

Conservazione delle risorse elettroniche per garantire il pubblico accesso

di

Olivia M.a. Madison
*Direttore dei Servizi Bibliotecari
dell'Università Statale dello Iowa*

Nel ventesimo secolo, le persone, quando hanno bisogno di informazioni, dovrebbero pensare che in primo luogo possono trovarle in biblioteca. Le biblioteche sono partners creativi nel processo di apprendimento e di ricerca insieme con le nostre università, comunità e paesi. I bibliotecari selezionano, organizzano, presentano e conservano le risorse informative per gli studiosi delle generazioni presenti e future. In tal modo, essi diventano gli intermediari attivi delle feconde risorse di conoscenza; precorrono le esigenze dell'utente e rispettano una differente comunità di idee.

L'accesso elettronico a raccolte librerie è ora uno standard e un segno di riferimento sempre più importante nella valutazione di collezioni librerie. Oggi la maggior parte delle singole biblioteche è responsabile della costruzione e del mantenimento di due biblioteche - una è reale e l'altra è virtuale - nelle quali si realizza la fusione tra il tradizionale e l'elettronico. La stessa attenta pianificazione e le risorse che disponiamo nelle nostre strutture e collezioni reali sono necessarie a disegnare, costruire e mantenere le nostre strutture elettroniche. Il cemento e i mattoni virtuali di queste biblioteche elettroniche sono costituiti dal contenuto scientifico e dalle infrastrutture tecnologiche. Forse George Soete sancisce l'operazione che noi compiamo nel mantenere queste nuove biblioteche molto succintamente quando scrive, "quando noi ci confrontiamo con il problema della conservazione dell'informazione digitale, ci confrontiamo con l'essenza stessa del significato di una biblioteca nel ventesimo secolo".

Le nostre collezioni elettroniche si sono evolute rapidamente durante gli ultimi tre decenni dalla mediazione per ricerche condotte dai bibliotecari sui servizi di indicizzazione a distanza all'accesso al singolo web ad un ricco assortimento di risorse full-text alloggiate su complicati network commerciali e server locali. In particolar modo con l'evoluzione delle collezioni digitali in full-text, i cambiamenti nella conservazione di questi fragili materiali sono stati enormi, dovuti sia alla complessità dei diritti di proprietà che a quella tecnologica. Paradossalmente, nelle nostre biblioteche ciò che si sta potenziando fisicamente sono proprio i materiali elettronici con CD-ROM di risorse specializzate, di base-dati locali, di materiali locali unici e materiali digitalizzati localmente che rimangono fuori dalle restrizioni del copyright. Dipendiamo sempre più dai fornitori commerciali di contenuto, che si spera saranno depositari affidabili di risorse scientifiche e di strumenti di consultazione - depositari, i quali dovrebbero fornire un accesso durevole e la necessaria conservazione di quelle raccolte che noi abbiamo acquisito tramite accordi contrattuali.

In questo paper parlerò brevemente della situazione attuale relativamente alle questioni di accesso e di conservazione di risorse elettroniche e proseguirò con degli esempi di standards in fase di sviluppo, esempi di alcuni modelli di sviluppo e di conservazione esistenti, ed esempi di progetti recenti d'archiviazione istituzionale. Io spero di offrirvi attraverso questi esempi un quadro della complessità dei cambiamenti per quanto riguarda la conservazione delle risorse elettroniche e di sottolineare che non esistono degli approcci semplici e diretti. Piuttosto dobbiamo trarre vantaggio da un insieme complesso di circostanze, da una estesa varietà di partners, e dalla rapida crescita tecnologica e dei nuovi standards.

