

Metadati

Antonella De Robbio

Sommario

- ⌘ Cosa sono i metadati?
- ⌘ I tipi di metadati
- ⌘ I linguaggi di Markup: HTML, XML...
- ⌘ Lo schema RDF
- ⌘ Esempi di metadati basati su RDF: RSS, RDDL
- ⌘ Verso il Semantic Web
- ⌘ Il Dublin Core

Metadati: una visione ingenua

- ⌘ **Metadati = dati che descrivono altri dati**
 - ☒ i record del catalogo di una biblioteca
 - ☒ i record di un database
- ⌘ **Spesso piu' facili da trattare dei dati che rappresentano**
 - ☒ Il formato dei metadati e' una decisione del creatore/catalogatore,
 - ☒ il formato dei dati dipende da molti fattori

Con i metadati si descrivono e si etichettano gli oggetti

- ⌘ Noi usiamo già i metadati nella vita reale
 - ☒ per trovare informazioni o cose
 - ☒ per organizzare informazioni o cose
- ⌘ dal libro Cent'anni di solitudine ...

accanto ad una visione ingenua...

... vi sono percezioni più scientifiche

... visioni del metadato correlate alle tecniche dell'Information Retrieval

... interpretazioni meramente tecniche ...

... nel nostro settore bibliotecomico ...

Metadati come etichette descrittive

⌘ Registrano le informazioni sull'oggetto (catalografiche o altro)

☒ Autore

☒ Titolo

☒ Copyright

☒ Informazioni di autenticazione

☒ Protezioni e permessi d'accesso

☒ Prezzo

⌘ esprimono:

☒ Struttura e semantica dei dati

A cosa servono i metadati

⌘ Utilizzo di metadati per le seguenti attività

☒ Catalogazione dei dati

☒ Elaborazione automatica dei dati da parte di agenti software

☒ Controllo degli accessi

☒ Filtraggio dei flussi informativi

Tipico metadato tradizionale

⌘ Il **MARC**

Machine Readable Cataloguing

⌘ da un formato di scambio ad un formato di lavoro cooperativo

⌘ metadato entro gli OPAC

⌘ USMARC, MARC21, DEMARC, ANNAMARC, UNIMARC, ...

⌘ Esperienza CORC (MARC21/DublinCore)

d2m : Dublin Core to MARC converter - Netscape

File Modifica Visualizza Vai Comunicatore ?

Segnalibri Indirizzo <http://www.bibsys.no/meta/d2m/> Arg correlati

d2m : Dublin Core to MARC converter

MARC format:

Presentation:

Show parsed metadata: ☐

Enter URL:

Enter HTML metadata:

Enter Reset Help

Documento: Operazione completata

008010812s2000 k

720 \$a Solimine
\$h Giovanni

513 \$a Relazione 11. Seminario Angela Vinay

546 \$a italian

260 \$b AIB - Associazione Italiana Biblioteche - Sezione Veneto
\$c 2000

787 \$a <http://www.aib.it/aib/sezioni/veneto/vinay11.htm>

787 \$a <http://www.aib.it>

680 \$a Biblioteche, Biblioteconomia, Economia, Finanziamenti,
Cooperazione, Bilancio, Servizi, Costi, Ricavi, Acquisti librari

655 \$a Text

24510\$a AIB. Sezione Veneto. Seminario Angela Vinay. 11.
BibliotECONOMIA: L'economia della cooperazione bibliotecaria

856 \$u <http://www.aib.it/aib/sezioni/veneto/vinay11/solimine00.htm>
\$q text/html

<http://www.bibsys.no/meta/d2m/distr/>

d2m: Dublin Core to MARC conversion - Netscape

File Modifica Visualizza Vai Comunicatore ?

Indietro Avanti Ricarica Home Ricerca Guida Stampa Sicurezza Interrompi

Segnalibri Indirizzo <http://www.bibsys.no/exec/d2m> Arg correlati

<http://www.aib.it/aib/sezioni/veneto/vinay11/solimine00.htm>

USMARC record:

```

008010812s
720 $a Solimine, Giovanni
260 $b AIB - Associazione Italiana Biblioteche - Sezione Veneto
7860 $n http://www.aib.it
653 $a Biblioteche, Biblioteconomia, Economia, Finanziamenti,
Cooperazione, Bilancio, Servizi, Costi, Ricavi, Acquisti librari
245 $a AIB. Sezione Veneto. Seminario Angela Vinay. 11.
BibliotECONOMIA: L'economia della cooperazione bibliotecaria
856 $u http://www.aib.it/aib/sezioni/veneto/vinay11/solimine00.htm
$q text/html

