

Équipe d'ingénierie de l'Internet (IETF)
Request for Comments : 7669
 Catégorie : Information
 ISSN : 2070-1721

J. Levine, Taughannock Networks
 octobre 2015
 Traduction Claude Brière de L'Isle

Allocation d'identifiants d'objet numériques aux RFC

Résumé

Le présent document décrit la façon dont les identifiants d'objet numérique (DOI, *Digital Object Identifier*) sont alloués aux RFC passées et à venir. Le DOI est un système largement utilisé qui alloue des identifiants univoques aux documents numériques qui peuvent être interrogés et gérés de façon cohérente.

Statut de ce mémoire

Le présent mémoire n'est pas une spécification sur la voie de la normalisation de l'Internet ; il est publié à des fins d'information.

Le présent document a été produit par l'équipe d'ingénierie de l'Internet (IETF). Il représente le consensus de la communauté de l'IETF. Il a subi une révision publique et sa publication a été approuvée par le groupe de pilotage de l'ingénierie de l'Internet (IESG). Plus d'informations sur les normes de l'Internet sont disponibles à la Section 2 de la [RFC5741].

Les informations sur le statut actuel du présent document, tout errata, et comment fournir des réactions sur lui peuvent être obtenues à <http://www.rfc-editor.org/info/rfc7669>

Notice de droits de reproduction

Copyright (c) 2015 IETF Trust et les personnes identifiées comme auteurs du document. Tous droits réservés.

Le présent document est soumis au BCP 78 et aux dispositions légales de l'IETF Trust qui se rapportent aux documents de l'IETF (<http://trustee.ietf.org/license-info>) en vigueur à la date de publication de ce document. Prière de revoir ces documents avec attention, car ils décrivent vos droits et obligations par rapport à ce document. Les composants de code extraits du présent document doivent inclure le texte de licence simplifié de BSD comme décrit au paragraphe 4.e des dispositions légales du Trust et sont fournis sans garantie comme décrit dans la licence de BSD simplifiée.

Table des Matières

1. Introduction.....	1
2. Structure et résolution des DOI.....	2
3. DOI pour les RFC.....	3
4. Processus d'allocation des DOI.....	3
4.1 Obtention d'un préfixe de DOI.....	3
4.2 Allocation rétroactive des DOI.....	3
4.3 Allocation de DOI aux nouvelles RFC.....	3
4.4 Utilisation des DOI dans les RFC.....	3
4.5 Travaux futurs possibles.....	4
5. Internationalisation.....	4
6. Considérations sur la sécurité.....	4
7. Références pour information.....	4
Membres de l'IAB au moment de l'approbation.....	4
Adresse de l'auteur.....	5

1. Introduction

Le système d'identifiant d'objet numérique (DOI, *Digital Object Identifier*) alloue des identifiants univoques aux documents numériques qui peuvent être interrogés et gérés de façon cohérente. La structure des DOI est définie par la norme ISO 26324:2012 [ISO-DOI] et est mis en œuvre par un groupe d'agences d'enregistrement coordonnées par la Fondation DOI internationale.

Chaque DOI est associé à des métadonnées bibliographiques sur l'objet, incluant un ou plusieurs URI où l'objet peut être

trouvé. Les métadonnées sont mémorisées dans une base de données publiques avec des entrées restituées via HTTP.

Les DOI sont largement utilisés par les éditeurs et les consommateurs des journaux techniques et autre matériel technique publié en ligne.

La page 15 de [CITABILITY] indique que (noter que des citations ont été omises) : "les adresses normales de la Toile ne sont pas fiables pour localiser les ressources en ligne, parce que elles peuvent bouger, changer ou disparaître entièrement. Mais les identifiants permanents sont fixes, avec une infrastructure qui permet que la localisation de l'élément soit mise à jour. Le résultat est que l'identifiant peut fournir un accès persistant aux données. DataCite fournit un tel service, et les DOI (utilisés par DataCite) sont de loin l'identifiant le plus couramment mentionné par les personnes interrogées, suivis de près par les descripteurs (*handles*) (sur lesquels le système de DOI est construit). Il y avait une nette préférence des personnes interrogées pour les DOI parce que c'est un système déjà utilisé et compris par les éditeurs pour les publications traditionnelles et dont la barrière à franchir est probablement moins haute que pour un système entièrement nouveau."

Certains éditeurs de livres scolaires acceptent les DOI comme références dans les documents publiés, et certaines versions de BibTeX peuvent automatiquement restituer les données bibliographiques pour un DOI et les formater. Les DOI peuvent avoir d'autres avantages, comme de rendre plus facile de trouver les versions en ligne gratuites des RFC plutôt que des copies sur un mur payant quand on suit les références ou qu'on utilise des index de documents.

Les avantages des DOI s'appliquent également aux documents provenant de tous les flux de soumission de RFC, de sorte que des DOI sont alloués à toutes les RFC.

