

A LOCKSS: Restituire ai bibliotecari la custodia del contenuto dei periodici

Nel corso dell'ultimo decennio, le biblioteche hanno sempre più modificato l'accesso ai periodici dalla carta stampata al digitale. La preferenza degli utenti per il contenuto online, la domanda dei lettori per contenuti più estesi, e l'aumento dei costi degli spazi di alloggiamento nelle biblioteche hanno contribuito a questo cambiamento. Tuttavia, gli attuali modelli di distribuzione degli editori obbligano gli utenti ad accedere ai contenuti conservati su un server web centralizzato e mantenuto dall'editore. L'approccio LOCKSS (per "molte copie rendono il materiale sicuro", o Lots of Copies Keep Stuff Safe [1], [8]) aiuta le biblioteche a divenire nuovamente le custodi degli asset dei periodici, mantenendo al tempo stesso le restrizioni stipulate dall'editore sugli accessi e le licenze.

Il nuovo modello di accesso

La pubblicazione su web facilita l'accesso, permette una disseminazione più rapida e fornisce benefici di costo e spazio per le biblioteche. Lo sviluppo di questo modello ha fatto emergere due sottoprodotti significativi. Gli editori desideravano restringere l'accesso agli utenti che avevano sottoscrizioni, cosa che era realizzata in modo molto semplice obbligandoli a identificarsi su un server centrale, tramite metodi come il sistema di autenticazione Athens e restrizioni di indirizzo IP. Esistono molte possibili condizioni alle quali si può avere accesso al contenuto: sottoscrizioni annuali, offerte a paniere, accesso limitato ai contenuti recenti, e aggregatori di contenuti; tutte sono largamente utilizzate e le variazioni nelle licenze si applicano a ciascuna opzione. L'aggiornamento dinamico delle condizioni di accesso e di utilizzo erano ottenute in modo ragionevolmente semplice tramite una singola sorgente.

Gli utenti dei periodici e i bibliotecari erano preoccupati che questo modello centralizzato fosse caratterizzato da un importante difetto. Gli mancava la ridondanza caratteristica del modello fisico, dove il numero di copie distribuite tra le biblioteche costituiva una misura di sicurezza contro la possibilità di perdere l'accesso al contenuto. In questo ambiente digitale, cosa accade se l'editore di un periodico fallisce o cambia proprietà? Come possono le biblioteche minimizzare i rischi per questi precari materiali digitali? Per rispondere a questi problemi, sono state sviluppate diverse proposte per l'archiviazione dei periodici ([7], [10]). Il deposito legale del contenuto pubblicato fornisce alcune basilari rassicurazioni, ma resta limitato ai materiali che si riferiscono al paese considerato, e l'accesso può essere molto ristretto (spesso all'accesso in un unico luogo, oppure a un numero molto piccolo di copie contemporaneamente). Si stanno sviluppando servizi di archiviazione non-profit di terze parti ([2], [3]). L'approccio LOCKSS permette un'archiviazione personalizzata per ogni biblioteca partecipante, il che facilita lo sviluppo di collezioni persistenti e ben mantenute, e ciò corrisponde agli obiettivi delle biblioteche.

Il sistema LOCKSS

Il sistema LOCKSS, in sviluppo presso le biblioteche dell'Università di Stanford a partire dal 1999, è un software a codice aperto che permette alle biblioteche di ottenere, mantenere, e accedere a copie locali di periodici. Ciò significa che esse possiedono, invece che avere in prestito, questi contenuti. Il software LOCKSS è per lo più reso disponibile come una "piattaforma CD" utilizzata per trasformare un computer in un apparecchio dedicato. La piattaforma CD automatizza in modo significativo la configurazione del sistema LOCKSS. Un amministratore di sistema deve identificare un gruppo di proprietà della rete come prerequisiti, ma tutte le altre opzioni di installazione sono standardizzate e preconfigurate. Il software LOCKSS non è mai installato nel senso tradizionale sul disco fisso. L'unica cosa che persiste su disco a ogni riavvio è il contenuto che la macchina ottiene e conserva. Ogni volta che il computer è avviato, il sistema operativo è reinstallato dal CD su un file system di memoria salvaguardato dalla partizione di swap. Scritto interamente in linguaggio Java, il software LOCKSS può essere distribuito liberamente e, come stabilito dai termini della licenza BSD, il codice sorgente è aperto e liberamente modificabile dai programmatori a condizione che le modifiche siano esse stesse libere e aperte.

Il “daemon” standalone LOCKSS funziona su ogni sistema compatibile con UNIX. Un certo numero di fattori ha orientato lo sviluppo del software LOCKSS ([4], [9]), per fare sì che il sistema sia sostenibile per un lungo periodo di tempo.

