

A Uniform Resource Name (URN) Formal Namespace
for the New Zealand Government

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2006).

Abstract

This document describes a Uniform Resource Name (URN) Namespace Identification (NID) convention as prescribed by the World Wide Web Consortium (W3C) for identifying, naming, assigning, and managing persistent resources and XML artefacts for the New Zealand Government.

1. Introduction

The New Zealand Government has already adopted XML as its primary means of storing and exchanging data. The New Zealand Government publishes documents, schemas, and other government artefacts.

The New Zealand Government now wishes to define a namespace convention and structure for its agencies by creating and managing globally unique, persistent, location-independent identifiers for their schema resources and XML artefacts.

This is a natural extension of the development of the Dublin Core based New Zealand Government metadata standard (New Zealand Government Locator Service, or NZGLS) used by government agencies to create metadata and made operational to the public through an all-of-government portal.

The New Zealand Government wishes to provide guidance on namespaces to its agencies so that they use a portion of the adopted namespace to minimise the risk of their creating different (and potentially conflicting) namespace structures. This issue potentially extends to

data exchange beyond government into the private sector of New Zealand, thus placing the government under an obligation to provide guidance in the assignment and management of additional namespaces.

The New Zealand Government wishes to register the country NID, NZL, with the Name Specific String (NSS) split into two parts; the first part being a specific sub-type <nz-specifier> and the second part as a <nz-specifier defined string>.

As part of the URN structure, the New Zealand Government wishes to define and subsequently manage the "govt" specifier. It will also assign additional specifiers requested by other New Zealand organisations in accordance with the rules and processes proposed herein.

The New Zealand Government hoped to make use of the two-letter Namespace Identifier (NID) combination for its ubiquitous country code, NZ. But since there is as yet no process to register these (see RFC 3406 [1]) the government has opted to request its well-known alternative three-letter country code (see ISO 3166 [3]).

This namespace specification requests a formal namespace (see [6] for more information about formal namespaces).

Please note that this paper includes a discussion on the use of diacritic marks, in particular, Maori macrons. Maori is an official language of New Zealand. In recognition of the established practice of publishing RFCs for a global audience in ASCII text where diacritic marks are unable to be recognised, the text has been presented without macrons.

2. Specification Template

Namespace ID:

"NZL".

Registration Information:

Version Number: 1

Date: 2005-03-31

Declared registrant of the namespace:

State Services Commission
New Zealand Government
100 Molesworth Street
Wellington,
New Zealand

Email: e-GIF@ssc.govt.nz

Declaration of structure:

The identifier has a hierarchical structure as follows:

urn:nzl:<nz-specifier>[:<nz-specifier defined string>]+

+ denotes one or more occurrences of nz-specifier defined strings all delimited by a colon.

For example:

urn:nzl:govt:registering:dogs:registration:1-0

urn:nzl:govt:registering:firearms:form:1-3

urn:nzl:govt:registering:recreational_fishing:form:1-0

The <nz-specifier> and <nz-specifier defined string> can comprise any UTF-8 characters compliant with URI syntax and must not contain the ":" character (see STD 66, RFC 3986 [2]). The exclusion of the colon from the list of other characters means that the colon can only occur as a delimiter between string values. The values come from the terms listed in the NZGLS.

The State Services Commission (SSC) will take responsibility for the <nz-specifier> "govt" and its sub level <nz-specifier defined string> terms; e.g., "registering".

The SSC will take responsibility to assign other <nz-specifiers> to organisations who apply and can satisfy the SSC that they have the capability to manage the sub level and its applicable <nz-specifier defined string(s)>.

Relevant ancillary documentation:

The function and noun syntax used in the <nz-specifier defined string> is based on and taken from the NZGLS (<http://www.e.govt.nz/standards/nzgl/thesauri/>).

Identifier uniqueness considerations:

Identifiers in the <nz-specifier> "govt" are defined and assigned in the requested namespace by the SSC after ensuring that the URNs to be assigned are unique. Uniqueness is achieved by checking against the registry of previously assigned names.

The SSC will ensure that the URNs to be assigned to other organisations applying for other <nz-specifier(s)> (e.g., mil, co, org) are unique by checking against the registry of previously assigned names.

The SSC will develop and publish the process for doing this, which, where applicable, is consistent with the process it uses for moderating the .govt.nz Top Level Domain (TLD).

