

Indice

1. Gli standard	3
1.1 Introduzione	3
1.2 Standard	6
1.2.1 Cos'è uno standard	6
1.2.2 Standard e catalogazione	7
1.3 Metadati	9
1.4 Standard di catalogazione	12
1.4.1 Dublin Core	12
1.4.2 ISBD	23
1.4.3 Unimarc	28
1.4.4 MAG	40
2. Informatizzazione di archivi	50
2.1 Progetto DAM	53
2.2 Progetto ADMV	55
3. Progetto Ricordi	57
3.1 L'archivio storico Ricordi	57
3.2 Obiettivi del progetto	58
3.3 Progettazione del database	59
3.4 Uniformità agli standard	71
3.4.1 Analisi dello standard Unimarc in corrispondenza al database Ricordi	71

3.4.2 Analisi dello standard MAG in corrispondenza al database Ricordi	89
3.5 Sviluppi futuri del progetto Ricordi	98
4. Conclusioni	99
Bibliografia	101

1. Gli standard

1.1 Introduzione

Il patrimonio culturale, in tutte le sue diverse forme, soddisfa la necessità fondamentale di garantire la memoria storica. Depositari e divulgatori di questa memoria sono gli archivi, le biblioteche e i musei che, pur rappresentando contesti diversi tradizionalmente separati e autonomi, ne conservano il supporto fisico e ne diffondono la conoscenza. Scopo principale di questi sarà la progettazione e la messa in opera di misure per la conservazione e fruizione dei beni. Molto presto si è capito che, per realizzare questo scopo, era necessario rappresentare i documenti secondo schemi simbolici omogenei e condivisibili, di conseguenza la documentazione ha posto, tra i suoi compiti istituzionali, la catalogazione dei documenti secondo modelli uniformi, allo scopo di dare atto della loro esistenza e dell'eventuale loro disponibilità, nonché per diffondere le informazioni in esse contenute.

Considerando i suoi tratti generali, la catalogazione comprende almeno due fasi distinte: la descrizione dei documenti e la determinazione dei punti di accesso ad essi. La prima consiste nella descrizione analitica delle informazioni relative agli aspetti fisici e al contenuto dei documenti dando atto delle loro peculiarità materiali, e indicando anche gli elementi che consentono di precisarne il contenuto. La seconda, invece, consta di una fase di analisi, di intestazione e di indicizzazione e ha lo scopo di individuare alcune parole-chiave che permetteranno in seguito di scegliere sia le rappresentazioni catalografiche sia, in seconda istanza, i relativi documenti. Infatti un utente cercherà un documento partendo da alcuni termini verbali – nella maggior parte dei casi il nome dell'autore, il titolo dell'opera oppure il suo argomento – i quali assumono il ruolo di dispositivi che ci collegano ai documenti e ci permettono in questo modo la loro fruizione. Lo scopo di queste attività è quello di ottenere delle rappresentazioni catalografiche che siano capaci di rispondere ai bisogni informativi dell'utenza.

Le rappresentazioni simboliche prodotte grazie alla descrizione e all'indicizzazione vengono riunite generalmente nei cataloghi sotto forma di serie di informazioni strutturate e disposte secondo uno o più criteri di ordinamento. La consultazione di questi cataloghi consente a chiunque di rintracciare i documenti necessari e di avere cognizione di quelli che sono disponibili. In questo modo è possibile instaurare un sistema di mediazione incentrato sul catalogo che unisce efficacemente le esigenze informative degli utenti ai documenti stessi.

Con lo sviluppo delle nuove tecnologie dell'informazione gli enti interessati nella conservazione e divulgazione dei beni culturali hanno la possibilità, ma anche il dovere, di migliorare e modernizzare i servizi che offrono e le tipologie dei contenuti. E' in questo senso che si può vedere la tendenza sempre più frequente di archiviazione dei cataloghi in formato elettronico e la digitalizzazione dei documenti.

Sono innumerevoli, ormai, le esperienze e i progetti di acquisizione digitale di materiali – diversi per natura, caratteri intrinseci ed estrinseci, età, pregio e competenza – che da anni biblioteche, archivi, musei e altre istituzioni culturali e scientifiche hanno attivato, sia a scopo conservativo sia per consentire un più facile e rapido accesso ai documenti. Tali esperienze e progetti appaiono quanto mai diversificati, tanto in rapporto ai contenuti quanto al loro grado di fungibilità rispetto agli originali dai quali derivano, ma diversi sono soprattutto le tecniche e gli strumenti impiegati per creare/acquisire i dati in formato digitale, identificarli e descriverli sulla base di un sistema più o meno coordinato di metadati, archivarli in un modo sicuro e rispondente a principi di economicità d'accesso, distribuirli nelle diverse e continuamente rinnovate modalità offerte dall'evoluzione tecnologica.

Un aspetto vitale nella costruzione di collezioni digitali è assicurare l'interoperabilità delle informazioni e delle descrizioni dei diversi elementi culturali. Cosa si intende con interoperabilità? Una convincente definizione la fornisce il Joint Information System Committee (JISC – <http://www.jisc.ac.uk/>): “Interoperabilità è la capacità di un sistema o di un prodotto di operare in combinazione con altri sistemi o prodotti senza richiedere un impegno particolare da parte dell'utente”. In particolare, essere interoperabili significa: “essere attivamente coinvolti nel processo volto ad assicurare che i sistemi, le procedure e la cultura di una organizzazione siano gestite in modo tale da massimizzare le occasioni di scambio e di riutilizzo dell'informazione, sia internamente che all'esterno”.

L'UK Interoperability Focus (<http://www.ukoln.ac.uk/interopfocus/about/>) individua cinque categorie di interoperabilità:

- **Interoperabilità tecnica:** presuppone lo sviluppo e l'applicazione di standard per la comunicazione, il trasferimento, l'immagazzinamento e la rappresentazione dei dati, come ad esempio i protocolli, i formati e i linguaggi di marcatura.
- **Interoperabilità semantica:** presuppone lo sviluppo di dispositivi, innanzi tutto logici, che stipulino le corrispondenze fra termini indicanti funzioni, ruoli e concetti equivalenti in sistemi

di ricerca appartenenti ad ambiti disciplinari diversi. Possiamo prendere tre casi come esempi. Il caso più semplice è costituito dal riconoscimento dell'equivalenza fra termini diversi che esprimono tuttavia un analogo tipo di responsabilità in documenti di differente natura come “autore”, “creatore”, “compositore”, e così via. Un altro è il caso rappresentato da termini identici, utilizzati tuttavia in ambiti disciplinari diversi per indicare concetti o funzioni che non possono essere considerati come “editore” o “fonte”. Vi è, infine, l'eventualità che vengano utilizzati, come chiavi di ricerca equivalenti, funzioni o concetti espressi da termini che hanno, nei diversi ambiti disciplinari, campi di applicazione differenti. Un esempio può essere rappresentato dal “soggetto” che, nel caso dei beni museali, per i quali esiste anche un campo “titolo”, svolge una funzione equivalente a quella svolta dal “titolo” nei documenti bibliografici, per i quali, a loro volta, il campo “soggetto” svolge una funzione affatto diversa.

- **Interoperabilità politica/umana:** presuppone la scelta consapevole da parte delle istituzioni di rendere disponibili le proprie risorse informative sia mediante l'introduzione di nuove procedure amministrative per il controllo degli accessi e per la riscossione dei diritti sui dati, sia con l'immissione di nuovo personale specializzato o la riqualificazione di quello esistente, sia per mezzo della sensibilizzazione della propria utenza.
- **Interoperabilità multidisciplinare:** presuppone l'avvio di iniziative congiunte da parte di biblioteche, archivi e musei, finalizzate da un lato alla maggior conoscenza reciproca e alla individuazione di problemi e obiettivi comuni, e dall'altro lato alla realizzazione di prodotti e di sistemi di ricerca che permettano di sperimentare le soluzioni adottate.
- **Interoperabilità internazionale:** presuppone che tutti i fattori esaminati in precedenza vengano considerati alla luce delle problematiche connesse allo scambio internazionale dei dati, tenendo conto in primo luogo delle barriere linguistiche e delle diversità culturali.

Di conseguenza, per ottenere l'interoperabilità delle informazioni e delle descrizioni dei diversi elementi culturali, è indispensabile lo sviluppo di standard descrittivi (internazionali e nazionali) e di un linguaggio di comunicazione condiviso, affidabile e sufficientemente sofisticato.

1.2 Standard

1.2.1 Cos'è uno standard

Nella vita di ogni giorno esistono norme più o meno esplicite, di cui abbiamo maggiore o minore consapevolezza, che comunemente condividiamo e rispettiamo e che ci consentono di svolgere proficuamente le nostre attività sociali, tra queste sono da includere anche tutti gli standard. Le loro caratteristiche sono molto varie sia per quel che riguarda il loro ambito di applicazione, sia invece per quel che concerne la loro diffusione e accettazione. Non siamo di fronte ad un fenomeno moderno, infatti è possibile leggere dal punto di vista dell'evoluzione degli standard molti momenti della storia della tecnica, ed è possibile delineare una successione cronologica che giunge in epoca moderna all'elaborazione delle odierne norme di unificazione. La successione storica segue a grandi linee la serie delle tappe del progresso tecnologico, determinando una sorta di parallelismo secondo il quale ad una particolare epoca tecnologica trova riscontro una relativa tipologia di standard e ad una maggiore complessità del sistema tecnologico corrisponde una altrettanto articolata struttura di norme unificate. Tuttavia, per le epoche passate non si dovrebbe parlare propriamente di standard nel senso moderno del termine, quanto piuttosto di elementi e norme unificate e unificanti, poiché gli standard moderni contemplano un'elevata quantità di caratteristiche vincolanti che ne limitano l'elaborazione, l'uso e la validità. A scopo esemplificativo, possiamo individuare tre epoche tecnologiche contraddistinte dai relativi proto-standard.

Epoche tecnologiche	Agricola	Industriale	Dell'informazione
Standard	Misura	Conformità	Compatibilità

Non è facile dare una definizione di standard, comunque possiamo affermare che uno standard si esprime attraverso delle regole più o meno complesse o comunque tramite delle direttive coerentemente espresse in un documento, mentre la standardizzazione include l'analisi, la formulazione e l'esplicitazione di tale disciplina. Ogni standard ha lo scopo di guidare nella produzione di un oggetto, nell'erogazione di un servizio o nel compimento di un'attività. Si può dire che uno standard è una regola, ma non è vero il contrario, cioè non tutte le regole che governano l'operare dell'uomo sono da considerarsi degli standard, o meglio non lo sono propriamente, poiché non tutte rispondono ad altri elementi che invece caratterizzano delle norme tecniche uniformi.

Altre caratteristiche importanti che devono essere garantite sono: la stabilità nel tempo, l'uso ripetuto e la larga diffusione. La durata temporale di ogni standard può essere interpretata come una

sorta di *ciclo di vita*: quando uno standard non sembra più adeguato a risolvere i problemi per cui era stato formulato, lo si sostituisce con uno nuovo in grado di soddisfare le esigenze più attuali.

Seguendo la definizione di De Vries (De Vries, 1999), si può dire che comunemente gli standard vengono emanati per risolvere dei problemi di *armonizzazione*. La standardizzazione si occupa cioè di individuare delle regole che, rispondendo alle caratteristiche appena esposte, facilitino le relazioni tra gli individui e/o tra le organizzazioni. Queste entità sono interessate al rispetto degli standard poiché si trovano a dover operare insieme o a fornire strumenti o servizi per il bene comune: devono perciò condividere oggetti, principi e procedure o offrire prodotti fruibili generalmente.

Possiamo ricorrere alla definizione elaborata da De Vries (De Vries, 1999): “La standardizzazione è l’attività che comprende la definizione e la registrazione di una serie limitata di soluzioni in grado di risolvere problemi di armonizzazione attuali o potenziali. Queste soluzioni sono dirette a vantaggio della parte o delle parti coinvolte nel problema e intendono bilanciare i loro rispettivi bisogni. Inoltre esse sono volte ed esigono un uso ripetuto o continuo durante un certo periodo di tempo da parte di un numero cospicuo di soggetti ai quali sono state destinate.”

Nell’ambito catalografico avremo quindi bisogno di definire dei modelli di documentazione standard per preservare a lungo termine il valore dei dati raccolti in un catalogo e per permettere il loro scambio fra le istituzioni

1.2.2 Standard e catalogazione

La rappresentazione catalografica di un documento si compone di due fasi, la prima, di descrizione, è volta ad esporre “in ordine fisso e in modo omogeneo [...] una serie di dati relativi agli elementi di natura fisica, testuale e storica [...] ritenuti essenziali e capaci di permetterne l’esatta individuazione patrimoniale e scientifica” (Petrucchi 2001), la seconda è volta a determinare i percorsi attraverso i quali un utente potrà recuperare la notizia catalografica. La scelta dei dati descrittivi è spesso condizionata da implicazioni di ordine amministrativo, poiché risponde alla necessità di inventariare e tutelare il patrimonio informativo e culturale di una comunità, inoltre tende al raggiungimento della massima condivisione possibile delle informazioni e dei documenti stessi che le tramandano. Questo obiettivo viene perseguito attraverso la formulazione e l’adozione di schemi specifici astratti che contengono le informazioni necessarie ad una corretta ed efficace descrizione e li esprimono sotto forma di schede catalografiche a stampa o di record bibliografici in

forma elettronica. Proprio a questo scopo sono nati i cosiddetti *standard internazionali per la descrizione bibliografica* che a dispetto del loro nome non si interessano soltanto di libri, ma di tutti i documenti che posseggono al pari del libro la qualità di supporti e media informativi.

Il bisogno di un modello standard per la catalogazione dei documenti bibliografici è emerso a partire dalla metà del XIX secolo (Solmine, 1995). Da allora è iniziato un movimento di elaborazione e di adozione degli standard catalografici che gradatamente ha coinvolto tutto il settore della biblioteconomia e della documentazione e che, durante gli ultimi trent'anni, ha ricevuto un notevole impulso.

Molto presto ci si è resi conto che servivano norme specifiche per la catalogazione dei vari materiali (ad esempio fotografie, materiale musicale, materiale cartografico, oggetti,...), diverse da quelle dedicate ai documenti verbali, questo perché ogni materiale ha delle caratteristiche bibliografiche particolari. Ad esempio un libro a stampa possiede un *titolo proprio* con il quale può essere facilmente identificato, mentre la maggior parte delle opere musicali invece non hanno un *titolo proprio* ma ne posseggono uno che consiste nel nome di una forma o di un genere musicale unito alle informazioni riguardanti l'organico vocale-strumentale, il numero d'opera, la tonalità e quant'altro. Solo l'insieme di queste informazioni rende identificabile una composizione.

Alcuni dei principali standard di catalogazione sono:

- Dublin Core – schema di metadati descrittivi
- ISBD – standard di descrizione catalografica
- Unimarc – formato per la rappresentazione e lo scambio di dati catalografici
- MAG – schema di metadati amministrativi e gestionali

Questi definiscono quali metadati debbano essere utilizzati per la descrizione di un documento e quale debba essere la loro struttura. Prima di analizzarli definiamo il termine *metadato*.

1.3 Metadati

Il termine *metadata* fu usato per la prima volta nella seconda metà degli anni Sessanta da Jack E. Myers (cfr. Miller 1996). Il termine *metadata* sta per “dati che caratterizzano altri dati in un modo referenziale, dati intorno a dati. In modo simile, parole intorno a parole. Nell’elaborazione dati, sono gli elementi di definizione che forniscono informazione o documentazione su altri dati gestiti all’interno di una applicazione o di un ambiente. Ad esempio, [...] elementi o attributi dei dati (nome, dimensioni, tipo, etc.), dati relativi ai record o alle strutture dei dati (lunghezza, campi, colonne, etc.) e dati relativi ai dati (dove sono localizzati, come sono associati, loro proprietà, etc.). I metadati possono includere informazione descrittiva sul contesto, la qualità e la condizione, o le caratteristiche dei dati.” (cfr. <http://www.metadata.com/word.htm>)

I metadati quindi sono "dati sui dati", ovvero informazioni, generalmente strutturate e scandite in campi, relative a documenti primari "a testo pieno" (*full-text*), che ne permettono una più efficiente organizzazione e recupero. La loro funzione è permettere o comunque facilitare il raggiungimento dei seguenti obiettivi:

- **Ricerca**, individuare l'esistenza di un documento.
- **Localizzazione**, rintracciare una particolare manifestazione del documento.
- **Selezione**, analizzare, valutare e filtrare una serie di documenti.
- **Interoperabilità semantica**, permettere la ricerca in ambiti disciplinari diversi grazie a una serie di equivalenze fra descrittori.
- **Gestione risorse**, gestire le raccolte di documenti grazie all'intermediazione di banche dati e cataloghi.
- **Disponibilità**, ottenere informazioni sulla effettiva disponibilità del documento.

I campi di una collezione di metadati sono costituiti da informazioni che descrivono le risorse informative a cui si applicano, con lo scopo di migliorarne la visibilità e facilitarne l’accesso. I metadati infatti permettono il recupero di documenti primari indicizzati attraverso sintetiche stringhe descrittive contenute nei record, schede in cui vengono rappresentate le caratteristiche più significative o le proprietà peculiari dei dati in questione, così che la loro essenza possa essere catturata da un’unica concisa descrizione, che, in modo sintetico e standardizzato, fornisce a sua volta una strada di ritorno ai dati stessi. Come appare evidente non si tratta di niente di più o di

diverso da quanto viene garantito, ad esempio, dalle canoniche schede catalografiche bibliotecarie, veri e propri "metadati" rispetto ai "dati" costituiti dai libri e dai periodici conservati sugli scaffali delle biblioteche, cui tali schede si riferiscono. Cataloghi, bibliografie e repertori bibliografici vari, sia in formato cartaceo che elettronico, non sono altro che insiemi organizzati di metadati costruiti e mantenuti da professionisti dell'informazione per individuare, selezionare, localizzare e permettere il recupero di documenti primari più ingombranti e meno maneggevoli delle loro "immagini" o "surrogati" sintetici, ovvero delle schede, dei record, degli item bibliografici.

Il lavoro necessario per estrarre i metadati dai dati è ricompensato dalla loro maggiore manipolabilità rispetto ai dati originari. L'autore produce i dati; l'indicizzatore ne estrae i metadati e li organizza; il lettore interroga i metadati e individua i dati a lui utili in modo più veloce ed efficace di quanto riuscirebbe a fare manipolando direttamente i dati.

Nei progetti di digitalizzazione e nelle attività di gestione degli archivi di oggetti digitali, i metadati rivestono un'importanza crescente, tanto da venire considerati parte costituente della definizione di Oggetto Digitale. “Un oggetto digitale è definito [...] come un qualcosa (es. un'immagine, una registrazione audio, un documento testuale) che è stato codificato in modo digitale e integrato con metadati tali da supportarne l'individuazione, l'uso e l'immagazzinamento”. (cfr. California Digital Library, 2001)

I metadati possono essere distinti in tre categorie in base alla loro funzione:

- **Metadati descrittivi:** hanno lo scopo di facilitare il recupero e l'identificazione della risorsa; sono costituiti da descrizioni dei documenti fonte, o dei documenti nati in formato digitale.
- **Metadati gestionali-amministrativi:** hanno lo scopo di supportare la gestione della risorsa all'interno di una collezione e quindi garantirne l'acquisizione, archiviazione e fruizione, l'utilizzo sulla base di eventuali diritti e licenze, la conservazione e l'uso futuro, la certificazione dell'autenticità e integrità. Nel mondo del digitale, data la labilità dell'informazione elettronica, questi tipi di metadati hanno grande rilevanza ai fini della conservazione degli oggetti digitali.
- **Metadati strutturali:** hanno lo scopo di collegare fra loro i componenti di oggetti informativi complessi, inoltre forniscono dati di identificazione e localizzazione del documento.

L'importanza dei metadati gestionali-amministrativi e strutturali diventa sempre più netta via via che emergono le problematiche poste sia dalla gestione quotidiana degli oggetti digitali (acquisire,

gestire archivi di file digitali), sia dalla necessità di garantire la cosiddetta conservazione digitale (cioè la conservazione e l'accesso a lungo termine degli oggetti digitali). Di minore importanza è a prima vista la funzione di questo tipo di metadati ai fini della ricerca e del recupero della risorsa, tuttavia sono utili nell'organizzazione degli oggetti digitali parte di una collezione (si pensi ad esempio alla sequenza delle immagini delle pagine di un libro digitalizzato), che diventa indispensabile per una corretta fruizione degli oggetti digitali stessi. Inoltre la presenza di questo tipo di metadati accanto agli oggetti digitali è da considerarsi necessaria e preliminare alla applicazione delle diverse possibili metodologie proposte per affrontare il problema della conservazione nel tempo.

In generale tutti gli oggetti informativi, indipendentemente dalla forma fisica o intellettuale che essi assumono, hanno tre caratteristiche, contenuto, contesto e struttura, le quali possono essere ricomposte attraverso il metadato.

- Il **contenuto** stabilisce un nesso tra ciò che l' oggetto contiene in sé o nei suoi dintorni ed è intrinseco all'oggetto informativo stesso.
- Il **contesto** indica il *chi, cosa, perché, dove, come*, tutti aspetti associati con la creazione dell'oggetto e comunque estrinseci all'oggetto informativo stesso.
- La **struttura** si riferisce all'insieme formale di associazioni situate all'interno o tra gli oggetti informativi. Tali insiemi possono essere intrinseci o estrinseci agli oggetti.

Per poter essere interpretati ed elaborati, i record di metadati devono avere una loro sintassi ed essere quindi strutturati in modo da essere riconosciuti in quanto appartenenti ad uno schema predefinito, a cui sia possibile accedere sia dall'uomo che dalla macchina. E' infatti necessario conoscere in modo preciso gli elementi e le loro caratteristiche, controllare la validità dei metadati secondo le regole fissate nello schema a cui appartengono, compararli con elementi di altri schemi per poi trasferirli in database per la gestione e l'accesso. Alcuni di questi schemi di metadati sono stati riconosciuti standard a livello internazionale.

1.4 Standard di catalogazione

1.4.1 Dublin Core

Definizione e finalità

Il *Dublin Core* (DC) nasce con lo scopo di identificare e normalizzare una serie minima di metadati che siano in grado di raggiungere il massimo obiettivo gestionale dei documenti elettronici con il minimo impegno descrittivo. Si tratta di uno schema semplice, costituito da alcuni elementi che sono presenti nella maggior parte delle risorse. E' in grado di mettere in comunicazione le varie comunità per la documentazione del proprio materiale, anche convertendo i metadati da loro prodotti, che possono essere, nelle specifiche applicazioni da loro implementate, molto ricchi e dettagliati. E' dunque un linguaggio comune, definito, nel mondo internazionale del web, con il termine inglese 'pidgin', per mettere in evidenza che si tratta di un vocabolario limitato ed essenziale, adatto ad una comunicazione di base, ma che consente un recupero con un discreto livello di precisione a fronte della disparità di oggetti trattati.

