Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 9, 2014

J. Levine
Taughannock Networks

February 5, 2014

Assigning Digital Object Identifiers to RFCs
draft-iab-doi-00

Abstract

The Digital Object Identifier (DOI) is a widely used system that assigns unique identifiers to digital documents that can be queried and managed in a consistent fashion. We propose a method to assign DOIs to past and future RFCs.

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 August 9, 2014.

Table of Contents

1. Introduction

The Digital Object Identifier (DOI) is a widely used system that assigns unique identifiers to digital documents that can be queried and managed in a consistent fashion. The structure of DOIs is defined by ISO 26324:2012 [ISO-DOI] and is implemented by a group of registration agencies coordinated by the International DOI Foundation.

Each DOI is accompanied by metadata about the object, such as one or more URIs where the object can be found. The DOI system also provides many features not relevant to RFCs, such as the ability to update the metadata after the DOI is assigned, and for organizations to maintain local caches of metadata, e.g., a university or corporate library that tracks its copies of purchased documents so subsequent users don't buy them again.

The wide use of DOIs suggests that even though RFCs can be downloaded directly from the IETF for free, organizations that use DOIs can have trouble locating non-DOI documents. DOIs with metadata that points to the existing free online RFCs would make RFCs easier to find. Some scholarly publishers accept DOIs as references in published documents, so DOIs would make RFCs easier to cite.

2. Structure and resolution of DOIs

DOIs are an application of the handle system defined by RFCs [RFC3650], [RFC3651], and [RFC3652]. A DOI for an RFC might be

    10.123456/rfc1234

The first part of a DOI is the number 10, which means a DOI within the handle system, a dot, and a unique number assigned to a publisher, in this example 123456. This part is the DOI prefix. Following that is a slash and a text string assigned by the publisher, called the DOI suffix. A reasonable way to assign DOIs

would be to use the familiar series names and numbers, e.g., rfc1234, bcp100, or std11. (DOIs are case-insensitive.)

Although the handle system has its own protocol described in [RFC3652], the usual way to look up a DOI is to use web lookup. CNRI provides a Firefox plugin that adds a "doi:" URI scheme. Lacking that, one can use a public http proxy, usually http://dx.doi.org, so the sample DOI above could be looked up at:

http://dx.doi.org/10.123456/rfc1234

Whenever a publisher assigns a DOI, it provides the metadata for the object (henceforth called a document, since that what they are in this context) to its registration agency which then makes it available to clients that look up DOIs. Publishers have considerable flexibility as to what actually resides at the URI(s) that a DOI refers to. Sometimes it's the document itself, while for commercial publishers it's typically a page with the abstract and bibliographic information, and some way to buy the actual document. Since some RFCs are in multiple formats (e.g., Postscript and text) an appropriate URI would be that of the RFC Editor's info page that has the RFC's abstract and links to the document in various formats. Hence the URI above would be set to redirect to

http://www.rfc-editor.org/info/rfc1234

More information on the structure and use of DOIs is in the DOI Handbook [DOI-HB].

3. DOIs for RFCs

Once the RFC series has DOIs assigned, it would be a good idea to include the DOI in the boilerplate for each RFC, perhaps next to the ISSN. Online databases and indexes that include RFCs would be updated to include the DOI, e.g. the ACM Digital Library. (A practical advantage of this is that the DOI would link directly to the IETF, rather than perhaps to a copy of an RFC behind a paywall.)

Since RFCs are immutable, existing RFCs still wouldn't mention their own DOIs within the RFC itself, but putting the DOIs into indexes would still provide value.

4. The process of assigning DOIs

There are three phases to assigning DOIs to RFCs, getting a DOI prefix, retroactively assigning DOIs to existing documents, and updating the publication process to assign DOIs as new RFCs are published.

4.1. Getting a DOI prefix

There are ten registration agencies [DOI-RA] that assign DOI prefixes. Most of them serve specialized audiences or limited geographic areas, but there are a few that handle scholarly and technical materials. All registration agencies charge for DOIs to defray the cost of maintaining the metadata databases. The prices are fairly low, on the order of $660/year for membership, 15 cents per document deposit fees for a bulk upload of the backfile (the existing RFCs), and $1/per document to deposit them as they are published.

4.2. Retroactively assigning DOIs

Other than paying the submission fees, assigning DOIs to all of the existing RFCs is primarily a software problem. We'd need tools to extract or create the metadata for all of the RFCs and submit it to the registration agency using the agency's online API. Where we are aware of indexes and databases that include RFCs, we would try to get them to include the DOI.

4.3. Assigning DOIs to new RFCs

As new RFCs are published, the publication process will add steps to collect and submit the metadata to the registration agency.

5. Informative References

[DOI-HB] International DOI Foundation, "DOI Handbook", April 2012, <http://www.doi.org/hb.html>.

[DOI-RA] International DOI Foundation, "DOI Registration Agencies", July 2013, <http://www.doi.org/registration_agencies.html>.

[ISO-DOI] International Organization for Standardization (ISO), "ISO 26324:2012 Information and documentation -- Digital object identifier system", 2012, <http://www.iso.org/iso/catalogue_detail?csnumber=43506>.

[RFC3044] Rozenfeld, S., "Using The ISSN (International Serial Standard Number) as URN (Uniform Resource Names) within an ISSN-URN Namespace", RFC 3044, January 2001.

[RFC3650] Sun, S., Lannom, L., and B. Boesch, "Handle System Overview", RFC 3650, November 2003.

[RFC3651] Sun, S., Reilly, S., and L. Lannom, "Handle System Namespace and Service Definition", RFC 3651, November 2003.

[RFC3652] Sun, S., Reilly, S., Lannom, L., and J. Petrone, "Handle System Protocol (ver 2.1) Specification", RFC 3652, November 2003.

Author's Address

John Levine
Taughannock Networks
PO Box 727
Trumansburg, NY 14886

Phone: +1 831 480 2300
Email: standards@taugh.com
URI: http://jl.ly