

DOI: uno strumento per costruire la biblioteca digitale

di Zeno Tajoli

1 Introduzione

La biblioteca digitale è un settore ancora in crescita e in sviluppo, ma già presente con varie esperienze concrete. Vi sono già delle riflessioni sedimentate sul passaggio dalla biblioteca "fisica" alla biblioteca digitale, una buona definizione può essere presa da De Robbio [1]:

«Si può dire allora che il passaggio dalla biblioteca fisica alla biblioteca virtuale sia segnato da particolari momenti di snodo nei quali la biblioteca che si immette nel web:

- fruisce dell'informazione che va a recuperare;
- diffonde l'informazione recuperata alla sua utenza.

Il passaggio che porta una biblioteca virtuale a divenire biblioteca digitale passa attraverso capacità organizzative e tecniche. Gradualmente, attraverso il web e con gli strumenti che la tecnologia mette a disposizione, la biblioteca diviene "attore", in quanto produce informazione che va a collocarsi fuori della sue mura, in luoghi accessibili da utenze remote.

Due quindi le condizioni del suo essere virtuale e/o digitale, nel ruolo di "client" e/o di "server", in spazi ove pulsano quattro attività principali:

- fruizione
- diffusione
- produzione
- conservazione».

Certo in futuro potranno esserci sorprese e novità, ma alcune importanti sperimentazioni sono già state fatte e gli insegnamenti dei primi prototipi si stanno realizzando in applicazioni accessibili a un pubblico più vasto.

Si può facilmente constatare che l'ambiente d'elezione per la diffusione di queste novità rimane ancora l'ambito universitario e di ricerca. Proprio in quest'ambito si è affacciato e si sta affermando il DOI (Digital Object Identifier) di cui saranno illustrate le caratteristiche e la storia.

Il DOI, come sarà più chiaro in seguito, non nasce con l'idea di limitarsi agli articoli scientifici, ma sta avendo successo proprio in quest'ambito perché già da tempo gli articoli sono in numero tale da porre diversi problemi su come renderli accessibili. Già Doyle [2, p. 5] formulava queste considerazioni: «In gran parte gli editori di

ZENO TAJOLI, CILEA, via Raffaello Sanzio 4, 20090 Segrate (Milano), e-mail tajoli@cilea.it.

Tutte le citazioni sono tradotte in italiano dall'autore. Per tutti i siti web l'ultima consultazione è stata effettuata il 25 ottobre 2004. Si ringraziano Susanna Mornati e Anna Maria Tammaro per i preziosi suggerimenti forniti e l'incoraggiamento ricevuto. Il presente lavoro nasce all'interno del Master internazionale a distanza in "Information studies" organizzato dall'University of Northumbria e dall'Università degli studi di Parma.

riviste scientifiche hanno raggiunto il loro primo risultato nel passaggio da un passato cartaceo a un futuro digitale. Ormai siamo in grado di immagazzinare gli articoli pubblicati nelle riviste sul web, pronti per essere inviati direttamente sul PC dei ricercatori.

Ora dobbiamo affrontare una sfida più difficile: collegare tutti questi articoli in un insieme coerente, facilmente navigabile da parte dei ricercatori che devono fare uso dell'informazione contenuta in essi.

Si può affermare che questo insieme coerente sarà basato su quattro elementi fondamentali che staranno alla base d'ogni robusto sistema di gestione dei collegamenti:

- *identificativi persistenti*: l'oggetto dei collegamenti (articoli di riviste scientifiche in questo caso) richiede identificativi ben definiti o nomi per collegamenti non ambigui;
- *validazione dei metadati*: certificare che i metadati riferiti a un articolo siano corretti e precisi;
- *mappe di collegamento dai metadati all'identificativo*: i metadati relativi a un articolo devono portare a un identificativo univoco e persistente;
- *manipolazione del collegamento*: trasformare un identificativo in uno specifico collegamento (o in un insieme di collegamenti)».

Il DOI si pone come risposta a queste necessità. Il suo ruolo sempre maggiore è visibile anche dal crescente numero d'istituzioni che se ne occupano [3].

2.1 Esigenze alla base del DOI

Il DOI nasce per rispondere alle esigenze degli editori nel contesto dei documenti digitali e della distribuzione in Internet delle opere tutelate da copyright.