- Guasti e obsolescenza dei supporti, dell’ hardware e del software: per minimizzare il rischio di obsolescenza il sistema LOCKSS opera su un nutrito gruppo di piattaforme hardware e software. L’ambiente distribuito rimuove la dipendenza dalla centralizzazione, minimizzando i rischi di guasto hardware. Il protocollo di audit LOCKSS rimuove la necessità di costose procedure di backup.
- Errori di comunicazione e guasti della rete: le comunicazioni di rete possono essere inaffidabili. Al fine di correggere gli errori di contenuto che possono avvenire durante l’acquisizione o la migrazione tra hardware, il protocollo di audit LOCKSS automatizza la loro scoperta e riparazione.
- Errori dell’operatore, o attacchi esterni o interni: operatori di sistema benintenzionati ma inesperti possono danneggiare per errore il contenuto o rendere vulnerabile un sistema. Questi errori accidentali possono essere evitati limitando l’amministrazione di sistema alle fasi di installazione e amministrazione di archivio durante le operazioni giornaliere. In modo simile, minimizzando il numero di applicazioni software installate, e restringendo sia l’accesso fisico sia quello virtuale al computer si riduce la possibilità di attacchi esterni o interni.
- Guasti economici e organizzativi: Le risorse economiche delle biblioteche fluttuano e un investimento costante non può essere garantito. La condivisione della responsabilità del contenuto intellettuale dei periodici elettronici tramite la custodia distribuita del contenuto riduce l’impatto dei guasti.

Raccolta e Accesso al Contenuto

Quando si archivia del materiale, il sistema LOCKSS recupera una copia identica del contenuto del periodico dal sito web dell’editore, e in seguito la archivia nel computer LOCKSS locale. Il contenuto è raccolto in unità gestibili, spesso corrispondenti alle unità di sottoscrizione utilizzate dagli editori. Queste prendono il nome di Unità di Archivio (UA). Tipicamente, una singola unità di archivio raccolta dal sistema LOCKSS corrisponde a un volume completo del periodico. Per ogni piattaforma di pubblicazione si utilizza un plugin LOCKSS, per assicurarsi che il daemon LOCKSS raccolga correttamente solo il materiale rilevante per una singola unità di archivio. Una serie di regole corrispondenti a URL, note come “espressioni regolari”, è utilizzata per definire i contenuti che costituiscono un’unità di archivio. Il sistema LOCKSS raccoglie dal sito dell’editore una particolare interpretazione del contenuto, invece che i documenti sorgente. Il risultato è che quando gli utenti accedono al materiale archiviato il “look and feel” corrisponde a quello inteso dall’editore originario. Il sistema LOCKSS è indipendente dai formati; ogni formato che può essere trasmesso tramite il web può essere archiviato. Questo è particolarmente rilevante nel settore emergente dei periodici dove insiemi di dati supplementari e formati di presentazione innovativi proliferano, a fianco dei formati testuali più tradizionali.

Il sistema LOCKSS contiene un piccolo server web tramite il quale il contenuto è reso accessibile. Per come è progettato, il software LOCKSS è stato concepito per essere integrato con un server proxy istituzionale, il che significa che l’accesso ai contenuti è trasparente per gli utenti: non c’è nessuna interruzione di servizio tra l’accesso al contenuto nel luogo originale e quello di LOCKSS. In risposta alle richieste delle biblioteche, LOCKSS sta sviluppando i componenti che permetteranno di raggiungere i contenuti preservati tramite i sistemi di risoluzione link delle biblioteche.

Informazioni e Risorse:

[1] LOCKSS. www.lockss.org

[2] Portico. www.portico.org

[3] CLOCKSS. www.clockss.org

[4] David S. H. Rosenthal, Thomas S. Robertson, Tom Lipkis, Vicky Reich, Seth Morabito, "Requirements for Digital Preservation Systems: A Bottom-Up Approach", D-Lib Magazine, Volume 11 Number 11 November 2005. www.dlib.org/dlib/november05/rosenthal/11rosenthal.html

[5] OpenURL. www.exlibrisgroup.com/sfx_openurl.htm

[6] LOCKSS Alliance. www.lockss.org/lockss/LOCKSS_Alliance

[7] Anne R. Kenney, Richard Entlich, Peter B. Hirtle, Nancy Y. McGovern, and Ellie L. Buckley, "E-Journal Archiving Metes and Bounds: A Survey of the Landscape", CLIR Report, September 2006. 120 pp. www.clir.org/pubs/abstract/pub138abst.html