Identifier persistence considerations:

The New Zealand Government is committed to maintaining uniqueness and persistence of all resources identified by assigned URNs.

Given that the URN sought is NZL (the long-held ISO 3166 Alpha-3 representation of the country) and that the country's independence from any other jurisdiction expected to continue indefinitely, the URN should also persist indefinitely.

Likewise, the <nz-specifier> "govt" has a very long life expectancy and can be expected to remain unique for the foreseeable future. The assignment process guarantees that names are not reassigned. The binding between the name and its resource is permanent.

The SSC will ensure that other organisations applying to manage other <nz-specifier> Second Level Name (2LN) sub-levels of the NZL URN namespace (e.g., mil, co, org) uniquely assign the namespace at this level.

Process of identifier assignment:

Under the "NZL" NID, the New Zealand Government will manage the <nz-specifier> "govt" and leverage the existing NZGLS thesaurus for identifier resources to maintain uniqueness.

The process of assigning URNs at the <nz-specifier> sub-level will be managed by the SSC of the New Zealand Government. (The SSC has managed and maintained the NZGLS thesauri since its inception in 2002 and has moderated the TLD .govt.nz).

The SSC will develop and publish the process for doing this, which is consistent with the process it uses for moderating the .govt.nz TLD, where applicable. The process for marketing the ".govt.nz" TLD can be found at these links:

<http://www.e.govt.nz/moderation/mod-policy/chapter1.html>

and

<http://www.e.govt.nz/moderation/mod-policy/chapter2.html>

The process is drawn from the 2LD policies and procedures of the New Zealand Office of the Domain Name Commissioner, <http://dnc.org.nz> (and specifically <http://www.dnc.org.nz/story/30043-35-1.html>).

Other New Zealand organisations may apply to the SSC to delegate specifiers for resolution and management assigned by them. Delegation of this responsibility will not be unreasonably withheld provided that the processes for their resolution and management are robust and are followed.

Organisations who apply to have a <nz-specifier> assigned to them must satisfy the SSC that they have the capability to manage the 2LN sub-level and its applicable <nz-specifier defined string(s)> responsibly. The policies and procedures in the links above will be provided to applicants as a guide and will be used by the SSC to determine the applicant's capability.

Process of identifier resolution:

For the <nz-specifier> "govt", the SSC will maintain lists of assigned identifiers on its web pages at <http://www.e.govt.nz/>.

The SSC will require other organisations that apply to manage other <nz-specifier> sub-levels to follow this practice unless there are specific reasons (e.g., security) not to do so.

Rules for Lexical Equivalence:

The lexical equivalence of the NZL namespace-specific strings (NSSs) is defined as an exact, but not case-sensitive, string match. Best Practice guidelines will specify:

- a) NZL in either uppercase or lowercase (The New Zealand government will assign names as case-insensitive, to ensure that there will not be two NZL URNs differing only by case.)
- b) The first letter of each <nz-specifier> and <nz specifier defined string> in uppercase or the whole value in lowercase.
- c) Any identifier in NZL namespaces can be compared using the normal mechanisms for percent-encoded UTF-8 strings.

Note that textual data containing diacritic marks (such as Maori macrons) will not be treated as lexically equivalent to textual data without diacritic marks; i.e., a distinction will be made. It is important to note that a macron can change the meaning of a word in the Maori language.

The following explanation provides guidance in this respect.

NSS is any UTF-8-encoded string that is compliant with the URN syntax (i.e., following the encoding rules for 8-bit characters). Since Maori is an official language in New Zealand and its use of diacritic marks (in this case macrons) invokes the requirement to percent-encode reserved characters, the following extract from RFC 3986 [4] is applicable.

When a new URI scheme defines a component that represents textual data consisting of characters from the Universal Character Set [UCS], the data should first be encoded as octets according to the UTF-8 character encoding [STD63]; then only those octets that do not correspond to characters in the unreserved set should be percent-encoded. For example, the character A would be represented as "A", the character LATIN CAPITAL LETTER A WITH GRAVE would be represented as "%C3%80", and the character KATAKANA LETTER A would be represented as "%E3%82%A2".

As described above, UTF-8 allows the use of diacritic marks such as New Zealand Maori macrons.

In the New Zealand context, the word "Maori" carries a diacritic mark over the "a". A URI including the macronised word "Maori" would be percent-encoded as M%C4%81ori.