Uno dei problemi centrali per gli editori è proprio la gestione dei diritti e la necessità di un sistema di controllo per opere in formato elettronico. La dizione usata inizialmente è Electronic Copyright Management System, ECMS, uno dei primi ad approfondire questi argomenti è stato Gervais [4]. Attualmente per indicare l'argomento e i sistemi informatici correlati si preferisce parlare di Digital Right Management (DRM). La presenza di un identificativo univoco per tutti gli oggetti digitali di cui si vogliono gestire i diritti è un'esigenza fondamentale per un funzionante sistema di DRM. Tuttavia i DRM non si esauriscono in questo e anche altre operazioni sono necessarie; ad esempio il controllo degli usi. Altri sono gli usi possibili per degli identificativi univoci: possono essere anche usati nelle operazioni di gestione interna rendendole più sicure e veloci.

Nel mondo della stampa tradizionale gli identificativi sono l'ISBN e l'ISSN. Tuttavia questi standard sono troppo legati alle caratteristiche della stampa su carta.

Il DOI è stato sviluppato pensando a questi punti [5, p. 3-6]:

- l'identificativo deve operare con diversa granularità a seconda delle necessità;
- gli URL non sono fatti per identificare ma indicano una localizzazione;
- bisogna gestire la transazione tra i nuovi e i vecchi supporti;
- queste caratteristiche sono considerate necessarie: univocità, interoperabilità, flessibilità, un vasto supporto.

2.2 Lo standard Z39.84-2000

Il DOI è anche un vero e proprio standard ANSI/NISO, lo Z39.84-2000. Essendo uno standard vi sono fonti d'informazione che risultano immediatamente come autorevoli: *in primis* lo standard medesimo [6] poi le pubblicazioni ufficiali della fondazione no-profit che gestisce il DOI, la International DOI Foundation (IDF). Queste pubblicazioni sono il manuale esplicativo dello standard e delle sue applicazioni [7] e il sito web dedicato al DOI [8].

In se stesso il DOI è un insieme di lettere senza un significato immediatamente rilevabile. La struttura di base è assai semplice:

Prefisso	Separatore	Suffisso
10.xxxxx	/	yyyyy

es: 10.1000/123456

- Tutti i DOI iniziano con “10.”;
- poi vi è un insieme di caratteri (una “stringa”) correlata all’organizzazione che ha creato quello specifico DOI;
- un carattere “/” come separatore;
- una stringa.

«È cruciale che la combinazione di prefisso e suffisso sia univoca in modo da garantire il corretto funzionamento di tutto il sistema correlato” [7, capitolo 2.3]. In generale una spiegazione completa e precisa della struttura del DOI è presente nel capitolo 2 del manuale riguardante il DOI.

Se il DOI in se stesso è abbastanza semplice, quello che ci sta dietro è abbastanza complesso e giustifica pienamente un manuale complesso. In genere nel testo l’insieme viene chiamato “the DOI system”.

2.3 Il dibattito riguardo il DOI

I punti di discussione sono stati numerosi. La proposta iniziale venne fatta dalla Association of American Publishers (AAP) [9]. Il DOI viene proposto come una lingua franca che permetta di collegare tra loro i diversi sistemi di identificazione (ISBN, ISSN, PII, SICI ecc.). Notevole enfasi viene data al fatto che non sarebbe stata la AAP a occuparsi del DOI e del sistema complessivo ma un’apposita fondazione, nata nel 1998, la International DOI Foundation. Le caratteristiche dell’agenzia, come viene prefigurata dalla proposta di AAP del 1996, sono così sintetizzabili [10]:

« 1. Sviluppare e mantenere un sistema dove i singoli oggetti (protetti dal copyright) possono essere identificati da identificativi univoci; [...] gestire e mantenere l’applicazione di un simile sistema a livello internazionale.

2. Creare e mantenere un sito web che permetta a ogni richiedente di accedere alle informazioni sull’oggetto desiderato con puntamento al database corrispondente; questo deve includere la gestione dei diritti e ogni informazione pertinente (...); il sito web gestisce tutte le richieste del pubblico e genera a una risposta all’utente.

3. Definire e rendere pubblici il formato, le funzionalità e le specifiche alle quali il database (...) deve essere compatibile. Il sito web e le specifiche del database non saranno proprietarie.

