

DL: nuove dimensioni informative e nuovi linguaggi

Daniela Canali

**Interoperabilità delle biblioteche digitali
Costituzione, diffusione e condivisione
delle risorse informative**

Università Roma Tre

Roma, 9-10 ottobre 2006

Argomenti

- Il cammino verso le biblioteche digitali
- Come cambiano i circuiti della comunicazione
 - Dinamicità delle interconnessioni
- Evoluzione organizzativa: e-resources e flussi di lavoro
- Salto del paradigma
 - Sistemi di comunicazione primaria
 - Sistemi di comunicazione secondaria
 - Archiviazione sostenibile



Il cammino verso le biblioteche digitali

Digital Library

Una o più collezioni di informazioni in formato elettronico gestite tramite l'uso di tecnologie digitali per l'acquisizione, la memorizzazione, la conservazione e l'accesso, **con relativi servizi associati** per comunità di utenti che la utilizzano tramite supporto tecnologico

Digital Library - fasi

prima generazione (a partire dalla fine degli anni 60)

- era del *record bibliografico* e dei *cataloghi online*, con la nascita del formato MARC e delle varie iniziative di digitalizzazione nazionali ed internazionali;

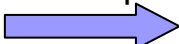
seconda generazione (dai primi anni 90)

- epoca della *sperimentazione*, fortemente influenzata dalla massiccia diffusione di Internet in ambiente statunitense, specie presso la Cornell University e la Library of Congress. In questa fase la biblioteca utilizza la rete come un mezzo aggiuntivo per espletare i servizi online tramite l'utilizzo di modalità di interconnessione (Telnet, per esempio). I cambiamenti non sono limitati al catalogo: è in questo periodo infatti che inizia la digitalizzazione delle raccolte e la conservazione digitale;

terza generazione (dalla metà degli anni 90)

- fase del *consolidamento* e dell'*istituzionalizzazione* via Internet tramite l'ampia disponibilità di web opac. Dalla metà alla fine degli anni 90 si moltiplicano la crescita e l'estensione dei web opac che offrono accesso all'intera gamma di collezioni della biblioteca, oltre a rendere fruibili le collezioni di altre istituzioni, a distanza in modalità remota;

generazione attuale

- enfaticizzazione della *prospettiva funzionale* della biblioteca digitale come componente di un **sistema esteso di servizi** - mediazione tra risorse informative diverse e distribuite e la comunità di utenti in continua evoluzione - **ambiente di servizio consistente** in uno spazio informativo online, in cui gli utenti possono cercare, localizzare, accedere e soprattutto **utilizzare** l'informazione  profilazione utenza e personalizzazione

Quali cambiamenti?

Politici

Organizzativi

Finanziari

Tecnologici

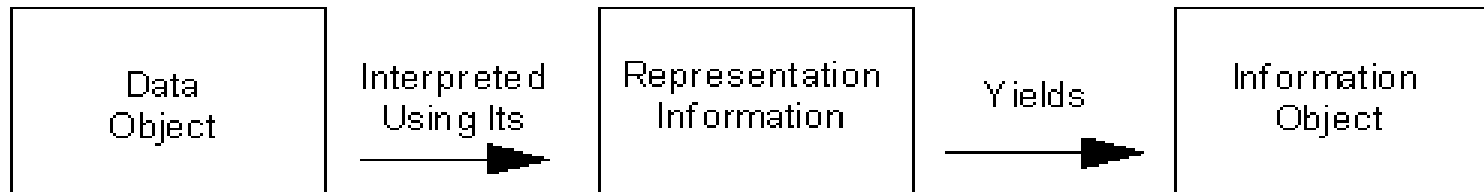
Risorse umane

Culturali

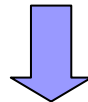


Change Management
Information Policy

Rappresentazione del processo di costruzione di un *data object*



information object + **rappresentazione formale dell'informazione** = *data object*



In ambiente digitale consiste di una sequenza di bits combinata con tutti gli altri dati necessari a rendere il flusso di bits comprensibile

