



Utilizzo del DOI (*Digital Object Identifier*) nei progetti di *digital humanities*

Giovanni Salucci

Nel mondo delle *Digital Humanities* (DH) i risultati prodotti sono spesso nella forma di banche dati, siti internet, blog o pubblicazioni digitali di varia forma; per gli autori è difficile ottenere il riconoscimento dei crediti e garantirne la tracciabilità. In questo articolo si descrive il sistema del DOI (utilizzato come identificatore persistente assegnato di norma a monografie e articoli scientifici per la loro identificazione e tracciabilità) proponendone l'adozione in alcuni contesti di progetti di DH per favorirne la identificazione e diffusione, attraverso anche un ampio set di metadati descrittivi.

In the world of Digital Humanities (DH), the results produced often take the form of databases, websites, blogs, or various digital publications. For authors, it is difficult to obtain recognition and ensure traceability of these results. This article describes the DOI system (used as a persistent identifier typically assigned to monographs and scientific articles for their identification and traceability), proposing its adoption in some DH project contexts to promote identification and dissemination, also through a wide set of descriptive metadata.

Parole chiave: DH, Informatica umanistica, DOI, Metadati, Identificatori persistenti

Keywords: DH, Digital Humanities, DOI, Metadata, Persistent Identifiers, PID

Sommario: Introduzione - Cos'è il DOI e a che cosa serve - Le diverse agenzie di registrazione e i *repositories* - Vantaggi e svantaggi nell'utilizzo del DOI nei progetti di DH - Alcuni possibili casi di utilizzo - Conclusioni

Peer review

Submitted 01/12/2022

Accepted 20/12/2022

Published 25/01/2023

Open access

© 2023 | Attribution - Non commercial - Non derivatives (IT)

Cita come

Giovanni Salucci, *Utilizzo del DOI (Digital Object Identifier) nei progetti di digital humanities* in Rivista DILEF - II, 2022/2 (gennaio-dicembre), pp. 308-319. 10.35948/DILEF/2023.4307

DOI

10.35948/DILEF/2023.4307

Introduzione

Nicola Barbuti (Barbuti 2019) afferma che «nella rivoluzione digitale in corso, solo da poco ci si è resi conto dei rischi connessi alla volatilità dei dati digitali e alla mancanza di requisiti che garantiscano la conservazione e il trasferimento della memoria dei processi di vita», e questo si applica non solo alla conversione di opere analogiche, ma soprattutto all'informazione che nasce digitale.

Nell'ambito della ricerca scientifica e della realizzazione di prodotti scientifici, sia l'archiviazione dei dati sia la pubblicazione dei risultati, specialmente nella forma di articoli in rivista, si appoggiano su solide basi metodologiche (con standard internazionali per formati e supporti) e su infrastrutture tecnologiche con linee guida informatiche condivise. Ad esempio, nelle riviste accademiche, gli articoli sono disponibili in rete in formato PDF, HTML e/o XML (JATS) ad accesso libero o dietro abbonamento, e sono ricercabili grazie ad un set di metadati con cui vengono indicizzati nei principali aggregatori scientifici e banche dati settoriali. Inoltre, gli articoli e i metadati sono archiviati in *repositories*¹ che garantiscono la preservazione a lungo termine; per preservazione (digitale) in questo contesto si intende l'insieme di tutte le attività necessarie a garantire nel tempo e nello spazio l'accesso alla risorsa digitale.

Anche per le monografie, per i volumi scientifici collettanei e per gli altri formati con cui vengono pubblicati i risultati della ricerca scientifica (quali ad esempio blog o siti internet) accanto alle edizioni cartacee sono realizzate edizioni digitali che utilizzano supporti e metadati standard per la diffusione e la fruizione da parte degli interessati. Identificatori, metadati e classificazioni tematiche sono anche alla base del funzionamento della catena di distribuzione dei libri (Attanasio 2022); il primo scopo della gestione dei metadati è quello di aumentare l'efficienza della filiera distributiva, e come visto in precedenza questa necessità vale anche per gli articoli scientifici e gli altri prodotti della ricerca, quelli ad esempio generati nei progetti di *digital humanities* (DH).