4. Creare linee guida e istruzioni per quelli che vorranno partecipare al sistema (...); fornire la necessaria assistenza per realizzare la loro interfaccia pubblica e connettersi al *network*».

Sia il lavoro di Paskin [10] che quello di Green e Bide [11] appaiono nell’aprile 1997 e discutono la proposta del DOI all’interno di articoli che trattano in generale dei diversi identificativi esistenti (SICI, BICI, PII ecc.) e del loro uso da parte degli editori. In entrambi gli articoli viene sottolineata la capacità del DOI di inglobare al suo interno gli altri identificativi, il suo proporsi come *metastandard*, rendendo così più facile un’integrazione tra diverse soluzioni. Una prima risposta del mondo delle biblioteche alla proposta del DOI avviene nell’ottobre 1997 da parte di Lynch [12] a cui risponde poco dopo Arms [13].

Lynch esprime notevoli preoccupazioni circa la natura del sistema, che gli appare molto commerciale e legato a una gestione del commercio elettronico di testi

digitali in modalità “paga per vedere”. Pone inoltre cinque specifiche questioni: 1. Quale è lo scopo del sistema? Che tipo di oggetti possono essere identificati con esso? Chi è autorizzato ad assegnare gli identificativi e come queste organizzazioni saranno identificate, registrate e autorizzate?

2. Quali sono le regole per assegnare nuovi identificativi? Quando due manifestazioni di un’opera sono considerate abbastanza simili da ricevere lo stesso identificativo all’interno del sistema e secondo quali criteri sono abbastanza diverse da ricevere due identificativi distinti? Quali gruppi beneficeranno delle distinzioni derivanti dall’assegnazione degli identificativi?

3. Come una persona può determinare l’identificativo di un’opera? Può ricavarlo dall’opera medesima o deve consultare un database mantenuto da terzi? A quale tipologia d’oggetti sono applicabili gli identificativi? Ci sarà un metodo automatico per costruire l’identificativo all’interno del sistema o qualcuno dovrà prendere una specifica decisione per assegnare un identificativo a un oggetto? Se sarà così, chi prenderà la decisione e perché?

4. Come viene risolto l’identificativo? O più esattamente, come un utente passa dall’identificativo all’opera identificata o ad altri identificativi o a metadati che permettono di localizzare l’opera e di accedervi? Inoltre, qual’è il ruolo di un possibile database proprietario nel risolvere l’identificativo? L’operatore o gli operatori di questo servizio di risoluzione avranno il monopolio della risoluzione? Quali sono le barriere alla partenza di un nuovo servizio di risoluzione? Quali sono le politiche dei servizi di risoluzione in campi come la *privacy* degli utenti e la raccolta di statistiche?

5. Quanto è persistente l’identificativo nel tempo? Si potrà risolvere un identificativo dopo che l’opera avrà cessato di essere sul mercato editoriale?

Nella risposta Arms [13] sottolinea che il DOI è sì un’iniziativa dell’Association of American Publishers, ma rimane aperta a eventuali altri partecipanti. Osserva inoltre che:

1. Poiché il DOI è un’iniziativa della AAP, identificherà materiali degli editori. Con il tempo emergerà chi va considerato un editore.

2. Ogni editore decide a suo giudizio quali sono gli elementi per cui assegna un DOI distinto. La granularità di cosa identifica un DOI è arbitraria. L’esperienza dirà quali saranno i criteri più utili.

3. Il DOI non è né sarà un indicatore costruito sulla base di informazioni esterne come SICI. Probabilmente verrà diffuso come metadato in formato elettronico all’interno di un raggruppamento standard di metadati come Dublin Core.

4. Il passaggio dal DOI al lavoro identificato dal DOI, la sua risoluzione, sarà sempre gratis per l’utente, questo è un principio base del DOI. Ora è la CNRI a provvedere alla risoluzione; ma questa situazione non durerà a lungo. Una fondazione no-profit vaglierà le politiche che verranno proposte per i servizi di risoluzione. Il sistema non manterrà i dati identificativi degli utenti.

5. Il problema non sono gli identificativi, ma la disponibilità del contenuto quando non avrà più valore commerciale. Bisognerà individuare un ente per la conservazione a lungo termine e fornirgli anche i DOI delle opere depositate.