```

Documento: Operazione completata

Accesso alle risorse elettroniche: il campo 856

856 Localizzazione elettronica ed	856\$k Password
accesso	856\$l Login/login
Indicatore 1: Metodo di	856\$m Contact for access assistance
accesso	856\$n Nome of location of host in
# = Non sono fornite	subfield \$a
informazioni	856\$o Operating system
0 = Email	856\$p Port
1 = FTP	856\$q Electronic format type
2 = Remote login (Telnet)	856\$r Setngs
3 = Dial-up	856\$s File size
4 = HTTP	856\$t Terminal emulation
7 = Metodo specificato in \$y	856\$u Uniform resource locator
856\$a Host name	856\$v Hours access method available
856\$b Access number	856\$w Record control number
856\$c Compression information	856\$x Nonpublic note
856\$d Path	856\$y Access method
856\$f Electronic name	856\$z Public note
856\$g Uniform resource name	
856\$h Processor of request	
856\$i Instruction	
856\$j Bits per second	

MARC e Ruolo dell'autore

☞ Acquaforista	☞ Designer della legatura	• Libraio
☞ Annotatore	☞ Designer della sovraccoperta	• Librettista
☞ Antecedente bibliografico	☞ Destinatario	• Litografo
☞ Arrangiatore	☞ Detentore del privilegio	• Miniature
☞ Artista	☞ Direttore	• Narratore
☞ Autore di sceneggiatura, etc.	☞ Direttore d'orchestra, di coro, di gruppo musicale, ecc.	• Paroliere
☞ Autore incerto	☞ Disegnatore dei caratteri	• Precedente proprietario
☞ Calcografo	☞ Donatore	• Produttore
☞ Calligrafo	☞ Editore	• Recensore
☞ Cartografo	☞ Fabbrikante della carta	• Responsabile del montaggio
☞ Censore	☞ Falsario	• Rubricatore
☞ Commentatore	☞ Finanziatore / sponsor	• Sceneggiatore
☞ Compositore	☞ Firmatario	• Scrittore del materiale allegato
☞ Compositore di musica	☞ Fotografo	• Segretario
☞ Copista	☞ Grafico	• Sottoscrittore
☞ Coreografo	☞ Illustratore	• Stampatore
☞ Correttore	☞ Incisore	• Stampatore di tavole
☞ Correttore di opera manoscritta	☞ Interprete	• Tecnico della registrazione
☞ Datore del privilegio	☞ Intervistato	• Titolare del copyright
☞ Dedicante	☞ Intervistatore	• Traduttore
☞ Dedicatario	☞ Legatore	• Xilografo
☞ Designer del libro		
☞ Designer dell'ex-libris		

Tipi di metadati

- ☞ vi sono differenti suddivisioni tipologiche in gruppi differenti
- ☞ scuole di pensiero diverse
- ☞ categorie diverse di metadati: semantici, descrittivi, di presentazione, amministrativo- gestionali (MAG), di utilizzo, tecnici, per utenze particolari ...
- ☞ Differenti funzionalità di metadati conducono a differenti tipologie

Il Getty Standards Program suddivide i metadati in 5 categorie

- ☞ **Amministrativi gestionali MAG** (usati per la gestione e amministrazione delle risorse informative)
- ☞ **Descrittivi** (dal MARC al Dublin Core)
- ☞ **Conservazione** (compresa migrazione)
- ☞ **Tecnici** (comportamento dei metadati e funzionamento dei sistemi)
- ☞ **Utilizzo** (relativi al livello e al tipo di utilizzo dell'utente)

I metadati possono essere

- ☞ **Interni alla risorsa**
 - ☞ utili ai motori di ricerca
- ☞ **Esterni alla risorsa**
 - ☞ entro collezioni di metadati
 - ☞ esposizione di metadati
 - ☞ metadati in banche dati integrate con altri contenitori
 - ☞ collezioni a modello distribuito (DoIS)
 - ☞ servizi di accesso ai contenuti

I formati di metadati sono

- ⌘ Numerosi
- ⌘ Differenti per
 - ☒ scopo
 - ☒ tipologia
 - ☒ funzione
- ⌘ Possono essere
 - ☒ automatici
 - ☒ creati manualmente

La legge del metadato

- ⌘ **Non esiste solo il Dublin Core**
- ⌘ **Anche il MARC è un metadato**
- ⌘ *Ad ogni metadato il suo utente*
- ⌘ *Ad ogni comunità di utente il proprio metadato specifico*

; -)

Comunità di metadati ...

