Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 27, 2016

S. Leonard
Penango, Inc.

September 24, 2015

URN Namespaces for XML Namespaces and RDF IRIs
draft-seantek-xmlns-rdf-urns-01

Abstract

XML segregates elements into namespaces, which can be used to mix tags with different semantics in a composite XML document. XML namespaces are identified by URIs (XML 1.0) or IRIs (XML 1.1). Similarly, RDF contains "nodes" that are identified by "URI references" (RDF 1.0) or "IRIs" (RDF 1.1). This document defines URNs specifically for XML namespaces and RDF nodes.

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 March 27, 2016.

Copyright Notice

Copyright (c) 2015 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.

1. Introduction

XML segregates elements into namespaces, which can be used to mix tags with different semantics in a composite XML document. XML namespaces are identified by URIs [XML1.0] or IRIs [XML1.1]. Similarly, RDF identifies nodes by "URI reference" [RDF1.0] or "IRI" [RDF1.1]. This document defines URN namespace identifiers (NIDs) specifically for identifying XML namespaces and RDF nodes.

Experience suggests that several URN namespace registrations have been proposed over the years, where the primary (yet only occasionally stated) purpose is to create concise, targeted resource identifiers for XML or RDF use. An designer now has the option of choosing a concise, mnemonic identifier without the cost of maintaining and relying upon a long-lived network location (such as an HTTP URL), and without the hassle of registering a URN namespace identifier via IETF Consensus. One exemplary advantage of using an xmlns URN over an HTTP URL is that even if the organization hosting the original XML schema ceases to exist, the URI will remain functional without needing to commandeer back an embedded domain name.

A name in the urn:xmlns namespace uniquely and persistently identifies an abstract XML namespace resource. The abstract resource does not have any particular concrete representation (such as a type of content identified by Internet media type), although concrete representations (referenced by URIs) may be associated with it as discussed in Section 4. Abstract parts of the abstract resource can be identified with fragment identifiers.

A name in the urn:rdf namespace uniquely and persistently identifies an abstract RDF node resource ("URI reference" per [RDF1.0], or "IRI" per [RDF1.1]). The abstract resource does not have any particular concrete representation (such as a type of content identified by Internet media type), although concrete representations (referenced by URIs) may be associated with it. Abstract parts of the abstract resource can be identified with fragment identifiers.

Two distinct namespaces are valuable because each namespace represents a different abstract resource type, which can have type-specific concrete representations. For example, "urn:xmlns:foobar2015" could represent some foobar2015 XML namespace, with associated XML schema definitions. In contrast, "urn:rdf:foobar2015" could represent some RDF node (or some RDF node prefix), with associated description resources.

1.1. Definitions

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.

The term "divider" means a punctuation character such as colon ":" that serves to divide a name into parts.

The term "part" means the characters between divisions (dividers or the ends of the name).

While a name "divided" into "parts" can be seen as forming a tree or hierarchy, names in this specification are not hierarchical in the [RFC3986] sense, and have no hierarchical semantic content. Names SHOULD be treated as opaque identifiers that are only compared for equality.

The term "initial prefix" means the first part of a name.

The term "DIGIT" is from the ABNF production in [RFC5234], i.e., the US-ASCII characters 0 through 9.

2. Registration Template for xmlns NID

```
Namespace ID:
   xmlns

Registration Information:
   Version: 1
   Date: 2015-09-24

Declared registrant of the namespace:
   IETF

Declaration of syntactic structures:
   An xmlns name is any
   valid XML name corresponding to "Name" in Section 2.3
   of [XML1.0] (production 5), with the following restrictions:
      1. The name MUST be at least four characters.
      2. Colons MAY be used as intra-name dividers.
      3. Colons MUST NOT appear at the beginning or end of the name.
      4. Consecutive colons MUST NOT appear.
   and the following relaxation:
      5. The first part of the name preceding the first colon MAY
         be comprised of DIGITs (which are intended to correspond
         to registered IANA Private Enterprise Numbers);
         further discussion is in
```

"Process of identifier assignment".

