emeraldinsight

The Electronic Library
Digital object identifier system: an overview
Rajesh Chandrakar,

Article information:

Users who downloaded this article also downloaded:

TECHNISCHE UNIVERSITÄT
CHEMNITZ

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

Digital object identifier system: an overview

Rajesh Chandrakar

INFLIBNET Centre, Ahmedabad, India

Abstract

Purpose – This paper aims to describe the digital object identifier (DOI) system, an implementation of the Corporation for National Research Initiatives handle system, where a handle is designed to provide an efficient, extensible, and secured global name to an intellectual object.

Design/methodology/approach – Explains the overview of DOI system, its components with examples in addition to benefits of DOI to user communities.

Findings – The management of intellectual objects in a digital environment such as the internet, which is flooded with various kind of objects like research articles, e-books, electronic theses and dissertations etc. requires the existence of persistent, reliable identifiers for each distinguishable piece of content and associated services activated by these identifiers to manage access and other digital rights. The DOI is the essential part of the electronic publishing especially for the management and the access of the resources is concerned. The DOI system is the new technology developed for persistent identification and interoperable exchange of intellectual property on digital networks.

Originality/value – This paper offers a useful explanation of DOI and their implementation in simple way for the professionals.

Keywords Digital signatures, Document handling, Data handling, Electronic publishing

Paper type Technical paper

Introduction

The digital object identifier (DOI) is one of the members of the private identifier family, which are maintained by self-contained systems like digital repositories for allowing access to resources within an encapsulated environment and designed to work without human intervention (Scott-Wilson, 2004). The DOI is defined as "a character string used to identify intellectual property in the digital environment" (*Information Standards Quarterly*, 2004). The DOI is an initiative of the publishing community for protecting its assets in the digital environment. It provides a framework for managing intellectual content including activities such as linking users with content owners, facilitating electronic commerce, and enabling automated copyright management for all type of media. In addition, the DOI extends an extensible framework for managing content in any form and at any level of granularity in a digital environment. The International DOI Foundation (IDF), a not-for-profit membership organization established in 1998 controlled by an Executive Board elected by the members of the Foundation, supports the needs of the intellectual property community in the digital environment, by the development and promotion of DOI System as a common infrastructure for content management (www.doi.org).

Format-wise, the DOI is composed of the prefix and the suffix; both are separated by forward slash '/'. The prefix is a combination of the directory code < DIR > and the registrant code < REG >, which identifies the creating organization that wishes to

register DOIs. The suffix is made up of DOI suffix string < DSS > , which identifies or locates information about particular object, which is always unique to a given prefix. Thus, DOI is a combination of "Prefix" and "Suffix" without having limitation on the length either in DOI string or any of its components, which is always globally unique in nature. The DOI is assigned by the registration agencies (RAs), which is appointed by the IDF. The IDF has appointed various RA for dealing with the specific client community, where the CrossRef (www.crossref.org) is the first major DOI Registration Agency works for electronic publishing (Figure 1). Accordingly, the syntax of the DOI is:

$$< DIR > . < REG > / < DSS >$$

Valid value for Directory Code < DIR > , which is numeric in nature, is 10 assigned by IDF. Registrant's Code < REG > , which is separated from < DIR > by "." within the DOI prefix, assigned by IDF to the individual organization. It can be any character within Unicode. In the example < REG > 1108 is assigned to the Emerald Group Publishing Ltd. (EGPL) by the CrossRef Registration Agency.

Suffix or < DSS > , which is also known as item identifier, is assigned by the Registrant to the object being published by the registrant. < DSS > is case sensitive and any character within Unicode cannot start with * /where *, which is reserved for future use. In the above example < DSS > 146845202104338 is assigned by registrant "EGPL" to an article published in *Online Information Review* of EGPL entitled "Multi-script bibliographic database: an Indian perspective".