- ⌘ I metadati riflettono le esigenze delle comunità dai quali provengono
- ⌘ Differenti comunità producono documenti intellettuali peculiari all'utenza specifica che vi afferisce (matematici, fisici, chimici, ...)
- ⌘ Comunità diverse parlano e necessitano di linguaggi specifici con i quali formulano la propria produzione intellettuale
- ⌘ I metadati prodotti da queste comunità riflettono necessariamente la struttura e la forma dei documenti da esse prodotti

Standard, profilo e registro

- ☒ **Standard:** documento riconosciuto con generale consenso, o con approvazione organismo ufficiale
- ☒ **Profilo:** sottoinsieme di specifiche che caratterizzano una particolare applicazione di un formato, protocollo...fornisce maggiori specifiche per implementatori (Bath Profile Z39.50, DC Euler Profile)
- ☒ **Registro:** un sistema che registra la semantica, la struttura e i formati di scambio di ogni tipo di dato (database, messaggi, documenti)

Elaborazione automatica

- ⌘ Grazie ai metadati un programma può elaborare un documento come se ne conoscesse il significato
- ⌘ Tre scenari diversi (in ordine di difficoltà)
 - applicazione che gestisce metadati che conosce
 - applicazioni che si scambiano metadati concordati
 - **applicazioni che si scambiano metadati senza accordo preventivo (tipico di Internet)**

Ambiente Internet ...

- ⌘ Per internet occorre quindi sviluppare dei modelli che permettano alle applicazioni di interagire anche in assenza di un accordo preventivo
- ⌘ Per questo servono degli **standard** per:
 - ☒ rappresentare i metadati
 - ☒ scambiare i metadati
 - ☒ estrarre significati dai metadati tramite motori inferenziali

Il web semantico

- ⌘ Il progetto è stato lanciato nel 1998 dal W3 allo scopo di rendere le informazioni che popolano la rete
 - ☒ comprensibili alle macchine
 - ☒ processabili anche da programmi progettati in maniera indipendente
- ⌘ Il web non deve servire solo a visualizzare documenti ma anche e soprattutto alla integrazione e riutilizzo delle informazioni tra applicazioni diverse

Il web semantico

- ⌘ In un certo senso abbiamo bisogno di un web che sia computer-friendly per essere più user-friendly
- ⌘ Più le macchine saranno in grado di interpretare autonomamente le informazioni della rete più efficaci saranno i loro servizi a:
 - ☒ trovare informazioni pertinenti
 - ☒ prendere decisioni

Le tecnologie per il nuovo web

- ⌘ Il sw utilizza una serie di tecnologie di cui alcune sono ancora in fase sperimentale
- ⌘ Alcune tecnologie sono
XML RDF DC DTD XSL FO XPATH DOM
SAX JDOM XHTML JAXP TREX RDDL RSS
SOAP WSDL UDDI OIL ...

Risorse e URI

- ⌘ URI = Uniform Resource Identifier
- ⌘ Le risorse sono identificate da URI
- ⌘ La risorsa identificata da una URI puo' essere *astratta* = **non recuperabile in Rete**

Metadati basati su XML

- ⌘ XML Schema
- ⌘ RDF (Resource Description Framework)
- ⌘ RDF Schema

Metadati e infrastrutture tecniche

- ⌘ INTEROPERABILITA'
 - ➔ XML
 - Metalinguaggio con regole per costruire linguaggi di marcature
 - ➔ RDF
 - Infrastruttura che promuove interoperabilità dei metadata
 - ➔ ANSI/NISO Z39.50-1995
 - Protocollo di rete per la ricerca su un largo numero di database eterogenei

Definizione di risorsa

⌘ Dalla RFC 2396:

- ⌘ "A resource can be anything that has identity. Familiar examples include an electronic document, an image, a service (e.g., "today's weather report for Los Angeles"), and a collection of other resources. Not all resources are network "retrievable"; e.g., human beings, corporations, and bound books in a library can also be considered resources. The resource is the conceptual mapping to an entity or set of entities, not necessarily the entity which corresponds to that mapping at any particular instance in time. Thus, a resource can remain constant even when its content---the entities to which it currently corresponds---changes over time, provided that the conceptual mapping is not changed in the process".

Altre visioni di RDF

⌘ Sintassi XML standard per rappresentare *grafi etichettati diretti*

⌘ Sintassi XML standard per rappresentare *fatti in logica del primo ordine senza variabili*

Asserzione :=
(predicato,soggetto,oggetto)
Predicato = URI
Soggetto = URI
Oggetto = URI o costante
Oggetto =
Predicato(Soggetto)



XML è ...