The ABNF of a name is:

```
urn-xmlns-name = [DIGIT+ ":"] NoColonNameStartChar
                 *([":"] NoColonNameChar)
```

Where the productions NoColonNameStartChar and NoColonNameChar are respectively taken from NameStartChar (production 4) and NameChar (production 4a) in [XML1.0], with ":" omitted. Although ABNF is formally US-ASCII only, the domain of this production includes the whole Unicode range.

The name is used as the basis of the Namespace Specific String (NSS) as follows. When the NSS is encoded in a URN in a URI protocol slot, Unicode code points beyond U+007F are encoded as percent-encoded UTF-8. Conveniently, all XML name characters in the US-ASCII range are in the [RFC3986] unreserved set. When the NSS is encoded in a URN in a IRI protocol slot, Unicode code points beyond U+007F in the unreserved set are encoded as-is; they MUST NOT be percent-encoded.

Relevant ancillary documentation:
[XML1.0], [XML1.1].

Identifier uniqueness considerations:
The meaning of an identifier is registered in the registry, and thus is unique.

Identifier persistence considerations:
Once an identifier is registered, its meaning cannot be changed.

Process of identifier assignment:
Identifiers are registered with IANA on a First-Come, First-Served basis. One-character initial prefixes are reserved for further use. Two- and three-character initial prefixes are intended to correlate with language tags and regional codes; however, they have no such semantic content when used in an xmlns name. Whole number initial prefixes are intended to represent IANA Private Enterprise Numbers.
Registrants are free to register names with reserved two-character and three-character initial prefixes, such as "au:flag" or "en:us:ca:lax". Registrants are also free to register names with whole number prefixes, such as "20:10-250": these names have a particular registration process since they implicitly relate relate to IANA Private Enterprise Numbers.

The IANA Considerations section fully defines the registration processes.

Process for identifier resolution:
The registration for a particular identifier MAY include any number of URIs that a URN resolver MAY use to resolve the URN to return specific resources. The registered URIs are not equivalent to the registered URN, so an XML document that refers to that particular namespace MUST use the registered URN as the XML namespace URI. IANA will maintain the resolution database--see Section 4.
A URN resolver SHALL pass any [RFC3986] fragment component in the urn: URI or IRI through to the resolved URI if the registered URI does not have a fragment component. See [URNBIS].
If the registered URI has a fragment component, a URN resolver SHALL NOT pass any [RFC3986] fragment component in the urn: URI or IRI; the fragment component SHALL be ignored.

Rules for Lexical Equivalence:
The NSS is compared case sensitively.
If a URI and a IRI are compared against each other, the UTF-8 percent-encoded octets in the URI representing code points in the unreserved set beyond U+007F SHALL be treated as the Unicode code points in the IRI. An IRI that contains UTF-8 percent-encoded octets in the unreserved set beyond U+007F is not supposed to exist; it is a protocol error.
Fragments (delimited by the # character) are not considered part of the name, the NSS, or the URN, so a fragment would not affect lexical equivalence. Nevertheless, a urn: URI or IRI might be produced with a fragment component.

Conformance with URN Syntax:
The URN of this namespace conforms to new URN Syntax [URNBIS], old URN syntax [RFC2141], and Uniform Resource Identifier (URI): Generic Syntax [RFC3986].

Validation mechanism:
An XMLNS URN may be validated by looking it up in the IANA Registry.

Scope:
Global.

3. Registration Template for rdf NID

Namespace ID:
rdf

Registration Information:
Version: 1
Date: 2015-09-24

[[TODO: paste xmlns content and make rdf substitutions.]]

Scope:
Global.

4. IANA Considerations

This document requests the assignment of formal URN namespace IDs "xmlns" and "rdf".

This document requests the creation of an IANA registry called "urn:xmlns Names", and an IANA registry called "urn:rdf Names". The registries are First-Come, First-Served [RFC5226]. Each registration shall contain:

a. the name conforming to this document

   1) in Unicode characters and

   2) with characters beyond U+007F percent-encoded in UTF-8,

b. an optional description,

c. optional [RFC3986] conforming URIs that are not URNs that are to be used for URN resolution, and

d. contact information for the registrant.

