Uniform Resource Name (URN) Syntax
draft-ietf-urnbis-rfc2141bis-urn-02

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 set forth in BCP 66, for which RFC 3406bis is the companion revised specification document replacing RFC 3406.

Discussion

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

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 13, 2012.
Table of Contents

1. Introduction .............................................. 4
   1.1. Historical Perspective and Motivation .................. 4
   1.2. Background on Properties of URNs ..................... 6
   1.3. Objective of this Memo ............................... 7
   1.4. Requirement Language ................................. 8
2. URN Syntax ................................................. 8
   2.1. Namespace Identifier (NID) Syntax .................... 13
   2.2. Namespace Specific String (NSS) Syntax ............... 15
   2.3. Special and Reserved Characters ...................... 15
       2.3.1. Delimiter Characters .......................... 16
       2.3.2. The Percent Character, Percent-Encoding ....... 16
       2.3.3. Other Excluded Characters .................... 17
3. Support of Existing Legacy Naming Systems and New Naming
   Systems ...................................................... 18
4. URN Presentation and Transport ................................ 18
5. Lexical Equivalence of URNs ................................ 18
   5.1. Examples of Lexical Equivalence ...................... 19
6. Functional Equivalence of URNs ................................ 19
7. The 'urn' URI Scheme ....................................... 20
   7.1. Registration of URI Scheme 'urn' ..................... 20
8. Security Considerations ..................................... 22
9. IANA Considerations ........................................ 22
10. Acknowledgements ......................................... 23
11. References ................................................. 23
   11.1. Normative References .............................. 23
   11.2. Informative References .............................. 24
Appendix A. Handling of URNs by URL Resolvers/Browsers .......... 26
Appendix B. Collected ABNF (Informative) ....................... 26
Appendix C. Breakdown of NSS Syntax Evolution since RFC 2141
            (Informative) ...................................... 27
Appendix D. Changes since RFC 2141 (Informative) ............... 29
   D.1. Essential Changes from RFC 2141 .................... 29
   D.2. Changes from RFC 2141 to Individual Draft -00 ....... 29
   D.3. Changes from Individual Draft -00 to -02 .......... 30
   D.4. Changes from Individual Draft -02 to WG Draft -00 .... 30
   D.5. Changes from WG Draft -00 to WG Draft -01 .......... 30
   D.6. Changes from WG Draft -01 to WG Draft -02 .......... 31
Appendix E. How to Locate IETF Documents (Informative) ......... 32
1. Introduction

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.

To this end, URNs are designed as an intrinsic part of the more general framework of Uniform Resource Identifiers (URIs); '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.

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".

It is an explicit design goal that pre-existing systems of persistent identifiers are mapped into the URN framework. Ordinarily, each such traditional identifier system (namespace) -- standard or otherwise -- will occupy its own URN Namespace. However, shared URN Namespaces are possible (and in fact, already exist), but the identifier-driven mechanisms needed to distinguish the originating namespaces make registration and maintenance of such URN Namespaces more complicated.

URN (as a URI Scheme) as such does not have a specific scope. The applicability of the URN system, that is, the totality of the resources that URNs can be assigned to, is the union of all identifier systems that have an associated registered URN Namespace. Ideally every new namespace will thus extend the URN applicability.

1.1. Historical Perspective and Motivation

Since this RFC will be of particular interest for groups and individuals that are 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 E gives hints on how to obtain RFCs and related information.

Attempts to define generally applicable identifiers for network resources go back to the mid-1970s. 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) was 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] was 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, was generalized and consolidated in the Generic URI specification, RFC 2396 [RFC2396]. Most parts of these URI/URL documents were 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 traditionally was 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

This section aims at quoting requirements as identified in the past; it does not attempt to revise or redefine these requirements, but it gives some hints where more than a decade of experience with URNs has shed a different light on past views. The citations below are given here to make this document self-contained and avoid normative down-references to old work.

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 was 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. [...]
o Extensibility: The URN scheme permits future extensions.

o Independence: It is solely the responsibility of a name issuing authority to determine the conditions under which it will issue a name.

o 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 relax unrealistic detail requirements:

o 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.

o Simple comparison: A comparison algorithm for URNs is simple, local, and deterministic. [...] 

o 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. [...] 

Note:
In particular practice gained with active URN Namespaces has shown that this former goal is rather unrealistic, since usually preference is given to 1:1 usage of existing namespaces, which might not have this property. However, we hold that, at least, the rough kind of resource identified by a URN should be easily recognizable for humans.

o 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.

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

o 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 URI Scheme 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 have been 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. This is being undertaken in a companion document, RFC 3406bis [I-D.ietf-urnbis-rfc3406bis-urn-ns-reg].

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 BCP 14 [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.

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

\[
\text{URI} = \text{scheme} \mid \text{"\:" path-rootless [ "?" query ] [ "#" fragment ]}
\]

\[
\text{path-rootless} = \text{segment-nz} \ast ( \text{"/" segment} )
\]
segment-nz = 1*pchar
segment = *pchar

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

In the case of URNs, we have:

scheme = "urn"

and for <path-rootless>, only a single segment is used, but 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.

Note that it is common practise in several existing URN Namespaces
(and fully supported by this syntax) to use additional colon(s) as
separator character(s) in order to introduce further level(s) of
hierarchy into the NSS syntax, where needed. (See also
Section 2.3.1 below.)

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. 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 URN 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
syntax and semantics of these <NSS> values are ordinarily specified
by an existing persistent identifier system (namespace); for
instance, in the 'ISBN' URN Namespace, each NSS must be a valid ISBN.
Some URN Namespaces may have strict rules for well formed NSSs, while
some others may be far more relaxed. There may also be significant
differences regarding the identifier assignment process. The overall
specification requirements and registration procedures for URN
Namespaces are the subject of a dedicated companion document, BCP 66,
which has been updated for conformance to BCP 26 and alignment with
implementation experience RFC 3406bis
[I-D.ietf-urnbis-rfc3406bis-urn-ns-reg].
Notes:

RFC 2141 was published before the URI Generic Syntax was finalized and therefore had to defer the decision on whether <query> and <fragment> components are applicable to URNs. RFC 2141 therefore has reserved the use of bare (unencoded) question mark ("?") and hash ("#") characters in URNs for future usage in conformance with the generic URI syntax.

URNs have been in use for more than a decade. Some user communities want 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 the context of their focal URNs. Therefore, this document allows the designers of selected URN Namespaces to specify the use of the <fragment> component with URNs belonging these Namespaces, whereas the specification of usage of the <query> component is set aside to future standardization efforts for URN resolution. Thus, this draft allows both of these components in the general syntax.

ISSUE:

Regarding fragment identifiers, Section 3.5, para 1 of RFC 3986, indicates that "The fragment identifier ... allows indirect identification of a secondary resource by reference to a primary resource and additional identifying information. The identified secondary resource may be some portion or subset of the primary resource, some view on representations of the primary resource, or some other resource defined or described by those representations." RFC 3986 continues in specifying that the details of the interpretation of fragment identifiers are specific to the media types returned upon resolution of an URI. The entirety of the purposes mentioned in the above quote obviously only can be achieved fully if the "consumer" of the URI becomes aware of the fragment identifier as part of the requested URI, since, e.g., secondary resources might consist in representations might only be available in particular media types. However, RFC 3986 subsequently (in the penultimate paragraph of Section 3.5) specifies that the evaluation of fragment identifiers be a client-side matter and browsers are to strip them from request URIs sent in information retrieval protocols.

Based on this, contemporary web browsers do not communicate fragment identifiers to the web server but perform fragment selection locally on the returned (HTML) resource. To make things even more complicated, the most popular media type (HTML) does only allow to set markers (which are anchor points in the serialized media stream and used by browsers to identify a specific position in the content) and does not allow browsers to
regularly identify actual, conceptional fragments of the media delivered -- like, e.g., the "proper content" of a web page, excluding navigation bars etc. -- so that in practice users have got accustomed to understanding a "fragment" as actually designating a *position* in the media, not a *part* of it.

Therefore, potential usage of <fragment> components in URNs is rather limited and has to be considered very seriously by designers of URN Namespaces that would like to make use of them. URN Namespaces that rely on (unmodified) browser resolution via HTTP/HTML cannot rely on the usage of fragment identifiers to steer the resolution process. Thus, the use of fragment identifiers only seems to be useful for URN Namespaces that are intended to either (a) exclusively make use of resolution systems / clients that can cope with handing off a full-featured URN (including a possible fragment identifier) to the resolution service, or (b) exclusively employ HTML/HTTP based resolution systems / clients, i.e., where the resolution results are returned as HTML such that web browsers can perform the fragment selection, or as some other media type that better supports the identification and actual selection of embedded fragments, even in off-the-shelf web browsers -- perhaps possible for certain variants of XML-based media types.

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 therefore MUST NOT appear unencoded in a NSS. This rule guarantees backwards compatibility with existing URN Namespaces and improves the compatibility of URN syntax with general URI parsers.

The <query> part MUST NOT be present in any *assigned* URN. This specification reserves its use for future standardization related to URN services and resolution. A <query> 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 indicate the requested URN resolution service and/or particular service aspects of the intended resolution response, e.g., to select the kind 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 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 such that there is a need to make reference of parts of such resources in typical network access environments without incurring the effort to assign and maintain different (assigned) NSSs in such cases.