Il DOI viene presentato ufficialmente alla Fiera di Francoforte del 1997; successivamente il dibattito sul DOI si svolge su una *mailing-list* dedicata, Discuss-doi [14], e su articoli della stampa specializzata. Se la nascita del DOI avviene sulla spinta di interessi commerciali, in particolare per la gestione di diritti dei formati digitali delle opere, Erickson [15] ricorda che lo scopo primario del DOI è di essere un identificativo persistente e univoco.

Due possono essere considerati i punti di partenza delle successive discussioni: il primo è espresso da Bide [5], che riassume la situazione del DOI all’inizio del 1998,

illustra le problematiche in discussione, fornisce alcuni scenari d'uso e ne analizza richieste e problemi. Gli scenari da lui ipotizzati sono: l'utente di periodici scientifici, l'utilizzatore di diritti per nuovi usi, l'utente finale, la retrodigitalizzazione. Pone anche la questione di come gestire il DOI in una prospettiva di lungo periodo. Il secondo è presentato da Rosenblatt [16], che ha partecipato al gruppo di lavoro dell'AAP. Rosenblatt fa capire chiaramente che gli editori aspirano unicamente a un mezzo per gestire il copyright in ambiente digitale, fornendo così esempi che possono essere classificati nell'ambito dell'*e-commerce*. Inoltre sottolinea che per gli editori è essenziale decidere in autonomia a che livello di specificità assegnare i DOI. Se per alcuni editori è sufficiente assegnare un unico DOI a una monografia, per altri invece può essere necessario assegnare un DOI a ogni capitolo di una monografia.

Successivamente Caplan [17] sottolinea che la natura commerciale del DOI porta anche conseguenze spiacevoli per l'utenza. Ad esempio cliccando su un DOI può apparire un *form* per il pagamento on-line. Propone anche di lasciare il DOI agli editori e a chi si occupa di gestire diritti. Le biblioteche farebbero meglio a creare un proprio identificativo.

Poiché il DOI è un stringa con nessun significato intrinseco, deve essere accompagnato da un insieme più o meno ricco di metadati. Questo tema è proposto in particolare da Morris [18 e 19]; la soluzione che la comunità DOI seleziona è di adottare il *framework* <indecs>. Questa decisione è annunciata da Barman e altri nel 1999 [20] e Paskin [21] rende pubblici i primi approcci del DOI al *framework* <indecs>. In particolare illustra gli elementi basilari dei metadati che accompagnano il DOI. La spiegazione completa di come ora il *framework* <indecs> viene usato nella comunità del DOI è disponibile nel capitolo 4 e nell'appendice 4 del manuale riguardante il DOI [7]. In Paskin [22] il tema degli identificativi viene affrontato a livello più generale, il DOI viene inserito all'interno di un contesto più ampio. Vengono affrontati e illustrati tutti i problemi del settore e lo stato dell'arte nel 1999. Fino a questo momento cliccando uno specifico DOI si ha come risultato uno specifico URL, la risoluzione è nella modalità più semplice, 1:1. Ma, come evidenzia nel 2000 Paskin, [22] la completa implementazione del DOI richiede una risoluzione uno a molti: «Il DOI è un identificativo azionabile: cliccare su un DOI genera una risposta. Una risposta "uno a uno" implica un'azione semplice e ripetibile, un DOI riceve come risposta un URL. Questa è stata l'implementazione iniziale. Tuttavia le esigenze reali degli utenti sono più complesse, ad un identificativo possono corrispondere diversi attributi e servizi e la somma di essi non può ridursi a un singolo URL. Il sistema alla base del DOI può incapsulare queste complessità, ma è richiesto il passaggio da un modello "uno ad uno" ad un modello "uno a molti". Questo è stato uno degli obiettivi del DOI sin dagli inizi dell'International DOI Foundation ed è una delle ragioni principali del ruolo di IDF come consorzio di studio e sviluppo.

Il problema di creare una funzionalità "uno a molti" non è specifico del DOI, ma il DOI è la prima implementazione pratica su larga scala.

"Uno a molti" implica che un DOI può essere usato per generare diverse azioni. Per rendere predicibile l'azione (significa scegliere uno o più d'uno tra i possibili risultati e manipolarli di conseguenza) bisogna aggiungere delle informazioni addizionali al parametro dell'azione. Gli strumenti per compiere questa specificazione sono i componenti del DOI: metadati e risoluzione.

