IETF URNbis WG
Internet-Draft
Obsoletes: 2141 (if approved)
Intended status: Standards Track
Expires: May 28, 2011

A. Hoenes, Ed.
TR-Sys

November 24, 2010

Uniform Resource Name (URN) Syntax
draft-ietf-urnbis-rfc2141bis-urn-00

Abstract

Uniform Resource Names (URNs) are intended to serve as persistent, location-independent, resource identifiers. This document serves as the foundation of the 'urn' URI Scheme according to RFC 3986 and sets forward the canonical syntax for URNs, which subdivides URNs into "namespaces". A discussion of both existing legacy and new namespaces and requirements for URN presentation and transmission are presented. Finally, there is a discussion of URN equivalence and how to determine it. This document supersedes RFC 2141.

The requirements and procedures for URN Namespace registration documents are currently set forth in RFC 3406, which is also being updated by a companion, revised specification dubbed RFC 3406bis.

Discussion

This draft version has been obtained by importing the text from RFC 2141 into modern tools and making a first rounds of updating steps. It is a chartered initial work item of the URNbis WG in the IETF; the aim is to bring URN RFCs in alignment with STD 66, STD 68, BCP 26, and the requirements from emerging distributed national and international URN resolution systems, and advance them on the IETF Standards Track.

Comments are welcome on the urn@ietf.org mailing list (or sent to the document editor). The home page of the URNbis WG is located at <http://tools.IETF.ORG/wg/urnbis/>.

Status of This Memo

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 May 28, 2011.

Table of Contents

1. Introduction

'urn' is a particular URI Scheme (according to STD 66, RFC 3986 [RFC3986] and BCP 35, RFC 4395 [RFC4395]) that is dedicated to forming a hierarchical framework for persistent identifiers.

Uniform Resource Names (URNs) are intended to serve as persistent, location-independent, resource identifiers and are designed to make it easy to map other namespaces (that share the properties of URNs) into URI-space. Therefore, the URN syntax provides a means to encode character data in a form that can be sent in existing protocols, transcribed on most keyboards, etc.

The first level of hierarchy is given by the classification of URIs into "URI Schemes", and for URNs, the second level is organized into "URN Namespaces". Henceforth both terms are used in this capitalization to distinguish them from the more general common meaning of "scheme" and "namespace".

1.1. Historical Perspective and Motivation

For the intended audience of this RFC, which is expected to include groups interested in persistent identifiers in general and not in continuous contact with the IETF and the RFC series, this section gives a brief outline of the evolution of the matter over time. Appendix A gives hints on how to obtain RFCs and related information.

Attempts to define generally applicable identifiers for network resources go back to the mid-1970 years. Among the applicable RFCs is RFC 615 [RFC0615], which subsequently has been obsoleted by RFC 645 [RFC0645].

The seminal document in the RFC series regarding URIs (Uniform Resource Identifiers) for use with the World Wide Web (WWW) has been RFC 1630 [RFC1630], published in 1994. In the same year, the general concept or Uniform Resource Names has been laid down in RFC 1737 [RFC1737] and that of Uniform Resource Locators in RFC 1736 [RFC1736].

The original formal specification of URN Syntax, RFC 2141 [RFC2141] has been adopted in 1997. That document was based on the original specification of URLs (Uniform Resource Locators) in RFC 1738 [RFC1738] and RFC 1808 [RFC1808], which later on, in 1998, has been generalized and consolidated in the Generic URI specification, RFC 2396 [RFC2396]. Most parts of these URI/URL documents have been superseded in 2005 by STD 66, RFC 3986 [RFC3986]. Notably, RFC 2141 makes (essentially normative) reference to a draft version of RFC 2396.

Over time, the terms "URI", "URL", and "URN" have been refined and slightly shifted according to emerging insight and use. This has been clarified in a joint effort of the IETF and the World Wide Web Council, published 2002 for the IETF in RFC 3305 [RFC3305].

The wealth of URI Schemes and URN Namespaces needs to be organized in a persistent way, in order to guide application developers and users to the standardized top level branches and the related specifications. These registries are maintained by the Internet Assigned Numbers Authority (IANA) [IANA] at [IANA-URI] and [IANA-URN], respectively. Registration procedures for URI Schemes originally had been laid down in RFC 2717 [RFC2717] and guidelines for the related specification documents were given in RFC 2718 [RFC2718]. These documents have been obsoleted and consolidated into BCP 35, RFC 4395 [RFC4395], which is based on, and aligned with, RFC 3986.

> Note that RFC 2141 predates RFC 2717 and, although the 'urn' URI scheme is listed in [IANA-URI] with a pointer to RFC 2141, this registration has never been performed formally.

Similarly, the URN Namespace definition and registration mechanisms originally have been specified in RFC 2611 [RFC2611], which has been obsoleted by BCP 66, RFC 3406 [RFC3406]. Guidelines for documents prescribing IANA procedures have been revised as well over the years, and at the time of this writing, BCP 26, RFC 5226 [RFC5226] is the normative document. Neither RFC 4395 nor RFC 3406 conform to RFC 5226.

Early documents specifying URI and URN syntax, including RFC 2141, made use of an ad-hoc variant of the original Backus-Naur Form (BNF) that never has been formally specified.

Over the years, the IETF has shifted to the use of a predominant formal language used to define the syntax of textual protocol elements, dubbed "Augmented Backus-Naur Form" (ABNF). The specification of ABNF also has evolved, and now STD 68, RFC 5234 [RFC5234] is the normative document for it (that also will be used in this RFC).

1.2. Background on Properties of URNs

RFC 1738 [RFC1738] defined the purpose of URNs as follows:

- The purpose or function of a URN is to provide a globally unique, persistent identifier used for recognition, for access to characteristics of the resource or for access to the resource itself.

Section 2 of RFC 1738 [RFC1738] listed the functional requirements for URNs (quote slightly edited to reflect the time passed since that RFC had been written and the actual definition of the URN scheme that has happened):

- Global scope: A URN is a name with global scope which does not imply a location. It has the same meaning everywhere.
- Global uniqueness: The same URN will never be assigned to two different resources.
- Persistence: It is intended that the lifetime of a URN be permanent. That is, the URN will be globally unique forever, and may well be used as a reference to a resource well beyond the lifetime of the resource it identifies or of any naming authority involved in the assignment of its name.
- Scalability: URNs can be assigned to any resource that might conceivably be available on the network, for hundreds of years.
- Legacy support: The URN scheme permits the support of existing legacy naming systems, insofar as they satisfy the other requirements described here. [...]
- Extensibility: The URN scheme permits future extensions.
- Independence: It is solely the responsibility of a name issuing authority to determine the conditions under which it will issue a name.
- Resolution: URNs will not impede resolution. [...]

The URN syntax described below also accommodates the fundamental "Requirements for URN Encoding" in Section 3 of RFC 1738 [RFC1738], as far as experience gained has not lead to lessen unrealistical detail requirements:

- Single encoding: The encoding for presentation for people in clear text, electronic mail and the like is the same as the encoding in other transmissions.
- Simple comparison: A comparison algorithm for URNs is simple, local, and deterministic. [...]
- Human transcribability: For URNs to be easily transcribable by humans without error, they need to be short, use a minimum of special characters, and be case insensitive. [...]

- Transport friendliness: A URN can be transported unmodified in the common Internet protocols, such as TCP, SMTP, FTP, Telnet, etc., as well as printed paper.

- Machine consumption: A URN can be parsed by a computer.

- Text recognition: The encoding of a URN needs to enhance the ability to find and parse URNs in free text.

1.3. Objective of this Memo

RFC 2141 does not seamlessly match current Internet Standards. The primary objective of this document is the alignment with the URI Standard [RFC3986] and guidelines [RFC4395], the ABNF Standard [RFC5234] and the current IANA Guidelines [RFC5226] in general.

Further, experience from emerging international efforts to establish a general, distributed, stable URN resolution service are expected to be taken into account during the draft stage of this document.

For advancing the URN specification on the Internet Standards-Track, it needs to be based on documents of comparable maturity. Therefore, to further advancements of the formal maturity level of this RFC, it deliberately makes normative references only to documents at Full Standard or Best Current Practice level.

Thus, this replacement document for RFC 2141 should make it possible to advance the URN framework on the Internet Standard maturity ladder. All other related documents depend on it; therefore this is the first step to undertake.

Out of scope for this document is a revision of the URN Namespace Definition Mechanisms document, BCP 66 [RFC3406]. This is going to be undertaken in a companion document, RFC 3406bis.

1.4. Requirement Language

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 RFC 2119 [RFC2119].

2. URN Syntax

This document defines the URI Scheme 'urn'. Hence, URNs are specific URIs as specified in STD 66 [RFC3986]. The formal syntax definitions below are given in ABNF according to STD 68 [RFC5234] and make use of some "Core Rules" specified in Appendix B of that Standard and several generic rules defined in Appendix A of RFC 3986.

The syntax definitions below do, and syntax definitions in dependent documents MUST, conform to the URI syntax specified in RFC 3986, in the sense that additional syntax rules must only constrain the general rules from RFC 3986. In other words: a general URI parser based on RFC 3986 MUST be able to parse any legal URN, and specific semantics can be obtained from URN-specific parsing.

> NOTE: The remainder of this Section still requires substantial work! To give a starting point for WG discussion, within this entire Section, much of the elaborations and editorial comments from the Individual I-D predecessor of this draft are kept. This will be cleaned up after discussion.

URNs conform to the <path-rootless> variant of the general URI syntax specified in Section 3 of [RFC3986] :

```
URI = scheme ":" path-rootless [ "?" query ] [ "#" fragment ]

path-rootless = segment-nz *( "/" segment )

segment-nz    = 1*pchar
segment       = *pchar

pchar = unreserved / pct-encoded / sub-delims / ":" / "@"
```

In the case of URNs, we have:

```
scheme     = "urn"
```

and the following additional syntax rule is superimposed on <path-rootless> to establish a level of hierarchy called "Namespace":

```
urn-path   = NID ":" NSS
```

Here "urn" is the URI scheme name, <NID> is the Namespace Identifier, and <NSS> is the Namespace Specific String. The colons are REQUIRED separator characters.

Per RFC 3986, the URN Scheme name (here "urn") is case-insensitive.

The Namespace ID (also a case-insensitive string) determines the syntactic structure and the semantic interpretation of the Namespace Specific String. Generic details on NID syntax can be found below in Section 2.1, and the NSS syntax is elaborated upon in Section 2.2.

Each particular namespace is based on a specific document that must normatively describe (among other things) the details of the <NSS> values allowed in conjunction with the respective <NID>. The

specification requirements and registration procedures for URN Namespaces are the subject of a dedicated document, currently RFC 3406 [RFC3406] -- to be updated for conformance to BCP 26 and alignment with implementation experience, in RFC 3406bis.

Note:

> RFC 2141 has deferred the decision on whether <query> and <fragment> components are applicable to URNs and reserved the use of bare (unencoded) question mark ("?") and hash ("#") characters in URNs for future usage in conformance with the generic URI syntax.
>
> There is evidence of desire to be able to use these components (which are split off by the high-level parsing rules of RFC 3986), or at least the <fragment> component, in URNs belonging to selected namespaces. Thus, this draft version tentatively aims at allowing these components in the general syntax.
>
> The considerations below reflect the current thinking based on implementation experience and preliminary discussion.

The syntax of <query> and <fragment> are defined in RFC 3986. Question mark and hash sign remain reserved as separator characters for these URI components and cannot appear unencoded in an NSS. This way, backwards compatibility with existing URN namespaces is guaranteed and compatibility with general URI parsers is improved.

The <query> part MUST NOT be present in any *assigned* URN. This specification reserves its use for future standardization related to URN resolution. This part can only be added to an assigned URN and appear in a URI reference [RFC3986] to a URN that is intended to be used with URN resolution services, and, in accordance with the general specification of this part in RFC 3986, its purpose is restricted to designate service aspects of the intended resolution response, e.g., to select the kind and amount of metadata sought about the given object that is identified by the basic, assigned URN.

The <fragment> part is not generally allowed in URNs. It is only applicable to URN Namespaces that specifically opt to support its usage. Thus, a URN Namespace registration document MAY specify the usage of <fragment> with URNs of that particular URN Namespace. Absent a registered namespace definition based on this document and RFC 3406bis that explicitly specifies its usage, URNs assigned within a particular URN Namespace MUST NOT contain a fragment identifier.

The use of fragment identifiers may be useful if the URN Namespace is based on an existing identifier scheme that designates objects of

reasonable complexity that there's a need to make reference of parts of such resources in typical network access environments.