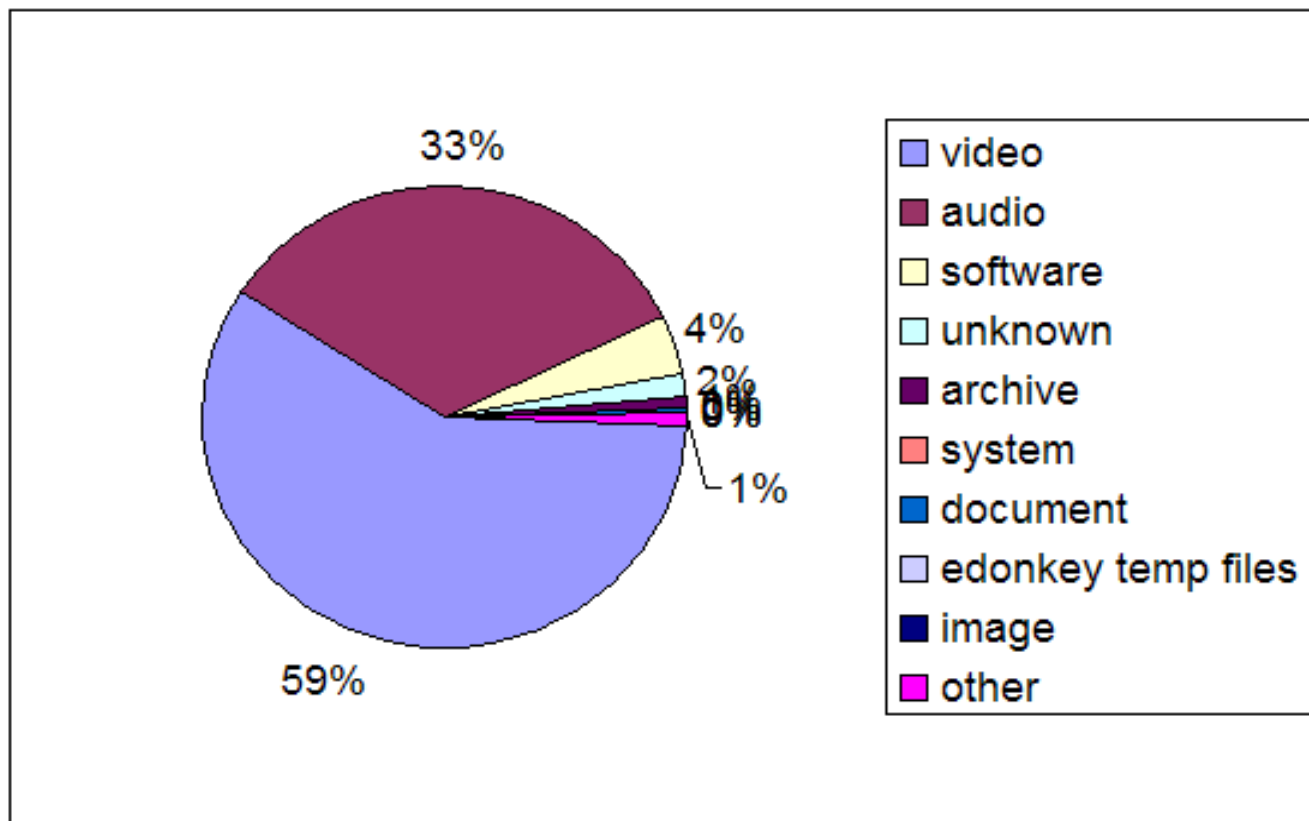
Digital objects

- I documenti di una Digital Library non sono le sole descrizioni dei documenti, in analogia con i dati catalografici, ma i sono **documenti compositi** (testi, video, suoni, immagini, dati, programmi, mappe, ...) che costituiscono un **unico insieme** in cui si colloca **ogni scheda e relativo materiale**, ovvero un **insieme che associa catalogo e collezione**
- Un oggetto digitale può contenere altri materiali informativi ed è composto da **entità**, la **minima componente logica cui l'utente ha accesso**
- Ciascuna descrizione può riferirsi all'oggetto digitale o ad una sua entità: nella complessità e nella eterogeneità dei formati, la distinzione fra varie tipologie di documenti tende quindi a svanire

Digital objects

- Leggibilità → *formati*
- Riproducibilità → *modifiche*
- Contenuto dinamico → *versioning*
- Affidabilità → *qualità*

Source: How much information 2003



Evoluzione della metadescrizione

- Nuove concezioni di documenti solo digitali → **digital born**
- Nuovi approcci metadescrittivi per il trattamento orientato alla raccolta, gestione, rilascio tecnologico e diffusione

Nuovi tipi di risorse digitali e trattamento dell'informazione a valore aggiunto

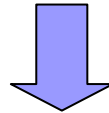
Oltre gli aspetti di creazione e di gestione di appositi strumenti di consultazione (standard descrittivi, qualitativi, rete sindetica)