The DOI system

According to the *DOI Handbook* (www.doi.org) the definition of the DOI system is, "The integrated system – comprising enumeration, description, resolution and policy making – managed by the IDF, providing an identifier i.e. an implemented system of labels through a numbering scheme in an infrastructure using a specification". The DOI system gives persistent identification and interoperable exchange of intellectual property on digital networks. The DOI System is an implementation of the handle system (www.handle.net) which is a general-purpose distributed information system, designed to provide an efficient, extensible, and secured global name service for use on networks such as the internet (Lannom, 2000). It is a defined collection of internet protocols and software libraries that are used to implement handle system functionality. The system was developed by the Corporation for National Research Initiatives (CNRI) (http://cnri.reston.va.us) and funded by the Defense Advanced Projects Agency to the Computer Science Technical Reports Project. The handle system manages handle which are persistence and always unique for digital objects regardless of there presence (Scott-Wilson, 2004). The DOI system is made up of four

Figure 1.
Structure of DOI

DOI	Prefix		Suffix
<Handle>	<Naming Authority>		<Item Identifier>
DOI	<DIR>	<REG>	<DSS>
DOI 10.1108/146845202104338	10	1108	146845202104338

components: identifier, resolution, metadata or description, and policy. The value of DOI system lies in its combination of resolution, metadata and policy.

Identifier
Identifiers have become popular in the digital environment. The identifier is the alphanumeric string of characters that are assigned to an object in the form of prefix and suffix. Once created, the prefix retains its uniqueness forever, regardless of copyright authority or ownership, which is provided by the Registration Agency. Assigning the suffix depends upon the organization – a number of organizations assign a sequentially numeric suffix according to an internal corporate framework and some intellectually meaningful strings on the basis of certain other identifiers such as PII (publisher item identifier), SICI (serial item and contribution identifier), and ISBN (International Standard Book Number) (Figure 2).

Resolution
Resolution is an internet process, which associates a DOI to one or more current values of information about the identified object. It is used to mean the act of submitting one or more pieces of current information related to the identifier. For example, in the case of the domain name system (DNS), the resolution is from domain name (www.inflibnet.ac.in), to a single IP 202.141.130.66, which is registered under the Registration Agency for India i.e. National Centre for Software Technology (NCST), Mumbai (now merged with C-DAC and

a) DOI for the "Author's Licensing and Collecting Society's Byline Service"

10.054/1418EC1N2LE

Prefix   Suffix

b) DOI (incorporating a SICI) from an article in the "Journal of the American Society for Information Science", published by John Wiley & Sons:

10.1001/(SICI)1097-4571(199806)49:8<693::AID-ASI4>3.0.CO;2-O

DIR REG   DSS

c) DOI for an article from the Journal of American Medical Association:

10.1001/PUBS.JAMA(278)3,JOC7055-ABST;

Prefix   Suffix

d) DOI for the article "ABO Blood Group System" from Encyclopedia of Immunology Online, 2nd edition, published by Academic Press:

10.1006/rwei.1999.0001

DIR REG   DSS

Source: American National Standards Institute (2000)

Figure 2.
Examples of DOI embedded with various identifiers

renamed as C-DAC, Mumbai), which is then used to communicate with that internet host. But, in the case of the handle/resolution system, the resolution is from a DOI/handle i.e. 10.1108/14684520210438705, to one or more pieces of typed data.

Metadata or description
The third DOI component is aimed at interoperability of metadata that describes the object being identified. Every identifier is registered with an accompanying set of metadata describing the object being referred in the string, where the metadata is provided keeping in view the usefulness of DOI string. Actually, both humans and machines need more information that just an identifier to make use of an information resource, especially outside the system in which the identifier was assigned.

Policy
The most important component of the DOI system is the policy and government. Norman Paskin, Director of the International DOI Foundation, has emphasized that persistence is a function of organizations, not technology. A persistence organization is necessary to ensure the ongoing support and management of a persistence identifier (www.doi.org).

Handle/resolution server
The handle system allows handles to be resolved in a distributed fashion, using dedicated clients, common clients such as web browsers using special extensions or plug-ins, or unextended clients going through various proxies. In all cases, communication with the handle system is carried out using handle system protocols, and in all cases, those protocols have both a formal specification and some specific implementations (www.doi.org). Figure 3 gives an example of the handle/resolution server.

Stepwise description of "an example of handle/resolution system"
Step 1. In the above example step 1 represents that the client is in need of one article entitled "Multi-script bibliographic database: an Indian perspective" written by Rajesh Chandrakar published with *Online Information Review*, which is assigned a DOI number 10.1108/14684520210438705 by EGPL (Figure 4). Where "10.1108" is a prefix represents the publisher code and "14684520210438705" is a suffix represents the unique identifier for said article. Arrow in the example indicates that the client sends a DOI request (handle) to the resolution server (handle server) (http://dx.doi.org) for resolution (Figure 5).

