.4 of **[RFC3978]**.)

Avoid using names that are either very general purpose or associated in the community with some other application or protocol. Avoid scheme names that are overly general or grandiose in scope (e.g., that allude to their "universal" or "standard" nature.)

Organizations that desire a private name space for URI scheme names are encouraged to use a prefix based on their domain name, expressed in reverse order. For example, a URI scheme name of com-example-info might be registered by the vendor that owns the example.com domain name.

4. Guidelines for Provisional URI/IRI Scheme Registration

Provisional registration can be an intermediate step on the way to permanent registration, e.g., before the scheme specification is finalized. Provisional registration is also appropriate for schemes that are known to be used, but where a definitive specification is not available. There is no time limit for provisional registration.

While the guidelines in **Section 3** are REQUIRED for permanent registration, they are

RECOMMENDED for provisional registration. For a provisional registration, the following are REQUIRED:

- The scheme name meets the syntactic requirements of **Section 3.8** and the encoding requirements of **Section 3.6**.
- There is not already an entry with the same scheme name. (In the unfortunate case that there are multiple, different uses of the same scheme name, the IESG may approve a request to modify an existing entry to note the separate use.)
- Contact information identifying the person supplying the registration is included. Previously unregistered schemes discovered in use may be registered by third parties (even if not on behalf of those who created the scheme). In this case, both the registering party and the scheme creator SHOULD be identified.
- If no permanent, citable specification for the scheme definition is included, credible reasons for not providing it should be given.
- The scheme definition SHOULD include a clear Security Considerations (**Section 3.7**) or explain why a full security analysis is not available (e.g., in a third-party scheme registration).
- If the scheme definition does not meet the guidelines laid out in **Section 3**, the differences and reasons SHOULD be noted.

5. Guidelines for Historical URI/IRI Scheme Registration

In some circumstances, it is appropriate to note a URI scheme that was once in use or registered but for whatever reason is no longer in common use or the use is not recommended. In this case, it is possible for an individual to request that the scheme be registered (newly, or as an update to an existing registration) as 'historical'. Any scheme that is no longer in common use MAY be designated as historical; the registration should contain some indication to where the scheme was previously defined or documented.

6. URI/IRI Scheme Registration Procedure

6.1. General

The URI/IRI registration process is described in the terminology of **[RFC5226]**. The registration process is an optional mailing list review, followed by "Expert Review". The registration request should note the desired status. The Designated Expert will evaluate the request against the criteria of the requested status. In the case of a permanent registration request, the Designated Expert may:

- Accept the specification of the scheme for permanent registration.
- Suggest provisional registration instead.
- Request IETF review and IESG approval; in the meanwhile, suggest provisional registration.

URI/IRI scheme definitions contained within other IETF documents (Informational, Experimental, or Standards-Track RFCs) must also undergo Expert Review; in the case of Standards-Track documents, permanent registration status approval is required.

The registration procedure for URI schemes is intended to be very lightweight for non-contentious registrations. For the most part, we expect the good sense of submitters and reviewers, guided by these procedures, to achieve an acceptable and useful consensus for the community.

In exceptional cases, where the negotiating parties cannot form a consensus, the final arbiter of any contested registration shall be the IESG.

If parties achieve consensus on a registration proposal that does not fully conform to the strict wording of this procedure, this should be drawn to the attention of a relevant member of the IESG.

6.2. Registration Procedures

Someone wishing to register a new URI/IRI scheme MUST:

1. Check the IANA URI scheme registry to see whether or not there is already an entry for the desired name. If there is already an entry under the name, choose a different URI scheme name, or update the existing scheme definition.
2. Prepare a URI/IRI scheme registration template, as specified in **Section 6.4**. The scheme registration template may be contained in an Internet Draft, submitted alone, or as part of some other permanently available, stable, protocol specification. The template may also be submitted in some other form (as part of another document or as a stand-alone document), but the contents will be treated as an "IETF Contribution" under the guidelines of **[RFC3978]**.
3. Send a copy of the template or a pointer to the containing document (with specific reference to the section with the template) to the mailing list **uri-review@ietf.org**, requesting review. In addition, request review on other relevant mailing lists as appropriate. For example, general discussion of URI/IRI syntactical issues could be discussed on uri@w3.org; schemes for a network protocol could be discussed on a mailing list for that protocol. Allow a reasonable time for discussion and comments. Four weeks is reasonable for a permanent registration requests.
4. Respond to review comments and make revisions to the proposed registration as needed to bring it into line with the guidelines given in this document.
5. Submit the (possibly updated) registration template (or pointer to document containing it) to IANA at iana@iana.org, specifying whether 'permanent' or 'provisional' registration is requested.