Le caratteristiche principali del Dublin Core lo rendono un formato di metadati che può essere applicato per descrivere oggetti in molteplici tipologie di supporti e interoperabile tra i vari enti. Queste sono:

- **Semplicità di creazione ed utilizzo:** si è cercato di mantenere l'insieme degli elementi del Dublin Core il più limitato e semplice possibile in modo tale da permettere anche a coloro che non sono specialisti di poter creare dei semplici record per la descrizione delle risorse con facilità ed economicità e utilizzare tali record per il ritrovamento di quelle stesse risorse in rete.
- **Interoperabilità semantica:** il Dublin Core fornisce un insieme di elementi comune il cui significato è universalmente compreso e supportato; questo permette la condivisione e il recupero delle informazioni tra enti di diversi paesi in quanto viene utilizzata la medesima terminologia e pratica descrittiva.
- **Ambito d'azione internazionale:** anche se il Dublin Core è stato sviluppato originariamente in lingua inglese, successivamente sono state create, e continuano ad esserlo, numerose versioni in molte altre lingue, compresa anche la lingua italiana, la cui traduzione è stata curata dall'Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane e per le Informazioni Bibliografiche - ICCU [S4]. Inoltre, le attività del DCMI (*Dublin Core Metadata Initiative*), l'organizzazione

del Dublin Core metadata dedita a promuovere l'adozione di standard di metadati interoperabili e a sviluppare dei *vocabolari* di metadati specializzati, coinvolgono diversi paesi.

- **Estendibilità:** è possibile l'estensione dell'insieme standard di elementi, permettendo così di aggiungere delle informazioni specifiche, necessarie ad un determinato ambito. Le varie comunità possono dunque creare ed amministrare set di metadati specializzati aggiuntivi, i quali potrebbero essere utilizzati in congiunzione con quelli del Dublin Core per andare incontro alle esigenze dell'interoperabilità.

Uno dei compiti del DCMi è quello di sviluppare dei *vocabolari controllati* di metadati specializzati a descrivere le risorse. Un cosiddetto vocabolario controllato è un insieme limitato di termini che vengono attentamente definiti e usati coerentemente. Esso viene utilizzato per facilitare le ricerche e per migliorarne i contenuti, viene effettuato cioè un controllo della terminologia per evitare che metadati incoerenti o imprecisi, dovuti ad esempio all'uso di sinonimi, possano degradare la qualità della ricerca. I vocabolari controllati possono anche ridurre gli errori di ortografia che possono avvenire durante la registrazione dei metadati.

Caratteristiche principali

Il Dublin Core standard attuale include due livelli: Semplice o Non qualificato (Unqualified) e Qualificato (Qualified). Il Dublin Core Semplice è formato da quindici elementi che sono caratteristiche che una risorsa può avere come Titolo, Editore o Soggetto. Il Dublin Core Qualificato include, oltre gli elementi del "Non qualificato", un elemento aggiuntivo, "Audience", ed un gruppo di sottoelementi di raffinamento o qualificatori che vengono utilizzati per meglio definire il significato dei vari elementi e che si rivelano utili per il ritrovamento delle risorse. Nelle dichiarazioni fatte per descrivere le risorse, gli elementi e i qualificatori hanno la funzione che nel linguaggio naturale hanno rispettivamente i nomi e gli aggettivi. Gli aggettivi possono essere ignorati da un'applicazione che vuole limitarsi al trattamento e alla comparazione fra elementi, oppure possono essere opportunamente trattati sfruttandone il valore semantico.

Gli aspetti essenziali del Dublin Core Non qualificato sono i seguenti:

- Un record DC Semplice è composto da una o più proprietà e dai valori ad esse associati.
- Ogni proprietà è un attributo della risorsa che si sta descrivendo.

- Ogni proprietà deve essere uno dei quindici elementi di cui il DC Semplice è formato.
- Le proprietà possono essere ripetute.
- Ogni valore è una stringa letterale.
- Ogni stringa può avere una lingua associata (per es. it).

Per quanto riguarda invece il Dublin Core Qualificato:

- Un record DC Qualificato è composto da uno o più proprietà e dai valori ad esse associati.
- Ogni proprietà è un attributo della risorsa che si sta descrivendo.
- Ogni proprietà può essere:
 - uno dei 15 elementi del DC,
 - l'ulteriore elemento del DC Qualificato, cioè *audience*
 - uno dei qualificatori.
- Le proprietà possono essere ripetute.
- Ogni valore è una stringa letterale.
- Ogni valore può avere uno *schema di codifica* associato.
- Ogni schema di codifica ha un nome.
- Ogni stringa può avere una lingua associata (per es. it).

Creazione e sviluppo

Il Dublin Core nasce nel marzo del 1995 in seguito ad un workshop organizzato congiuntamente dal National Center for Supercomputing Applications (NCSA) e dall'Online Computer Library Center (OCLC) e tenutosi presso la sede di OCLC a Dublin, nell'Ohio per discutere riguardo le difficoltà nel trovare le risorse desiderate attraverso il Web con l'obiettivo di sviluppare un insieme di metadati per descrivere le informazioni in rete facilitandone di conseguenza la ricerca e il recupero. Da allora il Dublin Core ha subito una serie di revisioni successive che hanno portato, partendo dal set minimo che fu proposto nel dicembre del 1996 costituito da quindici elementi base e che

divenne ufficialmente la versione 1.0 del formato, passando per la versione 1.1 del Dublin Core Metadata Element Set, alla versione attuale del Dublin Core Semplice che continua a mantenere intatto il set minimo dei 15 elementi che però vengono precisati utilizzando un insieme di attributi più ampio e specifico rispetto al passato. Inoltre per espandere ed affinare tale insieme di elementi è stato definito e sviluppato parallelamente il cosiddetto *Dublin Core Qualificato*.

Nel 2001 il Dublin Core diventa standard ANSI/NISO (Z39.85) e sotto la guida del NISO (*National Information Standards Organization*), nel 2003 viene recepito come norma ISO 15836:2003. L'emissione della norma ISO costituisce un riconoscimento ufficiale per l'uso del set Dublin Core che, sin dalla prima conferenza di Dublin, è stato tradotto in oltre 20 lingue ed utilizzato in tutto il mondo per integrare diverse tipologie di informazioni.

Dettagli tecnici

Il Dublin Core è costituito da un set di 15 elementi, ognuno dei quali è definito utilizzando un insieme di dieci attributi dello standard ISO/IEC 11179 [ISO11179] per la descrizione dei dati. Questi includono:

- **Nome** – Etichetta assegnata al dato
- **Identificatore** – Identificativo univoco assegnato al dato
- **Versione** – Versione del dato
- **Registrazione di autorità** – Entità autorizzata a registrare il dato
- **Lingua** – Lingua nella quale il dato è indicato
- **Definizione** – Indicazione che rappresenta chiaramente il concetto e la natura essenziale del dato.
- **Obbligatorietà** – Indica se il dato è richiesto sempre o solo in alcuni casi (contiene un valore)
- **Tipo di dato** – Indica la tipologia del dato che può essere rappresentato nel valore del dato stesso
- **Occorrenza massima** – Indica un limite alla ripetibilità del dato

- **Commento** – Un’osservazione che concerne l’applicazione del dato

Fortunatamente, sei dei dieci attributi sopra citati sono comuni a tutti gli elementi del DC. Questi sono, con i loro rispettivi valori:

Versione	1.1
Registrazione di autorità	Dublin Core Metadata Initiative
Lingua	En
Obbligatorietà	opzionale
Tipo di dato	stringa di caratteri
Occorrenza massima	illimitata

Le definizioni previste includono sia la forma concettuale che rappresentativa degli elementi del Dublin Core. L’attributo Definizione "cattura" il concetto semantico e gli attributi Tipo di dato e Commento "catturano" la rappresentazione del dato. Ogni definizione del Dublin Core si riferisce alla risorsa che si sta descrivendo. Una risorsa è "qualcosa che ha identità". Ai fini dei metadati del Dublin Core, una risorsa sarà tipicamente una risorsa informativa o di un servizio, ma può essere applicata in un contesto più ampio. La serie degli elementi previsti dal progetto Dublin Core è la seguente:

- **Title/Titolo**

Identificatore: Title

Definizione: Un nome dato alla risorsa.

Commento: In particolare, un Title sarà un nome con il quale la risorsa è formalmente conosciuta. E’ possibile assegnare anche titoli alternativi, come sottotitoli, titoli paralleli, etc. replicando il tag con il nome: Title.Alternative

- **Creator/Autore o creatore**

Identificatore: Creator

Definizione: Un’entità principalmente responsabile della creazione del contenuto della risorsa.

Commento: L’entità responsabile può essere una persona, un’organizzazione o un servizio. In particolare, il nome del Creator dovrebbe essere usato per indicare l’entità. E’ possibile specificare se l’entità è una persona o un’organizzazione replicando il tag con il nome: Creator.PersonalName o: Creator.CorporateName. È utile anche fornire l'indirizzo e-mail dell'autore, replicando il tag nella forma: Creator.PersonalName.Address o

Creator.CorporateName.Address. Per facilitare la ricerca è preferibile porre il nome personale nella forma: Cognome, Nome

- **Subject/Soggetto e Parole chiave**

Identificatore: Subject and keywords

Definizione: L'argomento della risorsa.

Commento: Subject può essere espresso tramite parole chiave o frasi chiave o codici di classificazione che descrivono l'argomento della risorsa. E' consigliabile scegliere per tale elemento un valore da un vocabolario controllato o da uno schema di classificazione formale facendo precedere il termine utilizzato dall'indicazione: Scheme = "nome del sistema scelto". Il tag è replicabile quante volte è necessario.

- **Description/Descrizione**

Identificatore: Description

Definizione: Una spiegazione del contenuto della risorsa.

Commento: Un elemento Description può includere, ma non è limitato solo a: un abstract, un indice, un riferimento al contenuto di una rappresentazione grafica o un testo libero del contenuto.

- **Publisher/Editore**

Identificatore: Publisher

Definizione: Un'entità responsabile della produzione della risorsa, disponibile nella sua forma presente.

Commento: Tale entità può essere una persona, un'organizzazione o un servizio. Normalmente il nome di un Publisher è usato per indicare l'entità stessa. È utile anche fornire l'indirizzo e-mail dell'editore, replicando il tag nella forma: Publisher.Address.

- **Contributor/Autore di contributo subordinato**

Identificatore: Contributor

Definizione: Un'entità responsabile della produzione di contributi al contenuto della risorsa.

Commento: Normalmente il nome di un Contributor è usato per indicare l'entità stessa. Gli autori secondari possono essere persone o enti, distinguibili nella forma: Contributors.PersonalName o Contributors.CorporateName. Se necessario, si può fornire l'indirizzo e-mail dell'autore, replicando il tag nella forma: Contributors.PersonalName.Address. Per facilitare la ricerca è preferibile porre il nome personale nella forma: Cognome, Nome. Il tag è replicabile quante volte è necessario.

- **Date/Data**

Identificatore: Date

Definizione: Una data associata a un evento del ciclo di vita della risorsa.

Commento: Normalmente Date è associata alla creazione o alla disponibilità della risorsa. Si consiglia di usare un numero di 8 caratteri nella forma GG-MM-AAAA. Molti altri schemi possono essere usati ma essi devono poter essere interpretati in modo univoco. Il tag è replicabile per distinguere la data di prima pubblicazione dalla data di ultimo aggiornamento nelle forme: Date.Creation e Date.X-MetadataLastModified.

- **Type/Tipo**

Identificatore: Resource Type

Definizione: La natura o il genere del contenuto della risorsa.

Commento: Il campo Type include termini che descrivono categorie generali, funzioni, generi o livelli di aggregazione per contenuto. Si raccomanda di selezionare un valore da un vocabolario controllato (per esempio il DCMI Type Vocabulary). Per descrivere la manifestazione fisica o digitale della risorsa, si usa l'elemento Format.

- **Format/Formato**

Identificatore: Format

Definizione: La manifestazione fisica o digitale della risorsa.

Commento: Normalmente Format può includere il tipo di supporto e/o le dimensioni della risorsa. Può essere usato per determinare il software, l'hardware o altro apparato necessario alla visualizzazione o all'elaborazione della risorsa. Esempi di dimensioni includono grandezza e durata. Si raccomanda di selezionare un valore da un vocabolario controllato (ad esempio la lista di Internet Media Types [MIME] che definisce i formati dei supporti elettronici, vedi <http://www.isi.edu/in-notes/iana/assignments/media-types/media-types>).

- **Identifier/Identificatore**

Identificatore: Resource Identifier

Definizione: Un riferimento univoco alla risorsa nell'ambito di un dato contesto.

Commento: Si raccomanda di identificare la risorsa per mezzo di una sequenza di caratteri alfabetici o numerici secondo un sistema di identificazione formalmente definito. Esempi di tali sistemi di identificazione includono l'Uniform Resource Identifier (URI), l'Uniform Resource Locator (URL), il Digital Object Identifier (DOI) e l'International Standard Book Number (ISBN).

- **Source/Fonte**

Identificatore: Source

Definizione: Un riferimento ad una fonte dalla quale è derivata la risorsa in oggetto.

Commento: La risorsa in oggetto può essere derivata dalla risorsa espressa nel campo Source in tutto o in parte. Si raccomanda che la risorsa venga identificata utilizzando una sequenza di caratteri alfabetici o numerici secondo un sistema di identificazione che sia definito formalmente.

- **Language/Lingua**

Identificatore: Language

Definizione: La lingua del contenuto intellettuale della risorsa.

Commento: Per i valori dell'elemento Language si raccomanda di utilizzare quanto stabilito dal RF 3066 che insieme allo standard ISO 639 definisce un tag codice di lingua di due o tre caratteri seguito opzionalmente da altri sottotag. Ad esempio, "it" o "ita" per l'italiano, "en" o "eng" per l'inglese, o "en-GB" per l'inglese usato nel Regno Unito.

- **Relation/Relazione**

Identificatore: Relation

Definizione: Un riferimento alla risorsa correlata.

Commento: Si raccomanda di far riferimento alla risorsa per mezzo di una sequenza di caratteri alfabetici o numerici in conformità ad un sistema di identificazione formalmente definito.

- **Coverage/Copertura**

Identificatore: Coverage

Definizione: L'estensione o scopo del contenuto della risorsa.

Commento: Normalmente Coverage include la localizzazione spaziale (il nome di un luogo o le coordinate geografiche), il periodo temporale (l'indicazione di un periodo, una data o un range di date) o una giurisdizione (ad esempio il nome di un'entità amministrativa). Si raccomanda di selezionare un valore da un vocabolario controllato (ad esempio il Thesaurus of Geographic Names [TNG] cfr. <http://www.getty.edu/research/tools/vocabulary/tgn/>) e, se possibile, di utilizzare i nomi di luogo o i periodi di tempo piuttosto che identificatori numerici come serie di coordinate o range di date.

- **Rights/Gestione dei diritti**

Identificatore: Rights Management

Definizione: Informazione sui diritti esercitati sulla risorsa.

Commento: Normalmente un elemento Rights contiene un'indicazione di gestione dei diritti sulla risorsa, o un riferimento ad un servizio che fornisce questa informazione. L'informazione sui diritti spesso comprende i diritti di proprietà intellettuale, Copyright e vari diritti di proprietà. Se l'elemento Rights è assente, non si può fare alcuna ipotesi sullo stato di questi o altri diritti in riferimento alla risorsa.

(cfr. *Dublin Core Metadata Element Set, Version 1.1*)

Tutti questi elementi sono comuni alla normali registrazioni bibliografiche e coprono gli ambiti generici del contenuto (titolo, soggetto, descrizione, fonte, linguaggio, relazione, copertura), della proprietà intellettuale (creatore, editore, autore di contributo subordinato, gestione dei diritti) e della risorsa elettronica nelle sue caratteristiche formali (data, tipo, formato, identificatore).

Supponiamo di voler creare un record DC per il libro di Witten, Bainbridge su Greenstone. Un possibile record Dublin Core è il seguente:

DC.Title	<i>How to Build a Digital Library</i>
DC.Creator.PersonalName	<i>Witten, Ian H.</i>
DC.Creator.PersonalName2	<i>Bainbridge, David</i>
DC.Subject	<i>Greenstone</i>
DC.Subject	Information Storage and Retrieval
DC.Subject	Documentation (CCS)
DC.Subject	H. Information Systems (CCS)
DC.Subject	H.4 Information Systems Application (CCS)
DC.Description	Il libro descrive come creare una
DC.Publisher	Morgan Kaufmann Publishers
DC.Contributor.CorporateName	University of Waikato in New Zeland
DC.Date	2003-01-01 (ISO8601)
DC.Type	Text
DC.Format	text/html (IMT)
DC.Identifier	http://pc-erato2.isti.cnr.it/savino/Greenstone-book.html
DC.Source	1-55860-790-0 (ISBN)
DC.Language	it (ISO639-1)
DC.Date.X-MetadataLastModified	2004-10-22 (ISO8601)

Al fine di garantire funzioni essenziali come l'adozione più ampia possibile dello schema, l'adattabilità a risorse e contesti più vari ed anche la possibilità di raffinamenti ulteriori, nelle specifiche di applicazione dello schema è prevista la massima flessibilità di uso. Tutti gli elementi infatti sono opzionali, ripetibili e presentabili in qualsiasi ordine. Tradotto in più di 50 lingue (l'elenco multilingue degli elementi è visibile al sito della Dublin Core Metadata Initiative), il formato è estensibile mediante la definizione di ulteriori elementi opportunamente identificati da un prefisso che ne indica lo schema di appartenenza. Ciò vale per aggiungere, in specifiche applicazioni che utilizzano Dublin Core, ulteriori metadati tipici di singoli settori disciplinari non coperti dallo schema elementare, ed anche metadati tecnici e amministrativi, utili per la gestione delle risorse.

Oltre gli elementi, il vocabolario di Dublin Core possiede dei qualificatori. Questi sono stati introdotti dalla necessità di esprimere in modo più preciso certi valori identificati dagli elementi. Lo schema di base formato dai 15 elementi è stato così arricchito da due tipi di qualificatori per rendere più precisi, per raffinare e contestualizzare gli elementi stessi: i qualificatori di schema (SCHEME) e i qualificatori di raffinamento (TYPE). I qualificatori di schema specificano gli schemi o i vocabolari controllati utilizzati per esprimere il valore assegnato ad un elemento, i qualificatori di raffinamento forniscono precisazioni sul tipo e la natura di un valore contenuto in un elemento. L'intento che sta alla base della predisposizione dei qualificatori è quello di rendere più precisi i valori espressi negli elementi da chi crea i metadati, ma non di estenderne il significato. Infatti se questo principio non fosse rispettato, l'interoperabilità ne sarebbe compromessa. Si distinguono così i due modelli: Dublin Core Semplice (solo elementi) e Dublin Core Qualificato (elementi e qualificatori).

Il Dublin Core può essere rappresentato utilizzando vari formati con sintassi differenti, inclusi:

- HTML (Hyper Text Markup Language)
- RDF/XML (Resource Description Framework utilizzando eXtensible Markup Language)
- XML (eXtensible Markup Language).

L' HTML può essere utilizzato per esprimere sia il Dublin Core semplice che qualificato. Esso fornisce un formato per rappresentare dei concetti fondamentali del Dublin Core che è facilmente comprensibile sebbene ci siano delle limitazioni per applicazioni più complesse che richiedono l'uso dei qualificatori. In tali casi spesso è più sensato utilizzare XML o RDF/XML che permettono schemi multipli di metadati leggibili dagli esseri umani e analizzabili dalle macchine. Indipendentemente dalla sintassi utilizzata, bisogna comunque tenere presente che i concetti del

Dublin Core sono ugualmente applicabili a ogni formato di file, purchè il metadato sia in una forma che può essere interpretata sia dai motori di ricerca che dagli esseri umani.

La struttura HTML del metatag è composta da due parti:

- il **nome** del metatag
- il suo **contenuto**

espressi nella forma:

<META NAME="DC.Element name" CONTENT="Content value">

Es. <META NAME="DC.Title" CONTENT="The Dark Side of the Moon">

nelle quali la parte di sinistra è predefinita dalla norma, mentre quella di destra deve essere compilata dall'autore.

La sintassi del Dublin Core Qualificato è:

- per i qualificatori **SCHEME**
<META NAME="DC.Element name" SCHEME="Scheme value" CONTENT="Content value">
Es. <META NAME="DC.Date" SCHEME="ISO8601" CONTENT="2006-03-30">
- per i qualificatori **TYPE**
<META NAME="DC.Element name.TYPE identifier" CONTENT="Content value">
Es. <META NAME="DC.Creator.PersonalName" CONTENT="Alighieri, Dante">

Esempio

Esempio di uso del Dublin Core con rappresentazione in HTML

A map in the United States Library of Congress
on-line American Memory Collection



```

<HTML>
<HEAD>
<META NAME = "DC.Title" CONTENT = "Novi Belgii Novæque Angliæ:nec non partis Virginie
tabula multis in locis emendata "/>
<META NAME = "DC.Subject" CONTENT = "Middle Atlantic States -- Maps -- Early works to
1800 -- Facsimiles" SCHEME = "LCSH"/>
<META NAME = "DC.Description" CONTENT = "An historical map showing the coast of New
Jersey as perceived in the seventeenth century"/>
<META NAME = "DC.Source" CONTENT = "G3715 1685 .V5 1969" SCHEME = "LCCN"/>
<META NAME = "DC.Language" CONTENT = "nl"/>
<META NAME = "DC.Relation" CONTENT = "isPartOf
http://lcweb2.loc.gov/ammem/gmdhtml/dsxpimg.html"/>
<META NAME = "DC.Coverage" CONTENT = "New Jersey" SCHEME = "LCSH"/>
<META NAME = "DC.Creator" CONTENT = "Nicolaum Visscher"/>
<META NAME = "DC.Publisher" CONTENT = "Library of Congress, United States"/>
<META NAME = "DC.Contributor" CONTENT = "Historic Urban Plans"/>
<META NAME = "DC.Rights" CONTENT = "http://www.loc.gov/rights_statement.htm"/>
<META NAME = "DC.Date" CONTENT = "1996-04-17" SCHEME = "ISO 8601"/>
<META NAME = "DC.Type" CONTENT = "image.photograph"/>
<META NAME = "DC.Format" CONTENT = "image/gif" SCHEME = "IMT"/>
</HEAD>
<BODY>
...[corpo del documento]
</HTML>

```

1.4.2 International Standard Bibliographic Description: ISBD

Definizione e finalità

Gli standard internazionali per la descrizione bibliografica sono manuali standard, riconosciuti a livello internazionale, di descrizione catalografica: essi offrono regole comuni per descrivere ed identificare qualsiasi tipo di documento, facilitando così lo scambio internazionale di registrazioni bibliografiche. In sostanza “lo scopo principale degli ISBD è definire in tutto il mondo i termini per una descrizione catalografica compatibile, per favorire lo scambio internazionale di registrazioni bibliografiche tra le agenzie bibliografiche nazionali, da una parte, e l'intera comunità bibliotecaria e informativa dall'altra. Gli ISBD, specificando elementi della descrizione bibliografica e prescrivendo l'ordine in cui questi elementi vanno presentati e la punteggiatura con cui essi dovrebbero essere scanditi, propongono:

- di rendere interscambiabili registrazioni provenienti da fonti diverse, in modo che le registrazioni prodotte in un paese possano essere accolte in cataloghi di biblioteche o altri elenchi bibliografici in ogni altro paese;

- di facilitare l'interpretazione delle registrazioni stesse al di là delle barriere linguistiche, in modo che registrazioni prodotte per gli utenti di una lingua possano essere interpretate dagli utenti di altre lingue;
- di favorire la conversione delle registrazioni bibliografiche in forma leggibile dalla macchina”.

