



InformaticaUmanistica

Biblioteche Digitali

Pasquale Savino

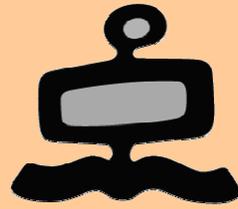
ISTI - CNR



UNIVERSITÀ DI PISA

Programma del corso

- ◆ Introduzione alle Biblioteche Digitali
- ◆ **Metadati**
- ◆ Esempi di Biblioteche Digitali (con esercitazioni)
- ◆ Architettura e tecnologie di base delle Biblioteche Digitali
- ◆ Progettazione di una Biblioteca Digitale (con esercitazioni)
- ◆ Nuove tendenze



InformaticaUmanistica

Metadati



UNIVERSITÀ DI PISA

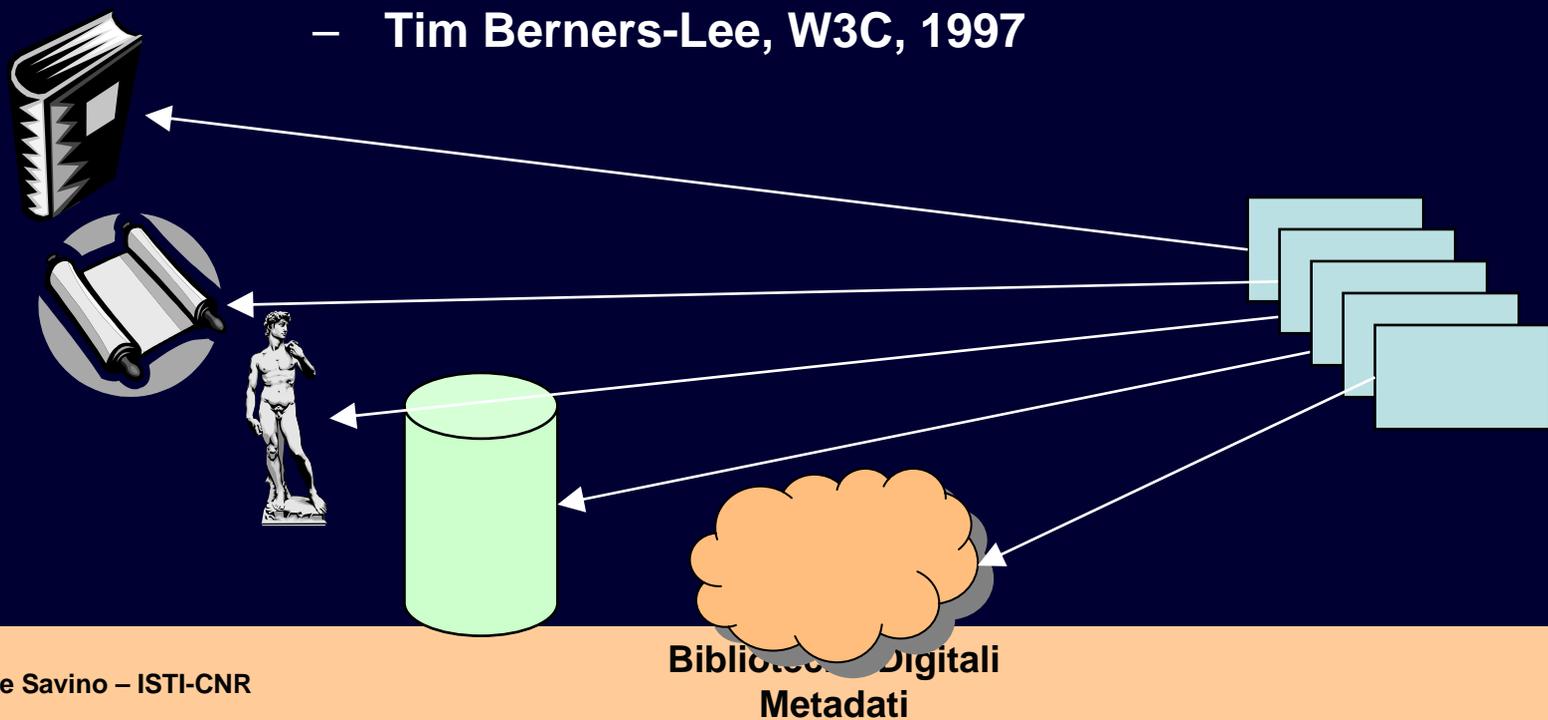
Sommario

- ◆ **Definizione**
- ◆ **XML**
- ◆ **I vari tipi di metadati**
- ◆ **Modelli di metadati**
 - MARC21
 - Dublin Core
 - MPEG-7
 - ECHO

Metadati - Definizione

Cosa sono i metadati?

