



Rivista del digitale nei beni culturali

ICCU-ROMA

Implementing Persistent Identifiers: Overview of concepts, guidelines and recommendations / Hans-Werner Hilse and Jochen Kothe

Mario Sebastiani

ICCU

Hans-Werner Hilse and Jochen Kothe sono due ricercatori della Biblioteca Statale e Universitaria di Gottinga, esperti in programmazione di computer e gestione di risorse digitali, ai quali, nel marzo del 2005, il CERL – Consortium of European Research Libraries (Consorzio delle biblioteche europee di ricerca) ha commissionato un rapporto che fornisce una panoramica degli “identificatori persistenti” o, per meglio dire, dei “sistemi di identificazione persistente” di risorse digitali, sviluppati dalla metà degli anni ‘90 in qua. Il CERL, come è noto, produce lo Hand Press Book Database (HPB), catalogo collettivo della stampa manuale in Europa; l’ICCU, Istituto Centrale per il Catalogo Unico delle biblioteche italiane, è membro del CERL insieme ad altre 43 tra le più importanti biblioteche ed istituzioni europee. Proprio l’evidenza – maturata nel corso dell’attività di aggiornamento e manutenzione di questo database – che giorno dopo giorno aumentano i link “rotti” o “spezzati” (“broken” in inglese) contenuti nei record del database, ha indotto il CERL a commissionare questo rapporto, come base da cui partire per individuare un sistema adeguato di identificazione persistente. Al CERL, nel supportare il rapporto, si è associata la European Commission on Preservation and Access (ECPA), un organismo costituito nel 1994 presso l’Accademia Reale Olandese di Arti e Scienze, per promuovere varie attività nell’ambito della preservation e dell’accesso.

Come è noto, i comuni indirizzi web – quelli come <http://digitalia.sbn.it/> tanto per capirci – sono intrinsecamente instabili. Può sempre accadere infatti che la risorsa alla quale puntano sia diventata repentinamente inaccessibile per una grande varietà di motivi. Ad esempio la risorsa è stata spostata su un altro server. Oppure la risorsa è sempre presente sullo stesso server ma per motivi tecnici – ad es. la sostituzione del software gestionale – le directory del server e quindi i riferimenti locali della risorsa, sono state variamente riorganizzate e/o ridenominate. O, ancora, può essere accaduto che, per motivi amministrativi, commerciali, giuridici o altro, al server sia stato assegnato un nuovo nome di dominio, cioè sia stata modificata la parte iniziale dell’indirizzo web assegnata secondo le regole internazionali del sistema a nomi di dominio (DNS - Domain Name System). In tutti questi casi l’indirizzo web originario della risorsa non funzionerà più; cliccando sui link con questo indirizzo, si otterrà come risposta una pagina vuota con la dichiarazione “404 – Not found” (nel protocollo http, hypertext transfer protocol, il codice 404 segnala appunto l’impossibilità di recuperare una data risorsa).

Tutti i comuni link web sono esposti alla possibilità di diventare, prima o poi, dei link “rotti”, cioè riferimenti ipertestuali che indirizzano l’utente verso il nulla. Questo perché tanto lo standard DNS che lo standard URL – Uniform Resource Locator, localizzatore uni-

forme di risorse, gli standard alla base dei comuni indirizzi web, non distinguono la funzione di localizzazione di una risorsa elettronica da quella di denominazione. Per i comuni indirizzi web, denominare una risorsa equivale a localizzarla e viceversa. Chiaramente l'instabilità degli indirizzi URL costituisce l'aspetto più problematico del World Wide Web. La "rottura" dei link rende vuote le citazioni a risorse web che possono essere contenute in documenti cartacei, elettronici, record di database, siti web ecc. È per questo che dalla metà degli anni '90 è cominciato lo studio e la sperimentazione di nuovi standard in grado di distinguere tra funzione di localizzazione e funzione di denominazione delle risorse elettroniche *on-line*. Questi standard sono meglio noti col nome di "identificatori persistenti" e col tempo, per via della crescita sempre più ampia delle iniziative di diffusione via web delle risorse elettroniche, l'esigenza di nuovi strumenti gestionali con queste caratteristiche si è fatta sempre più pressante.

Secondo gli autori, al momento attuale è impossibile indicare come preferibile un sistema piuttosto che un altro. Semmai è importante che le istituzioni responsabilizzino loro stesse e i loro utenti relativamente al fatto che «la digital preservation non ha solo a che fare con la migrazione verso nuovi supporti e nuovi formati e con la conservazione delle funzionalità, ma che il punto di partenza sta in un requisito fondamentale, vale a dire che i documenti possano essere identificati e localizzati senza ambiguità da coloro che ne hanno bisogno» (p. 1). Però occorre anche evitare di coltivare l'illusione che la tecnica possa risolvere da sola il problema dell'identificazione persistente; la gestione delle risorse elettroniche deve necessariamente trovare un riflesso nell'organizzazione amministrativa dell'ente di gestione perché «al nocciolo, la persistenza è un compito amministrativo che non può essere sostituito dalla tecnologia» (p. 45).

Il rapporto descrive questi sistemi:

- URN – Uniform Resource Name;
- NBN – National Bibliographic Number;
- Handle system;
- DOI – Digital Object Identifier;
- ARK – Archival Resource Key;
- PURL – Persistent URL;
- OpenURL.

Il rapporto è poi completato da una consistente bibliografia. Il testo, in formato PDF, può essere scaricato dai siti CERL (<http://www.cerl.org/>) e ECPA (<http://www.knaw.nl/ecpa/>) ed è fruibile nei limiti di una licenza Creative Commons indicata all'interno del documento. Su Digitalia web (<http://digitalia.sbn.it/>) è presente un link al documento nella sezione di documentazione.

Implementing Persistent Identifiers : Overview of concepts, guidelines and recommendations / Hans-Werner Hilse and Jochen Kothe. - London : Consortium of European research Libraries; Amsterdam: European Commission on Preservation and Access, 2006. - 57 p. - ISBN 90-6984-508-3.