(cfr. *ISBD(M)*, 1991).

Creazione e sviluppo

Gli ISBD sono stati redatti da apposite commissioni della IFLA (International Federation of Library Associations and Institutes) sulla scorta delle risoluzioni stabilite durante l' *International meeting of cataloguing experts* che si svolse a Copenhagen nel 1969. Inizialmente fu elaborato uno schema per la descrizione dei libri monografici verbali, ISBD(M), che fu pubblicato nel 1974. Presto però ci si rese conto che era necessario redigere uno standard generale, ISBD(G) pubblicato nel 1977, che servisse come base per la redazione di una serie di norme specifiche, capaci di rispondere alla grandissima varietà tipologica dei documenti. Oggi esistono numerosi standard ISBD, tutti creati per materiali specifici, che offrono un valido e riconosciuto mezzo di descrizione bibliografica. Questi sono:

ISBD (G), lo standard generale di riferimento per tutti i tipi di materiale documentario

ISBD (M), lo standard di riferimento per le monografie

ISBD (S), lo standard di riferimento per i periodici

ISBD (NBM), lo standard di riferimento per il materiale non librario

ISBD (CM), lo standard di riferimento per il materiale cartografico

ISBD (PM), lo standard di riferimento per le opere musicali a stampa

ISBD (A), lo standard di riferimento per il libro antico

ISBD (CF), lo standard di riferimento per computer files che dal 1997 è diventato ISBD (ER), lo standard di riferimento per le risorse elettroniche.

Nel 1978 il Cataloguing Committee stabilì che ogni cinque anni l'ISBD Review Group debba compiere revisioni a questi standard, come aggiunte di esempi e definizioni, interpretazioni di casi dubbi e correzione di refusi, oppure come la produzione di un nuovo standard o la cessazione di uno esistente.

Dettagli tecnici

Ogni manuale ISBD è concepito in maniera da raccogliere un insieme coerente di indicazioni per la propria categoria di pubblicazioni, ma senza alcun intento di rendersi esclusiva. Tutte le ISBD si fondano sull' ISBD generale, ISBD (G).

Ogni ISBD è suddiviso in 8 aree:

- Area del titolo e dell'indicazione di responsabilità
- Area dell'edizione
- Area specifica del materiale
- Area della pubblicazione, distribuzione, ecc.
- Area della descrizione fisica
- Area della collezione
- Area delle note
- Area del numero standard e delle condizioni di disponibilità

Ogni singola area è poi suddivisa in una serie di elementi alcuni dei quali si differenziano da ISBD a ISBD. Di seguito riportiamo le aree e i rispettivi elementi dell' ISBD (G)

Area	Punteggiatura convenzionale che precede (o racchiude) gli elementi	Elemento
1. Area del titolo e della formulazione di responsabilità		1.1 Titolo proprio
	[]	1.2 Indicazione generale del materiale (facoltativa)
	=	*1.3 Titolo parallelo
	:	*1.4 Complementi del titolo
		1.5 Formulazioni di responsabilità
	/	Prima formulazione
2. Area dell'edizione	;	*Successiva formulazione
		2.1 Formulazione di edizione
	=	*2.2 Formulazione parallela di edizione (facoltativa)
		2.3 Formulazioni di responsabilità relative all'edizione
	/	Prima formulazione
	;	*Successiva formulazione
	,	*2.4 Ulteriore formulazione di edizione
		2.5 Formulazioni di responsabilità relative a una ulteriore formulazione
	/	di edizione
	;	Prima formulazione
		*Successiva formulazione

3. Area del tipo e
dell'estensione della risorsa
(facoltativo)

4. Area della pubblicazione,
distribuzione, etc.

4.1 Luogo di pubblicazione,
distribuzione etc.
Primo luogo
; *Successivo luogo
: *4.2 Nome dell'editore,
distributore etc.
[] *4.3 Formulazione di funzione del
distributore
, 4.4 Data di pubblicazione,
distribuzione etc.
(*4.5 Luogo di manifattura
: *4.6 Nome del manifatturiere
,) 4.7 Data di manifattura

5. Area della descrizione
fisica

5.1 Indicazione specifica del
estensione del documento
: 5.2 Altri particolari fisici
; 5.3 Dimensioni
+ *5.4 Formulazione del materiale
allegato(facoltativa)

6. Area della serie

6.1 Titolo proprio della serie o
sottoserie
= *6.2 Titolo parallelo della serie o
sottoserie
: *6.3 Complementi del titolo della serie
o sottoserie
6.4 Formulazioni di responsabilità
relative alla serie o sottoserie
/ Prima formulazione
; *Successiva formulazione
, 6.5 International Standard Serial
Number of series or sub-series
; 6.6 Numerazione all'interno della
serie i sottoserie

7. Area delle note

8. Area del numero standard
(o equivalente) e delle
condizioni di disponibilità

= *8.1 Numero standard (o equivalente)
8.2 Titolo chiave
: *8.3 Condizioni di disponibilità e/o
prezzo (facoltativo)
() *8.4 Condizione (facoltativo)

Particolare importanza nella descrizione ISBD acquista la punteggiatura tecnica, la quale ha lo scopo di separare tra loro le aree e gli elementi in modo da consentire il loro riconoscimento e il loro trattamento automatico tramite computer per fornire una prospettazione dei dati facilmente leggibile dall'occhio umano. L'uso di questa è descritta nello schema sopra. Si deve aggiungere che ogni area, eccetto la prima, è preceduta da punto, spazio, linea lunga, spazio (. –).

Esempi

Ecco degli esempi di descrizione secondo i vari standard ISBD:

- **ISBD (M)** – monografie

I promessi sposi / Alessandro Manzoni. - 5.ed. - Mondadori : Milano, 1977. - 3 v. ; 18cm.-
(Collana Classici)

- **ISBD (S)** – periodici

Canadian journal of African studies = Le journal canadien des études africaines / Committee on African Studies in Canada ; editor Donald C. Savage. – No. 1 (Mar. 1967). – Montreal (2141 Sherbrooke Street West, H3H 1G6) : Loyola College, [1967]. – 25 cm.

Two no. per year. – Title proper taken from cover. – Text in English and French.

ISSN 0008-3968 = Canadian journal of African studies : \$5. per year (Great Britain £1.50, France FF 23)

- **ISBD (NBM)** – materiale non librario

Mike Oldfield boxed [Sound recording]. – London : Virgin Records, 1976. – 4 sound discs (52, 40, 37, 38 min.) : 331 rpm., quad. ; 30 cm, in container.

Includes pamphlet (12p.).

V BOX 1 : £7.99

Record 1: Tubular bells / composed by Mike Oldfield, except The Sailor's Hornpipe (Trad. arr. Oldfield). Originally released: 1973. – Re-mixed in quad. by Phil Newell, assisted by Alan Perkins
Record 2: Hergest Ridge / composed by Mike Oldfield. Originally released: 1974. – Re-mixed in quad. by Mike Oldfield

Record 3: Ommadawn / composed by Mike Oldfield ; lyrics to "On Horseback" by Mike Oldfield and William Murray. Originally released: 1975. – Re-mixed in quad. by Mike Oldfield and Phil Newell

Record 4: Collaborations. Includes variations on original compositions by David Bedford, with improvisations by Oldfield on traditional and other themes. – Recorded between 1974 and 1976. – Re-mixed in quad. by Phil Newell and Mike Oldfield.

- **ISBD (CM)** – materiale cartografico

Foglio 21 : (Trento) / Giacomo Roberti. – Firenze : Istituto geografico militare, 1952. – 104 p. ; 26 cm + 2 c. geogr. ripieg

- **ISBD (PM)** – musica a stampa

Concerto in mi maggiore per violino, archi e organo (o cembalo) [Musica a stampa] : La Primavera : F. I n. 22 / Antonio Vivaldi ; a cura di Gian Francesco Malpiero. – [Partitura]. – Milano : Ricordi, 1950. – 1 partitura (28 p.) ; 27 cm. – (Istituto Italiano Antonio Vivaldi ; 76). – Num. ed.: P.R.434.

- **ISBD (A)** – libri antichi

Trattato della pittura / di Lionardo da Vinci ; nouvarnente data in luce, colla vita dell'istesso autore, scritta da Rafaell du Fresnev ; Si sono giunti, I tre libri della pittura ; ed, Il trattate della statua di Leon Battista Alberti ; colla vita del medesimo. - E di nuovo ristampato, corretto, ed a maggior perfezione condotto. - In Napoli : nella stamperia di Francesco Ricciardo : a spese di Noccola e Vincenzo Rispoli, 1733. - [8], 115, [9], 55, [9]p., 2 f. di lamine : ill., vitratti ; 2° (36 cm). - "Osservazioni di Niccolo' Pussino supra la pittura" p.54-55 bis.

- **ISBD (ER)** – risorse elettroniche

Amazon.com. – [Seattle, Wash] : Amazon.com Inc., c1996. – Servizio in linea. – Modalità di accesso: World Wide Web. URL: <http://www.amazon.com>. – Tit. dell'home page. – Servizio di acquisto in linea di libri, audiovisivi, CD musicali, con informazioni bibliografiche e recensioni degli utenti. – Aggiornato frequentemente. – Descrizione della risorsa al 19 aprile 1999.

1.4.3 Unimarc

Definizione e finalità

Per consentire la gestione dei dati bibliografici da parte di un calcolatore occorre organizzare il contenuto delle registrazioni bibliografiche in una struttura logica “leggibile dalla macchina e formalizzata sulla base di regole precise: questo tipo di struttura si definisce come formato”. (cfr. Scolari, 2000) Quindi possiamo dire che un formato è uno schema di organizzazione e disposizione strutturata dei dati.

Per definire generalmente la natura e lo scopo dei formati di scambio delle informazioni bibliografiche ricorriamo alla definizione proposta da Antonio Scolari: “[...] i formati di scambio dei dati, detti anche formati di interscambio o di comunicazione, sono dedicati allo scambio di dati tra differenti sistemi. Formati di questo tipo debbono essere il più possibile accettabili e accettati da sistemi eterogenei, debbono cioè essere indipendenti dall’hardware e dal software di gestione. Queste osservazioni valgono per i sistemi di gestione di basi di dati di qualsiasi tipo e a maggior

ragione per i sistemi di gestione dei dati bibliografici” (Scolari, 2000). In pratica questi formati definiscono uno schema che prevede dei campi ordinati in modo da contenere tutte le informazioni previste dalle normative catalografiche.

Un ruolo fondamentale tra i formati di scambio di registrazioni bibliografiche è occupato dal cosiddetto Unimarc (*Universal Marc*), diretta estensione del Marc (*MAchine Readable Cataloguing*) che da anni ormai ha assunto validità internazionale per la diffusione dei dati bibliografici. Scopo dell’Unimarc è specificare i content designators (tags, indicatori e codici di sottocampo) da assegnare ai records bibliografici e il formato logico e fisico dei records. L’Unimarc copre una vasta gamma di materiali: monografie, seriali, materiali cartografici, musica, registrazioni sonore, immagini, video e immagini in movimento, libri rari e risorse elettroniche. Il formato si caratterizza per l’analiticità della codifica, che distingue gli elementi descrittivi dagli elementi organizzatori del record, per la stabilità e la garanzia di manutenzione, per l’applicabilità a una vasta gamma di materiali e per la possibilità di certificazione. Si è voluto creare un formato che si basasse sulle regole internazionali ISBD, che infatti è stato incorporato nel formato. Inoltre è stato effettuato uno notevole sforzo di razionalizzazione che si riflette nella suddivisione nel raggruppamento dei campi in blocchi secondo una logica precisa.

Caratteristiche principali

L’Unimarc e più in generale un formato di scambio, in quanto insieme di regole per la codifica e la rappresentazione dei dati catalografici in un contesto automatizzato, è composto da tre elementi: i dati stessi, i codici che li identificano e la specifica struttura in cui i dati sono inquadrati per garantirne il riconoscimento, l’elaborazione e lo scambio fra applicazioni e sistemi.

Un formato standard è costituito di base da:

- **Struttura fisica:** definisce le regole di costruzione della registrazione indicando la tipologia dei campi di cui si compone e la possibile ripetibilità dei campi stessi. Ogni registrazione bibliografica può essere suddivisa in una sequenza di campi di diversa lunghezza, di cui alcuni sono ripetibili e altri no.
- **Identificatori di contenuto:**
 - Etichette: identificano il tipo di campo che le segue e di solito sono formate da tre caratteri.

- Indicatori dopo l'etichetta e all'inizio dei dati: sono di tipo numerico (al massimo due), forniscono informazioni sui contenuti del campo, creano collegamenti fra esso e altri campi delle registrazioni e indicano alla macchina di manipolare i dati.
 - Codici di sottocampi: sono all'interno dei campi e introducono e identificano gli elementi che costituiscono i sottocampi. I codici sono solitamente rappresentati da lettere, a volte numeri preceduti dal simbolo \$.
- **Contenuto delle registrazioni:** costituito da dati contenuti in ciascun campo e sottocampo.

Il formato Unimarc si basa sullo standard ISO 2709 creato dall'International Organization for Standardization per l'interscambio di dati bibliografici su supporti magnetici. Perché una registrazione sia conforme a questo standard è necessario che abbia:

- **Guida o Leader** (Etichetta del record): 24 caratteri con informazioni codificate sul record.
- **Directory** (Indice): riporta un numero di occorrenze uguale al numero dei campi presenti nella registrazione.
- **Datafields** (Campi di dati): indicatori, identificatori di sottocampo, codici di sottocampo, dati (max. 9999 caratteri per campo e 99999 per record).
- **Field terminator** (Separatore di registrazione): codice di fine record (@).

Un record ISO 2709 è dunque un file lineare di testo, ciò consente la sua esportabilità e lo scambio tra database gestiti da software diversi.

Creazione e sviluppo

La storia dell'Unimarc inizia negli anni '60 quando la Library of Congress pensa di convertire il catalogo cartaceo in un catalogo su elaboratore, viene quindi sviluppato il primo formato MARC (*MACHine Readable Cataloguing*), da cui nel 1973 deriverà la prima edizione di ISO 2709. Questo inizia a diffondersi nei vari paesi: nel '75 fu pubblicato UKMARC (per la Gran Bretagna) e INTERMARC (per paesi francofoni e slavi), in Italia si diffuse ANNAMARC (Automazione Nella NAzionale), negli Stati Uniti USMARC. La prima bozza di Unimarc fu creata dall'IFLA nel 1975 e due anni dopo apparve la prima edizione ufficiale. Inizialmente Unimarc è stato usato per lo scambio di informazioni su nastro magnetico, ma da allora è stato adottato da diversi ambienti di

elaborazione e scambi di dati. Nel corso del tempo sono state fatte molte correzioni e aggiunte per poter stare al passo con il continuo sviluppo nel campo della tecnologia e della bibliografia. Nel 1980, per armonizzare lo standard alle revisioni di ISBD, uscì la seconda edizione di Unimarc e nel 1987 ci fu una nuova descrizione dello standard con l'aggiunta di campi specifici dedicati alle registrazioni sonore, video, al materiale grafico, alla musica a stampa. Negli anni '90 l'IFLA costituisce il PUC (Permanent Uniamrc Committee) con il compito di seguire la costante evoluzione del formato, ma anche di promuoverne la diffusione.

Negli anni '90 si moltiplicano i progetti europei per la conversione in Unimarc dai Marc nazionali. In Italia già dal 1985 la BNI aveva adottato Unimarc come formato di scambio delle registrazioni. In Unimarc sono anche i dati del catalogo della Biblioteca nazionale centrale di Firenze. Unimarc tende ad affermarsi non solo come formato di scambio, ma anche come formato di gestione "interno".

Dettagli tecnici

Come già accennato, il formato Unimarc utilizza come standard di codifica del dato l'ISO 2709, in base ad esso ogni record deve essere formato da:

- **Un Record Label o Leader**
- (guida di 24 caratteri)
- Una **Directory** (indice)
- **Data fields** (zona dei dati)
- **Record Terminator**

RECORD LABEL	DIRECTORY	DATA FIELDS	R/T
--------------	-----------	-------------	-----

R/T = Record Terminator

Ogni record deve iniziare con una Record Label di 24 caratteri contenente i dati relativi alla struttura del record richiesti per elaborarlo e solo indirettamente per identificare l'unità bibliografica. I dati codificati presenti in questo sono:

0-4	Lunghezza del record
5	Stato del record
6	Tipo di documento descritto
7	Livello bibliografico
8	Livello gerarchico (rispetto ad altro record)
9	Non definito
10	Lunghezza degli indicatori
11	Lunghezza degli identificatori di sottocampo
12-16	Posizione del carattere iniziale del primo campo
17	Livello di codifica Unimarc
18	Tipo di catalogazione (conformità a ISBD)
20-23	Struttura della directory

Ecco un esempio di Record Label:

00515nam0#2200193####|450#

nuovo record, di 515 caratteri, dove il carattere iniziale del primo campo occupa la posizione 193esima. Il record è relativo a un testo a stampa monografico, non è collegato gerarchicamente ad altri records, è pienamente codificato secondo UNIMARC e la catalogazione è pienamente conforme a ISBD.

Di seguito al Record Label ogni record Unimarc presenta una Directory che consiste di varie stringhe (entries) ognuna delle quali è costituita da:

- Un **Tag** costituito da 3 cifre che identifica l'etichetta di campo.
- Un numero di 4 cifre che indica la **lunghezza del campo** (il numero di caratteri, compresi indicatori, sottocampi, dati e il marcatore di fine campo).
- Un numero di 5 cifre che indica la **posizione del carattere iniziale del campo**.

Directory entry 1			Directory entry 2	Other directory entries	
Tag	Length of Field	Starting Position			F/T

F/T = Field Terminator

Esempio: 0010011000000005001700011...

001= tag (3 car.)

0011= lunghezza campo 001 (4 car.)

0000= posizione di partenza (5 car.)

005= tag (3 car.)

0017= lunghezza del campo 005 (4 car.)

00011= posizione di partenza (5 car.)

.....

Dopo l'indice compaiono in campi di dati (Data fields), che sono la parte del record in cui vengono riportate generalmente le vere e proprie informazioni bibliografiche, in alcuni tipi di campi in forma testuale, in altri in forma codificata. All'interno di un campo, i dati sono generalmente articolati in sottocampi e possono essere introdotti da un massimo di due indicatori che ne specificano il contenuto o il trattamento.

La codifica dei dati è ottenuta mediante enumeratori che si compongono dei seguenti elementi:

- **Tag:** un numero di tre cifre, es. *700*, che definisce la tipologia del dato bibliografico.
- **Indicatori:** 2 numeri di una cifra ciascuno che seguono il Tag, es. *700#0*, che raffinano la definizione del campo o che mostrano come questo dovrebbe essere trattato per la produzione del catalogo. Gli spazi vuoti sono rappresentati col simbolo #.
- Ogni campo si divide in ulteriori sottocampi codificati da appositi **identificatori di sottocampo** a due caratteri: \$, carattere alfabetico minuscolo o un numero da 0 a 9 (es. *700#0\$a ... \$b ...*, etc.). Lo scopo è di raffinare ulteriormente la definizione dei dati.
- Ogni campo è seguito da un **segno di chiusura:** @.

I dati sono le informazioni bibliografiche fornite nel record. Ci sono due tipi di dati:

- **Dati codificati:** sono le informazioni presentate mediante l'uso di codici numerici o alfabetici. Ad esempio *1012#\$aita\$clat* (contiene testo in lingua originale latina e traduzione in italiano).
- **Dati descrittivi:** sono le informazioni presentate mediante l'uso di termini correnti nei linguaggi naturali. Ad esempio *50010\$a[...]\$mlatino e italiano* (contiene l'opera dal titolo [...], in latino e in italiano).

I campi, identificati da tre caratteri numerici, sono disposti in blocchi funzionali (0-1 contengono i dati codificati, 2-8 contengono i dati bibliografici):

- **Blocco 0 - Identificazione**

Etichette: 000-099

Descrizione: Dedicato alle informazioni di identificazione e codificate

Esempi: 001 numero di sistema del record nel proprio archivio

005 data e ora dell'ultima modifica

- **Blocco 1 – Informazioni codificate**

Etichette: 100-199

Descrizione: Dedicato alle informazioni di identificazione e codificate

Esempi: 101 lingua di pubblicazione

102 paese di pubblicazione

- **Blocco 2 – Informazioni descrittive**

Etichette: 200-299

Descrizione: Dedicato alla descrizione, suddiviso secondo le aree ISBD

Esempi: 200 titolo e formulazione di responsabilità

215 descrizione fisica

- **Blocco 3 – Note**

Etichette: 300-399

Descrizione: Dedicato alle note, suddivise sulla base della funzione e della tipologia

Esempi: 304 note riferite al titolo e alla formulazione di responsabilità

327 nota di contenuto

- **Blocco 4 – Legami**

Etichette: 400-499

Descrizione: Dedicato ai legami fra registrazioni

Esempi: 453 tradotto come

455 riproduzione di

- **Blocco 5 – Titoli in relazione**

Etichette: 500-599

Descrizione: Dedicato alle varie forme di titoli

Esempi: 512 titolo di copertina

532 titolo per esteso

- **Blocco 6 – Analisi semantica**

Etichette: 600-699

Descrizione: Dedicato all'analisi semantica (soggetti, classificazioni, termini di thesaurus)

Esempi: 610 termini di soggetto non controllati

676 Classificazione Decimale Dewey

- **Blocco 7 – Responsabilità intellettuale**

Etichette: 700-799

Descrizione: Dedicato alla responsabilità intellettuale

Esempi: 700 autore personale – responsabilità principale

712 ente collettivo – responsabilità secondaria

- **Blocco 8 – Dati internazionali**

Etichette: 800-899

Descrizione: Dedicato alla fonte bibliografica

Esempi: 801 fonte della registrazione

856 localizzazione e accesso elettronici (URL)

- **Blocco 9 – Informazioni ad uso locale**

Etichette: 900-999

Descrizione: Non definito dallo standard e utilizzabile per inserire dati locali

Unimarc prevede un elevato numero di campi, solo pochissimi dei quali sono obbligatori, questi sono:

001 Identificatore del record

100 Dati codificati generali

101 Lingua del documento

200 \$a Titolo proprio

801 Fonte del record

120, 123, 206 per materiali cartografici

Una caratteristica assai importante di Unimarc è la possibilità della gestione dei legami bibliografici fra record differenti ai quali è dedicato l'intero blocco 4: è questa un'innovazione assai importante e qualificante del formato rispetto ad altri Marc nazionali. Ci sono due tecniche per creare il legame:

- di **incapsulamento** (embedding), consiste nell'incorporare interi campi del record collegato, inclusi i rispettivi tag, indicatori e sottocampi, in altrettante occorrenze di un sottocampo \$1 del campo di legame. Esempio: Legame dal record di una monografia a quello della serie di cui la monografia fa parte *410 #1 \$1 001 0000876*

\$1 200 1# \$aCambridge greek and latin classics

- per **sottocampi standard**, fondata sulla citazione in appositi sottocampi del blocco 4 delle informazioni presenti nell'altro record. Esempio: *410 1\$Titolo della serie\$0numero di sistema del record collegato*

Esempio

Esempio di record Unimarc (preso da <http://www.ifla.org/VI/3/p1996-1/unimarc.htm>): Romanzo di Alain-Fournier *Le Grand Meaulnes*, tradotto in inglese *The lost domain*.

```
001 0192122622@
010##$a0-19-212262-2$d£12.95@
020##$aUS$b59-12784@
020##$aGB$bb5920618@
100##$a19590202d1959#####|||y0engy0103#####ba@
1011#$aeng$cfre@
102##$aGB$ben@
105##$aac#####000ay@
2001#$a{NSB}The {NSE}lost domain$fAlain-Fournier$gtranslated from the French by Frank
Davison$gafterword by John Fowles$gillustrated by Ian Beck@
210##$aOxford$cOxford University Press$d1959@
215##$aix,298p,10 leaves of plates$cill, col.port$d23cm@
311##$aTranslation of: Le Grand Meaulnes. Paris : Emile-Paul, 1913@
454#1$1001db140203$150010$a{NSB}Le {NSE}Grand Meaulnes$1700#0$aAlain-
Fournier$f1886- 1914$1210##$aParis$cEmile-Paul$d1913@
50010$a{NSB}Le {NSE}Grand Meaulnes$mEnglish@
606##$aFrench fiction$2lc@
676##$a843/.912$v19@
680##$aPQ2611.O85@
700#0$aAlain-Fournier,$f1886-1914@
702#1$aDavison,$bFrank@
801#0$aUK$bWE/N0A$c19590202$gAACR2@
98700$aNov.1959/209@
```

Descrizione

001 0192122622@

001 (l'identificatore del record) è un unico numero o una combinazione di lettere e numeri che serve ad identificare il record in un file. E' pressochè l'unico campo a non avere indicatori.