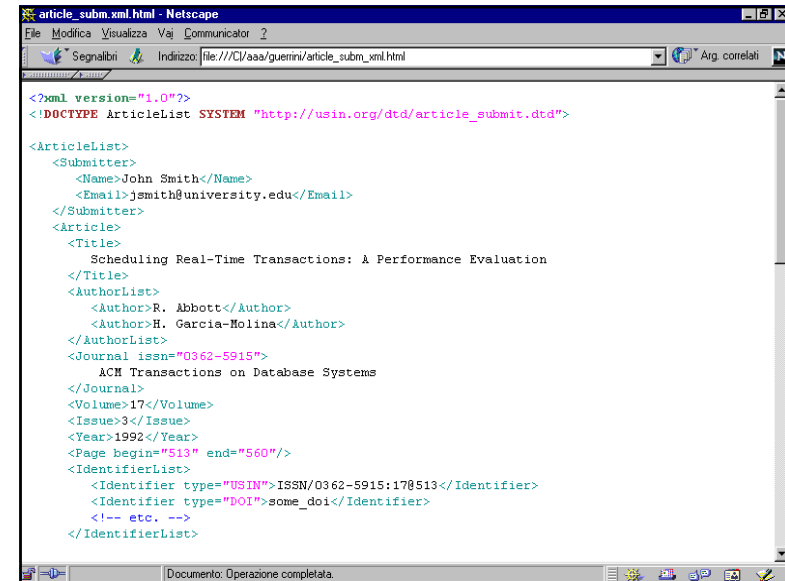
- ⌘ XML = eXtensible Markup Language
- ⌘ XML può essere definito in molti modi diversi
- ⌘ XML è una sintassi per marcare i documenti
- ⌘ XML è una famiglia di tecnologie correlate
- ⌘ XML consente di
 - ☒ aggiungere metadati (semantica) ai propri dati
 - ☒ separare il contenuto dalla presentazione
 - ☒ definire in maniera rigorosa il linguaggio di markup utilizzato
 - ☒ utilizzare un formato neutrale e testuale

... ma da solo XML non è sufficiente ...

- ⌘ XML è troppo flessibile mentre è necessario un modello estremamente semplice per esprimere delle conoscenze processabili automaticamente
- ⌘ HTML è estremamente semplice ma non flessibile e perciò può solo rappresentare informazioni senza tener conto della semantica

XML non è facile come HTML

- ⌘ Può essere un limite?
- ⌘ Con HTML chiunque poteva pubblicare in Web
- ⌘ XML è complesso e non tutti saranno in grado di utilizzarlo
- ⌘ Ritorno all'epoca pre-Web?
- ⌘ O selezione obbligata di ciò che verrà pubblicato?



```
<?xml version="1.0"?>
<!DOCTYPE ArticleList SYSTEM "http://usin.org/dtd/article_submit.dtd">

<ArticleList>
  <Submitter>
    <Name>John Smith</Name>
    <Email>jsmith@university.edu</Email>
  </Submitter>
  <Article>
    <Title>
      Scheduling Real-Time Transactions: A Performance Evaluation
    </Title>
    <AuthorList>
      <Author>R. Abbott</Author>
      <Author>H. Garcia-Molina</Author>
    </AuthorList>
    <Journal issn="0362-5915">
      ACM Transactions on Database Systems
    </Journal>
    <Volume>17</Volume>
    <Issue>3</Issue>
    <Year>1992</Year>
    <Page begin="513" end="560"/>
    <IdentifierList>
      <Identifier type="USIN">ISSN/0362-5915:170513</Identifier>
      <Identifier type="DOI">some_doi</Identifier>
    <!-- etc. -->
    </IdentifierList>
  </Article>
</ArticleList>
```

```
<?xlm version=1.0"?>
```

```
<libro>
```

```
<autore>Antonella De Robbio</autore>
```

```
<titolo>Diritto d'autore</titolo>
```

```
</libro>
```

```
<?xlm version=1.0"?>
```

```
<libro>
```

```
<titolo>Antonella De Robbio</titolo>
```

```
<autore>Diritto d'autore</autore>
```

```
</libro>
```

```
<?xlm version=1.0"?>
```

```
<libro autore="Antonella De Robbio">
```

```
<titolo>Diritto d'autore</titolo>
```

```
</libro>
```

...esempio

- ⌘ Nell'esempio precedente i tre documenti hanno lo stesso contenuto informativo seppure presentato con strutture diverse
- ⌘ Per noi umani i nomi degli elementi sono di grande aiuto per interpretare il contenuto
- ⌘ Ma cosa dire del seguente esempio, ovvero come un calcolatore vede il documento

```
<?xmlversion=1.0"?><v><x>  
aaaa</x><y>bbbb</y></v>
```

- ⌘ Cosa si può dire? che aaaa è y di bbbb o che bbbb è x di aaaa? Dal punto di vista della macchina i nomi degli elementi non aiutano, e non aiuta l'indentazione; quello che serve è una struttura rigida per specificare la "direzione" della relazione
- ⌘ A questo scopo il W3C ha inventato RDF (Resource Description Framework).
- ⌘ RDF aiuta la macchina a stabilire le relazioni

Resource Description Framework (RDF)

- ⌘ RDF è un modello di metadati
 - ⌘ RDF è una piattaforma per il trattamento dei metadati
 - ⌘ RDF costituisce una ipotesi di soluzione individuata dal W3C al problema della rappresentazione dei metadati
 - ⌘ RDF è complementare a XML
- <<http://www.w3.org/RDF/>>