Necessità di specifici
modelli di riferimento della meta-informazione per le risorse digitali

Tali modelli permettono di **costruire formati di interscambio** indispensabili al funzionamento avanzato di applicazioni tecnologiche orientate alla rete e in grado di rispondere alle **nuove necessità gestionali e di fruizione** determinate dalla natura eterogenea degli oggetti digitali e dalla **dinamicità delle interconnessioni**

I **nuovi modelli multi livello** tengono conto delle numerose interdipendenze indispensabili nella gestione dei documenti digitali: sebbene la fase di inserimento e quella di fruizione si presentino sempre più interconnesse i **meta elementi descrittivi si moltiplicano** e assumono **pesi e ruoli diversi nei vari momenti del ciclo di vita dell'oggetto digitale**



Ruolo delle Mappature

- componente progettuale della fattibilità
- base per uno schema di funzionamento nelle applicazioni tecnologiche avanzate
- Tre funzionalità logiche:
 - gestione dell'oggetto (DC)
 - valorizzazione del contenuto
 - personalizzazione per la comunità di riferimento



Come cambiano i circuiti della comunicazione

Circuiti della comunicazione

Trasferimento della conoscenza

■ Circuiti commerciali

- Autore

- Distribuzione

 - Editore

 - tradizionale/cartaceo: distributore / libraio

 - digitale: provider / aggregatore / vendor

- **Circuiti alternativi**

 - open access

 - disseminazione

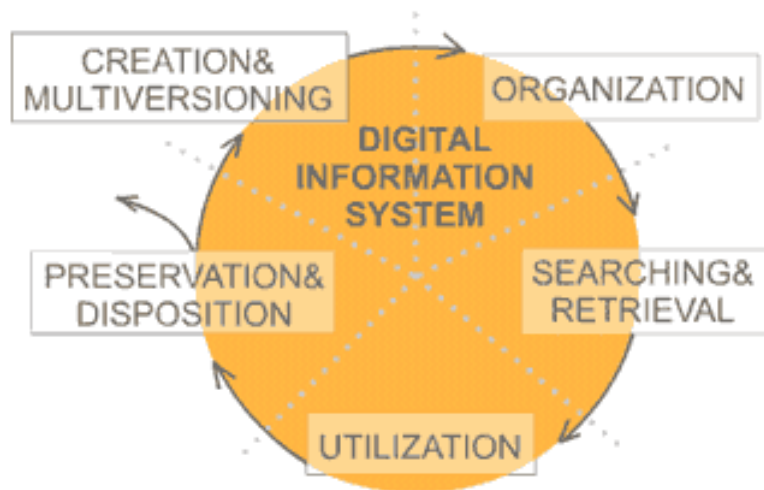
 - depositi istituzionali

 - ...



Dinamicità delle interconnessioni

Ciclo di vita di un *information object* in un sistema informativo digitale



Gli oggetti entrano in un **sistema informativo** in **forma digitale** nativa o attraverso la conversione da altro formato. Versioni diverse dello stesso oggetto possono essere utilizzate per la **conservazione, la ricerca, la disseminazione** o per altri sviluppi (*Creation and multi-versioning*). **L'autore può inserire in questo caso alcuni metadati amministrativi**. Gli oggetti vengono **organizzati manualmente o automaticamente** nella struttura del sistema informativo digitale e vengono creati **metadati aggiuntivi** tramite la **registrazione, la catalogazione ed i processi di indicizzazione** (*Organization*) finalizzati alla **ricerca** e al **recupero** da parte degli utenti. Il sistema crea metadati relativi agli **algoritmi** di recupero, alle transazioni degli utenti e all'efficacia (*Searching and retrieval*). Una volta recuperati, gli oggetti vengono **utilizzati, riprodotti, modificati**: per monitorare tali processi sono utili **metadati** relativi alle annotazioni degli utenti, ai diritti e al controllo della versione (*Utilization*). Infine gli oggetti subiscono **processi di refreshing, migrazione e controlli di integrità** che ne assicurano la disponibilità continua: i **metadati** documentano sia la **conservazione** sia la **disponibilità** (*Preservation and disposition*).



Evoluzione organizzativa

digital objects e flussi di
lavoro

Flussi...