Accesso e conservazione

Fornire accesso e conservare materiali librari sono stati per lungo tempo i compiti essenziali e solitari delle singole biblioteche. Mentre ci sono stati progetti e iniziative di collaborazione e cooperazione, queste funzioni sono state portate avanti a lungo nell'isolamento istituzionale. Durante l'ultimo decennio del ventesimo secolo queste tradizionali incombenze sono cambiate clamorosamente. Le biblioteche non mantengono più fisicamente alcune delle loro risorse più interessanti e più adoperate, mentre acquistano l'accesso a materiali spesso attraverso licenze complicate e associandosi attraverso complesse forme consortili d'acquisto. L'importanza di un depositario di fiducia (fornitore di contenuti o singola biblioteca) acquista un nuovo significato in questo nuovo contesto digitale. Noi acquistiamo generosamente accessi a records di pubblicazioni scientifiche, e ci aspettiamo che i rispettivi fornitori di contenuto (e spesso anche depositari) assicurino un accesso duraturo a quel record. Esistono ora dei documenti legali che delineano i nostri rispettivi ruoli in questi nuovi accordi giuridici. Di conseguenza, non siamo più sicuri di quello che possediamo. La mancanza di sistemi tecnologici affidabili e commerciali che preservino i record con le risorse elettroniche scientifiche ha condotto entrambi, sia il fornitore di contenuti che l'acquirente, a pubblicare e a mantenere la versione a stampa delle risorse elettroniche come copie di conservazione. Molte biblioteche, infatti, stanno iniziando ad introdurre dei titoli elettronici tra i propri acquisti e a trattare le copie stampate corrispondenti (che rappresentano un ulteriore costo d'acquisto) come copie di "conservazione". Spesso in questi accordi legali, quando vi è una discontinuità di accesso ad un singolo titolo (sia per la cessazione di un abbonamento sia per la chiusura di un magazzino) si ricorre ad un CD-Rom. Se però da un lato, esso riproduce l'argomento, non provvede a fornire l'accesso remoto ai potenziali link attivi ad altre risorse elettroniche attuali sul Web.

Senza riguardo al tipo di deposito digitale andiamo prefigurando (sia esso una biblioteca universitaria, un archivio, un'agenzia nazionale bibliografica, una associazione non profit, una società scientifica o una società commerciale), noi affrontiamo tutti le medesime sfide per quanto riguarda la conservazione, che sono:

Il formato fragile - i materiali digitali sono particolarmente vulnerabili perché sono memorizzati su fragili supporti magnetici e ottici che si deteriorano rapidamente e possono venire meno improvvisamente.

Obsolescenza della tecnologia - i materiali digitali diventano illeggibili ed inaccessibili se i dispositivi di playback necessari per richiamare le informazioni dai media diventano obsoleti o se il software che traduce le informazioni digitali non è più disponibile.

Le questioni legali che persistono intorno alla copiatura e all'accesso - biblioteche, archivi ed altre istituzioni culturali, hanno limitato e reso incerti i diritti per copiare le informazioni digitali a scopo di recupero o di conservazione, per convertire l'informazione in modo tale che essa rimanga accessibile alla tecnologia corrente, e per fornire un accesso pubblico¹.

Un problema scottante che si pone a tutti noi nel nostro ambiente, da un lato decentralizzato, dall'altro soggetto all'accentramento dei costi è quello se i depositi (privati, pubblici o commerciali) siano disposti o meno a sopportare i costi continui della rigenerazione dei dati e dell'aggiornamento della struttura degli archivi elettronici, del software e dell'hardware secondo gli standard correnti. Dal momento che le singole biblioteche potrebbero decidere di ritirare i libri più vecchi poco usati e i periodici, gli editori possono decidere di fare altrettanto con i loro archivi elettronici, con i periodici obsoleti e le monografie superate dal ridotto utilizzo abbinato e dai costi tecnologici.

Le biblioteche attualmente stanno facendo affari con un ambiente commerciale globale florido con piani gestionali adeguati ai profitti previsti per gli azionisti; fusioni e cambi di ge-

¹ *LC21 : A Digital Strategy for the Library of Congress.* (Washington, D.C.: National Academy Press, 2000), 106-107.

stione sfavorevoli con conseguente pratiche monopolistiche e prezzi e pratiche di accesso non competitive; case editrici che vanno in fallimento con potenziali perdite di accesso continuo alle loro holding elettroniche; numerose aziende imprenditoriali dot.com che si sono avventurate nel settore commerciale dei fornitori di informazioni con infrastrutture a rischio capitale; e imprese non profit che iniziano con contributi di sostegno, ma non sono economicamente scalabili appena il sostegno diventa instabile. Cosa significa per il ruolo delle biblioteche questo continuare a fornire l'accesso a risorse elettroniche che possono o non possono essere duplicate da copie stampate per la conservazione?