URN Namespaces will deal with various kinds of fragments. For instance, publications can be divided into smaller parts -- journals consist of volumes, issues and articles, and books may contain chapters. These logical fragments are usually not fragments in the sense of the deliberations in the URI Generic Syntax, and if so, `<fragment>` MUST NOT be used. However, namespaces MAY have internal means for identification of logical fragments such as journal articles. For instance, the ISBN (International Standard Book Number) system allows assignment of ISBN numbers to book chapters if they are available as separate items. Namespace specific fragment identification practices are beyond the scope of this document, since they do not rely on URI Generic Syntax, and their application is the primary RECOMMENDED way to deal with fragment identification. If a namespace lacks this possibility, a URN Namespace definition SHOULD define syntactical parts of its NSSs that amend the original identifiers of the underlying namespace in a readily parseable way and serve to allow assignment of URNs in that namespace to the intended abstract fragments. A URN Namespace registration MAY forbid all kind of fragment identification (even if it were possible on the basis of URI Generic Syntax), if the application rules and syntax of the identifier does not allow identification of fragments. ISSN (International Standard Serial Number) is an example of this kind of identifier / namespace.

The use of `<fragment>` as specified in RFC 3986 is possible if and only if (a) the URN Namespace is based on an existing identifier scheme that designates objects of reasonable complexity that there is a need to make reference of parts of such resources in typical network access environments; and (b) these parts will be identified in the canonical manner of the media type(s) delivered upon URN resolution. Direct resolution to them SHOULD be possible and sustainable.

If in a given namespace URNs are never assigned to a particular manifestation of a resource (for instance, a PDF version of a book), but can be transferred from one manifestation to the next or apply to all of them, `<fragment>` usage is forbidden. This applies also to the situation when identified resources are works (without any references to physical embodiments of the work).
The use of <fragment> SHOULD NOT be opted for if the underlying namespace provides for the intrinsic possibility to identify such parts or if there is a readily usable method to construct NSSs by combining the existing identifiers with a component (or components) to identify such parts in an easily discernable manner.

Whether the URI Generic Syntax is applied or not, there are various ways in which fragment identifiers can be generated:

(a) Fragment identifiers (if any) are assigned individually to the relevant fragments of a larger entity during the URN assignment process. If a URN Namespace opts for this model, its specification SHOULD describe the additional syntax restrictions to be adhered to 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 automated 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 SHOULD 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 (i) be consistent with all potential resolution schemes and (ii) not put any undue constraints on any potential resolution scheme, Namespace Identifiers are ASCII strings with the syntax:

NID = (ALPHA / DIGIT) 0*30(ALPHA / DIGIT / "-") (ALPHA / DIGIT)
Note:
The above definition is slightly more restrictive than it was in RFC 2141, to better reflect common practice for "handle"-like identifiers in other IETF protocols (a.k.a. "LDH" syntax) and requirements from RFC 3406bis. RFC 3406bis contains further syntax restrictions on NID strings.

ISSUE:
The above rule still allows NIDs that contain multiple adjacent hyphens or have the form of decimal numbers or decimal number ranges.

Should this be further restricted _in this document_ or is it sufficient to defer to the additional (NID kind specific) rules in RFC 3406bis and the common sense of URN Namespace authors and the designated IANA experts? Anyhow, such restrictions would be fully backward compatible -- as is the above tightened rule -- 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[bis].

The document editor senses the low level of discussion of this issue as an indication that this Issue can be closed.

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.

Note:
This reservation is carried over unchanged from RFC 2141, for historical reasons.

ISSUE:
Further possible reservations and/or details are out of scope for this document, but might be within the scope of RFC 3406bis. It has been suggested that no additional reservations should be codified and the final decision in any case should be left to the common sense of URN Namespace authors and the designated IANA experts.

The document editor senses the low level of discussion of this issue as an indication that this Issue can be closed.
2.2. Namespace Specific String (NSS) Syntax

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

The format of this single canonical form follows:

\[
\text{NSS} = 1^*\text{pchar} \quad ; \text{or equivalent: } \quad \text{NSS} = \text{segment-nz}
\]

(<pchar> and <segment-nz> are defined in Section 3.3 of RFC 3986.)

Note: The informational Appendix C expands on the evolution of the NSS syntax specification since RFC 2141.

ISSUE (for the record):
In comparison to RFC 2141, essentially now "&" and "˜" are allowed in the NSS syntax, in full conformance with the generic URI syntax. On the other hand, the <reserved> characters are no more part of the formal syntax -- unfortunately (or erroneously) these were included in the formal syntax rules of RFC 2141 and only excluded after that fact in the prose, which at least in one instance has lead to a URN Namespace definition document that allows <reserved> in the formal NSS syntax but does _not_ properly exclude their use in the prose. The interpretation of "%" was ambiguous in RFC 2141; it is now only allowed (in the formal syntax and in the prose) in <pct-encoded> constructs.

The document editor senses that this change of the NSS syntax has found consensus and that hence this Issue is regarded as closed.

Depending on the rules governing a namespace, valid identifiers in a namespace might contain characters that are not members of the URN character repertoire above (<pchar>). In order to achieve conformance with this NSS specification, such strings MUST be translated into canonical NSS format before embedding them into a URN, using them as protocol elements, or otherwise passing them on to other applications. Translation is done by encoding each character outside the URN character repertoire 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 latter two characters form the hexadecimal representation of that octet. (See Section 2.3.2 below for more details.)

2.3. Special and Reserved Characters

The remaining printable characters not included in the <pchar> repertoire 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:

\[ \text{gen-delims} = ":" / "/" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" / ":" ]

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

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

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

\textbf{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 \text{<query>} and \text{<fragment>} parts in the generic URI syntax; they are restricted to this role in URNs as well, although the \text{<path>} in URNs only admits a single \text{<segment>} and hence "/" is not allowed. Therefore, these characters MUST NOT appear literally in the \text{<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 \text{<authority>} part, which is absent in URNs. However, for conformance with the generic URI syntax, they are not allowed literally in the \text{<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, Percent-Encoding

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 characters together semantically 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 repertoire 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 -- cf. RFC 3406bis [I-D.ietf-urnbis-rfc3406bis-urn-ns-reg].

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 repertoire, and if used in an URN, MUST be percent-encoded.

```plaintext
excluded = CTL / SP ; control characters and space
/ DQUOTE ;"
/ "#" ; from <gen-delims>
/ "%" ; see above
/ "/" ; from <gen-delims>
/ "<" / ">"
/ "<" ; from <gen-delims>
/ ";" ; from <gen-delims>
/ "\" ; from <gen-delims>
/ "|" ; from <gen-delims>
/ "^" ;
/ \" ;
/ "\{ / "|" / "\}" ;
/ %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 URNs, in either unencoded or percent-encoded form.

In a textual context for a URN, the NSS part 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 NSS.

The more general issue of discerning URNs in non-structured text is not specific to URNs, but a general issue for recognizing URIs (by humans or automata), and hence out of scope of this document.

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

Any identifier to be used as a URN MUST be expressed in conformance with the URI and URN syntax specifications ([RFC3986], this document). If names from (existing or newly devised) 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.

On the other hand, every namespace specific string in a given URN Namespace MUST be based on an identifier that conforms to the requirements of the identifier system to which the URN Namespace is assigned; in the simplest form, if the syntactical rules admit, the NSS can be the original identifier. For instance, every legal NSS in the ISBN Namespace must be a valid ISBN.

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).

Note: Such transformation for the purpose of presentation, if done blindly without NID-specific knowledge of special character usage, might introduce ambiguity, because in the cases described above in the second paragraph of Section 2.3.2, the unescaped and percent-escaped form of the same character might carry different semantics in NSSs of some URN Namespaces.

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;
4. remove the <query> part of the URI, if present.

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. Note also that <fragment> MUST NOT be removed, since there is no lexical equivalence between the "base" URN and one which uses <fragment> -- the former identifies the resource as the whole; the latter just a part of it.

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
7- urn:foo:a123,456?xyz
8- urn:foo:a123,456#xyz

URNs 1, 2, 3, and 7 are all lexically equivalent. URN 4 is not lexically equivalent to any of the other URNs of the above set. The same holds for URN 8.
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 URN 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, 3406bis, 3629 (STD 63), and 3986 (STD 66) ([RFC2169] [RFC2276] [RFC2608] [RFC3401] [RFC3402] [RFC3403] [RFC3404] [I-D.ietf-urnbis-rfc3406bis-urn-ns-reg] [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, an update is in progress as RFC 3406bis [I-D.ietf-urnbis-rfc3406bis-urn-ns-reg]).

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, there may be characters allowed by the syntax and semantics of the identifier system underlying the URN Namespace but not contained in the US-ASCII charset. Such characters MUST first be represented in Unicode and encoded in UTF-8 according to STD 63. Any octets outside the allowed character set MUST then be percent-encoded.

Note that it is perfectly possible that the syntax and semantics of an underlying identifier system does not admit specific characters allowed by the syntax rules in RFC XXXX.

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 protocols.

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; an update is in progress in the URNbis WG as RFC 3406bis
[I-D.ietf-urnbis-rfc3406bis-urn-ns-reg]).

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 particular characters generically allowed in the Namespace
Specific String, any security considerations resulting from such
assignment are outside the scope of this document. It is REQUIRED by
BCP 66 (currently [RFC3406], to be replaced by RFC 3406bis
[I-D.ietf-urnbis-rfc3406bis-urn-ns-reg]) 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.
IANA is asked to again make available the URN Namespace Registry [IANA-URN] in a generic form (i.e. HTML) at the generic URI given in the Reference, and to make the XML and TXT versions available from that HTML version. (This state already had been achieved, but something seems to have been lost in 2011.)

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]. The experiences gathered during the first (more than a) decade of URN usage were also helpful, so individuals and organizations which have implemented and used URNs are also acknowledged.

Many individuals in the URNbis working group have participated in the detailed discussion of this memo. Particular thanks for detailed review comments and text suggestions go to Juha Hakala and Mykyta Yevstifeyev.

11. References

11.1. Normative References


11.2. Informative References

[I-D.ietf-urnbis-rfc3406bis-urn-ns-reg]
Hoenes, A., "Uniform Resource Name (URN) Namespace Definition Mechanisms",
draft-ietf-urnbis-rfc3406bis-urn-ns-reg-02 (work in progress), March 2012.

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

[IANA-URI] IANA, "URI Schemes Registry",

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


Appendix A. 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 B. Collected ABNF (Informative)

As a service to implementers specifically interested in URN syntax, the complete ABNF for URNs is collected here, including the referenced rules from [RFC5234] and [RFC3986]. In case of (unexpected) inconsistencies, these documents remain normative for the respective productions.

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

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

scheme        = ALPHA *( ALPHA / DIGIT / "+" / "-" / "." )
path-rootless = segment-nz *( "/" segment )
query         = *( pchar / "/" / "?" )
fragment      = *( pchar / "/" / "?" )

segment-nz    = 1*pchar
segment       = *pchar
pchar         = unreserved / pct-encoded / sub-delims / ":" / "@"

unreserved    = ALPHA / DIGIT / "-" / "." / "_" / "." / "~"
pct-encoded   = "%" HEXDIG HEXDIG
sub-delims    = "!" / "$" / ";" / ":" / "@" / "'" / "(" / ")" / "[" / "]" / "{" / "}" / "\" / ";" / "#" / "_" / "~"
```