Sia la preservazione che la diffusione necessitano a monte della identificazione univoca e persistente, cioè durevole nel tempo, della risorsa digitale cui si riferiscono. Gli identificatori univoci sono essenziali per la gestione dell'informazione in ogni contesto digitale; se anche persistenti sono un elemento fondamentale all'interno della infrastruttura della ricerca scientifica per i ricercatori, per le loro istituzioni di appartenenza e finanziatori, per gli oggetti a cui si riferiscono (Meadows 2019). Un identificatore assegnato da un primo soggetto in un certo momento può essere utilizzato da un secondo soggetto in un differente contesto e in un altro momento, senza che il secondo soggetto debba conoscere o chiedere autorizzazione al primo soggetto. Il concetto di persistenza estende questo concetto di riutilizzo alla dimensione temporale, svincolando la responsabilità dell'assegnatario e delegandola

ad una infrastruttura che possa garantirne la durata. In questo articolo analizzeremo il DOI (*Digital Object Identifier*) - la cui struttura, infrastruttura e funzionamento di base sono ben descritti in Paskin 2005 – valutando il suo possibile utilizzo come identificativo univoco e persistente nei progetti di DH.

Cos'è il DOI e a che cosa serve

Il DOI è allo stesso tempo sia il codice che identifica una entità (oggetto fisico o logico) nella rete digitale sia la infrastruttura applicativa nata a fine degli anni Novanta per gestire l'identificazione degli oggetti digitali nella rete; il DOI rientra quindi a pieno titolo nell'insieme dei cosiddetti PIDs (*Persistent IDentifiers*, in italiano: identificatori persistenti), e quindi anche ad esso possono applicarsi le riflessioni generali introdotte in precedenza.

Il DOI rappresenta un sistema interattivo per la identificazione persistente e lo scambio interoperabile di informazioni in rete; si può applicare a risorse di differenti tipologie, principalmente per condividerne le proprietà all'interno di una comunità di interesse, o per identificarle univocamente in vista della gestione della proprietà intellettuale. Altro elemento da tenere in considerazione è che lo stesso DOI può coprire l'intero ciclo di vita di una risorsa, dalla ideazione alla pubblicazione, attraverso la modifica e l'adeguamento dei metadati associati allo stesso oggetto durante le fasi di realizzazione.

Il DOI, pur essendo potenzialmente un sistema di identificazione persistente di qualunque risorsa digitale, di fatto ha trovato il suo campo principale di utilizzo nel settore della editoria accademica, che per prima lo ha utilizzato massicciamente già dal 1998. Le edizioni, cartacee o digitali, hanno un utilizzo preminente nell'editoria accademica in quanto rappresentano la sorgente primaria e le fondamenta delle discussioni e proposte scientifiche in tutte le discipline. Utilizzare una citazione non ambigua di un testo pubblicato è una necessità di qualunque pubblicazione scientifica, dato che rende rintracciabili e trasparenti le fonti alla base dei processi utilizzati all'interno della pubblicazione; la trasformazione dal cartaceo al digitale non ha modificato questa necessità, ma ha reso più difficoltosa questa operazione, con la moltiplicazione della disponibilità di una risorsa, attraverso URL (*Uniform Resource Locator*, cioè indirizzo di una risorsa nella rete) più o meno affidabili.

L'utilizzo del DOI è quindi prerogativa degli editori (non solo gli editori commerciali, ma in generale chiunque abbia la responsabilità della pubblicazione o diffusione dei contenuti); nel corso degli anni il DOI ha esteso il suo campo di uso, dapprima limitato a monografie e articoli, ad altri “prodotti” della ricerca scientifica quali i capitoli, i report, i preprint, le peer review, le presentazioni a convegni, ecc.

