

Network Working Group
Request for Comments: 2611
BCP: 33
Category: Best Current Practice

L. Daigle
Thinking Cat Enterprises
D. van Gulik
ISIS/CEO, JRC Ispra
R. Iannella
DSTC Pty Ltd
P. Faltstrom
Tele2/Swipnet
June 1999

URN Namespace Definition Mechanisms

Status of this Memo

This document specifies an Internet Best Current Practices for the Internet Community, and requests discussion and suggestions for improvements. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (1999). All Rights Reserved.

Abstract

The URN WG has defined a syntax for Uniform Resource Names (URNs) [RFC2141], as well as some proposed mechanisms for their resolution and use in Internet applications ([RFC2168, RFC2169]). The whole rests on the concept of individual "namespaces" within the URN structure. Apart from proof-of-concept namespaces, the use of existing identifiers in URNs has been discussed ([RFC2288]), and this document lays out general definitions of and mechanisms for establishing URN "namespaces".

1.0 Introduction

Uniform Resource Names (URNs) are resource identifiers with the specific requirements for enabling location independent identification of a resource, as well as longevity of reference. There are 2 assumptions that are key to this document:

Assumption #1:

Assignment of a URN is a managed process.

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

Assumption #2:

The space of URN namespaces is managed.

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

The purpose of this document is to outline a mechanism and provide a template for explicit namespace definition, along with the mechanism for associating an identifier (called a "Namespace ID", or NID) which is registered with the Internet Assigned Numbers Authority, IANA.

Note that this document restricts itself to the description of processes for the creation of URN namespaces. If "resolution" of any so-created URN identifiers is desired, a separate process of registration in a global NID directory, such as that provided by the NAPTR system [RFC2168], is necessary. See [NAPTR-REG] for information on obtaining registration in the NAPTR global NID directory.

2.0 What is a URN Namespace?

For the purposes of URNs, a "namespace" is a collection of uniquely-assigned identifiers. 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, ISBNs and ISSNs are both collections of identifiers used in the traditional publishing world; while there may be some number (or numbers) that is both a valid ISBN identifier and ISSN identifier, using different designators for the two collections ensures that no two URNs will be the same for different resources.

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

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

3.0 URN Namespace Definition Template

Definition of a URN namespace is accomplished by completing the following information template. Apart from providing a mechanism for disclosing 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)
- 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.

Information in the template is as follows:

Namespace ID:

Assigned by IANA. In some contexts, a particular one may be requested (see below).

Registration Information:

This is information to identify the particular version of registration information:

- registration version number: starting with 1, incrementing by 1 with each new version
- registration date: date submitted to the IANA, using the format YYYY-MM-DD as outlined in [IS08601].

Declared registrant of the namespace:

Required: Name and e-mail address.

Recommended: Affiliation, address, etc.

Declaration of syntactic structure:

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

Relevant ancillary documentation:

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

Answers might include, but are not limited to:

- RFCs outlining syntax of the namespace
- Other of the defining community's (e.g., ISO) documents outlining syntax of the identifiers in the namespace
- Explanatory material introducing the namespace

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 which are individually responsible for respecting uniqueness rules
- identifiers are assigned sequentially
- information is withheld; 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 [RFC2276]) such as NAPTR. 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

Rules for Lexical Equivalence:

If there are particular algorithms for determining equivalence between two identifiers in the underlying namespace (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.

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 hex encoding the symbol.

For example, see the section on SICIs in [RFC2288].

Validation mechanism:

Apart from attempting resolution of a URN, a URN namespace may provide mechanism for "validating" a URN -- i.e., determining whether a given string is currently a validly-assigned URN. For example, even if an ISBN URN namespace is created, it is not clear that all ISBNs will translate directly into "assigned URNs".

A validation mechanism might be:

- a syntax grammar
- an on-line service
- an off-line service

Scope:

This section should outline the scope of the use of the identifiers in this namespace. Apart from considerations of private vs. public namespaces, this section is critical in evaluating the applicability of a requested NID. For example, a namespace claiming to deal in "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".

4.0 URN Namespace Registration, Update, and NID Assignment Process

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 or may request a particular identifier. The [RFC2434] document 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.

There are 3 categories of URN namespaces defined here, distinguished by expected level of service and required procedures for registration. Furthermore, registration maintenance procedures vary slightly from one category to another.

- I. **Experimental:** These are not explicitly registered with IANA. They take the form

X-<NID>

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

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

- II. **Informal:** These are registered with IANA and are assigned a number sequence as an identifier, in the format:

"urn-" <number>

where <number> is chosen by the IANA on a First Come First Served basis (see [RFC2434]).

Registrants should send a copy of the registration template (see section 3.0), duly completed, to the

urn-nid@apps.ietf.org

mailing and allow for a 2 week discussion period for clarifying the expression of the registration information and suggestions for improvements to the namespace proposal.

After suggestions for clarification of the registration information have been incorporated, the template may be submitted to:

iana@iana.org

for assignment of a NID.

The only restrictions on <number> are that it consist strictly of digits and that it not cause the NID to exceed length limitations outlined in the URN syntax ([RFC2168]).