A URN Namespace definition has two options to support fragment identifiers, and only one of these methods is possible within a given URN Namespace:

(a) Fragment identifiers (if any) are assigned individually to parts of a larger entity during the URN assignment process. If a URN Namespace opts for this model, its specification MUST describe the additional syntax restrictions to be adhered and the particulars of the (per-URN) assignment process.

(b) A specific set of fragment identifiers is generally applicable to all resources targeted by URNs of the specific URN Namespace. In this case, the specification document MUST specify a finite set of <fragment> values, or precise, generic rules for the formation of syntactically valid fragment identifiers for the particular URN Namespace. The specification SHOULD indicate the treatment of syntactically valid <fragment> values in case they are not semantically valid for a given base URN. Absent such specification, the default is to ignore such fragment identifiers.

URN resolver clients MUST pass a given <fragment> part of a URN unchanged to the resolver service. The default URN resolution behavior is to ignore any <fragment> part if either the applicable URN Namespace definition did not specify its use, or if no specific related information was available for the basic resource in case (b) above, or if that basic URN plus fragment identifier has not been assigned in case (a) above.

2.1. Namespace Identifier (NID) Syntax

The following is the syntax for the Namespace Identifier. To (A) be consistent with all potential resolution schemes and (B) not put any undue constraints on any potential resolution scheme, Namespace Identifiers are ASCII strings with the syntax:

```
NID = ( ALPHA / DIGIT ) 0*31 ( ALPHA / DIGIT / "-" )
```

Note for discussion:
The above definition is taken from RFC 2141. Should this be further restricted, e.g., to avoid possible confusion caused by multiple adjacent hyphens and NIDs looking like a numerical value or a numerical range? Does it really make sense to allow single-letter NIDs? Such restrictions would be fully backward compatible because no NIDs have been defined so far that would violate these

restrictions. Hyphens have been used only in the naming pattern for "Informal Namespace IDs" per RFC 3406.

Namespace Identifiers are case-insensitive, so that for instance "ISBN" and "isbn" refer to the same namespace.

To avoid confusion with the URI Scheme name "urn", the NID "urn" is permanently reserved by this RFC and MUST NOT be used or registered.

2.2. Namespace Specific String (NSS) Syntax

Note:
In order to make visible the migration path from RFC 2141 and the influence of the evolution of URI syntax from RFC 2396 to RFC 3986 on it, at this draft stage, the subsequent syntax description is highly annotated and expanded. After discussion, a substantial consolidation is expected.

As already required by RFC 1737, there is a single canonical representation of the NSS portion of an URN.

Note:
If the DISCUSSes above and below can be affirmed (allowing optional <query> and <fragment> components as well as "&" and "~" in the path), the syntax below could be simplified very much to:

```
NSS   = 1*pchar   ; or equivalent:    NSS   = segment-nz
```

The format of this single canonical form follows:

```
      NSS         = 1*URN-char

      URN-char    = trans / pct-encoded

      trans       = ALPHA / DIGIT / u-other
   ; NO?          / reserved
   ; Issue: This lead to ambiguity in RFC 2141 wrt "%".

      u-other     = ":" / "@"
                    ; those from RFC 3986 <gen-delims>
                    ; specifically allowed in <pchar>.
   ; From RFC 3986:
   ;    gen-delims  = ":" / "/" / "?" / "#" / "[" / "]" / "@"

                  / "!" / "$" /       "'" / "(" / ")"
                  / "*" / "+" / "," / ";" / "="
                    ; this is RFC 3986 <sub-delims> except "&".
   ; From RFC 3986:
   ;    sub-delims  = "!" / "$" / "&" / "'" / "(" / ")"
   ;                / "*" / "+" / "," / ";" / "="
   ; Issue: can/should "&" be allowed ?
   ; If we allow <query> and <fragment> according to the
   ; generic URI syntax, there seems to be no more need to exclude "&".

                  / "-" / "." / "_"   ; <unreserved> except "~"
   ; From RFC 3986:
   ;    unreserved  = ALPHA / DIGIT
   ;                / "-" / "." / "_" / "~"
   ; Issue: can/should "~" be allowed as well ?

   ; If we allow "&" and "~" , <trans> becomes <pchar> ,
   ; greatly simplifying the syntax rules and parsers!

   ; from RFC 2141:
   ;    reserved  = '%" / "/" / "?" / "#"         ; SIC!
   ;                ^ ^
```

Depending on the rules governing a namespace, valid identifiers in a namespace might contain characters that are not members of the URN character set above (<URN-char>). Such strings MUST be translated into canonical NSS format before using them as protocol elements or otherwise passing them on to other applications. Translation is done by encoding each character outside the URN character set as a sequence of octets using UTF-8 encoding STD 63 [RFC3629], and the "percent-encoding" of each of those octets as "%" followed by two <HEXDIG> characters. The two characters form the hexadecimal representation of that octet.

2.3. Special and Reserved Characters

The remaining printable characters left to be discussed above comprise the generic delimiters and the reserved characters, which are restricted for special use only. These characters are discussed below, giving the specifics of why each character is special or reserved.

2.3.1. Delimiter Characters

RFC 3986 [RFC3986] defines the general delimiter characters used in URIs:

```
gen-delims = ":" / "/" / "?" / "#" / "[" / "]" / "@"
```

From among the <gen-delims>, ":" and "@" are also included in <pchar> and hence allowed in the path components of URIs.

The at-character ("@") in generic URIs only has a specific meaning when contained in the <authority> part, which is absent in URNs. Hence, "@" is available in the <NSS> part of URNs.

With URNs, the colon (":") is used as a delimiter character not only between the scheme name ("urn") and the <NID>, but also between the latter and the <NSS>, and many existing URN namespaces additionally use ":" to further subdivide a single RFC 3986 path segment in the <NSS> in a hierarchical manner.

Note: Using ":" as a sub-delimiter in the path in favor of "/" is attractive because it avoids possible complications that could arise from accidental inappropriate use of relative URI references [RFC3986] for URNs.

The characters "/", "?", and "#" separate path components and the <query> and <fragment> parts in the generic URI syntax; they are restricted to this role in URNs as well, although the <path> in URNs only admits a single and hence "/" is not allowed. Therefore, these characters MUST NOT appear in the <NSS> part of a URN in unencoded form. Namespaces that need these characters MUST employ in their URNs the appropriate percent-encoding for each such character.

The square brackets ("[" and "]") also play a particular role when contained in the <authority> part, which is absent in URNs. However, for conformance with the generic URI syntax, they are not allowed literally in the <NSS> component of URNs. If a specific URN namespace reflects semantics that require these characters, they MUST be percent-encoded in the respective URNs.

2.3.2. The Percent Character

The percent character ("%") is reserved in the URN syntax for introducing the escape sequence for an octet that is either not a printable ASCII character or reserved for special purposes, as described in this section. The presence of a "%" character in a URN MUST always be followed by two <HEXDIG> characters, which three together semanticaly form an abstract <pct-encoded> octet. Literal use of the "%" character in an underlying namespace MUST therefore be encoded as "%25" in URNs for that namespace.

Namespaces MAY designate one or more characters from the URN character set as having special meaning for that namespace. If the namespace also uses that character in a literal sense as well, the character used in a literal sense MUST be encoded with "%" followed by the hexadecimal representation of that octet. Further, a character MUST NOT be percent-encoded if the character is not a reserved character. Therefore, the process of registering a namespace identifier shall include publication of a definition of which characters have a special meaning to that namespace.

2.3.3. Other Excluded Characters

The following list is included only for the sake of completeness. It includes the characters discussed in Sections 2.3.1 and 2.3.2. Any octets/characters on this list are explicitly NOT part of the URN <NSS> character set, and if used in an URN, MUST be percent-encoded.

```
           excluded = CTL / SP        ; control characters and space
                    / DQUOTE          ; "
                    / "#"             ; from <gen-delims>
                    / "%"             ; see above
   ; DISCUSS!       / "&"             ; DISCUSS -- see above!
                    / "/"             ; from <gen-delims>
                    / "<" / ">"
                    / "?"             ; from <gen-delims>
                    / "["             ; from <gen-delims>
                    / "\"
                    / "]"             ; from <gen-delims>
                    / "^"
                    / "`"
                    / "{" / "|" / "}"
   ; DISCUSS!       / "~"             ; DISCUSS -- see above!
                    / %x7F            ; DEL (control character)
                    / %x80-FF         ; non-ASCII
```

The NUL octet (0 hex) is renowned for a long history of trouble in implementations. It MUST NOT be used, in either unencoded or percent-encoded form.

In textual context, a URN ends when an octet/character from the excluded character set (<excluded>) is encountered. The character from the excluded character set is NOT part of the URN.

[Does that still make sense? -- it collides with possible query / fragment parts!]

3. Support of Existing Legacy Naming Systems and New Naming Systems

Any namespace (existing or newly devised) that is proposed as a URN namespace and fulfills the criteria of URN namespaces MUST be expressed in this syntax. If names in these namespaces contain characters other than those defined for the URN character set, they MUST be translated into canonical form as discussed in Section 2.2.

4. URN Presentation and Transport

The URN syntax defines the canonical format for URNs and all URN transport and interchanges MUST take place in this format. Further, all URN-aware applications MUST offer the option of displaying URNs in this canonical form to allow for direct transcription (for example by cut-and-paste techniques). Such applications MAY support display of URNs in a more human-friendly form and may use a character set that includes characters that aren't permitted in URN syntax as defined in this RFC (that is, they may replace %-notation by characters in some extended character set in display to humans).

5. Lexical Equivalence of URNs

For various purposes such as caching, it is often desirable to determine whether two URNs are the same without resolving them. The general-purpose means of doing so is by testing for "lexical equivalence" as defined below.

Two URNs are lexically equivalent if they are octet-by-octet equal after the following preprocessing:

1. normalize the case of the leading "urn" scheme name;
2. normalize the case of the NID;
3. normalize the case of any percent-encoding.

Note that percent-encoding MUST NOT be removed. It is an implementation detail not affecting interoperability whether a URN comparison function internally prefers normalization (in the above 3 steps) to lower or to upper case.

Some namespaces may define additional lexical equivalences, such as case-insensitivity of the NSS (or parts thereof). Additional lexical equivalences MUST be documented as part of namespace registration, MUST always only have the effect of eliminating some of the false negatives obtained by the procedure above, i.e. they MUST NOT say that two URNs are not equivalent if the procedure above says they are equivalent.

5.1. Examples of Lexical Equivalence

The following hypothetical URN comparisons highlight the lexical equivalence definitions:

1- URN:foo:a123,456
2- urn:foo:a123,456
3- urn:FOO:a123,456
4- urn:foo:A123,456
5- urn:foo:a123%2C456
6- URN:FOO:a123%2c456

URNs 1, 2, and 3 are all lexically equivalent. URN 4 is not lexically equivalent to any of the other URNs of the above set. URNs 5 and 6 are only lexically equivalent to each other.

6. Functional Equivalence of URNs

Functional equivalence is determined by practice within a given namespace and managed by resolvers for that namespace. Thus, it is beyond the scope of this document. Namespace registrations must include guidance on how to determine functional equivalence for that namespace, i.e., when two URNs are identical within a namespace.

On the other hand, it is permissible to have two different URNs -- even from different URN namespaces -- be assigned to a particular resource. This can only be detected by resolving the URNs and analysis of the resolution responses; hence, this is out of scope for this memo.

7. The ’urn’ URI Scheme

At the time of publication of RFC 2141, no formal registration procedure for URI Schemes had been established yet, and so IANA only informally has registered the ’urn’ URI Scheme with a reference to [RFC2141].

Section 7.1 below contains the URI scheme registration template for the ’urn’ scheme, in accordance with RFC 4395 [RFC4395].

Note: In order to be usable as a standalone text (after being extracted from this RFC), the template below does not contain formal anchors to the references listed in section 11, but instead gives the common document designations in prose. However, for compliance with editorial policy, it needs to be noted here:

This registration template refers to RFCs 2196, 2276, 2608, 3401 through 3404, 3406, 3629 (STD 63), and 3986 (STD 66) ([RFC2169] [RFC2276] [RFC2608] [RFC3401] [RFC3402] [RFC3403] [RFC3404] [RFC3406] [RFC3629] [RFC3986]).

7.1. Registration of URI Scheme 'urn'

[RFC Editor: Please replace "XXXX" in all instances of "RFC XXXX" below by the RFC number assigned to this document.]

URI scheme name: urn

Status: permanent

URI scheme syntax:

See Section 2 of RFC XXXX.