La registrazione di un DOI, in sintesi, comporta tre elementi:

- il codice DOI assegnato (di cui la prima parte -prefisso- è assegnata all'editore dall'agenzia di registrazione, la seconda -suffisso- è generata autonomamente dall'editore purché univoca e nel rispetto di alcune semplici regole²), che non cambia una volta assegnato
- l'URL a cui il DOI punta, che dà accesso alla risorsa digitale, e che va mantenuto aggiornato nel tempo;
- l'insieme dei metadati che descrivono la risorsa, a cominciare dalla tipologia di risorsa cui il DOI è assegnato.

È possibile che una risorsa accademica, così come qualunque altra risorsa sul web, cambi posizione sulla rete: se fosse stata citata o *linkata* direttamente, attraverso l'URL, si genererebbe un cosiddetto "link rotto"; citata attraverso il DOI, invece, rimane accessibile in quanto, proprio per come funziona il DOI, sebbene la risorsa abbia cambiato l'URL, il suo identificativo DOI resta lo stesso. Naturalmente, la risorsa sarà rintracciabile purché chi ha la gestione di quel DOI abbia provveduto ad aggiornare il puntamento dell'URL.

Per completezza, occorre sottolineare come mantenere aggiornato il *mapping* (cioè la corrispondenza) tra il DOI e l'indirizzo URL a cui risolvere non è automaticamente garanzia di ottenere la risorsa digitale associata al DOI; come mostrato da Klein e Balakireva (2020) la persistenza dell'identificatore è condizione necessaria ma non sufficiente per la persistenza della risorsa: se l'archiviazione sul web non soddisfa ai requisiti di persistenza, è possibile che il DOI funzioni, puntando al giusto URL, ma che ugualmente la risorsa digitale risulti indisponibile (ad esempio perché danneggiata).

Le diverse agenzie di registrazione e i *repositories*

Per assegnare e registrare i DOI, una istituzione scientifica (o casa editrice) deve appoggiarsi ad una agenzia di registrazione, accreditata al consorzio DOI, che rilascia il prefisso da usare per le registrazioni; per una realtà italiana, la scelta della agenzia per registrare i DOI si limita a tre opzioni: mEDRA, Crossref e Datacite. È possibile associarsi anche a più di una agenzia, anche se ovviamente a ciascuna risorsa può essere assegnato un solo DOI; usare più di una agenzia complica la gestione (per la predisposizione di metadati con tracciati differenti) e aumenta i costi, come descriveremo più avanti.

È possibile assegnare un DOI a diverse tipologie di risorse, ma non tutte sono gestite allo stesso modo dalle varie agenzie; questo dipende proprio dalla storia di formazione di ciascuna agenzia, che è nata e si sviluppa avendo un proprio *core business* e una propria organizzazione e scopo, differente da caso a caso. Ad esempio, mEDRA³ specializzata nella distribuzione libraria ed e-commerce, consente di

registrare DOI alle monografie (nei vari formati) e agli articoli di riviste, ma a differenza di Crossref⁴ non consente di registrare DOI per i preprint o le peer review, che sono di fatto altri elementi importanti della editoria accademica. Datacite⁵ invece è specializzata nella registrazione di DOI da assegnare a *dataset*, mappe, rilevazioni scientifiche, e consente un'ampia copertura delle tipologie di oggetti.

Nella scelta della agenzia occorrerà tenere conto della tipologia di oggetti gestita⁶, dello schema dei metadati previsti per ciascun caso e del costo di registrazione.

Il costo per ottenere il prefisso e quindi essere ammessi a registrare un DOI è variabile, dato che ciascuna agenzia applica criteri diversi; Datacite prevede una quota associativa e una tariffa per la registrazione dei DOI che dipendono entrambe dal numero di DOI registrati⁷; Crossref richiede il pagamento di una quota associativa annuale che dipende dal fatturato del soggetto richiedente e poi in aggiunta una tariffa di 1 euro (o 25 centesimi di euro, a seconda della tipologia) per ciascuna registrazione⁸; mEDRA applica un listino che prevede una quota annuale in base al numero di DOI e poi un costo variabile per ciascun DOI registrato in eccesso⁹.