Step 2. The resolution server consists of a collection of handle services owned by various registration agencies. Where one global handle service, is responsible for knowing the locations and name space of all the public local services (LHS). Respectively, each of the LHS knows how to access the global handle service and helps a resolution query to enter the resolution server at any point and routes to the specific service and server as well. In the above example, the CrossRef, which is a registration agency (RA) for IDF in the area of electronic publishing and EGPL is one of the members for assigning the DOI number to the articles. Thus, here in the above
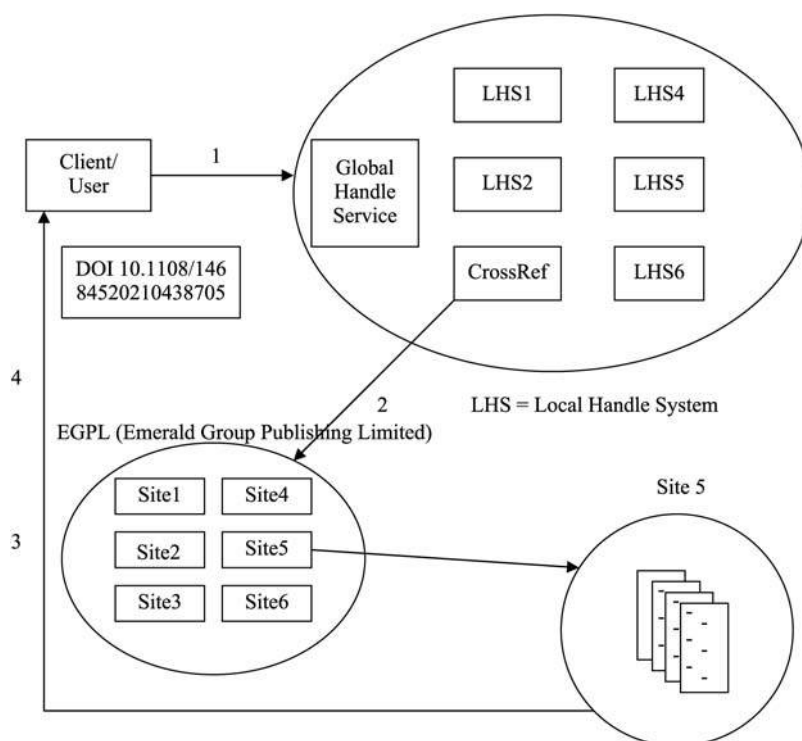
Figure 3.
Handle/resolution server

example, 10.1108/14684520210438705 is a DOI number assigned by EGPL to the said article.

Steps 3 and 4. Each handle can be associated with one or more pieces of typed data. Steps in the example shows that the DOI 10.1108/14684520210438705 is associated with EGPL and resolves to a URL http://iris.emeraldinsight.com/vl = 6347518/ cl = 145/nw = 1/rpsv/cgi-bin/cgi?ini = xref&body = linker&reqdoi = 10.1108/ 14684520210438705 by returning in to the user's browser (Figure 6).

Benefits of the DOI system

By integrating an identifier into a DOI, the identifier becomes actionable as a standard hyperlink (but, unlike URL with persistence), and can function in DOI applications across a variety of platforms. A variety of different identifier systems become readily interoperable when incorporated into DOIs. DOIs may be assigned to ISBN entities (books) to achieve this; DOIs may be used to identify related entities or linked material in any form. Using one system enables one set of tools to be applicable across many platforms, media, standard identification schemes, etc. and promotes interoperable

DOI 10.1108/14684520210438705 Multi-scri t biblio ra hic database: an Indian ers ective