010##\$a0-19-212262-2\$d£12.95@

Questo campo contiene l'ISBN (\$a) e il prezzo (\$d). Dal momento che, in questo caso, gli indicatori non hanno alcuna funzione sono entrambi vuoti (##).

020##\$aUS\$b59-12784@

020##\$aGB\$b5920618@

Questi campi mostrano che al dato è stato assegnato un numero di catalogazione nazionale. Il primo (59-12784) è quello assegnato dalla Library of Congress, di conseguenza il codice (\$a) che identifica il paese sarà US.

100##\$a19590202d1959#####|||y0engy0103#####ba@

	1	2	3
0123456789	0123456789	0123456789	012345
100##	19590202d1	959#####	y0engy0103 #####ba@

Questo è un campo a lunghezza fissa dove il significato di caratteri dipende dalla loro posizione:

- | | |
|----------|---|
| cp 0-7 | Mostrano che il record è stato aggiunto al file il 2 febbraio 1959. |
| cp 8-12 | Mostrano che il record è una monografia (d) pubblicata nel 1959. |
| cp 13-16 | Non sono usati per le monografie e perciò sono vuoti (#####). |
| cp 17-19 | Possono includere codici che mostrano l'utenza a cui è diretto il documento, es. per i bambini di anni 9-14 (codice d). In questo caso abbiamo tre caratteri () che indicano che l'informazione non è stata fornita. |
| cp 20 | Indica se il documento è una pubblicazione ministeriale. Il carattere "y" significa "no". |
| cp 21 | Mostra se il record è stato modificato (1) o no (0). |
| cp 22-24 | Lingua di catalogazione, in questo caso inglese. |
| cp 25 | E' il codice di traslitterazione, "y" indica che non è stata effettuata. |
| cp 26-33 | Contengono i codici che identificano i set di caratteri usati. "01" rappresenta basic Latin, 03 extended Latin set. I quattro spazi bianchi |

indicano che non sono stati usati set in aggiunta.

cp 34-35 Indica lo script del titolo del documento che può essere diverso dal titolo del record. In questo caso è Latin.

1011#\$aeng\$cfre@

Questo campo dà dettagli riguardo i linguaggi coinvolti. Il valore del primo indicatore (1) mostra che il documento è una traduzione, in particolare in inglese (\$a) dal francese (\$c).

102##\$aGB\$ben@

Il campo 102 riguarda il paese di pubblicazione e contiene il codice standard ISO indicante il paese, nel nostro caso Gran Bretagna. Oltre a questo è presente il codice per le località (es. “en” per Inghilterra).

105##\$aac#####000ay@

Campo di dati per i libri e le altre pubblicazioni monografiche.

2001#\$a{NSB}The {NSE}lost domain\$fAlain-Fournier\$gtranslated from the French by Frank Davison\$gafterword by John Fowles\$gillustrated by Ian Beck@

Campo del titolo. Prima e dopo l’articolo “The” del titolo c’è un carattere speciale qui rappresentato con {NSB}e{NSE}, per permettere l’ordinamento alfabetico del titolo sotto la lettera “L” di “Lost domain” anziché la lettera “T”.

210##\$aOxford\$coxford University Press\$d1959@

Campo contenente i dettagli di pubblicazione, distribuzione ecc. In questo caso sono necessari solo tre sottocampi: \$a per il luogo, \$c per l’editore e \$d per la data.

215##\$aix,298p,10 leaves of plates\$cill, col.port\$d23cm@

Questo campo contiene la descrizione fisica.

311##\$aTranslation of: Le Grand Meaulnes. Paris : Emile-Paul, 1913@

E’ una nota riferita a campi di collegamento ed è o prodotta dal computer o inserita dal catalogatore.

454#1\$1001db140203\$150010\$a{NSB}Le{NSE}GrandMeaulnes\$1700#0\$aAlainFournier\$f1886-1914\$1210##\$aParis\$coxford University Press\$d1913@

Questo è un campo di collegamento. In questo caso collega la traduzione al documento originale. Ogni sottocampo \$1 possiede i contenuti di un campo: 001 identificatore del record, 500 titolo uniforme, 700 autore, 210 dettagli di pubblicazione.

50010\$a{NSB}Le {NSE}Grand Meaulnes\$mEnglish@

Titolo uniforme. Il primo indicatore ha la stessa funzione che assume anche nel campo 200. Il sottocampo \$m (lingua) permette al catalogo di raggruppare assieme tutte le traduzioni in inglese di quest'opera.

606##\$aFrench fiction\$2lc

Questo campo indica il soggetto. Il codice \$2 mostra che il thesaurus usato appartiene alla lista Library of Congress Subject Headings.

676##\$a843/.912\$v19@

Numero di Classificazione Decimale Dewey dalla XIX edizione.

680##\$aPQ2611.O85@

Numero di classificazione della Library of Congress.

700#0\$aAlain-Fournier,\$f1886-1914@

Il tag significa “autore personale – responsabilità principale”. Il secondo indicatore è 0, ciò significa che il nome dell'autore è inserito sotto il nome, invece che sotto il cognome. Il sottocampo \$f contiene la data di nascita e di morte dell'autore.

702#1\$aDavison,\$bFrank@

Il tag significa “autore personale – responsabilità secondaria”. Il secondo indicatore è 1 e ciò significa che il nome è inserito tramite il cognome. Il nome proprio è nel sottocampo \$b.

801#0\$aGB\$bWE/N0A\$c19590202\$gAACR2@

Questo campo contiene i dettagli della creazione del record. E' specialmente utile nei cataloghi unificati che contengono registrazioni provenienti da enti differenti. Il sottocampo \$a contiene il codice del paese e il \$b quello dell'ente creatrice del record. \$c è la data di creazione e \$g contiene i dettagli del codice di catalogazione usato, in questo caso *Anglo-American cataloguing rules*, 2nd edition.

98700\$aNov.1959/209@

987 è un campo ad uso locale e in questo caso contiene la collocazione fisica.

Naturalmente un utente non consulterà mai un record Unimarc di questo tipo. E' per questo che Unimarc fornisce una gamma di formati utili a rappresentare le registrazioni in forma leggibile.

Ecco un esempio :

843.912 (DC19)

Alain Fournier, 1886-1914

[Le Grand Meaulnes. *English*]. The lost domain / Alain-Fournier;
translated from the French

by

Frank

Davison; afterword by John Fowles; illustrated by Ian Beck. - Oxford:
Oxford University Press, 1959.

- ix, 298p, 10 leaves of plates; ill, col. port; 23cm

Translation of: Le Grand Meaulnes. Paris : Emile-Paul, 1913

ISBN 0-19-212262-2: £12.95

1.Ti 2.The lost domain 3.Davison, Frank 4.French fiction B59-20618

Pressmark: Nov.1959/209

L'Unimarc sta subendo la spinta verso una trasposizione in XML per adeguare maggiormente la distribuzione delle registrazioni bibliografiche alle applicazioni telematiche. Il futuro dei formati di archiviazione digitale è indirizzato verso le applicazioni di XML.

1.3.4 MAG

Definizione e finalità

L'acronimo MAG sta per Metadati Amministrativi e Gestionali. Questo standard è stato creato per raccogliere i metadati amministrativi e gestionali relativi agli oggetti digitali prodotti in un progetto di digitalizzazione in particolare relativamente a:

- modalità e politica di accesso alle risorse digitali;
- aspetti organizzativi e di gestione degli oggetti digitali con i relativi servizi;
- strategie di conservazione di lungo periodo degli oggetti medesimi.

Nel lavoro di definizione di un primo set standard di metadati si è tenuto conto delle principali iniziative e dei più significativi progetti realizzati in ambito internazionale. Sulla base di tali esperienze è stato messo in evidenza che:

- più che per particolari tipologie di documenti fonte (es. periodici, musica a stampa o manoscritta, carte geografiche etc.), il set di metadati di base viene definito per tipologie di oggetti digitali; una prima distinzione può riguardare:
 - immagini statiche,
 - testi prodotti con tecnologia ocr,

- suono,
 - immagini in movimento e oggetti multimediali,
 - *born digital*;
- i metadati ed il loro sistema di gestione devono essere completamente indipendenti da specifiche piattaforme hardware e software, al fine di favorirne un impiego generalizzato;
 - devono invece essere coerenti con le funzioni previste nel modello logico-funzionale standard dell'archivio degli oggetti digitali cui si fa riferimento (es. le funzioni di Immissione, Archiviazione, Gestione, Accesso, Amministrazione, Pianificazione della conservazione, nel modello OAIS).

Caratteristiche principali

Il MAG è un “Application profile” conforme agli standard internazionali in quanto permette l’uso di metadati mantenuti e definiti da altri schemi (Dublin Core e NISO) in combinazione con metadati specifici definiti per una particolare applicazione. Scopo del set MAG è quello di produrre uno schema XML basato sul modello METS utilizzato dalla Library of Congress e di predisporre un set minimo di metadati gestionali da applicare nei progetti di digitalizzazione. Il modello fornisce le specifiche formali per la fase di raccolta e di trasferimento dei metadati e dei dati digitali nei rispettivi archivi. Ogni formato di metadati utilizzato è associato ad un Namespace che fissa in modo non ambiguo la terminologia e ad un XML Schema che ne fissa la struttura sintattica. Con questo Schema XML all’interno di un progetto di digitalizzazione ogni postazione di lavoro è in grado di produrre per ogni oggetto digitalizzato un file guida “standard” ovvero conforme allo schema XML che:

- raccoglie tutte le informazioni sull’oggetto digitalizzato (metadati);
- contiene la mappa di tutti i file generati contestualmente dalla digitalizzazione e relativi all’oggetto digitalizzato.

Grazie alla standardizzazione operata in fase di acquisizione (indipendente da un determinato hardware e da un determinato software) è possibile popolare gli archivi - "i magazzini del digitale" - in maniera standardizzata .

Creazione e sviluppo

Lo standard di Metadati Amministrativi Gestionali (MAG) versione 0 è stato creato nel 2001 dal Gruppo di studio sugli standard e le applicazioni di metadati nei beni culturali costituito dall'ICCU, ente responsabile della diffusione delle normative e degli standard bibliografici, nel 2000. Il Gruppo, di cui fanno parte rappresentanti di biblioteche, musei e archivi è nato con l'obiettivo di coordinare a livello nazionale le implementazioni di metadati nei progetti di digitalizzazione nei diversi settori dei beni culturali e di raccordare le iniziative italiane con quelle europee e internazionali. Questo si è suddiviso al suo interno in due sottogruppi:

- Sottogruppo Metadati descrittivi
- Sottogruppo Metadati amministrativi gestionali – MAG

Dopo la versione 0 è stata creata una versione 1.0 nel 2002, 1.5 nel 2004, 2.0 nel 2005, l'ultima è la 2.0.1 che è stata pubblicata nel marzo 2006. Da luglio 2003 si è formalmente costituito un Gruppo di lavoro permanente, il Comitato MAG che prosegue le attività del Gruppo di studio sugli standard e le applicazioni di metadati nei beni culturali con particolare riferimento alle attività connesse alla diffusione ed evoluzione del set di Metadati Amministrativi Gestionali. I suoi obiettivi principali sono:

- Diffusione dello Schema MAG
- Completamento dello schema per le fasi AIP (Archival Information Package) e DIP (Dissemination Information Package) dello schema MAG
- Mantenimento ed evoluzione dello schema
- Produzione di manuali e linee guida
- Assistenza agli implementatori
- Formazione e promozione
- Rapporti con altri progetti e agenzie (Progetti europei, Dublin Core, etc.)
- Sviluppo di schemi e profili di applicazione di metadati descrittivi nei diversi settori della cultura

Dettagli tecnici

Analizziamo ora la sintassi e la struttura del MAG, che si può definire ad albero. L'elemento radice di ciascun record MAG è il metadigit, che si compone di una serie di sottoelementi. Questi sono:

- **gen**, contiene informazioni generali sul progetto e sul tipo di digitalizzazione effettuata
- **bib**, contiene metadati descrittivi che identificano la risorsa analogica (l'oggetto fonte)
- **stru**, contiene metadati che descrivono la struttura logica dell'oggetto digitalizzato
- **img**, raccoglie i dati relativi alle immagini fisse
- **ocr**, descrive un file di testo ottenuto mediante riconoscimento ottico automatico del contenuto
- **doc**, descrive un file di testo corretto manualmente e eventualmente codificato in qualche linguaggio formale
- **audio**, descrive i dati relativi a un file contenente una traccia audio
- **video**, descrive i dati relativi a un file contenente uno stream video
- **dis**, raccoglie i dati circa la fruibilità dell'oggetto digitale

Di questi solo i primi due sono obbligatori. Ognuno di questi elementi ha dei sottoelementi alcuni dei quali si compongono di altri elementi.

Le tipologie dei valori contenuti negli elementi possono essere raggruppate in due categorie:

- **tipi semplici**

xsd:string	stringa di caratteri
xsd:positiveInteger	numero intero positivo
xsd:double	numero reale in doppia precisione
xsd:datetime	data e ora
xsd:time	ora
xsd:anyURI	Uniform Resource Identifier di una risorsa in rete
xsd:restriction	restrizione di un tipo semplice
xsd:enumeration	elenco di possibili valori
xsd:ID	l'attributo di questo tipo ha come valore un identificatore univoco

xsd:IDRef	l'attributo di questo tipo ha per valore il riferimento a un xsd:ID dichiarato in altro punto del documento
xsd:float	rappresenta un numero a singola precisione
dc:elementType	elemento Dublin Core come definito in http://dublincore.org/schemas/xmls/simpledc20020312.xsd

- **tipi complessi**

xsd:sequence	sequenza ordinata di sotto-elementi
xsd:choice	alternativa tra più sotto-elementi diversi

La sintassi del MAG può essere schematizzata così:

nome	nome dell'elemento o del sotto-elemento
descrizione	descrizione del contenuto
path	percorsi degli elementi all'interno dei quali gli elementi si possono usare
tipo	assume la forma namespace di riferimento: tipo di dati - cfr. la prima colonna nelle due tabelle in alto
obbligatorietà	M (Obbligatorio), O (Opzionale), MA (Obbligatorio se applicabile)
ripetibilità	Sì/No
attributi	eventuali attributi del tag, con indicazione del tipo e dell'obbligatorietà
valori	applicabile solo a tipi semplici - elenco di valori possibili o esempi d'uso
componenti	applicabile solo a <code><xsd:sequence></code> o <code><xsd:choice></code> - lista dei sotto-elementi
Note	eventuali note d'uso o problemi ancora aperti

Esempio

Esempio di record MAG di una digitalizzazione parziale di un testo a stampa: immagini + ocr (preso da <http://www.bncf.firenze.sbn.it/progetti/mag/CFI0404365.xml>)

```
<?xml version="1.0" encoding="UTF-8" ?>
```

```

<metadigit xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:niso="http://www.niso.org/pdfs/DataDict.pdf"
xmlns:xlink="http://www.w3.org/TR/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://www.iccu.sbn.it/metaAG1.pdf"
xsi:schemaLocation="http://www.iccu.sbn.it/metaAG1.pdf metadigit.xsd"
version="1.0">

```

L'elemento <metadigit> è la radice di ciascun record MAG. Essendo di tipo xsd:sequence avrà al suo interno una sequenza ordinata di sottoelementi che sono: <gen>, <bib>, <stru> (opzionale), (opzionale), <audio> (opzionale), <video> (opzionale), <ocr> (opzionale), <doc> (opzionale), <dis> (opzionale). Nel nostro caso avremo solo, oltre agli elementi obbligatori, e <ocr>. L'attributo del metadigit è version che è una stringa di caratteri indicante la versione del MAG utilizzato, in questo caso 1.0.

```

<gen>
  <stprog>http://www.bncf.firenze.sbn.it/progetti/arsbni/index.html</stprog>
  <agency>IT:BNCF</agency>
  <access_rights>1</access_rights>
  <completeness>1</completeness>
</gen>

```

La sezione <gen> raccoglie informazioni generali sul progetto e sul tipo di digitalizzazione. Anche questa è composta da una serie di sottoelementi, in questo caso abbiamo:

- <stprog>: un riferimento a una risorsa di rete che descrive gli standard di progetto, è di tipo xsd:anyURI cioè Uniform Resource Identifier di una risorsa in rete
- <agency>: agenzia responsabile del processo di digitalizzazione, è una stringa di caratteri. Si raccomanda l'uso della sintassi Unimarc definita per il campo 801, cioè codice del paese in due caratteri, IT sta per Italia e codice dell'ente per intero oppure secondo una sigla riconosciuta, BNCF indica la Biblioteca Nazionale Centrale di Firenze.
- <access_rights>: riguarda le condizioni di accesso all'oggetto analogico originale, permette solo due valori: 0 (uso riservato all'interno dell'istituzione), 1 (uso pubblico).
- <completeness>: livello di completamento della digitalizzazione, anche questo permette solo due valori: 0 (digitalizzazione completa), 1 (digitalizzazione parziale).

```

<bib level="m">
  <dc:identifier xsi:type="SBN">CFI0404365</dc:identifier>
  <dc:identifier xsi:type="BNI">99-2276</dc:identifier>

```

```

<dc:title>Canzoni / Giacomo Leopardi ; versione in prosa, note e postfazione di
Marco Santagata</dc:title>
<dc:creator>Leopardi, Giacomo</dc:creator>
<dc:publisher>Milano : A. Mondadori, 1998.</dc:publisher>
<holdings ID="H1">
  <library>Bibl. Nazionale Centrale di Firenze</library>
  <inventory_number>CF005360457</inventory_number>
  <shelfmark>GEN.A05.5826</shelfmark>
</holdings>
</bib>

```

L'elemento <bib> contiene metadati descrittivi che identificano la risorsa analogica. Ha un attributo, level che indica il tipo di documento ("m" sta per monografia). I sottoelementi di cui è composto sono i tag Dublin Core più <holdings> e <piece>, in questo caso:

- <dc:identifier>: identificatore univoco alla risorsa analogica nell'ambito di un dato contesto. In questa versione del MAG abbiamo l'attributo xsi:type, nella 2.0 è stato eliminato.
- <dc:title>: nome dato alla risorsa analogica.
- <dc:creator>: entità che ha la responsabilità principale della produzione del contenuto della risorsa analogica.
- <dc:publisher>: entità responsabile della produzione della risorsa analogica, disponibile nella sua forma presente.
- <holdings>: inventario e collocazione della risorsa analogica digitalizzato. Ha un attributo, ID e tre sottoelementi:
 - <library>: istituzione che possiede la risorsa analogica.
 - <inventory_number>: numero di inventario della risorsa analogica.
 - <shelfmark>: collocazione dell'unità fisica.

```

<img>
  <sequence_number>1</sequence_number>
  <nomenclature>Frontespizio: r</nomenclature>
  <file Location="URL" xlink:href="./CFI0404365/IMG/A0000003.TIF" />
  <md5>984df2d193f132ad7fa53efccd6b49ab</md5>
  <image_dimensions>
    <niso:imagelength>2630</niso:imagelength>
    <niso:imagewidth>1542</niso:imagewidth>
  </image_dimensions>
  <image_metrics>
    <niso:samplingfrequencyunit>2</niso:samplingfrequencyunit>
    <niso:samplingfrequencyplane>2</niso:samplingfrequencyplane>
    <niso:photometricinterpretation>BlackIsZero</niso:photometricinterpretation>

```

```

    <niso:bitpersample>1</niso:bitpersample>
  </image_metrics>
  <dpi>300</dpi>
  <format>
    <niso:name>TIFF</niso:name>
    <niso:mime>image/tiff</niso:mime>
    <niso:compression>CCITT Group 4</niso:compression>
  </format>
  <scanning>
    <niso:scanningsystem>
      <niso:scanner_manufacturer>Zeutschel</niso:scanner_manufacturer>
      <niso:scanner_model>OS5000</niso:scanner_model>
      <niso:capture_software>Volumina 3.0</niso:capture_software>
    </niso:scanningsystem>
  </scanning>
  <datetimecreated>2001-01-31T12:05:52</datetimecreated>
</img>

```

La sezione raccoglie i dati relativi a ciascuna immagine. Essa contiene una serie di sottoelementi, in questo esempio essi sono:

- <sequence_number>: numero progressivo che identifica un file multimediale all'interno del file XML.
- <nomenclature>: descrizione di un'immagine.
- <file>: link al file multimediale. Nell'esempio ci sono due attributi: "Location" e "xlink:href" opzionale, identifica una risorsa remota, deve essere una URI.
- <md5>: impronta del file, serve a monitorare l'integrità del file. Le regole per l'individuazione dell'impronta devono essere definite localmente.
- <image_dimensions>: dimensioni dell'immagine, le sue componenti sono:
 - <niso:imagelenght>: lunghezza dell'immagine, cioè la dimensione verticale espressa in pixel.
 - <niso:imagewidth>: larghezza dell'immagine, cioè la dimensione orizzontale espressa in pixel.
- <image_metrics>: caratteristiche tecniche principali dell'immagine, secondo lo standard NISO, ha una serie di sottoelementi:
 - <niso:samplingfrequencyunit>: unità di misura del campionamento usata da <niso:xsamplingfrequency> e <niso:ysamplingfrequency>. Ha tre possibili valori: 1 (nessuna unità di misura definita) 2 (inch, pollice) 3 (centimetro).