Relazioni "uno a molti" sono prerequisiti per costruire una rete intelligente di relazioni tra entità identificate; una simile rete è necessaria per descrivere il concetto di "web semantico" elaborato dal W3C. Simili considerazioni si applicano a tutti quei modelli che descrivono relazioni complesse».

Questo cambiamento avviene alla partenza della prima agenzia che fornisce servizi commerciali con il DOI, CrossRef. Atkins e altri [24] riportano le prime esperienze sulla gestione dei DOI e le considerazioni che era possibile fare dopo una prima fase di sperimentazione. Il punto della situazione, dopo un anno d'attività, viene fatto da Brand [25] nel 2001. La possibilità che a un DOI vengano associate diverse copie in un rapporto "uno a molti", porta anche a usare il DOI per la soluzione del problema della "copia appropriata". Di questo si occupa l'articolo di Beit-Arie e altri [26]. In quest'articolo il DOI e CrossRef vengono inseriti all'interno della strutturazione di servizi estesi collegati tra loro tramite l'uso dello standard OpenURL. In particolare sono testati degli OpenURL costituiti dal solo DOI. Una volta che il DOI giunge al *resolver*¹ insieme alle informazioni sull'utente, il *resolver* interroga il database di CrossRef e ottiene i metadati collegati a quel DOI. Con il DOI, i metadati collegati e le informazioni sull'utente, il *resolver* costruisce i link alle copie appropriate. In [26] è sottolineato che il successo di questa operazione è dovuto all'uso di standard robusti e condivisi.

In parallelo, a partire dal 2000, Paskin emerge come referente ufficiale del DOI, cosa dovuta anche alla sua posizione di presidente dell'IDF. In [26] riassume la situazione del DOI nell'anno 2000 e le sue prospettive. In questo intervento sottolinea come il DOI sia una risposta ai problemi delle citazioni delle opere in formato digitale. Sottolinea che gli usi stabiliti in 500 anni di stampa su carta non possono essere abbandonati da un giorno all'altro ma che le risorse digitali hanno delle esigenze di carattere diverso. Nel 2002 propone un *Internet draft* per il DOI, l'*Uniform Resource Identifier (URI) scheme for Digital Object Identifiers* [28]. Il draft chiede la stringa DOI come prefisso registrato, descrive la sintassi dell'identificativo come definito in [6] ed elenca i caratteri riservati. Il draft passa 4 diverse versioni, tuttavia l'IDF non riesce a completare tutto il percorso necessario per essere uno standard di Internet, per diventare cioè un RFC. Questo ha significato che dopo 6 mesi dall'ultima modifica, l'*Internet draft* è stato cancellato dal sito ufficiale dell'Internet Engineering Task Force². Peranto ufficialmente non esiste più. L'ultimo articolo riassuntivo sul DOI è sempre di Paskin ed esce nel 2003 [29]. Paskin fa diverse considerazioni sulla situazione: appare chiaro che il DOI ha avuto un certo successo come identificativo permanente (nel giugno 2003 erano stati rilasciati 10 milioni di DOI), tuttavia rileva che servizi più sofisticati sono ancora su piccola scala; inoltre il DOI non è ancora preso seriamente in considerazione come identificativo per tutti quei contenuti digitali diversi dai testi (musica, video ecc.). Tuttavia la comunità DOI si è ampliata, non vi sono solo editori al suo interno; vi sono biblioteche nazionali e non, uffici per la documentazione governativa e della Comunità Europea.

Paskin, dopo un excursus storico sul DOI, esamina le caratteristiche salienti del DOI e le riassume con questi termini: Persistenza, Interoperabilità, Estensibilità, Efficienza, Aggiornamento dinamico. Successivamente Paskin affronta la gestione del sistema, in particolare come finanziare l'attività della IDF. Se all'inizio l'IDF è stata finanziata dai membri, per il futuro è necessario che essa venga finanziata tramite i servizi forniti dal DOI. Probabilmente le entrate del 2003 per la prima volta copriranno tutte le spese di IDF.