Inoltre, anche un accesso affidabile come è tradizionalmente quello delle informazioni governative, può cambiare drammaticamente nell'arco di una notte - come si è visto per gli Stati Uniti dall'11 settembre, 2001. I vari enti governativi e dipartimenti hanno iniziato a rimuovere i dati sensibili o le informazioni classificate dai siti Internet governativi. In alcuni casi gli enti federali e governativi stanno disponendo i loro siti dietro a dei *firewalls* ad accesso ristretto o stanno sospendendo l'attività di interi siti. Per esempio, la Nuclear Regulatory Commission degli Stati Uniti ha interrotto il proprio sito Web, che includeva le informazioni circa i progetti di centrali nucleari. Il Department of Transportation's Office of Pipeline Safety degli Stati Uniti ha recentemente limitato l'accesso al suo National Pipeline Mapping System, che fornisce l'ubicazione delle condutture di gas naturale e quelle dove si potrebbe mettere a rischio l'acqua potabile. Anche Google, uno dei maggiori motori di ricerca Internet, volontariamente ha rimosso copie dei siti che il suo staff ha giudicato tali da poter costituire una minaccia per la sicurezza.

Standard per conservazione del digitale

Con la crescita della dipendenza culturale e pratica dai mezzi elettronici, la quale consente ad una grande varietà di utenti delle nostre comunità un accesso ad ampio spettro, nasce finalmente un momento di autonomia per creare gli standard e le politiche necessari che vadano incontro a queste esigenze. Inoltre, la digitalizzazione offre un nuovo formato per la conservazione e la riconversione in formato elettronico di libri, manoscritti, e altri materiali per fornire l'accesso in ordine alla salvaguardia dei materiali originali o per provvedere alla loro totale sostituzione dovuta alla perdita dei prodotti originali. Dovrebbero risultarne standards e tecnologie che potrebbero fornire un largo assortimento di strumenti atti a soddisfare in modo appropriato le necessità di accesso e quelle di conservazione.

Attualmente sono in corso numerosi progetti pilota ed iniziative locali, regionali, nazionali ed internazionali. Le descrizioni o gli aggiornamenti sono forniti per tre standard che sono destinati a dimostrare la molteplicità di standards necessari a far fronte alle complessità della conservazione digitale.

Attributi d'un deposito digitale affidabile e OAIS

Nel marzo 2000, le Research Libraries Group (RLG) ed OCLC hanno iniziato ad analizzare e a proporre le caratteristiche di un deposito digitale per le organizzazioni di ricerca. Ciò potrebbe apparire come un piano di certificazione per gli archivi digitali o i depositi - la cosa risponde ad un precedente rapporto RLG del 1996 sull'archiviazione delle informazione digitale. Il risultato dello studio è una prima bozza di rapporto RLG-OCLC che ipotizza le caratteristiche di un deposito digitale sicuro (come una biblioteca², una biblioteca nazionale, o un archivio) e sollecita un commento da parte del pubblico. Il rapporto definisce le due funzioni essenziali della conservazione a lungo termine: il mantenimento a lungo termine d'un bytestream e l'accesso ai propri contenuti che tengano conto del tempo e dei cambiamenti

² John Garrett and Donald Waters, *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information* (Mountain View, CA: Commission on Preservation and Access and RLG, 1996). Disponibile alla url: <http://www.rlg.org/ArchTF/index.html>

tecnologici³. Il rapporto presuppone che per assicurare la conservazione a lungo termine di una risorsa digitale e scientifica siano necessarie delle infrastrutture serie e decentralizzate e una comunità scientifica che esiga un clima generale di affidabilità dei sistemi, dei protocolli, dei principi e delle organizzazioni che supportano tali infrastrutture.

Il rapporto propone la seguente definizione:

Un deposito digitale affidabile è quello il cui compito è di fornire l'accesso a lungo termine alle risorse digitali amministrare; che accetti la responsabilità del mantenimento a lungo termine delle risorse digitali a nome dei relativi depositanti a beneficio degli utenti attuali e futuri; che disegni il proprio sistema(i) conformemente a convenzioni e a standards comunemente accettati per assicurare lo sviluppo gestionale, accesso e sicurezza dei materiali depositati all'interno di esso; che stabilisca metodologie per valutare un sistema che incontri le aspettative di affidabilità della collettività; che sia subordinato apertamente ed esplicitamente alle garanzie a lungo termine richieste dai committenti e dagli utenti; e le cui politiche, pratiche e prestazioni possano essere verificate e misurate⁴.

In quella occasione il rapporto studiò i potenziali processi di certificazione per archivi digitali che avrebbero creato un clima generale di fiducia sulle possibilità dei depositi di conservare l'informazione digitale. Tale studio individua due modelli: il modello di verifica ed il modello standard. Il modello di verifica è usato frequentemente all'interno dei sistemi di deposito, spesso coinvolgendo agenzie governative o enti legislativi che creano linee guida per biblioteche singole e nazionali che ricevono e forniscono accesso alle proprie risorse. Il modello standard, d'altra parte, è quello di cui la maggior parte delle biblioteche fa uso consueto. Le biblioteche seguono in generale gli standards accreditati, come prestito interbibliotecario ISO, standards di catalogazione nazionali ed internazionali, le linee guida per la creazione di metadati, etc. Istituzioni o agenzie parimenti "certificano" il prodotto o il servizio prendendolo e/o usandolo⁵.

