

5.1 Introduzione ai linguaggi di marcatura

Insegnamento di Informatica

Elisabetta Ronchieri

Corso di Laurea di Economia, Università di Ferrara

I semestre, anno 2014-2015



Argomenti

Concetto di marcatura

Tipi di marcatura

Metadati



Argomenti

Concetto di marcatura

Tipi di marcatura

Metadati



Esempio di marcatura

Data la seguente frase

Thousands of geniuses live and die undiscovered - either by themselves or by others (migliaia di geni vivono e muoiono senza essere scoperti, o da loro stessi o dagli altri) by Mark Twain.

elementi del linguaggio di marcatura (markup) sono:

- ▶ gli spazi, delimitatori di parole;
- ▶ la virgola, delimitatore di frase;
- ▶ il trattino per evidenziare qualcosa;
- ▶ le parentesi per aggiungere una precisazione;
- ▶ il punto, delimitatore di periodo.

Sono tutte aggiunte al testo scritto che lo rendono più fruibile.



Cosa é la marcatura

- ▶ La marcatura fa parte del nostro sistema di scrittura.
- ▶ Fornisce una esplicita interpretazione di un testo.
- ▶ La marcatura é un insieme di convenzioni per decidere quali parti del documento vanno trattate in modo diverso dalle altre parti, come il titolo del testo.



Con la digitalizzazione

- ▶ Il formato digitale del testo ha consentito di sviluppare tipologie di marcatura focalizzate all'elaborazione da parte di un calcolatore.
- ▶ Informatizzando l'impaginazione e la stampa dei testi, la marcatura in un testo elettronico permette di controllarne la formattazione, la stampa e la gestione.
- ▶ Tramite l'inserimento di specifiche sequenze di caratteri o codici (detti markup tags o tags) nel testo i programmi possono distinguere le varie parti di un documento.



Linguaggio di marcatura

- ▶ Un linguaggio di marcatura é un insieme di convenzioni (sistema formale) per scambiare e pubblicare informazioni in formato testo in modo strutturato.
- ▶ É nato per segnalare le caratteristiche logiche di un documento e delle sue parti attraverso opportune istruzioni.
- ▶ Deve specificare:
 - ▶ le marcature permesse;
 - ▶ come distinguere tra marcatura e testo;
 - ▶ il significato della marcatura.
- ▶ Ogni linguaggio di marcatura favorisce la riusabilit , la flessibilit  e l'apertura ad applicazioni complesse.



Argomenti

Concetto di marcatura

Tipi di marcatura

Metadati



Tipi di marcatura

Esistono diverse marcature a seconda del sistema di elaborazione, dell'applicazione e dello scopo a cui il documento é soggetto:

Nessuna : tipica dei sistemi di scrittura piú antichi.

Puntuazionale : é definita dai sistemi di scrittura e fornisce informazioni sintattiche.

Presentazionale : si occupa dell'effettiva impaginazione e presentazione dei documenti.

Procedurale : definisce comandi per indicare come il testo debba essere elaborato.

Dichiarativa : ha lo scopo di definire la struttura logica di un testo.

Referenziale : permette di riferirsi ad entit  esterne al documento e di sostituirle tali entit  durante l'elaborazione.



Nessuna marcatura

Esempio:

le previsioni del tempo non sono buone questa settimana

Marcatura puntuazionale

- ▶ Consiste nell'usare un insieme prefissato di segni per fornire informazioni sintattiche sul testo.
- ▶ Le regole di punteggiature sono stabili e note agli autori.
- ▶ Esistono problemi nell'usare la punteggiatura:
 - ▶ incertezze strutturali legate per esempio all'uso della virgola, punto o punto e virgola;
 - ▶ incertezze grafiche per esempio per l'uso delle virgolette aperte e chiuse o neutre;
 - ▶ ambiguità procedurali essendo per esempio il punto usato sia per delimitare un periodo che per indicare un'abbreviazione.



Marcatura presentazionale

- ▶ Consiste nell'indicare effetti grafici o altro per rendere piú chiara la presentazione del contenuto.
- ▶ Esempi sono:
 - ▶ l'uso di interlinee;
 - ▶ l'uso di pallini per le liste;
 - ▶ il cambio di pagina all'inizio di una nuova sezione.



Marcatura puntuazionale e presentazionale

Esempio:

Le previsioni del tempo non sono buone questa settimana.
Previste temperature sotto lo zero ad alta quota e
nevicate in pianura.

Aggiunte maiuscole, minuscole, punteggiatura, spazi e ritorni a capo.



Marcatura procedurale

- ▶ Indica le procedure da eseguire per la visualizzazione del contenuto.
- ▶ Specifica le istruzioni che devono essere eseguite per la successiva visualizzazione nei Browser.
- ▶ Gestisce i testi da un punto di vista estetico, sfruttando le capacità del sistema di presentazione per avere l'effetto voluto.
- ▶ Un esempio é HTML (Hyper Text Markup Language).



Marcatura procedurale

Esempio:

```
<HTML>
<HEAD>
<TITLE>Monumenti Storici</TITLE>
</HEAD>
<CENTER><H1>In Italia</H1></CENTER>
<FONT SIZE=+2>
<CENTER>Emilia Romagna</CENTER>
</FONT>
</HTML>
```

Aggiunte istruzioni o comandi che il sistema di lettura esegue sul testo.