RDF

- ⌘ Resource Description Framework
- ⌘ W3C Recommendation, dal Febbraio 1999
 - ☑ si basa su XML e spazi di denominazione
- ⌘ Indipendente dalla piattaforma
- ⌘ Eredita'
 - ☑ PICS
 - ☑ biblioteche digitali, SGML e rappresentazione della conoscenza
- ⌘ Scopo principale
 - ☑ rendere le etichette elaborabili da agenti software oltre che leggibili dall'uomo

Primi passi con RDF

- ⌘ I dati sono descritti da *asserzioni*
- ⌘ Un'asserzione e' composta da
 - ☑ Una *risorsa*
 - ☑ Pagina Web o parte di essa, intero sito Web...
 - ☑ ...sottoalbero XML ...un libro ... qualunque cosa dotata di URI
 - ☑ una *proprietà*
 - ☑ e.g. Titolo, Data, Tipo
 - ☑ un *valore*
 - ☑ es. stringa, numero, frammento XML, un'altra risorsa

Scopo e finalità di RDF

- ⌘ realizzare un meccanismo per la descrizione di risorse
- ⌘ definizione di un meccanismo
 - ☒ neutrale rispetto ai domini
 - ☒ adattabile alla descrizione di informazioni di qualsiasi dominio
- ⌘ permettere la definizione di semantiche per dati in formato XML
 - ☒ secondo modalità standardizzate e che mirano l'interoperabilità

Cosa offre RDF

- ⌘ RDF offre l'interoperabilità tra applicazioni che scambiano sul Web informazioni comprensibili dalle macchine
- ⌘ il modello RDF consente di rappresentare metadati in una sintassi per la codifica e la trasmissione di metadati
- ⌘ scambio secondo modalità che favoriscano l'interoperabilità fra server e client Web, sviluppati in modo indipendente

RDF può essere usato

- ⌘ da agenti software intelligenti per facilitare la condivisione e lo scambio di conoscenza
- ⌘ nella valutazione di contenuto
- ⌘ nel descrivere collezioni di pagine che rappresentano un unico "documento" logico
- ⌘ per descrivere i diritti di proprietà intellettuale di pagine Web
- ⌘ per esprimere le preferenze sulla riservatezza da parte di un utente così come le politiche di riservatezza di un sito Web

Il modello RDF

- ⌘ Il modello dei dati RDF offre una struttura concettuale astratta per la definizione e l'uso di metadati
- ⌘ Per creare e scambiare questi metadati è necessaria anche una sintassi concreta.
- ⌘ A grandi linee la sintassi di RDF usa XML (Extensible Markup Language) all'interno di un modello che poggia su tre concetti base: **il soggetto, l'oggetto e il predicato**

il soggetto, l'oggetto e il predicato

- ⌘ Il **Soggetto** è la risorsa descritta,
- ⌘ L'**Oggetto** è quell'affermazione che attribuisce ad una caratteristica della risorsa descritta un determinato valore.
- ⌘ Il **Predicato** è la caratteristica o l'aspetto specifico utilizzato per la descrizione del soggetto, che determina la proprietà della risorsa descritta.

Esempio

- ⌘ composizione di una frase
soggetto predicato e oggetto
(RDF fa lo stesso: è uno schema)
- ⌘ struttura da riempire di significati
lo catalogo un documento

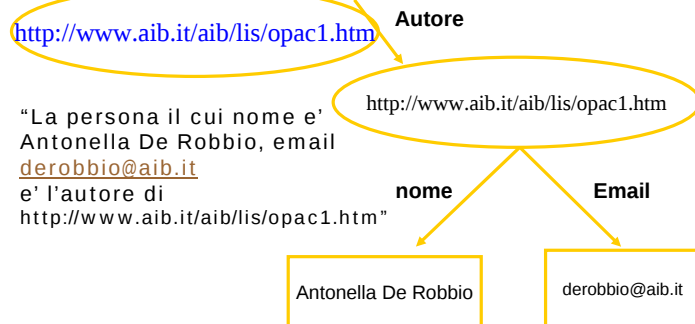
RDF

- ⌘ RDF definisce delle terne:
Risorsa TipoDiProprietà Valore
- ⌘ Ciascuna terna definisce una proprietà della risorsa
- ⌘ Per **risorsa** si intende qualunque entità identificabile tramite una URI
<http://www.aib.it/aib/lis/opac1.html>
- ⌘ **TipoDiProprietà**: autore:
- ⌘ **Valore**: "Antonella De Robbio"

Rappresentazione di RDF

- ⌘ Le terne di RDF possono essere rappresentate in molti modi, anche graficamente
- 
- ⌘ XML è usato per la sua sintassi rigorosa

Esempio di modello RDF



L'elemento RDF

- ⌘ L'elemento RDF è un semplice involucro che delimita i confini in un documento XML entro cui il contenuto è esplicitamente da intendere mappabile in una istanza di modello di dati RDF