Il rapporto richiede che un'infrastruttura seria sia capace di supportare un sistema archivistico digitale distribuito e suggerisce di prestare considerazione all'Open Archival Information Systems (OAIS) Reference Model, che è stato sviluppato dal Consultive Committee for Space Data Systems. Questo modello può essere applicato sia ai depositi tradizionali che agli altri ed è stato utilizzato con competenza tecnica attraverso un ampio arco disciplinare. L'OAIS Reference Model serve ad unificare i gruppi di concetti per un archivio OAIS. Richiede un team organizzativo di persone e un sistema tecnologico che si siano fatti carico di conservare informazione e di renderla disponibile alle proprie comunità.

OAIS è stato usato con successo all'interno del progetto di NEDLIB per sviluppare il relativo sistema di deposito delle pubblicazioni elettroniche, del British Library's digital Storage Project, e nel lavoro di sviluppo della Library of Congress, dei National Archives, dell'Harvard University, della Stanford University, di RLG e OCLC. La National Library of Australia lo ha usato come un modello generico per rafforzare le funzioni ed le relazioni nel proprio archivio PANDORA, di cui si parlerà più avanti. Il suo successo relativamente alle proprie funzioni è di modello di riferimento e non di modello per la progettazione di un sistema.⁶

Il rapporto delinea le caratteristiche di un deposito digitale affidabile (basato essenzialmente sul modello OAIS), che comprende:

³ *Attributes of a Trusted Digital Repository: Meeting the Needs of Research Resources – AN RLG-OCLC Report.* Draft for Public Comment. (Mountain View, CA: RLG, August 2001), I.

⁴ *Ibidem*, 12.

⁵ *Ibidem*, 14-15.

⁶ *Ibidem*, 23-24.

Negozia per accogliere le informazioni adatte dai produttori di informazione e dai proprietari di diritti

Ottiene un sufficiente controllo dell'informazione provvedendo alla conservazione a lungo termine del supporto

Segue politiche e procedure documentate che assicurino che l'informazione sia conservata contro tutte le ragionevoli eventualità e che permetta di diffondere le informazioni come copie autentiche dell'originale

Fa sì che le informazioni protette siano disponibili per le comunità designate

Istruzione di ICOLC della prospettiva corrente e delle pratiche preferite

L'International Coalition of Library Consortia (ICOLC) è un'organizzazione ufficiosa che ha iniziato a riunirsi nel 1997. Include sessanta consorzi bibliotecari negli Stati Uniti, nel Regno Unito, nei Paesi Bassi, nel Canada, in Germania, in Israele ed in Australia quindi ne sono membri 5.000 biblioteche di tutto il mondo. La coalizione serve soprattutto alle istituzioni d'istruzione superiore facilitando la discussione fra i propri membri su questioni di interesse comune. ICOLC organizza delle riunioni per tenere i propri membri informati circa le nuove risorse informative elettroniche e le politiche dei prezzi dei fornitori e dei providers elettronici; fornisce occasioni di incontro tra i fornitori d'informazione e le varie comunità per discutere i loro prodotti; e per coinvolgerli in un dialogo con i membri della associazione intorno a questioni di reciproca preoccupazione.

Nel mese di marzo 1998, l'ICOLC ha pubblicato il proprio "Statement of Current Perspective and Preferred Practices for the Selection and Purchase of Electronic Information".⁷ Il comunicato stabilisce per la prima volta una prospettiva internazionale sulle licenze consortili e sull'acquisto da parte delle biblioteche di informazioni elettroniche. Il documento ripercorre le questioni attuali e future dell'ambiente dell'informazione elettronica quali le crescenti aspettative degli utenti delle biblioteche a costituire in un ambiente stabile, un giusto uso, per l'archiviazione delle informazioni, strategie economiche, e valutazioni quantitative dell'information delivery elettronico. La sezione principale riguarda negoziazioni contrattuali, quotazioni, inserimento dei dati e l'archiviazione, piattaforme di sistema, termini contrattuali, contenuto informativo e sua gestione, e certificazione dell'utente.