2. Structure et résolution des DOI

Les DOI sont une application du système de descripteurs (*handle*) défini par les [RFC3650], [RFC3651], et [RFC3652]. Par exemple, un DOI pour une RFC pourrait être : 10.17487/rfc1149

La première partie d'un DOI est le nombre 10, qui signifie un DOI dans le système de descripteurs, suivi par un point et un nombre unique alloué à un éditeur, dans ce cas, 17487. Cette partie est le préfixe du DOI. Ensuite, il y a une barre oblique et une chaîne de texte allouée par l'éditeur, appelée le suffixe DOI.

Les DOI sont traités comme des identifiants opaques. Les suffixes DOI alloués aux RFC sont actuellement fondés sur le champ "doc-id" de l'index des RFC en XML (*rfc-index.xml*), mais le suffixe des futures RFC pourrait se fonder sur quelque chose d'autre si les circonstances changent. Donc, la façon fiable de trouver le DOI pour une RFC n'est pas de deviner, mais de le chercher dans l'index des RFC ou sur le site de la Toile de l'éditeur des RFC < <https://www.rfc-editor.org/> >. Les références de RFC créées à partir des entrées dans les bibliothèques usuelles *bibxml* auront leurs DOI inclus automatiquement.

Bien que le système des descripteurs ait son propre protocole décrit dans la [RFC3652], la façon usuelle de chercher un DOI est d'utiliser la recherche sur la Toile. Une proposition d'URN "doi:" n'a jamais été largement mise en œuvre, de sorte que la façon standard de chercher un DOI est d'utiliser le mandataire public HTTP à < <https://dx.doi.org/> >. L'exemple de DOI ci-dessus pourrait ressembler à : <https://dx.doi.org/10.17487/rfc1149>

Chaque fois qu'un éditeur alloue un DOI, il fournit les métadonnées bibliographiques pour l'objet (qu'on appellera ici un document, car c'est ce dont il s'agit dans ce contexte) à son agence d'enregistrement qui va le rendre disponible aux clients qui cherchent les DOI. Les métadonnées du document sont normalement téléchargées à l'agence d'enregistrement en XML en utilisant une API fondée sur HTTP. Le logiciel d'utilisateurs ou d'éditeur peut restituer les métadonnées en allant chercher l'URL du DOI et en utilisant une négociation de contenu HTTP standard pour demander *application/citeproc+json*, *application/rdf+xml*, ou un autre format bibliographique.

Les éditeurs ont une souplesse considérable quant à ce qui réside aux URI auxquels un DOI se réfère. Parfois, c'est le document lui-même, tandis que pour les éditeurs commerciaux c'est normalement une page avec le résumé, des informations bibliographiques, et des façons d'acquérir le document réel. Parce que certaines RFC sont sous plusieurs formats (par exemple, Postscript et text) un URI approprié est celui de la page d'informations de l'éditeur des RFC qui a le résumé du document et les liens avec le ou les documents dans divers formats. Donc, l'URI ci-dessus, quand on va le chercher via une demande HTTP qui accepte *text/html*, redirige sur : <https://www.rfc-editor.org/info/rfc1149>

Plus d'informations sur la structure et l'utilisation des DOI sont dans le manuel du DOI [DOI-HB].

3. DOI pour les RFC

Avec des DOI alloués à chaque RFC, il est utile d'inclure des informations de DOI dans la bibliographie XML comme éléments de "seriesInfo", afin que les moteurs de rendu puissent les afficher si désiré. Les bases de données en ligne et les index qui incluent les RFC devraient être mis à jour pour inclure le DOI, par exemple, la bibliothèque numérique ACM. (Un avantage pratique de cela est que le DOI va relier directement à l'éditeur des RFC, plutôt que peut-être à une copie d'une RFC derrière un mur payant.)

Comme les RFC sont immuables, les RFC existantes ne mentionnent pas encore leur propre DOI en leur sein, mais mettre leurs DOI dans les index est précieux.

4. Processus d'allocation des DOI

Il y a trois phases dans l'allocation des DOI aux RFC : obtenir un préfixe de DOI, allouer rétroactivement des DOI aux documents de RFC existants, et mettre à jour le processus de publication pour allouer des DOI lorsque de nouvelles RFC sont publiées.

4.1 Obtention d'un préfixe de DOI

Il y a dix agences d'enregistrement [DOI-RA] qui allouent les préfixes de DOI. La plupart d'entre elles desservent un public spécialisé ou des zones géographiques limitées, mais il y en a quelques unes qui traitent le matériel éducatif et technique. Toutes les agences d'enregistrement tarifient les DOI pour se défrayer du coût de tenu des bases de données de métadonnées.

L'éditeur des RFC a choisi CrossRef, une agence largement utilisée par les éditeurs de journaux. Les prix associés à l'adhésion à CrossRef sont de l'ordre de 660,00 \$ par an pour l'adhésion, les frais de dossier sont de 0,15 \$ par document pour un chargement brut du fichier (la RFC existante) et 1,00 \$ par document pour le déposer à sa publication.

Le préfixe de DOI de l'éditeur des RFC est 10.17487.