La riflessione di Paskin passa poi in rassegna le attività del DOI circa i metadati, in particolare il *framework* <indec> e ONIX. Infine si sofferma sull'iterazione DOI/OpenURL; gli esperimenti descritti in [26] sono proseguiti con notevole successo. Si può dire che il DOI si riserva l'identificazione a livello globale, che viene utilizzata dal

¹ Uno software dedicato al trattamento degli OpenURL come SFX, OpenURL Resolver e altri.

² <www.ietf.org>.

resolver OpenURL per contestualizzarla localmente e offrire dunque i servizi disponibili a quel particolare utente che si trova in quel particolare punto della rete.

Successivamente non ci sono stati degli articoli particolarmente innovativi sul DOI, tuttavia si ritiene utile quanto scritto da Vitiello e da Attanasio. Vitiello [30] ha fatto il punto sul tema degli identificativi. Esaminando la situazione tra la fine del 2003 e l'inizio del 2004, ha constatato che il DOI era l'identificativo di maggior successo tra quelli proposti per gestire oggetti digitali. Ha tuttavia fatto presente che lo scenario dell'identificazione delle risorse digitali rimaneva incerto e che evoluzioni future avrebbero potuto cambiare pesantemente la situazione. L'intervento di Attanasio [31] descrive un'interessante applicazione operativa del DOI facendo riferimento all'esperienza dell'editrice Casalini che vuole porsi sul mercato come aggregatore di contenuti digitali provenienti dai piccoli editori italiani. Usando il DOI e i servizi dell'agenzia di registrazione mEDRA³, Casalini aggrega più facilmente i contenuti che gli giungono dai piccoli editori e può proporre servizi innovativi; ad esempio unire, secondo specifiche richieste, testi provenienti da diverse fonti per costruire un pacchetto per gli studenti di un particolare corso universitario e gestire il pagamento dei relativi diritti.

3 Alcune considerazioni e conclusioni

Il DOI è dunque partito da una base relativamente semplice e tecnicamente non troppo complessa per realizzare servizi abbastanza sofisticati e diversificati. La collaborazione con altri standard come <in&dec> e OpenURL è stata uno dei suoi punti di forza.

La situazione nel campo degli identificativi è chiaramente ancora abbastanza fluida e qualche sorpresa può sempre capitare. Nel prossimo futuro un rivale del DOI sarà probabilmente l'identificativo OAI, compreso nel protocollo OAI-PMH [32]. A differenza del DOI ha una parte significativa, non è un identificativo completamente opaco. La parte significativa indica il sito web in cui si trova l'archivio che contiene l'opera collegata all'identificativo. Questo rende più semplice la costruzione di servizi di risoluzione e di identificazione, basandosi sul servizio di risoluzione alla base di Internet, il Domain Naming System (DNS). Tutto questo è indicato nel recente intervento di Powell, Young e Hickey [33]. In questo intervento viene spiegato come costruire dei Permanent URL (PURL) partendo da degli identificativi OAI costruiti secondo le apposite linee guida [34].

Questo permetterebbe all'identificativo OAI di offrire le stesse caratteristiche del DOI; tuttavia il fatto di avere una parte significativa legata al sistema DNS pone dei problemi nel medio-lungo periodo. Se il gestore dell'archivio cambia indirizzo web e non costruisce dei sistemi di reindirizzamento, ricompare il problema dei "link rotti". Il DOI centralizza questa manutenzione, dunque è necessario pagare un'entità centralizzata con questo compito. L'identificativo OAI pone questo costo a carico di ogni singolo gestore di archivio OAI-PMH.

Inoltre appare chiaro che gli autori non aggiungeranno nei riferimenti bibliografici né il DOI né l'identificativo OAI né altri. Dopo tutto l'ISBN esiste da 30 anni ma non viene messo nelle citazioni. Le citazioni riceveranno i link dinamici basati sugli identificativi da parte dei fornitori di servizi. Essi dovranno sviluppare *parser* per l'estrazione dei metadati dalle citazioni e con essi risalire all'identificativo. Probabilmente il successo o meno di un qualsiasi identificativo dipenderà in maniera significativa da chi fornirà il servizio migliore per questa operazione. Inoltre nella scelta tra DOI e identificativo OAI peserà la misura in cui si estenderà l'uso di diffondere la ricerca scientifica secondo il modello degli Open Archives.

3 <www.medra.org>.