Upon receipt of a URI/IRI scheme registration request, the following steps MUST be followed:

1. IANA checks the submission for completeness; if sections are missing or citations are not correct, IANA may reject the registration request.
2. IANA checks the current registry for a entry with the same name; if such a registry exists, IANA may reject the registration request.
3. IANA requests Expert Review of the registration request against the corresponding guidelines (from this document.)
4. The Designated Expert may request additional review or discussion, as necessary.
5. If Expert Review recommends registration 'provisional' or 'permanent' registration, IANA adds the registration to the appropriate registry.
6. Unless Expert Review has explicitly rejected the registration request within two weeks, IANA should automatically add the registration in the 'provisional' registry.

Either based on an explicit request or independently initiated, the Designated Expert or IESG may request the upgrade of a 'provisional' registration to a 'permanent' one. In such cases, IANA should move the corresponding entry from the provisional registry.

6.3. Change Control

Registrations may be updated in each registry by the same mechanism as required for an initial registration. In cases where the original definition of the scheme is contained in an IESG-approved document, update of the specification also requires IESG approval.

Provisional registrations may be updated by the original registrant or anyone designated by the original registrant. In addition, the IESG may reassign responsibility for a provisional registration scheme, or may request specific changes to a scheme registration. This will enable changes to be made to schemes where the original registrant is out of contact, or unwilling or unable to make changes.

Transition from 'provisional' to 'permanent' status may be requested and approved in the same manner as a new 'permanent' registration. Transition from 'permanent' to 'historical' status requires IESG approval. Transition from 'provisional' to 'historical' may be requested by anyone authorized to update the provisional registration.

6.4. URI/IRI Scheme Registration Template

This template describes the fields that must be supplied in a URI/IRI scheme registration request:

Resource Identifier (RI) Scheme name:
> See **Section 3.8** for guidelines.

Status:
> This reflects the status requested, and should be one of 'permanent', 'provisional', or 'historical'.

Scheme syntax:
> See **Section 3.2** for guidelines.

Scheme semantics:
> See **Section 3.3** and **Section 3.4** for guidelines.

Encoding considerations:
> See **Section 3.3** and **Section 3.6** for guidelines.

Applications/protocols that use this scheme name:
> See **Section 3.5**.

Interoperability considerations:
> If the person or group registering the scheme is aware of any details regarding the scheme that might impact interoperability, identify them here. For example: proprietary or uncommon encoding methods; inability to support multibyte character sets; incompatibility with types or versions of any underlying protocol.

Security considerations:
> See **Section 3.7** for guidelines.

Contact:
> Person (including contact information) to contact for further information.

Author/Change controller:
> Person (including contact information) authorized to change this.

References:
> Include full citations for all referenced documents. Registration templates for provisional registration may be included in an Internet Draft; when the documents expire or are approved for publication as an RFC, the registration will be updated.

7. The "example" Scheme

There is a need for a URI/IRI Scheme name that can be used for examples in documentation without fear of conflicts with current or future actual schemes. The URI/IRI Scheme "example" is hereby registered as a Permanent URI/IRI Scheme for that purpose.

Scheme name:
> example

Status:
> permanent

Scheme syntax:
> The entire range of allowable syntax for URI/IRI schemes specified in **[RFC3987bis]** is allowed for "example" URI/IRIs.

Scheme semantics:
> URI/IRIs in the "example" scheme should be used for documentation purposes only. The use of "example" URIs/IRIs must not be used as locators, identify any resources, or specify any particular set of operations.

Encoding considerations:
> See Section 2.5 of **[RFC3986]** for guidelines.

Applications/protocols that use this URI scheme name:
> The "example" URI should be used for documentation purposes only. It MUST not be used for any protocol.

Interoperability considerations:
> None.

Security considerations:
> None.

Contact:
> N/A

Author/Change controller:
    IETF
References:
    This RFC XXXX. RFC Editor Note: Replace XXXX with this RFC's reference.

8. IANA Considerations

Previously, the former "URL Scheme" registry was replaced by the Uniform Resource Identifier scheme registry. The process was based on **[RFC5226]** "Expert Review" with an initial (optional) mailing list review.