Hoenes                 Expires September 13, 2012              [Page 26]
In the case of URNs, the above rules are subject to more specific restrictions:

scheme = "urn"
  ; specific, fixed (assigned) value

urn-path = NID "::" NSS
  ; to be superimposed on <path-rootless>

NID = ( ALPHA / DIGIT ) 1*31( ALPHA / DIGIT / "-" )
  ; RFC 3406[bis] contains more specific rules

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

The above rules make use of the following "Core Rules" from Appendix B.1 of [RFC5234]:

ALPHA = %x41-5A / %x61-7A   ; A-Z / a-z
DIGIT = %x30-39              ; 0-9
HEXDIG = DIGIT / "A" / "B" / "C" / "D" / "E" / "F"

Appendix C. Breakdown of NSS Syntax Evolution since RFC 2141
(Informative)

In order to make visible the detailed migration path from RFC 2141 and the influence of the evolution of URI syntax from RFC 2396 to RFC 3986 on it, this appendix provides a highly annotated and expanded version of the NSS syntax provided in Section 2.2:

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

In particular, the breakdown below serves to provide evidence of that this syntax correctly reflects the addition of "&" and "~" to the repertoire of characters allowed in the NSS portion of URNs previously allowed by RFC 2141; it expands on the syntax specified in RFC 2141 after translation to standard ABNF.

NSS = 1*URN-char

URN-char = trans / pct-encoded
  ; Note that <pct-encoded> from RFC 3986 here replaces the
  ; explicit, expanded form used in RFC 2141.
trans = ALPHA / DIGIT / u-other
; Note that RFC 2141's <other> has been disambiguated here
; into <u-other>.
; RFC 2141 also said:
; / reserved
; This caused an ambiguity in RFC 2141 with respect to "%", which
; now is resolved here by omission of this dangling alternative.
; After adoption of the generic URI syntax from RFC 3986, there
; is no more need to deal here with the higher-level separator
; characters "/", "?", and "#" contained in <reserved>
; (beyond "%", which is fully taken care of by <pct-encoded>),
; which are part of RFC 3986's <gen-delims>, as shown below.

; From RFC 2141:
;     reserved = "%" / "/" / "?" / "#" ; SIC!
;     ^  ^

      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 = "!" / "¥" / "‘" / "’" / "(" / ")"
;                / "+" / "," / ";" / ":" / "="
;     The URNbis WG arrived at unanimous consensus that "&" can be
;     allowed without harm to backward compatibility for existing
;     URN Namespaces.
; From RFC 3986:
;     unreserved = ALPHA / DIGIT
;                / ":" / "," / ";" / "="
;     The URNbis WG arrived at unanimous consensus that "=" can be
;     allowed without harm to backward compatibility for existing
;     URN Namespaces.

; Since we now allow "&" and "=" <trans> becomes <pchar>,
; greatly simplifying the syntax rules and parsers!

; From RFC 3986:
;     segment-nz = 1*pchar
;     pchar = unreserved / pct-encoded / sub-delims / ":" / "@"
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.2.

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

Added new s1.2 to Introduction, with excerpts from RFC 1737 to provide background on URN functional and syntax requirements. Renumbered previous s1.2 and s1.3 to s1.3 and s1.4, respectively.

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.

D.5. Changes from WG Draft -00 to WG Draft -01

Reflect WG consensus on adding "&" and "˜" to the set of characters allowed in the NSS part of URNs, thus aligning URN syntax with generic URI syntax from RFC 3986.

Moved breakdown of NSS syntax evolution from s2.2 to new Appendix C.

Avoid "[URN] character set" in favor of "character repertoire" to minimize potential clashes with IETF terminology on charsets.

s2.3.3: URN recognition in text documents is regarded out of scope.

The previous version was ambiguous on whether eventual query and/or fragment parts were regarded as part of the NSS; after closer inspection of the syntax, clarification has been added that the <urn-path> syntax is indeed superimposed on the <segment-nz> ABNF rule for
URNs, and hence does not cover the trailing higher level parts (query, fragment) according to the URI syntax.

Filled in Appendix B contents.

Numerous editorial and grammar improvements.

D.6. Changes from WG Draft -01 to WG Draft -02

Added note at the beginning of Section 1.2 highlighting the purpose of this section. The URNbis charter excludes a revision of RFC 1738, and hence the changes suggested on the list to alter and update this section have been dismissed.

Added hint to URN Namespace designers in Section 2 that ":" is customarily used in URN Namespaces to provide further level(s) of hierarchical subdivision of NSSs.

Reworked text on fragment identification issues and resulting specification, mostly based on Juha Hakala’s evaluation of the consensus evolving from the list discussion.

Modified ABNF rule for NIDs to better align it with rules for similar identifiers used in IETF protocols. The new rule now prohibits a trailing hyphen, but defers further restricting rules on NID syntax (based on the kind of NID) to RFC 3406bis.

More clearly documented and marked (still open / already closed) ISSUES. The related text will be removed in the next draft version, whence it should have been transferred into the IETF issue tracking system.

Text of Section 3 revised, based on Juha’s suggestion.

In Section 5, added removal of <query> part (but not <fragment> part) to canonicalization steps for the purpose of determining lexical equivalence of URNs (Juha’s comment). Also added examples showing this.

Elaborated a bit more on Encoding Consideration in the URI Scheme registration template (Juha’s comments).

Numerous editorial corrections and improvements.
Appendix E. 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/>.

Author’s Address

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

EMail: ah@TR-Sys.de
Using International Standard Serial Numbers as Uniform Resource Names
draft-ietf-urnbis-rfc3044bis-issn-urn-00

Abstract

The International Standard Serial Number, ISSN, has been the authoritative identifier for continuing resources (which include serials) for more than three decades. Since 2001, the URN (Uniform Resource Name) namespace "ISSN" has been reserved for ISSNs. The namespace registration was performed in RFC 3044. This document redefines how the revised ISSN standard can be supported within the URN framework, taking into account in particular the latest revision of the ISSN standard in the ISO framework (ISO 3297:2007). Moreover, additional syntax related information required by the RFC 2141[bis] has been included. An updated namespace registration is provided. This document replaces RFC 3044.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 27, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents
Table of Contents

1. Introduction ................................................. 4
2. Conventions used in this document .......................... 4
3. Fundamental Namespace and Community Considerations ...... 5
   3.1. The URN:ISSN Namespace ................................. 5
   3.2. Community Considerations for ISSNs ....................... 5
   4.1. Overview / Namespace considerations ..................... 5
   4.2. ISSN Structure ......................................... 8
   4.3. Encoding Considerations ................................ 8
   4.4. Resolution of ISSN-based URNs .......................... 8
      4.4.1. General ............................................ 9
      4.4.2. Practical Aspects .................................. 9
   4.5. Additional considerations ................................ 10
      4.5.1. ISSN-L (or "linking ISSN") ......................... 10
      4.5.2. Updating and management of URLs corresponding to
             resources identified by URN:ISSN ...................... 12
5. URN Namespace Registration and Use ......................... 14
   5.1. URN Namespace ID Registration for the International
        Standard Serial Number (ISSN) ......................... 14
6. Security Considerations ...................................... 17
7. IANA Considerations ......................................... 17
8. Acknowledgements ............................................ 18
9. References .................................................. 18
   9.1. Normative References .................................... 18
   9.2. Informative References .................................. 18
Author’s Address ............................................... 19
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. In August 2011 there were 44 registered URN namespaces (see [IANA-URN]), one of which belongs to ISSN, International Standard Serial Number, as specified 2001 in RFC 3044 [RFC3044].

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 and 3044 ([RFC3187], [RFC3044]), respectively.

The RFC at hand replaces RFC 3044; all ISSN information has been updated and the namespace registration revised to make it compliant with 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].
3. Fundamental Namespace and Community Considerations

3.1. The URN:ISSN Namespace

ISSN is an authoritative standard identifier system for continuing resources and in particular serial publications. Therefore, any useful and deployable method for identifying these entities for network-wide reference and making their metadata available on the Internet needs to be based on ISSNs. ISSNs are authoritatively referenced in a centrally managed database called the "ISSN Register" which can be used as the basis for URN:ISSN resolution services.

3.2. Community Considerations for ISSNs

ISSNs are assigned under the auspices of the International ISSN International Centre and national ISSN Centres. ISSN assignment is a well managed and understood process, but as in any process administered by humans errors do take place. While some errors may happen in the ISSN Register itself and are readily corrected, most errors happen in the outside world through the use of inappropriate ISSN in external references or the resources themselves.

