

## DOI: The Persistence of Memory

The Digital Object Identifier (DOI) story began in the fall of 1994, when the Association of American Publishers, the trade association, established its Enabling Technologies Committee; that committee launched the DOI in October 1997 at the Frankfurt Book Fair. The committee's charge was to focus on content identification for the management of intellectual property in the digital publishing world. The various meanings of that "management" include, but were not limited to, management of access rights and copyright, and electronic commerce. Here is a definition of DOI for our purposes:

*The DOI is a unique, persistent identifier of intellectual property in the digital environment.*

Using this tool, an article or paper or image published online would be given a DOI just before or at the time of its creation, and the DOI would remain with it throughout its life span. DOIs are created by and meant for the publishing community, not the individual web user.

**The DOI is unique:** its prefix identifies its publisher, much as the first numbers of an ISBN identify a publisher. That is followed by a slash (/) and the suffix, which identifies the piece down to a level of "granularity," as it is often referred to, that the publisher can decide on—a book, an article, an issue, an image, even a paragraph within a larger document. What a DOI looks like is nicely described in many of the bibliography documents, but no more clearly than on the Whatis DOI page at <http://whatis.com/doi.htm>. The syntax of the DOI is being examined by NISO, the National Information Standards Organization, with an eye toward possible creation of standards, at <http://www.niso.org/commitar.html>.

**The DOI is persistent:** it stays for the life span of the piece, which means that while the URL or other web address may change, the persistence of the DOI means that the object will always be found—no more 404s. The way that happens is through software called the Handle System. The [Corporation for National Research Initiatives](#) (CNRI) a not-for-profit organization in Reston, Virginia founded by Robert Kahn and Vinton Cerf, both founders of the Net as we know it, adapted its [Handle System](#) for the DOI. (A new, Java-based version of the Handle System is expected in April 2000.) What Handle does is connect the DOI to the document's current location on the web or to a location that points to the document, such as a request form. In technical terms, it "resolves" the DOI to the URL where the document lives. If the URL changes, the publisher informs Handle's directory. The DOI does not change, and so all the links to that document are then mapped to the new URL.

**The DOI identifies a piece of digital intellectual property.** How publishers choose to define "a piece" is a lot like trying to define a piece of chocolate. It can be macro: an entire online book or an entire online issue of a journal; or it can be micro: a single image, a single chart; Toblerone bars to Reese's Pieces. It is possible that information about how to access a particular digital object may include fees. Sometimes, for some

objects, it will, but as Clifford Lynch eloquently explains, that is not its entire purpose, even as foreseen by the publishing community.

[The International DOI Foundation](#), a nonprofit organization based in New York and Geneva and founded in 1997, maintains a rich web site, <http://www.doi.org>, that explains DOI, the Handle technology that supports it, links to papers and other publications exploring the implications of DOI, and offers position papers for comment. One such is the paper by its director, Norman Paskin, at [http://www.doi.org/sun\\_pap2.html](http://www.doi.org/sun_pap2.html).

Publishers may register with the Foundation and as of October 1998, according to a *Journal of Electronic Publishing* article, 30 had done so. As recently as a *Library Hotline* article of March 29, 1999, only a week or two before this article was written, Academic Press' IDEAL, its online collection of scientific journals, announced its inclusion of DOIs. There is a letter of intent to pursue an agreement between the DOI Foundation and the international ISBN agency. DOI was constituted as a tool for publishers, but obviously its use has implications for librarians. No library is a member of the Foundation, but library representatives are on the NISO committee that will examine standards. The Coalition for Networked Information's Clifford Lynch has written as lucidly about DOIs as he has about numerous other information technology topics.

Publishers see the DOI as a tool not only to automate the management of material to which they hold the rights, but also as a way to link customers with information that the customers desire even if the rights have been sold or have been taken over by another publishing entity. Will it enable publishers to some day charge for online material? That was part of the intent of its creators. But DOI may also hold the opportunity for keeping digital access alive through multiple generations of paths, URLs, and servers. As Lynch notes, "The DOI as it currently seems to be evolving is likely to be a useful tool to permit consumers to acquire content from publishers on the net with some confidence about who they are doing business with."