Se già non è semplice confrontare i costi di registrazioni a causa di queste differenze, la situazione è ancora più complessa se volessimo prevedere il costo totale di questa operazione; infatti per un computo totale occorre sommare, oltre al costo di attribuzione del DOI da versare all'agenzia visto in precedenza, anche (o soprattutto) il costo di *data-entry* necessario per la predisposizione dei metadati di registrazione e la trasmissione all'agenzia ai fini della registrazione. Le operazioni di *data-entry* dei metadati e di registrazione possono essere svolte manualmente oppure in modalità semi-automatizzata; le agenzie di registrazione infatti mettono a disposizione interfacce web e API¹⁰ con le quali effettuare rispettivamente l'inserimento dati manuale oppure la registrazione diretta tramite un flusso dati predisposto dal registrante. In questo caso, la generazione del flusso è una operazione svolta solitamente attraverso software specializzati che hanno un costo di realizzazione o almeno di mantenimento, da conteggiare quindi nel computo totale.

Il costo per la compilazione di metadati di qualità, cioè attendibili e al massimo livello di completezza, è un elemento da tenere in considerazione nella scelta dell'agenzia con cui registrare i DOI. Nella registrazione di una risorsa, infatti, l'insieme dei metadati previsti (obbligatori e facoltativi) varia molto da agenzia ad agenzia, sia nel tracciato (la struttura e il numero degli elementi previsti) sia nella sintassi di riempimento di ogni elemento. In linea di massima, Datacite è l'agenzia che prevede un numero inferiore di dati obbligatori; mEDRA si posiziona in seconda posizione ma gestisce poche tipologie di oggetti; infine Crossref rappresenta l'agenzia che prevede (e pretende) un set più ampio e completo di metadati. Per una realtà che debba registrare da una decina fino ad un centinaio di DOI l'anno, il prezzo da versare all'agenzia è più o meno lo stesso e si aggira intorno ai cinquecento euro; questa cifra,

come detto, rappresenta solo i costi fissi di adesione e registrazione dei DOI all'agenzia scelta, e non il costo totale dell'operazione.

Negli ultimi anni si sono resi disponibili alcuni *repositories* istituzionali o privati, tematici o generalisti, che consentono non solo di archiviare gli oggetti digitali, ma di assegnare loro in automatico un DOI al momento del deposito.

Tra tutti, segnaliamo Zenodo¹¹, gestito dal CERN inizialmente per conservare i dati prodotti dall'acceleratore di Ginevra, e poi finanziato dalla comunità Europea per favorire il diffondersi della *Open Science*, è stato aperto al deposito di dati e risultati della ricerca scientifica prodotti da qualunque soggetto pubblico o privato. Altri esempi sono Figshare¹² oppure i vari *repositories* e *preprint server* tematici ArXiv, BioRxiv, ChemRxiv, MedRxiv e altri che stanno nascendo. In tutti questi casi l'accesso al servizio è attualmente gratuito, ma tutto il materiale caricato deve essere distribuito ad accesso aperto con licenze *Creative Commons*; il DOI (e la relativa gestione) viene assegnato al *repository* che ne curerà il corretto puntamento alla risorsa digitale: la risorsa stessa sarà conservata nel *repository* che si occuperà della persistenza e durata nel tempo.