Given that the "govt" namespaces will draw from the NZGLS thesaurus (which does not at present utilise diacritic marks), the "govt" <nz-specifier> will not utilise UTF-8's percent-encoding convention for diacritic marks. An "a" with a diacritic mark will be presented simply as an "a". There is no mapping or equivalence table. Therefore, the requirement to distinguish between terms that have diacritic marks and those that do not will not arise in the <nz-specifier> "govt".

Other organisations may use diacritic marks with certain conditions. Organisations that apply to manage other <nz-specifier> sub-levels of the NZL URN namespace could utilise UTF-8's diacritic functionality provided that they have the applicable processes to separate Maori language terms using macrons from those that do not, in order to ensure uniqueness in accordance with rule c) above.

Conformance with URN Syntax:

No special considerations.

Validation mechanism:

None other than names being derived from the NZGLS thesaurus "dictionary".

Scope:

Global, but primarily of national interest.

3. Namespace Considerations

The SSC undertook a preliminary study of the URI alternatives against the key requirements. The options were narrowed down to five. These were a private URI scheme, URL, PURL, IRI, and URN. URN was considered the most appropriate URI against the criteria.

Consultation on the preliminary study was actively sought from the Internet Society of NZ (InternetNZ), the NZ Computer Society, applicable vendors, and government agencies. Publication on the e-government web site allowed for public participation.

Points that should be noted are:

- a) With respect to the NID, the New Zealand Government is the first known jurisdiction to apply its globally known ISO 3166 Alpha-3 country code to become a URN. One objective of the ISO 3166 Alpha-2 and 3-letter country codes was to provide uniqueness.

- b) The namespace follows the logical structure of the NZGLS as shown in the examples above.

4. Community Considerations

Providers of government information for data exchange benefit by the publication of the namespace because it provides much-needed guidance on generating target namespaces for schema development using a process that reflects what they already know; namely, metadata creation in NZGLS. The identifiers under the "govt" specifier will track the terms used in the New Zealand government thesaurus.

Consequently, New Zealanders will ultimately benefit since the exchange of more structured information will potentially improve online experiences in areas such as forms design.

Any citizen or organisation with Internet web browser capability will be entitled to access the namespace and its associated application, registration, and resolution services. While the assignment of identifiers will be managed by the SSC, additional specifiers (such as mil, co, org, and their <nz-specifier defined string(s)>) can be openly applied for and registered by anyone following an approved namespace governance process and proof of the applicant's bona fide association with the intended specifier (i.e., no squatting or hoarding).

5. IANA Considerations

This document includes a URN NID registration for NZL for entry in the IANA registry of URN NIDs (see RFC 2434 [5] for more information).

6. Security Considerations

No serious security implications are envisaged beyond the potential threat of spoofing. The application, registration and assignment of subsequent specifiers will leverage existing government processes to authenticate the applicants and their association with the proposed specifier application.

7. Acknowledgements

Since the specification described in this document is derived from STD 66, RFC 3986 and RFC 3406, the acknowledgements in those documents still apply. In addition, the authors wish to acknowledge Leslie Daigle and Ted Hardie for their suggestions and review.

8. References

8.1. Normative References

- [1] Daigle, L., van Gulik, D., Iannella, R., and P. Faltstrom, "Uniform Resource Names (URN) Namespace Definition Mechanisms", BCP 66, RFC 3406, October 2002.
- [2] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax", STD 66, RFC 3986, January 2005.
- [3] ISO 3166, "Country name codes", ISO 3166-1:1997.

8.2. Informative References

- [4] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.
- [5] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.
- [6] URI Planning Interest Group, W3C/IETF (See acknowledgments) September 2001,
<<http://www.w3.org/TR/2001/NOTE-uri-clarification-20010921/>>.

Authors' Addresses

Ferry Hendrixx
Information & Communications Technology (ICT) Branch
State Services Commission
PO Box 329
Wellington
New Zealand

Phone: +64 4 495 6600
EMail: ferry.hendrixx@ssc.govt.nz

Colin Wallis
Information & Communications Technology (ICT) Branch
State Services Commission
PO Box 329
Wellington
New Zealand

Phone: +64 4 495 6600
EMail: colin.wallis@ssc.govt.nz

Full Copyright Statement

Copyright (C) The Internet Society (2006).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM 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.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgement

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).