Continuing resources, including serials, most often consist of component parts such as volumes, issues, articles. The ISSN standard does not allow augmentation of the ISSN of the resource with (URI) fragments for identification of component parts. If a fragment identifier is added to an ISSN, the resulting namespace specific string will not be an ISSN.

For all the communities interested by the identification of continuing resources and their contents, URN:ISSN-based identification and resolution services offer efficient, reliable and persistent access to resources and/or resource-related services. The users will not need special tools for this as Web browsers are sufficient to display bibliographic information or when appropriate, an access point to the resources themselves.

The next chapter presents an overview of the application of the URN:ISSN namespace and the principles, and systems used, for the resolution of ISSN-based URNs.

4. International Standard Serial Numbers

4.1. Overview / Namespace considerations

Each ISSN is a unique identifier for a specific serial or other continuing resource in a defined medium.
ISSN are applicable to serials and other continuing resources, whether past, present, or to be produced in the foreseeable future, whatever the medium of production. Continuing resources are issued over time with no predetermined conclusion, they include serials and ongoing integrating resources. ISSN are assigned to the entire population of serials and to ongoing integrating resources.

Serials are resources for which additional information is supplied indefinitely in a succession of discrete parts. All serials are eligible for an ISSN. Also eligible for ISSN assignment are those bibliographic resources issued in successive issues or parts which bear numbering and that also bear other characteristics of a serial (e.g. frequency in the title), but whose duration is limited (e.g. the newsletter of an event).

Ongoing integrating resources are resources that are updated over time and with no predetermined conclusion, for which the updates are integrated into the resources and do not remain discrete. Those ongoing integrating resources which are eligible for an ISSN must be updated indefinitely, and/or have an update statement. Advertising and individual home pages, online diaries, personal weblogs, and web sites consisting exclusively of links are not eligible for an ISSN.

Individual monographs, technical reports, sound and video recordings, printed music publications, audiovisual works and musical works have their own identifier systems. Such items may carry an ISSN in addition to their own standard numbers when they are part of a continuing resource.

Only one ISSN is assigned to a continuing resource in a defined medium. This ISSN is permanently linked to the so called key title, a standardized form of title derived from information appearing on the continuing resource. A key title is unique to a particular continuing resource. Titles which would otherwise not be unique are made unique by the addition of qualifying elements. In cases where the title changes sufficiently (as per specific rules defined in the ISSN Manual) to warrant creating a new key title, a new ISSN is assigned. In cases where the medium of the continuing resource changes, a new ISSN and a new key title are assigned.

Changes in publisher, country, language, frequency, subject scope or any other characteristic of a given continuing resource do not warrant the assignment of a new ISSN. Title changes which are deemed minor are registered in the ISSN metadata as "variant titles".

When a new ISSN is assigned to a continuing resource (because of a significant title change or of a medium change), both the "former" and "new" ISSN are deemed valid and identify two distinct entities:
each of them identifies the continuing resource in its incarnation in a given time interval, under a particular key title and/or physical medium. "Dead" continuing resources are dead in the sense that they are no longer updated, but they continue to be accessible in library shelves or as archives on servers and their continuing identification is an obvious need for the whole chain of stakeholders.

In such cases, ISSNs, through the metadata stored in the ISSN records of the ISSN Register are reciprocally linked. In fact, one of the major aspects of the ISSN Register is its linking structure through which various incarnations of continuing resources are reciprocally linked using the ISSN as pointer. There are different categories of such links (for former and successor titles, other medium editions, other language editions, supplements etc.). A given ISSN may thus be linked directly to a number of other ISSN which in turn may be linked to other ISSNs etc. We can thus define the concepts of directly and indirectly linked ISSNs.

Allocation of blocks of ISSN

The International Centre is responsible for the allocation of blocks of ISSN to National Centres. Each Centre receives limited blocks of numbers. In using blocks of ISSN, National ISSN Centres adhere to the following procedures:

I. Report all ISSN assigned by their centre to the ISSN Register;

II. Use ISSN within their assigned block consecutively and use up one block completely before starting another block;

III. Ensure that ISSN assignments made in advance of publication or production of a continuing resource are recorded in the ISSN Register by determining if publication or production of the resource has occurred and creating the appropriate ISSN records.

Although it is possible, on the basis of internal management procedures of the ISSN Register, to determine in a majority of cases that a given ISSN is part of a given block of ISSN allocated to a given ISSN National Centre, this feature cannot be used for ISSN resolution mechanisms for the following reasons:

- The country of publication of continuing resource may change over time (which implies that the responsibility of a given ISSN may be switched from one ISSN National Centre to another and that the ISSN block from which the given ISSN was used may not correspond to the actual country of publication)

- A significant number of ISSN were initially assigned in a
"multinational" framework where the block of ISSN was not attached to a given country

- There exists a significant number of "multinational" publications where "national" responsibility for ISSN assignment and management has necessarily to be defined on somewhat arbitrary basis which may vary over time

For similar or identical reasons, although metadata attached to an ISSN in the framework of the ISSN Register defines current National ISSN Centre responsible for the management of the ISSN and its corresponding ISSN record, this information cannot and should not be used to infer a resolution path.

4.2. ISSN Structure

An ISSN consists of eight digits. These are the Arabic numerals 0 to 9, except that an upper case X can sometimes occur in the final position as a check digit (when representing the number "10"). Since ISSN are likely to be used in the same context as codes designed for other purposes, a distinction must be preserved in the form of presentation. An ISSN therefore appears as two groups of four digits, separated by a hyphen:

ISSN 0317-8471

ISSN 1050-124X

The check digit is always located in the extreme right (low order) position, and is calculated on a modulus 11 basis using weights 8 to 2.

4.3. Encoding Considerations

Embedding ISSNs within the URN framework does not present encoding problems, since all of the characters that can appear in an ISSN (the 10 digits (0 to 9), the hyphen and capital letter X) are valid in the namespace-specific string (NSS) part of the URN. percent-encoding, as described in RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn], is never needed. In order to improve readability of the NSS, central hyphens SHOULD always be used.

Example: URN:ISSN:1234-1231

4.4. Resolution of ISSN-based URNs
4.4.1. General

For URN resolution purposes, all elements, including the check digit and the central hyphen, must be taken into account.

If a local resource stores and manages ISSN without a central hyphen, it should be programmatically added for the constitution of URN:ISSN.

Applications, such as the national bibliography or the open archive of a university, can use the URN as the persistent address of the resource. There is just one place (the URN registry) where the address is mapped to one or more physical locations.

4.4.2. Practical Aspects

Persistence is one of the key features for any persistent identifier system. There are three inter-related aspects of persistence that need to be discussed: persistence of the resource itself, persistence of the identifier, and persistence of the URN-based resolvers.

Persistence of the resources: continuing resources are complex objects which evolve over time. In their (mostly) paper incarnations, they have been stored on library shelves sometimes for centuries. Bibliographic records mediate identification and access. If a continuing resource is available on print only, its URN:ISSN will resolve to the bibliographic record in the ISSN register.

The ISSN Register has identified (at the beginning of 2012) almost 100,000 online continuing resources which may or may not have print equivalents. Furthermore, vast digitization efforts are now undertaken over the world to create electronic archives of printed continuing resources (these initiatives have often a dual aim of long term preservation and economies in shelving space); efforts are also under way to manage the long term preservation of online continuing resources.

All these efforts which have as a goal the persistence of the continuing resources will be all the more successful if they benefit from a standardized identification layer. This obviously also has an impact on the management of contents (volumes, issues, and first and foremost articles) where linking frameworks which appeared during the last ten years (CROSSREF or Open URL) make heavy use of the ISSN.

Persistence of the identifier: The ISSN as an identifier is persistent in the sense that once assigned, an ISSN will never be re-assigned to a different continuing resource.

Persistence of the resolvers: URN resolvers are not static. The
services they’ll supply will change over time, due to changes in technical infrastructure. For instance, new URN resolution services may be added or modified over time. Persistence of the resolvers themselves is mainly an organizational issue, related to the persistence of organizations maintaining them. As URN:ISSN resolution services will be based on the ISSN Register, which is itself a persistent resource which has been maintained for almost 35 years, we may thus assume that URN:ISSN resolution services will be persistent.

The ISSN Register will initially support four resolution services specified in RFC 2483 [RFC2483], namely I2L, I2Ls, I2C and I2Cs. Only I2C and I2Cs (URI to URC(s); delivery of descriptive metadata related to the resource) are valid for non-networked resources. Descriptive metadata can only be supplied in the MARC21 format.

Due to the structure of the ISSN, it is assumed that URN:ISSN resolution can only be reliably achieved through a central service, based on the ISSN Register which in turn can benefit from automated linking with other local resources using the ISSN as an identifier. Only a combination of the authority of the centralized ISSN Register and of local data can guarantee both reliability and persistence.

4.5. Additional considerations

4.5.1. ISSN-L (or "linking ISSN")

In the framework of URN:ISSN resolution, the ISSN-L is a very important new feature.

The ISSN-L (or "linking ISSN") is an important modification introduced in the latest revision of the ISSN standard in the ISO framework.

The ISSN-L has been defined to meet the need for a collocation, or grouping mechanism that brings together the various medium versions of a continuing resource, and thus facilitate content management.

The ISSN-L is an ISSN designated by the ISSN Network to group the different media versions of a continuing resource.