Vantaggi e svantaggi nell'utilizzo del DOI nei progetti di DH

I dati rappresentano un primo risultato della ricerca, ma formano anche la base per nuove ipotesi, utilizzati da altri ricercatori in contesti magari differenti da quelli originali per i quali erano stati prodotti o raccolti (Roueda et al. 2016). Possono quindi essere utilizzati vantaggiosamente da ricercatori, studiosi e istituzioni, purché siano resi disponibili e nel rispetto di adeguate politiche di diffusione. La disponibilità, rintracciabilità e riproducibilità dei risultati della ricerca scientifica sono fattori decisivi per l'avanzamento della conoscenza, e il DOI costituisce un elemento importante in questo scenario. L'uso di identificativi persistenti quali il DOI costituisce il prerequisito necessario anche nella applicazione dei cosiddetti principi FAIR (*Findable, Accessible, Interoperable and Reusable*) introdotti nell'ambito della scienza aperta (Wilkinson et al. 2016).

Nei settori scientifici STM (scienza, tecnica, medicina) l'assegnazione del DOI ai *dataset* è una pratica abbastanza diffusa, spesso arrivando ad assegnare singoli DOI a ciascun *record* della raccolta. Il progetto UNITE¹³ ad esempio è una banca dati sviluppata in internet e un ambiente di sequenziamento molecolare per la identificazione dei funghi; contiene oltre un milione di sequenze raggruppate in oltre 400.000 possibili specie, a ciascuna delle quali è stato assegnato un DOI per promuoverne l'utilizzo non ambiguo. I promotori del progetto hanno utilizzato Datacite come agenzia, e se andiamo a vedere un esempio di registrazione (la tassonomia TH000120¹⁴) nei metadati di registrazione troviamo il titolo, il nome dell'Istituzione, nomi e cognomi dei curatori, la data di rilascio, la licenza per la

condivisione del record e un abstract generico, che viene ripetuto uguale per tutti i record della stessa collezione. In presenza di una banca dati così ampia, se si adotta il livello di granularità che assegna un DOI a ciascun *record*, è normale che nella generazione dei metadati si predisponga un tracciato minimo e laddove possibile si individuino dei valori comuni tra tutti i *record*, di modo da ridurre il più possibile il costo di generazione dei metadati stesso.

Sempre nel settore STM i dati degli esperimenti scientifici sono alla base del metodo sperimentale, e la loro condivisione per la verifica della riproducibilità dell'esperimento è alla base della scienza stessa; si pensi ad esempio ai dati raccolti nel progetto VLIZ¹⁵ a tema marino, oppure a quelli prodotti in ambito spaziale da ESAC e descritti in Masson et al. (2021).

Nel settore delle *digital humanities*, ci sono esempi di progetti a cui sono stati assegnati DOI, ma la registrazione generalmente si riferisce all'intera collezione (registrata come "banca dati"), o a gruppi di *record*/schede. Nell'ambito della linguistica internazionale, "Twenty-two Historical Encyclopedias Encoded in TEI: a New Resource for the Digital Humanities"¹⁶ oppure progetti italiani, ad esempio, "Lingua e lessico del gaming italiano"¹⁷ oppure "L'ittionimia nel Vocabolario dei Dialetti Salentini"¹⁸. Ma quali sono i vantaggi di utilizzare il DOI in un progetto di *digital humanities*? Innanzi tutto, il DOI garantisce la persistenza, e questo si applica sia alla sua identificazione che alla sua citabilità.

In secondo luogo il DOI favorisce la interoperabilità tra sistemi informatici, grazie alla presenza dei metadati con cui i sistemi sono in grado di scambiare informazioni tra loro e di essere poi in grado di utilizzarle. I metadati vengono utilizzati non solo da aggregatori accademici, ma anche dai motori di ricerca generalisti, come indicato in un recente studio preliminare sulla presenza dei DOI in Google¹⁹.

Il DOI inoltre costituisce un sistema estensibile, in quanto consente di creare delle relazioni stabili tra risorse di varie tipologie, collegate tra di loro; inoltre consente la caratterizzazione di dati e metadati da poter applicare in una volta sola ad un insieme di formati differenti (si pensi ad esempio ai vari supporti o formati con cui viene pubblicato un articolo scientifico o un volume).