URI scheme semantics:

'urn' URIs, known as Universal Resource Names (URNs), serve as persistent, location-independent, resource identifiers for concrete and abstract objects that have network accessible instances and/or metadata.

URNs are structured hierarchically into URN Namespaces, the management of which is delegated to namespace-specific authorities. Each such URN namespace is founded in an independent specification and registered with IANA, following the guidelines and procedures of BCP 66 (at the time of this registration: RFC 3406).

Encoding considerations:

All URNs are ASCII strings conforming to the general URI syntax from STD 66. As described in Sections 2.2 and 2.3.2 of RFC XXXX, characters needed by the URN namespace specific semantics but not contained in the US-ASCII charset MUST be encoded in UTF-8 according to STD 63; any octets outside the allowed character set MUST then be percent-encoded.

Applications/protocols that use this URI scheme:

URNs that serve to identify abstract resources for protocol purposes are expected to be recognized directly by the implementations of these portocols.

In general, resolution systems for URNs are specified on a per-namespace basis. If appropriate for the namespace, these systems resolve URNs to (possibly multiple) URIs that allow the network access to the identified object or metadata on it.

"Architectural Principles of Uniform Resource Name Resolution" (RFC 2276) explains the basic concepts. Some resolution systems laid down in IETF specifications are:

* Trivial HTTP-based URN Resolution (RFC 2169)

* Dynamic Delegation Discovery System (DDDS, RFCs 3401-3404)

* Service Location Protocol (SLPv2, RFC 2608)

Interoperability Considerations:

Persistence and stability of URNs require appropriate resolution systems.

Security Considerations:

See Section 8 of RFC XXXX.

Contact:

The IETF URNbis working group.
This registration will be discussed on the following IETF lists:
urn and uri-review (AT ietf.org).

Author / Change controller:

The authors of RFC XXXX.
Change control is with the IESG.

References:

RFC XXXX.

Procedures for the specification and registration of URN namespaces are detailed in BCP 66 (at the time of this writing: RFC 3406; the URNbis WG is chartered to provide a RFC 3406bis).

8. Security Considerations

This document specifies the syntax and general requirements for URNs, which are the specific URIs that use the 'urn' URI scheme. As such, the general security considerations of STD 66 [RFC3986] apply. However, each URN namespace will have specific security considerations, according to the semantics and usage of the underlying namespace. While some namespaces may assign special meaning to certain of the characters of the Namespace Specific String, any security considerations resulting from such assignment are outside the scope of this document. It is REQUIRED by BCP 66 [RFC3406] that the process of registering a namespace identifier include any such considerations.

9. IANA Considerations

IANA is asked to update the existing informal registration of the 'urn' URI Scheme by the template in Section 7.1 above and list this RFC as the current normative reference in [IANA-URI].

IANA is asked to add a note to [IANA-URN] that 'urn' is a permanently reserved formal namespace identifier string that cannot be registered, in order to avoid confusion with the 'urn' URI scheme.

10. Acknowledgements

This document is heavily based on RFC 2141, the author of which has laid the foundation for this work; that RFC contained the following Acknowledgements:

> Thanks to various members of the URN working group for comments on earlier drafts of this document. This document is partially supported by the National Science Foundation, Cooperative Agreement NCR-9218179.

This document also heavily relies on and acknowledges the work done for STD 66 [RFC3986] and earlier RFCs that are being quoted informally, in particular RFC 1737 [RFC1737].

Your name could go here ...

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.

[RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, November 2003.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

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

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

11.2. Informative References

[IANA] IANA, "The Internet Assigned Numbers Authority", <http://www.iana.org/>.

[IANA-URI] IANA, "URI Schemes Registry", <http://www.iana.org/assignments/uri-schemes/>.

[IANA-URN] IANA, "URN Namespace Registry", <http://www.iana.org/assignments/urn-namespaces/>.

[RFC0615] Crocker, D., "Proposed Network Standard Data Pathname syntax", RFC 615, March 1974.

[RFC0645] Crocker, D., "Network Standard Data Specification syntax", RFC 645, June 1974.

[RFC1630] Berners-Lee, T., "Universal Resource Identifiers in WWW: A Unifying Syntax for the Expression of Names and Addresses of Objects on the Network as used in the World-Wide Web", RFC 1630, June 1994.

[RFC1736] Kunze, J., "Functional Recommendations for Internet Resource Locators", RFC 1736, February 1995.

[RFC1737] Sollins, K. and L. Masinter, "Functional Requirements for Uniform Resource Names", RFC 1737, December 1994.

[RFC1738] Berners-Lee, T., Masinter, L., and M. McCahill, "Uniform Resource Locators (URL)", RFC 1738, December 1994.

[RFC1808] Fielding, R., "Relative Uniform Resource Locators", RFC 1808, June 1995.

[RFC2141] Moats, R., "URN Syntax", RFC 2141, May 1997.

[RFC2169] Daniel, R., "A Trivial Convention for using HTTP in URN Resolution", RFC 2169, June 1997.

[RFC2276] Sollins, K., "Architectural Principles of Uniform Resource Name Resolution", RFC 2276, January 1998.

[RFC2396] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax", RFC 2396, August 1998.

[RFC2608] Guttman, E., Perkins, C., Veizades, J., and M. Day, "Service Location Protocol, Version 2", RFC 2608, June 1999.

[RFC2611] Daigle, L., van Gulik, D., Iannella, R., and P. Faltstrom, "URN Namespace Definition Mechanisms", BCP 33, RFC 2611, June 1999.

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

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

[RFC3305] Mealling, M. and R. Denenberg, "Report from the Joint W3C/IETF URI Planning Interest Group: Uniform Resource Identifiers (URIs), URLs, and Uniform Resource Names (URNs): Clarifications and Recommendations", RFC 3305, August 2002.

[RFC3401] Mealling, M., "Dynamic Delegation Discovery System (DDDS) Part One: The Comprehensive DDDS", RFC 3401, October 2002.

[RFC3402] Mealling, M., "Dynamic Delegation Discovery System (DDDS) Part Two: The Algorithm", RFC 3402, October 2002.

[RFC3403] Mealling, M., "Dynamic Delegation Discovery System (DDDS) Part Three: The Domain Name System (DNS) Database", RFC 3403, October 2002.

[RFC3404] Mealling, M., "Dynamic Delegation Discovery System (DDDS) Part Four: The Uniform Resource Identifiers (URI)", RFC 3404, October 2002.

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

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

Appendix A. How to Locate IETF Documents (Informative)

Request For Comments (RFCs) are available from the RFC Editor site using the canonical URIs <http://www.rfc-editor.org/rfc/rfcNNNN.txt> or <ftp://ftp.rfc-editor.org/in-notes/rfcNNNN.txt> (where 'NNNN' is the serial number of the RFC), and from numerous mirror sites. Additional metadata for any RFC, including possible Errata, are available from <http://www.rfc-editor.org/info/rfcNNNN> (where 'NNNN' again is the serial number of the RFC). A HTML-ized version and a PDF facsimile of each RFC are available from the IETF Tools site at <http://tools.ietf.org/http/rfcNNNN> and <http://tools.ietf.org/pdf/rfcNNNN>, respectively.

Current Internet Draft documents are available via the search engines at <http://www.ietf.org/id-info/> and <http://www.rfc-editor.org/idsearch.html>; archival copies of older IETF documents can be found at <http://tools.ietf.org/id/>.

Appendix B. Handling of URNs by URL Resolvers/Browsers

The URN syntax has been defined so that URNs can be used in places where URLs are expected. A resolver that conforms to the current URI syntax specification [RFC3986] will extract a scheme value of "urn" rather than a scheme value of "urn:<nid>".

An URN MUST be considered an opaque URI by URL resolvers and passed (with the "urn:" tag) to a URN resolver for resolution. The URN resolver can either be an external resolver that the URL resolver knows of, or it can be functionality built into the URL resolver.

To avoid confusion of users, a URL browser SHOULD display the complete URN (including the "urn:" tag) to ensure that there is no confusion between URN Namespace identifiers and URI Scheme names.

Appendix C. Collected ABNF (Informative)

As a service to implementers specifically interested in URN syntax, after consolidation of Section 2, the complete ABNF for URNs will be collected here, including the referenced rules from [RFC5234] and

[RFC3986]. In case of (unexpected) inconsistencies, these documents remain normative for the respective productions.

T.B.D.

...

Appendix D. Changes since RFC 2141 (Informative)

D.1. Essential Changes from RFC 2141

[RFC Editor: please remove the Appendix D.1 headline and all subsequent subsections starting with Appendix D.2.]

T.B.D. (after consolidation of this memo)

D.2. Changes from RFC 2141 to Individual Draft -00

Abstract amended: URI scheme, replacement for 2141, point to 3406. Use contemporary boilerplate. Added transient "Discussion" section.

s1: added new 1st para (URI scheme) and 3rd para (hierarchy).
s1.1 (Historical Perspective) added for background & motivation.
s1.2 (Objective) added.
s1.3 (2119 keywords) added -- used now throughout normative text.

s2 (URN Syntax): Shifted from BNF to ABNF; explain relationship to 3986 and gaps, how the gaps could be bridged, distinguish between URI generics and URN specifics; got rid of references to immature documents (1630, 1737).
s2.1 (NID syntax): Use ABNF and RFC 5234 terminals (core rules); removed reference to an old draft of 2396; clarified prohibition to use "urn" as NID.
s2.2 (NSS syntax): Shifted from BNF to ABNF; made ABNF consistent with subsequent textual description; exposition much expanded, showing relationship with 3986 and resulting incompatibilities; proposed how to bridge gaps, to make parsing more uniform among URIs; updated i18n considerations and pointer to UTF-8 specification.
s.2.3, s2.3.*: reworked and much expanded, along the grouping of delimiter characters from 3986 in new s2.3.1 (including old s.2.3.2); made text fully consistent with ABNF in s2.2; consistent usage of term "percent-encoded"; old s.2.3.1 became s2.3.2; old s3.4 became s3.3.3, providing complete, annotated list of excluded characters, ordered by ascending code point; and restating design decisions needed to be made to close gaps to 3986.

s3 through s6: only minor editorial changes.

s7: formal registration of 'urn' URI scheme added, using 4395 template.

s8: Security Cons. slightly amended.

s9: new: IANA Cons. added wrt s7.1 and prohibition of NID "urn".

s10: Acknowledgments amended.

s11: References split into Normative and Informative; updated refs and added many; only FS and BCP allowed as Normative Refs to further promotion of document.

Added Appendices A through D.

D.3. Changes from Individual Draft -00 to -02

Updated "Discussion" on front page to point to dedicated urn list.

Numerous editorial improvements and additions for clarification, in particular in the Introduction. No technical changes.

More Informative References; missing details supplied in D.1.

D.4. Changes from Individual Draft -02 to WG Draft -00

Added new s2.1 with excerpts from RFC 1737 to Introduction to provide background on URN functional and syntax requirements.

Supplied text in s2 regarding the envisioned use of query and fragment parts, based on various discussion -- including a preliminary evaluation in PersID.

Changed "SHOULD never" to "MUST NOT" for NUL character in NSS.

Various editorial and grammar fixes; corrected STD / BCP numbers.

Author's Address

Alfred Hoenes (editor)
TR-Sys
Gerlinger Str. 12
Ditzingen D-71254
Germany

EMail: ah@TR-Sys.de

IETF URNbis WG
Internet-Draft
Obsoletes: 2288,3187
(if approved)
Intended status: Standards Track
Expires: June 20, 2011

M. Huttunen
J. Hakala
The National Library of Finland
A. Hoenes, Ed.
TR-Sys

December 17, 2010

Using International Standard Book Numbers as Uniform Resource Names
draft-ietf-urnbis-rfc3187bis-isbn-urn-00

Abstract

The International Standard Book Number, ISBN, is a widely used identifier for monographic publications. Since 2001, there has been a URN (Uniform Resource Names) namespace for ISBNs. The namespace registration was performed in RFC 3187 and applies to the ISBN as specified in the original ISO Standard 2108-1992. To allow for further growth in use, the successor ISO Standard, ISO 2108-2005, has defined an expanded format for the ISBN, known as "ISBN-13". This document defines how both the old and new ISBN standard can be supported within the URN framework and the syntax for URNs defined in RFC 2141[bis]. An updated namespace registration is included, which describes how both the old and the new ISBN format can share the same namespace.

This document replaces RFC 3187; it also obsoletes and moves to Historic status the predecessor thereof, RFC 2288.

Discussion

Comments are welcome and should be directed to the urn@ietf.org mailing list or the authors.

Status of This Memo

This Internet-Draft will expire on June 20, 2011.

