

METAMAG

1 IL PRODOTTO

Metamag è un prodotto che permette l'acquisizione, l'importazione, l'analisi e la catalogazione di oggetti digitali per materiale documentale (quali immagini oppure file di testo (OCR)), e contestualmente consente la creazione ed il trattamento dei relativi metadati.

Con le risorse digitali infatti i metadati gestionali assumono una sempre più ampia rilevanza: la vita stessa della risorsa digitale dipende da questi metadati.

La prima istanza è riferita a risorse disponibili nell'ambito bibliografico, ma può estendersi al trattamento di quanto realizzato in archivi, musei, istituzioni culturali, ecc.

L'obiettivo è di costituire un valido strumento per:

- supportare le numerose campagne di digitalizzazione in atto
- garantire l'omogeneità descrittiva indispensabile all'integrazione delle informazioni sul materiale digitale nei servizi attualmente in essere
- favorire l'evoluzione dei sistemi della cultura verso la valorizzazione dei prodotti culturali.

2 IL CONTESTO

I riferimenti prioritari di MetaMAG sono gli standard descrittivi, nazionali e internazionali, di scambio dati e di servizio affermati nell'ambito dei beni culturali e di seguito elencati:

- MAG (per la descrizione standardizzata dei metadati MAG: Metadati Amministrativi Gestionali, in formato XML)¹
- Z39.50 (per l'interoperabilità in fase di ricerca sul catalogo bibliografico)
- OAI: Open Archive Initiative (per la raccolta dei metadati da parte di un livello centrale).

La scelta degli standard consente

- una omogeneità descrittiva dei documenti digitali
- di ricostituire la visibilità dei cataloghi bibliografici di riferimento a prescindere dalle scelte architettoniche e di standard operate nei singoli sistemi.

Questa ultima caratteristica costituisce l'elemento caratterizzante del prodotto e ne determina la possibilità di distribuzione in vari contesti catalografici per favorire sia la derivazione automatica degli elementi descrittivi che la integrazione con i servizi di informazione bibliografica locale e centrale.

3 FUNZIONALITÀ PREVISTE

La gestione dei metadati è consentita attraverso una unica interfaccia amichevole che consente attualmente di:

1. Gestire ed attivare funzioni di amministrazione dell'applicazione e controllo accessi alle varie funzionalità previste
2. Configurare l'accesso al catalogo Bibliografico di riferimento e le modalità di cattura (derivazione) automatica dei dati da riportare sulla sezione BIB del DB Metadati e sviluppare funzionalità di interrogazione Client verso il DB Catalogo per

¹ Prodotto nell'ambito del Gruppo di studio sugli standard e sulle applicazioni di metadati nei beni culturali promosso dall'ICCU, <http://www.bncf.firenze.sbn.it/progetti/mag/index.html>

- l'individuazione della scheda di catalogo e la memorizzazione dei dati catturati: è possibile ricercare sul catalogo sia tramite bid che tramite altri punti di accesso come Autore, Titolo, Soggetto e selezionare la notizia da cui recuperare i parametri BIB
3. Attribuire gli elementi comuni del fondo oggetto di digitalizzazione per la gestione dell'anagrafe digitale (sezione GEN)
 4. Gestire i Metadati tramite emissione di schermi per il completo trattamento dei metadati in formato MAG con pre-impostazione di dati relativi a set di oggetti digitali (ad es. immagini)
 5. Attivare funzioni aggregate per la fruizione degli oggetti digitali (ad es. visualizzazione) e delle relative informazioni
 6. Interrogare il Database dei Metadati con possibilità di richiamo alle funzioni di Gestione Metadati (modifica, cancellazione, inserimento)²
 7. Esportare i file XML del Database Metadati in formato MAG-SCHEMA

4 CARATTERISTICHE DELLA SOLUZIONE

La soluzione è 'web-oriented' e 'server-oriented'. In questo modo l'utente non deve installare niente sul proprio PC ma utilizzare un Web Browser. Tramite un indirizzo (Intranet) accederà alle funzioni messe a disposizione del server tramite pagine HTML.

La tecnologia utilizzata è Java in termini sia di pagine JSP che di servlet Java Beans residenti sul server. La realizzazione in Java permette la sua portabilità in ambienti diversi come Windows, Macintosh, Linux, Unix.

Il software realizzato è basato su librerie disponibili con licenza Open Source.

Sul server risiede anche il database che consente la tenuta dei metadati MAG e di quanto è necessario per una corretta acquisizione e archiviazione dei dati digitali e metadati nell'ambiente custom.