The updated template has an additional field for the status of the scheme, and the procedures for entering new name schemes have been augmented. **Section 6** establishes the process for new URI/IRI scheme registration.

IANA is requested to update the name of the registry "URI Schemes" to "URI/IRI Schemes". The registry should be updated to point to this document. For the tables within that registry "Permanent URI Schemes" should become "Permanent URI/IRI Schemes", "Provisional URI Schemes" should become "Provisional URI/IRI Schemes", and "Historical URI Schemes" should become "Historical URI/IRI Schemes".

The example URI scheme "example" is hereby registered. (See the template above for registration.)

9. Security Considerations

All registered values are expected to contain accurate security consideration sections; 'permanent' registered scheme names are expected to contain complete definitions.

Information concerning possible security vulnerabilities of a protocol may change over time. Consequently, claims as to the security properties of a registered URI/IRI scheme may change as well. As new vulnerabilities are discovered, information about such vulnerabilities may need to be attached to existing documentation, so that users are not misled as to the true security properties of a registered URI scheme.

10. Acknowledgements

Many thanks to Patrick Faltstrom for his comments on this version.

Many thanks to Paul Hoffmann, Ira McDonald, Roy Fielding, Stu Weibel, Tony Hammond, Charles Lindsey, Mark Baker, and other members of the uri@w3.org mailing list for their comments on earlier versions.

Parts of this document are based on **[RFC2717]**, **[RFC2718]** and **[RFC3864]**. Some of the ideas about use of URIs were taken from the "Architecture of the World Wide Web" **[W3CWebArch]**.

Appendix A. Changes Since RFC 4395

1. Significant edits to be clear that a "URI scheme" and an "IRI scheme" are the same thing.
2. Added the "example:" URL Scheme.
3. Allow for IRI-specific scheme registration.
4. Clarify that the URI scheme registry is also the IRI scheme registry.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).

[RFC2141] Moats, R., "URN Syntax," RFC 2141, May 1997 (TXT, HTML, XML).

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs," BCP 26, RFC 5226, May 2008 (TXT).

[RFC3978] Bradner, S., "IETF Rights in Contributions," RFC 3978, March 2005 (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).

[RFC3987bis] Duerst, M., Masinter, L., and M. Suignard, "Internationalized Resource Identifiers (IRIs)," September 2010.

11.2. Informative References

[RFC2717] Petke, R. and I. King, "Registration Procedures for URL Scheme Names," BCP 35, RFC 2717, November 1999 (TXT, HTML, XML).

[RFC2718] Masinter, L., Alvestrand, H., Zigmond, D., and R. Petke, "Guidelines for new URL Schemes," RFC 2718, November 1999 (TXT, HTML, XML).

[RFC3406] Daigle, L., van Gulik, D., Iannella, R., and P. Faltstrom, "Uniform Resource Names (URN) Namespace Definition Mechanisms," BCP 66, RFC 3406, October 2002 (TXT).

[RFC3864] Klyne, G., Nottingham, M., and J. Mogul, "Registration Procedures for Message Header Fields," BCP 90, RFC 3864, September 2004 (TXT).

[RFC3987] Duerst, M. and M. Suignard, "Internationalized Resource Identifiers (IRIs)," RFC 3987, January 2005 (TXT).

[RFC4395] Hansen, T., Hardie, T., and L. Masinter, "Guidelines and Registration Procedures for New URI Schemes," BCP 35, RFC 4395, February 2006 (TXT).

[RFC4690] Klensin, J., Faltstrom, P., Karp, C., and IAB, "Review and Recommendations for Internationalized Domain Names (IDNs)," RFC 4690, September 2006 (TXT).

[W3CWebArch] W3C Technical Architecture Group, "Architecture of the World Wide Web, Volume One," December 2004.

Authors' Addresses

Tony Hansen
AT&T Laboratories
200 Laurel Ave.
Middletown, NJ 07748
USA
Email: tony+urireg@maillennium.att.com

Ted Hardie
Panasonic Wireless Research Lab
10900 Tantau Ave.
Cupertino, CA
USA
Phone: +1 408 628 5864
Email: ted.ietf@gmail.com

Larry Masinter
Adobe
345 Park Ave.
San Jose, CA 95110
US
Phone: +1 408 536 3024
Email: masinter@adobe.com
URI: http://larry.masinter.net