Only one ISSN-L is designated regardless of how many different medium versions of a continuing resource exist. A continuing resource will be associated with only one ISSN-L.
4.5.1.1. Designation of the ISSN-L

The designation of the ISSN-L is carried out either by a centre of the ISSN Network or is performed automatically as records are added to the ISSN Register. It is done either by those ISSN National Centres that are able to undertake this responsibility, or by the International Centre. The records produced by these National Centres include the ISSN-L in the ISSN records under their responsibility.

4.5.1.2. Rules for the designation of the ISSN-L

The first ISSN assigned, in the ISSN Register, to any medium version of a continuing resource is designated by default to function also as the ISSN-L and applies to all other media versions of that resource identified in the ISSN Register. An ISSN-L is designated for each continuing resource identified in the ISSN Register, even if the continuing resource is issued in only one medium. Only one ISSN-L is designated regardless of how many different medium versions of a continuing resource exist.

In the framework of URN:ISSN resolution, whether an ISSN is submitted as an ISSN-L or as an ISSN should be considered as having no practical impact as the response should always include by default basic resolution data for all ISSN which may be linked through a common ISSN-L.

For efficient practical resolution purposes, it should not be assumed that the requesting service has an unambiguous knowledge of either:

- The medium version associated to a given ISSN
- The ISSN which is designated to function as ISSN-L linking the different medium versions

The URN:ISSN resolution service should make no such assumption concerning the knowledge of the requesting service. The URN:ISSN resolution should make available sufficient authoritative metadata so as to allow the requesting service to obtain the expected response, even if the ISSN submitted is not used fully adequately by the requestor. URN:ISSN resolution metadata should allow the requesting service to check and correct if necessary its potentially incorrect assumptions, so as to avoid the following situations:

- An ISSN would be left unresolved (for instance because a "print" ISSN was sent instead of the "online" ISSN and I2L service is requested)
The requesting service would be left unaware of the existence of other ISSN linked through a common linking ISSN-L, because it would have submitted for resolution an ISSN not designated as ISSN-L.

The requesting service would have to perform several successive URN:ISSN resolution requests for all ISSN linked through a common ISSN-L.

Examples:

URN:ISSN:1234-1231 identifies the current print edition of "Medical News".

URN:ISSN:1560-1560 identifies the current online edition of "Medical News".

The ISSN-L linking both media versions of "Medical News" happens to be ISSN-L 1234-1231 (i.e. based on the ISSN 1234-1231, designated as such in the framework of the management of the ISSN Register).

The resolution of URN:ISSN:1234-1231 should be equivalent to the resolution of URN:ISSN 1560-1560, i.e; in both cases one should find a reference to the other medium version.

4.5.2. Updating and management of URLs corresponding to resources identified by URN:ISSN

As already indicated, continuing resources are complex objects or sets of objects which evolve over time and can be partly or fully duplicated, published, archived, remixed. Various entities (publishers, issuing bodies, libraries, content aggregators, archiving institutions, subscription services...) may be partly or fully responsible over time for the online management of these objects.

Their stewardship may be ambiguously implemented for various reasons:

- It may extend over a restricted sub-set of the resource (only from a given year for instance)

- It may express itself over time through various technical implementations which may translate into server name and URL changes

- It may or may not be associated with an adequate stewardship over the appropriate identification of the objects) (inadequate ISSN
being used for media versions, title changes not taken into account...)

- The ISSN assigned may or may not be used in a consistent and standardized machine processable form in the objects themselves or in external reference lists. Even if an appropriate ISSN is used in the stored metadata, it may be duplicated at the level of all the sub-objects (volumes, issues, articles...) making it impractical to ascertain the adequate entry point of the continuing resource itself as a whole.

In some cases, continuing resources are not processed or managed as such and only their constituent parts (for instance, volumes, issues, articles...) are made directly accessible as evolving sets.

Last but not least, commercial publications may restrict access to authenticated users only.

This practically means that "resolution" (in the sense of localization) of continuing resources can best be achieved in the framework of long term coordinated efforts, but cannot be guaranteed in all cases. The best results will of course be obtained with "preservation silos" or big entities. Concerning the "long tail" of small publications, preservation initiatives are best equipped to handle the link between identification and access to the resources themselves.

This means that URN:ISSN resolution will not be able to offer "full resolution" (i.e. reliable access to the resource itself) in all cases, even if the resource is "online".

On the other hand, URN:ISSN extended resolution services could be offered on a systematic basis thanks to the metadata stored in the ISSN Register and its potential linking with external resources, such as:

- basic metadata stored in the ISSN Register identifying or describing the resource, including, as stated above, metadata from the ISSN records linked through a common ISSN-L; in particular, the minimum set of data should include the category of the ISSN as defined above, so as to allow for instance and adequate processing of "cancelled ISSNs".

- access to the direct or indirect linking structure associating the ISSN with other continuing resources

- linking to administrative metadata such as archiving and preservation information about the resource or in a more general
and wider way, to any kind of relevant ancillary information: publisher, issuing body, availability through third party outfits, such as union catalogues etc., external evaluation and authentication data, in fact to any other party or service offering relevant ISSN based metadata.

URN:ISSN resolution services can be both human readable and machine processable so as to support semantic web compatible services.

It should finally be noted that as the ISSN Register is a subscription based resource, URN:ISSN resolution cannot be a fully open service.

5. URN Namespace Registration and Use

The formal URN Namespace Identifier Registration for the pre-2007 version of the International Standard Serial Number (ISSN) standard was done in "RFC 3044 [RFC3044]."

The revised ISSN standard does not require a new namespace, but the registration is renewed here. The registrant organization has moved from a former address to a new one in Paris. Moreover, the description of the NSS and resolution details have been amended.

5.1. URN Namespace ID Registration for the International Standard Serial Number (ISSN)

This registration describes how the International Standard Serial Number (ISSN) can be supported within the URN framework.

_NAMESPACE ID: _ ISSN

This Namespace ID has already been assigned to the International Standard Serial Number in January 2001 when the namespace was initially registered.

_REGISTRATION INFORMATION:_

Version: 2 Date: 2012-02-23

_DECLARED REGISTRANT OF THE NAMESPACE:_

_REGISTERING ORGANIZATION:_ Centre international d’Enregistrement des Publications en Série - CIEPS-ISSN - ISSN International Centre

_DESIGNATED CONTACT PERSON:_
Name: Ms. Francoise Pelle

Affiliation: Director, ISSN International Centre

Email: issnic@issn.org

45 rue de Turbigo, 75003 PARIS, FRANCE

Web URL: <http://www.issn.org/ >

_Declaration of syntactic structure of NSS part: _

__The ISSN syntax is as follows:

NNNN-NNNC

where N is a Digit character [0..9]

C is either a Digit character or letter "X" [0..9,X]

C is the check character

Example 1: URN:ISSN:1234-1231

Example 2: URN:ISSN: 0259-000X

_Relevant ancillary documentation: _

The ISSN (International Standard Serial Number) is a unique machine-readable identification number, which identifies unambiguously any continuing resource. This number is defined in ISO Standard 3297: 2007. ISSNs have been in use for more than 30 years and they have deeply affected the handling of continuing resources and their contents. 88 countries are officially ISSN members (at the beginning of 2012).

The administration of the ISSN system is carried out at two levels: International Centre and National Centres.

The ISSN International Centre is located in Paris (France). The main functions of the Centre are:

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

- To maintain and publish the ISSN Register.
o To allocate blocks of ISSNs to ISSN National Centres.

o to assign ISSN to international publications and to serials issued in countries with no National Centre.

Detailed information about ISSN usage can be found from the ISSN Users’ Manual.


_Conformance with URN Syntax:_

Legal ISSN characters are 0-9 and hyphen and X. No percent-encoding is needed. Hyphen carries no semantic content but SHOULD NOT be dropped from the NSS.

_Rules for Lexical Equivalence of NSS part:_

ISSN numbers are usually printed with the letters 'ISSN' and a single blank preceding the ISSN proper (for instance: ISSN 1234-1231). Any data preceding the ISSN MUST NOT be included in the NSS. No percent encoding is needed.

Prior to comparing the NSS of two ISSN-based URNs for equivalence, all hyphens, if present, MUST be removed and letter 'X' capitalized.

Note that, according to RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn], the prefix "URN:ISSN:" 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:_

ISSN is a unique and persistent identifier. An ISSN, once it has been assigned, MUST NOT be re-used for another continuing resource. 'ISSN' URNs inherit the uniqueness and persistence properties from ISSNs.

_Process of identifier assignment:_

Assignment of ISSNs is controlled, and 'ISSN' URNs immediately inherit this property. There are three levels of control: the ISSN international Centre, national ISSN Centres, and finally all the stakeholders responsible for a correct use of the ISSN system.
Process for identifier resolution: See Section 4.3 of RFC3044.

_Validation mechanism:_

The check digit helps to assure the correctness of an ISSN number assigned for a continuing resource when it has been entered or processed. Applications processing bibliographic data such as integrated library systems MAY use the check digit to check the correctness of the ISSN string. If the number is found to be wrong due to, e.g., a typing error made by a publisher, the correct ISSN SHOULD nevertheless be used while the wrong number may be stored alongside for reference. Although the resource itself may only contain the wrong number, national bibliographies and systems used by relevant communities often will contain both the wrong and correct ISSN number.

_Scope:_

ISSN is a global identifier system used for identification of continuing resources. It is very widely used and supported by the publishing and scholarly publication communities.

6. Security Considerations

This document proposes means of encoding ISSNs within the URN framework. An ISSN-based URN resolution service is depicted here, but in a generic level only; 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 ISSNs. Issues regarding intellectual property rights associated with objects identified by the ISSNs are also beyond the scope of this document.

Access control mechanisms may be implemented to limit access to some or all URN resolution services available in the URN Register. Such mechanisms, if any, will be discussed separately.