La figura seguente indica l'architettura logica del prodotto

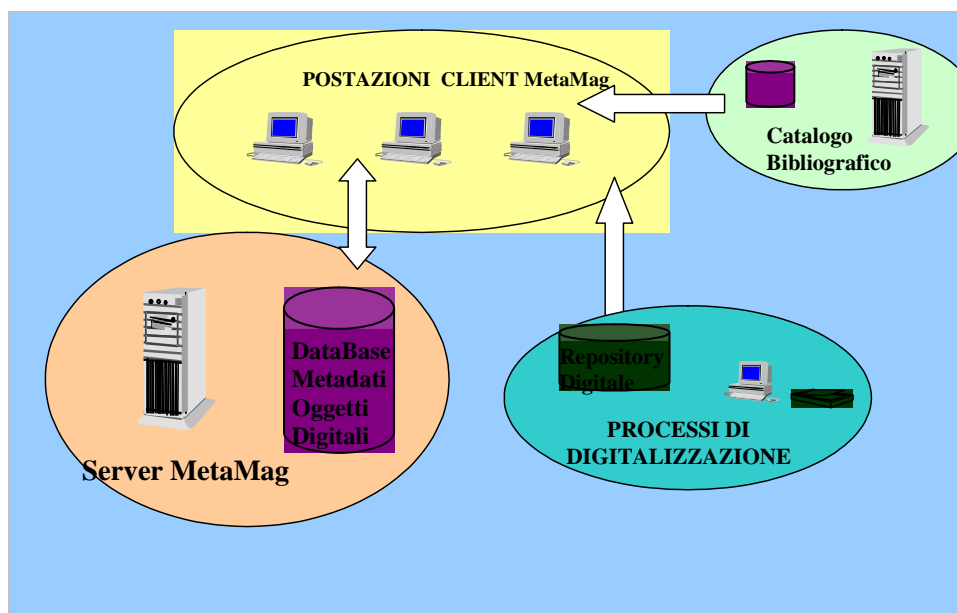


Figura A Architettura logica di riferimento del prodotto

L'architettura di riferimento è costituita da un ambiente tipicamente Intranet (ad es. una LAN) con:

² I metadati di tipo descrittivo (cioè i Dublin Core identificati con BIB nella struttura MAG) possono non essere memorizzati nel database dei MAG perché comunque presenti nel catalogo bibliografico da cui sono derivati.

- Postazioni Client su P.C. caratterizzate da:
 - Accesso agli oggetti digitali da trattare tramite Web Browser
- un Server di LAN contenente il web server (Apache Tomcat), l'application server (Tomcat), il software MetaMAG e il Database dei Metadati (RDBMS MySQL) con disponibile:
 - Connessione JDBC con il catalogo locale dei Metadati per la loro tenuta (Interrogazione, inserimento, variazione, Export XML)
 - Connessione al catalogo bibliografico tramite TCP/IP in modalità Z39.50³ per la derivazione dei metadati BIB a fronte di ricerche attivate dal Client.

5 IL FLUSSO DI LAVORO

La configurazione prevista è indicata nello schema seguente; sono previste due fasi principali di elaborazione:

- *Acquisizione*, il cui output è la produzione di oggetti digitali e/o l'analisi tecnica nel caso di oggetti pre-esistenti, accompagnati da file XML
- *Archiviazione*, che prevede la memorizzazione dei dati XML su database, l'archiviazione degli oggetti digitali sul repository, l'aggiornamento dell'OPAC.

