

The IETF XML Registry

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 (2004). All Rights Reserved.

Abstract

This document describes an IANA maintained registry for IETF standards which use Extensible Markup Language (XML) related items such as Namespaces, Document Type Declarations (DTDs), Schemas, and Resource Description Framework (RDF) Schemas.

1. Introduction

Over the past few years, the Extensible Markup Language (XML) [W3C.REC-xml] has become a widely used method for data markup. There have already been several IETF Working Groups that have produced standards that define XML Document Type Definitions (DTDs), XML Namespaces [W3C.REC-xml-names], and XML Schemas [W3C.REC-xmlschema-1]. Each one of these technologies uses Uniform Resource Identifiers (URIs) [RFC2396] and other standardized identifiers to identify various components.

For example, while it has been the practice within some standards that use Document Type Definitions (DTDs) to forego the use of the PUBLIC identifiers in favor of 'well known' SYSTEM identifiers, it has proven to be more trouble than its worth to attempt to standardize SYSTEM identifiers. The result is that several IETF standards that have simply created non-resolvable URIs in order to simply identify but not resolve the DTD for some given XML document.

This document seeks to standardize and improve these practices by creating an IANA maintained registry of XML element identifiers so that document authors and implementors have a well maintained and

authoritative location for their XML elements. As part of this standard, the IANA will maintain:

- o the public representation of the document,
- o the URI for the elements if one is provided at the time of registration,
- o a registry of Public Identifiers as URIs.

In the case where the registrant does not request a particular URI, the IANA will assign it a Uniform Resource Name (URN) that follows [RFC3553].

2. 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 BCP 14, RFC 2119 [RFC2119].

3. Registerable Documents

3.1. The Assigned/Registered URI

All elements (except PUBLIC identifiers) in this registry will require a URI in order to be registered. If the registrant wishes to have a URI assigned, then a URN of the form

urn:ietf:params:xml:<class>:<id>

will be assigned where <class> is the type of the document being registered (see below). <id> is a unique id generated by the IANA based on any means the IANA deems necessary to maintain uniqueness and persistence. NOTE: in order for a URN of this type to be assigned, the item being registered MUST have been through the IETF consensus process. Basically, this means that it must be documented in a RFC. The RFC 3553 [RFC3553] URN registration template is found in Section 6.

The IANA will also maintain a file server available via at least HTTP and FTP that contains all of the registered elements in some publicly accessible file space in the same way that all of the IANA's registered elements are available via <http://www.iana.org/assignments/>. While the directory structure of this server is up to the IANA, it is suggested that the files be organized by the <class> and the individual files have the <id> as their filename.

Implementors are warned that they should not programatically rely on those resources being available or the directory structure remaining static for any reason. It is explicitly recognized that some software tools attempt to download DTDs, schema, etc., 'on the fly' and that developers should understand when this is done and when to not reference IANA network resources as a 'schema download repository'. This is the reason that the IANA will not register or provide SYSTEM identifiers.

3.2. Registerable Classes

The list of types of XML elements that can be registered with the IANA are:

`publicid` -- An XML document that contains a DOCTYPE declaration or any other external reference can identify that reference via both a PUBLIC identifier and a SYSTEM identifier. The SYSTEM identifier is system-specific information that enables the entity manager of an XML system to locate the file, memory location, or pointer within a file where the entity can be found. It should also be noted that a system identifier could be an invocation of a program that controls access to an entity that is being identified. Thus, they are not registered items. In many cases, SYSTEM identifiers are also URIs. However, in these cases, the URI is still only used for system-specific information. In the case where a PUBLIC Identifier is also a URI, it is possible for the SYSTEM Identifier to contain the same URI but this behavior is not recommended unless its side effects are well known and understood to not cause any unacceptable harm.

A PUBLIC identifier is a name that is intended to be meaningful across systems and different user environments. Typically, it will be a name that has a registered owner associated with it, so that public identifiers will be guaranteed unique and no two entities will have the same public identifier. In practice, PUBLIC identifiers are typically Formal Public Identifiers [ISO.8879.1986] but they are not restricted to just that set. As said in [RFC3151]:

"Any string which consists only of the public identifier characters (defined by Production 13 of Extensible Markup Language (XML) 1.0 Second Edition) is a legal public identifier."

Therefore, it is legal for a PUBLIC identifier to be a URN if it adheres to the character set restrictions.