- <niso:samplingfrequencyplane>: piano focale di campionamento. Ha tre possibili valori: 1 (camera/scanner focal plane): quando non sono definite le dimensioni dell'oggetto che si sta digitalizzando (per es. quando si riproduce con una fotocamera), 2 (object plane): quando l'oggetto e la riproduzione hanno la stessa dimensione (per es. quando si riproduce con uno scanner), 3 (source object plane): quando la dimensione della riproduzione è maggiore dell'oggetto originale (per es. quando si riproduce da un microfilm).
- <niso:photometricinterpretation>: interpretazione fotometrica dei bit del campione.
- <niso:bitpersample>: numero di bit per ciascun campione, profondità-colore. I valori che può assumere sono:
 - 1 (bitonale, bianco e nero)
 - 4 (4-bit scala di grigi)
 - 8 (8-bit scala di grigi o gamma di 256 colori)
 - 8,8,8 (24-bit RGB)
 - 16,16,16 (48-bit TIFF, HDR)
 - 8,8,8,8 (32-bit CMYK)
- <dpi>: pixel per inch, risoluzione spaziale in entrambe le direzioni orizzontale e verticale (pixel quadrato).
- <format>: tipologia dell'immagine e modalità di compressione, ha tre componenti:
 - <niso:name>: formato dell'immagine.
 - <niso:mime>: tipologia MIME del file-immagine, ha dei valori predefiniti.
 - <niso:compression>: algoritmo di compressione dell'immagine, anche questo ha valori predefiniti.
- <scanninig>: modalità di svolgimento della scansione. Ha delle componenti opzionali, nell'esempio è presente <niso:scanningsystem> che descrive il sistema di acquisizione dell'immagine. Si compone di tre sottoelementi:
 - <niso:scanner_manufacturer>: produttore dell'apparecchiatura di acquisizione.
 - <niso:scanner_model>: marca e modello dell'apparecchiatura di acquisizione.
 - <niso:capture_software>: software di acquisizione dell'immagine.
- <datetimecreated>: data e ora di creazione del file


```

<ocr>
  <sequence_number>1</sequence_number>
  <nomenclature>Frontespizio: r</nomenclature>
  <file Location="URL" xlink:href="./CFI0404365/OCR/A0000003.txt" />
  <md5>94b5e3b55502a3c7885a2c34b2cc68f5</md5>
  <source Location="URL" xlink:href="./CFI0404365/IMG/A0000003.TIF" />
  <format>
    <niso:name>text</niso:name>
    <niso:mime>text/plain</niso:mime>
  </format>
  <software_ocr>Cuneiform 99</software_ocr>
  <datetimecreated>2001-01-31T14:36:10</datetimecreated>
</ocr>
</metadigit>

```

La sezione <ocr> contiene metadati relativi a file di testo ottenuti mediante riconoscimento ottico automatico del contenuto, ha una serie di componenti, nell'esempio sono:

- <sequence_number>: numero progressivo che identifica un file multimediale all'interno del file XML.
- <nomenclature>: descrizione del contenuto del testo.
- <file>: link al file di testo. Nell'esempio ci sono due attributi: "Location" e "xlink:href" opzionale, identifica una risorsa remota, deve essere una URI.
- <md5>: impronta del file, serve a monitorare l'integrità del file. Le regole per l'individuazione dell'impronta devono essere definite localmente.
- <source>: link al file sorgente usato per il riconoscimento ottico (per es. l'immagine TIFF ottenuta con lo scanner).
- <format>: tipologia dell'immagine e modalità di compressione. Nell'esempio ha due sottoelementi:
 - <niso:name>: formato del file di testo (es, text).
 - <niso:mime>: tipologia MIME del file testo. Può assumere solo una serie precisa di valori tra cui "text/plain".
- <software_ocr>: tipo di software usato per il riconoscimento dei caratteri.
- <datetimecreated>: data e ora di creazione del file.

2. Informatizzazione di archivi

In questi ultimi anni le biblioteche e molte altre "istituzioni della memoria" hanno dato vita a progetti di digitalizzazione. La programmazione di un'attività di digitalizzazione dalla portata più o meno ampia deve considerare innanzitutto gli scopi che si intendono perseguire e gli esiti attesi. Anche se in linea di principio tali attività rientrano in programmi generali di sfruttamento delle possibilità offerte dallo sviluppo tecnologico, gli scopi che un singolo istituto si propone di raggiungere possono essere quanto mai diversificati. Si può infatti andare dalla creazione di copie d'archivio destinate a preservare la memoria di determinati documenti al trasferimento in formato digitale di dati già conservati su altro supporto (ad es. il microfilm); dall'integrazione dei dati risultanti dall'attività su piattaforme digitali preesistenti all'intento primario di distribuire i dati digitali localmente o mediante accesso remoto (OPAC, Internet); dalla necessità di un aggiornamento tecnologico complessivo destinato a rinnovare la qualità dei servizi offerti e a costituire magari fonte di introiti, alla sperimentazione di metodi e tecniche propria di molti progetti-pilota. Altre questioni riguardano le previsioni circa l'ampiezza dell'eventuale utenza e le sue, per quanto prevedibili, aspettative; un conto è infatti offrire un servizio destinato soltanto all'accesso locale, un altro conto predisporre un sistema di distribuzione che garantisca all'utenza la continuità world wide del servizio. E ancora: i dati distribuiti avranno caratteristiche tali da poter essere in qualche misura sostitutivi dell'originale, oppure serviranno soltanto all'identificazione e al riconoscimento dei singoli documenti? E tali dati sono destinati a essere utilizzati soprattutto da un pubblico generico o si rivolgono invece a settori specializzati del mondo della ricerca? E l'accesso ai dati potrà essere consentito agli utenti senza mediazione alcuna oppure sarà necessario attivare un servizio assistito? Le risposte a questi ed altri simili interrogativi, qualora essi vengano esplicitati nella fase preparatoria, serviranno a circoscrivere il campo di applicazione del progetto, a commisurarne meglio i risultati in rapporto alle risorse, ad accrescerne l'efficacia e in certa misura a evitare inutili appesantimenti delle procedure, che si traducono inevitabilmente in aggravio dei costi.

In generale gli obiettivi di un progetto di digitalizzazione possono essere così riassunti:

- promuovere un nuovo utilizzo da parte degli studiosi mediante la diffusione su larga scala di collezioni locali o uniche;
- migliorare l'uso dei documenti attraverso tecniche di manipolazione dell'immagine;

- creare collezioni virtuali attraverso l'integrazione di vari formati o di materiali distribuiti in luoghi diversi;
- limitare la consultazione diretta di documenti originali in particolari condizioni di criticità;
- facilitare l'accesso a materiale di difficile fruizione.

Un progetto di digitalizzazione può essere esemplificato in cinque fasi:

- **Ricognizione e individuazione dei fondi:** consiste nell'individuazione dei fondi che si vogliono trattare e nella quantificazione dei materiali in rapporto alle tipologie alle quali appartengono. E' una delle fasi più importanti in quanto vengono studiate le caratteristiche fisiche dei documenti dalle quali dipenderà il modo di operare.
- **Organizzazione dell'indicizzazione dei dati e definizione della struttura dei metadati:** consiste nell'inventariare, descrivere e catalogare i materiali. L'identificazione dei record digitali dovrà essere infatti tale da consentire la rapida e precisa associazione degli stessi con quelli descrittivi e/o di catalogazione.
- **Digitalizzazione:** consiste nella digitalizzazione del materiale scelto.
- **Integrazione dei dati in un sistema preesistente o costruito ad hoc:** consiste nell'inserimento dei dati nel database.
- **Post-produzione e manutenzione**

Come già accennato, le finalità generalmente perseguite con l'attuazione di progetti di digitalizzazione sono di due specie:

- Assicurare la **conservazione** della memoria
- Favorire la **distribuzione** dei dati su scala più o meno ampia

Nel primo caso la preoccupazione principale sarà quella di ottenere dati in un formato il più possibile “trasparente” ai futuri cambiamenti tecnologici – in altre parole, di archiviare i dati in formati non soltanto largamente diffusi, ma che riducano quanto più possibile le perdite di informazione legate ai diversi tipi e fattori di compressione.

Nel secondo caso, a tale assicurazione di fondo dovranno aggiungersi altre considerazioni strettamente legate agli scopi del progetto e all'impiego delle risorse disponibili. Possiamo distinguere tre casi:

- **Distribuzione locale di tipo stand alone:** non sarà necessario nessun sistema servente, ma soltanto una o più postazioni di accesso, magari dedicate; in questo caso la complessità del sistema sarà assai ridotta, in quanto le procedure potranno in sostanza ridursi alle fasi di acquisizione e archiviazione, mentre la distribuzione dei dati potrà avere carattere manuale o automatico, a seconda che il supporto di archiviazione viva on line o off line. Un'architettura di questo genere non necessita, al livello più basso, di programmi per la gestione dei metadati e l'associazione degli stessi con i dati digitali, ma può servirsi di semplici elenchi che descrivano i contenuti dell'archivio.
- **Distribuzione locale attraverso una rete:** sono necessari l'appoggio a una stazione servente destinata a tale scopo e la presenza di programmi per la gestione delle comunicazioni e dei dati. Converrà in questo caso valutare la convenienza strategica di inserire il progetto in un'architettura preesistente, nel caso questa esista, o invece di crearne una che abbia un carattere dedicato, e che sia perciò dipendente dal server principale soltanto per il passaggio delle comunicazioni.
- **Distribuzione globale con possibilità di accesso remoto tramite OPAC e Internet:** l'accesso tramite OPAC richiede architetture hardware e software fortemente strutturate e altamente complesse, il cui sviluppo è in buona parte regolamentato da norme e indicazioni emesse da istituti preposti al coordinamento e al controllo, come l'ICCU. La finalità in questo caso è principalmente, in Italia, la distribuzione dei dati all'interno di SBN (Servizio Bibliotecario Nazionale) per mezzo di protocolli di comunicazione ad alto grado di standardizzazione e di controllo, come ad esempio il protocollo Z3950. La scelta o l'esigenza di inserire un progetto di digitalizzazione in un ambiente già così notevolmente strutturato e normalizzato ridurrà la questione relativa ai contenuti tecnici al rispetto delle procedure di integrazione dei dati nel sistema esistente. Molto più aperta la seconda soluzione, ovvero Internet, la quale non offre per contro le garanzie di omogeneità e la standardizzazione assicurate dalla prima. La stazione servente dovrà in questo caso possedere funzioni di web server del tutto automatizzate, in modo da garantire la regolare e continua distribuzione dei dati. Resta a carico di questa soluzione l'interfacciamento delle modalità di accesso con il database contenente i dati; su questo versante le soluzioni tecnologiche sono quanto mai diversificate, e ciò appare un vantaggio soprattutto

nel caso di realtà e progetti dalle ridotte dimensioni, nelle quali simili sistemi possono essere configurati con un alto grado di specificità e precisione.

Passiamo ora all'analisi di due progetti di digitalizzazione: il progetto DAM del Teatro Alla Scala di Milano e il progetto ADMV Archivio Digitale della Musica Veneta. Le principali differenze tra i due si riscontrano nelle finalità: il primo è volto essenzialmente ad un uso interno del Teatro, il secondo invece punta allo scambio delle risorse informative degli enti partecipanti e a facilitare l'accesso integrato dell'utenza a tali risorse attraverso una disseminazione selettiva dell'informazione. Questo accesso integrato implica una pianificazione concettuale d'alto livello che permetta all'utente di ottenere tutti i documenti pertinenti ad un determinato soggetto, indipendentemente dal loro supporto, da chi lo possiede e dalla loro localizzazione geografica.

2.1 Progetto DAM

Il progetto DAM è nato nel giugno del 1999 con lo scopo di digitalizzare e mettere a disposizione il materiale conservato negli archivi del Teatro Alla Scala di Milano. Il termine "DAM" sta per *Digital Asset Management*, ossia gestione delle risorse digitali; nel caso del Teatro Alla Scala esso assume anche il significato di Depositi Archivi Magazzini. Il materiale d'interesse di questo progetto è costituito da tutto il patrimonio storico-artistico scaligero tra cui partiture, libretti, fotografie, costumi, attrezzeria, bozzetti, figurini, locandine, audio.

Gli obiettivi principali del progetto sono:

- Rendere fruibile e facilmente accessibile gli archivi da qualsiasi postazione interna autorizzata attraverso un'applicazione Intranet chiamata "LaScalaDAM".
- Assicurare l'accesso dal Teatro e dalla tournée a tutte le informazioni relative all'attività artistica del Teatro alla Scala dal 1950 ad oggi: dati di locandina di opera, balletto, concerti, recital e altre manifestazioni, oltre che a tutti i file multimediali relativi al materiale di scena presente negli archivi scaligero.
- Rendere fruibili parte dei dati anche al pubblico esterno, tramite postazioni DAM all'interno del Sistema Scala e una sezione "archivio" sul sito Internet del Teatro.

Prima fase del progetto è stata quella di realizzare 9 archivi digitali, per la gestione, la conservazione e la catalogazione del patrimonio artistico: costumi, accessori ai costumi, attrezzi di

scena, locandine, fotografie, bozzetti, figurini, audio e partiture. Questo ha rappresentato il nuovo punto di partenza per la successiva realizzazione di un'applicazione intranet centralizzata - La Scala Dam - che ha reso gli archivi e tutti i dati di locandina degli spettacoli scaligeri, dal 1950 ad oggi, disponibili non solo agli utenti attuali (Responsabili degli archivi e dei magazzini), ma anche alla Sovrintendenza, alle Direzioni ed a tutti i Reparti del Teatro. Dal 2000 in poi, con i primi archivi digitali consegnati, fino ai giorni nostri, il Teatro alla Scala ha così potuto attingere sempre più dai contenuti multimediali inseriti e catalogati al fine di soddisfare un vasto fronte di esigenze, fra cui, le principali sono:

- Ricerca di asset archiviati/conservati per riutilizzo su nuovi allestimenti: la rapidità della ricerca e la completezza delle informazioni fornite hanno permesso un miglioramento della percentuali di riutilizzo e snellito di molto il processo già in essere.
- Selezione e spedizione ad editori di contenuti relativi agli spettacoli della stagione in corso e/o di quelle precedenti.
- Alimentazione snella, supportata da un workflow di autorizzazione e di pubblicazione per alimentare intere aree del sito web istituzionale (la Stagione, l'Archivio Web).

L'ultima fase del progetto, partita nell'aprile 2004, ha comportato un upgrade tecnologico che consente di centralizzare tutti i contenuti, distribuiti in precedenza sui vari archivi locali, in un'unica base dati.

Il DAM nasce principalmente per uso interno al Teatro, ne consegue che:

- Non tutti gli utenti hanno accesso a tutto il materiale. Sono stati cioè stabiliti dei profili personalizzati per alcune macro-categorie di utenti caratterizzati da particolari permessi d'accesso. Tutto ciò che è pubblicato su Internet è accessibile a tutti ma è solo una minima parte del materiale. Ovviamente i responsabili dei singoli archivi hanno piena visibilità su quanto è di loro competenza, ad esempio i responsabili dell'archivio fotografico avranno accesso agli ingrandimenti delle foto e alla scheda che ne descrive l'identificazione, le caratteristiche e la collocazione fisica.
- Non sono stati presi provvedimenti per permettere l'interoperabilità con collezioni o con progetti di altre istituzioni. Quindi non sono state attuate procedure di standardizzazione sia riguardo al contenuto che ai formati di scambio.

2.2 Progetto ADMV

L'idea del progetto ADMV Archivio Digitale della Musica Veneta risale al 1996, ma l'anno ufficiale di inizio è il 1999. Ormai si può parlare di ADM, omettendo cioè l'aggettivo "veneta", in quanto lo stato attuale del progetto è uno stato di transizione verso obiettivi che potrebbero andare oltre la connotazione geografica originaria che l'ha caratterizzato fino a questo momento. ADMV nasce come cooperazione fra la Discoteca di Stato, la Biblioteca Nazionale Universitaria di Torino e la Biblioteca Nazionale Marciana di Venezia per permettere la valorizzazione delle loro collezioni e la creazione di un servizio di biblioteca digitale innovativo.

I materiali su cui si scelse di lavorare sono:

- I fondi manoscritti contariniani della Biblioteca Nazionale Marciana con partiture settecentesche di compositori come Cavalli, Monteverdi, Marcello, Perez, Hasse, Scarlatti.
- Le partiture autografe e non di Vivaldi, partiture di autori in prevalenza veneti come Marcello, Galuppi e Stradella e le immagini dei "Balletti sabaudi" della Biblioteca Nazionale Universitaria di Torino.
- Le esecuzioni di tali musiche trasmesse dalle collezioni sonore della Discoteca di Stato.

Gli obiettivi iniziali del progetto, nonché punti di partenza per la selezione del materiale, sono tre:

- **Obiettivo di natura culturale:** riunire virtualmente collezioni omogenee che altrimenti sarebbero disperse.
- **Obiettivo di servizio:** creare un servizio di biblioteca digitale che possa permettere la connessione funzionale fra la ricerca bibliografica e l'individuazione di un record descrittivo di una partitura, la visualizzazione dell'immagine digitalizzata della partitura, e l'ascolto del documento sonoro digitalizzato relativo a quell'opera.
- **Sostenibilità:** capacità di attrarre finanziamenti e nuovi partner per mantenersi nel tempo, per ottenere ciò è necessario rivolgere i servizi ad un target il più ampio e variegato possibile.

Si è detto che uno degli obiettivi del progetto è quello di realizzare un servizio di biblioteca digitale. Questa va intesa come un'architettura a supporto di una varietà di servizi di distribuzione di informazioni digitali. Il contesto della biblioteca digitale è ampio e variegato, vi è una mole di

dati da trattare per le diverse funzioni di gestione delle risorse elettroniche che vengono eseguite essenzialmente in modo automatico. Gli oggetti digitali sono accessibili in un contesto distribuito e si rende necessario controllare la loro autenticità, l'effettiva disponibilità a fronte di un identificatore ad esse associato, consentire il loro utilizzo sulla base di condizionamenti tecnici e giuridici e garantirne la conservazione e l'uso nel tempo. Per questa ragione è necessario da una parte uno stretto collegamento fra tutte le informazioni che sono utili al trattamento delle risorse (fra metadati descrittivi e gestionali ed anche fra i vari tipi di metadati gestionali), dall'altra un buon livello di standardizzazione nella definizione degli elementi che identificano le caratteristiche tecniche ed amministrative degli oggetti digitali. Per questo, caratteristiche fondamentali di ADMV sono:

- **Architettura totalmente distribuita:** la base dati ADMV non esiste fisicamente, essa si presenta come un insieme virtuale di archivi locali che condividono la stessa struttura del record e che sono consultabili unitariamente dall'utente finale. In questo modo ciascun polo ha piena autonomia organizzativa nella gestione delle collezioni, ha il suo catalogo, le sue attrezzature tecnologiche, i suoi applicativi. Tutto ciò rende il sistema più aperto, e utilizzabile da istituzioni diverse (biblioteche, archivi sonori e video, etc.), i quali possono aderirvi anche in momenti successivi.
- **Cooperazione basata non sulla condivisione di un software ma sulla condivisione di standard:** ADMV non prescrive l'uso di software specifici, ma obbliga i partner ad adottare gli stessi standard sia per i record bibliografici che per gli oggetti digitali. Per quanto riguarda i primi, gli applicativi di catalogazione utilizzati dai partner, pur differenti, dispongono di funzionalità di export in formato Unimarc standard, per i secondi è necessario che gli oggetti digitali siano corredati di metadati amministrativi e gestionali in formato MAG, a prescindere dall'applicativo utilizzato per produrli. Relativamente agli standard di risoluzione o compressione delle immagini o del suono invece non ci sono vincoli, ma piuttosto indicazione di "buone pratiche".
- **Condivisione del medesimo strumento per la ricerca:** per l'OPAC è previsto un applicativo unico, nel senso di uno stesso strumento replicato nei tre siti. Esso è la chiave di accesso e lo strumento centrale di erogazione dei servizi, è cioè il mezzo che consente, dal lato dell'utente finale, la riconduzione a virtuale unità delle collezioni dei singoli partner, che possono essere indagate simultaneamente o singolarmente e vengono viste dall'utente in maniera integrata, come se fossero una sola.

3 Progetto Ricordi

A seguito della sempre maggiore necessità di informatizzazione degli archivi, la Biblioteca Digitale Italiana ha realizzato un programma nazionale per coordinare e promuovere le attività di digitalizzazione. La Direzione Generale per i Beni Librari e gli Istituti Culturali nel 1999 ha commissionato uno Studio di Fattibilità per fornire un quadro di riferimento necessario alla formulazione di un progetto nazionale per la nascita di una biblioteca digitale, che potesse inserirsi a pieno titolo tra le analoghe iniziative europee ed extraeuropee. Tale studio, nell'analizzare i diversi interventi di digitalizzazione di alcune istituzioni culturali, mise in evidenza come in Italia, a fronte dell'esistenza di numerose iniziative, mancasse una cornice coordinata di intenti e finalità. A seguito di tali considerazioni, nel 2001 venne istituito un Comitato Guida composto da rappresentanti delle biblioteche statali e regionali, dei musei, dell'università, della ricerca con il compito di definire le linee guida del programma, fissandone i principi fondamentali e delineando il quadro di riferimento culturale e scientifico entro cui collocare le iniziative esistenti ed avviare quelle nuove. Fu avviato quindi un programma coordinato per definire le priorità degli interventi, indicare standard e tecnologie comuni da adottare, raccordare le attività di livello nazionale con le iniziative internazionali.

Diversi sono i progetti che sono stati avviati a seguito di tali direttive, tra questi abbiamo il progetto *La musica in rete. Digitalizzazione dell'Archivio Storico Ricordi* avviato nel 2004. Il progetto prevede il completo riversamento digitale dei diversi fondi custoditi nell'Archivio, facendo fronte così a esigenze sia di carattere conservativo, sia scientifiche e di ricerca, sia editoriali e di valorizzazione dell'Archivio e della cultura italiana.

3.1 L'Archivio Storico Ricordi

Casa Ricordi è la più antica casa editrice musicale italiana in attività con quasi duecento anni di storia. Essa venne fondata nel 1808 da Giovanni Ricordi. Il suo Archivio, ora in deposito presso la Biblioteca Braidense, rappresenta una delle più importanti raccolte musicali private.

L'Archivio Storico Ricordi nasce dall'esigenza di raccogliere e catalogare documenti, come manoscritti autografi, libretti, bozzetti, lettere, fotografie, manifesti, funzionali al lavoro della

società. Nel corso di quasi due secoli, la raccolta continua ad arricchirsi, dando vita a un inestimabile patrimonio che rappresenta la storia della musica e del teatro italiano.

Attualmente comprende:

- 3.593 partiture dal '700 al primo '900 di cui 2.246 autografe,
- 15.000 lettere di musicisti e librettisti,
- 10.000 bozzetti e figurini,
- 9.000 libretti,
- 4.000 foto d'epoca,
- manifesti liberty firmati dai grandi della grafica,
- piante di scena,
- appunti.

