

L'applicazione dei metadati gestionali nel processo di digitalizzazione dell'Archivio Sonoro - Audiovisivo della Discoteca di Stato

Francesco Baldi

L'utilizzo dei metadati per la gestione degli oggetti digitali nel campo dell'archiviazione dei documenti sonori e audiovisivi presenta diversi aspetti critici.

Nell'ambito della Discoteca di Stato, siamo ancora in una fase di studio per la messa a punto di un set di metadati: ci sono infatti alcuni nodi concettuali che devono essere ancora sufficientemente approfonditi.

Alcune considerazioni di carattere generale:

Se si paragona lo sviluppo dei set di metadati in ambito audiovisivo con il più tradizionale ambito bibliografico ci accorgiamo che indubbiamente il primo è un settore relativamente giovane.

Sebbene alcuni standard, nell'ambito della produzione audiovisiva siano già stati sviluppati e alcuni di essi siano già operativi, non molti risultano consolidati.

L'esigenza è duplice, da un lato garantire dei protocolli che ben si adattino alle necessità del singolo ente, dall'altro fare in modo che questi sistemi di metadati sviluppati a livello locale abbiano la possibilità di essere condivisi su scala nazionale e internazionale applicando gli standard internazionali.

Si pensi ad esempio al lavoro del gruppo di MPEG-7

Naturalmente nessuno nega che il contenuto necessiti dei metadati per essere pienamente accessibile, specialmente in un'era in cui i contenuti si vanno dilatando in modo esponenziale.

Nel settore audiovisivo ci confrontiamo con un'enorme quantità di informazioni di cui solo una piccola porzione è data in formato testuale.

Tra l'altro questo contenuto può essere distribuito su supporti che possono andare dal CD al DVD al vinile, da vecchi supporti come i dischi a 78 giri ai nastri magnetici. Oppure i contenuti audiovisivi possono essere diffusi direttamente attraverso la rete.

Inoltre mentre i set di metadati per la ricerca del contenuto alfanumerico – testuale sono stati messi a punto da decenni, le tecniche per la ricerca del contenuto multimediale, vale a dire suono, fotogrammi, immagini in movimento, non sono ancora state messe a punto, o nel migliore dei casi sono allo stato embrionale.

Possiamo aggiungere anche che, il settore che gestisce i contenuti audiovisivi presenta diversi problemi riconducibili alla delicata gestione dei diritti d'autore.

Un altro problema sul tappeto è che i metadati per gli audiovisivi debbono colloquiare con quelli messi a punto per i più tradizionali ambiti bibliografici, in quanto il mercato tende sempre più a una forte ibridizzazione del prodotto culturale informativo.

E' chiaro che ciò fa sì che gli schemi dei metadati dovrebbero tendere ad abbracciare tutte le forme di espressione culturale.

Tendenzialmente andiamo verso un supporto in cui testo immagine e suono sono utilizzati in modo sempre più integrato.

Per questo motivo i set di metadati devono tendere alla multifunzionalità.

Un esempio di metadati multifunzionali è quello che attualmente stanno mettendo in piedi le grandi case discografiche che tentano di razionalizzare la produzione e la distribuzione del materiale sonoro in un mercato caratterizzato da crescente complessità (internet, reti satellitari, oltre ai tradizionali canali di distribuzione) e da un'innovazione tecnologica impetuosa.

Sempre per quanto riguarda i metadati multifunzionali, un ruolo chiave è rivestito dai più importanti operatori del commercio on line (come Amazon.com, che dovendo distribuire libri, musica e video su un piano di parità si è posto il problema di sviluppare dei metadati multifunzionali)

In generale, uno dei maggiori problemi è la difficoltà di sincronizzare obiettivi e finalità nell'ambito degli organismi (commissioni) di standardizzazione dei progetti pilota e di altre iniziative nazionali e internazionali.

Allo stato attuale non esiste uno scambio di informazioni puntuale e sistematico per le iniziative riguardanti gli schemi di metadati per il materiale audiovisivo.