Infine, il DOI consente un aggiornamento dinamico dei metadati, per mantenere aggiornate le informazioni sulla risorsa. L'aggiornamento è a cura dell'ente "titolare" del DOI assegnato, e la propagazione delle modifiche apportate ai metadati, di solito, si conclude nell'arco di una settimana dalla modifica.

Tutte queste caratteristiche, applicate al contenuto di un progetto di ricerca e dei risultati ottenuti, consentono di facilitare la diffusione e la conoscenza di quanto realizzato indicizzando le risorse negli aggregatori e motori di ricerca sia specialistici che generalisti, e di garantire identificabilità, responsabilità e persistenza. Sarà quindi più semplice per gli utenti e i ricercatori trovare e citare correttamente il progetto di cui si è registrato il DOI, utilizzarlo nel rispetto delle licenze di uso assegnate e

attribuirne la corretta responsabilità. Inoltre, sempre più spesso i finanziatori della ricerca chiedono di essere inseriti nei crediti, per documentare e tracciare le fonti di finanziamento e favorirne la rendicontazione; questo è possibile, in automatico, se al momento della registrazione del DOI sono state correttamente compilate le apposite sezioni dei metadati amministrativi relativi ai finanziamenti.

Per quanto riguarda gli svantaggi e i punti critici, occorre tenere presente che utilizzare i DOI ha comunque una barriera di ingresso, costituita sia da aspetti finanziari (per essere assegnatari di un prefisso DOI è richiesto un investimento minimo di circa 400-500 euro/anno) ma soprattutto tecnici e scientifici (per la individuazione e la gestione di metadati di qualità).

Inoltre, per garantire una preservazione a lungo termine della risorsa, occorre adottare una politica di archiviazione digitale adeguata; senza di questa, ancora una volta l'operazione di assegnazione del DOI può rappresentare solamente una operazione di facciata.

Anche utilizzando i *repository* istituzionali e tematici, che offrono contestualmente la registrazione del DOI e l'archiviazione della risorsa, occorre considerare il costo di realizzazione di metadati di qualità; registrare infatti DOI con pochi metadati, oppure addirittura con informazioni errate o parziali, rappresenta una scelta poco lungimirante, difficilmente giustificabile. Infine, occorre valutare attentamente la soluzione scelta, in quanto ciascun *repository* è gestito da un soggetto (pubblico o privato) di cui occorre valutare la stabilità e affidabilità nel tempo; affidarsi interamente ad un *repository* riduce certamente i costi complessivi ma fa perdere la propria autonomia, delegandola ad un soggetto esterno.

Alcuni possibili casi di utilizzo

L'utilizzo dei DOI per identificare e rendere citabile una edizione digitale è il più noto ed immediato caso possibile, ma certamente non rappresenta una novità, perché come notato (Carlquist 2004) una edizione digitale dovrebbe includere testo ricercabile, immagini, metadati significativi, apparati critici e indici, e consentire ricerche, scaricamento di dati in formato XML così come la stampa o la pubblicazione in altri formati elettronici. In questa sede interessa proporre possibili utilizzi del DOI nei casi delle banche dati sviluppate in progetti di *digital humanities*.

Un primo esempio di applicazione è rappresentato dal progetto DaLiB²⁰ che si occupa della produzione, fruizione e circolazione libraria nel Mediterraneo antico: Grecia, Egitto, Roma. Il progetto si articola nella costruzione di un reticolo di relazioni tra luoghi (fisici o virtuali); ciascun luogo (di rilevanza libraria) sarà caratterizzato da una scheda approfondita con informazioni storiche, geografiche, archeologiche, bibliografiche e arricchito da materiale multimediale, come immagini, fotografie, trascrizioni di testi. Alla costruzione di ciascuna scheda, per la natura

interdisciplinare del progetto, collaborano studiosi e ricercatori di diverse aree scientifiche, ciascuno per la propria competenza. Infine, il numero di schede prodotte sarà dell'ordine delle centinaia; ci sono dunque le condizioni perfette per l'attribuzione di singoli DOI a ciascuna scheda, valorizzando così la presenza in rete e la diffusione di ciascun "luogo".