- **Selezione**
- **Acquisizione**
- **Implementazione e integrazione**
- **Promozione e servizio utente**
- **Manutenzione**
 - monitoraggio risorsa: contenuto e licenza
 - controllo statistico
 - gestione problemi di accesso (rete interna e fornitori)
 - SW/HW locale, consorzio, fornitore/ aggregatore
 - gestione problemi utenti (locali e remoti)
 - gestione contratti con consorzio e /o fornitore
 - ...
- **Conservazione ...**



Conservazione *mission* delle biblioteche...

...anche delle biblioteche
digitali?

Conservazione digitale

Questione non solo tecnologica

- continuità/refreshing
- affidabilità /autenticità
- compatibilità legali (diritto di autore)
- sostenibilità economica

- **Gruppo di lavoro per la conservazione delle memorie digitali** a livello europeo – obiettivi
 - ricognizione stato dell'arte in materia di conservazione digitale
 - sviluppo di un piano d'intervento necessario a diffondere e dar seguito ai principi contenuti nella Risoluzione del Consiglio d'Europa del 25 giugno 2002 – Preservare la memoria del futuro, preservare i contenuti digitali per le generazioni future (2002/C 162/02)
- Referenti principali del WG i progetti ERPANET e MINERVA, sotto la guida della Commissione Europea



Salto del paradigma

Sistemi di gestione di risorse elettroniche

- Sistemi che funzionano da depositi
 - le risorse risiedono fisicamente (*trusted digital repositories*) con forte componente conservativa (DSpace, E-prints)
- Sistemi senza finalità di conservazione
 - indirizzano semplicemente alla risorsa informativa



Comunicazione primaria

trusted digital repositories

self-archiving
...intermediazione...

■ E-LIS Eprints in Library and Information Science

E-LIS is an open access archive for scientific or technical documents, published or unpublished, on Librarianship, Information Science and Technology, and related areas. E-LIS relies on the voluntary work of individuals from a wide range of backgrounds and is non-commercial. It is not a funded project of an organization. It is community-owned and community-driven. We serve LIS researchers by facilitating their self-archiving, ensuring the long-term preservation of their documents and by providing word-wide easy access to their papers.

It is a free-access international archive.

- [Submission demonstration](#)

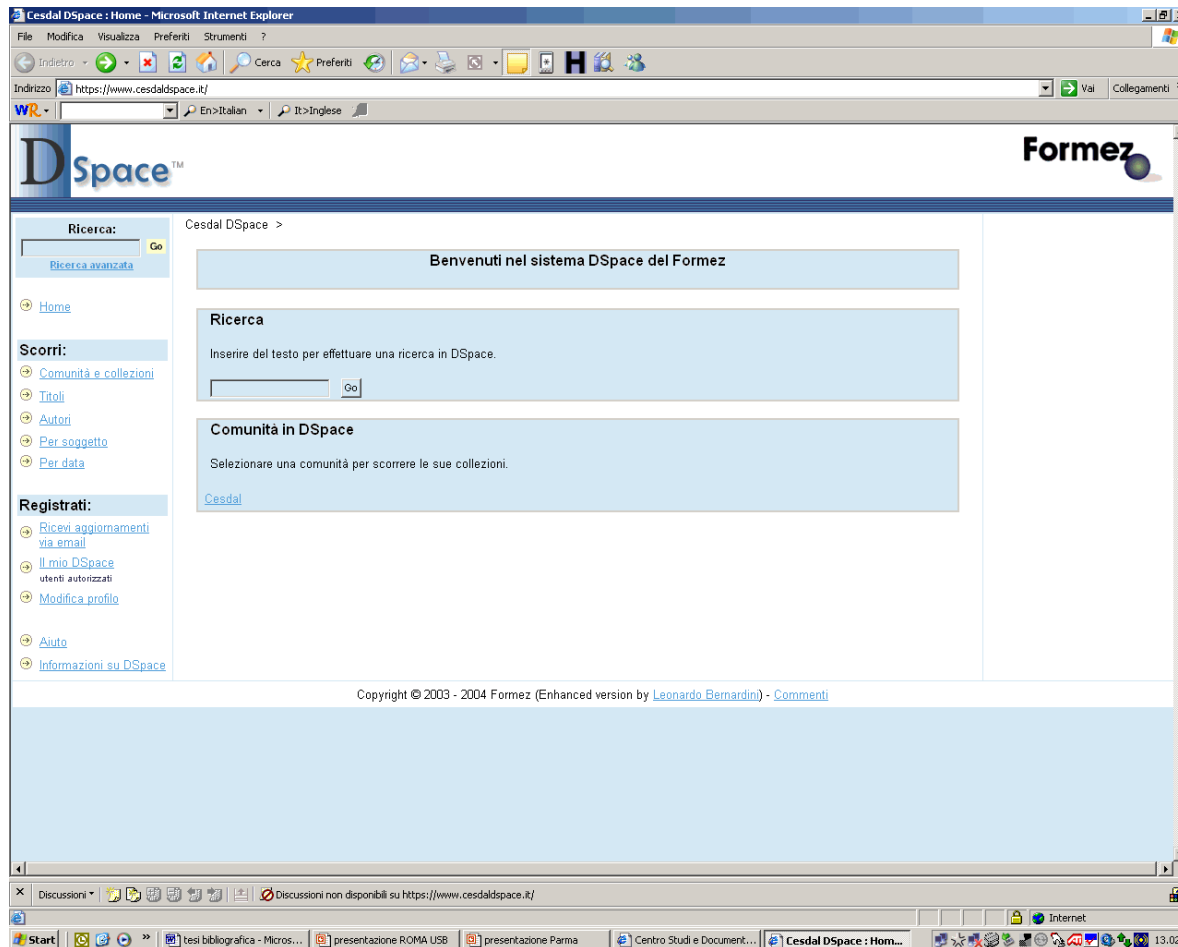
- Operativo dal 2003
- Attualmente ha **4282** documenti
- E-LIS è basato su protocollo e standard OAI
 - metadati in formato standard OAI con uscita in Dublin Core
 - raccolti da alcuni harvester in servizi internazionali come OAIster
 - progetto OAIster12 creato dal servizio di produzione della biblioteca digitale dell'Università del Michigan in linea con l'iniziativa Metadata Harvesting della Mellon Foundation. L'obiettivo è rendere ricercabili un gran numero di *repositories* liberamente accessibili e utili alla ricerca


