

Elena Giglia

DOI outreach conference, Milano 21 novembre 2014



Abstract

The DOI outreach conference focused on services enabled by DOIs and demonstrated by use cases showing applications and external projects within the scholarly communication system, the dataset world, the teaching and learning environment, and the entertainment market. With the value added services presented during the conference and summarized by this report, DOI is going to be in the next future a fundamental tool in the digital communication.

Premessa

Il DOI, Digital Object Identifier, è uno standard nato alla fine degli anni Novanta dall'esigenza comune di identificare univocamente e in modo persistente gli oggetti in rete [1]. Si tratta di un "identificativo digitale", *digital identifier*, di un oggetto – e non, distinzione sottile ma necessaria, di un "identificativo di un oggetto digitale". Il DOI può essere assegnato, in virtù della sua fine granularità, a un testo digitale, a un set di dati, a un'immagine, a un oggetto fisico (come ad esempio un DVD): a titolo di curiosità, persino le "Philosophical Transactions" di Henry Holdenburg, 1665, hanno ora il loro DOI [2]. A partire dal lancio nel 2000, a oggi si contano più di 100 milioni di DOI assegnati e circa due miliardi di risoluzioni di DOI per anno, indice del fatto che il servizio risponde a precise necessità della rete.

Con questi dati, utili a fornire il contesto di riferimento, si è aperta la 2014 DOI Outreach Conference, organizzata da mEDRA a Milano lo scorso 21 novembre, le cui presentazioni sono disponibili sulla pagina della International DOI Foundation [3].

La rete è mobile

Norman Paskin, primo direttore della International DOI Foundation, ha posto l'attenzione su un dato di fondo: gli accessi a Internet avvengono ormai in maggioranza da dispositivi mobili, e questo deve far ripensare profondamente una serie di paradigmi, compresa l'offerta di servizi, che vanno rimodulati in base al comportamento degli utilizzatori. È una notazione valida non solo per il mondo dell'editoria scientifica ma anche per tutti gli intermediari/facilitatori dell'accesso all'informazione, biblioteche per prime [4].

Paskin ha poi evidenziato il triplice ruolo della fondazione DOI, che è quello di offrire una infrastruttura tecnologica, atta alla risoluzione dei milioni di DOI e al mantenimento del *data model* sottostante; una infrastruttura per così dire sociale, che consiste nell'obbligo alla persistenza delle risoluzioni stesse e al loro *backup*; e infine una sorta di infrastruttura di facilitazione, per affrontare temi comuni quali l'interoperabilità e la standardizzazione (ricordiamo che il DOI è uno standard ISO dal 2012, ISO 26324-2012).

"Interoperabilità" e "standardizzazione" hanno fatto da filo conduttore per l'intera giornata, con un *focus* sui servizi a valore aggiunto forniti dall'uso del DOI nell'ambito della comunicazione scientifica ma non solo: Raymond Drewry ha illustrato le potenzialità del DOI e l'ottimo riscontro ottenuto dalla nuova piattaforma EIDR (<<http://eidr.org/>>), registro unico per materiale audiovisivo, dai film alle serie televisive, per l'industria dell'intrattenimento. In questo caso l'utilizzo del DOI è stato di fondamentale supporto al mercato.

Il meraviglioso mondo di DOI

Le relazioni della prima sessione hanno mostrato, da diversi punti di vista, l'importanza degli identificativi univoci per ogni entità coinvolta nella catena della produzione della comunicazione scientifica.

Ed Pentz, di CrossRef (<<http://www.crossref.org/>>), ha sottolineato l'uso dinamico del DOI per la creazione di link persistenti nei riferimenti bibliografici degli articoli scientifici, un servizio molto apprezzato sia da lettori sia da editori per evitare il temuto messaggio "Error 404 – Not found". Il CrossRef DOI assicura per i circa 70 milioni di record presenti la persistenza del link grazie alla partnership con gli oltre 5.000 "enti", per la maggior parte editori, che associano DOI ad oggetti - siano essi articoli, capitoli, interi volumi, e che si impegnano in un clima di mutua fiducia a mantenere attivo il collegamento segnalando ogni variazione alla DOI Foundation.

Fra i servizi implementati di recente si segnalano CrossRef-Cited by (<<http://www.crossref.org/citedby/index.html>>), che permette di tracciare con la precisione assoluta legata al DOI le citazioni in ingresso, con il vantaggio di non avere barriere disciplinari né di tipologia di contenuto, e Text and data mining services (<<http://www.crossref.org/tdm/>>), teso a identificare diritti e sito di accesso al testo integrale del lavoro scientifico da analizzare

Particolare interesse riveste il nuovo FundRef (<<http://www.crossref.org/fundref/index.html>>), un registro di identificativi di enti finanziatori della ricerca di estrema utilità per collegare l'ente finanziatore o il numero di *grant* (finora desumibile solo dalla lettura della sezione *Acknowledgments*) ai prodotti della ricerca generati.

Attraverso i servizi attivabili via DOI si possono connettere quindi in modo persistente identificativi, metadati e servizi, il che si traduce a sua volta in un circolo virtuoso fra collegamenti persistenti, la possibilità di trovare il contenuto, e la fiducia nel sistema di editoria scientifica.

I dati sono parte della ricerca

Il consorzio Datacite (<<https://www.datacite.org/>>), fin dai suoi esordi, si è focalizzato sulla ricercabilità e accessibilità dei set di dati, e assegna a oggi circa 4 milioni di DOI. Jan Brase ha ricordato come per "dati" si intenda tutto ciò che sta alla base della ricerca - indipendentemente dalla estrema eterogeneità dei dati stessi, che possono essere studi medici, video, risultati di esperimenti, immagini, modelli: per questo motivo è necessario che siano identificabili, ricercabili e citabili univocamente attraverso il DOI.