Figure 4.
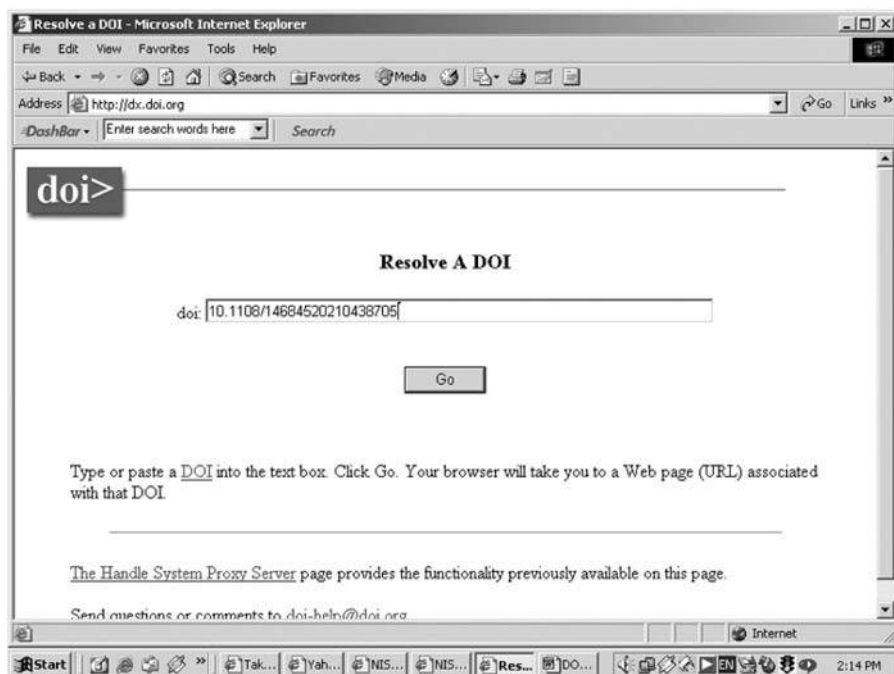An example of
handle/resolution system

Figure 5.
Handle/resolution server
(step 1)

transactions of intellectual property. On the downside, the DOI, in the first instance, seems to be an issue for publishers, when multiple copies of a digital object available on different servers, but it is not as difficult in practice as the resolution system takes the query to the concerned server which has been registered for.

Some other benefits of the DOI are as follows:

· globally accepted unique identifier of electronic resources;

· persistence;

· provides multiple instances;

· extensible namespace;

· international organizational support;

· works under distributed service model;

· works with secured name service;

· part of distributed administration service; and

· provides efficient resolution service.

Conclusion

The DOI is a system for interoperably identifying and exchanging intellectual property in the digital environment. A DOI assigned to content enhances a content producer's ability to trade electronically. It provides a framework for managing content in any form
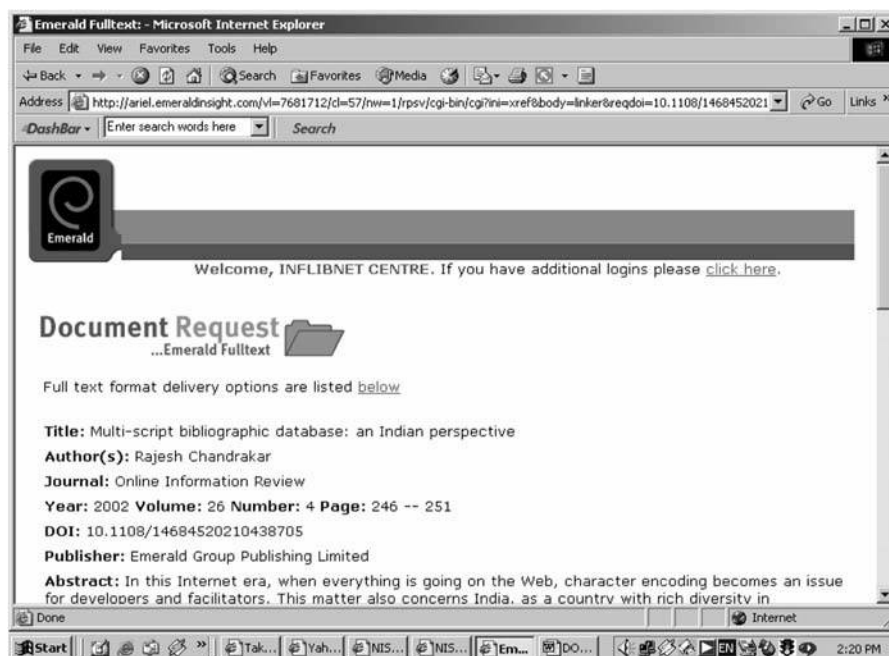
Figure 6.
Resolution (step 4)

at any level of granularity, for linking customers with content suppliers, for facilitating electronic commerce, and enabling automated copyright management for all types of media. The International DOI Foundation, a non-profit organization, manages development, policy and licensing of the DOI to registration agencies and technology providers and advises on usage and development of related services and technologies.