Caratteristiche marcatura procedurale

- ▶ É basata sulla procedura per visualizzare il documento.
- ▶ É dipendente dal sistema di visualizzazione, ciascuno dei quali impone le proprie regole e sintassi per carattere, dimensione, corsivi, grassetti, margini, interlinea.
- ▶ Non é contestuale.
- ▶ Non é possibile inserire vincoli sulla correttezza di un documento.



Marcatura dichiarativa

- ▶ Identifica strutturalmente il tipo di ogni elemento del contenuto.
- ▶ Dichiarare l'appartenenza di un blocco di testo ad una determinata classe di strutture.
- ▶ Lascia la scelta del tipo di rappresentazione, da applicare al testo, al software che di volta in volta lo riprodurrà.
- ▶ Gestisce i contenuti.
- ▶ Un esempio é XML (eXtensible Markup Language).



Marcatura dichiarativa

Esempio:

```
<?xml version="1.0"?>
<nitf>
<head> <title>Monumenti Storici</title> </head>
<body>
<body.head>
<headline><h1>Torre di Pisa</h1></headline>
<dateline>
<location>Pisa, Italia</location>
<story.date>Dicembre 29 2014</story.date>
</dateline>
</body.head>
<body.content> </body.content>
</body>
</nitf>
```

Aggiunte descrizioni sugli elementi del documento.



Caratteristiche marcatura dichiarativa

- ▶ É basata sul ruolo (come titolo, paragrafo o citazione) all'interno del testo.
- ▶ É indipendente dal sistema di visualizzazione.
- ▶ É basata su categorie (ruoli) che riflettono le caratteristiche di ogni elemento automaticamente.
- ▶ É contestuale, specificando le regole di correttezza dei documenti codificati.



Marcatura referenziale

- ▶ Consiste nel far riferimento ad entità esterne al documento per fornire significato o effetto grafico agli elementi del documento.
- ▶ Per esempio una sigla nota viene sostituita dalla parola intera durante la stampa.

Argomenti

Concetto di marcatura

Tipi di marcatura

Metadati



Cosa sono i metadati

Definizione semplice.

I metadati sono informazioni strutturate relative ai dati, interpretabili da parte di un computer.

Dalla definizione si deduce che:

- ▶ i metadati forniscono informazioni relative ai dati;
- ▶ le informazioni hanno una struttura ben definita;
- ▶ l'informazione può essere interpretata ed utilizzata dal programma di un computer.



Cosa sono i metadati

Definizione piú precisa.

I metadati sono informazioni strutturate appartenenti a diverse tipologie, che vengono utilizzate per effettuare un insieme di operazioni su un insieme di risorse.

- ▶ La risorsa può essere:
 - ▶ digitale come un documento elettronico, un'immagine e un servizio;
 - ▶ fisica come libri, persone e opere;
 - ▶ astratta.



Cosa sono i metadati

- ▶ Dati che forniscono informazioni su una fonte informativa.
- ▶ Informazioni che caratterizzano i dati.
- ▶ Dati utili ad identificare caratteristiche condivise da piú documenti.
- ▶ Dati strutturati su dati.



Tipi di metadati

Esistono tre classi di metadati:

1. Descrittivi che descrivono le risorse per facilitarne il recupero e l'identificazione.
 - ▶ Tipicamente usati per descrivere le risorse di una biblioteca digitale.
 - ▶ Possono includere nel caso di biblioteche contenenti documenti il titolo, l'autore e la data di pubblicazione.
2. Strutturali che descrivono la struttura delle risorse e le relazioni che esistono tra risorse diverse o tra parti di una risorsa stessa.
 - ▶ Per una biblioteca digitale di documenti, questi metadati possono descrivere la struttura di un documento costituito da capitoli e sezioni, oppure possono descrivere i riferimenti tra documenti diversi quali le citazioni.
3. Amministrativi o gestionali che forniscono una descrizione delle risorse da usare come supporto alla gestione delle stesse.
 - ▶ Possono includere gli utenti che possono accedere alla risorsa o a parti di essa, una descrizione degli ultimi backup effettuate.



Dove descrivere i metadati

- ▶ I metadati possono essere incorporati nella risorsa digitale o possono essere conservati separatamente.
- ▶ Conservare i metadati con la risorsa assicura che non vengano persi e che siano aggiornati insieme.
- ▶ Conservare i metadati separatamente può semplificarne la gestione e facilitarne la ricerca e il recupero.

Marcatura e metadati

- ▶ La marcatura dichiarativa é lo strumento che associa ad una risorsa dei metadati.
- ▶ I metadati possono essere descritti in un qualsiasi linguaggio di marcatura dichiarativa.



Schema di metadati

- ▶ Uno schema di metadati é un insieme di elementi di metadati sviluppati con uno specifico scopo come quello di descrivere un tipo particolare di risorsa informativa.
- ▶ La definizione o il significato degli elementi di uno schema é detta semantica dello schema.
- ▶ I valori dati agli elementi dello schema sono una istanza dello schema.
- ▶ Uno schema di metadati specifica i nomi degli elementi e la loro semantica.



Dublin Core

- ▶ Il Dublin Core é uno schema di metadati sufficientemente generico per essere utilizzato in domini diversi e applicato a qualsiasi risorsa di informazione accessibile nel World Wide Web.
- ▶ I domini Dublin Core sono espressi in XML o RFC.
- ▶ Usato oggi per catalogare o collezionare dati su internet.
- ▶ Usato da ricercatori, librari e curatori di musei.

Ulteriori dettagli:

<http://dublincore.org/documents/usageguide/index.shtml>