In termini di archiviazione, il comunicato richiede che i venditori forniscano "una licenza perpetua quando il consorzio acquista il contenuto. I consorzi e le loro biblioteche membro dovrebbero poter attingere entro misure ragionevoli al contenuto dell'archivio che hanno acquistato o preso a leasing (per fare le copie di riserva)".⁸ Ancora, "quando un fornitore di informazioni dà accesso ai dati del proprio sito Web (piuttosto che attraverso il montaggio locale dei dati), il fornitore dovrebbe garantire la disponibilità continua del contenuto. Questa disponibilità non deve obbligare il fornitore a un accesso in tempo reale. Per esempio, può essere possibile rifornire il consorzio con copie di files di dati in un formato appropriato, depositare presso terzi files di dati, o altri mezzi appropriati."⁹

Metadati Dublin Core

La Dublin Core Metadata Initiative è ha preso il via nel 1995 con un seminario ad inviti patrocinato da OCLC a Dublino, nell'Ohio. Allo stato attuale si è iniziato lo sviluppo di un gruppo fondamentale standard di descrittori bibliografici/di accesso di files elettronici, che possono essere inseriti specificatamente in qualsiasi file dato, e da questo possono essere

⁷ International Coalition of Library Consortia (ICOLC), *Statement of Current Perspectives and Preferred Practices for the Selection and Purchase of Electronic Information*. Disponibile alla url: <http://www.library.yale.edu/consortia/statement.html>

⁸ *Ibidem*.

⁹ *Ibidem*.

cercati per identificazione e a scopo di potenziale recupero. Letteralmente si tratta di dati reperibili e descrittivi che descrivono dati che potrebbero essere inclusi in un file o esterni ad esso. I dati possono essere creati dall'"autore" del file elettronico o da un creatore esterno. I metadati spesso sono visti come un altro elemento distintivo per catalogare un record tradizionale e fornire un meccanismo che consenta di ricercare direttamente attraverso i complessi files di dati testuali. Gli elementi Dublin Core sono titolo, soggetto, descrizione, fonti, lingua, relazioni, copertura, creatore, editore, autore di contributo, diritti, data, tipo, formato, e identificatore. Questi elementi sono stati fissati molto attentamente in rapporto agli elementi previsti dai requisiti funzionali dei record bibliografici dell'IFLA. Questo particolare schema di metadati che porto alla vostra attenzione è quello dell'inizio di ottobre 2001 NISO (National Information Standards Organization) che annunciava che il gruppo di elementi Dublin Core era stato approvato come uno standard ANSI (Z39.85-2001) - dandogli quindi lo status di standard nazionale negli Stati Uniti con implicazioni internazionali.¹⁰

Esempi dello sviluppo di modelli esistenti

Al momento non ci sono singoli standard di conservazione o gruppi di tecnologie che funzionino per tutti i depositi. Infatti, data la vasta gamma di compiti e di collezioni dei nostri depositi, sarebbe stato naïve presupporre un approccio dilettantistico. Piuttosto, come accennato all'inizio, abbiamo bisogno di un largo gruppo di modelli standard che rispondano ai nostri obiettivi di conservazione e di accesso. I due esempi forniti consentono di dimostrare la varietà di modelli necessari e di proporre quelli che stanno sviluppandosi, emergendo e evolvendosi. Un esempio rappresenta un'iniziativa volontaria di collaborazione per conservare l'accesso al contenuto integrale di pubblicazioni scientifiche e l'altro rappresenta un approccio piuttosto unico al contenuto dei providers attraverso l'ottenimento di un uso perpetuo dei materiali elettronici.

Garantire l'accesso continuo attraverso LOCKSS™

LOCKSS rappresenta un emozionante metodo di collaborazione per garantire l'accesso agli e-journal che rispecchia il metodo decentrato tradizionale di responsabilità istituzionale per conservare le singole collezioni. Il nome è un acronimo per "Lots of Copies Keep Stuff Safe™" si tratta di un'iniziativa partita dalla Stanford University degli Stati Uniti, con il supporto della National Science Foundation, della Mellon Foundation, e della Sun Microsystems. L'esperimento Beta è iniziato nell'aprile 2001 con un gruppo mondiale di biblioteche e la partecipazione di editori, e con la produzione di una versione prevista per il 2002.