Tra questi sono conservati gli originali di 23 delle 28 opere scritte da Verdi, tutte le opere di Puccini (con la sola eccezione di *Rondine*), che si affiancano a moltissime opere di grandi autori di musica classica: Rossini, Bellini, Donizetti, Paganini, Liszt, fino ai più moderni come Respighi, Malipiero, Casella, Pizzetti, Nono, Maderna, Donatoni.

3.2 Obiettivi del progetto

Il progetto prevede di raccogliere, catalogare e digitalizzare i materiali originali, presenti nell'Archivio. Per avviare il progetto sono state scelte due opere significative:

- *La Bohème* di Giacomo Puccini andato in scena al Teatro Regio di Torino nel 1896
- *Falstaff* di Giuseppe Verdi in scena al Teatro alla Scala nel 1893.

La scelta è stata spinta dal desiderio di creare, a partire dalla singola opera musicale, percorsi che investono le problematiche della creazione artistica, quelle relative all'allestimento teatrale, la

cultura letteraria e figurativa, fino alla storia dell'imprenditoria italiana e quella della società italiana.

I piani di lavoro per Falstaff e Bohème prevedono la digitalizzazione di:

- partiture autografe
- disegni di scene e costumi degli allestimenti originali
- libretti manoscritti e a stampa
- carteggi fra Ricordi e gli autori
- carteggi con i librettisti, direttori d'orchestra, e collaboratori delle prime rappresentazioni
- foto originali dell'epoca

Oltre a questo lavoro di digitalizzazione è necessaria la creazione di un archivio digitale conforme agli standard internazionali e ampliabile in futuro con materiali di altre opere. Per la realizzazione dell'archivio si è deciso di creare un database multimediale.

3.3 Progettazione del database

Una *base di dati* è una collezione di dati, utilizzati per rappresentare le informazioni di interesse per un sistema informativo. Un *sistema di gestione di basi di dati* (in inglese *Data Base Management System*, abbreviato con DBMS) è un sistema software in grado di gestire collezioni di dati che siano grandi, condivise e persistenti, assicurando la loro affidabilità e privacy.

Un *modello dei dati* è un insieme di concetti utilizzati per organizzare i dati di interesse e descriverne la struttura in modo che essa risulti comprensibile ad un elaboratore. Due principali tipi di modelli sono:

- **Modelli logici:** utilizzati per l'organizzazione dei dati, ad essi fanno riferimento i programmi. Sono indipendenti dalle strutture fisiche. Esempi di questo modello sono: relazionale, reticolare, gerarchico, a oggetti.

- **Modelli concettuali:** permettono di rappresentare i dati in modo indipendente da ogni sistema, cercando di descrivere i concetti del mondo reale. Sono utilizzati nelle fasi preliminari di progettazione, il più noto è il modello Entity-Relationship.

Attualmente il modello relazionale è quello più diffuso. Proposto da E. F. Codd nel 1970 (E. F. Codd, 1970) per favorire l'indipendenza dei dati è stato reso disponibile come modello logico in DBMS reali nel 1981. Si basa sul concetto matematico di *relazione*, il quale fornisce al modello una base teorica che permette di dimostrare formalmente proprietà di dati e operazioni.

Una relazione può essere vista come una tabella con righe, dette *tuple*, e colonne contenenti dati di tipo specificato, come ad esempio interi e stringhe. Alle colonne sono associati dei nomi, detti *nomi di attributo* la coppia nome di attributo-dominio è detta *attributo*. L'insieme degli attributi di una relazione ne costituisce lo *schema*.

Il primo passo verso la creazione di un database è raccogliere tutte le informazioni di interesse. Nel caso dell'Archivio Ricordi, i materiali da catalogare sono:

- | | |
|-----------------------|--------------------------|
| • Lettere | • Libretti e frontespizi |
| • Articoli | • Partiture |
| • Giornali | • Disposizioni sceniche |
| • Foto costumi | • Foto attrezzi |
| • Note sartoria | • Tavole attrezzi |
| • Bozzetti e figurini | • Fotografie |

Una volta raccolti i materiali che costituiranno il database, si studiano le loro caratteristiche e si stabilisce che relazioni ci sono tra essi. Questa fase è chiamata *progettazione concettuale*. Il suo scopo è quello di rappresentare le specifiche informali della realtà di interesse in termini di una descrizione formale e completa, ma indipendente dai criteri di rappresentazione utilizzati nei sistemi di gestione di basi di dati. Il prodotto di questa fase viene chiamato *schema concettuale* e fa riferimento ad un *modello concettuale* dei dati. Tali modelli ci permettono di descrivere l'organizzazione dei dati a un alto livello di astrazione, senza tenere conto degli aspetti implementativi.

Uno dei modelli più utilizzati nell'ambito della progettazione concettuale è il modello Entità-Relazione (ER). Tale modello ha una rappresentazione grafica, il diagramma ER e dei costrutti principali:

- **Entità:** rappresentano classi di oggetti che hanno proprietà comuni e esistenza autonoma (es. persona, città, ecc.). Le *istanze di entità* sono oggetti appartenenti a una certa entità (es. io , Milano, ecc.). (vedi *Figura 1*)



Figura 1

- **Relazioni o associazioni:** rappresentano legami logici, significativi per l'applicazione di interesse, tra due o più entità. Le *istanze di associazione* sono combinazioni delle istanze delle entità che prendono parte ad una associazione. (vedi *Figura 2*)

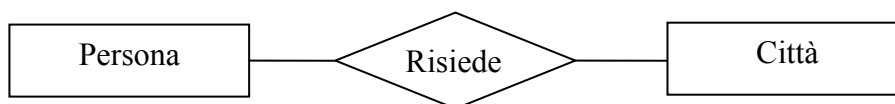


Figura 2

- **Attributi:** descrivono le proprietà elementari di entità o relazioni che sono di interesse ai fini dell'applicazione. (vedi *Figura 3*)

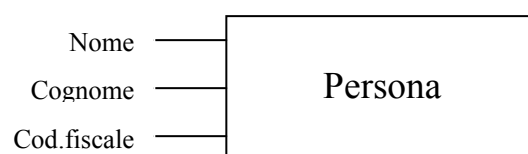


Figura 3

- **Cardinalità delle relazioni:** vengono specificate per ciascuna partecipazione di entità a una relazione e descrivono il numero minimo e massimo di occorrenze di relazione a cui una occorrenza dell'entità può partecipare. In uno schema ER, le cardinalità minima e massima si specificano tra parentesi (vedi *Figura 4*)

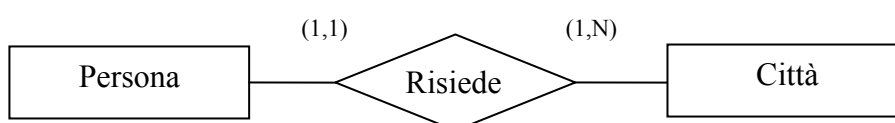


Figura 4

Nell'esempio illustrato in *Figura 4*, una persona deve avere la residenza almeno in una città e massimo in una città, nella città deve risiedere almeno una persona ma non ci sono vincoli massimi. Osservando le cardinalità massime è possibile classificare le relazioni in base al tipo di corrispondenza che viene stabilita tra le occorrenze delle entità coinvolte. Le relazioni aventi cardinalità massima pari a uno per entrambe le entità, vengono definite *relazioni uno a uno*. In maniera analoga, le relazioni aventi una entità con cardinalità massima pari a uno e l'altra pari a N sono denominate *relazioni uno a molti*. Infine le relazioni aventi cardinalità massima pari a N per entrambe le entità coinvolte vengono definite *relazioni molti a molti*.

Altri due concetti fondamentali sono:

- **La chiave:** è un insieme di attributi che distingue tra loro le tuple di una relazione. Si può dire che un insieme X di attributi di una relazione R è chiave di R se verifica che:
 - Qualsiasi sia lo stato di R, non esistono due tuple distinte di R che abbiano lo stesso valore per tutti gli attributi in X.
 - Nessun sottoinsieme proprio di X verifica la proprietà precedente.

Inoltre una chiave non può avere valori nulli e una relazione può avere più di un insieme X che verifica le proprietà precedenti. In questo caso, se il sistema non supporta più chiavi, si dovrà scegliere una chiave primaria tra le candidate. Esistono anche le *chiavi esterne*. Date due relazioni R e R' tali che R abbia un insieme di attributi X e R' abbia come chiave un insieme Y di attributi, Y è chiave esterna di R su R' se Y è un sottoinsieme di X. In altre parole, se una relazione R ha tra i suoi attributi un insieme di attributi che costituisce la chiave di una relazione R', allora tale insieme di attributi è una chiave esterna di R su R'. R' è detta *relazione riferita*. Le chiavi esterne permettono di collegare tra loro tuple di relazioni diverse e costituiscono un meccanismo, detto per valore, per modellare le associazioni tra entità. Una tupla che deve riferire un'altra tupla include tra i suoi attributi uno o più attributi il cui valore è il valore della chiave della seconda tupla.

- **L'integrità referenziale:** L'integrità referenziale rappresenta un importante vincolo di integrità semantica. Se una tupla t riferisce come valori di una chiave esterna i valori $v_1, \dots, v_n$ allora deve esistere nella relazione riferita una tupla t' con valori di chiave $v_1, \dots, v_n$.

Ai fini del database Ricordi sono state introdotte altre entità oltre ai materiali veri e propri, questo per permettere di relazionarli tra loro, per fornire maggiori informazioni all'utente e per poter soddisfare al meglio le sue possibili richieste. L'elenco delle entità è:

- Composizione
- Titolo composizione
- Lettera
- Persona
- Articolo
- Giornale
- Personaggio
- Costume
- Nota sartoria
- Bozzetto figurino
- Libretto frontespizio
- Partitura
- Allestimento
- Disposizioni sceniche
- Attrezzo
- Tavola attrezzo
- Fotografia
- Movimento
- Titolo movimento
- Ruolo
- Scansione attrezzo
- Scansione articolo
- Scansione partitura
- Scansione libretto frontespizio
- Scansione lettera
- Scansione costume

In realtà in fase di progettazione non è stato creato un vero e proprio schema ER, infatti la fase di progettazione concettuale è stata accorpata a quella di progettazione logica.

La *Progettazione logica* consiste nella traduzione dello schema concettuale nel modello di rappresentazione dei dati adottato dal sistema di gestione di base di dati a disposizione. Il prodotto di questa fase viene denominato *schema logico* e fa riferimento ad un *modello concettuale*. Tale modello ci consente di descrivere i dati secondo una rappresentazione ancora indipendente da dettagli fisici, ma concreta perché disponibile nei sistemi di gestione di base di dati.

Sono presenti due fasi principali:

- Fase di ristrutturazione
- Fase di traduzione

La *fase di ristrutturazione* consiste nell'eliminazione dallo schema ER di tutti i costrutti che non possono essere direttamente rappresentati nel modello relazionale:

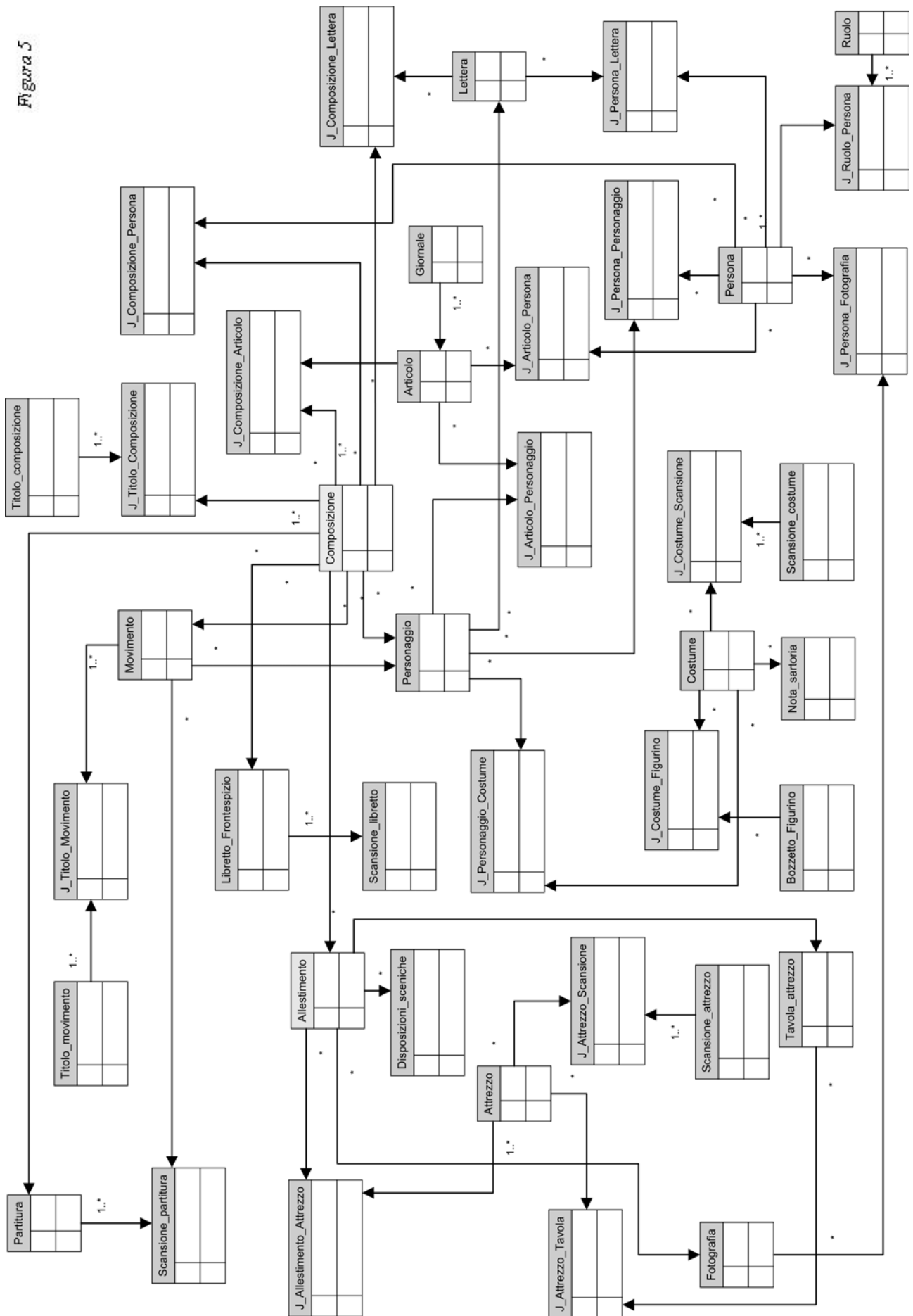
- Identificatori esterni
- Attributi composti e multivalore
- Gerarchie di generalizzazione

La *fase di traduzione* consiste nella traduzione con regole di trasformazione di entità, attributi e associazioni dello schema ER in relazioni del modello relazionale. Le regole di trasformazione sono:

- **Traduzione delle entità:** Ogni entità viene trasformata in relazione avente come attributi tutti gli attributi dell'entità corrispondente e come chiave la chiave dell'entità.
- **Traduzione relazioni/associazioni:** Si devono distinguere due casi:
 - Associazione binaria uno a uno e uno a molti. Si deve aggiungere la chiave dell'entità lato molti nell'entità lato uno. Tale chiave è detta chiave esterna.
 - Associazione binaria molti a molti e associazione n-aria. Si deve generare una nuova relazione che avrà come chiavi le chiavi delle entità.

Lo schema logico che è stato creato per il database Ricordi è rappresentato in *Figura5*. Punti di partenza dello schema sono i concetti di composizione e allestimento. Per composizione si intende la versione base dell'opera così come è stata concepita dall'autore, di conseguenza sarà indipendente dalla sua realizzazione, da eventuali traduzioni, trascrizioni, ecc. Per allestimento invece si intende una serie di rappresentazioni caratterizzate dallo stesso programma musicale, produzione, costumi, scenografia, cast e appartenenti alla stessa stagione. Di conseguenza l'entità composizione sarà legata all'entità movimento, libretto-frontespizio, personaggio, in quanto parti di questa e invariabili da rappresentazione a rappresentazione, all'allestimento dato che una composizione potrà avere zero o più allestimenti e alla partitura. E' stata creata un'entità titolo-composizione perché una composizione può avere più titoli diversi, es. traduzioni, titoli completi o non completi, con o senza articolo iniziale, ecc. Inoltre la composizione sarà relazionata alle entità lettera e articolo per il fatto che questa può essere stata citata e all'entità persona per poter identificare i vari autori, editori, ecc.

Figura 5



L'allestimento invece sarà legato a tutto ciò che caratterizza le rappresentazioni, quindi all'entità attrezzo, tavola attrezzo, disposizioni sceniche e fotografia. Nella figura si può vedere che, oltre alle tabelle relative alle entità trovate precedentemente, ne sono presenti alcune il cui nome inizia con "J", sono tabelle di join. Queste sono state create per poter trasformare le associazioni molti a molti in relazioni.

Le varie relazioni avranno poi degli attributi che ne descrivono le proprietà utili ai fini degli scopi del database. Si è deciso di utilizzare come chiavi degli identificatori univoci. Di seguito è riportato l'elenco delle relazioni con i rispettivi attributi.

Legenda

ENTITA'
Chiave interna
Chiave esterna
Attributo

RUOLO
ID ruolo

COMPOSIZIONE
ID composizione
Genere
Data pubblicazione
Paese pubblicazione
Lingua originale

TITOLO COMPOSIZIONE
ID titolo composizione
Titolo
Lingua del titolo
Titolo originale
Lingua originale

J TITOLO COMPOSIZIONE
ID composizione
ID titolo composizione

J COMPOSIZIONE LETTERA
ID composizione
ID lettera

PERSONA
ID persona
Nome
Cognome
Data nascita
Data morte

ARTICOLO
ID articolo
ID giornale
Titolo
Lingua del testo
Soggetto

J COMPOSIZIONE PERSONA
ID composizione
ID persona

J COMPOSIZIONE ARTICOLO
ID composizione
ID articolo

LETTERA
ID lettera
Numero fogli
Luogo
Data
Tecnica
Segnatura
Lingua del testo

MOVIMENTO
ID movimento
ID composizione
Lingua originale
Forma della composizione
Organico

PERSONAGGIO
ID personaggio
ID composizione
Nome

J ARTICOLO PERSONA
ID articolo
ID persona

ALLESTIMENTO
ID allestimento
ID composizione
Data
Luogo

SCANSIONE LETTERA
ID scansione lettera
ID lettera
Scansione
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore scansione

ATTREZZO
ID attrezzo
Nome
Genere
Materiale
Colore
Misure
Segnatura

GIORNALE
ID giornale
Titolo
Data
Segnatura
Paese pubblicazione

J PERSONA PERSONAGGIO
ID persona
ID personaggio
ID allestimento

J ARTICOLO PERSONAGGIO
ID articolo
ID personaggio

COSTUME
ID costume
Segnatura
Descrizione
Colore

SCANSIONE COSTUME
ID scansione costume
Scansione
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore scansione

TITOLO MOVIMENTO
ID titolo movimento
Titolo
Lingua del titolo
Titolo originale
Lingua originale

NOTA SARTORIA
ID nota sartoria
ID costume
Descrizione
Segnatura
Misure
Lingua del testo

BOZZETTO FIGURINO
ID bozzetto
Data creazione
Materiale del supporto principale
Materiale del supporto secondario
Colore
Tecniche
Misure esterne
Misure interne
Segnatura
Scansione fronte
Scansione retro
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore
Descrizione

FOTOGRAFIA
ID fotografia
ID allestimento
Data
Luogo
Descrizione
Misure
Colore
Segnatura
Dedica
Lingua dedica
Autore dedica
Scansione fronte
Scansione retro
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore scansione

J COSTUME SCANSIONE
ID costume
ID scansione costume

J COSTUME FIGURINO
ID costume
ID bozzetto

SCANSIONE LIBRETTO
ID scansione libretto
ID libretto
Scansione
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore

SCANSIONE PARTITURA
ID scansione partitura
ID partitura
Scansione
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore scansione

J PERSONA LETTERA
ID persona
ID lettera

J PERSONAGGIO COSTUME
ID personaggio
ID costume
ID allestimento

TAVOLA ATTREZZO
ID tavola attrezzo
ID allestimento
Data creazione
Materiale del supporto principale
Materiale del supporto secondario
Colore
Tecniche
Segnatura
Scansione
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore scansione

DISPOSIZIONI SCENICHE
ID disposizioni sceniche
ID allestimento
Misure
Segnatura
Scansione
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore
Lingua del testo
Lingua originale

LIBRETTO FRONTESPIZIO
ID libretto
ID composizione
Segnatura
Data rappresentazione
Luogo rappresentazione
Misure
Tecnica
Lingua del testo
Lingua originale
Data pubblicazione
Luogo pubblicazione

SCANSIONE ARTICOLO
ID scansione articolo
ID articolo
Scansione
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore scansione

J ATTREZZO TAVOLA
ID attrezzo
ID tavola attrezzo

J ALLESTIMENTO ATTREZZO
ID allestimento
ID attrezzo

J ATTREZZO SCANSIONE
ID attrezzo
ID scansione attrezzo

J PERSONA FOTOGRAFIA
ID persona
ID foto

J TITOLO/MOVIMENTO
ID movimento
ID titolo movimento

J MOVIMENTO PERSONAGGIO
ID movimento
ID personaggio

J RUOLO PERSONA
ID ruolo
ID persona

J LETTERA PERSONAGGIO
ID lettera
ID personaggio

SCANSIONE ATTREZZO
ID scansione attrezzo
Scansione
Bit per sample
Pixel per inch
Formato immagine
Algoritmo di compressione
Dimensioni immagine originale
Dimensioni file
Data scansione
Luogo scansione
Operatore scansione

PARTITURA
ID partitura
ID composizione
Tipo di partitura
Lingua del testo
Lingua originale del testo
Forma della composizione
Organico
Data creazione
Data pubblicazione
Luogo pubblicazione
Tecnica
Misure
Segnatura
Note autore
Lingua note

Ultima fase di progettazione di un database è la *progettazione fisica*. In questa fase lo schema logico viene completato con la specifica dei parametri fisici di memorizzazione dei dati. Il prodotto di questa fase viene denominato *schema fisico* e fa riferimento a un *modello fisico dei dati*. Tale modello dipende dallo specifico sistema di gestione di basi di dati scelto e si basa sui criteri di organizzazione fisica dei dati in quel sistema.

Gli aspetti importanti da considerare sono:

- Le caratteristiche fisiche di realizzazione della base di dati
- Le sue dimensioni
- Il tipo di richieste a cui dovrà rispondere

- Le risorse già disponibili

Quindi si potrà procedere con la definizione dello schema fisico dei dati e l'implementazione del database. Nel caso del progetto Ricordi è stato scelto di implementare il database con Oracle. Questa fase non è stata ancora ultimata.