[8] Vicky Reich, "Editors' Interview with Victoria Reich, Director, LOCKSS Program", Volume 10, Number 1, Feb 15 2006. www.rlg.org/en/page.php?Page_ID=20894

[9] David S. H. Rosenthal, Thomas Lipkis, Thomas Robertson, Seth Morabito, "Transparent Format Migration of Preserved Web Content", D-Lib Magazine, Volume 11, Number 1, January, 2005. www.dlib.org/dlib/january05/rosenthal/01rosenthal.html

[10] Maggie Jones, "Review and Analysis of the CLIR Report E-Journal Archiving Metes and Bounds: A Survey of the Landscape", February 2007. www.jisc.ac.uk/whatwedo/programmes/programme_preservation/ejournalarchiving.aspx

L' Alleanza LOCKSS

L' Alleanza LOCKSS [6], fondata nel 2005, è un' organizzazione costituita da più di 90 biblioteche che intende offrire a queste istituzioni un forum in cui discutere esperienze e preoccupazioni riguardo a LOCKSS e all'archiviazione dei periodici in generale. I membri beneficiano altresì di opportunità strategiche per aiutarli a determinare le priorità di lungo periodo e le direzioni per l'evoluzione del software e programma LOCKSS. Ogni membro paga una quota di iscrizione annuale, basata sulla dimensione e sulle disponibilità finanziarie dell' istituzione. Sebbene il contenuto dei periodici ad accesso aperto sia disponibile a tutti, l'iscrizione all'Alleanza dà accesso a contenuti riservati agli abbonamenti, a supporto tecnico continuo, e a una collaborazione diretta con il gruppo di sviluppo di LOCKSS.

Informazioni Chiave su LOCKSS

L'approccio LOCKSS è stato lodato sia dalle biblioteche sia dagli editori perché fornisce alla comunità accademica un robusto ed economico strumento per la conservazione del contenuto dei periodici. Più di 125 biblioteche fanno parte dell'Alleanza LOCKSS. A queste si aggiungono altre 60 e più biblioteche che utilizzano LOCKSS ma non partecipano all'Alleanza, per un totale di più di 185. La lista completa delle biblioteche partecipanti si trova qui: <http://www.lockss.org/lockss/Libraries>. Giunto al suo decimo anno, LOCKSS è basato su software a codice aperto che ha ottenuto il premio "Peer Reviewed Award for Best Research" dall'Association for Computing Machinery nel 2004. Più di 300 editori hanno contribuito contenuti di periodici per la conservazione nel sistema LOCKSS. La lista completa degli editori e dei titoli si trova qui: http://lockss.org/lockss/Publishers_and_Titles. Il Programma Pilota LOCKSS nel Regno Unito, finanziato da JISC e CURL, è un' iniziativa d'avanguardia per la realizzazione di un' Alleanza LOCKSS nazionale e per la conservazione collettiva dei contenuti dei periodici in rapporto alle priorità nazionali. L'attività è stata diretta da un gruppo di supporto dedicato all'interno del Digital Curation Centre, e ha consentito l'incontro dell'organizzazione JISC e del suo agente di negoziazione, "Content Complete", con un variegato gruppo di trenta biblioteche di educazione superiore. I fondi delle biblioteche sono stati suddivisi in base ad accordi comuni, e i periodici elettronici accessibili da queste istituzioni sono aumentati di numero (da circa 7.500 a più di 20.000 titoli), e sono disponibili tramite numerose combinazioni di editori e aggregatori. Il Programma Pilota LOCKSS nel Regno Unito è discusso nel dettaglio qui: <http://www.ijdc.net/ijdc/article/view/49/55>.

Conclusioni

In risposta alle preoccupazioni dei bibliotecari, gli editori stanno sempre più partecipando agli sforzi per sviluppare strategie efficaci di archiviazione dei periodici. Il sistema LOCKSS è una componente importante nell'infrastruttura di distribuzione dei periodici, consentendo alle biblioteche di ottenere la custodia dei patrimoni per i quali hanno pagato, e al tempo stesso conformarsi agli accordi di licenza che hanno siglato con gli editori. L'approccio LOCKSS assicura che le biblioteche non solo siano responsabili dell'accesso a breve termine, ma che siano anche coinvolte a diversi livelli nel nuovo modello di archiviazione dei periodici. È probabile che le tecnologie e gli accordi di licenza continueranno a svilupparsi e ad evolversi, permettendo sia alle biblioteche sia agli editori di ottenere, proteggere e mantenere i diritti, le condizioni di accesso e i benefici finanziari che entrambi si aspettano e meritano.