Il concetto, che sottende al sistema di LOCKSS, è basato su regole semplici: acquistare lotti di copie; divulgarle per il mondo, in modo che alcune siano facili da trovare, e che contestualmente sia difficile trovarle tutte; prestare o copiare le vostre copie quando altre biblioteche ne hanno bisogno e collaborare soltanto con biblioteche competenti e di fiducia. Alle biblioteche partecipanti, inoltre, vengono richieste soltanto una tecnologia e un'esperienza informatica di base.¹¹

Il progetto include quarantacinque *test beta* per le biblioteche dei cinque continenti e quarantacinque editori che hanno appoggiato il *test beta*. Lo stato di avanzamento del progetto è disponibile al link: <http://lockss.standard.edu/projectstatus.htm>. Il sistema è destinato a assicurare l'accesso al contenuto pubblicato disponibile attraverso Internet. Esso si fonda sul concetto tradizionale delle singole biblioteche, attraverso l'acquisto e la conservazione

¹⁰ Marilyn Geller E-mail datata 5 ottobre 2001 (10:00:54-0400) che annuncia l'approvazione del Dublin Core Metadata Element Set approvato dall'ANSI.

¹¹ "LOCKSS: Protecting & Preserving Web Documents." In: Sun Microsystem Laboratories, *Ten Years of Impact*. Disponibile alla url: <http://research.sun.com/features/tenyears/LOCKSS.html>

delle loro collezioni, con una collaborazione tale da assicurare che ci sia più di una copia elettronica per un gruppo di pubblicazioni scientifiche identificate.

L'obiettivo del progetto LOCKSS è quello di permettere alle biblioteche di prendere in custodia il materiale al quale si abbonano - nello stesso modo in cui viene fatto per il cartaceo - e conservarlo permanentemente. In ogni ricerca, gli utenti hanno un primo accesso a copie di giornali elettronici pubblicati. Se questa copia, per una serie di motivi, non è disponibile, l'utente è indirizzato verso una copia collocata nel sito LOCKSS. Attraverso un sofisticato sistema di preferenze, il sistema LOCKSS nasconde permanente le copie di contenuto on-line - avendo abbastanza copie da assicurare un accesso continuo in tutto il mondo. Questo farà in modo da assicurare che i link e le ricerche provenienti da individui autorizzati continuino ad individuare il materiale pubblicato anche se esso non è più disponibile presso l'editore. E quando una copia d'un giornale on-line è persa o danneggiata, il sistema LOCKSS provvederà ad identificarla e a sostituirla.¹²

Mentre LOCKSS non rappresenta un processo tradizionale di conservazione, per scienziati, bibliotecari e editori che sono interessati al fatto che il record scientifico digitale possa scomparire, malgrado tutte le migliori intenzioni e un'attenta gestione delle licenze, esso ha invece la capacità di andare incontro a tutte le possibili aspettative degli utenti - come:

Fornire alle future generazioni di scienziati l'accesso a tutta la letteratura corrente per la ricerca scientifica, l'insegnamento e la formazione.

Assicurare che i bibliotecari di oggi e quelli del futuro abbiano un meccanismo poco costoso e duraturo da poter controllare per fornire alle loro comunità un accesso di lunga durata alla letteratura primaria.

Fornire agli editori di oggi e del futuro un'assicurazione che i valori editoriali e le marche delle loro pubblicazioni siano disponibili soltanto per lettori autorizzati e certificati.¹³

JSTOR

JSTOR rappresenta un modello unico nel campo dell'archiviazione elettronica il cui scopo è di offrire un archivio retrospettivo di pubblicazioni scientifiche edite e, in avvenire, di quelle pubblicazioni edite solo in forma elettronica. La *mission* di JSTOR è di creare un deposito affidabile di numeri arretrati di periodici scientifici editi. I contratti inerenti al sistema di concessione delle licenze sono firmati con gli editori partecipanti, i quali consentono a JSTOR di creare delle versioni digitali delle loro pubblicazioni e di cedere a JSTOR i "diritti perpetui" di queste versioni elettroniche. Se una singola biblioteca sceglie di abolire un titolo di JSTOR, JSTOR potrà sostituirlo con una versione CD-Rom dei materiali acquistati. JSTOR inoltre ha un accordo con il Center for Research Libraries (CRL) in Chicago per il quale il centro manterrà le possibili copie di conservazione delle pubblicazioni a stampa.

In principio JSTOR ha iniziato il suo servizio in modo da controbilanciare i costi legati all'immagazzinamento in memoria con accesso elettronico. Questo diviene un importante vantaggio, ma rappresenta anche un modo per assicurare l'accesso ad Internet a quegli utenti che sono diventati abituali.