Sintassi

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <rdf:description
    about="http://www.math.unipd.it/~derobbio/home/antohp.htm">
    <s:Creator>Antonella De Robbio</s:Creator>
  </rdf:description>
</rdf:RDF>
```

PROPRIETA' - VALORE
 (in questo caso costante)

SPAZIO DI DENOMINAZIONE =
 VOCABOLARIO DI DOMINIO =
 ELENCO DI PROPRIETA'

RISORSA

Schema RDF

- ⌘ Lo schema RDF è come una sorta di dizionario
- ⌘ In RDF il significato, fondamentale per la comprensione delle asserzioni, è espresso attraverso il riferimento ad uno schema
- ⌘ Attraverso la condivisione degli schemi, RDF favorirà il riutilizzo delle definizioni di metadati

Schema RDF

- ⌘ RDF disporrà di un sistema di classi, analogo a quelli utilizzato dai sistemi di programmazione e modellazione orientata agli oggetti
- ⌘ Una collezione di classi viene detta schema
- ⌘ Le classi sono organizzate in una gerarchia, e offrono meccanismi di estensibilità attraverso un raffinamento in sottoclassi
- ⌘ realizzazioni di nuovi schemi, attraverso modifiche incrementali rispetto allo schema di base.

Aree di applicazione di RDF

- ☒ ricerca di risorse al fine di migliorare le capacità dei motori di ricerca
- ☒ nella catalogazione per descrivere il contenuto e le relazioni del contenuto disponibili in un particolare sito Web
- ☒ nella biblioteca digitale
- ☒ associato alla firma digitale rappresenterà la chiave di volta per la costruzione di un " Web affidabile " per il commercio elettronico

RDF/XML e W3C

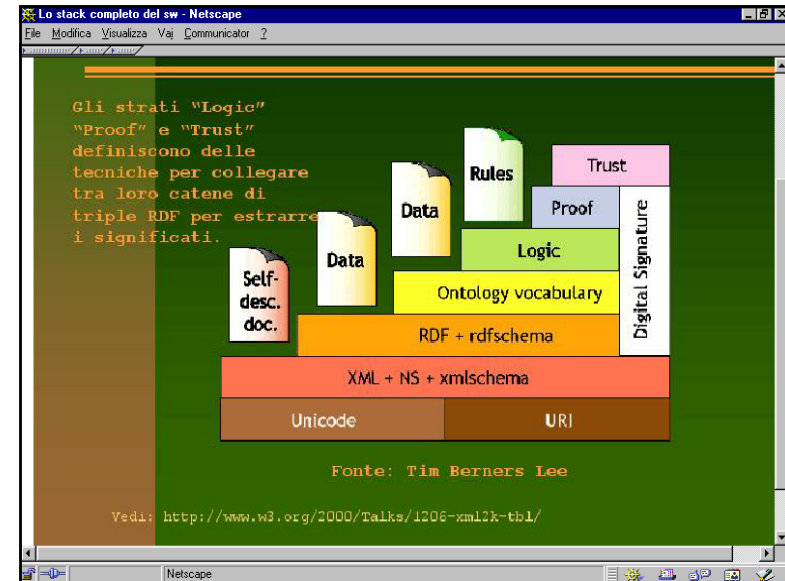
- ⌘ L'RDF Core Working Group ha rilasciato un Working Draft delle Specifiche di RDF/XML Syntax (Revisionate).
<http://www.w3.org/TR/2001/WD-rdf-syntax-grammar-20011218/>
- ⌘ Il documento aggiorna la grammatica della Specifica di Resource Description Framework (RDF) Model and Syntax ed è ora basato su XML Infoset e RDF Model Theory. Maggiori informazioni disponibili sulla Semantic Web Activity.
<http://www.w3.org/2001/sw/>

RDF da solo non è sufficiente

- ⌘ Perché RDF è solo un framework
- ⌘ Può definire solo la struttura della proposizione ma non dice niente sul contenuto
- ⌘ E' necessario che vengano definiti dei set di metadati per esprimere i **Tipi di Proprietà** usati nelle terne RDF

Lo stack XML/RDF/DC

- ⌘ Il modello attuale prevede l'utilizzo di tre standard come strati diversi (layers comunicativi)
 - ☒ XML stabilisce la sintassi di base (struttura fisica)
 - ☒ RDF definisce la struttura delle relazioni (struttura logica)
 - ☒ DC definisce i contenuti (struttura per i contenuti)
- ⌘ Il modello esposto non è l'unico possibile



RDF la terra di mezzo

“RDF, the middle layer, will only be visible to the wizards who invent new ways of cataloging, describing and organizing information - the librarian of the future.”

Goldfarb, Prescod - The XML Handbook (pag. 53)

Dublin Core

- ⌘ Abbreviazione di **Dublin Metadata Core Element Set**
- ⌘ Formato di metadata definito dal **OCLC/NCSA Metadata Workshop** (marzo 1995)
- ⌘ Dublin Core è anche uno standard
- ⌘ DC definisce un set minimo di elementi per descrivere le risorse di rete a sintassi indipendente di
- ⌘ 15 elementi con qualificatori
- ⌘ Il set minimo degli elementi DC comprende: title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage, rights

Scopi del Dublin Core

- ⌘ Nato per descrivere pagine web attraverso etichette da applicare alla risorsa informativa
- ⌘ Convivenza con altri schemi di metadati
- ⌘ Facilita i motori nella raccolta delle informazione in rete
 - limitazione del rumore
 - inserimento del metadato entro i TAG in HTML e XML

Perché è nato il Dublin Core

- ⌘ Creazione di metadati da parte degli stessi creatori delle risorse
- ⌘ Rappresentazione degli aspetti essenziali relativi alla descrizione di una risorsa/oggetto
- ⌘ Definizione della semantica della descrizione (HTML non basta, non ne tiene conto)
 - ➡ Miglioramento “resource discovery”
 - ➡ Collaborazione attraverso consenso

Caratteristiche principali del Dublin Core

- ⌘ Ciascun elemento può essere:
 - ☑ opzionale
 - ☑ ripetibile
- ⌘ Indipendenza della sintassi
- ⌘ Estensione
- ⌘ Interdisciplinarietà
- ⌘ Internazionale

Dublin Core

- ⌘ 15 etichette
- ⌘ Metadato di tipo descrittivo
- ⌘ record bibliografici
- ⌘ Campi tipizzati
 - ☑ vocabolario controllato
 - ☑ ISO 8601 (date)
 - ☑ tipi di risorse
 - ☑ URLs, ISBNs
 - ☑ RC 1766 (lingue)

Elementi Dublin Core (1)

⌘ Titolo	⌘ Pubblicatore
⌘ Creatore (persona, organizzazione, servizio)	⌘ Contributore
⌘ Soggetto (lista di parole chiave dal vocabolario controllato)	⌘ Data
	⌘ Tipo (e.g. collezione, evento, immagine, oggetto fisico, software, suono, ...)
	⌘ Formato

Elementi Dublin Core (2)

⌘ Identificatore	⌘ Diritti (proprietà intellettuale, copyright)
⌘ Fonte	
⌘ Lingua (lingua)	⌘ Descrizione (riassunto, indice, o testo libero)
⌘ Relazione (collegamento a una risorsa correlata)	
⌘ Copertura (es. luogo, tempo, a giurisdizione)	

