Network Working Group
Internet-Draft
Expires: May 7, 2002

G. Klyne
MIMEsweeper Group
Nov 6, 2001

A URN sub-namespace for language tags
draft-klyne-urn-ietf-lang-01

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on May 7, 2002.

Copyright Notice

Copyright (C) The Internet Society (2001). All Rights Reserved.

Abstract

This specification defines a form of URN to be used for language tags defined or registered according to RFC 3066. The URN namespace used is 'ietf', defined by RFC 2648 and extended by RFC????.

Table of Contents

1. Introduction 3
1.1 Background 3
2. Registration template 3
3. Examples 4
4. IANA considerations 4
5. Security considerations 5
6. Acknowledgements 5
References 5
Author's Address 6
A. Revision history 6
A.1 draft-klyne-urn-ietf-lang-00 6
A.2 Todo 6
Full Copyright Statement 7

1. Introduction

This specification defines a form of URN to be used for language tags defined or registered according to RFC 3066 [4]. The URN namespace used is 'IETF', defined by RFC 2648 [3] and extended by RFC???? [7] to include names for IETF protocol parameters.

1.1 Background

RFC 3066 [4] defines a construction and registry for tags used to identify human languages. These tags can be used to describe the language used by human-readable text, or other data intended for human prsentation.

There are situations in which it is desired to express a language tag in the form of a URI, but no definitive URI form has been defined.

For example, W3C has defined RDF [5], a generic XML-based metadata format that uses URIs to identify objects and the relationships between them. There has been some requirement to use language tags with RDF. RFC 3066 language tags have the desired semantics, but it is not clear how they should be represented in RDF. The normal form of identification used by RDF (and the Web in general) is a URI or URI-reference as defined by RFC 2396 [2].

This specification defines a way to embed RFC 3066 language tags in a urn: form of URI, which can be used to identify a language in contexts where a URI is prefered to a text string or token.

2. Registration template

The URN sub-namespace for language tags is defined as follows.

Registry name:
language

Specification:
RFC 3066 [4].

Repository:
http://www.iana.org/assignments/language-tags
See also non-registered language tags defined by RFC 3066 [4].

Index value:
The language tag name is the registry index value. RFC 3066 allows this tag name to contain uppercase letters, lowercase letters, digits, and dash ("-"). Language tags are case-insensitive.

Some allowed unregistered language tag values are defined by reference to ISO standard 639 [6].

URN formation:
The URN for a language tag is formed as: "urn:ietf:params:language:<tag-name>", where <tag-name> is the language registry index value, expressed using lower case letters.

RFC 2141 [1] defines the format of URNs. Allowable characters include all of those noted above.

URNs are defined by RFC 2141 [1] as lexically equivalent if they are identical following case normalization of the urn scheme name, the namespace name and any %-escaping used. Language tags are defined such that upper- and lower-case ASCII characters are not distinguished. In forming a URN, all ASCII characters in the language tag must be expressed in lower case.

3. Examples

This table lists some language tags, and the corresponding urn: URIs.

MN (Mongolian)
urn:ietf:params:language:mn

en-US (American English)
urn:ietf:params:language:en-us

zh-yue (Cantonese)
urn:ietf:params:language:zh-yue

sgn-GB (British sign language)
urn:ietf:params:language:sgn-gb

i-klingon
urn:ietf:params:language:i-klingon

i-default
urn:ietf:params:language:i-default

x-private
urn:ietf:params:language:x-private

4. IANA considerations

This document calls for the creation of a new IETF sub-namespace per RFC???? [7]. Registration details are in the preceding section.

5. Security considerations

No security considerations are introduced by this specification beyond those already inherrent in the use of language tags [4].

6. Acknowledgements

The author gratefully acknowledges the contributions of: Martyn Horner, [[[...]]]

References

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

[2] Berners-Lee, T., Fielding, R. and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax", RFC 2396, August 1998.

[3] Moats, R., "A URN Namespace for IETF Documents", RFC 2648, August 1999.

[4] Alvestrand, H., "Tags for the Identification of Languages", BCP 47, RFC 3066, January 2001.

[5] Lassila, O. and R. Swick, "Resource Description Framework (RDF) Model and Syntax Specification", W3C Recommendation rdf-syntax, February 1999, <http://www.w3.org/TR/REC-rdf-syntax>.

[6] International Organization for Standardization, "Code for the representation of names of languages, 1st edition", ISO Standard 639, 1988.

[7] Mealling, M., Masinter, L., Hardie, T. and G. Klyne, "An IETF URN Sub-namespace for Registered Protocol Parameters", Internet draft draft-mealling-iana-urn-02, August 2001, <http://search.ietf.org/internet-drafts/draft-mealling-iana-urn-02.txt>.

Author's Address

Graham Klyne
MIMEsweeper Group
1310 Waterside
Arlington Business Park
Theale, Reading RG7 4SA
UK

Phone: +44 118 903 8000
Fax: +44 118 903 9000
EMail: Graham.Klyne@MIMEsweeper.com

Appendix A. Revision history

[[[This section to be removed on final publication]]]

A.1 draft-klyne-urn-ietf-lang-00

00a 18-Sep-2001: document initially created.

00b 24-Sep-2001: Small editorial fixes.

01a 06-Nov-2001: Revise background material. Editorial fixes.

A.2 Todo

- Align with urn:ietf:params: draft as it develops.
- Replace RFC???? with appropriate RFC numbers, when issued.
- Remove [[[editorial comments]]].

Full Copyright Statement

Copyright (C) The Internet Society (2001). 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.