

DOI

Resources

■ International DOI Foundation
www.doi.org

A nonprofit standards group which governs the DOI system

■ Corporation for National Research Initiatives
www.cnri.reston.va.us

A research organization and the DOI system's technology partner

Registration Agencies

■ Content Directions, Inc.
www.contentdirections.com

A consulting and services firm advancing DOI among publishers and vendors

■ CrossRef
www.crossref.org

A nonprofit organization that offers a collaborative linking service, using DOIs, for over 80 journal publishers representing over 5,000 publications

■ Enpia
www.enpia.com

A consulting firm providing specialized services to the media industries

■ Learning Objects Network (LON)
www.learningobjectsnetwork.com

An Internet infrastructure company specializing in digital asset management

The broad distribution of paid content is a key goal of the infocommerce industry, and the Digital Object Identifier (DOI), a system spearheaded in the mid-nineties by book and journals publishers, is now gearing up to serve in this cause.

DOI is a system to identify content with a high level of granularity (article, chapter, table, data record), and provides an Internet-based management system that allows publishers to manage access, pricing and payment for each unique identified “chunk” of content. Think of a DOI number as sort of an ISBN number, but for chunks of information rather than entire publications.

From the beginning, DOI was designed to work for multiple media, but it was only in May of 2001 that a prototype, unveiled at *BookExpo America*, revealed its capability to fully and seamlessly integrate e-commerce into mainstream publishing processes.

This elegant implementation takes full advantage of hypertext links, so an article published by one publisher can contain DOI links to other articles, database records, or other media such as photographs, video and software, from other publishers. This approach also provides for the ability to charge for access and automate copyright management.

The result is a much more powerful and rewarding experience for information users, and a full and true implementation of “infocommerce” for

publishers — the ability to seamlessly and ubiquitously distribute and vend content on an à la carte basis with minimal transactional overhead.

The DOI system is designed around the Web. It assigns a unique, permanent identification number to every piece of

content, and maintains giant lookup tables in a DOI Directory that connect each identifier to a Web URL address. The URL could be the content itself (if the content is available for free), a password page for subscription-based access or another mechanism to permit à la carte purchasing.

Another interesting aspect of DOI is its “multiple resolution” capability. A traditional hypertext link has one-to-one resolution, in other words, one link takes you to one place. A single DOI link, by contrast, can offer access

to multiple pieces of related content, and they can all be in different physical locations.

The metadata concept is key to understanding the positioning and significance of DOIs. Metadata has been around for a while, albeit without the fancy name. A library catalog card is a well-known example of metadata since it contains a highly summarized reference to a book, along with a locator for the physical book. You also may have heard of “meta tags” on Web pages. A meta tag is simply a summarized description of what is contained on the Web page with which the meta tag is associated.

All it requires is a modest cost investment and a radically new understanding of what it means to be in the content business

OBJECT IDENTIFIER

It is also important to understand that the DOI methodology depends on a network of central servers distributed throughout the Internet, somewhat like the Domain Name System (DNS) that routes URLs today.

These servers translate or “resolve” the DOI number to an Internet address — the giant lookup table mentioned earlier. They are fed by Registration Agencies that function similarly to Domain Name registrars. They are certified by the International DOI Foundation (IDF), and serve publishers that are already using DOIs.

Ultimately there may be many registration agencies, but currently there are four:

- CrossRef
- Content Directions
- Enpia
- Learning Objects Network.

See sidebar, page 8, for more details.

So why is DOI important?

First it provides an elegant mechanism for a user to move via a hypertext link to another publisher’s site, and pay for the linked content.

Just as importantly for both user and publisher, DOI numbers are “persistent links.” To move your data from one server to another, all you need to do is update the master DOI database and the world is in synch with your change. That’s a huge improvement over having to identify and chase down everyone

who may have linked to your content.

DOI will also accelerate data-text integration and make it easier and more convenient for third parties to resell content. That’s because content will be available in clearly identified, easily accessible units. As importantly as creating these standard “part numbers” for content, a standardized, central system means that cross-references based on DOI numbers become possible, a sure stimulant of usage, and thus revenue.

DOI also allows publishers to track customer needs on a granular level to develop new, targeted information products. Taking publishers far beyond the simple migration of today’s print products to the Web, DOI will enable them to

repackage content into customized “recombinant” information products tailored for highly specific needs.

Although a common lament has been that the Web has reduced information to a commodity, DOI

is a compelling example of how the Web can enhance the value of information. All it requires is a modest cost investment and a radically new understanding of what it means to be in the content business. ●

To see a DOI demo, go to www.contentdirections.com (“See the DOI in action”)

DOI Key Characteristics

- Must be unique
- Are “opaque strings;” they have no meaning (unlike ISBNs, for example, where you can identify the publisher and language from the coded identifier)
- Unlimited in number, with no chance of running out
- Can incorporate any existing identification scheme (such as ISSN or ISBN), or any proprietary numbering scheme that may already be in use
- Associated with metadata (data about data) which greatly facilitates the discovery of content by search engines and portals
- Incorporates Web URLs
- No restriction on granularity
- Is permanent and persistent; once assigned to a piece of content, the DOI would never change (even if moved to another URL or website)

How To Get Started

1. Determine what content your customers want online — this is the content that must be identified with DOI
2. Obtain a Publisher Prefix. Current cost: a one-time \$1,000 fee from International DOI Foundation (IDF)
3. Choose a numbering scheme, either a standard or proprietary legacy numbering scheme, or an entirely new one.
4. Define the metadata within the production process
5. Assign DOIs within the production process
6. Register DOIs and metadata



DOI in action.