Sembra che soprattutto le biblioteche di ricerca acquistino i titoli di JSTOR per avere l'opportunità di distribuire l'accesso ai backfiles di periodici. Ogni anno, dati gli accordi con gli editori, archivi arretrati suppletivi di pubblicazioni implementano questo archivio elettronico. JSTOR attualmente esplora e converte le edizioni cartacee in forma elettronica con un "arco di tempo" determinato che va dai tre - ai cinque anni.

¹² *Ibidem.*

¹³ *Ibidem.*

In termini semplici, una biblioteca continua ad abbonarsi al periodico corrente e acquista l'equivalente elettronico per la parte retrospettiva attraverso i files elettronici di JSTOR. La biblioteca può optare per il mantenimento degli abbonamenti cartacei correnti fino a quando lo desidera e può scegliere di non rilegarli o di mantenere il duplicato di quelle edizioni tramite l'accesso elettronico a JSTOR.

Per esempio, la collezione *Arti & Scienze* di JSTOR comprende centoventisette periodici, i quali rappresentano all'incirca settecentosettanta volumi che richiederebbero circa mille- duecento piedi lineari di scaffali se venissero conservati in forma cartacea. JSTOR quantifica un risparmio di centoventicinquemila dollari per una raccolta completa a scaffale aperto, escludendo i costi annuali in corso.¹⁴ Le biblioteche non sono state veloci nel ritirare le collezioni duplicate a stampa, tuttavia le cose stanno cambiando. Alla Iowa State University, date le misure di sicurezza di conservazione che JSTOR sta assicurando, a causa di problemi di spazio, stiamo cominciando a selezionare i duplicati a stampa. Noi li abbiamo coinvolti nella nostra facoltà decisionale chiedendogli di valutare l'attuabilità delle immagini scannizzate - specialmente per quanto riguarda le formule matematiche ed i grafici - prima di decidere quali fossero le copie a stampa da ritirare per poterle riprodurre.

Progetti di archiviazione recenti

Mentre i principali progetti di biblioteche digitali fatti da biblioteche nazionali stanno iniziando a diventare luoghi comuni, essi frequentemente sono caratterizzati da regole variabili. Ciò si deve essenzialmente alle diverse definizioni nazionali di record nazionale, di risorse tecniche e di risorse umane e di possibili cooperatori. Alle conferenze generali dell'I-FLA, molti relatori hanno descritto i progetti nazionali prescelti. Due progetti sono stati descritti qui sotto e sono stati scelti per la loro differente metodologia. Inoltre, descritto di seguito vi è l'unico progetto di archiviazione privato per l'intera Internet, il quale dimostra come una sola persona possa fare la differenza per tutti noi.

PANDORA: Collezione Nazionale Australiana di pubblicazioni selezionate on-line.

Il progetto di PANDORA rappresenta il metodo identificato da parte di una biblioteca nazionale al fine di conservare i propri record elettronici nazionali. Nel 1996, la National Library ha iniziato a costituire un archivio di pubblicazioni selezionate on-line che erano nate su supporto digitale. Nel delineare questa iniziativa, si è sviluppato un gruppo di linee guida per la selezione e un insieme di principi economici per guidare il progetto e definirne gli obiettivi. Il modello economico è disponibile al sito <http://pandora.nla.gov.au/pandora/bpm.html>. L'unità elettronica del progetto è responsabile della gestione delle pubblicazioni on-line, che in parte include:

Selezione

Relazioni con l'editore/creatore

Controllo di qualità e problem solving dei quesiti

Catalogazione nel Database bibliografico nazionale

La procedura di selezione ha vaste implicazioni giacché ne risulta un impegno a conservare ogni titolo dato per l'uso dei posteri. Le pubblicazioni digitali, per essere selezionate, dovrebbero riguardare l'Australia o gli Australiani o essere scritte da un australiano o avere come argomento di interesse l'Australia. Catalogare richiede delle regole durature per attribuire titoli oltre a politiche e procedure relative ai metadati per l'accesso e la gestione della conservazione futura. Con questo progetto, gli Australiani stanno lavorando in cooperazione con il progetto CEDARS nel Regno Unito ed allo sviluppo di altri standards di metadati. A partire dall'agosto 2000, c'erano seicentocinquanta titoli nell'Archivio PANDORA e approssimativamente vengono selezionati e archiviati circa trentacinque titoli al mese.¹⁵

¹⁴ Kevin M. Guthrie, "Archiving in the Digital Age: There's a Will, But Is There a Way?" Scheduled for publication in *EDUCAUSE Review* in the November/December 2001 issue.