```
<head>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
<meta name="DC.Creator" content="Solimine, Giovanni">
<meta name="DC.Date" content="2000-08-10">
<meta name="DC.Description" content="Relazione 11. Seminario Angela Vinay">
<meta name="DC.Format" content="text/html">
<meta name="DC.Identifier"
content="http://www.aib.it/aib/sezioni/veneto/vinay11/solimine00.htm">
<meta name="DC.Language" content="italian">
<meta name="DC.Publisher" content="AIB - Associazione Italiana Biblioteche -
Sezione Veneto">
<meta name="DC.Relation"
content="http://www.aib.it/aib/sezioni/veneto/vinay11.htm">
<meta name="DC.Source" content="http://www.aib.it">
<meta name="DC.Subject" content="Biblioteche, Biblioteconomia, Economia,
Finanziamenti, Cooperazione, Bilancio,
Servizi, Costi, Ricavi, Acquisti librari">
<meta name="DC.Type" content="Text">
<meta name="DC.Title" content="AIB. Sezione Veneto. Seminario Angela
Vinay. 11. BibliotECONOMIA: L'economia della cooperazione
bibliotecaria">
<title>11&deg; Seminario Angela Vinay - BibliotECONOMIA</title>
</head>
```

I qualificatori DC

Sono di due tipi

scheme=contesto per interpretazione di un valore
basato su standard esterni (DDC) e regole di
codifica

ISO 8601 per data, ISO 639/1 per la lingua
type=raffina la semantica di un elemento

DC.CREATORE.MUSICISTA

Ogni elemento ha un qualificatore raccomandato

DC e lo spazio dei nomi

- ⌘ DC può essere esteso con altri elementi all'interno del modello DC
- ⌘ Si possono utilizzare altri elementi definiti in namespaces diversi
- ⌘ I namespaces servono ad identificare in maniera univoca il set di metadati che stiamo utilizzando

Progetto Euler per la matematica

- ⌘ Nasce nel 1998 nel quadro di "Telematics for Libraries"
- ⌘ Consente l'accesso integrato a specifiche risorse informative in rete nel campo della matematica
- ⌘ Realizza una Biblioteca digitale di matematica europea con integrazione delle eterogenee risorse già esistenti
- ⌘ Integra differenti contenitori informativi: basi dati bibliografiche (MATHZbl), OPACS, riviste elettroniche, archivi di pre-prints, letteratura grigia e risorse Internet controllate
- ⌘ Descrizione delle risorse eterogenee provenienti dai vari contenitori informativi basate su DC
- ⌘ Utilizzo di chiave di de-duplicazione
- ⌘ One-shop point = unico punto d'accesso utilizzando un'unica interfaccia disponibile sul WWW.

PICS

- ⌘ Platform for Internet Content Selection
www.w3.org/PICS/
- ⌘ Recommendation dal 1996
- ⌘ Condivisione e interscambio di valutazioni su siti e pagine Web
 - ☒ etichette PICS trasportate negli header HTTP o come marcatori META
 - ☒ from document server
 - ☒ or an independent *label bureau*
 - ☒ firmati elettronicamente e associati con un *digest*

Esempio