3.4 Uniformità agli standard

Per permettere l'esportazione e l'integrazione di questo database ad altri archivi è necessario un formato standard di scambio. Si è scelto quindi di utilizzare:

- **Unimarc** per i record bibliografici. La scelta è stata effettuata sulla base del fatto che l'Unimarc, oltre ad essere uno dei più importanti formati di scambio di registrazioni bibliografiche, da anni ha assunto validità internazionale e molti sono gli enti che hanno già provveduto ad uniformare i loro archivi digitali a tale formato. Inoltre l'Unimarc si caratterizza per la stabilità e la garanzia di manutenzione, per l'applicabilità ad una vasta gamma di materiali e per la corrispondenza con le regole internazionali ISBD e il Dublin Core.
- **MAG** per gli oggetti digitali. In questo caso si è scelto questo standard perché permette la descrizione di metadati amministrativi e gestionali specifici per gli oggetti digitali. Inoltre è conforme agli standard internazionali in quanto permette l'uso di metadati mantenuti e definiti da altri schemi in combinazione con metadati specifici definiti per una particolare applicazione.

3.4.1 Analisi dello standard Unimarc in corrispondenza con il database Ricordi

Il modo di procedere adottato è stato quello di confrontare le entità e gli attributi, precedentemente individuati, con i campi Unimarc, cercare le possibili corrispondenze ed eventualmente aggiungere nuove proprietà che non erano state prese in considerazione, ma presenti nello standard.

Innanzitutto si deve notare che l'Unimarc prevede un elevato numero di campi, solo pochissimi dei quali sono obbligatori, questi sono:

001 Identificatore del record

100 Dati codificati generali

101 Lingua del documento

200 \$a Titolo proprio

801 Fonte del record

Di conseguenza si dovrà provvedere affinché ogni record, e quindi ogni tupla o elemento della tabella del database abbia questi campi.

In questa sezione non verranno analizzati tutti i campi e sottocampi Unimarc, ma solo quelli di interesse per la realizzazione dell'archivio digitale Ricordi.

Record label

Descrizione: Contiene le informazioni generali.

Occorrenza: E' obbligatorio all'inizio di ogni record. Non è ripetibile.

Elementi: Sono identificati dalla posizione dei caratteri che vanno da 0 a 23.

Posizioni 0-4. Lunghezza della registrazione. Di solito viene calcolato automaticamente.

Posizione 5. Stato della registrazione. E' identificato da un unico carattere:

c = record corretto;

d = record eliminato;

n = nuovo record;

o = precedentemente emesso a livello gerarchico più alto;

p = precedentemente emesso come incompleto.

Posizione 6. Tipo di registrazione:

a = materiale a stampa

b = materiale manoscritto

c = musica a stampa

d = musica manoscritta

k = immagini a due dimensioni

r = oggetti tridimensionali

Posizione 7. Livello bibliografico. Sono definiti quattro possibili valori:

a = documento analitico

m = documento monografico

s = serie

c = raccolta

Posizione 8. Codice di livello gerarchico. Indica la relazione gerarchica tra i record:

= non definita

0 = nessuna gerarchia

1 = record di livello più alto

2 = record di livello inferiore a quello più alto

Posizione 9. Non definito

Posizione 10. Lunghezza degli indicatori

Posizione 11. Lunghezza degli indicatori di sottocampo

Posizioni 12-16. Indirizzo di partenza dei dati. Indicano la posizione del primo campo di dati relativo all'inizio del record.

Posizione 17. Livello di codifica. Indica il livello di completezza del record:

= completamente codificato

1 = codificato senza esame del documento

2 = non completamente codificato

3 = non completamente codificato e non è detto che verrà completato.

Posizione 18. Tipo di catalogazione descrittiva:

= interamente conforme a ISBD

i = parzialmente conforme a ISBD

n = non conforme a ISBD

Posizione 19. Non definito

Posizioni 20-23. Mappa della directory

Corrispondenze con il database: Essendo un campo obbligatorio dovrà essere presente in tutti i record.

Blocco 0 – Identificazione

Descrizione: Dedicato alle informazioni di identificazione e codificate.

001 Identificatore del record

Descrizione: Questo campo contiene i caratteri associati unicamente al record, es. il numero di controllo dato al record.

Occorrenza: Obbligatorio, non ripetibile.

Indicatori: Conformemente ad ISO 2709 non possiede indicatori.

Sottocampi: Conformemente ad ISO 2709 non possiede sottocampi.

Corrispondenze con il database: Essendo obbligatorio per ogni record, questo campo sarà presente in ogni tabella. Si potrebbe farlo coincidere con l'identificativo (ID) in quanto univoco per ogni record.

Blocco 1 – Informazioni codificate

Descrizione: Dedicato alle informazioni di identificazione e codificate.

100 Dati codificati generali

Descrizione: Il campo contiene dati codificati di lunghezza fissa.

Occorrenza: Obbligatorio, non ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito).

Sottocampi:

\$a Dati codificati generali. Non ripetibile. Tutti i dati in questo campo sono identificati dalla posizione dei caratteri. Convenzionalmente le posizioni sono numerate da 0 a 35 e sono:

\$a/0-7 Data di registrazione. (obbligatorio) Costituito da 8 caratteri numerici. La data viene scritta nella forma YYYYMMDD dove YYYY rappresenta l'anno, MM il mese e DD il giorno.

\$a/22-24 Lingua della catalogazione. (obbligatorio) Costituito da un codice di 3 caratteri. La lista dei codici è contenuta nell'Appendice A del manuale Unimarc.

\$a/26-29 Set di caratteri. (obbligatorio) Dei codici identificano i principali set di caratteri usati. La lista dei codici è contenuta nell'Appendice J del manuale Unimarc.

Non sono stati descritte tutte le posizioni in quanto non utili ai fini del database. Nelle posizioni vuote verrà usato il carattere “|”.

Corrispondenze con il database: Essendo obbligatorio dovrà essere presente in ogni record. Non essendoci attributi relazionabili a questo campo, si dovrà provvedere all'inserimento di tre nuovi attributi, Data di registrazione, Lingua di catalogazione e Set di caratteri per ogni entità del database.

101 Lingua del documento

Descrizione: Il campo contiene informazioni relative alla lingua del documento, alle sue parti, al titolo e alla lingua originale se esso è una traduzione.

Occorrenza: Obbligatorio se l'opera ha una lingua. Non ripetibile.

Indicatori: Indicatore 1: Indicatore di traduzione. Specifica se il documento è o no una traduzione o se ne contiene una. 0 Il documento è in lingua originale, 1 Il documento è una traduzione dell'opera originale, 2 Il documento contiene traduzioni
Indicatore 2 bianco (non definito)

Sottocampi:

\$a Lingua del testo. Ripetibile quando il testo è in più lingue.

\$c Lingua originale dell'opera. Ripetibile quando il testo è in più lingue.

\$h Lingua del libretto, ecc. Lingua del testo quando il documento include il contenuto vocale/testuale dell'opera stampata come testo. Ripetibile.

Corrispondenze con il database: Non tutti i record del database prevedono una lingua, perciò le entità a cui può essere applicato questo campo sono: Composizione, Lettera, Articolo, Nota sartoria, Libretto frontespizio, Partitura, Disposizioni sceniche, Foto, Movimento.

102 Paese di pubblicazione o produzione

Descrizione: Il campo contiene i codici di identificazione del paese di pubblicazione o produzione del documento.

Occorrenza: Opzionale. Non ripetibile

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito)

Sottocampi:

\$a Paese di pubblicazione. Contiene un codice di identificazione del paese in cui il documento è stato prodotto o pubblicato. La lista dei codici è presente nell'Appendice B del manuale Unimarc.

\$b Località di pubblicazione. Contiene un codice di identificazione della località in cui il documento è stato prodotto o pubblicato. E' possibile usare un codice nazionale o internazionale.

Corrispondenze con il database: Può essere utile utilizzare questo campo per le entità: Composizione, Giornale, Libretto frontespizio, Partitura.

106 Dati relativi il contenuto: materiale testuale - caratteristiche fisiche

Descrizione: Il campo contiene dati relativi alle caratteristiche fisiche del materiale testuale.

Occorrenza: Opzionale. Non ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito)

Sottocampi:

\$a Dati codificati per il materiale testuale: mezzo fisico

d = a caratteri grandi

e = formato giornalistico

g = micro stampa

h = manoscritto

j = a caratteri piccoli

r = a caratteri regolari

z = altro

Corrispondenze con il database: Questo campo può risultare utile soprattutto per distinguere i testi manoscritti da quelli stampati. Le entità interessate sono: Lettera, Articolo, Nota sartoria, Libretto frontespizio, Disposizioni sceniche.

116 Dati relativi il contenuto: arte grafica, immagini

Descrizione: Il campo contiene i dati relativi a immagini (es. stampe o foto)

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito)

Sottocampi:

\$a Dati per l'arte grafica. Non ripetibile. Tutti i dati in questo campo sono identificati dalla posizione dei caratteri. Convenzionalmente le posizioni sono numerate da 0 a 17 e sono:

\$a/0 Materiale specifico. Un codice di un carattere è usato per identificare la classe relativa al materiale cui il documento appartiene. I codici usati sono:

a = collage
b = disegno
c = pittura
d = riproduzione fotomeccanica
e = fotonegativo
f = riproduzione fotografica
h = immagine
i = stampa
k = disegno tecnico
z = altro

\$a/1 Materiale di supporto principale. Un codice di un carattere è usato per identificare la classe relativa al materiale di supporto principale del documento. I codici usati sono:

a = tela	j = gesso
b = cartoncino	k = legno compresso
c = cartone	l = porcellana
d = vetro	m = pietra
e = sintetico	n = legno
f = pellami	u = non conosciuto
g = tessuti	v = collana mista
h = metallo	z = altro
i = carta	

\$a/2 Materiale di supporto secondario. Un codice di un carattere è usato per identificare la classe relativa al materiale di supporto secondario del documento. I codici usati sono gli stessi del materiale del supporto principale con un codice in più: y = nessun supporto secondario.

\$a/3 Colore. Un codice di un carattere è usato per identificare il colore. I codici usati sono:

a = un colore, monocromatico
b = bianco e nero
c = multicolore
d = colorato a mano
u = non conosciuto
v = misto
x = non applicabile
z = altro

\$a/4-9 Tecniche (disegni). Un codice di due caratteri è usato per specificare la tecnica in cui il disegno è stato eseguito. Ci sono tre posizioni per questo codice, quelle che non vengono utilizzate devono essere lasciate bianche (##). I codici sono:

aa = matita	bb = macchia
ab = grafite	bc = crayon
ac = matita colorata	bd = seppia
ad = inchiostro di china	be = inchiostro
ae = inchiostro di china lavierung	bf = caseina
af = carboncino	bg = doratura
ag = gessetto	bh = encaustico
ah = gessetto nero	bi = acrilico
ai = sanguigna	bj = collage
aj = acquarello	bk = punta d'argento
ak = tempera	bl = pennello
al = guazzo	uu = non conosciuto
am = pastello	vv = misto
an = olio	xx = non applicabile
ba = pennarello	zz = altro

\$a/10-15 Tecniche (stampe). Un codice di due caratteri è usato per specificare la tecnica in cui il disegno è stato eseguito. Ci sono tre posizioni per questo codice, quelle che non vengono utilizzate devono essere lasciate bianche (##). I codici sono:

ba = xilografia	be = eliotipia
bb = xilografia a chiaroscuro	bf = cromolitografia
bc = xilografia a linea bianca	bg = incisione su linoleum
bd = camaiu	bh = incisione all'acquaforte

bi = litografia	cf = mezzatinta
bj = fotolitografia	cg = monotipo
bk = zincografia	ch = silkscreen
bl = algraphy	ci = steel engraving
bm = acquatinta	cj = grafica al computer
bn = reservage	ck = fotocopia
ca = vernis-mou	uu = non conosciuto
cb = incisione	vv = misto
cc = incisione a crayon	xx = non applicabile
cd = incisione a bulino	zz = altro
ce = puntasecca	

\$a/16-17 Funzione. Nel nostro caso queste posizioni verranno riempite con i caratteri “xx” in quanto il codice non è applicabile al nostro database.

Corrispondenze con il database: Nell’archivio Ricordi sono presenti molte immagini, per questo motivo questo campo è molto utile, in particolare per le entità: Bozzetto figurino, Disposizioni sceniche, Tavola attrezzo, Fotografia.

125 Dati relativi al contenuto: registrazioni sonore e musica a stampa

Descrizione: Il campo descrive il formato della stampa musicale.

Occorrenza: Opzionale. Non ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito)

Sottocampi:

\$a Formato della musica a stampa. Non ripetibile. Formato da due caratteri:

\$a/0 Tipo di partitura. Identificato da un codice di un carattere:

- a = partitura completa
- b = partitura da studio
- c = spartito
- d = partitura per voci, per coro, senza accompagnamento
- e = riduzione pianistica

g = a parti strette (es. inni)

u = sconosciuto

x = non applicabile

z = altro

\$a/1 Indicatore delle parti. Un codice da un carattere indica l'esistenza delle parti di una partitura:

a = esistono le parti

u = sconosciuto

x = non applicabile

y = parti non presenti

Corrispondenze con il database: Il campo dovrà essere utilizzato per le partiture.

128 Dati relativi al contenuto: esecuzioni musicali e partiture

Descrizione. Il campo descrive la forma della composizione e fornisce i dettagli relativi a strumenti e voci.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito).

Sottocampi:

\$a Forma della composizione. Un codice di due caratteri identifica la forma della composizione (es. balletti, overture, concerti), la lista completa dei codici è presente sul manuale Unimarc. Il sottocampo è ripetibile.

\$b Strumenti e voci per ensembles

\$c Strumenti e voci per solisti. Questi due sottocampi usano gli stessi codici. Ogni sottocampo è costituito da due o quattro caratteri ed è ripetibile. I primi due caratteri indicano il tipo di strumento o voce, gli altri due specificano il numero degli strumenti o voci. Se il numero è minore di dieci, la terza posizione conterrà uno 0. La lista degli strumenti con i rispettivi codici è presente sul manuale Unimarc.

Corrispondenze con il database: Questo campo può essere utilizzato per le partiture.

Blocco 2 – Informazioni descrittive

Descrizione: Dedicato alla descrizione, suddiviso secondo le aree ISBD.

200 Titolo e formulazione di responsabilità

Descrizione: Il campo contiene il titolo e altre informazioni riguardo questo e le formulazioni di responsabilità. Corrisponde alle aree ISBD Titolo e Formulazione di responsabilità.

Occorrenza: Obbligatorio. Non ripetibile.

Indicatori: Indicatore 1: Indicatore di titolo significativo. Indica se il titolo viene usato come accesso al documento: 0 Titolo non significativo, 1 Titolo significativo

Indicatore 2 : bianco (non definito)

Sottocampi:

\$a Titolo proprio. Obbligatorio per tutti i record

\$b Indicazione generale del materiale

\$c Titolo proprio di altro autore

\$d Titolo parallelo

\$e Complemento del titolo

\$f Prima indicazione di responsabilità

\$g Successive indicazioni di responsabilità

\$h Numero di una parte

\$i Nome di una parte

\$z Lingua del titolo parallelo

Corrispondenze con il database: Essendo un campo obbligatorio dovrà essere applicato a tutti i record del database.

205 Edizione

Descrizione: Il campo contiene le informazioni riguardanti le edizioni e corrisponde all'area Edizione di ISBD.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito)

Sottocampi:

\$a Formulazione dell'edizione

\$b Formulazione aggiuntiva di edizione

\$d Formulazione parallela di edizione

\$f Formulazione di responsabilità relativa all'edizione

\$g Seconda e successive formulazioni di responsabilità relative all'edizione

Corrispondenze con il database: Questo campo verrà utilizzato per tutte le entità che possono essere edite: Giornale, Libretto frontespizio, Partitura.

208 Materiale specifico: musica a stampa

Descrizione: Il campo contiene un'esposizione relativa al formato della musica a stampa o manoscritta. Il campo corrisponde all'Area Specifica Musica a Stampa dell'ISBD(PM).

Occorrenza: Opzionale. Non ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito)

Sottocampi:

\$a Esposizione relativa alla musica a stampa

\$d Esposizione parallela

Corrispondenze con il database: Il campo è applicabile alle partiture

210 Pubblicazione, distribuzione, ecc.

Descrizione: Il campo contiene le informazioni relative alla pubblicazione, distribuzione, ecc. del documento. Corrisponde all'area Pubblicazione, Distribuzione, ecc. di ISBD.

Occorrenza: Opzionale. Non ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito).

Sottocampi:

\$a Luogo di pubblicazione, distribuzione, ecc.

\$b Indirizzo dell'editore, distributore, ecc.

\$c Nome dell'editore, distributore, ecc.

\$d Data di pubblicazione, distribuzione, ecc.

\$e Luogo di stampa

\$f Indirizzo del tipografo

\$g Nome del tipografo

\$h Data di stampa

Corrispondenze con il database: Questo campo verrà utilizzato per quelle entità che possono essere pubblicate distribuite e stampate: Giornale, Libretto frontespizio, Partitura.

Blocco 3 – Note

Descrizione: Dedicato alle note, suddivise sulla base della funzione e della tipologia

300 Note generali

Descrizione: Il campo contiene note riguardanti ogni aspetto dell'unità bibliografica.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito).

Sottocampi:

\$a Testo della nota. Non ripetibile.

Corrispondenze con il database: Questo campo può essere applicato a tutte le entità del database in quanto tratta di note di carattere generale.

303 Note generali riguardanti le informazioni descrittive

Descrizione: Note riferite alla descrizione del documento.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito)

Sottocampi:

\$a Testo della nota. Non ripetibile

Corrispondenze con il database: Questo campo può essere applicato a tutte le entità del database in quanto tratta di note di carattere generale.

307 Note riguardanti la descrizione fisica

Descrizione: Il campo contiene note riguardanti la descrizione fisica del documento.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito).

Sottocampi:

\$a Testo della nota. Non ripetibile

Corrispondenze con il database: Questo campo può essere applicato a tutte quelle entità che hanno struttura fisica: Lettera, Articolo, Giornale, Costume, Nota sartoria, Bozzetto figurino, Libretto frontespizio, Partitura, Disposizioni sceniche, Tavola attrezzo, Fotografia.

Blocco 5 – Titoli in relazione

Descrizione: Dedicato alle varie forme di titoli.

510 Titolo parallelo

Descrizione: Il campo contiene un titolo parallelo (es. il titolo in una lingua diversa) per il quale si vuole creare una nota o un punto di accesso.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: Indicatore di titolo significativo. Specifica se il titolo parallelo può essere usato come punto di accesso al documento: 0 Titolo parallelo non significativo, 1 Titolo parallelo significativo.

Indicatore 2: bianco (non definito)

Sottocampi:

\$a Titolo parallelo. Il titolo principale del documento in una lingua e/o in una scrittura riferito al titolo in \$200. Non ripetibile.

\$b Altre informazioni sul titolo. Sottotitoli e altre informazioni subordinate al titolo parallelo. Ripetibile.

\$z Lingua del titolo. Identifica la lingua del titolo parallelo. Non ripetibile.

Corrispondenze con il database: Applicabile alle tabelle Titolo composizione e Articolo.

Blocco 6 – Analisi semantica

Descrizione: Dedicato all'analisi semantica (soggetti, classificazioni, termini di thesaurus)

600 Nome di persona usato come soggetto

Descrizione: Il campo contiene il nome di una persona che fa parte dei soggetti del documento.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: Indicatore di forma del nome. Questo indicatore specifica se il nome è inserito tramite nome di battesimo o tramite cognome: 0 Nome inserito tramite nome di battesimo, 1 Nome inserito tramite cognome.

Sottocampi:

\$a Elemento principale. Porzione di nome usata come elemento per l'ordinamento. Non ripetibile

\$b Altra parte del nome. Il resto del nome.

\$c Aggiunte al nome diverse dalle date. Ogni aggiunta al nome (diversa dalle date) che non è parte integrante del nome (es. titoli). Ripetibile.

\$f Date. Date unite al nome. Non ripetibile.

Corrispondenze con il database: Applicabile alle entità Lettera e Articolo.

604 Nome e titolo usato come soggetto

Descrizione: Questo campo contiene autore e titolo di un'opera che fa parte dei soggetti del documento.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito).

Sottocampi:

\$l Dati di legame. E' seguito dal campo 500 Titolo uniforme incapsulato e dal campo 7— Responsabilità intellettuale incapsulato.

Corrispondenze con il database: Anche questo campo può essere usato per le lettere e gli articoli.

605 Titolo usato come oggetto

Descrizione: Il campo contiene un titolo che è parte dei soggetti di un documento.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: bianco (non definito)

Sottocampi:

\$a Elemento principale. Il titolo. Non ripetibile

\$h Numero di sezione o parte. Numero di una parte del documento cui il titolo si riferisce. Ripetibile

\$i Nome della sezione o parte. Il nome della parte. Ripetibile

\$q Versione (o data della versione). Identifica la versione dell'opera rappresentata dal documento. Non ripetibile

\$r Mezzo di esecuzione (per musica). La strumentazione, ecc. del documento. Ripetibile

\$s Designazione numerica (per musica). Un numero assegnato dal compositore o altri per distinguere l'opera. Ripetibile

\$u Chiave (per musica). La chiave musicale usata come parte del titolo. Non ripetibile

\$w Arrangiamento (per musica). Non ripetibile

Corrispondenze con il database: Questo campo è applicabile alle tabelle Articolo, Lettera, Partitura, Libretto.

610 Termini di soggetto non controllati

Descrizione: Il campo è usato per i termini che non derivano da alcuna lista controllata di soggetti.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Indicatore 1: Livello del termine di soggetto. E' usato per distinguere i descrittori principali da quelli secondari. Un termine è considerato principale se riguarda il soggetto principale del documento. Un termine che riguarda invece gli aspetti meno importanti è detto secondario. 0 Livello non specificato, 1 Termine principale, 2 Termine secondario.

Indicatore 2: bianco (non definito).

Sottocampi:

\$a Termine di soggetto. Ripetibile

Corrispondenze con il database: Dato che questo campo non presenta particolari vincoli possiamo applicarlo, quando necessario, a tutte le tabelle del database.

Blocco 7 – Responsabilità intellettuale

Descrizione: Dedicato alla responsabilità intellettuale

700 Autore personale responsabilità principale

Descrizione: Il campo contiene il nome della persona considerata autore dell'opera in forma di punto di accesso.

Occorrenza: Opzionale. Non ripetibile.

Indicatori: Indicatore 1: bianco (non definito). Indicatore 2: Indicatore della forma del nome. Specifica se il nome è inserito tramite nome di battesimo o cognome. 0 Nome inserito tramite nome di battesimo, 1 Nome inserito tramite cognome

Sottocampi:

\$a Elemento principale. Porzione di nome usata come elemento per l'ordinamento. Non ripetibile

\$b Altra parte del nome. Il resto del nome.

\$c Aggiunte al nome diverse dalle date. Ogni aggiunta al nome (diversa dalle date) che non è parte integrante del nome (es. titoli). Ripetibile.

\$f Date. Date unite al nome. Non ripetibile

Corrispondenze con il database: E' possibile applicare il campo alla tabella Persona.

702 Autore personale – responsabilità secondaria

Descrizione: Il campo contiene il nome di una persona avente responsabilità secondaria nella creazione di un'opera.