After submitting a new registration, the registration SHALL be effective and entered into the registry after 120 hours (5 days). Registrants or their successors may update their entries from time to time. After submitting an updated registration, the registration SHALL be effective and entered into the registry after 48 hours (2 days). During the waiting period, a registrant MAY withdraw the proposed registration for any reason.

The registration template SHALL be encoded in UTF-8.

If a registrant attempts to register a name that is confusingly similar to other registered names (such as only differing by case, or differing by code points but generating the same or confusingly

similar visual representations), the registrants of the prior names are to receive a warning notification of the registration. IANA SHOULD implement a modern algorithm to detect such confusingly similar names.

If a registrant attempts to register a name that contains a whole number initial prefix, the number MUST correspond to an existing IANA Private Enterprise Number. The registrant of the corresponding IANA Private Enterprise Number is to receive a notification of the impending registration. The PEN registrant MAY veto the impending registration during this time. Otherwise, the registration will succeed.

IANA is to maintain these registries at designated URIs on its website, currently www.iana.org. Regardless of other formatting, IANA will designate URIs of the form: {pre}/{name}, where {pre} is IANA's chosen name registry prefix over HTTP or HTTPS, and {name} is the name (with UTF-8 percent-encoding). The response to an HTTP GET request at one of these URIs SHALL be a text/plain document with UTF-8 encoding ("charset=UTF-8") formatted as follows:

```
urn:{xmlns or rdf}:{name}
Description: {description}
Contact Information: {contact information}

{URI 0}
{URI 1}
{...}
```

The Description and Contact Information fields MAY span multiple lines by ending the line and including one space (U+0020) on the subsequent line. Semantically, the line-ending and space mean "end of line" only. URN resolvers MAY use this database (whether IANA's customary form, or the form specified by this section) when resolving URNs to URIs to access particular resources.

The purpose of these registered URIs is to assist protocol designers, systems analysts, and software engineers with understanding the nature of the XML namespace or RDF node, rather than for general consumption. For example, one registered URI could forward to a text/html resource that is a standards document describing a protocol that uses the URN. [[NB: compare with xml.resource.org for I-Ds.]] Consequently, the traffic burden imposed on IANA is expected to be negligible.

5. Security Considerations

XML processors use XML namespaces to validate XML content. This document is not expected to introduce any additional security considerations beyond those inherent in XML processing.

RDF processors use RDF URI references (RDF 1.0) or IRIs (RDF 1.1) to identify nodes (subjects, predicates, and objects). This document is not expected to introduce any additional security considerations beyond those inherent in RDF processing.

6. References

6.1. Normative References

[RDF1.0] Klyne, G. and J. Carroll, "Resource Description Framework (RDF): Concepts and Abstract Syntax", World Wide Web Consortium Recommendation REC-rdf-concepts-20040210, February 2004, <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210>.

[RDF1.1] Cyganiak, R., Wood, D., and M. Lanthaler, "RDF 1.1 Concepts and Abstract Syntax", World Wide Web Consortium Recommendation REC-rdf11-concepts-20140225, February 2014, <http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>.

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

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

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

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

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

[URNBIS] Saint-Andre, P. and J. Klensin, "Uniform Resource Names (URNs)", draft-ietf-urnbis-rfc2141bis-urn-13 (work in progress), September 2015.

[XML1.0] Bray, T., Paoli, J., Sperberg-McQueen, M., Maler, E., and F. Yergeau, "Extensible Markup Language (XML) 1.0 (Fifth Edition)", World Wide Web Consortium Recommendation REC-xml-20081126, November 2008, <http://www.w3.org/TR/2008/REC-xml-20081126>.

[XML1.1] Bray, T., Paoli, J., Sperberg-McQueen, M., Maler, E., Yergeau, F., and J. Cowan, "Extensible Markup Language (XML) 1.1 (Second Edition)", World Wide Web Consortium Recommendation REC-xml11-20060816, August 2006, <http://www.w3.org/TR/2006/REC-xml11-20060816>.

6.2. Informative References

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

Author's Address

Sean Leonard
Penango, Inc.
5900 Wilshire Boulevard
21st Floor
Los Angeles, CA 90036
USA

Email: dev+ietf@seantek.com
URI: http://www.penango.com/