7. IANA Considerations

IANA is asked to update the existing registration of the Formal URN Namespace ’ISSN’ 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 is part of the URNBIS effort to revise the basic URN RFCs. 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 PERSID project <http://www.persid.org/> is a significant impulse to this work.

9. References

9.1. Normative References

[I-D.ietf-urnbis-rfc2141bis-urn]
Hoenes, A., "Uniform Resource Name (URN) Syntax",
draft-ietf-urnbis-rfc2141bis-urn-01 (work in progress),
October 2011.

[I-D.ietf-urnbis-rfc3406bis-urn-ns-reg]
Hoenes, A., "Uniform Resource Name (URN) Namespace Definition Mechanisms",
draft-ietf-urnbis-rfc3406bis-urn-ns-reg-01 (work in progress),
October 2011.


9.2. Informative References

[I-D.ietf-urnbis-rfc3187bis-isbn-urn]
draft-ietf-urnbis-rfc3187bis-isbn-urn-02 (work in progress),
February 2012.

[IANA-URI]
IANA, "URI Schemes Registry",

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


Author’s Address

Pierre Godefroy
ISSN International Centre
45 rue de Turbigo
Paris, 75003
France

Phone: +33-1-44-88-22-18
Email: godefroy@issn.org
Abstract

The International Standard Book Number, ISBN, is a widely used identifier for monographic publications. Since 2001, the URN (Uniform Resource Name) namespace "ISBN" has been reserved for ISBNs. The namespace registration was performed in RFC 3187 and applied only to the ISBN as specified in the ISO Standard 2108-1992, now known as "ISBN-10". 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 of these ISBN standard versions can be supported within the URN framework. Moreover, additional syntax related information required by RFC 2141[bis] has been included. An updated namespace registration is provided. It 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

This draft is based on individual work started in 2008. When the URNnBIS working group was launched, revision of the ISBN namespace registration was included in its charter.

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

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.
Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 19, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.
Table of Contents

1. Introduction ........................................ 4
2. Conventions used in this document .................. 5
3. Fundamental Namespace and Community Considerations ..... 5
   3.2. Community Considerations for ISBNs ............... 5
   4.1. Overview / Namespace Considerations ............... 6
      4.1.1. ISBN-10 Structure .......................... 6
      4.1.2. ISBN-13 Structure .......................... 7
   4.2. Encoding Considerations .......................... 8
   4.3. Resolution of ISBN-based URNs ..................... 8
      4.3.1. General .................................... 8
      4.3.2. Practical Aspects .......................... 9
   4.4. Additional Considerations ......................... 12
5. URN Namespace Registration and Use ................... 13
   5.1. URN Namespace ID Registration for the International
        Standard Book Number (ISBN) .................... 13
6. Security Considerations ................................ 17
7. IANA Considerations .................................. 17
8. Acknowledgements ..................................... 17
9. References ........................................... 18
   9.1. Normative References ............................ 18
   9.2. Informative References .......................... 19
Appendix A. Draft Change Log ............................ 20
   A.1. draft-hakala-rfc3187bis-isbn-urn-00 to
        draft-ietf-urnbis-*-00 .......................... 20
   A.2. draft-ietf-urnbis-rfc3187bis-isbn-urn-00 to -01 .... 20
   A.3. draft-ietf-urnbis-rfc3187bis-isbn-urn-01 to -02 .... 20
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. In February 2012, there were 45 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. 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), mainly due to the identifier's limited popularity. Moreover, the URN resolution process for SICIs would be complicated.

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", "SHALT NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].


"URN:ISBN" is used as a shorthand for "ISBN-based URN".

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 for identifying these entities for Internet-wide reference and making their metadata available on the Internet needs to be based on ISBNS.

3.2. Community Considerations for ISBNS

ISBNS are assigned under the auspices of the International ISBN Agency [ISBNORG] and national/regional ISBN agencies. ISBN assignment is a well managed and understood process, but as in any process administered by humans, errors do take place. If so, there are procedures in place for fixing the incorrect ISBNS.

Books are finite objects, which may consist of component parts such as chapters or short stories / novellas. Such component parts may be given their own ISBNS if and only if they are available separately. The ISBN standard does not allow augmentation of the ISBN of the book with (URI) fragments for identification of the book’s physical component parts. If a fragment identifier is added to an ISBN, the resulting namespace specific string will not be an ISBN; it could be another identifier such as a national bibliography number (NBN).

In late 90s there was an attempt to develop BICI (Book Item and Contribution Identifier), but the standard was neither completed nor implemented. BICIs would have been based on ISBNS, and the idea was to generate them programatically for e-books containing structured metadata. The applications needed for this failed to materialize.
The materials identified by an ISBN may exist only in printed or other physical form, not electronically. And even if an electronic book exists, access rights may be limited. ISBN-based URN resolution services are expected to support a wide variety of information services related to books. Hence, when the identified manifestation of the book cannot be supplied, the applicable URN:ISBN resolver SHOULD supply descriptive and rights metadata about the relevant manifestation; the resolver MAY also provide links to other manifestations of the same work, or to related works.

National libraries are among the key organizations providing URN resolution services for books. Many of them are currently digitizing their historical printed books collections. As a rule, a digitized book does not get an ISBN, especially if the original printed book did not have one. Instead, national bibliography numbers are often used for identification. In such cases, the applicable resolver MAY resolve the ISBN of the printed original by pointing to the digital copy.

For library users and Internet-based supply chain management for the delivery of monographs, URN:ISBN-based identification and resolution services offer efficient, reliable and persistent access to resources and/or resource-related services. The users will not need special tools for this; Web browsers are sufficient.

The next section presents an 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. ISO 2108 requires that each product form (e.g. hardcover, paperback, PDF) has its own ISBN.

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. Note that these hyphens can be removed; ISBNs with and without the hyphens are lexically equivalent. The parts are as follows (in this order):

- a group identifier that specifies a group of publishers, based on national scope, geographic scope, or some other criteria;
o the publisher identifier;

o the title identifier; and

o a modulo 11 check digit, using X instead of 10; the details of the calculation are specified in the ISO Standard [ISO1].


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 length, but the other parts have variable length. These parts are as follows (in this order):

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

o 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;

o the registrant element;

o the publication element; and

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


Further, the 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:


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. Percent-encoding, as described in RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn], is never needed. In order to improve readability of the NSS, hyphens MAY be used.


4.3. Resolution of ISBN-based URNs

4.3.1. General

For URN resolution purposes, all elements except the check digit (0-9 for ISBN-13, and 0-9 or X for legacy ISBN-10) 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 combinations of prefix and registration group elements are
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 persistent address of the resource. There is just one place (the URN registry) where the address is mapped to one or more physical locations.

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 [ISBN_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".

Any national URN registry can resolve URN:ISBNs with foreign registration group element values if a) there is a URN:ISBN
resolution service for that country, b) the national resolution service is aware of the existence of the foreign service and how to find it, and c) the two resolution services can communicate with one another. The PERSID project (<http://www.persid.org/>) developed such an infrastructure for the URN:NBN namespace.

Alternatively, instead of linking the national resolvers together, it is also possible to build international resolvers that copy resolution data from several national services, or to create a way station which will enable the resolvers to communicate with one another. We can assume that the network of URN:ISBN resolvers will grow, and at the same time the set of services they support will also grow and become more diverse. Such development might make these union resolvers and way stations more important.

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, the German, Austrian, and Swiss national registries, in this order --, which collectively is 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 on 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 respective 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 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] can 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, 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 might be able to provide bibliographic information only, while the publisher can provide the book itself, on its own terms. Users can expect Resolution services to co-exist and complement one another. The same ISBN can 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 metadata, whereas a service based on the DOI system provided by the publisher may supply the book, parts of the book or various other services.

Persistence is one of the key features for any persistent identifier system. There are three inter-related aspects of persistence that need to be discussed: persistence of the resource itself, persistence of the identifier, and persistence of the URN-based resolvers.

ISBNs are assigned to manifestations (physical embodiments) of books. Printed books tend to be persistent, so even after 500 years, a URN:ISBN identifying a printed book can resolve to a bibliographic description of the book, which MAY contain the location of the book.

With digital books things get more complicated. According to ISO 2108, each product form must have a separate ISBN, but digital manifestation will not be long-lived. Anyone who tries to use a 100-year old e-book will probably be disappointed. Manifestations of an
e-book should be interlinked (using, for instance, the work level metadata record) so as to make a user aware of the existence of these product forms. This will enable the user to retrieve the form that matches his / her interests best. Some users may prefer a modern manifestation although it might not have the original look and feel, while other users may want the original manifestation which is authentic but might require digital archaeology for access.

Manifestations of e-books, like other e-resources, are not required to be persistent per se, but require successive migrations into new file formats. Employing URN:ISBN will support information architectures that enable persistent access to the relevant intellectual content (work), independent of its form, although, according to the ISO Standard, ISBNs should not be used to identify the works themselves.

URN resolvers are not static. The services they’ll supply will change over time, due to changes in technical infrastructure. For instance, implementation of long term preservation systems will enable and necessitate a set of new URN resolution services.

Persistence of resolvers themselves is mainly an organizational issue, related to the persistence of organizations maintaining them. As URN:ISBN resolution services will be supplied (among others) by the national libraries to enable access to their legal deposit collections, we may assume that URN:ISBN resolution services will be persistent.

4.4. Additional Considerations

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

- Product form and the means of delivery are irrelevant to the decision whether a product needs an ISBN or not. If the content meets the requirements of the standard, it gets an ISBN, no matter what the file format of the delivery system.

- Each product form (manifestation) 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.

Mistakes can happen. For instance, an ISBN has sometimes been reused for another book. These reasonably rare kind of human error do 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. The error described above SHOULD only lead into the retrieval of two 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.