Copyright Notice

Table of Contents

1. Introduction

One of the basic permanent URI schemes (cf. RFC 3986 [RFC3986], [IANA-URI]) is 'URN' (Uniform Resource Name) as originally defined in RFC 2141 [RFC2141] and now being formally specified in RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn]. Any identifier, when used within the URN system, needs its own namespace. As of this writing, there are roughly 40 registered URN namespaces (see [IANA-URN]), one of which belongs to ISBN, International Standard Book Number, as specified 2001 in RFC 3187 [RFC3187].

Since 2007, there have been two variants of ISBN in use; an outdated one based on ISO 2108-1992 [ISO1] and a new one defined in ISO 2108-2005 [ISO2]. These versions shall subsequently be called "ISBN-10" and "ISBN-13", respectively, in this document. For the time being, both ISBNs may still be printed on a book, but the ISBN-13 is the actual identifier. If what is said in this document applies to both ISBN versions, the generic term "ISBN" is used.

As part of the validation process for the development of URNs, the IETF URN working group agreed that it is important to demonstrate that a URN syntax proposal can accommodate existing identifiers from well established namespaces. One such infrastructure for assigning and managing names comes from the bibliographic community. Bibliographic identifiers function as names for objects that exist both in print and, increasingly, in electronic formats. RFC 2288 [RFC2288] investigated the feasibility of using three identifiers (ISBN, ISSN and SICI, see below) as URNs, with positive results; however, it did not formally register corresponding URN namespaces. This was in part due to the still evolving process to formalize criteria for namespace definition documents and registration, consolidated later in the IETF into RFC 3406 [RFC3406]. That RFC, in turn, is now being updated as well into RFC 3406bis [I-D.ietf-urnbis-rfc3406bis-urn-ns-reg].

URN Namespaces have subsequently been registered for both ISBN (International Standard Book Number) and ISSN (International Serial Standard Number) in RFCs 3187 [RFC3187] and 3044 [RFC3044], respectively, but not for SICI (Serial Item and Contribution Identifier), due to both the identifier's limited popularity and its complicated URN resolution process.

Guidelines for using ISBN-10s (based on ISO 2108-1992) as URNs and the original namespace registration have been published in RFC 3187 [RFC3187]. The RFC at hand replaces RFC 3187; sections related to ISBN-13 have been added, all ISBN-10 information has been updated and the namespace registration revised to make it compliant with both ISBN versions and stipulations of RFC 3406bis

[I-D.ietf-urnbis-rfc3406bis-urn-ns-reg], the work-in-progress successor of RFC 3406 [RFC3406], which in turn had replaced the legacy RFC 2611 [RFC2611] applied in the initial registration.

2. Conventions used in this document

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 RFC 2119 [RFC2119].

"ISBN-10" refers to the original, 10-digit ISBN scheme specified in ISO 2108-1992 [ISO1].

"ISBN-13" refers to the current, 13-digit ISBN scheme specified in ISO 2108-2005 [ISO2].

3. Fundamental Namespace and Community Considerations

3.1. The URN:ISBN Namespace

ISBN is a well established standard identifier system for monographic publications. Therefore, any useful and deployable method to identify these entities for network-wide reference and make their metadata available on the Internet needs to be based on that system.

Section 4 below, and there in particular Section 4.1, presents a detailed overview of the history and the structure of the ISBN namespace, related institutions, and the identifier assignment principles used, and Section 4.3 gives an overview of existing and emerging resolution systems for the URN:ISBN namespace.

3.2. Community Considerations for ISBNs

ISBNs are assigned under the auspices of the International ISBN Agency [ISBNORG].

ISBNs identify finite objects, but sometimes these objects might be so large that resolution into a hierarchical system is appropriate. The materials identified by an ISBN may exist only in printed or other physical form, not electronically. In such a case, the URN:ISBN resolver should nevertheless be able to supply bibliographic data, possibly including information about where the physical resource is stored in the owning institution's holdings. There may be other resolution services supplying a wide variety of information resources or services related to the identified books.

National libraries and large publishers are the key organizations providing persistent URN resolution services for resources identified with ISBNs, independent of their form.

For library users and Internet-based supply chain management for the delivery of monographic work, URN-based identification and resolution services offer more efficient, uniform, and reliable access to resources in general. No special tools are needed for this; Web browsers are sufficient.

Section 4 below, and in particular Section 4.3 therein, presents a detailed overview of the application of the URN:ISBN namespace and the principles, and systems used, for the resolution of ISBN-based URNs.

4. International Standard Book Numbers

4.1. Overview / Namespace Considerations

An International Standard Book Number (ISBN) identifies a product form or edition of a monographic publication.

4.1.1. ISBN-10 Structure

The ISBN-10 is defined by the ISO Standard 2108-1992 [ISO1]. It is a ten-digit number (the last "digit"can be the letter "X" as well) that is divided into four variable length parts usually separated by hyphens when printed. The parts are as follows (in this order):

- a group identifier that specifies a group of publishers, based on national, geographic, or some other criteria;
- the publisher identifier;
- the title identifier; and
- a modulo 11 check digit, using X instead of 10; the details of the calculation are specified in [ISO1].

ISBN-10 was in use from 1970s until ISBN-13 replaced it in January 2007.

4.1.2. ISBN-13 Structure

ISBN-13 is defined by the ISO Standard 2108-2005 [ISO2]. The ISBN-13 is a thirteen-digit number that is divided into five parts usually separated by hyphens when printed. The first and the last part have a fixed lenght, but the other parts have variable length. These parts are as follows (in this order):

- a prefix element of ISBN-13 is a 3 digit prefix specified by the International ISBN Agency; at the time of this writing, legal values were 978 and 979; future versions of the standard may define additional values;

- a registration group element that specifies the registration group; it identifies the national, geographic, language, or other such grouping within which one or more ISBN Agencies operate;

- the registrant element;

- the publication element; and

- a modulo 10 check digit; the details of the calculation are specified in [ISO2].

4.1.3. Relation between ISBN-10 and ISBN-13

The structural differences between the ISBN-10 and ISBN-13 are the prefix element (which did not exist in the old ISBN) and the check digit calculation algorithm, which was modulo 11 in ISBN-10 and is now modulo 10.

Terminology in ISBN-10 differs substantially from the terminology applied in ISBN-13. In this document, ISBN-13 terminology shall be used from now on; for a reader used to ISBN-10 terminology, the following mapping may be useful:

- ISBN-10 group identifier <-> ISBN-13 registration group element;

- ISBN-10 publisher identifier <-> ISBN-13 registrant element;

- ISBN-10 title identifier <-> ISBN-13 publication element.

Any ISBN-10 can be converted to ISBN-13 form, and retrospective conversion is indeed a recommended practice in ISO 2108-2005. Any application that processes ISBN-based URNs should however be prepared to deal with both ISBNs, since ISBN-10 numbers may not be converted to the new form. ISBN-13s using prefix element 979 can not be converted back to ISBN-10, since in these ISBNs group identifiers will be re-assigned. New books may still have ISBN-10 alongside ISBN-13 for practical reasons, but only as long as the prefix element in ISBN-13 is 978.

4.2. Encoding Considerations

Embedding ISBNs within the URN framework does not present encoding problems, since all of the characters that can appear in an ISBN are

valid in the namespace-specific string (NSS) part of the URN. %-encoding, as described in RFC 2141 [RFC2141] and RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn], is never needed.

Example 1: URN:ISBN:978-0-395-36341-6

Example 2: URN:ISBN:951-0-18435-7

Example 3: URN:ISBN:951-20-6541-X

4.3. Resolution of ISBN-based URNs

4.3.1. General

For URN resolution purposes, all elements except the check digit (now 0-9, previously 0-9 or X) must be taken into account. The registration group and registrant element assignments are managed in such a way that the hyphens are not needed to parse the ISBN unambiguously into its constituent parts. However, the ISBN is normally transmitted and displayed with hyphens to make it easy for humans to recognize these elements without having to make reference to or have knowledge of the number assignments for registration group and registrant elements. In ISBN-10, registration group element codes such as 91 for Sweden were unique. In ISBN-13, only the combination of prefix and registration group elements is guaranteed to be unique. 978-951 and 978-952 both mean Finland, but 979-951 and 979-952 almost certainly will not (once they will be assigned in the future); at the time of this writing, registration group element(s) for Finland are not yet known for ISBNs starting with 979.

The Finnish URN registry is maintained by the national library. The service is capable of resolving ISBN-based URNs. URNs starting with URN:ISBN:978-951 or URN:ISBN:978-952 are mapped into appropriate URL addresses in a table maintained within the registry. Applications, such as the national bibliography or the open archive of a university, can use the URN as the address of the resource. There is just one place (the registry) where the location information must be kept up to date.

ISBN-13 prefix / registration group element combinations (and the corresponding ISBN-10 registration group identifiers, if any) usually designate a country, but occasionally a single combination / ISBN-10 group identifier is used to indicate a language area. For instance, "978-3" (or "3" in ISBN-10) is utilised in Germany, Austria, and the German speaking parts of Switzerland. As of this writing, there are two regional registration groups: "978-976" is used in the Caribbean community and "978-982" in the South Pacific (see [PREFIX]).

Note that the prefix and registration group element combination "979-3" has not yet been assigned. There is no intention to allocate the registration group elements in the same way as was done with ISBN-10.

The registrant element may or may not be used for resolution purposes, depending on whether individual publishers have set up their resolution services.

The publication element shall enable targeting the individual publication.

4.3.2. Practical Aspects

Due to the lack of URN support in, e.g., web browsers, the URNs are usually expressed as URLs when embedded in documents. The Finnish URN registry is located at <http://urn.fi/>, and URNs are therefore expressed in the form http://urn.fi/<URN>. For example, the URI <http://urn.fi/URN:ISBN:978-952-10-3937-9> identifies Sami Nurmi's doctoral dissertation "Aspects of Inflationary Models at Low Energy Scales".

The Finnish URN registry can not resolve URN:ISBNs with non-Finnish registration group element values until other countries establish their registries, and all these services become aware of each other and their respective registration group responsibility domains and are able to communicate with each other. Thus the Finnish registry can deal with URN:ISBN instances with registration group element value 91 (indicating Sweden) if and only if the Swedish registry exists, its address is known to the Finnish peer and the Swedish service is capable of receiving and processing requests from other registries.

If a registration group element does not identify a single country but a language area, there are at least two means for locating the correct national bibliography. First, it is possible to define a cascade of URN registries - for instance, German, Austrian and Swiss national registries, in this order - which should collectively be aware of resolution services such as national bibliographies for ISBN-13s starting with "978-3". If the German registry is not able to find an authoritative resolution service, the request could be passed to the Austrian one, and if there are still no hits, finally to the Swiss service.

Second, the registrant element ranges assigned to the publishers in Germany, Austria and Switzerland by the ISBN Agencies could be defined directly into the national registries. This method would be more efficient than cascading, since the correct resolution service

would be known immediately. The choice between these two and possible other options should be made when the establishment of the European network of URN registries reaches this level of maturity.

In some exceptional cases -- notably in the US and in the UK, where international companies do a significant portion of publishing -- the information provided by the group identifier may not always be fully reliable. For instance, some monographs published in New York by international publishing companies may get an ISBN with the registration group element "3". This is technically appropriate when the headquarters or one of the offices of the publisher is located in Germany.

Information about such a book may not always be available in the German national bibliography, but via the Library of Congress systems. Unfortunately, the German/Austrian/Swiss URN registries that should in this case be contacted may not be aware of the appropriate resolution service.

However, the problem posed by the international publishers may well be less severe than it looks. Some international publishers (Springer, for example) give the whole production to the national library of their home country as legal deposit, no matter which country the book was published. Thus everything published by Springer in New York with registration group element "3" should be resolvable via the German national bibliography. On the other hand, when these companies give their home base also as a place of publication, the "home" national library requires the legal deposit.

A large union catalogue, such as WorldCat maintained by OCLC [OCLC-WC] could be used to complement the resolution services provided in the national level, or as the default service, if no national services exist or are known to the registry from which the query originates.

Due to the semantic structure of ISBN-13, even the registrant element can be used as a "hint". Technically, it is possible to establish a number of URN resolution services maintained by different kinds of organizations. For instance, "978-951-0" is the unique ISBN registrant element of the largest publisher in Finland, Sanoma-WSOY. Resolution requests for ISBNs starting with "978-951-0" can be passed to and dealt with the publisher's server, if and when it is made URN-aware. In such a case, resolving the same URN in multiple locations may provide different services; the national bibliography may be able to provide bibliographic information only, while the publisher can also provide the book itself, on its own terms. Different resolution services may co-exist and complement one another. Same ISBN may be resolved both as URN and as a Digital Object Identifier (DOI)

[DOIHOME]. URN-based services hosted by, e.g., a national library, might provide only bibliographic data, whereas a service based on the DOI system provided by the publisher might deliver the book, parts of the book or various services related to the work.