Registrations may be updated by the original registrant, or an entity designated by the registrant, by updating the registration template, submitting it to the discussion list for a further 2 week discussion period, and finally resubmitting it to IANA, as described above.

- III. Formal: These are processed through an RFC review process. The RFC need not be standards-track. The template defined in section 3.0 may be included as part of an RFC defining some other aspect of the namespace, or it may be put forward as an RFC in its own right. The proposed template should be sent to the

urn-nid@apps.ietf.org

mailing list to allow for a 2 week discussion period for clarifying the expression of the registration information, before the IESG progresses the document to RFC status.

A particular NID string is requested, and is assigned by IETF consensus (as defined in [RFC2434]), with the additional constraints that the NID string must

- not be an already-registered NID
- not start with "x-" (see Type I above)
- not start with "urn-" (see Type II above)
- not start with "XY-", where XY is any combination of 2 ASCII letters (see NOTE, below)
- be more than 2 letters long

NOTE: ALL two-letter combinations, and two-letter combinations followed by "-" and any sequence of valid NID characters, are reserved for potential use as countrycode-based NIDs for eventual national registrations of URN namespaces. The definition and scoping of rules for allocation of responsibility for such namespaces is beyond the scope of this document.

Registrations may be updated by updating the RFC through standard IETF RFC update mechanisms. Thus, proposals for updates may be made by the original authors, other IETF participants, or the IESG. In any case, the proposed updated template must be circulated on the urn-nid discussion list, allowing for a 2 week review period.

URN namespace registrations will be posted in the anonymous FTP directory "ftp://ftp.isi.edu/in-notes/iana/assignments/URN-namespaces/".

5.0 Example

The following example is provided for the purposes of illustration of the URN NID template described in section 3.0. Although it is based on a hypothetical "generic Internet namespace" that has been discussed informally within the URN WG, there are still technical and infrastructural issues that would have to be resolved before such a namespace could be properly and completely described.

Namespace ID:

To be assigned

Registration Information:

Version 1

Date: <when submitted>

Declared registrant of the namespace:

Required: Name and e-mail address.

Recommended: Affiliation, address, etc.

Declared registrant of the namespace:

Name: T. Cat
E-mail: leslie@thinkingcat.com
Affiliation: Thinking Cat Enterprises
Address: 1 ThinkingCat Way
Trupville, NewCountry

Declaration of structure:

The identifier structure is as follows:

URN:<assigned number>:<FQDN>:<assigned US-ASCII string>

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

Relevant ancillary documentation:

Definition of domain names, found in:

P. Mockapetris, "DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION", RFC1035, November 1987.

Identifier uniqueness considerations:

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

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

Identifier persistence considerations:

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

Same note as above.

Process of identifier assignment:

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

e.g. urn:<assigned number>:thinkingcat.com:001203

Process for identifier resolution:

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

Rules for Lexical Equivalence:

FQDNs are case-insensitive. Thus, the portion of the URN

urn:<assigned number>:<FQDN>:

is case-insensitive for matches. The remainder of the identifier must be considered case-sensitive.

Conformance with URN Syntax:

No special considerations.

Validation mechanism:

None specified.

Scope:

Global.

6.0 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.

7.0 References

- [RFC2168] Daniel, R. and M. Mealling, "Resolution of Uniform Resource Identifiers using the Domain Name System", RFC 2168, June 1997.
- [RFC2169] Daniel, R., "A Trivial Convention for using HTTP in URN Resolution", RFC 2169, June 1997.
- [ISO8601] ISO 8601 : 1988 (E), "Data elements and interchange formats - Information interchange - Representation of dates and times"
- [RFC2288] Lynch, C., Preston, C. and R. Daniel, "Using Existing Bibliographic Identifiers as Uniform Resource Names", RFC 2288, February 1998.
- [NAPTR-REG] Mealling, M., "Assignment Procedures for NAPTR DNS URI Resolution", Work in Progress.
- [RFC2141] Moats, R., "URN Syntax", RFC 2141, May 1997.
- [RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.
- [RFC1737] Sollins, K. and L. Masinter, "Functional Requirements for Uniform Resource Names", RFC 1737, December 1994.
- [RFC2276] Sollins, K., "Architectural Principles of Uniform Resource Name Resolution", RFC 2276, January 1998.

8.0 Authors' Addresses

Leslie L. Daigle
Thinking Cat Enterprises

EMail: leslie@thinkingcat.com

Dirk-Willem van Gulik
ISIS/STA/CE0 - TP 270
Joint Research Centre Ispra
21020 Ispra (Va)
Italy.

Phone: +39 332 78 9549 or 5044
Fax: +39 332 78 9185
EMail: Dirk.vanGulik@jrc.it

Renato Iannella
DSTC Pty Ltd
Gehrmann Labs, The Uni of Queensland
AUSTRALIA, 4072

Phone: +61 7 3365 4310
Fax: +61 7 3365 4311
EMail: renato@dstc.edu.au

Patrik Faltstrom
Tele2/Swipnet
Borgarfjordsgatan 16
P.O. Box 62
S-164 94 Kista
SWEDEN

Phone: +46-5626 4000
Fax: +46-5626 4200
EMail: paf@swip.net

9.0 Full Copyright Statement

Copyright (C) The Internet Society (1999). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.