DSpace

■ DSpace

- progetto congiunto del MIT e della HP per la realizzazione di un *digital repository* per catturare, distribuire e conservare la produzione intellettuale in formato digitale (rapporti tecnici e di lavoro, atti di conferenze, immagini, ecc.).
- piattaforma software *open source* che consente di:
 - catturare e descrivere lavori digitali usando un semplice modulo di sottomissione (*workflow*)
 - distribuire sul web i prodotti digitali di un'istituzione attraverso un sistema di *search and retrieval*
 - conservare a lungo termine gli oggetti digitali

Cesdal/DSpace



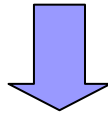


Comunicazione secondaria

Sistemi di segnalazione delle risorse

Esperienze di ricerca - Procedura

- Tramite *template weblog* si genera l'inizializzazione di una scheda di metadati successivamente integrata da un operatore umano



gestione dell'informazione

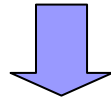
non semplicemente *appesa* come un *post-it* nel mondo di un determinato weblog ma viva, circolante, che genera conoscenza

- La tecnologia di pubblicazione web è utilizzata per l'invio del **messaggio di notifica o di segnalazione di esistenza di una nuova risorsa informativa**
- Procedure di completamento e di validazione a garanzia della qualità dei dati

- Sviluppo in ambiente XML di *template* weblog per la segnalazione dell'esistenza di una risorsa ed il caricamento dei dati minimi da parte dell'utente (acquisizione intermedia dei dati):
 - Una versione relativa al *messaggio generico* dedicata alla mera funzione comunicativa del blog in cui l'utente è invitato ad inserire dati minimi personali ed il messaggio che intende inviare al sistema
 - Una versione arricchita di metadati di inizializzazione più spinta

- Comunque necessario **controllo** e **convalida** da parte di un amministratore del sistema a livello server, a garanzia della **qualità dei dati** e della loro trasformazione da **raw material** (dati semplici) in **rich material** (dati a valore aggiunto)

Il ricercatore produce la risorsa informativa



la consegna a chi si occupa di metterla in rete e contestualmente procede alla segnalazione della localizzazione

senza dover effettuare l'*upload* della risorsa come richiesto dai sistemi di *open access*

Questo tipo di sistemi presenta il vantaggio di non dover gestire materialmente le risorse poiché viene alimentato dalla **segnalazione di identificatori web** attraverso cui reperire la risorsa da descrivere, senza doverla caricare e mantenere nel proprio server

- evita i problemi di conservazione

MA

- non riesce ad aggirare i problemi legati alla permanenza dell'url (se non tramite sistemi di reindirizzamento che gestiscono la risorsa, tipo PURL)



Grazie per l'attenzione!

daniela.canali@sanita.it