Persistence of resolution services is largely dependent on persistence of organizations providing them. Thus some services, independent on base technology chosen, may disappear or their content may change much sooner than some peer solutions.

4.4. Additional Considerations

The basic guidelines for assigning ISBNs to electronic resources are the following:

- Format/means of delivery is irrelevant to the decision whether a product needs an ISBN or not. If the content meets the requirement, it gets an ISBN, no matter what the format of the delivery system.

- Each format of a digital publication should have a separate ISBN. The definition of a new edition is normally based on one of the two criteria:

  * A change in the kind of packaging involved: the hard cover edition, the paperback edition and the library-binding edition would each get a separate ISBN. The same applies to different formats of digital files.

  * A change in the text, excluding packaging or minor changes such as correcting a spelling error. Again, this criterion applies regardless of whether the publication is in printed or in digital form.

Although these rules seem clear, their interpretation may vary. As already RFC 2288 [RFC2288] pointed out,

> The choice of whether to assign a new ISBN or to reuse an existing one when publishing a revised printing of an existing edition of a work or even a revised edition of a work is somewhat subjective. Practice varies from publisher to publisher (indeed, the distinction between a revised printing and a new edition is itself somewhat subjective). The use of ISBNs within the URN framework simply reflects these existing practices. Note that it is likely that an ISBN URN may resolve to many instances of the work (many URLs).

These instances may be fully identical, or there may be some minor differences between them. Publishers have also in some occasions re-used the same ISBN for another book. This reasonably rare kind of human error does not threaten or undermine the value of the ISBN system as a whole. Neither do they pose a serious threat to the URN resolution service based on ISBNs. An error should only lead into the retrieval of two or more bibliographic records describing two different monographic publications. Based on the information in the records, a user can choose the correct record from the result set.

Most national bibliographies and especially the Books in Print correct ISBN mistakes. The systems then provide cross references "incorrect ISBN -> correct ISBN"). This should be taken into account in the URN resolution process. Further details on the process of assigning ISBNs can be found in section 5 (Namespace registration) below.

5. URN Namespace Registration and Use

The formal URN Namespace Identifier Registration for the pre-2005 version of the International Standard Book Number (ISBN) was done in RFC 3187 [RFC3187].

The new ISBN standard does not require a new namespace, but the registration is renewed here, as the registrant organization has moved from Staatsbibliothek zu Berlin - Preussischer Kulturbesitz to The International ISBN Agency, London, U.K, and the syntax and resolution details are amended.

5.1. URN Namespace ID Registration for the International Standard Book Number (ISBN)

This registration describes how International Standard Book Numbers (ISBN) can be supported within the URN framework.

[RFC Editor: please replace "XXXX" in all instances of "RFC XXXX" below by the RFC number assigned to this document.]

Namespace ID: ISBN

This Namespace ID has already been assigned to the International Standard Book Number in January 2001 when the namespace was registered for the first time.

Registration Information:

Version: 2
Date: 2010-12-17

Declared registrant of the namespace:

Registering Organization: The International ISBN Agency

Designated Contact Person:
Name: Mr. Brian Green
Affiliation: Director, The International ISBN Agency
Email: brian@isbn-international.org
Postal: EDItEUR, 39-41 North Road, London, N7 9DP, U.K.
Web URL: <http://www.isbn-international.org/>

Declaration of syntactic structure of NSS part:

The namspace-specific string of 'ISBN' URNs is either an ISBN-13 (see Section 4.1.2 of RFC XXXX) or an ISBN-10 (see Section 4.1.1 of RFC XXXX); the former is preferred.

Example 1: URN:ISBN:978-0-395-36341-6
Example 2: URN:ISBN:951-0-18435-7
Example 3: URN:ISBN:951-20-6541-X

Relevant ancillary documentation:

The ISBN (International Standard Book Number) is a unique machine-readable identification number, which marks any edition of a book unambiguously. This number is defined in ISO Standard 2108. The number has been in use now for 30 years and has revolutionised the international book-trade. 170 countries and territories are officially ISBN members, and more of them are joining the system.

The administration of the ISBN system is carried out on three levels:
International agency,
Group agencies,
Publisher levels.

The International ISBN agency is located in London. The main functions of the Agency are:

* To promote, co-ordinate and supervise the world-wide use of the ISBN system.

* To approve the definition and structure of group agencies.

* To allocate group identifiers to group agencies.

* To advise on the establishment and functioning of group agencies.

* To advise group agencies on the allocation of international publisher identifiers.

* To publish the assigned group numbers and publisher prefixes in up-to-date form.

Information about ISBN usage in general can be found from the ISBN FAQ, available at <http://www.isbn-international.org/en/faqs>.

Conformance with URN Syntax:

Legal ISBN characters are 0-9 and hyphen for ISBN-13 and 0-9, hyphen, and X for ISBN-10. No percent-encoding is needed.

[[Editorial Note: Need to discuss new specification requirements from the RFC 2141bis draft!]]

Rules for Lexical Equivalence of NSS part:

ISBN numbers are usually printed with the letters 'ISBN' and a single blank preceding them (for instance: ISBN 951-746-795-8). The data preceding the actual number must be removed before the ISBNs are analysed. The ISBN serves directly as the namespace-specific string (NSS) of 'ISBN' URNs.

Prior to comparing the NSS of two ISBN-based URNs for equivalence, all hyphens MUST be removed and letter 'X' capitalized. Prior to comparing a URN based on ISBN-10 with a URN based on ISBN-13, the ISBN-10 MUST be converted to the ISBN-13 form. This step is necessary since the ISBN-10s may or may not be already converted to the new form; libraries SHOULD keep the old ISBN since it is the one printed in books published prior to 2007, while publishers may convert the old identifiers originally assigned in ISBN-10 form and use the equivalent ISBN-13s in unchanged reprints of the books, which according to the ISBN assignment rules should not receive a new ISBN.

Note that, according to RFC 2141bis, the prefix "URN:ISBN:" is case-insensitive; generic URI parsing and comparison software frequently uses lower case as the canonical (normalized) form.

The URNs are equivalent if the normalized forms obtained this way compare equal.

Identifier uniqueness and persistence considerations:

ISBN is a unique and persistent identifier. An ISBN, once it has been assigned, must never be re-used for another book. Moreover,

a single manifestation of a book must never get a new ISBN. 'ISBN' URNs inherit the uniqueness and persistence properties from the underlying ISBN namespace.

There may be multiple manifestations of a single literary work such as a novel. In such case each manifestation shall receive a different ISBN. ISO has developed a new standard, ISTC (International Standard Text Code, ISO 21047-2009) that enables identification of textual works. See <http://www.istc-international.org/> for more information. In the standard itself, annex E describes the relations between ISBN and other publication identifiers and ISTC.

Process of identifier assignment:

Assignment of ISBNs is controlled, and 'ISBN' URNs immediately inherit this property. There are three levels of control: the international agency, group agencies that typically operate in the national level, and finally each publisher is responsible of using the ISBN system correctly. Small publishers may demand ISBN numbers one at a time by contacting the ISBN group agency. Large publishers receive ISBN blocks from which they allocate ISBNs to the books according to the ISBN assignment rules.

Process for identifier resolution:

See Section 4.3 of RFC XXXX.

Validation mechanism:

The check digit helps to assure the correctness of an ISBN number assigned for a book when it has been entered or processed by a human. Applications processing bibliographic data such as integrated library systems typically can check the correctness of both ISBN-10 and ISBN-13 (and make conversions between the two). If the number is wrong due to, e.g., a typing error made by a publisher, a correct ISBN is usually assigned afterwards. Although the book will only contain the wrong number, national bibliography and system used by the book trade often will contain both the wrong and new, correct ISBN number.

Scope:

ISBN is a global identifier system used for identification of monographic publications. It is very widely used and supported by the publishing industry.

6. Security Considerations

This document proposes means of encoding ISBNs within the URN framework. An ISBN-based URN resolution service is depicted here both for ISBN-10 and ISBN-13, but only in a fairly generic level; thus questions of secure or authenticated resolution mechanisms are excluded. It does not deal with means of validating the integrity or authenticating the source or provenance of URNs that contain ISBNs. Issues regarding intellectual property rights associated with objects identified by the ISBNs are also beyond the scope of this document, as are questions about rights to the databases that might be used to construct resolvers.

7. IANA Considerations

IANA is asked to update the existing registration of the Formal URN Namespace 'ISBN' using the template given above in Section 5.1, which follows the outline specified in RFC 3406bis [I-D.ietf-urnbis-rfc3406bis-urn-ns-reg].

8. Acknowledgements

This draft version is the outcome of work started in 2008 and brought to the IETF in 2010 to launch a much larger effort to revise the basic URN RFCs as a part of project PersID (http://www.persid.org). PersID is developing tools for establishing an European network of URN resolvers concentrating on bibliographic identifiers. The aim in the IETF is to bring these RFCs in alignment with the current URI Standard (STD 63, RFC 3986), ABNF, and IANA guidelines. The discussion in PersID has contributed significantly to this work.

Leslie Daigle has provided valuable guidance in the initial draft stage of this memo.

Your name could go here ...

9. References

9.1. Normative References

[I-D.ietf-urnbis-rfc2141bis-urn]
Hoenes, A., "Uniform Resource Name (URN) Syntax", draft-ietf-urnbis-rfc2141bis-urn-00 (work in progress), November 2010.

[I-D.ietf-urnbis-rfc3406bis-urn-ns-reg]
Hoenes, A., "Uniform Resource Name (URN) Namespace Definition Mechanisms",

draft-ietf-urnbis-rfc3406bis-urn-ns-reg-00 (work in progress), December 2010.

[ISO1] ISO, "Information and documentation - The International Standard Book Number (ISBN)", ISO 2108-1992, 1992.

[ISO2] ISO, "Information and documentation - The International Standard Book Number (ISBN)", ISO 2108-2005, 2005.

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

9.2. Informative References

[DOIHOME] International DOI Foundation, "The Digital Object Identifier System", <http://www.doi.org/>.

[IANA-URI] IANA, "URI Schemes Registry", <http://www.iana.org/assignments/uri-schemes/>.

[IANA-URN] IANA, "URN Namespace Registry", <http://www.iana.org/assignments/urn-namespaces/>.

[ISBNORG] International ISBN Agency, "", <http://www.isbn-international.org/>.

[OCLC-WC] OCLC WorldCat, "WorldCat.org: The World's Largest Library Catalog", <http://www.worldcat.org>.

[PREFIX] International ISBN Agency, "ISBN Prefix Ranges", <http://www.isbn-international.org/page/ranges>.

[RFC2141] Moats, R., "URN Syntax", RFC 2141, May 1997.

[RFC2288] Lynch, C., Preston, C., and R. Jr, "Using Existing Bibliographic Identifiers as Uniform Resource Names", RFC 2288, February 1998.

[RFC2611] Daigle, L., van Gulik, D., Iannella, R., and P. Faltstrom, "URN Namespace Definition Mechanisms", BCP 33, RFC 2611, June 1999.

[RFC3044] Rozenfeld, S., "Using The ISSN (International Serial Standard Number) as URN (Uniform Resource Names) within an ISSN-URN Namespace", RFC 3044, January 2001.

[RFC3187] Hakala, J. and H. Walravens, "Using International Standard Book Numbers as Uniform Resource Names", RFC 3187, October 2001.

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

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

Appendix A. Draft Change Log

[[RFC-Editor: Whole section to be deleted before RFC publication.]]

A.1. draft-hakala-rfc3187bis-isbn-urn-00 to draft-ietf-urnbis-*-00

- formal updates for a WG draft;
- RFC 2288 now obsoleted and made Historic;
- added references to rfc2141bis and rfc3406bis;
- Sect.3 reorganized and amended: Namespace/Community Considerations;
- registrartion template adapted to rfc3406bis [-00];
- numerous editorial fixes and improvements.