Occorrenza: Opzionale. Ripetibile.

Indicatori: Uguali a quelli del campo 700.

Sottocampi: Uguali a quelli del campo 700.

Corrispondenze con il database: E' possibile applicare il campo alla tabella Persona.

Blocco 8 – Dati internazionali

Descrizione: Dedicato alla fonte bibliografica.

Corrispondenze con il database: Non sono presenti corrispondenze.

Blocco 9 – Informazioni ad uso locale

Descrizione: Non definito dallo standard e utilizzabile per inserire dati locali.

Corrispondenze con il database: E' possibile utilizzare questi campi per l'inserimento di quelle informazioni che non rientrano negli altri blocchi ma che sono utili ai fini del database.

3.4.2 Analisi dello standard MAG in corrispondenza con il database Ricordi

Come già accennato, si è scelto di utilizzare questo standard per i metadati relativi agli oggetti digitali. Nel database Ricordi, gli oggetti digitali presenti sono immagini in quanto scansioni di partiture, foto, lettere, articoli, ecc., di conseguenza i metadati utili saranno i medesimi per tutti i documenti.

In questa sezione non verranno analizzati tutti gli elementi MAG ma solo quelli di interesse per la realizzazione dell'archivio digitale Ricordi.

Elemento <metadigit>

Descrizione: Elemento radice di ciascun record MAG (obbligatorio, non ripetibile).

Tipo: xsd:sequence

Attributi: Version (xsd:string) – dato che la versione che si sta usando è la 2.0, questo sarà un valore di default.

Componenti: <gen>, <bib>,

Commento: La scelta delle componenti è stata effettuata sulla base del fatto che <gen> e <bib> sono obbligatorie, è opzionale ma è la sezione contenete i metadati per le immagini.

Sezione <gen>

Descrizione: Contiene le informazioni generali sul progetto e sul tipo di digitalizzazione effettuata (obbligatorio, non ripetibile).

Tipo: xsd:sequence

Attributi:

Creation (xsd:datetime, opzionale) – data creazione del record

Last_update (xsd:datetime, opzionale) – data ultimo aggiornamento

Componenti: <stprog>, <agency>, <access_rights>, <completeness>

<stprog>

Descrizione: Riferimento a una risorsa di rete che descrive gli standard di progetto (obbligatorio, non ripetibile)

Tipo: xsd:anyURI

<agency>

Descrizione: Ente responsabile della digitalizzazione (obbligatorio, non ripetibile).

Tipo: xsd:string

<access_rights>

Descrizione: Riguarda le condizioni di accesso all'oggetto analogico originale (obbligatorio, non ripetibile).

Tipo: xsd:enumeration

Valori:

0 (uso riservato all'interno dell'istituzione)

1 (uso pubblico)

<completeness>

Descrizione: Livello di completamento della digitalizzazione (obbligatorio, non ripetibile)

Tipo: xsd:enumeration

Valori:

0 (digitalizzazione completa)

1 (digitalizzazione parziale)

Sezione <bib>

Descrizione: Contiene i metadati descrittivi che identificano la risorsa analogica (obbligatorio, non ripetibile).

Tipo: xsd:sequence

Attributi:

Level

a = spoglio

m = monografia

s = seriale

c = raccolta prodotta dall'istituzione

Componenti: I tag Dublin Core

<dc:identifier>

Descrizione: Identificatore univoco alla risorsa analogica (obbligatorio, ripetibile)

Tipo: dc:elementType

Commento: Dato che per i documenti analogici si è deciso di utilizzare lo standard Unimarc, questa sezione è pressoché inutile, perciò è stato preso in considerazione l'unico elemento obbligatorio.

**Sezione **

Descrizione: Contiene i dati relativi alle immagini (opzionale, ripetibile)

Tipo: xsd:sequence

Componenti: <sequence_number>, <nomenclature>, <file>, <md5>, <filesize>, <image_dimension>, <image_metrics>, <format>, <scanning>, <datetimecreated>, <note>

<sequence_number>

Descrizione: Numero progressivo che identifica un file multimediale all'interno del file XML (opzionale, non ripetibile).

Tipo: xsd:positiveInteger

<nomenclature>

Descrizione: Descrizione di un'immagine (opzionale, non ripetibile)

Tipo: xsd:string

<file>

Descrizione: Link al file multimediale (obbligatorio, non ripetibile).

Tipo: link

<md5>

Descrizione: Impronta del file, serve a monitorare l'integrità del file (obbligatorio, non ripetibile)

Tipo: xsd:restriction of xsd:string (length = 32)

<filesize>

Descrizione: Dimensione del file in bytes (opzionale, non ripetibile)

Tipo: xsd:positiveInteger

<image_dimension>

Descrizione: Dimensioni dell'immagine (obbligatorio, non ripetibile)

Tipo: xsd:sequence

Componenti: <niso:imagelength>, <niso:imagewidth>

<niso:imagelength>

Descrizione: Dimensione verticale dell'immagine in pixel (obbligatorio, non ripetibile)

Tipo: xsd:positiveInteger

<niso:imagewidth>

Descrizione: Dimensione orizzontale dell'immagine in pixel (obbligatorio, non ripetibile)

Tipo: xsd:positiveInteger

<image_metrics>

Descrizione: Caratteristiche tecniche principali dell'immagine, secondo lo standard NISO (obbligatorio se applicabile, non ripetibile).

Tipo: xsd:sequence

Componenti: <niso:samplingfrequencyunit>, <niso:samplingfrequencyplane>,
<niso:xsamplingfrequency>, <niso:ysamplingfrequency>, <niso:bitpersample>,
<niso:photometricinterpretation>

<niso:samplingfrequencyunit>

Descrizione: Unità di misura del campionamento (obbligatorio, non ripetibile)

Tipo: xsd:enumeration

Valori:

1 = nessuna unità di misura,

2 = inch, pollice,

3 = centimetro

<niso:samplingfrequencyplane>

Descrizione: Piano focale del campionamento (obbligatorio, non ripetibile)

Tipo: xsd:enumeration

Valori:

1 = camera/scanner focal plane (quando non sono definite le dimensioni dell'oggetto che si sta digitalizzando)

2 = object plane (quando l'oggetto e la riproduzione hanno la stessa dimensione)

3 = source object plane (quando la dimensione della riproduzione è maggiore dell'oggetto originale).

<niso:xsamplingfrequency>

Descrizione: Frequenza di campionamento nella direzione orizzontale (obbligatorio se applicabile, non ripetibile)

Tipo: xsd:positiveInteger

<niso:ysamplingfrequency>

Descrizione: Frequenza di campionamento nella direzione verticale (obbligatorio se applicabile, non ripetibile)

Tipo: xsd:positiveInteger

<niso:bitpersample>

Descrizione: Numero di bit per ciascun campione, profondità-colore (obbligatorio, non ripetibile)

Tipo: xsd:enumeration

Valori:

1 = bitonale, bianco e nero

4 = 4-bit scala di grigi

8 = 8-bit scala di grigi o gamma di 256 colori

8,8,8 = 24-bit RGB

16,16,16 = 48-bit TIFF, HDR

8,8,8,8 = 32-bit CMYK

<niso:photometricinterpretation>

Descrizione: Interpretazione fotometrica dei bit campione (obbligatorio, non ripetibile)

Tipo: xsd:enumeration

Valori:

WhiteIsZero

BlackIsZero

RGB

Palette color

Transparency Mask

CMYK

YcbCr

CIELab

<format>

Descrizione: Tipologia dell'immagine e modalità di compressione (obbligatorio, non ripetibile)

Tipo: xsd:sequence

Componenti: <niso:mime>, <niso:compression>

<niso:mime>

Descrizione: Tipologia MIME del file (obbligatorio, non ripetibile)

Tipo: xsd:enumeration

Valori:

image/jpeg

image/tiff

image/gif

image/png

image/vnd.djvu

application/pdf

<niso:compression>

Descrizione: Algoritmo di compressione (obbligatorio, non ripetibile)

Tipo: xsd:enumeration

Valori:

Uncompressed

CCITT 1D

CCITT Group 3

CCITT Group 4

LZW

JPG

PNG

DJVU

<scanning>

Descrizione: Modalità di svolgimento della scansione (opzionale, non ripetibile)

Tipo: xsd:sequence

Componenti: <niso:scanningagency>

<niso:scanningagency>

Descrizione: Persona, società o ente produttore dell'immagine (opzionale, non ripetibile)

Tipo: xsd:string

<datetimecreated>

Descrizione: Data e ora di creazione del file (opzionale, non ripetibile)

Tipo: xsd:datetime

<note>

Descrizione: Qualsiasi tipo di annotazione all'immagine (opzionale, non ripetibile)

Tipo: xsd:string

Commento: Questa sezione è quella che interessa le caratteristiche delle scansioni, di conseguenza è quella fondamentale per il database in progetto. Ciò nonostante non sono stati

utilizzati tutti gli elementi che il MAG mette a disposizione questo perché non tutti sono risultati di interesse o applicabili agli oggetti digitali presenti nell'archivio.

3.5 Sviluppi futuri del progetto Ricordi

Come visto, i materiali costituenti il database per l'Archivio Ricordi sono immagini e descrizioni bibliografiche. Questa potrebbe essere però una situazione temporanea in quanto, per una maggiore completezza, il database potrebbe essere integrato con materiali audio. In questo modo si permetterebbe all'utente anche l'ascolto di un'opera, o parte di questa, oltre che la sua visualizzazione in partitura. La realizzazione di tutto ciò implica però un ulteriore lavoro ricerca e studio riguardante innanzitutto la scelta del materiale. In questa rientrano i problemi relativi ad esempio alla scelta di quale realizzazione dell'opera mettere a disposizione, se metterne a disposizione più d'una, se tutta o in parte, ecc. Altro aspetto rilevante è la messa a punto di un set di metadati per la gestione degli oggetti audio in modo da garantire da un lato, dei protocolli che ben si adattino alle necessità del singolo ente, dall'altro fare in modo che questo possa essere condiviso su scala nazionale e internazionale applicando gli standard internazionali. Inoltre questi metadati dovranno relazionarsi con quelli bibliografici, gestire la segmentazione del documento sonoro per consentire, ad esempio, l'indicizzazione del documento digitale in termini di brani e parti all'interno di ciascun brano, gestire le relazioni tra le vari oggetti digitali presenti nell'archivio e gestire i dati sul copyright.

Un'altra possibile integrazione del database potrebbero essere le ricerche per contenuto. Infatti, per il momento, l'accesso ai materiali contenuti nel database è mediato da una ricerca di tipo catalografico. Per permettere le ricerche per contenuto è necessario un sistema in grado riconoscere quando un frammento musicale è presente in un brano, o se un frammento è simile a una parte di una composizione (ovvero “suona” simile al frammento cercato). Un utente dovrebbe cioè riuscire a rintracciare un documento archiviato, attraverso un frammento melodico che potrà essere audio pre-registrato, fischiettato attraverso un microfono, suonato con uno strumento MIDI, rappresentato tramite un gruppo di note posizionate su un pentagramma attraverso un'interfaccia grafica, una stringa di caratteri ASCII o con altri formati simbolici.

4. Conclusioni

Il progetto *La musica in rete. Digitalizzazione dell'Archivio Storico Ricordi* come pure il progetto ADMV rientrano nella generale convergenza di archivi, biblioteche e musei che, da alcuni anni, è divenuta parte essenziale delle politiche pubbliche. Oltre a questa convergenza ce ne è anche una di tipo tecnologico, cioè relativa agli strumenti che consentono di produrre e allo stesso tempo di accedere in rete ad una varietà enorme di informazioni. Ciò a cui si vuole arrivare è un unico sistema che raccolga materiali eterogenei per forma, contenuti e provenienza e che li metta a disposizione, attraverso la rete, di un pubblico sempre più vasto e variegato. Quello che prima era inevitabilmente distinto o fruibile separatamente, dovrà essere fruibile simultaneamente e in modo integrato.

Nello *Studio di fattibilità per la realizzazione della biblioteca digitale* sono state prese in esame le esperienze internazionali in atto, e dal loro prevedibile sviluppo si sono tratte considerazioni di carattere biblioteconomico, politico ed organizzativo, le quali sono state poi rapportate alla situazione italiana. Da questo studio sono state ricavate le seguenti considerazioni:

- Anche in Italia, come nel resto del mondo, stanno sviluppandosi progetti digitali ma si ha la sensazione che, ferma restando la validità tecnologica e culturale di molti di questi progetti ciascuno di essi sia partito per impulso proprio, senza alcuna attenzione a iscrivere l'iniziativa in un quadro più ampio di progettazione coordinata e condivisa.
- La maggioranza dei progetti ha come promotore e attore principale una singola struttura generando il legittimo dubbio che, non solo non si sia rivolta attenzione ad analoghe iniziative in campo nazionale, ma neppure a livello internazionale.
- E' molto difficile conoscere la misura e la durata degli investimenti per tali progetti.

Ciò contrasta con l'imperativo della cooperazione e della condivisione delle esperienze, che sembra rappresentare il "leitmotiv" del dibattito internazionale a proposito del digitale.

La proposta che emerge da tale *Studio* non è solo una soluzione tecnico-informativa che, abbastanza semplicisticamente, raccolga in un unico grande "magazzino" il digitale prodotto in Italia, ma che sia anche orientata al coordinamento, al supporto e alla gestione del digitale. Nella situazione italiana saranno soprattutto le nuove iniziative ad avvalersi di tali funzioni di servizio. In

tal modo le singole iniziative, costruite secondo criteri standard e di qualità, potranno confluire in quella realtà che possiamo chiamare fin da ora la Biblioteca Digitale Italiana (BDI).

Il progetto BDI e il successivo scenario di integrazione di sistemi definiscono una strategia nazionale del settore. Il successo di tale strategia dipenderà dalla capacità di affrontare tre ordini di problemi:

- condivisione degli obiettivi;
- partecipazione dei servizi bibliotecari e degli istituti culturali;
- definizione di un nuovo ambiente che sia in grado di far cooperare fra loro istituti diversi (archivi, biblioteche, musei e altre agenzie coinvolte nel digitale), culture e linguaggi diversi senza sacrificare le differenze.

In prospettiva, e sulla base delle esperienze già in atto, si dovranno determinare gli elementi informativi irrinunciabili per la descrizione, affiancati da quelli utili per la comunicazione all'esterno, inserendoli in formati di scambio che permettano la cumulabilità e il consolidamento delle informazioni. Le nostre istituzioni dovranno sviluppare gli strumenti e standard necessari non solo per lo scambio delle risorse informative, ma anche per facilitare l'accesso integrato dell'utenza a tali risorse attraverso una disseminazione selettiva dell'informazione. L'accesso integrato implica peraltro una pianificazione concettuale d'alto livello che permetta all'utente di ottenere tutti i documenti pertinenti ad un determinato soggetto, indipendentemente dal loro supporto, da chi lo possiede e dalla loro localizzazione geografica. La progettazione di una politica della comunicazione del patrimonio, considerato parte viva delle memorie collettive ed individuali, nonché elemento di crescita e sviluppo sostenibile, non può limitarsi all'accumulo delle realizzazioni spontanee ed alla frammentazione delle risorse. L'interoperabilità nell'area d'azione comune richiede la definizione condivisa e partecipata dei progetti. Dato che l'informazione culturale è molto ricca e complessa, gli standard di contenuto (per esempio lessici e thesauri, bibliografie, liste d'autorità, descrizione scientifica di materiale d'archivio e d'opere) sono altrettanto importanti degli standard tecnologici, ma per evitare duplicazioni e sovrapposizioni d'azioni è necessario concertare regole uniformi per la catalogazione e formati di scambio per la comunicazione, come suggerito nel presente lavoro.

Bibliografia

AA.VV. *Studio di fattibilità per la realizzazione della biblioteca digitale – Sezione prima – La situazione attuale*, www.iccu.sbn.it/BDI-SDF.pdf

AA.VV. Atti del convegno internazionale *Futuro delle memorie digitali e patrimonio culturale*, ottobre 2003, www.iccu.sbn.it/PDF/atti_firenze_2003.pdf

AA.VV., *Linee di indirizzo per i progetti di digitalizzazione del materiale fotografico*, 2004, http://www.iccu.sbn.it/PDF/Linee_guida_fotografie.pdf

Archivio Digitale della Musica Veneta, <http://marciana.venezia.sbn.it/admv.htm>

Atzeni Paolo, Ceri Stefano, Paraboschi Stefano, Torlone Riccardo, *Basi di dati – Modelli e linguaggi di interrogazione*, McGraw-Hill, Milano, 2002

Associazione italiana biblioteche, homepage <http://www.aib.it>

Biblioteca Digitale Italiana – Progetti avviati <http://www.librari.beniculturali.it/genera.jsp?s=31>

Bergamin Giovanni, *Introduzione ai metadati*, <http://www.unicam.it/ateneo/organizzazione/biblioteche/documenti/Bergamin.pdf>

Bertocci Emilio, *La catalogazione descrittiva*, <http://www.provincia.sp.it/Biblioteche/All.3.PDF>

Biblioteca Nazionale Braidense, <http://www.braidense.it/ricordi.html>

Buttà Basilio, *Metadati negli Open Archive : tecniche di conversione dei formati dei record nel formato MARC21 utilizzato da CDSware*, 2004, <http://eprints.rclis.org/archive/00003022>

California Digital Library, *Digital Object Standard: Metadata, Content and Encoding*, May 18, 2001

Codd E. F., *A Relational Model of Data for Large Shared Data Banks*, Reprinted from *Communications of the ACM*, Vol. 13, No. 6, June 1970

De Robbio Antonella, *Seminario Angela Vinay. bibliotECONOMIA: la frontiera digitale - Vademecum on-the-fly per progetti di digitalizzazione*, aprile 2003, <http://eprints.rclis.org/archive/00000137/02/Vinay2003.pdf>

De Robbio Antonella, *Seminario nazionale: Metadati per la comunicazione scientifica*, aprile 2001, www.iccu.sbn.it/PDF/Derobbio.pdf

De Vries Henk J., *Standardization: a business approach to the role of National Standardization Organizations*, Boston-Dordrecht-London, Kluwer Academic Publishers, 1999

Dublin Core metadata element set, versione 1.1 / traduzione italiana [a cura di] Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche (ICCU), <http://www.iccu.sbn.it/dublinco.html>

Dublin Core Metadata Initiative (DCMI) home page, <http://dublincore.org>

Falchetta Piero, *Guida breve alla digitalizzazione in biblioteca*, www.bibliotecheoggi.it/2000/20000905201.pdf.

Fondazione Italiana Accenture, <http://www.fondazioneaccenture.it/dam.html>

Format for Information Interchange. Geneva, ISO, 1996 (ISO 2709-1996)

Gambari Stefano e Guerrini Mauro, *Definire e catalogare le risorse elettroniche*, Editrice Bibliografica, Milano, 2002

Gilmour Ron, *XML: a guide for librarians*, American Library Association, Chicago, 2003

Haus Goffredo e Ludovico Luca A., *The Digital Opera House: an Architecture for Multimedia Databases*, Journal of Cultural Heritage, 2005.

International Federation of Library Associations and Institutions, homepage <http://www.ifla.org/>

ISBD(M), a cura dell'International Federation of Library Associations and Institutions. Edizione italiana a cura di Rossella Dini. Roma: Associazione Italiana Biblioteche, 1991

Istituto Centrale per il Catalogo Unico delle Biblioteche Italiane e per le Informazioni Bibliografiche – ICCU, homepage <http://www.iccu.sbn.it>

Joint Information System Committee, JISC – <http://www.jisc.ac.uk/>

Ludovico Luca A., *Applicazione DAM centrale - Manuale utente*

MAG Schema, <http://www.bncf.firenze.sbn.it/progetti/mag>

Mariani Maria Pia, *Seminario nazionale sui metadati - Verso l'interoperabilità tra sistemi: Biblioteche, Musei e Archivi. Il Gruppo di studio sugli standard e le applicazioni di metadati nei beni culturali - Intervento introduttivo*, www.iccu.sbn.it/PDF/Mariani.pdf.

Mazzola Merola Giovanna, *Seminario nazionale sui metadati - Verso l'interoperabilità tra sistemi: Biblioteche, Musei e Archivi. Il Gruppo di studio sugli standard e le applicazioni di metadati nei beni culturali - Intervento introduttivo*, www.iccu.sbn.it/PDF/Merola.pdf.

Miller Paul, *I am a name and a number*, "Ariadne Issue", 24, 21 June 2000

Peruch Stefano, *La Scala DAM*, http://www.fondazioneaccenture.it/pdf/Report05/Report_peruch.pdf

Petrucchi Armando, *La descrizione del manoscritto. Storia, problemi, modelli*, seconda ed. corretta e aggiornata, Roma, Carocci, 2001

Ramez A. Elmasri, Shamkant B. Navathe, *Sistemi di basi di dati : fondamenti*, Pearson Education Italia, Milano, 2004

Saltarelli Alberto, Tammaro Anna Maria, *La biblioteca digitale*, Editrice Bibliografica, Milano, 2000

Scolari Antonio, *Unimarc*, Roma: Associazione Italiana Biblioteche, 2000

Solmine Giovanni, *Introduzione allo studio della biblioteconomia. Riflessioni e documenti*, Manziana, Vecchierelli, 1995

Subrahmanian S., *Principles of multimedia database systems*, San Francisco, Calif., Morgan Kaufmann Publishers, 1998

Tangari Nicola, *Standard e documenti musicali*, Editrice Bibliografica, Milano, 2002

UK Interoperability Focus, <http://www.ukoln.ac.uk/interopfocus/about/>

UNIMARC manual: authorities format / International Federation of Library Associations and Institutions, IFLA Universal Bibliographic Control and International MARC Programme. IFLA 2001

Weston Paul Gabriele, *Gruppo di studio sui metadati descrittivi - Presentazione della tavola rotonda*, www.iccu.sbn.it/PDF/Weston.pdf

Ringraziamenti

Dato che questo lavoro conclude i miei tre anni (quasi quattro) di università, volevo ringraziare tutti coloro che mi sono stati accanto in questo periodo e che mi hanno aiutato ad arrivare a questo punto.

Un ringraziamento speciale a:

- Professor Haus, relatore di questa tesi, per l'aiuto, la disponibilità, l'incoraggiamento, nonché per aver ideato STCM.
- Adriano e Luca, correlatori della tesi, per l'aiuto e la pazienza.
- La mia famiglia, per avermi dato l'opportunità di compiere questi studi universitari, in particolare il papà, correttore ufficiale di bozze.
- Mauro, Albi, Vivi per essermi stati vicini ed avermi sopportato.
- Il LIM, Dante, la Marisa, Maddalena Novati, per avermi dato l'opportunità di nuove esperienze, per gli insegnamenti e l'amicizia.
- I vari professori, per avermi insegnato un sacco di cose interessanti e per la loro disponibilità.
- Tutti i compagni di STCM, in particolar modo: la Cri, Cecilia, Federico, Gabriele, Gummo, Gygus, Jeff, Juke, King, Luca, Marianna, Massimo, Renato, Smirne, la Sara, Savo, Walla.
- Diego, il Gigante, Neil e gli altri amici informatici e comunicatori digitali.
- Tutti gli altri.