- Alcune definizioni semplici...
- 'Structured data about data'.
 - **Dublin Core Metadata Initiative FAQ, 2003**
- Machine-understandable information about Web resources or other things.
 - **Tim Berners-Lee, W3C, 1997**



Una visione “funzionale” dei metadati

- Data associated with objects which relieves their potential users of having to have full advance knowledge of their existence or characteristics.
A user might be a program or a person.
 - **Lorcan Dempsey & Rachel Heery, 1998**
- Structured data about resources that can be used to help support a wide range of operations
 - **Michael Day, 2001**

A quali risorse possono essere associati dei metadati?

- I metadati possono essere associati a qualunque cosa – digitale, fisica, risorsa “astratta”
 - HTML documents
 - digital images
 - databases
 - books
 - museum objects
 - archival records
 - metadata records
 - Web sites
 - collections
 - services
 - physical places
 - people
 - institutions
 - abstract “works”
 - concepts
 - events

Che informazione contengono i metadati?

- I record di metadati includono
 - Record bibliografici appartenenti a cataloghi o ottenuti da servizi di indicizzazione e creazione di abstract
 - Descrizione del materiale archiviato
 - Record relativi ad oggetti contenuti nella documentazione del museo
 - Descrizione di oggetti digitali (documenti, immagini, video software, ...)
 - Descrizione di collezioni di oggetti digitali
 - Descrizione di servizi in rete
 - Descrizione di record di metadati

Quali operazioni possono essere effettuate sui metadati?

- Operazioni da parte degli utenti o da parte di tool software
- I metadati possono essere usati per supportare diverse funzioni, quali ad esempio
 - **resource disclosure & discovery**
 - **resource management, (incluso preservation)**
 - **intellectual property rights management**
 - **Autenticazione ed autorizzazione**
 - **Personalizzazione e localizzazione dei servizi**
- **Le diverse funzioni richiedono diversi tipi e classi di metadati**
 - **Non è possibile trovare una soluzione unica per tutti gli usi**
 - **È necessario specificare i requisiti funzionali per definire le caratteristiche dei metadati da utilizzare**

Elementi ed insiemi di elementi dei Metadati

- I Metadati descrivono gli attributi e le proprietà di una risorsa
- Ogni attributo o proprietà è descritto da un **metadata element**
 - Può essere identificato e definito in modo preciso
 - Ogni metadata element può essere rappresentato in diverse forme
- Un **metadata element set** è
 - Insieme coerente di elementi
 - Creato per uno scopo specifico, come una singola unità
- Schema
 - Rappresentazione strutturata di un element set

Autore = J Smith
Data = 2001-11-05
Titolo = Report

Resource1

I tipi di metadati

Tipi di metadati

- ◆ **Metadati descrittivi o per “resource discovery”**
 - Utilizzati per descrivere le risorse della Biblioteca Digitale
 - Utilizzati per facilitare la ricerca di documenti nella BD
 - Cataloghi, indici, ecc.
- ◆ **Metadati strutturali**
 - Descrivono la struttura degli oggetti della BD e le relazioni che esistono tra oggetti (o loro parti) ed altri oggetti
- ◆ **Metadati amministrativi**
 - Utilizzati per la gestione delle collezioni di documenti e per il controllo dell'accesso

Metadati per resource discovery

- Gli utenti desiderano
 1. Individuare le risorse in accordo a determinati criteri
 - Ad es. *Trova tutti i documenti scritti da ... che parlano di ...*
 2. (oppure) identificare una specifica risorsa
 - Essere in grado di distinguere tra diverse risorse simili tra loro
 3. Selezionare
 - **Scegliere una risorsa che sia appropriata ai bisogni**
 4. Localizzare la risorsa
 5. *Ottenere/accedere alla risorsa*
 6. *Utilizzare la risorsa*
 - ***open, read, display, run, play, copy, unpackage/repackage***
 - ***Interpretare il contenuto***
- I metadati per resource discovery sono utilizzati principalmente per le operazioni 1-4

Metadati per resource discovery

Continuum di complessità/funzionalità

Indici full-text, feature per immagini	Generato con tool software in modo automatico. Secondo alcune definizioni potrebbero non essere considerati metadati	discovery per contenuto, localizzazione
Forme semanticamente semplici (e.g. Dublin Core)	Permettono la descrizione di un gran numero di risorse In alcuni casi possono essere generati in modo automatico, in altri sono prodotti manualmente	discovery, identificazione, selezione, localizzazione
Forme complesse più ricche (e.g. MARC, ECHO, etc)	Relativo a tipi di risorse specifiche Spesso associato con domini/comunità specifici La creazione può richiedere un significativo intervento umano	discovery, identificazione, selezione, localizzazione, accesso, uso

Associazione di una risorsa ai metadati [1/3]