RIFERIMENTI BIBLIOGRAFICI

- [1] Antonella De Robbio. *La biblioteca nel web, il web nella biblioteca*. «Bibliotime», 2 (1999), n. 2, <<http://www.spbo.unibo.it/bibliotime/num-ii-2/derobbio.htm>>.
- [2] Mark Doyle. *Citing and linking in electronic scholarly publishing*. «Learned publishing», 13 (2000), n. 1, p. 5-14.
- [3] International DOI Foundation. *Registration Agencies*. <http://www.doi.org/registration_agencies.html>, 2004.
- [4] Daniel J. Gervais. *Electronic rights management and digital identifier systems*. «JEP: the journal of electronic publishing», 4 (1999), n. 3, <<http://www.press.umich.edu/jep/04-03/gervais.html>>.
- [5] Mark Bide. *In search of the unicorn: the Digital Object Identifier from a user perspective; BNFR Report 89*. London: Book Industry Communication, 1998. <<http://www.bic.org.uk/unicorn2.pdf>>.
- [6] ANSI/NISO. *Standard Z39.84-2000: standard for the Digital Object Identifier*. Bethesda: NISO Press, 2000. <http://www.niso.org/standards/standard_detail.cfm?std_id=480>.
- [7] International DOI Foundation. *The DOI handbook*. La prima versione è del 1999, quella ora disponibile è dell'aprile 2004. <<http://www.doi.org/hb.html>>.
- [8] International DOI Foundation. *The Digital Object Identifier system*. Il sito nasce nel 1997, l'ultima revisione è del settembre 2004. <<http://www.doi.org>>.
- [9] Association of American Publishers. *AAP request for proposal to develop a Digital Object Identifier*. In: *Technology-based intellectual property management: electronic commerce for content technology-based solutions to managing rights for digital copyrighted works, Washington, 1996*, edited by Kahin B., Arms K. Annapolis: IMA, 1996, vol. 2, p. 69-70.
- [10] Norman Paskin. *Information identifiers*. «Learned publishing», 10 (1997), n. 2, p. 135-156. Il brano riportato è una citazione da *Digital Object Identifiers: an overview*, <<http://www.handle.net/doi/overview.html>> ma questo sito non è più disponibile.
- [11] Brian Green – Mark Bide. *Unique identifiers: a brief introduction*. «Against the grain», 9 (1997), n. 2, p. 75-78, 85.
- [12] Clifford Lynch. *Identifiers and their role in networked information applications*. «ARL: a bimonthly newsletter of research library issues and actions», 1997, n. 194, <<http://www.arl.org/newsltr/194/identifier.html>>.
- [13] William Y. Arms. *Digital Object Identifiers (DOIs) and Clifford Lynch's five questions on identifiers*, «ARL: a bimonthly newsletter of research library issues and actions», 1997, n. 194, <<http://www.arl.org/newsltr/194/arms.html>>.
- [14] *Discuss-doi* [lista di discussione online]. <<http://www.doi.org/mail-archive/discuss-doi/maillist.html>>.
- [15] John Erickson. *Commentary on "Multiple roles of identifiers" (Nolan Pope)*. «Discuss-doi» (Internet discussion list), 30 Jan 1998. <<http://www.doi.org/mail-archive/discuss-doi/msg00022.html>>.
- [16] Bill Rosenblatt. *Solving the dilemma of copyright protection online*. «JEP: the journal of electronic publishing», 3 (1997), n. 2, <<http://www.press.umich.edu/jep/03-02/doi.html>>.
- [17] Priscilla Caplan. *DOI o don't we?* «Public-access computer systems review», 9 (1998), n. 1, p. 1-4. <<http://info.lib.uh.edu/pr/v9/n1/capln1.html>>.
- [18] Sally Morris. *DOI metadata discussions*. «Discuss-doi» (Internet discussion list), 2 Feb 1998. <<http://www.doi.org/mail-archive/discuss-doi/msg00024.html>>.