Thus, the identifier registered along with a DTD is its PUBLIC identifier. The only restriction being that it must adhere to the character set restrictions. In the case where the registrant does not provide one, the IANA will assign one of the form 'urn:ietf:params:xml:pi:<id>'. Registrants are encouraged to investigate RFC 3151 [RFC3151] as a recommended method for minting a URN that can also be represented as an FPI.

ns -- XML Namespaces [W3C.REC-xml-names] are named by a URI. They have no real, machine-parseable representation. Thus, the registered document will be either the specification or a reference to it. In the case where a URI is not provided by the registrant, the IANA will assign a URN of the form 'urn:ietf:params:xml:ns:<id>' which will be the XML Namespace's name.

schema -- XML Schemas [W3C.REC-xmlschema-1] are also identified by a URI but their contents are machine parseable. The IANA registered document will be the XML Schema file. The URN the IANA assigns can be used as the URI for the schema and is of the form 'urn:ietf:params:xml:schema:<id>'.

rdfschema -- The Resource Description Format (RDF) [W3C.CR-rdf-schema] is an XML serialization of a connected graph based data model used for metadata expression. RDF makes use of schemas for RDF that express grammars about relationships between URIs. These grammars are identified by URIs. The URN assigned by the IANA can be used as the identifying URI and is of the form 'urn:ietf:params:xml:rdfschema:<id>'.

4. Registration Procedures

Until the IANA requests or implements an automated process for the registration of these elements, any specifications must make that request part of the IANA considerations section of their respective documents. That request must be in the form of the following template:

URI

The URI or PUBLIC identifier that identifies the XML component. If the registrant is requesting that the IANA assign a URI then this field should be specified as "please assign".

Registrant Contact

The individual/organization that is the registration contact for the component being registered. Ideally, this will be the name and pertinent physical and network contact information. In the case of IETF developed standards, the Registrant will be the IESG.

XML

The exact XML to be stored in the registry. Unless the beginning and end of the file is obvious, the document should use the text "BEGIN" to mark the beginning of the file and "END" to mark the end of the file. The IANA will insert any text between those two strings (minus any page breaks and RFC formatting inserted by the RFC Editor) into the file kept in the repository.

5. Security Considerations

The information maintained by the IANA will be authoritative and will be a target for attack. In some cases, such as XML Schema and DTDs, the content maintained by the IANA may be directly input into software. Thus, extra care should be taken by the IANA to maintain the security precautions required for an important reference location for the Internet.

Beyond this concern, there are no other security considerations not already found with any other IANA registry.

6. IANA Considerations

This document seeks to create a rather large registry for which the IANA (at the direction of the IESG) will be primarily responsible. The amount of effort required to maintain this registry is not insignificant and the policies and procedures surrounding any approval process are non-trivial. The registry is on a First Come First Served basis, but a Specification is Required. Once the IETF has some experience with this registry, these policies may change.

RFC 3553 [RFC3553] specifies that any new registry requiring a name, to be assigned below the 'urn:ietf:params' namespace and must specify the structure of that space in template form. The IANA has created and will maintain this new sub-namespace:

Registry-name: xml

Specification: This document contains the registry specification. The namespace is organized with one sub-namespace which is the <id>.

Repository: To be assigned according to the guidelines found above.

Index value: The class name

7. Normative References

- [ISO.8879.1986] International Organization for Standardization, "Information processing - Text and office systems - Standard generalized markup language (SGML)", ISO Standard 8879, 1986.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2396] Berners-Lee, T., Fielding, R. and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax", RFC 2396, August 1998.
- [RFC3151] Walsh, N., Cowan, J. and P. Grosso, "A URN Namespace for Public Identifiers", RFC 3151, August 2001.
- [RFC3553] Mealling, M., Masinter, L., Hardie, T. and G. Klyne, "An IETF URN Sub-namespace for Registered Protocol Parameters", BCP 73, RFC 3553, June 2003.
- [W3C.CR-rdf-schema] Brickley, D. and R. Guha, "Resource Description Framework (RDF) Schema Specification 1.0", W3C CR-rdf-schema, March 2000, <<http://www.w3.org/TR/2000/CR-rdf-schema-20000327>>.
- [W3C.REC-xml] Bray, T., Paoli, J., Sperberg-McQueen, C. and E. Maler, "Extensible Markup Language (XML) 1.0 (2nd ed)", W3C REC-xml, October 2000, <<http://www.w3.org/TR/REC-xml>>.
- [W3C.REC-xml-names] Bray, T., Hollander, D. and A. Layman, "Namespaces in XML", W3C REC-xml-names, January 1999, <<http://www.w3.org/TR/REC-xml-names>>.
- [W3C.REC-xmlschema-1] Thompson, H., Beech, D., Maloney, M. and N. Mendelsohn, "XML Schema Part 1: Structures", W3C REC-xmlschema-1, May 2001, <<http://www.w3.org/TR/xmlschema-1/>>.

8. Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any intellectual property 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; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in BCP-11. Copies of claims of rights made available for publication 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 implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

9. Author's Address

Michael Mealling
VeriSign, Inc.
Mountain View, CA
USA

Email: michael@verisignlabs.com
URI: <http://www.research.verisignlabs.com>

10. Full Copyright Statement

Copyright (C) The Internet Society (2004). 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 assignees.

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.