¹⁵ Cliff Law, "PANDORA - Towards a National Collection of Selected Australian Online Publications." Paper presentato alla

Il Royal Swedish Web Archiw3e

Il Royal Swedish Web Archiw3e rappresenta un approccio globale incentrato sull'archiviazione di tutti i siti Web svedesi, a partire dall'agosto 2000, contiene sessantacinquemilioni di articoli. La Royal Library ha scelto questo tipo di approccio per la difficoltà di determinare quale tipo di informazioni verranno considerate importanti dalle generazioni future. Il progetto è attuabile economicamente perché la memoria dell'elaboratore sta diventando meno costosa, ed ora è possibile identificare e raccogliere queste pagine attraverso degli *snapshots* (fotografie istantanee) del Web e tramite un *robot raccoglitore di tecnologia*. In questo modo, una copia completa del Web svedese è memorizzata dopo ogni *snapshot*, il quale impiega un paio di mesi a completarla. Si accede all'archivio con la navigazione e con ricerche a testo libero, e un diritto di precedenza limitato è dato ai metodi bibliotecari più tradizionali come la catalogazione. A partire dall'agosto 2000, purtroppo non è più possibile accedere all'archivio a causa di una possibile infrazione della legge di copyright.¹⁶

Wayback Machine

L'Internet Archive ha appena annunciato l'apertura della propria Wayback Machine, uno strumento Internet libero che permette l'accesso a dieci miliardi di pagine Web catalogate (o oltre cento terabyte di dati) archiviate durante le sue continue esplorazioni su Internet effettuate dal 1996. L'archivio continua a svilupparsi ad un tasso di 10 terabyte al mese. Alla cerimonia inaugurale per la Wayback Machine tenutasi il 24 ottobre 2001, Brewster Kahle, il fondatore del non profit Internet Archive, disse "noi abbiamo creato l'Internet Archive (<http://web.archive.org/>) poiché ritenevamo che fosse essenziale conservare una testimonianza permanente di questo nuovo mezzo così significativo dal punto di vista storico per il pubblico. Fin qui, l'archivio ha catalogato più di dieci miliardi di pagine Web che altrimenti avrebbero potuto andare perdute, dandoci sia una registrazione delle origini e sia l'evoluzione di Internet, come pure gli *snapshots* (le istantanee) della nostra società insieme a quanto è avvenuto intorno al cambiamento di secolo.¹⁷

Dalla sua fondazione come un raccolta permanente di materiali digitali per il pubblico, l'Internet Archive ha memorizzato e registrato il materiale digitale per il pubblico. L'archivio ha collaborato con la Library of Congress e la Smithsonian Institution ed il risultato è la più grande base di dati esistente conosciuta. Brewster Kahle ha fondato l'Archive per sviluppare una biblioteca digitale ed altri prodotti culturali in forma digitale con il proposito di offrirli permanente e in libero accesso ai ricercatori, ai bibliotecari e al grande pubblico. Le raccolte speciali recenti includono un filmato dell'11 settembre ed un catalogo on-line; films archiviati dal 1903 al 1973; e le elezioni del 2000 negli Stati Uniti.¹⁸

Traduzione di Cinzia Fortuzzi

"Traduzione provvisoria; la traduzione uscirà in veste definitiva negli atti cartacei che saranno pubblicati dalla Editrice Bibliografica di Milano"

Conferenza Generale dell'IFLA, nell'agosto 2000. disponibile alla url: <http://www.ifla.org/IV/ifla66/papers/174-157e.htm>

¹⁶ Allan Arvidson, Krister Perssson, and Johan Mannerhiem, "The Kiulturarw3 Project – The Royal Swedish Web Archiw3e – An example of "complete" collection of web pages." Paper presentato alla 66ª Conferenza Generale dell'IFLA, nell'agosto 2000. Disponibile alla url: <http://www.ifla.org/IV/ifla66/papers/154-157e.htm>

¹⁷ Allan Arvidson, Krister Perssson, and Johan Mannerhiem, "The Kiulturarw3 Project – The Royal Swedish Web Archiw3e – An example of "complete" collection of web pages Paper presentato alla 66ª Conferenza Generale dell'IFLA, nell'agosto 2000. Disponibile alla url: <http://www.ifla.org/IV/ifla66/papers/154-157e.htm>

¹⁸ *Ibidem*.