

scaletta per seminario 14-12-05

## UN SISTEMA DI ARCHIVI AUDIOVISIVI

sul tema

### Accesso on line ai metadati e ai dati

a cura di

*Francesco Baldi – Discoteca di Stato e Museo dell'Audiovisivo*

novembre 2005

<b>Quadro legislazione in Italia e nell'Unione Europea sugli archivi audiovisivi</b>	1 rapporto <b>Maria Carla Cavagnis Sotgiu</b> Direttore OPIB
<b>Stato della legislazione regionale e locale (Regioni, Province, Comuni) sugli archivi audiovisivi</b>	1 rapporto <b>Vincenzo Bazzocchi</b> - Sovrintendenza per i Beni Librari e Documentari della Regione Emilia Romagna
<b>La legge sul deposito legale in Italia e in Europa</b>	1 rapporto <b>Anna Maria Mandillo - ICCU</b>
<b>La proprietà intellettuale ed economica nell'audiovisivo e nella multimedialità: lo stato attuale, le diverse proposte</b>	1 rapporto <b>Andrea Micciché</b> avvocato
<b>Modelli di archivi audiovisivi, loro gestione, professionalità</b>	2 rapporti <b>Ansano Giannarelli</b> - Università della Tuscia, Consigliere di amministrazione dell'Aamod; <b>Cecilia Angeletti</b> –Biblioteca Nazionale Braidense – Mediateca di Santa Teresa
<b>La catalogazione: standard, modelli, necessità formative</b>	2 rapporti <b>Antonella Pagliarulo</b> - Regesta, documentalista; <b>Cristina Magliano</b> –ICCU
<b>Accesso on line ai metadati e ai dati</b>	2 rapporti <b>Giovanni Bruno</b> - Regesta, amministratore unico; <b>Francesco Baldi</b> Discoteca di Stato Museo dell'Audiovisivo
<b>Restauro analogico e digitale: criteri di scelta e necessità formative</b>	2 rapporti <b>Mario Musumeci</b> Centro sperimentale di cinematografia; <b>Luciano D'Aleo</b> Discoteca di Stato-Museo dell'Audiovisivo

## **1 – I METADATI**

I metadati sono dati sui dati, le stesse schede di un catalogo o i record di un database si possono considerare dei metadati.

Il termine è stato coniato a metà degli anni 60 e le prime applicazioni sono avvenute nel campo della meteorologia e nei dati geospaziali. All'inizio i metadati venivano concepiti come interni al file, ma possono anche essere contenuti nel file o esterni in un database esterno con procedure che vanno ad associare il doc. ai metadati

Le proprietà associate ad un file (le informazioni su un documento di testo, le caratteristiche di un file audio) sono dei metadati memorizzati all'interno del file, le schede di un catalogo o i record di un database sono dei metadati esterni al dato.

Il record di catalogo è un metadato (informazione sull'informazione) il libro è il dato il record catalografico è un metadato (di tipo descrittivo) che permette un accesso e anche una gestione di questo documento.

I metadati stabiliscono le relazioni tra i doc. e i diversi file che vengono prodotti e i loro record descrittivi oggetto a cui è associato un record di catalogo al quale sono associati diversi oggetti digitali (file) tutti i metadati associati (come è stato prodotto il file, che trasformazioni ha avuto, da chi è stato realizzato) informazione generale

## **2. METADATI PER L'ACCESSO ALLE INFORMAZIONI DIGITALI**

Nel contesto digitale hanno la funzione di individuare la risorsa, consentire la sua fruizione e garantire la preservazione nel tempo memorizzando una serie di informazioni di carattere tecnico.

quindi l'oggetto digitale viene prodotto e vengono associati i metadati, gli oggetti digitali possono essere complessi (più file di testo di immagine o insieme dei vari file). Il collegamento tra il record catalografico e i vari file digitali e la fruizione dell'oggetto digitale è reso possibile attraverso i metadati (organizzazione delle pagine di un capitolo di un libro, ascolto dei vari file audio relativi ad un opera musicale)

i metadati sono degli strumenti descrittivi delle risorse digitali

categorie in cui si dividono i metadati:

- metadati tecnici
- metadati amministrativi e gestionali
- metadati descrittivi

Un libro anche se non è descritto da una scheda di catalogo o da una registrazione bibliografica può essere ugualmente individuato e consultato perché ha una sua fisicità. Una risorsa elettronica che non è descritta non è individuabile, è come se non esistesse.