Authors' Addresses

Maarit Huttunen
The National Library of Finland
P.O. Box 26
Helsinki, Helsinki University FIN-00014
Finland

EMail: maarit.huttunen@helsinki.fi

Juha Hakala
The National Library of Finland
P.O. Box 15
Helsinki, Helsinki University FIN-00014
Finland

EMail: juha.hakala@helsinki.fi

Alfred Hoenes (editor)
TR-Sys
Gerlinger Str. 12
Ditzingen D-71254
Germany

EMail: ah@TR-Sys.de

IETF URNbis WG
Internet-Draft
Obsoletes: 3406 (if approved)
Intended status: BCP
Expires: June 20, 2011

A. Hoenes
TR-Sys

December 17, 2010

Uniform Resource Name (URN) Namespace Definition Mechanisms
draft-ietf-urnbis-rfc3406bis-urn-ns-reg-00

Abstract

Uniform Resource Names (URNs) are intended to serve as persistent, location-independent, resource identifiers. To structure and organize their usage, the URN syntax specifies a hierarchy that horizontally divides the set of possible URNs into "URN Namespaces" that can be individually defined and managed. URN Namespaces in particular serve to map existing identifier systems into the URN system and thereby make available generic, network-based resolution services for the identified documents, artifacts, and other objects (and their metadata).

To actually leverage such synergetic advantage, URN namespaces need to be specified in a comparable manner, and their Namespace Identifiers (NIDs) need to be registered with IANA, so that naming conflicts are avoided and implementers of services can follow a structured approach in support of various namespaces, guided by the registry to the related documents and the particularities of specific namespaces, as described in these namespace registration documents.

This document serves as a guidleline for authors of URN Namespace definition and registration documents. It describes the essential content of such documents and how they shall be structured to allow readers familar with the scheme to quickly assess the properties of a specific URN Namespace. Further, this RFC describes the process to be followed to get a URN Namespace registered with IANA.

This document is a companion document to the revised URN Syntax specification, RFC 2141bis; it supersedes and replaces RFC 3406.

Discussion

This draft version has been obtained by importing the text from RFC 3406 into modern tools and making a first round of updating steps. It is an initial chartered work item of the URNBIS WG.

Discussion of this memo utilizes the urn@ietf.org mailing list.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Uniform Resource Names (URNs) are resource identifiers with the specific requirements for enabling location-independent identification of a resource, as well as longevity of reference. URNs are part of the larger Uniform Resource Identifier (URI) family (see the joint W3C/IETF memorandum, RFC 3305 [RFC3305], and the IETF STD 66, RFC 3986 [RFC3986]) with the specific goal of providing persistent naming of resources.

There are two assumptions that are key to this document:

Assumption #1: Assignment of a URN is a managed process.

> I.e., not all strings that conform to URN syntax are necessarily valid URNs. A URN is assigned according to the rules of a particular namespace (in terms of syntax, semantics, and process).

Assumption #2: The space of URN namespaces is managed.

> I.e., not all syntactically correct URN namespaces (per the URN syntax definition) are valid URN namespaces. A URN namespace must have a recognized definition in order to be valid.

The purpose of this document is to outline a mechanism and provide a template for explicit namespace definition, as well as provide the mechanism for associating an identifier (called a "Namespace ID", or NID), which is registered with the Internet Assigned Numbers Authority (IANA) [IANA] in the URN Namespaces registry maintained at [IANA-URN].

The URN Namespace definition and registration mechanisms originally have been specified in RFC 2611 [RFC2611], which has been obsoleted by BCP 66, RFC 3406 [RFC3406]. Guidelines for documents prescribing IANA procedures have been revised as well over the years, and at the time of this writing, BCP 26, RFC 5226 [RFC5226] is the normative document. This document is a revision of RFC 3406 based on the revised URN Syntax specification RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn] and RFC 5226.

The reader is referred to Section 1.1 of RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn] for a more detailed synopsis of the history of documents fundamental for URNs.

Note that this document restricts itself to the description of processes for the creation of URN namespaces. If "resolution" of any so-created URN identifiers is desired, a separate process of registration in a global NID directory, such as that provided by the

DDDS system [RFC3401], is necessary. See [RFC3405] for information on obtaining registration in the DDDS global NID directory.

1.1. Requirement Language

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 RFC 2119 [RFC2119]. In this document, these key words describe requirements for the process to be followed and the content to be provided in namespace definition documents and registration templates.

2. What is a URN Namespace?

For the purposes of URNs, a "namespace" is a collection of uniquely-assigned identifiers. That is, the identifiers are not ever assigned to more than 1 resource, nor are they ever re-assigned to a different resource. A single resource, however, may have more than one URN assigned to it for different purposes. A URN namespace itself has an identifier in order to:

- ensure global uniqueness of URNs,

- (where desired) provide a cue for the structure of the identifier.

For example, many identifier systems use strings of numbers as identifiers (e.g., ISBN, ISSN, phone numbers). It is conceivable that there might be some numbers that are valid identifiers in two different established identifier systems. Using different designators for the two collections ensures that no two URNs will be the same for different resources (since each collection is required to uniquely assign each identifier).

The development of an identifier structure, and thereby a collection of identifiers, is a process that is inherently dependent on the requirements of the community defining the identifier, how they will be assigned, and the uses to which they will be put. All of these issues are specific to the individual community seeking to define a namespace (e.g., publishing community, association of booksellers, protocol developers, etc.); they are beyond the scope of the IETF URN work.

This document outlines the processes by which a collection of identifiers satisfying certain constraints (uniqueness of assignment, etc.) can become a bona fide URN namespace by obtaining a NID. In a nutshell, a template for the definition of the namespace is completed for deposit with IANA, and a NID is assigned. The details of the process and possibilities for NID strings are outlined below.

3. URN Namespace (Registration) Types

There are three categories of URN namespaces defined here, distinguished by expected level of service and required procedures for registration. Registration processes for each of these namespace types are given in Section 4.

3.1. Experimental Namespaces

These are not explicitly registered with IANA. They take the form:

X-<NID>

No provision is made for avoiding collision of experimental NIDs; they are intended for use within internal or limited experimental contexts.

[[Editorial Note:
Has anybody ever seen usage of such experimental URN Namespaces? According to the observations of the author, three years of RFC 2611 and eight years of RFC 3406 have constantly seen "tentative grabbing" and subsequent usage of NIDs that the stakeholders later have tried to register with IANA as Formal NIDs (with varying success).
So should this kind of namespaces better be dropped and a kind of provisional NIDs be created? -- This would be in the spirit of BCP 100, RFC 4020 [RFC4020], and it would resemble the manner how URI Scheme registrations are dealt with (RFC 4395 [RFC4395], [IANA-URI]).
]]

3.2. Informal Namespaces

These are fully fledged URN namespaces, with all the rights and requirements associated thereto. Informal namespaces can be registered in global registration services. They are required to uphold the general principles of a well-managed URN namespace -- providing persistent identification of resources and unique assignment of identifier strings. Informal and formal namespaces (described below) differ in the NID assignment. IANA will assign an alphanumeric NID (following a defined pattern) to registered informal namespaces, per the process outlined in Section 4.

3.3. Formal Namespaces

A formal namespace may be requested, and IETF review sought, in cases where the publication of the NID proposal and the underlying namespace will provide benefit to some subset of users on the Internet. That is, a formal NID proposal, if accepted, must be functional on and with the global Internet, not limited to users in

communities or networks not connected to the Internet. For example, assume a NID is requested that is meant for naming of physics research. If that NID request required that the user use a proprietary network or service that was not at all open to the general Internet user, then it would make a poor request for a formal NID. The intent is that, while the community of those who may actively use the names assigned within that NID may be small (but no less important), the potential use of names within that NID is open to any user on the Internet.

It is expected that Formal NIDs may be applied to namespaces where some aspects are not fully open. For example, a namespace may make use of a fee-based, privately managed, or proprietary registry for assignment of URNs in the namespace, but it may still provide benefit to some Internet users if the services associated have openly-published access protocols.

In addition to the basic registration information defined in the registration template (in Appendix A), a formal namespace request must be accompanied by documented considerations of the need for a new namespace and of the community benefit from formally establishing the proposed URN namespace.

Additionally, since the goal of URNs is to provide persistent identification, some consideration as to the longevity and maintainability of the namespace must be given. The collective experience of the IETF community contains a wealth of information on technical factors that will prevent longevity of identification. Thus, the IESG may elect not to accept a proposed namespace registration if the IETF community consensus is that the registration document contains technical flaws that will prevent (or seriously impair the possibility of) persistent identification, and that it therefore should not be published as an RFC.

Consideration should be given to these aspects:

- the organization maintaining the URN namespace should demonstrate stability and the ability to maintain the URN namespace for a long time, and/or it should be clear how the namespace can continue to be usable/useful if the organization ceases to be able to foster it;

- it should demonstrate ability and competency in name assignment; this should improve the likelihood of persistence (e.g., to minimize the likelihood of conflicts);

- it should commit to not re-assigning existing names and allowing old names to continue to be valid, even if the owners or assignees

of those names are no longer members or customers of that organization; this does not mean that there must be resolution of such names, but that they must not resolve the name to false or stale information, and that they must not be reassigned.

These aspects, though hard to quantify objectively, should be considered by organizations/people considering the development of a Formal URN namespace, and they will be kept in mind when evaluating the technical merits of any proposed Formal URN namespace.

4. URN Namespace Registry: Processes for Registration and Update

Different levels of disclosure are expected/defined for namespaces. According to the level of open-forum discussion surrounding the disclosure, a URN namespace may be assigned an identifier or may request a particular identifier.

The IANA Considerations Guidelines document (BCP 26, RFC 5226 [RFC5226]) suggests the need to specify update mechanisms for registrations -- who is given the authority to do so, from time to time, and what are the processes. Since URNs are meant to be persistently useful, few (if any) changes should be made to the structural interpretation of URN strings (e.g., adding or removing rules for lexical equivalence that might affect the interpretation of URN IDs already assigned). However, it may be important to introduce clarifications, expand the list of authorized URN assigners, etc., over the natural course of a namespace's lifetime. Specific processes are outlined below.

The official list of registered URN namespaces is currently maintained by IANA at <http://www.iana.org/assignments/urn-namespaces/urn-namespaces.xhtml>.

[[NOTE: It would be preferable to restore the generic, most universally supported (HTML) form of the registry be identified by an implementation-neutral URL, as previously supported by IANA: <http://www.iana.org/assignments/urn-namespaces>. The content there should link to alternate forms (.xml, .txt), and those alternate versions should indicate the *other* versions; i.e., where currently the .txt version also says, "This registry is also available in XML and plain text formats.", it should better say: "This registry is also available in HTML and XML formats."]]

The registration is subdivided into two sub-registries, one for "Formal URN Namespaces" and one for "Informal URN Namespaces", and each entry there links to a stable repository of the registration document or (an escrow copy of) the filled-out registration template.

The registration and maintenance procedures vary slightly between the namespace types.

4.1. Experimental Namespaces: No Registration

The NIDs of Experimental Namespaces (Section 3.1) are not explicitly registered with IANA. They take the form:

X-<NID>

No provision is made for avoiding collision of experimental NIDs; they are intended for use within internal or limited experimental contexts exclusively.

As there is no registration, no registration/maintenance procedures are needed.

4.2. Informal Namespaces

The NIDs of Informal Namespaces are synthesized by IANA using an assigned sequence number and registered in their own sub-registry, as indicated in Section 4; they take the format:

"urn-" <number>

where <number> is the decimal representation of a natural number, with no leading zeroes. This sequence number is assigned by the IANA on a First-Come-First-Served [RFC5226] basis to registration requests for informal namespaces.

Registrants should send a copy of the registration template (as shown in Appendix A), duly completed, to the urn-nid@ietf.org mailing list for review and allow for a two-week discussion period for clarifying the expression of the registration information and suggestions for technical improvements to the namespace proposal. [NOTE: Longer time is needed in practice! Increase to 4 weeks?]

After suggestions for clarification of the registration information have been incorporated, the template may be submitted for assignment of a NID by email to iana@iana.org .

Registrations may be updated later by the original registrant, or by an entity designated by the registrant, by updating the registration template, submitting it to the discussion list for a further two-week discussion period, and finally resubmitting it to IANA in a message to iana@iana.org .

4.3. Formal Namespaces

Formal NIDs are assigned via IETF Review, as defined in BCP 26 [RFC5226]. The designated expert(s) for URN namespace registrations are nominated by the IESG, and their role adheres to the regulations in BCP 26, unless specified otherwise below.

This means that the Formal NID application is made via submission to the IETF of an Internet-Draft that contains the namespace definition and targets publication as an RFC of Informational or Standards Track category, which needs to be approved by the IESG after performing an IETF Last Call on the document and evaluating review comments. The applicant can be an individual or an IETF working group, in alignment with the designation of the Internet-Draft.