```
<html>
<head>
<META http-equiv="PICS-Label" content='
(PICS-1.1 "http://www.gcf.org/v2.5"
  labels on "1999.11.05T08:15-0500"
  until "2001.12.31T23:59-0000"
  for "http://w3.org/PICS/Overview.html"
  ratings (suds 0.5 density 0 color/hue 1))
'>
</head>
<body>
<p>qui va il documento</p>
</body>
</html>
```

Applicazioni di PICS

- ⌘ Protezione dei minori
 - ☒ CyberPatrol, EvalWeb, NetNanny, ...
- ⌘ Utente genera o ottiene un *profilo*
- ⌘ L'applicazione confronta rating e profilo e compie una delle seguenti azioni:
 - ☒ suggest
 - ☒ search
 - ☒ monitor
 - ☒ warn
 - ☒ block

Etichette Recreational Software Advisory Council (RSAC): Violenza

- ⌘ (category (transmit-as "v") (name "Violence") (label (name "**Conflict**") (description "Harmless conflict; some damage to objects") (value 0)) (label (name "**Fighting**") (description "Creatures injured or killed; damage to objects; fighting") (value 1)) (label (name "**Killing**") (description "Humans injured or killed with small amount of blood") (value 2)) (label (name "**Blood and Gore**") (description "Humans injured or killed; blood and gore") (value 3)) (label (name "**Wanton Violence**") (description "Wanton and gratuitous violence; torture; rape") (value 4)))

Problemi

- ⌘ Chi sceglie le etichette?
- ⌘ Chi ne definisce la sintassi?
- ⌘ Come determinarne il significato?
- ⌘ **A quale dato sono associate?**

Biblioteca digitale

- ⌘ Fornisce i propri documenti di metadati appropriati
- ⌘ Creare collezioni di metadati a fianco delle collezioni di documenti
- ⌘ Espone le proprie collezioni di metadati
- ⌘ Congloba le tecniche documentarie e biblioteconomiche tradizionali con le nuove applicazioni per il trattamento delle risorse elettroniche

Risorse utili

⌘ Metadata UKOLN

<http://www.ukoln.ac.uk/metadata/>

⌘ Sito IFLA

<http://www.ifla.org/II/metadata.htm>

⌘ Dublic Core

<http://dublincore.org/>

Risorse utili (2)

⌘ Mapping between metadata formats

<http://www.ukoln.ac.uk/metadata/interoperability/>

⌘ A review of metadata: a survey of current resource description formats Work Package 3 of Telematics for Research project DESIRE (RE 1004)

http://www.ukoln.ac.uk/metadata/desire/overview/rev_toc.htm

Risorse utili (3)

⌘ Gruppo di studio ICCU sugli standard e le applicazioni di metadati nei beni culturali

<http://www.iccu.sbn.it/metadati.htm>

☒ tradotto la versione 1.1 del "Dublin Core Metadata Element Set"

☒ ha implementato il Dublin Core sul sistema SBN-on line

☒ partecipa al progetto europeo DELOS NOE

⌘ Lista dei progetti che utilizzano i metadati in Italia

Il gruppo è articolato in due sottogruppi:

⌘ metadati descrittivi

☒ predisposizione di modello di raccolta dei dati descrittivi per individuare gli elementi di metadati ritenuti essenziali all'identificazione dell'oggetto digitale

☒ riferimento ad un modello logico (FRBR)

☒ mapping per confronto relativo ai Metadati o ai Profili

⌘ metadati gestionali-amministrativi e strutturali

☒ definizione di un set di metadati relativi

☒ 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

Alcuni generatori...

⌘ DC-dot

<http://www.ukoln.ac.uk/metadata/dcdot/>

⌘ META builder

<http://vancouver-webpages.com/META/mk-metas.html>

⌘ DONOR metadatagenerator

<http://www.kb.nl/coop/donor/mg-start.html>

⌘ Automatic RDF Metadata Generator

<http://www.scit.wlv.ac.uk/~ex1253/metadata.html>