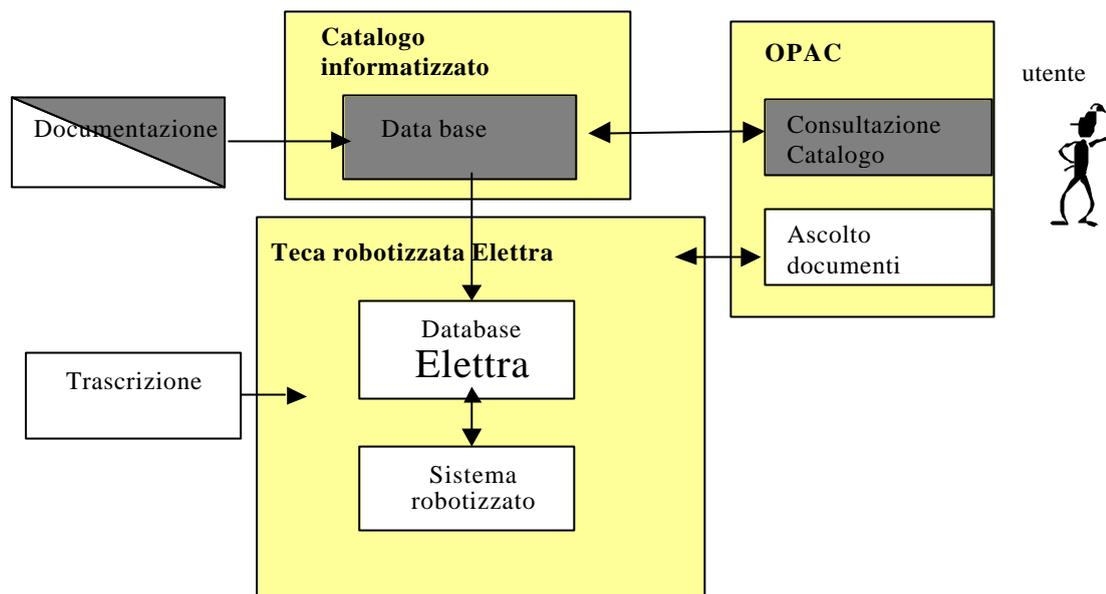
Le commissioni di standardizzazione, poi, hanno tempi diversi rispetto alle esigenze della produzione audiovisiva.

Un altro problema è la scarsa comunicazione tra gruppi di lavoro in ambiti diversi quali il settore tecnico, il settore archivistico e il settore della produzione.

La Discoteca di Stato sta avviando il progetto di trascrizione digitale del proprio archivio.
(sistema di digitalizzazione e gestione dei documenti sonori della Discoteca di Stato)

Questo progetto si colloca nell'ambito dell'iniziativa ADMV (Archivio Digitale della Musica Veneta), frutto della collaborazione tra la Discoteca di Stato, la Biblioteca Nazionale Marciana di Venezia e la Biblioteca Nazionale Universitaria di Torino.

Gli obiettivi del progetto sono il recupero dei documenti musicali e cartacei relativi alla musica veneta del '700 in possesso dei tre enti e la realizzazione di un sistema distribuito che permetta ad un utente esterno, connesso ad un terminale remoto tramite protocollo Z39.50 (standard di interrogazione di cataloghi tra biblioteche), di effettuare ricerche congiunte sui diversi cataloghi.



Il contributo di DS consiste principalmente nella messa a disposizione delle registrazioni sonore, mentre le due biblioteche contribuiscono con le descrizioni bibliografiche e le immagini digitalizzate di manoscritti e musica a stampa.

Fig. 1 - Architettura generale del sistema Elettra

La Discoteca di Stato intende fornire all'utenza, con modalità da definirsi in funzione delle leggi sui diritti d'autore, le funzionalità d'ascolto del segnale musicale, opportunamente compresso.

Per fornire questo servizio e per assolvere ai compiti istituzionali di preservazione e conservazione del patrimonio sonoro e audiovisivo nazionale, la Discoteca costituirà un archivio digitale.

A questo scopo sarà utilizzato il sistema elaborato per la realizzazione dell'archivio digitale della RAI, effettuando le modifiche e personalizzazioni necessarie per soddisfare esigenze diverse rispetto alla RAI, in particolare in materia di conservazione del patrimonio audiovisivo.

Questo sistema di trascrizione digitale è stato sviluppato dalla società Advanced Computer Systems ed è denominato "ELETTRA".

La DS possiede circa 250.000 documenti sonori. La crescita attuale dell'archivio è di circa 6000 nuovi supporti l'anno.

Le tipologie di supporti attualmente in possesso di DS che dovranno essere trascritti in digitale sono:

Dischi a 78, 45, 33 giri, Cassette audio, Nastri magnetici, Compact Disc, DAT, rulli di cera, lacche, matrici, registrazioni su filo magnetico

relativi a :

(musica classica, leggera, jazz, tradizioni popolari, teatro, voci storiche)

L'archivio digitale andrà gradualmente a sostituire i sistemi di consultazione ed archiviazione attualmente in uso presso la Discoteca di Stato.

Il nuovo sistema dovrà permettere un miglioramento del servizio fornito da DS all'utenza rispetto alla situazione attuale, dovrà consentire un risparmio economico sulla gestione conservativa dei supporti.

Il processo di trascrizione consiste nel riversamento di tutti i vecchi supporti analogici in formato digitale sotto forma di file che verranno poi caricati nell'archivio digitale.

Contestualmente alla digitalizzazione, vengono anche generate le versioni compresse del materiale e vengono scannerizzati i documenti cartacei allegati. (copertine, libretti d'opera, depliant, foto, ecc.)

La trascrizione ad alta qualità di tutti i supporti genererà file secondo lo standard europeo Broadcast Wave Format (BWF). Il formato di campionamento sarà generalmente 48 kHz 24 bit stereo salvo casi particolari da definirsi per tutti i supporti analogici.

Dovranno inoltre essere generate due copie compresse per il browsing: una ad alta qualità ad uso degli utenti presso DS ed una adatta allo streaming su internet.

Il bit rate e l'algoritmo di codifica verranno definiti al momento dell'entrata in servizio del sistema di trascrizione, in modo da poter utilizzare la tecnologia più promettente.