Libraries routinely correct ISBN mistakes. Their catalogs provide cross references ("incorrect ISBN -> correct ISBN"). This MUST 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. The registrant organization has moved from Staatsbibliothek zu Berlin - Preussischer Kulturbesitz to The International ISBN Agency, London, U.K. Moreover, the description of the NSS and resolution details have been 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.
Namespace ID: ISBN

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

Registration Information:

Version: 2
Date: 2012-02-16

Declared registrant of the namespace:

Registering Organization: The International ISBN Agency

Designated Contact Person:
Name: Ms. Stella Griffiths
Affiliation: Executive Director, The International ISBN Agency
Email: info@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 namespace-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.


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 2108:2005. ISBNs has been in use for more than 30 years and they have revolutionised the international book-trade. 170 countries and territories are officially ISBN members.

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

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.


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. Hyphen carries no semantic content and MAY be dropped from the NSS.

Rules for Lexical Equivalence of NSS part:

ISBN numbers are usually printed with the letters ‘ISBN’ and a single blank preceding the ISBN proper (for instance: ISBN 951-746-795-8). The data preceding the ISBN MUST NOT be included in the NSS. No percent-encoding is needed.

Prior to comparing the NSS of two ISBN-based URNs for equivalence, all hyphens, if present, 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 NOT be re-used for another book or another product form of the same book. A single product form (manifestation) of a book MUST NOT get a new ISBN. 'ISBN' URNs inherit the uniqueness and persistence properties from ISBNs. Please note that the same ISBN can be used as in another persistent identifier system, such as DOI or Handle. The resulting persistent identifier SHALL NOT render the URN:ISBN non-unique; however, it might provide different resolution services than URN:ISBN.

If there are multiple manifestations of a single literary work such as a novel, each one MUST 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. Applications processing bibliographic data such as integrated library systems MAY 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 SHOULD be 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.

Beyond the generic security considerations laid out in the underlying documents listed in the Normative References (Section 9.1), no specific security threats have been identified for ISBN-based URNs.

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. 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 participants of project PERSID (<http://www.persid.org/> ) contributed significantly to the standards work.

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

Stella Griffiths has advised and guided the development of this document, has verified the technical content from the director’s view of the International ISBN Agency, and provided valuable comments.

Larry Masinter, Subramanian Moonesamy, Julian Reschke, and other participants of the URNbis working group have provided review comments and text suggestions that have improved this document.

9. References

9.1. Normative References

[I-D.ietf-urnbis-rfc2141bis-urn]
Hoenes, A., "Uniform Resource Name (URN) Syntax",
draft-ietf-urnbis-rfc2141bis-urn-01 (work in progress),
October 2011.

[I-D.ietf-urnbis-rfc3406bis-urn-ns-reg]
Hoenes, A., "Uniform Resource Name (URN) Namespace
Definition Mechanisms",
draft-ietf-urnbis-rfc3406bis-urn-ns-reg-01 (work in
progress), October 2011.

[ISO1] ISO, "Information and documentation – The International

[ISO2] ISO, "Information and documentation – The International

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate

Resource Identifier (URI): Generic Syntax", STD 66,
9.2. Informative References


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;
- registration template adapted to rfc3406bis [-00];
- numerous editorial fixes and improvements.

A.2. draft-ietf-urnbis-rfc3187bis-isbn-urn-00 to -01

- discussion on persistence altered, based on list discussion;
- changes and amendments to discussion of URN resolution services;
- discussion of fragment part usage added;
- broken link to ISBN manual fixed based on feedback from [ISBNORG];
- various editorial fixes and enhancements.

A.3. draft-ietf-urnbis-rfc3187bis-isbn-urn-01 to -02

- addressed review comments by LM and SM;
- cleanup of requirements language, but
- kept RFC 2119 terms where non-canonical/non-intuitive behavior of
  resolver systems is specified;
- URLs for ISBN user manual (new public version) etc. updated;
- numerous editorial updates, fixes, and enhancements.

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
Abstract

Uniform Resource Names (URNs) are intended to serve as persistent, location-independent, resource identifiers. To structure and organize their usage, the URN syntax (RFC 2141bis) specifies a hierarchy that 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 metadata related to them).

To achieve these goals, 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 RFC serves as a guideline for authors of URN Namespace definition and registration documents and the process to be followed to register a URN Namespace with IANA. It describes the essential content of such documents and how they shall be structured to allow readers familiar with the scheme to quickly assess the properties of a specific URN Namespace.

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

Discussion

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

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 13, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.
Table of Contents

1. Introduction .............................................. 4
   1.1. Requirement Language and Terminology .................. 5
2. What is a URN Namespace? ................................. 5
3. URN Namespace (Registration) Types ..................... 7
   3.1. Experimental Namespaces ............................. 7
   3.2. Informal Namespaces ................................ 7
   3.3. Formal Namespaces ................................ 7
4. URN Namespace Registry: Processes for Registration and Update ........................................... 9
   4.1. Experimental Namespaces: No Registration ............. 10
   4.2. Informal Namespaces ................................ 10
   4.3. Formal Namespaces ................................ 11
   4.4. Registration Documents .............................. 12
      4.4.1. Namespace Considerations in Registration Documents .. 12
      4.4.2. Community Considerations in Registration Documents .. 13
      4.4.3. Security Considerations in Registration Documents .. 14
      4.4.4. IANA Considerations in Registration Documents .. 14
5. Security Considerations ................................. 15
6. IANA Considerations ...................................... 15
7. Acknowledgements ....................................... 16
8. References ........................................... 17
   8.1. Normative References ............................... 17
   8.2. Informative References ............................. 17
Appendix A. URN Namespace Definition Template ............. 18
Appendix B. Registration steps in practice ................ 24
Appendix C. Changes from RFC 3406 ......................... 25
   C.1. Essential Changes since RFC 3406 .................... 25
   C.2. Changes from RFC 3406 to URNBIS WG Draft -00 ......... 25
   C.3. Changes from URNBIS WG I-D -00 to -01 ................ 28
   C.4. Changes from URNBIS WG I-D -01 to -02 ................ 28
Appendix D. Issues in this Draft .......................... 28
1. Introduction

Uniform Resource Names (URNs) are resource identifiers adhering to the specific requirements of 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.

The URN Syntax (see below and RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn]) structures and organizes the entirety of URNs into a hierarchy that divides the set of possible URNs into "URN Namespaces" that can be individually defined, managed, and (optionally) further subdivided. 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).

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.

To actually leverage the potential synergetic advantage of this unification (structured embedding of existing namespaces into the URN framework), URN Namespaces need to be specified in a comparable manner, and their Namespace Identifiers (NIDs) need to be centrally registered, 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.

The purpose of this document is to outline a mechanism and provide a template for explicit URN 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 generic "resolution" of any so-created URN identifiers is desired, a separate process of registration in a global NID directory, such as that proposed by the DDDS system [RFC3401], is necessary. See [RFC3405] for information on obtaining registration in the DDDS global NID directory. There also is work in progress [Ref: t.b.d.] to establish an IANA registry for URN services, such that registration documents can unambiguously identify such services and discuss their applicability to the particular URN Namespace.

1.1. Requirement Language and Terminology

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 URN Namespace definition documents and registration templates.

For the purpose of this document, its subject is spelled "URN Namespace" (in headline case), whereas in other context, "namespace" is spelled in lower case, e.g. to designate a (standard or non-standard) identifier system on which a URN Namespace is based.

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 one resource. These resources may be stable (e.g., a doctoral dissertation or an abstract concept of a protocol) or dynamic (e.g., a continuously evolving web site of a periodical or a
specific protocol parameter registry subject to additions and maintenance). If the identified resource is a metadata record, such record may describe several objects (such as two versions of a book) or a collection of objects (such as a periodical with, say, monthly issues); in this case, these subordinate objects are not the identified resources. For each namespace, it must be clear what the identified resources are; if the namespace is heterogenous in this respect, the registration and resolution systems must unambiguously designate the kind of identified resource, for each identifier assigned in the namespace. Once assigned, URNs are never re-assigned to a different resource. A single resource, however, may have more than one URN assigned to it -- within a particular Namespace or among different Namespaces -- for different purposes, since the Namespaces are not mutually exclusive.

Such abstract namespace might be defined by some pre-established (standard or non-standard) identifier system that can be made "network-actionable" by embedding it into the URN framework using a specific URN Namespace. 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 (and making these designators an intrinsic syntactic part of URNs) 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, technology-specific vendor groups, 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 (types) 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. In both this Section and Section 4 these categories are ordered by increasing relevance/importance for the Internet and, accordingly, increasing strength of requirements for the definition and registration processes.

3.1. Experimental Namespaces

These are not explicitly registered with IANA.

No provision is made for avoiding collision of experimental NIDs; they are intended for use within internal or limited experimental contexts. However, as described below in Section 4.1, these are designated by a specific form of the NID to allow differentiation (without preexisting knowledge of details) from the other URN Namespace types.

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 to registered Informal Namespaces a simply structured, alphanumeric, ordinal NID (following a pattern defined in Section 4.2 below), 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 material. 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 however 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 with it 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, careful consideration must be given to the longevity and maintainability of the URN Namespace. 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.

In addition to the technical aspects of the Namespace and its resolution, consideration should be given to the following organizatorial aspects:

- the organization maintaining the URN Namespace should credibly 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;

- the organization(s) assigning URNs within the URN Namespace should demonstrate ability and competency in name assignment; this should improve the likelihood of persistence (e.g., to minimize the likelihood of conflicts);

- the organization(s) assigning URNs within the URN Namespace need to be committed to honor the scope, rules, and regulations outlined its registration document and the documents defining the underlying namespace and covering its identifier assignment and maintenance procedures (if any), and the organization maintaining the URN Namespace needs to have procedures in force that aim at ensuring this adherance at a very high confidence level; and
the involved organization(s) need to commit to not re-assign existing names; old names MUST continue to be valid, even if the owners or assignees of those names are no longer members or customers of such organization; this does not mean that there needs to be resolution of such names, but that they must not resolve such names to false or stale information and that they must not be reassigned.