Some see a "sinister" aspect to the DOI, as John Berry put it in a *Library Journal* editorial of November 1, 1997, and Stephanie Ardito wonders in "Inevitability: Death, Taxes, and Copyright" *Online*, Jan/Feb 1998 if printed text would also eventually be so encrypted that a page on a copier screen would not be copy-able without coins for the publisher. Patricia Schroeder, president and CEO of the Association of American Publishers responded to *LJ* in a letter published in *LJ*'s February 15, 1998 issue using a quotation from *Publishers Weekly*, "publishers and librarians see tremendous advantages in a universal approach to electronic labeling of all objects available electronically."

The DOI is a genuinely developing idea. None of its characteristics is yet set in stone. It has the potential to give a piece of information online a name it will hold on to forever, enabling it to be found beyond dead links, the HTTP protocol, and the current shape of the web. It also puts into the hands of publishers who own intellectual property a way of selling, leasing, and safeguarding access to that property. DOIs are not an either/or. They might be both.

## Bibliography

What is a DOI?

A fine one-page discussion from Whatis.

<http://whatis.com/doi.htm>

The Digital Object Identifier

The Foundation site: position papers, history, bibliography, links.

<http://www.doi.org>

"The Digital Object Identifier System: Digital Technology Meets Content Management"

by Norman Paskin, Director of the DOI Foundation

[http://www.doi.org/sun\\_pap2.html](http://www.doi.org/sun_pap2.html)

"Identifiers and Their Role in Networked Information Applications"

by Clifford Lynch, *Bulletin of the American Society for Information Science*, December 1997/January 1998.

A version of this appeared in the *ARL Newsletter*, available at

<http://www.arl.org/newsltr/194/identifier.html>

As he always does, Lynch sweeps through the hype and examines DOIs in the context of other identifiers like URLs, URNs, and PURL. He quarrels, as does Caplan, with the name DOI, but point out in some detail that "In a very real sense, there are no bad identifiers, but it is very possible to put identifiers to bad or inappropriate uses."

Caplan, Priscilla. "DOI or Don't We?"

*The Public-Access Computer Systems Review* 9, no. 1 (1998)

<http://info.lib.uh.edu/pr/v9/n1/capl9n1.html>.

Lively, refreshing, and a model of clarity in its introduction to DOI.

"The Digital Object Identifier: Solving the Dilemma of Copyright Protection Online"

by Bill Rosenblatt, *Journal of Electronic Publishing*, University of Michigan Press, v3, issue 2, December 1997.

<http://www.press.umich.edu/jep/03-02/doi.html>.

Written by a participant in the planning stages of DOI, from the publishers' perspective. Detailed history of decision-making, syntax, and DOI outlook at the end of 1997.

"Digital Object Identifiers: Promise and Problems for Scholarly Publishing"

by Lloyd Davidson and Kimberly Douglas,

*The Journal of Electronic Publishing*, v4, issue 2, December 1998.

<http://www.press.umich.edu/jep/04-02/davidson.html>.

A year after Rosenblatt, and a nuanced study of problems and opportunities; 37 footnotes and many links.

NISO's committee for DOI syntax

<http://www.niso.org/commitar.html>

"Metadata: Projects & Standards" by Jessica Milstead and Susan Feldman, *Online*; Wilton; Jan/Feb 1999, Volume: 23, Issue:1, pp32-40. Here is their take:

When a user clicks on a DOI, a message goes to the central directory where the current Web address is located. This address is the one used by the publisher; it may contain the content itself or further information about the content and how to obtain it. The publisher maintains the database, which sits behind this response screen, and it is here that the real set of metadata resides. This system may be used for a variety of purposes, but it was designed by publishers in order to control the distribution of their information. Thus, the metadata in a DOI database is likely to contain not just identification information for an item, but means of linking users of materials to rights holders, thus facilitating e-commerce.

*Publishers Weekly* carried a more recent article about academic publishing and the DOI, Calvin Reid, "STM Publishers Debut DOI Linking Service" *PW* November 29, 1999.

Prepared by [GraceAnne A. DeCandido](#) for the Public Library Association; reviewed April 2000. [ladyhawk@well.com](mailto:ladyhawk@well.com)

*The Public Library Association's Tech Notes project grew out of the desire to continue the work of Wired for the Future: Developing Your Library Technology Plan, by Diane Mayo and Sandra Nelson, published for PLA by ALA in 1999. Each of the Tech Notes, written by GraceAnne A. DeCandido, is a Web-published document of 1500-2000 words, providing an introduction and overview to a specific technology topic of interest to public libraries at a particular point in time. Topics were identified by PLA's Technology in Public Libraries Committee. Each Note is marked with the date of its completion and posting, and with the date, approximately one year later, when links and other information were reviewed.*