## **3 - INDICIZZAZIONE DEI CONTENUTI**

Aspetto importante per il recupero dell'informazione e per i servizi di consultazione che viene realizzato attraverso l'utilizzo dei metadati. L'indicizzazione o segmentazione permette di associare materiale multimediale agli elementi della catalogazione analitica, individuandolo all'interno di materiali già digitalizzati; ad esempio volendo documentare un singolo concerto (spoglio):

nel caso di documenti cartacei, permette di associare allo spoglio le immagini digitali di una parte delle pagine del libretto della opera madre nel caso di materiale sonoro o video, consente di individuare i segmenti (mediante marcatori temporali di inizio/ fine) dei file già acquisiti in trascrizione, associati all'opera madre

## **4 SERVIZI DI CONSULTAZIONE**

### **4.1 Servizi di accesso – Interrogazione delle banche dati – OPAC (Online Public Access Catalogue)**

I primi OPAC della fine degli anni settanta esibivano limitati punti di accesso e funzionalità di ricerca elementari, usando la stessa informazione bibliografica e i punti di accesso del catalogo cartaceo: (autore, titolo, soggetto, classe)

La seconda generazione deriva dall'utilizzo della tecnologia di information retrieval e dal miglioramento del software di gestione delle basi dati degli anni 80. Possibilità di utilizzare gli operatori booleani nelle ricerche e operatori di prossimità e di troncamento.

- È possibile combinare i termini di ricerca.
- introduzione dei Filtri di ricerca come lingua, data

La terza generazione vede l'introduzione delle interfacce grafiche (GUI - graphic user interface),

- Utilizzo del mouse,
- -ricerca per parole chiave e interrogazioni nel linguaggio naturale.
- Combinazione dei termini di ricerca,
- ricerca semplice e ricerca avanzata

Gli OPAC di ultima generazione sono interrogabili attraverso il Web e simultaneamente attraverso il protocollo Z39.50, presentano funzioni innovative: interattività servizi di richiesta e prenotazione del documento, richiesta di riproduzioni, document delivery (DD) e interlibrary loan (ILL)

### **4.2 Consultazione delle collezioni digitali**

La biblioteca digitale è la biblioteca dell'accesso remoto e deve dare la possibilità di

- Trovare facilmente tutto quello che è rilevante per chi cerca;
- Estendere la ricerca a tutti i documenti in qualche modo collegati

Vantaggi per l'utente attraverso il sistema di accesso:

- Trovare l'informazione senza dover attendere tempi proibitivi senza essere sommerso dalla gran quantità di informazione recuperata e senza avere problemi di orario e di spazio;
- gli utenti possono essere simultanei, quindi una stessa risorsa può essere utilizzata da più utenti contemporaneamente.
- Possibilità di effettuare ricerche su collezioni più ampie
- Avere uno spazio informativo a disposizione utilizzando collegamenti a risorse connesse con cui è possibile interagire.

Si possono utilizzare diverse modalità per accedere alle informazioni:

- Navigazione tra gli indici (browsing)
- Ricerca attraverso un'interfaccia per la consultazione (ricerca a campi fissi, ricerca a testo libero sul modello dei motori di ricerca di Internet)

- La terza funzionalità è quella di visualizzare e usare i documenti localizzati attraverso le due precedenti modalità di ricerca fino a costituirsi la propria biblioteca personale.

Queste funzionalità sono realizzate attraverso l'**Interfaccia utente**:

Per le realizzazioni più recenti coincide con l'interfaccia web e spesso è comune allo stesso OPAC della biblioteca.

L'OPAC diventa lo strumento integrato attraverso il quale è possibile recuperare informazioni bibliografiche e catalografiche e contenuti digitali.

#### **4.3 Infrastruttura per l'accesso**

L'infrastruttura per l'accesso alla biblioteca digitale è la strategia per il futuro.

Ciò che si può sperimentare oggi è la presenza di collezioni digitali non collegate. L'utente è costretto a passare da una banca dati all'altra, con diverse interfacce e modalità di ricerca.

Il servizio di accesso dovrà essere un servizio integrato: l'utente sarà in grado di utilizzare al meglio la collezione della biblioteca (locale e remota) e, in modo trasparente, dovrà avere la possibilità di andare a recuperare l'informazione nelle risorse di internet, filtrate o no dalla biblioteca digitale.

#### **4.4 Interoperabilità dei sistemi**

La possibilità di effettuare ricerche tra sistemi diversi presenta particolari problematiche dovute alle diversità dei sistemi di ricerca, dei formati usati per il contenuto degli standard descrittivi delle risorse.