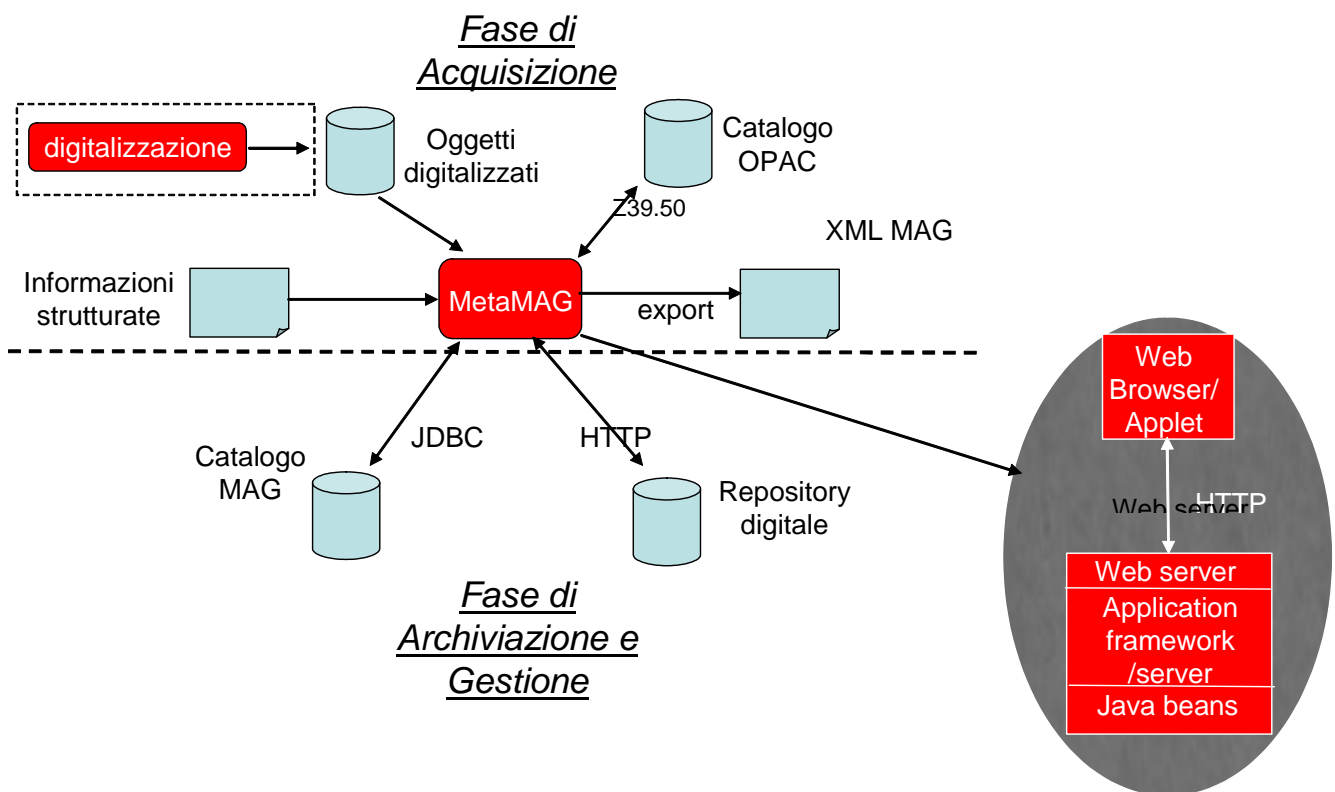


Figura B Configurazione del sistema MetaMAG

5.1 ACQUISIZIONE

La fase di acquisizione è composta da più sessioni di lavoro: nell'ambito di un singolo progetto di scansione (corrispondente ad una sezione GEN dei MAG) ogni sessione di lavoro è relativa ad una

³ Sono supportate le sintassi di trasferimento XML e SUTRS con alcune limitazioni ed eventuali adeguamenti da concordare a richiesta

opera diversa, cioè ad un nuovo BIB. Ogni sessione produce su una diversa cartella del server un file XML completo di tutti i componenti più le immagini associate.

Alla fine della fase di acquisizione sul server ci saranno tanti file XML quante sono state le sessioni, ognuno accompagnato da più oggetti digitali relativi.

5.1.1 Digitalizzazione

Il processo di digitalizzazione precede la fase di acquisizione (nella figura precedente è messa in un riquadro tratteggiato).

5.1.1.1 Digitalizzazione precedente all'acquisizione

Il processo di digitalizzazione precede (o ha preceduto in un periodo antecedente) la fase di acquisizione: l'utente effettuerà un 'upload' dei file digitali verso il server che esaminerà il contenuto dei file estraendone le informazioni. Inoltre tramite l'accesso al catalogo e il completamento dei dati da parte dell'utente sarà disponibile il file XML MAG e la memorizzazione nella banca dati.

5.2 ARCHIVIAZIONE

La fase di archiviazione comprende le seguenti fasi:

1. caricamento dei file XML nel database o catalogo metadati MAG e relative funzioni di tenuta
2. upload degli oggetti digitali nel repository su LAN (in genere un sistema NAS, Network Attached Storage): il trasferimento dei file digitalizzati su NAS può essere fatto in modalità batch; altre funzioni relative sono: fruizione, spostamento, rinomino, cancellazione di oggetti digitali nel repository digitale
3. aggiornamento del catalogo (OPAC) con l'informazione che consente il recupero degli oggetti digitali associati alla notizia bibliografica (ad es. identificativo della risorsa digitale ovvero tag UNIMARC 956).

Solo la prima fase è attualmente coperta dal prodotto MetaMAG. Le altre anche per la loro estrema variabilità e configurabilità, verranno gestite come 'add-on' al prodotto valutabili singolarmente.

6 INFORMAZIONI ED INTERFACCE

Le informazioni gestite da MetaMAG sono sia relative alla struttura definita dall'XML MAG Schema che locali e di supporto al progetto, per quanto concerne l'acquisizione, l'archiviazione e la gestione, e su queste informazioni e la loro struttura è possibile effettuare delle 'customizzazioni', in relazione alle esigenze del cliente e/o del progetto. Le informazioni locali sono mantenute, insieme a quelle presenti nell'XML dello schema MAG, nella base dati di MetaMAG.

L'interfaccia Web integrata di MetaMAG consente contestualmente di effettuare:

- sessioni di lavoro per operazioni di scansione, esame dei file digitali, inserimento informazioni, creazione dei file XML e immissione nella banca dati

- configurazione ed amministrazione, da utilizzare sia per impostare i parametri che devono rimanere fissi nell'ambito delle sessioni di lavoro, che per inserire/modificare/cancellare i dati presenti sulla banca dati.