Un importante risultato è il recente accordo fra Datacite e CrossRef per una maggiore interoperabilità fra i due sistemi, in modo da permettere l'attivazione di link bidirezionali fra l'articolo pubblicato e il relativo set di dati, e dare accesso al testo integrale e al data set con un solo passaggio [5].

Altrettanto strategico il progetto europeo ODIN (<<http://odin-project.eu>>), per l'interoperabilità fra DOI e ORCID (<<http://orcid.org/>>), l'identificativo univoco per gli autori, utile a creare un sistema della ricerca ancora più integrato.

DataCite è tra gli estensori degli 8 punti dei Data Citation Principles, fondamentali per creare una cultura della condivisione dei dati, ovvero Importance, Credit and Attribution, Evidence, Unique Identification, Access, Persistence, Specificity and Verifiability, Interoperability and Flexibility [6].

Identificatori univoci alla base di una valutazione della ricerca efficace

Susanna Mornati (CINECA) ha presentato un'applicazione pratica dei servizi offerti dagli identificatori univoci nell'ecosistema dei Research Information Management Systems, i nuovi strumenti che cercano di offrire una visione d'insieme della ricerca, a partire dal finanziamento fino ai risultati e all'impatto che questi hanno sulla comunità scientifica.

Il sistema di valutazione e valorizzazione della ricerca di CINECA, il nuovo IRIS, fa dell'interoperabilità e dell'integrazione di protocolli e formati standard il suo punto di forza.

Basato come *data model* sullo standard europeo CERIF (<<http://www.eurocris.org/Index.php?page=CERIFreleases&t=1>>), e su Linked Data come contesto di riferimento (<<http://linkeddata.org/>>), IRIS adotta gli identificativi DOI, Handle, ORCID e FundRef. Dalla sinergia fra questi elementi nasce un sistema di monitoraggio che permette intanto l'*import* dei dati bibliografici via DOI dalle principali banche dati, garantendo la qualità dei dati stessi nell'anagrafe della ricerca, e consente poi la navigazione fra numero del finanziamento, prodotti della ricerca ad esso collegati, prodotti di un singolo autore, citazioni e così via, per fornire al valutatore un cruscotto di informazioni funzionale ad una valutazione efficace. Grazie agli identificativi univoci il record diviene robusto, persistente, ricercabile e ritrovabile per il ricercatore, e valutabile in modo efficace per gli amministratori.

Progetti e servizi avanzati

Oltre alla fondamentale notazione dell'importanza che oggi ha assunto il concetto di *machine-readable* per dati, metadati e applicazioni, nelle altre sessioni sono state presentati sia utilizzi specifici del DOI - a Taiwan il consorzio AIRITI lo utilizza per corsi online e materiale didattico, con ottimi riscontri in termini di ricercabilità e di uso - sia progetti e servizi avanzati.

LCC, Linked Content Coalition (<<http://www.linkedcontentcoalition.org/>>), è una iniziativa delle agenzie di registrazione tesa alla gestione dei diritti del contenuto in rete, attraverso l'uso sinergico degli identificativi delle "creazioni" e delle "parti" portatrici di diritti sulle creazioni stesse. Di particolare importanza il nuovo standard ISNI, International Standard Name Identifier (<<http://www.isni.org/>>), che può agire come identificativo in sé o come "hub", una sorta di sovra-identificativo che associa e correla diversi identificativi di una "parte" - sia essa un autore, un editore, un ente - in modo da renderli sostituibili in sistemi differenti. Un primo frutto di questa coalizione è il progetto europeo RDI - Rights Data Network (<<http://www.rdi-project.org/>>), che mira all'identificazione univoca dei portatori di diritti su un oggetto digitale, che sia testo, audio, video.

La novità tecnica più interessante sembra la *multiple resolution*, basata sull'associazione di servizi all'identificativo: il DOI può avere in sé una serie di informazioni e può rilasciarle in modalità sensibile al contesto, ovvero secondo l'interesse specifico dell'utente. Invece di una risoluzione singola - come al momento: al DOI corrisponde un oggetto - in un futuro prossimo si potrà avere

una risoluzione multipla, per cui al DOI possono essere associati sia un oggetto sia i diritti connessi sia altri servizi, che l'utente sceglie da un elenco in base al suo specifico interesse o prospettiva.

Elena Giglia, Direzione Sviluppo Organizzativo, Innovazione e Servizi Bibliotecari - Università di Torino, e-mail: elena.giglia@unito.it

Note

[1] International DOI Foundation, <<http://www.doi.org>>. Su questo sito verranno messe a disposizione le slides della giornata.

[2] "Philosophical transactions", <<http://rstl.royalsocietypublishing.org/content/1/1-22.toc>>.

[3] DOI Outreach 2014, <http://www.doi.org/topics/outreach_mtgs.html>.

[4] J. Naughton, *The future of technology? Is in your hands*, "The Guardian", 2014 November, 2, <<http://www.theguardian.com/technology/2014/nov/02/future-of-technology-in-your-hands-smartphone>>; cita uno studio di B. Evans, *Mobile are eating the world*, 2014 October, 28, disponibile su Slideshare: <<http://ben-evans.com/benedictevans/2014/10/28/presentation-mobile-is-eating-the-world>>.

[5] *CrossRef and DataCite announce new initiative to accelerate the adoption of DOIs for data publication and citation*, 10 novembre 2014, <<https://www.datacite.org/CrossRefDataCiteinitiative>>.

[6] FORCE11, *Data citation principles*, <<https://www.force11.org/datacitation>>.