Before publication can be requested, however, the draft namespace specification document must undergo an Expert Review process [RFC5226] pursuant to the guidelines written here (as well as standard RFC publication guidelines). The template defined in Appendix A SHOULD be included as part of an RFC-to-be defining some other aspect(s) of the namespace, or it may be put forward as a namespace definition document in its own right. The proposed template (including a pointer to a readily available copy of the registration document) should be sent to the urn-nid@ietf.org mailing list for review. This list is monitored by the designated expert(s). The applicant has to allow for a two-week discussion period for clarifying the expression of the registration information, and SHOULD improve the namespace document and/or registration template based on the comments received, under the guidance of the designated expert(s), before the IESG reviews the document.

Working groups generally SHOULD seek early expert review for a namespace definition document, before they hand it over to the IESG, and individual applicants are also advised to seek expert comments early enough. The aforementioned list can be contacted for informal advice at any stage.

4.4. Registration Documents

The following subsections describe essential, MANDATORY parts of URN namespace registration documents, which will be focal in the expert Review process and IETF Review.

4.4.1. Namespace Considerations in Registration Documents

The namespace definition document MUST include a "Namespace Considerations" section that outlines the perceived need for a new namespace (i.e., where existing namespaces fall short of the proposer's requirements).

Considerations MUST include, directly or with the help of referenced stable (and preferably readily available) documents:

- URN assignment procedures;
- URN resolution/delegation;
- type of resources to be identified;
- type of services to be supported.

NOTE: It is expected that more than one namespace may serve the same "functional" purpose; the intent of the "Namespace Considerations" section is to provide a record of the proposer's "due diligence" in exploring existing possibilities, for the IESG's consideration.

[[Editorial Note: See the endnote of the next section!]]

4.4.2. Community Considerations in Registration Documents

The namespace definition document MUST also include a "Community Considerations" section that indicates the dimensions upon which the proposer expects its community to be able to benefit by publication of this namespace, as well as how a general Internet user will be able to use the space if they care to do so.

Potential considerations include:

- open assignment and use of identifiers within the namespace;
- open operation of resolution servers for the namespace (server);
- creation of software that can meaningfully resolve and access services for the namespace (client).

[[Editorial Note:
It is acknowledged that, in many cases, the Namespace Considerations and Community Considerations are closely intertwined. Further, the bulleted list above (from RFC 3406) seems to be more related to the items in the registration template entitled "Identifier uniqueness

considerations", "Identifier persistence considerations", "Process of identifier assignment", and "Process for identifier resolution" than to the primary objectives presented in the first paragraph above (also from RFC 3406).
In fact, namespace registration documents seen so far duplicate in the registration template material from the "Community Considerations" that addresses the above bullets.
Therefore: Should this specification now allow a combined section "Namespace and Community Considerations" that focuses on the (non-)utility of possible alternate namespace re-use and the *benefits* of an independent new namespace?
]]

4.4.3. Security Considerations in Registration Documents

According to the general procurements for RFCs, URN namespace definition documents must include a "Security Considerations" section (cf. BCP 72 [RFC3552]). That section has to identify the security considerations specific to the subject URN namespace. If the subject URN namespace is based on an underlying namespace, the registration can include substantive security considerations described in specifications related to that particular namespace by reference to these documents. For general security considerations regarding URN usage (and more generally, URI usage), for the sake of clarity and brevity, it should refer to the Security Considerations in STD 63 [RFC3986] and in the URN Syntax document [I-D.ietf-urnbis-rfc2141bis-urn].

4.4.4. IANA Considerations in Registration Documents

According to the general procurements for RFCs, URN namespace definitions documents must include an "IANA Considerations" section (cf. BCP 26 [RFC5226]). That section has to indicate that the document includes a URN Namespace registration that is to be entered into the IANA registry of Formal URN Namespaces.

Registration documents for formal URN namespaces will provide a particular, unique, desired NID string, and this will be assigned by the Standards/Protocol Action of the IESG that approves the publication of the registration document as an RFC. RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn] specifies that NID strings are ASCII strings that are interpreted in a case-insensitive manner, but the NID string SHALL be registered in the capitalization form preferred by the registrant. The proposed NID string MUST conform with the <nid> syntax rule in Section 2.1 of RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn] and it MUST adhere to the following additional constraints:

- not be an already-registered NID;
- not start with "X-" (see Section 4.1 above);
- not start with "urn-" (see Section 4.2 above);
- not start with "xy-", where xy is any combination of 2 ASCII letters (see NOTE below);
- be more than 2 letters long.

NOTE: All two-letter combinations as well as two-letter combinations followed by "-" and any sequence of valid NID characters are reserved for potential use as countrycode-based NIDs for eventual national registrations of URN namespaces. The definition and scoping of rules for allocation of responsibility for such namespaces is beyond the scope of this document.
Further, to avoid confusion, "urn" is not allowed as an NID string; IANA has permanently reserved this string to prohibit assignment.

Registrations may be revised by updating the RFC through standard IETF RFC update processes. In any case, a revised document, in the form of a new Internet-Draft, must be published, and the proposed updated template must be circulated on the urn-nid discussion list, allowing for a two-week review period before pursuing RFC publication of the new document.

5. Security Considerations

This document largely focuses on providing mechanisms for the declaration of public information. Nominally, these declarations should be of relatively low security profile, however there is always the danger of "spoofing" and providing mis-information. Information in these declarations should be taken as advisory.

6. IANA Considerations

This document outlines the processes for registering URN namespaces, and has implications for the IANA in terms of registries to be maintained, as previously defined in RFC 3406 [RFC3406]. This document replaces RFC 3406; it contains a revised description for the management of the "Uniform Resource Names (URN) Namespaces" IANA Registry that uses the policy designation terms from BCP 26, RFC 5226 [RFC5226], but does not introduce significant changes to the applicable procedures.

All references there to the predecessor, [RFC3406], should be replaced by references to this document.

Section 4.4.4 above describes the syntax rules for NIDs to which the registry needs to obey. As pointed out in Section 4.4.4 above and in RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn], the string "urn" is permanently reserved and MUST NOT be assigned as an NID.

In all cases of new namespace registration proposals, the IANA should provisionally assign the appropriate NID (informal or formal), as described throughout the body of this memo, once an IESG-designated expert has confirmed that the requisite registration process steps have been completed. These registrations become permanent and can be made publicly available once the registration document has been approved by the IESG for publications as a Standards Track or Informational RFC.

7. Acknowledgements

This document is heavily based on RFC 3406, the author of which are cordially acknowledged.

This document also been inspired by other recent documents that have updated important IANA registries, and the countless authors and contributors to these efforts are acknowledged anonymously.

Your name could go here ...

8. References

8.1. Normative References

[I-D.ietf-urnbis-rfc2141bis-urn]
Hoenes, A., "Uniform Resource Name (URN) Syntax", draft-ietf-urnbis-rfc2141bis-urn-00 (work in progress), November 2010.

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

[RFC3339] Klyne, G., Ed. and C. Newman, "Date and Time on the Internet: Timestamps", RFC 3339, July 2002.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

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

8.2. Informative References

[IANA] IANA, "The Internet Assigned Numbers Authority", <http://www.iana.org/>.

[IANA-URI] IANA, "URI Schemes Registry", <http://www.iana.org/assignments/uri-schemes/>.

[IANA-URN] IANA, "Uniform Resource Names (URN) Namespace Registry", <http://www.iana.org/assignments/urn-namespaces/>.

[RFC2276] Sollins, K., "Architectural Principles of Uniform Resource Name Resolution", RFC 2276, January 1998.

[RFC2611] Daigle, L., van Gulik, D., Iannella, R., and P. Faltstrom, "URN Namespace Definition Mechanisms", BCP 33, RFC 2611, June 1999.

[RFC3305] Mealling, M. and R. Denenberg, "Report from the Joint W3C/ IETF URI Planning Interest Group: Uniform Resource Identifiers (URIs), URLs, and Uniform Resource Names (URNs): Clarifications and Recommendations", RFC 3305, August 2002.

[RFC3401] Mealling, M., "Dynamic Delegation Discovery System (DDDS) Part One: The Comprehensive DDDS", RFC 3401, October 2002.

[RFC3405] Mealling, M., "Dynamic Delegation Discovery System (DDDS) Part Five: URI.ARPA Assignment Procedures", BCP 65, RFC 3405, October 2002.

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

[RFC3552] Rescorla, E. and B. Korver, "Guidelines for Writing RFC Text on Security Considerations", BCP 72, RFC 3552, July 2003.

[RFC4020] Kompella, K. and A. Zinin, "Early IANA Allocation of Standards Track Code Points", BCP 100, RFC 4020, February 2005.

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

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

Appendix A. URN Namespace Definition Template

Definition of a URN namespace is accomplished by completing the following information template.
Apart from providing a mechanism for disclosing the structure of the URN namespace, this information is designed to be useful for

- entities seeking to have a URN assigned in a namespace (if applicable) and

- entities seeking to provide URN resolvers for a namespace (if applicable).

This is particularly important for communities evaluating the possibility of using a portion of an existing URN namespace rather than creating their own.

Applications for Formal URN namespaces must also document "Namespace Considerations", "Community Considerations", "Security Considerations", and "IANA Considerations", as described in Section 4.4.

Information in the template is as follows (text in curly braces is tutorial and should be removed from filled-in templates):

Namespace ID:

{ If request is for an Informal NID, indicate so; the number will be assigned by IANA. In the case of a Formal NID registration, regularly a particular NID string will be requested. }

Registration Information:

{ This is information to identify the particular version of registration information: }
- version number:
  { starting with 1, incrementing by 1 with each new version }
- date:
  { date submitted to the IANA or date of approval of registration document, using the format outlined in "Date and Time on the Internet: Timestamps", [RFC3339]: YYYY-MM-DD }

Declared registrant of the namespace:

- Registering organization:
  Name: { ... }
  Address: { ... }
- Designated contact person:
  Name: { ... }
  { Address: ...
  (at least one of: Email, Phone, Postal address) }

Declaration of syntactic structure of NSS part:

[[Editorial Note: In the past, there has been iterated trouble in tentative registration documents with confusion between entire URN syntax and NSS syntax (only). Since the "urn:" prefix is fixed and the NID is fully determined by the "Namespace ID" clause above, in order to avoid error prone duplication, this version of the template tentatively restricts this clause to the NSS (namespace specific string) part of the new URNs.]]

{
This section should outline any structural features of identifiers in this namespace. At the very least, this description may be used to introduce terminology used in other sections. This structure may also be used for determining realistic caching/ shortcuts approaches; suitable caveats should be provided. If there are any specific character encoding rules (e.g., which character should always be used for single-quotes), these should be listed here.

Answers might include, but are not limited to:
- the structure is opaque (no exposition);
- a regular expression for parsing the identifier into components, including naming authorities;
- formal syntax of the NSS, preferably in ABNF (STD 68 [RFC5234]).
}

Relevant ancillary documentation:

{
This section should list any RFCs, standards, or other published documentation that defines or explains all or part of the namespace structure.

Answers might include, but are not limited to:
- RFCs that outline the syntax of the namespace;
- other documents of the defining community (e.g., ISO) that outline the syntax of the identifiers in the namespace;
- explanatory material that introduces the namespace.
}

Conformance with URN Syntax:

[[Editorial Note: This clause moved into vicinity of "syntax".]]

{
This section should outline any special considerations required for conforming with the URN syntax. This is particularly applicable in the case of legacy naming systems that are used in the context of URNs.

For example, if a namespace is used in contexts other than URNs, it may make use of characters that are reserved in the URN syntax.

This section should flag any such characters, and outline necessary mappings to conform to URN syntax. Normally, this will be handled by percent-encoding the symbol.
}

Rules for Lexical Equivalence of NSS part:

[[Editorial Note: This clause moved into vicinity of "syntax".]]

[[Editorial Note: In the past, there has been iterated trouble in tentative registration documents with regard to what rules can be imposed for lexical equivalence. Since the "urn:" prefix and the NID part both are invariably case-insensitive per RFC 3986 and RFC 2141[bis], in order to avoid repeated confusion, this version of the template tentatively restricts this clause to only the NSS part of the new URN namespace definition documents.]]

{
If there are particular algorithms for determining equivalence between two identifiers in the underlying namespace (and hence, in the URN string itself), rules can be provided here.

Some examples include:
- equivalence between hyphenated and non-hyphenated groupings in the identifier string;
- equivalence between single-quotes and double-quotes;
- namespace-defined equivalences between specific characters, such as "character X with or without diacritic marks".