References

Information Standards Quarterly (ISQ) (2004), "National Information Standards Organization (NISO)", *Information Standards Quarterly (ISQ)*, Vol. 16 No. 3, p. 16.

Lannom, L. (2000), "Handle system overview", paper presented at the 66th IFLA Council and General Conference, Jerusalem, 13-18 August, available at: www.ifla.org/IV/ifla66/papers/032-82e.htm

Scott-Wilson, E. (2004), "Identifiers and interoperability", in Gilchrist, A. and Mahon, B. (Eds), *Information Architecture: Designing Information Environments For Purpose*, Facet Publishing, London, pp. 161-73.

Further reading

American National Standards Institute (2000), ANSI/NISO Z39.84-2000: Syntax for the Digital Object Identifier. An American National Standard Developed by the National Information Standards Organization. Approved by ANSI, Bethesda, MD, p. 21.

Hondgon, C. (2004), "The International DOI Foundation: a persistent organization supporting a persistence identifier", *Information Standards Quarterly*, Vol. 16 No. 3, p. 16.

452

About the author

Rajesh Chandrakar is currently Scientific and Technical Officer at INFLIBNNET (Information and Library Network) Centre, Ahmedabad, India, where has been for the last eight years. He is working with the Database Development & Management Group, and the Networking and Quality Control Group. Rajesh holds a Bachelor degree in Science (Physics, Chemistry and Mathematics) from Government Model College of Science, Raipur, and a Masters degree in Library and Information Science from Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India. He also has the Postgraduate diploma in Computer Application from Pt. Ravishankar Shukla University, Raipur. He is Convener of MARC21 Core Group of INFLIBNET Centre. He has been nominated an Alternate representative of the Bureau of Indian Standards (BIS), Technical Committee MSD5 from INFLIBNET Centre and an Alternate Voting Representative to NISO (National Information Standards Organization). He was Joint-Convener of International CALIBER2005 held at Cochin University of Science and Technology, Kochi (Kerala). He also pursued the Training on Managing Academic Libraries as a Commonwealth Professional Fellow at Middlesex University, London during 2005. Rajesh Chandrakar can be contacted at: rchandrakar@yahoo.com or rajesh@inflibnet.ac.in

To purchase reprints of this article please e-mail: **reprints@emeraldinsight.com**
Or visit our web site for further details: **www.emeraldinsight.com/reprints**

This article has been cited by:

1. Nikolay A. Mazov, Vadim N. Gureyev. Modern Challenges in Bibliographic Metadata Identification 1-4. [Crossref]

2. SchopfelJoachim, Joachim Schopfel, RasuliBehrooz, Behrooz Rasuli. 2018. Are electronic theses and dissertations (still) grey literature in the digital age? A FAIR debate. *The Electronic Library* **36**:2, 208-219. [Abstract] [Full Text] [PDF]

3. KendallGraham, Graham Kendall, YeeAngelina, Angelina Yee, HardySteven, Steven Hardy. 2017. We should be just a number and we should embrace it. *The Electronic Library* **35**:2, 348-357. [Abstract] [Full Text] [PDF]

4. Christophe Boudry, Ghislaine Chartron. 2017. Availability of digital object identifiers in publications archived by PubMed. *Scientometrics* **110**:3, 1453-1469. [Crossref]

5. Juan Gorraiz, David Melero-Fuentes, Christian Gumpenberger, Juan-Carlos Valderrama-Zurián. 2016. Availability of digital object identifiers (DOIs) in Web of Science and Scopus. *Journal of Informetrics* **10**:1, 98-109. [Crossref]

6. Piotr Przybyła, Matthew Shardlow, Sophie Aubin, Robert Bossy, Richard Eckart de Castilho, Stelios Piperidis, John McNaught, Sophia Ananiadou. 2016. Text mining resources for the life sciences. *Database* **2016**. . [Crossref]

7. Ruth E. Duerr, Robert R. Downs, Curt Tilmes, Bruce Barkstrom, W. Christopher Lenhardt, Joseph Glassy, Luis E. Bermudez, Peter Slaughter. 2011. On the utility of identification schemes for digital earth science data: an assessment and recommendations. *Earth Science Informatics* **4**:3, 139-160. [Crossref]

8. Erin Carreiro. 2010. Electronic Books: How Digital Devices and Supplementary New Technologies are Changing the Face of the Publishing Industry. *Publishing Research Quarterly* **26**:4, 219-235. [Crossref]