4.2 Allocation rétroactive des DOI

À part de payer les frais de dossier, allouer les DOI à toutes les RFC existantes était principalement un problème de logiciel. La base de données interne du centre de production des RFC a été mise à jour pour inclure un champ DOI pour chaque RFC, le schéma pour rfc-index.xml a été mis à jour pour inclure un champ DOI, et les scripts qui créent les index XML et text ont été mis à jour pour inclure le DOI de chaque RFC. Un script de soumission de DOI spécialisé a extrait les métadonnées pour toutes les RFC à partir de l'index XML et les a soumises à l'agence d'enregistrement en utilisant l'API en ligne de l'agence.

4.3 Allocation de DOI aux nouvelles RFC

Lorsque les RFC sont publiées, le logiciel de publication alloue un DOI à chaque nouvelle RFC. Le script de soumission extrait les métadonnées pour les nouvelles RFC de l'index XML et soumet les informations sur les nouvelles RFC à l'agence d'enregistrement.

4.4 Utilisation des DOI dans les RFC

L'agence de DOI demande que les documents qui reçoivent un DOI incluent à leur tour le DOI lorsque possible lors d'une référence à des documents d'autres organisations. Les DOI peuvent être énumérés en utilisant le champ existant "seriesInfo" dans l'entité de référence xml2rfc, et il est aussi demandé aux auteurs de fournir des DOI pour les documents non RFC lorsque possible. Le centre de production des RFC peut ajouter les DOI manquants lorsque il est facile de le faire, par exemple, lorsque la même référence avec un DOI est apparue dans une RFC antérieure, ou une rapide recherche en ligne trouve le DOI. Lorsque la bibliothèque de citation inclut les DOI, le résultat (les références créées à partir de ces bibliothèques de citation) va inclure le DOI.

Le guide de style des RFC [RFC-STYLE] a été mis à jour pour décrire les règles pour inclure les DOI dans les sections Références des RFC.

4.5 Travaux futurs possibles

Comme il est généralement possible de restituer les informations bibliographiques pour un document à partir de son DOI (comme BibTeX peut le faire, décrit ci-dessus) il pourrait valoir la peine d'ajouter cette caractéristique à xml2rfc, afin qu'une référence avec seulement un DOI puisse automatiquement aller la chercher et l'afficher.

5. Internationalisation

Ajouter les DOI ne présente aucun problème d'internationalisation.

Comme les DOI sont opaques, les caractères utilisés dans un DOI particulier sont sans importance à part de s'assurer qu'ils peuvent être représentés quand nécessaire. Le système des descriptifs dit que c'est de l'UNICODE codé en UTF-8, mais en pratique tous les DOI apparaissent comme utilisant seulement des caractères ASCII imprimables. Les métadonnées pour chaque RFC sont téléchargées comme XML codé en UTF-8.

6. Considérations sur la sécurité

Le système DOI ajoute une nouvelle façon de localiser les RFC et une base de données bibliographiques contenant une description de chaque RFC. Les localisations et les informations bibliographiques existantes sont pour l'essentiel inchangées, de sorte qu'il n'y a pas de nouvelle dépendance sur le système de DOI.

Si CrossRef ou la base de données des DOI subissait une brèche de sécurité, il serait éventuellement possible que les utilisateurs soient dirigés sur des localisations autres que le site de la Toile de l'éditeur des RFC ou qu'ils restituent des données bibliographiques incorrectes, mais les RFC réelles resteraient intactes.

7. Références pour information

[CITABILITY] Kotarski, R., Reilly, S., Schrimpf, S., Smit, E., and K. Walshe, "Report on best practices for citability of data and on evolving roles in scholarly communication", 2012, <http://www.stm-assoc.org/2012_07_10_STM_Research_Data_Group_Data_Citation_and_Evolving_Roles_ODE_Report.pdf>.

[DOI-HB] International DOI Foundation, "DOI Handbook", DOI 10.1000/182, avril 2012, <<http://www.doi.org/hb.html>>.

[DOI-RA] International DOI Foundation, "DOI Registration Agencies", juillet 2015, <http://www.doi.org/registration_agencies.html>.

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

[RFC-STYLE] RFC Editor, "RFC Editor Style Guide", <<https://www.rfc-editor.org/styleguide/>>.

[RFC3650] S. Sun, L. Lannom, B. Boesch, "Généralités sur le système de descripteurs", novembre 2003. (*Information*), DOI 10.17487/RFC3650.

[RFC3651] S. Sun, S. Reilly, L. Lannom, "Espace de noms du système de descripteurs et définitions de service", novembre 2003. (*Info.*), DOI 10.17487/RFC3651.

[RFC3652] S. Sun et autres, "Spécification du protocole du système de descripteurs (version 2.1)", novembre 2003. (*Information*), DOI 10.17487/RFC3652.

Membres de l'IAB au moment de l'approbation

Jari Arkko (Président de l'IETF)
Mary Barnes
Marc Blanchet

Ralph Droms
Ted Hardie
Joe Hildebrand
Russ Housley
Erik Nordmark
Robert Sparks
Andrew Sullivan (Président de l'IAB)
Dave Thaler
Brian Trammell
Suzanne Woolf

Adresse de l'auteur

John Levine
Taughannock Networks
PO Box 727
Trumansburg, NY 14886

téléphone : +1 831 480 2300
mél : standards@taugh.com
URI: <http://jl.ly>