Un secondo esempio è costituito dal progetto OIM²¹ (Osservatorio degli Italianismi del mondo) promosso dall'Accademia della Crusca con il coinvolgimento di oltre 15 gruppi di ricerca italiani e internazionali. Il progetto mira a raccogliere e studiare le forme italiane diffuse nelle altre lingue nel mondo; per ciascuna forma viene quindi realizzata una scheda lessicografica italiana, e una serie di schede collegate, una per ciascuna lingua di rilevamento, con i relativi significati, fonti e caratterizzazioni linguistiche.

Anche in questo caso siamo in presenza di un sistema di informazione reticolare, collaborativa, e in rapida evoluzione, dato che la raccolta e il numero delle lingue oggetto di studio è in fase di incremento. Tutte queste caratteristiche favoriscono quindi l'adozione dei DOI; attraverso i metadati di registrazione, è possibile proprio mantenere aggiornato lo stato di conoscenza e avanzamento di studio di ciascuna forma, e diffonderne la presenza in rete. In questo caso il numero di schede e quindi di DOI da registrare è dell'ordine di diverse migliaia; quindi, l'operazione va ponderata e predisposta con una certa cura. Certamente è consigliabile procedere inizialmente con la costituzione di alcune collezioni (o raccolte) di forme, assegnando quindi un DOI ad un insieme di schede, ma a regime è evidente che la registrazione sarà effettuata a livello di singola forma, proprio per sfruttare a pieno i vantaggi offerti dal DOI.

Conclusioni

Il DOI è un sistema di identificazione persistente degli oggetti digitali, fisici o logici, molto diffuso nell'editoria accademica e nella archiviazione di *dataset* di esperimenti scientifici, ma scarsamente utilizzato sinora nei progetti di *digital humanities*.

I benefici nell'utilizzo consistono nella diffusione e nella tracciabilità di responsabilità e processi; nel caso di progetti di banche dati testuali quali ad esempio un dizionario o un glossario, calibrando adeguatamente le risorse è possibile stabilire il livello di dettaglio nei metadati e di granularità nel numero di risorse digitali a cui associare e registrare il DOI, per ottenere un giusto compromesso tra investimento e risultati ottenuti.

In sintesi, registrare DOI utilizzando un ampio set di metadati, correttamente compilati, consente di rendere identificabile e persistente (nella citabilità e nell'accesso se viene associata una corretta politica di archiviazione a lungo termine) una risorsa digitale realizzata nel corso di un progetto di ricerca di *digital humanities*;

la registrazione del DOI consentirà il corretto utilizzo della risorsa stessa, nel rispetto delle politiche di accesso e di tutela della responsabilità scientifica.