Creator = J Smith

Date = 2001-11-05

Title = Report

Le metadati possono essere contenuti all'interno della risorsa

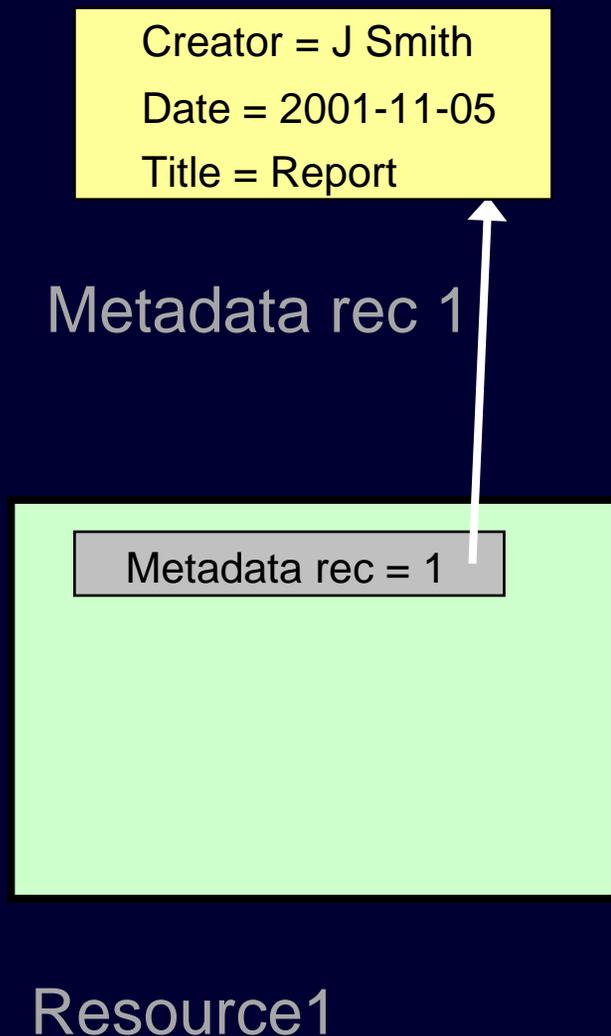
e.g. elementi `meta` in doc HTML;

Proprietà del documento nei word processor

- Le risorse possono permettere che i metadati siano contenuti al loro interno (embedding)?
- Il creatore dei metadati può avere i diritti di **scrittura** sulla risorsa?
- Può un servizio **estrarre** i metadati che sono contenuti nella risorsa?
- Possono esistere metadati per **aggregati di risorse**? Come vengono gestiti?

Resource1

Associazione di una risorsa ai metadati [2/3]



*Il record dei metadati è un oggetto separato
L'identificatore del record e' contenuto
nella risorsa*

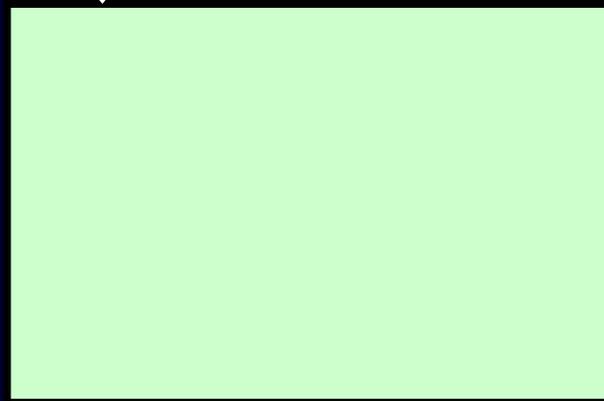
e.g. elementi `link` elements in doc HTML
I record di metadati possono essere remoti
rispetto alla risorsa

- Le risorse devono poter contenere i link
- Il creatore dei metadati può avere i diritti di **scrittura** sulla risorsa?
- Devono esistere dei meccanismi per seguire i link
- Cosa succede se una risorsa viene cancellata?
- Possono esistere metadati per **aggregati di risorse**? Come vengono gestiti?

Associazione di una risorsa ai metadati [3/3]

Doc = 1
Creator = J Smith
Date = 2001-11-05
Title = Report

Metadata rec 1



Resource 1

*Il record dei metadati è un oggetto separato
L'identificatore della risorsa e' contenuto nel
record*

Il record dei metadati può essere remoto rispetto alla risorsa

Non richiede che nessuna informazione sui metadati sia contenuta nella risorsa Non richiede che il creatore dei metadati possa modificare la risorsa

Il record dei metadati può essere creato indipendentemente dalla risorsa – possibilità di record multipli

I servizi utilizzano i metadati indipendentemente dalla risorsa

Il record dei metadati può continuare ad esistere anche se la risorsa viene cancellata

Il record dei metadati può descrivere qualunque oggetto che abbia un identificatore

Gestione dei metadati

Il record dei metadati viene usato separatamente dalla risorsa che viene descritta

I metadati sono essi stessi una risorsa

Il contenuto dei metadati può essere memorizzato nella forme più appropriate per il servizio: in un database, in forms, ecc.