E' possibile realizzare l'interoperabilità sviluppando la comune adesione ad un unico insieme di standard, oppure mantenendo le differenze delle diverse collezioni digitali ma realizzando dei servizi di ricerca che si interfaccino al contenuto e ai metadati di diverse biblioteche digitali.

I metodi utilizzati per garantire l'interoperabilità dei servizi di ricerca :

standard Web: URL, HTML, HTTP

Sistemi con maggiori funzionalità: Z39.50

### **5. COOPERAZIONE PER LO SVILUPPO DI NUOVI SERVIZI ALL'UTENZA E CONDIVISIONE DELLE RISORSE**

- scambio dei dati digitali
- accesso in remoto a risorse condivise
- biblioteca virtuale

Importante il valore che sta assumendo la cooperazione tra le varie istituzioni e le biblioteche che operano in questo campo allo scopo di unire gli sforzi in modo sinergico per ottimizzare i risultati.

Un insegnamento da considerare è rappresentato dal fatto che gli utenti delle risorse elettroniche non si preoccupano della provenienza delle informazioni, ma piuttosto del fatto che siano autorevoli ed autentiche.

### **6. STANDARDIZZAZIONE**

La standardizzazione dovrebbe favorire la circolazione delle informazioni e consentire l'utilizzo di soluzioni non proprietarie con conseguente risparmio di risorse organizzative e finanziarie.

Progettare un sistema aperto basato sugli standard rende possibili futuri sviluppi e più facile l'integrazione di nuove tecnologie. In questo campo i progressi tecnologici sono molto rapidi. In un

prossimo futuro potremmo avere nuovo hardware per l'acquisizione del digitale (scansione di immagini, cattura di contenuti audio e video, indicizzazione automatica, ecc.) che ci consentirà di svolgere le attività di digitalizzazione più velocemente e con maggiore definizione e qualità. Ci saranno ulteriori miglioramenti nelle tecnologie di rete con miglioramenti nella qualità dei servizi distribuiti. Le connessioni senza l'ausilio dei cavi diverranno più veloci ed economiche rendendo possibile l'accesso ai servizi offerti da più luoghi.

## **7 METADATI DEL GRUPPO ITALIANO DI STUDIO SUI METADATI PROMOSSO DALL'ICCU - IL PROGETTO MAG**

Il gruppo di studio italiano sugli standard e le applicazioni di metadati nei beni culturali promosso dall'ICCU, al quale partecipa anche la Discoteca di Stato, ha elaborato uno schema di metadati MAG (metadati amministrativi e gestionali) con l'obiettivo di dare delle indicazioni per la fase di raccolta e di trasferimento dei metadati e dei dati digitali nei rispettivi archivi rifacendosi al modello OAIS.

### Metadati descrittivi

Si fa l'ipotesi che gli oggetti digitalizzati abbiano metadati descrittivi già presenti in un archivio bibliografico standard.

La presenza di tali metadati nello schema è garanzia del corretto collegamento con tale archivio. La sintassi definita fa riferimento alle specifiche del set di elementi Dublin Core.

Metadati tecnici, per la definizione delle caratteristiche tecniche e per la descrizione del processo di digitalizzazione;

Metadati amministrativi, per la definizione dei metadati relativi alle immagini fisse è stato preso in considerazione lo standard NISO

I set di metadata sono finalizzati a garantire il collegamento univoco fra oggetti digitali e apparati catalografici; b) gestire le funzioni che permettono la fruizione, la conservazione, la tutela e la manutenzione delle collezioni digitali; esplicitare le informazioni relative ai progetti di digitalizzazione (contenuti, responsabili, standard adottati, informazioni tecniche, ecc.); identificare i files che costituiscono la collezione digitale; raccogliere i dati utili alla codifica degli stessi formulata in formato XML secondo gli standard definiti dal MAG-Schema.

Il Mag-schema prevede diverse sezioni:

1. **gen** informazioni generali sul progetto e sul tipo di digitalizzazione
2. **bib** metadati descrittivi sull'oggetto digitalizzato
3. **stru** metadati strutturali
4. **img** metadati specifici relativi alle immagini fisse
5. **ocr** metadati specifici relativi al riconoscimento ottico del testo
6. **doc** sezione utilizzata per descrivere ad esempio un file in formato pdf o rtf
7. **Audio** metadati specifici relativi alle risorse audio
8. **Video** metadati specifici relativi alle risorse video