Note that these are not normative statements for any kind of best practice for handling equivalences between characters; they are statements limited to reflecting the namespace's own rules.
}

Identifier uniqueness considerations:

{
This section should address the requirement that URN identifiers be assigned uniquely -- they are assigned to at most one resource, and are not reassigned.

(Note that the definition of "resource" is fairly broad; for example, information on "Today's Weather" might be considered a single resource, although the content is dynamic.)

Possible answers include, but are not limited to:
- exposition of the structure of the identifiers, and partitioning of the space of identifiers amongst assignment authorities that are individually responsible for respecting uniqueness rules;
- identifiers are assigned sequentially;
- information is withheld; that is, the namespace is opaque.
}

Identifier persistence considerations:

{
Although non-reassignment of URN identifiers ensures that a URN will persist in identifying a particular resource even after the "lifetime of the resource", some consideration should be given to the persistence of the usability of the URN. This is particularly important in the case of URN namespaces providing global resolution.

Possible answers include, but are not limited to:
- quality of service considerations.
}

Process of identifier assignment:

{
This section should detail the mechanisms and/or authorities for assigning URNs to resources. It should make clear whether assignment is completely open, or if limited, how to become an assigner of identifiers, and/or get one assigned by existing assignment authorities.

Answers could include, but are not limited to:
- assignment is completely open, following a particular algorithm;
- assignment is delegated to authorities recognized by a particular organization (e.g., the Digital Object Identifier Foundation controls the DOI assignment space and its delegation);
- assignment is completely closed (e.g., for a private organization).

}

Process for identifier resolution:

{
If a namespace is intended to be accessible for global resolution, it must be registered in an RDS (Resolution Discovery System, see RFC 2276 [RFC2276]) such as the DDDS (see RFC 3401 [RFC3401]). Resolution then proceeds according to standard URI resolution processes, and the mechanisms of the RDS. What this section should outline is the requirements for becoming a recognized resolver of URNs in this namespace (and being so listed in the RDS registry).

Answers may include, but are not limited to:
- the namespace is not listed with an RDS, this is not relevant;
- resolution mirroring is completely open, with a mechanism for updating an appropriate RDS;
- resolution is controlled by entities to which assignment has been delegated.

}

Validation mechanism:

{
Apart from attempting resolution of a URN, a URN namespace may provide mechanisms for "validating" a URN -- i.e., determining whether a given string is currently a validly-assigned URN. There are 2 issues here: 1) users should not "guess" URNs in a namespace; 2) when the URN namespace is based on an existing identifier system, it may not be the case that all the existing identifiers are assigned on Day 0. The reasonable expectation is that the resource associated with each resulting URN is somehow related to the thing identified by the original identifier system, but those resources may not exist for each original identifier. For example, even if a telephone number-based URN namespace was created, it is not clear that all telephone numbers would immediately become "valid" URNs, that could be resolved using

whatever mechanisms are described as part of the namespace registration.

Validation mechanisms might be:
- a syntax grammar;
- an on-line service;
- an off-line service.
}

Scope:

{
This section should outline the scope of the use of the identifiers in this namespace. Apart from considerations of private vs. public namespaces, this section is critical in evaluating the applicability of a requested NID. For example, a namespace claiming to deal with "social security numbers" should have a global scope and address all social security number structures (unlikely). On the other hand, at a national level, it is reasonable to propose a URN namespace for "this nation's social security numbers".
}

Appendix B. Illustration

B.1. Example Template

[[Editorial Note: Do we really need this any more? Such an almost-concrete example likely contradicts current IESG policy on usage of examples in RFCs.]]

The following example is provided for the purposes of illustrating the URN NID template described in Appendix A. Although it is based on a hypothetical "generic Internet namespace" that has been discussed informally within the URN WG, there are still technical and infrastructural issues that would have to be resolved before such a namespace could be properly and completely described.

Namespace ID:

To be assigned

Registration Information:

- version number: 1
- date: <when submitted>

Declared registrant of the namespace:

- Registering organization:
  Name: Thinking Cat Enterprises Name: Thinking Cat Example Enterprises
  Postal: 1 ThinkingCat Way
  Trupville, NewCountry
- Designated contact person:
  Name: L. Daigle
  Email: leslie@thinkingcat.example

Declaration of syntactic structure of NSS part:

The namespace specific string structure is as follows:

<FQDN>:<assigned string>

where FQDN is a fully-qualified domain name, and the assigned string is conformant to URN syntax requirements.

Relevant ancillary documentation:

Definition of domain names, found in:

P. Mockapetris, "DOMAIN NAMES - CONCEPTS AND FACILITIES", STD 13, RFC 1034, November 1987.

P. Mockapetris, "DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION", STD 13, RFC 1035, November 1987.

Conformance with URN Syntax:

No special considerations.

Rules for Lexical Equivalence of NSS part:

FQDNs are case-insensitive. Thus, the leading portion of the URN up to the colon after the FQDN is case-insensitive for matches. The remainder of the identifier must be considered case-sensitive.

Identifier uniqueness considerations:

Uniqueness is guaranteed as long as the assigned string is never reassigned for a given FQDN, and that the FQDN is never reassigned.

N.B.: operationally, there is nothing that prevents a domain name from being reassigned; indeed, it is not an uncommon occurrence. This is one of the reasons that this example makes a poor URN namespace in practice, and is therefore not seriously being proposed as it stands.

Identifier persistence considerations:

Persistence of identifiers is dependent upon suitable delegation of resolution at the level of "FQDN"s, and persistence of FQDN assignment.

Same note as above.

Process of identifier assignment:

Assignment of these URNs is delegated to individual domain name holders (for FQDNs). The holder of the FQDN registration is required to maintain an entry (or delegate it) in the DDDS. Within each of these delegated name partitions, the string may be assigned per local requirements.

E.g., urn:urn-<assigned number>:thinkingcat.example:001203

Process for identifier resolution:

Domain name holders are responsible for operating or delegating resolution servers for the FQDN in which they have assigned URNs.

Validation mechanism:

None specified.

Scope:

Global.

B.2. Registration steps in practice

The key steps for registration of informal or formal namespaces typically play out as follows:

A) Informal NID:

1. Complete the registration template. This may be done as part of an Internet-Draft.

2. Communicate the registration template to urn-nid@ietf.org for technical review -- as an email with a pointer to the submitted I-D or inline text containing the template.

3. Update the registration template (and/or document) as necessary from comments, and repeat steps 2 and 3 as necessary.

4. Once comments have been addressed (and the review period has expired), send a request to IANA with the revised registration template.

B) Formal NID:

1. Write an Internet-Draft describing the namespace and include the registration template, duly completed. Be sure to include "Namespace Considerations", "Community Considerations", "Security Considerations", and "IANA Considerations" sections, as described in Section 4.4.

2. Submit the Internet-Draft, and send a pointer to the I-D (perhaps using a copy of the I-D announcement) to urn-nid@ietf.org in order to solicit technical review.

3. Update the Internet-Draft as necessary from comments, and repeat steps 2 and 3 as needed.

4. If the Internet-Draft is the product of a working group in the IETF, follow the usual WG process to forward the document to the IESG for publication as an RFC. Otherwise, find a sponsoring Area Director willing to guide the draft through the IESG. The IESG (or the IETF at large in case an IETF-wide last call is deemed necessary) may request further changes (submitted as I-D revisions) and/or direct discussion to designated working groups, area experts, etc.

5. The IESG evaluation process includes a review by IANA, and if the IESG approves the document for publication as an RFC, IANA processing of the document will follow the regular work-flow between the RFC Editor and IANA. This way, the NID registration will be made public by IANA when the RFC is published.

Appendix C. Changes from RFC 3406

C.1. Essential Changes since RFC 3406

[RFC Editor: please remove the Appendix C.1 headline and all subsequent subsections of Appendix C starting with Appendix C.2.]

T.B.D. (after consolidation of this memo)

C.2. Changes from RFC 3406 to URNbis WG Draft -00

- Abstract: rewritten entirely;
- Section 1 (Introduction): added historical RFC information;
- Section 1.1 (Requirements Language): added;
- Section 3.1: added Note that challenges the utility of Experimental namespaces and raises question of whether formal "provisional" registrations would be useful;
- Section 4: text expanded and updated; background material added; added Note to challenge IANA website practices;
- Section 4.2 ff: changed "home" of URN-NID registration discussion list (it already had been moved to the IETF Secretariat servers);
- Section 4.2: added Note to challenge the 2-week review period; in current practice, that is almost always exceeded, and some regard it as too short;
- Section 4.3: largely clarified procedures as they happen in practice; adapted language for conformance with RFC 5226; use new home of URN-NID (as mentioned above); the registration template (Appendix A) now "SHOULD" be used;
- Section 4.3: split off new Section 4.4 on Registration Documents, because registrants essentially are encouraged to follow these guidelines for Informal namespaces as well, as far as practical; replaced "RFC" by "Registration Document"; Section 4.4 is subdivided for all mandatory sections;
- Section 4.4.1: made requirements a "MUST";
- Sections 4.4.1 and 4.4.2: added common Note that challenges the need to split Namespace and Community Considerations, based on observed problems in practice to separate the topics, and pointing to overlap with clauses in the registration template due to

bullets listed that are not so clearly related to the headlines under which they appear; suggestion is to avoid duplication, place factual stuff into the template and focus on rationale in these Considerations, perhaps in a common section;

- Section 4.4.3: added discussion of Security Considerations section; advice is to focus on namespace-specific considerations and refer to the SecCons in the "generic" RFCs for the general issues;

- Section 4.4.4: amended discussion of IANA Considerations section; this tries to reflect standing practice and codifies that Formal NIDs are generally proposed by the registrant; added Note that "urn" is permanently reserved and MUST NOT be assigned as a NID, to avoid confusion (as also specified in RFC 2141bis draft); wrt registration maintenance: got rid of wrong reference in RFC 3406 (to RFC 2606);

- Section 6 (IANA Considerations): updated and rephrased description of the role of this document, including a sketch of the history; added teat that tries to precisely describe what is expected from IANA on approval of this draft; added text on procedures and suggest a provisional assignment practice upon "thumbs-up" of the IANA Expert to protect prospective registrants from collateral damage on NID precedence in case the document suffers from delays unrelated to the registration template before it eventually gets approved;

- Section 7 (Acknowledgements): added;

- References: Updated and amended references; added pointers to chartered URNbis work items; removed entirely outdated example material related to legacy documents;

- Appendix A and B.1: added words on Security Considerations section;

- Appendix A (Registration Template): clarified role of text snippets in the Template: hint and commentary now all enclosed in curly braces, with not that these parts shall be removed when filling in the tempalte; indicate that Formal NIDs are normally proposed by registrant; changed date/time ref. from ISO 8601 to RFC 3339; use inherited term "percent-encoding";

- Appendix A -- structure: moved formal clauses on Conformance with URN Syntax and Rules for Lexical Equivalence to vicinity of namespace specific syntax clause, to which these are closely related;

- Appendix A -- changes of clauses: the Declaration of syntactic structure and Rules for Lexical Equivalence clauses now tentatively have been restricted to the NSS part only; this change is described in NOTEs and motivated by the observation of repeated confusion in past and present registration documents, which hopefully can be avoided (and the job of the Expert and reviewers made easier) by leaving discussion of the invariate parts that cannot be re-specified there at the single place where they belong to: the NID is fully specified in the initial clause, rules for the NID and the URI scheme name "urn" are inherited from RFC 2141[bis] and RFC 3986, respectively, and hence the new clause descriptions avoid conflict by taking these components out of scope of these clauses;

- Appendix B.1 (Example Template): facelifted a bit; concerns with IESG policy on examples in RFCs raised in a NOTE;

- Appendix B.2 (Registration steps in practice): updated and clarified description of procedure, in alignment to current practice;

- Appendix C: removed "Changes from RFC 2611"; added this change log;

- General: numerous editorial changes and enhancements, following contemporary RFC style.

Appendix D. Open Issues

Discuss consequences of RFC 2141bis (once consensus is achieved); if proposal for fragment part is adopted, details need to be described per namespace that wants to adopt these possibilities, and maybe the registration template needs a new clause where this will be specified -- or the information has to be assigned to existing clauses.

More elaboration on Services. Since RFC 2483 is considered outdated, but RFC 2483bis not yet a URNbis work item, we might need a registry for URN Services (initially populated from RFC 2483) that can be referred to in namespace registration documents, thus avoiding normative dependencies on a future RFC 2483bis.

Also see the Editorial Notes interspersed in the body of this draft.

What else?

Author's Address

Alfred Hoenes
TR-Sys
Gerlinger Str. 12
Ditzingen D-71254
Germany

EMail: ah@TR-Sys.de