Note

1. Ad esempio, l'iniziativa CLOCKSS (*Controlled Lots of Copies Keep Stuff Safe*) consiste di una collaborazione fra editori accademici pubblici e privati che offre un servizio di preservazione a lungo termine per le riviste accademiche; cfr: <https://clockss.org/>
2. Sul sito web del consorzio DOI è possibile reperire la guida tecnica operativa dove sono indicate le regole e le buone pratiche di assegnazione dei codici DOI https://www.doi.org/the-identifier/resources/handbook/2_numbering
3. mEDRA è una agenzia di registrazione italiana nata nel 2003 come iniziativa della Associazione Italiana Editori che ancora la controlla tramite la società di servizi EDISER srl. Cfr: <https://www.medra.org/>
4. CROSSREF è la principale agenzia del DOI (per numero di risoluzioni mensili) gestita da una Associazione di editori accademici pubblici e privati con base negli USA e diffusa in tutto il mondo, cfr. <https://www.crossref.org/>
5. Datacite è la più importante agenzia DOI a carattere no-profit, gestita da un ampio numero di enti e istituzioni della ricerca scientifica in tutto il mondo, cfr. <https://datacite.org/>
6. L'elenco delle tipologie di oggetti, a seconda delle agenzie, è nella documentazione ufficiale sul sito del DOI: <https://www.doi.org/the-identifier/resources/handbook/> e delle varie agenzie di registrazione.
7. Si veda il dettaglio delle quote associative sul sito di Datacite: <https://datacite.org/feemodel.html>
8. Si veda il dettaglio delle quote associative e dei costi di registrazione sul sito di Crossref: <https://www.crossref.org/fees/>
9. Si veda il dettaglio delle quote previste sul sito di mEDRA: <https://www.medra.org/it/terms.htm>
10. API (*Application Programming Interface*). Set di procedure software disponibili in un sistema informatico per dialogare con l'esterno; consentono generalmente di interrogare e manipolare i dati.
11. Si veda il sito web ufficiale: <https://zenodo.org/>
12. Si veda il sito web: <https://figshare.com/>
13. <https://unite.ut.ee/>
14. <https://search.datacite.org/works/10.15156/bio/th000120>
15. Si veda il sito del progetto: <https://www.vliz.be/en/data-projects>
16. <https://doi.org/10.5281/zenodo.4554112>
17. <https://doi.org/10.1285/i9788883051623p1>
18. <https://doi.org/10.1285/i22390359v51p69>

19. <https://www.crossref.org/blog/measuring-metadata-impacts-books-discoverability-in-google-scholar/>
20. <https://dalib.unifi.it>
21. <https://www.edizionidicrusca.it/scheda.asp?IDV=575>

Bibliografia

- Attanasio 2022 = Paolo Attanasio, *New Challenges in Metadata Management between Publishers and Libraries*, J LIS.It, 13(1), 2022, pp. 116–122; DOI: <https://doi.org/10.4403/jlis.it-12777>
- Barbuti 2019 = Nicola Barbuti, *Ripensare i formati, ripensare i metadati: prove “tecniche” di conservazione digitale*, Umanistica Digitale, 3(5) 2019, p. 123; DOI: <https://doi.org/10.6092/issn.2532-8816/9055>
- Carlquist 2004. Jonas Carlquist, *Medieval Manuscripts, Hypertext and Reading. Visions of Digital Editions*, Literary and Linguistic Computing, 19(1), April 2004, pp. 105–118; DOI: <https://doi.org/10.1093/lc/19.1.105>
- Klein e Balakireva 2020 = Martin Klein e Lyudmila Balakireva, *On the Persistence of Persistent Identifiers of the Scholarly Web* in M. Hall et al. (Eds.): TPD L 2020, LNCS 12246, pp. 102–115; DOI: https://doi.org/10.1007/978-3-030-54956-5_8
- Masson et al. 2021 = Arnaud Masson et al., *Google dataset search and DOI for data in the ESA space science archives*, Advances in Space Research, 67(8), 2021, pp. 2504–2516; DOI: <https://doi.org/10.1016/j.asr.2021.01.035>
- Meadows et al. 2019 = Alice Meadows et al., *Persistent identifiers: the building blocks of the research information infrastructure*, Insights, 32(9), 2019, pp.1–6; DOI: <https://doi.org/10.1629/uksg.457>
- Paskin 2005 = Norman Paskin, *Digital Object identifiers for scientific data*, Data Science Journal, 4-2005, pp. 12-20; DOI: <https://doi.org/10.2481/dsj.4.12>
- Rueda et al. 2016 = Laura Rueda et al., *DataCite: Lessons Learned on Persistent Identifiers for Research Data*, International Journal of Digital Curation, 11(2), 2016, pp. 39–47. DOI: <https://doi.org/10.2218/ijdc.v11i2.421>
- Wilkinson et al. 2016 = Mark D. Wilkinson et al., *The FAIR Guiding Principles for scientific data management and stewardship*, Sci Data 3, 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>