If the underlying namespace is based on an established standard, the standards body or the organization(s) in charge with the maintenance of the namespace should be involved in the process, either by performing the registration on their own, or by supporting the action of the registrant and asserting support of the registration document.

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. The kind of mandate upon which the organization aims to undertake this activity might give a strong indication for this evaluation, because it likely mirrors the trust that other parties (for instance states, international treaty organizations, professionals’ associations, etc.) put on the organization.

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 [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 [IANA-URN].

The registry 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 SHOULD take the form:

X-<nid>

where <nid> is a string of up to 30 characters, consisting solely of letters, decimal digits, and hyphen ("-") characters, as specified by the NID syntax specification in Section 2.1 of RFC 2141bis [I-D.ietf-urnbis-rfc2141bis-urn].

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

Note: The above form is no more considered MANDATORY, in order to accommodate experience and demonstrated evidence that, under specific circumstances, experimental prototype systems have to create and assign identifiers that the interested community perceives are infeasible to be changed once the Namespace gets formally registered. However, it is strongly RECOMMENDED to prefix eventually targeted NIDs by "X-" during experiments and tests.

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

Usage of Experimental URN Namespaces MUST be short-lived and within a private scope; it MUST NOT be disclosed to the Internet at large, e.g. by distribution of software versions that make use of such.

4.2. Informal Namespaces

The NIDs of Informal Namespaces are synthesized by the 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 four-week discussion period for clarifying the expression of the registration information and suggestions for technical improvements to the Namespace proposal.  

[[ Editorial NOTE: An even longer time is needed in practice! Should we further increase the upper limit to 8 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 four-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.

NIDs for Formal URN Namespaces MUST NOT have the forms indicated in the preceding two sections for the other two Namespace types. (The detailed formal rules are given below in Section 4.4.4.) Applicants, in concert with the IANA experts, should ensure that the sought NID strings are "proper" for the designated purpose, according to common sense (and applicable legal rules).

"IETF Review" (per [RFC5226]) 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. The actual choice should be properly considered by applicants, but it is RECOMMENDED that the registration documents for NIDs belonging to an established standard namespace aim at Standards-Track, whereas other applications aim at Informational RFC.
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, but 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 four-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). Part of the expected elaborations need to be the arguments why other identifier systems, in particular a specific/new URI Scheme would not be suitable for the intended purpose.

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! In particular, the above list (from RFC 3406) seems to be rather orthogonal to the primary purpose of such section (as indicated in the first paragraph), namely to provide evidence for the perceived need for the new Namespace. ]]

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 lists 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? ]]

---

Hoenes                Expires September 13, 2012                [Page 13]
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);
- not be equal to or start with "example" (see NOTE below);
- be more than 2 characters 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 future 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; To allow neutral example URNs in code and documentation, NID strings starting with "example" are set aside for use in documentation; IANA has permanently reserved these string to prohibit assignment.

Applicants and the IANA experts have to ensure that the sought NID strings are suitable and proper for the designated purpose and not misleading, according to common sense and applicable legal rules. The IETF Review process gives interested parties the opportunity to rise concerns if they want to challenge proposed strings; the final approval decision still remains with the IESG.

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 four-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.

Until recently, that registry has been available in HTML, XML, and plain text from the generic web page at <http://www.iana.org/assignments/urn-namespaces/> [IANA-URN].
[[ 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>. Yet, currently this URI and similar forms all resolve to an XML version. The content there should link to alternate forms (.xml, .txt), and those alternate versions should indicate the *other* versions; i.e., where the .txt version (currently only available at ftp.IANA.ORG) 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." Similarly, the XML form should point to the HTML and plain text forms. ]]

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

We would appreciate a reorganization of the Registry web page to make the registration templates for Informal URN Namespaces directly linked from the main page; this would make the page /assignments/urn-informal.htm page dispensable (for persistency’s sake, the web server should redirect requests to the /assignments/urn-namespaces page.

Section 4 of this document outlines the general procedures.

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. All strings starting with "example" are permanently reserved for use in code and documentation, and hence 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 authors 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.

Several individuals in the URNbis working group have participated in the detailed discussion of this memo. Particular thanks for detailed review comments and text suggestions go to Juha Hakala, Peter Saint-Andre, and Mykyta Yevstifeyev.

8. References

8.1. Normative References


8.2. Informative References


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:

{ 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 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:

- 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:

- 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:

- 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 2141bis, in order to avoid repeated confusion, this version of the template tentatively restricts this clause to only the NSS part of the newly specified URNs.
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.

However, namespaces that seek to provide higher-level lexical equivalence rules should preferably make use of established and standardized normalization procedures (like the methods leading to the various Unicode Normalization Forms, which would have to be applied before UTF-8 encoding) and not invent their own "magic"; in practice, the utility of such things is likely to be limited since test of lexical equivalence is a typical client-side pre-screening operation performed by applications that try to remain as general as possible and typically will not have built-in, NID-specific knowledge -- ultimately, functional (or semantical) equivalence of URNs can only be decided in the NID-specific assignment/resolution systems, and their internal rules can be handled much more flexibly than more complicated, nailed-down lexical equivalence rules that are unlikely to be implemented at large.

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, i.e. the precise kind of resources to which the URNs are assigned. 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. 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" and "Community Considerations" sections (or a combined section for these), "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);
o 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;

o 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;

o 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;

o Section 4.4.1: made requirements a "MUST";

o 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;

o 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;

o 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);

o Section 6 (IANA Considerations): updated and rephrased description of the role of this document, including a sketch of the history; added text 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 template; 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 invariable 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.
C.3. Changes from URNbis WG I-D -00 to -01

Usage of terminology strengthened.

Clarified role and usage of Experimental Namespaces.

Clarified NID strings for Formal Namespaces.

Added hint that recommends Std. Track RFCs for NID applications based on established standard namespaces, and Informational for others.

Changed standard review period from 2 to 4 weeks (pending discussion).

Resolved with IANA: simple, traditional and generic URL used by IANA for the URN Namespace registry. (Needed to be re-opened in -02!)

Numerous editorial enhancements and fixes.

C.4. Changes from URNbis WG I-D -01 to -02

General text edits based on evaluation of meeting and on-list comments.

Updated and tightened the organizational requirements for Formal Namespace requests.

Restored additional IANA Considerations -- due to observed defects.

Reserved NID strings "example.*" for documentation (as suggested by Larry Masinter, Peter Saint-Andre, and Julian Reschke).

Added text on possible "higher level" methods to establish lexical equivalence of URNs, with the caveats that such things are rather unlikely to get traction in general-purpose client software.

Removed historical Appendix B.1 (Example Template).

Various editorial enhancements and fixes.

Updated and expanded "Issues" Appendix (below) in preparation of usage of the IETF Issue Tracker.

Appendix D. Issues in this Draft

[ Appendix to be replaced by use of IETF Tools issue tracker. ]

For more details on the issues below, please also see the Editorial Notes interspersed in the body of this draft.
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.

Do registration documents need more guidance and be caused to be more precise in their elaboration on the applicability of Services? Since RFC 2483 is considered outdated, but RFC 2483bis not yet alife (nor 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.

Do we actually need Experimental Namespaces? [Regarded as CLOSED affirmatively at IETF 80.]

There are concerns regarding usage of "X-" NIDs, which is reported to having proven impractical in practice. This draft version contains tentative text to address these concerns; "X-" is now demoted to "SHOULD" level.

The syntax of the NID strings for the various NID types is given in an informal manner (as has been done in RFC 3406); is it worth the effort to introduce ABNF for this purpose? [The request for ABNF has been voiced only once; the document Editor regards this issue as CLOSED.]

Increase review/timeout periods for urn-nid list and IANA experts from 2 to 4 (or more) weeks? This draft version tentatively specifies 4 weeks.
Juha Hakala has argued that the assessment of the responsible organizations needed to assure their ability to properly operate the Namespace could never be performed within the present 2 weeks time span; 8 weeks might be an even better choice for the future upper limit for the review period. It has been pointed out that even 8 weeks are miniscule with regard to the expected lifetime of the to-be-registered Namespace and hence should not matter. In practice, the subsequent IESG evaluation of URN Namespace registration documents has typically needed much longer time.

Clarification of the desired content of the "Namespace Considerations" and "Community Considerations" sections in registration documents. Shall we admit a combined section for both topics? (so far supported by 2 postings) Cf. the NOTEs in Sections 4.4.1 and 4.4.2 for more details. [No feedback on the list since -01, so the draft text seems to have silent consensus and the issue is regarded as CLOSED.]
Shall other strings beyond "urn" also be 'reserved' in the NID registry? (e.g. "uri", "url", "urc", ...) There have been voices in favor of leaving the decision of what is acceptable and reasonable in practice to the common sense of prospective registrants and the designated IANA experts. This draft version reserves NID strings matching the RE "^example.*" for documentation.

Appendix A: Once RFC 2483 gets updated and an IANA registry for URN resolution services gets established, the "Process for identifier resolution" clause in the registration template should call out for enumerating the registered services that are applicable for the newly defined URN Namespace. How far can we go in this respect without an update to RFC 2483 at hands?

Do we really still need Appendix B.1? (There are lots of real-life examples now!)
[ Old B.1 removed, old B.2 became Appendix B; ==> CLOSED ]

Author’s Address

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

EMail: ah@TR-Sys.de