Il processo di costituzione del nuovo archivio digitale dovrebbe essere completato nell'arco di cinque anni.

Per ragioni di stato di conservazione e obsolescenza dei lettori, si intende dare priorità massima ai nastri magnetici e ai 78 giri.

Alcuni supporti presentano problemi specifici da affrontare durante la fase di trascrizione

per esempio le diverse curve di equalizzazione adottate per la registrazione dei 78 giri o le variazioni di velocità nominale di 78 giri al minuto che non sempre veniva applicata in fase di incisione.

Di tutto il materiale archiviato, il sistema dovrà produrre automaticamente una seconda copia che verrà poi gestita manualmente all'esterno dell'archivio Digitale per scopi di disaster recovery.

In questa fase saranno prodotti dei metadati gestionali che conterranno le informazioni tecniche relative al processo di trascrizione ed al formato dei file generati.

Alcuni di questi metadati saranno memorizzati direttamente all'interno degli oggetti digitali, mentre altri verranno memorizzati in un database opportunamente strutturato.

I metadati verranno applicati inoltre per gestire le relazioni tra gli oggetti digitali presenti nell'archivio Elettra e le descrizioni che si trovano nel catalogo tradizionale.

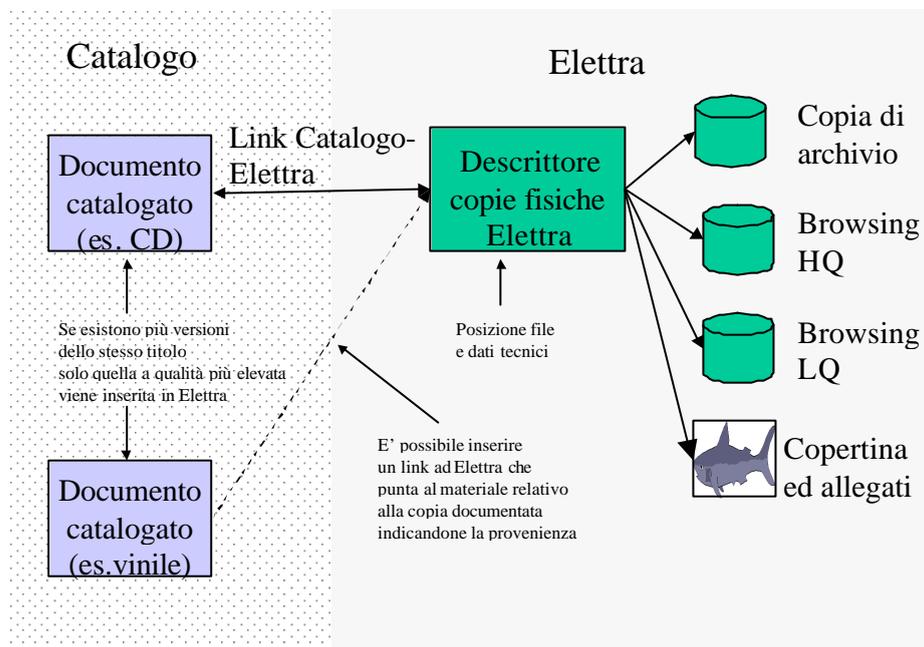


Fig. 2 - Interfacciamento Catalogo – Elettra.

Il processo di trascrizione sarà vincolato alla preventiva, o al più contestuale, immissione nel catalogo informatizzato delle schede catalografiche relative ai documenti da trascrivere.

Un altro aspetto importante è l'indicizzazione del documento digitale in termini di brani e parti all'interno di ciascun brano. Per esempio i movimenti di un concerto o le arie delle opere liriche.

Anche in questo caso per la gestione della segmentazione del documento sonoro audiovisivo in forma digitale dovremo usare i metadati.

Il servizio al pubblico verrà fornito tramite un OPAC appositamente predisposto che, per quanto riguarda le funzionalità di ascolto del materiale, si baserà su due formati compressi a diversa qualità a seconda che l'utente acceda attraverso reti a capacità limitata (internet) o in locale.

Il formato compresso per la consultazione locale deve essere indistinguibile all'ascolto dalla copia non compressa.

Il formato internet seguirà criteri di qualità legati alla capacità di fruizione stimata

L'interfaccia utente deve integrare, per quanto possibile, le funzionalità fornite dal catalogo informatico, ovvero ricerca, selezione e visualizzazione dei dati informativi dei documenti, con le funzionalità di Elettra relative all'ascolto dei documenti stessi e della visualizzazione degli allegati scannerizzati.

L'ascolto via internet relativo ai documenti di cui DS non possiede i diritti si considera limitato al diritto di cronaca (30 secondi?). Il sistema dovrà comunque permettere che l'ascolto venga effettuato sia su brani interi che su spezzoni, secondo criteri definiti dagli amministratori del sistema in accordo con le norme vigenti sul diritto d'autore.

La gestione dei dati sul copyright sarà controllata attraverso l'applicazione dei metadati.

L'ascolto deve essere possibile impiegando software comunemente utilizzati a questo scopo come Real player, Windows media player o Micromedia shockwave.