- [19] Sally Morris. *No subject*. «Discuss-doi» (Internet discussion list), 20 Feb 1998. <<http://www.doi.org/mail-archive/discuss-doi/msg00048.html>>.
- [20] David Bearman. et al. *A common model to support interoperable metadata*. «D-lib magazine», 5 (1999), n. 1, <<http://www.dlib.org/dlib/january99/bearman/01bearman.html>>.
- [21] Norman Paskin. *DOI current status and outlook May 1999*. «D-lib magazine», 5 (1999), n. 5, <<http://www.dlib.org/dlib/may99/05paskin.html>>.
- [22] Norman Paskin. *Toward unique identifiers*. «Proceedings of the IEEE», 87 (1999), n. 7, p. 1208-1227.
- [23] Norman Paskin. *From one to many*. <<http://www.doi.org/one2many.pdf>>, 2000.
- [24] Helen Atkins et al. *Reference linking with DOIs. A case study*. «D-lib magazine», 6 (2000), n. 2, <<http://www.dlib.org/dlib/february00/02risher.html>>.
- [25] Amy Brand. *CrossRef turns one*. «D-lib magazine», 7 (2001), n. 5, <<http://www.dlib.org/dlib/may01/brand/05brand.html>>.
- [26] Oren Beit-Arie. et al. *Linking to the appropriate copy: report of a DOI-based prototype*. «D-lib magazine», 7 (2001), n. 9, <<http://www.dlib.org/dlib/september01/caplan/09caplan.html>>.
- [27] Norman Paskin. *E-citation: actionable identifiers and scholarly referencing*. «Learned publishing», 13 (2000), n. 3, p.159-168.
- [28] Norman Paskin et al. *Uniform Resource Identifier (URI) scheme for Digital Object Identifiers (DOIs)*. <<http://www.watersprings.org/pub/id/draft-paskin-doi-uri-00.txt>>, 2002. Il link indicato non è quello ufficiale della IETF, si tratta di un sito che mantiene gli *Internet draft* anche dopo che sono scaduti.
- [29] Norman Paskin. *DOI: a 2003 progress report*. «D-lib Magazine», 9 (2003), n. 6, <<http://www.dlib.org/dlib/june03/paskin/06paskin.html>>.
- [30] Giuseppe Vitiello. *L' identificazione degli identificatori*. «Biblioteche oggi», 22 (2004), n. 2, p. 67-80. Disponibile anche a: <<http://www.bibliotecheoggi.it/2004/20040206701.pdf>>.
- [31] Piero Attanasio. *The use of DOI in eContent value chain*. <http://reg.medra.org/documents/BC_mEDRA_Casalini.pdf>, 2004.
- [32] *The Open Archives Initiative protocol for metadata harvesting*. <<http://www.openarchives.org/OAI/openarchivesprotocol.html>>, 2004.
- [33] Andy Powell – Jeff Young – Thomas Hickey. *The PURL-based Object Identifier (POI)*. <<http://www.ukoln.ac.uk/distributed-systems/poi/>>, 2004.
- [34] *OAI identifier format*. <<http://www.openarchives.org/OAI/2.0/guidelines-oai-identifier.htm>>, 2004.

DOI: an instrument for building the digital library

by Zeno Tajoli

The Digital Object Identifier (DOI) is born to address to the needs of publishers in the context of digital documents and of the distribution via Internet of works protected by copyright. The syntax of DOI is quite simple; one of the most important characteristics is to be a “dump number”.

The initial proposal sorts from the Association of American Publishers in 1996. DOI came as a common base that allows to connect among different identification systems, like ISBN, ISSN, PII, SICI ect. With notable emphasis the fact is stressed that AAP did not intend to deal with DOI and with the DOI system but a specific foundation, born in 1998, the International DOI Foundation. The discussion about it was developed in particular between 1997 and 2002, when the infrastructure was built that served to manage and to promote the DOI. Several interventions specified the characteristics and the role of DOI. It was decided to use the <indecs> framework to describe the set of the metadata elements connect with DOI. Norman Paskin, director of IDF, emerges as referring officer of DOI. In the beginning resolution was “one to one”, but DOI requests a resolution “one to many”. The change was done when the agency CrossRef started to offer its services to a wider public. This passage facilitates further developments, such as the integration of the standard OpenURL.

It is suggested that a possible competitor for the DOI will be the OAI identifier. However it is less robust than the DOI. The future developments of the situation will depend from the quality of related services and from the proliferation of open archives.

ZENO TAJOLI, CILEA, via Raffaello Sanzio 4, 20090 Segrate (Milano), e-mail <tajoli@cilea.it>.

Bollettino **AIB**, ISSN 1121-1490, vol. 45 n. 1 (March 2005), p. 75-84.