

MOBILE DEVELOPER

The background is a solid blue color with a faint, complex network of white lines and dots, resembling a web or a data structure. A large, thin white arc is positioned in the upper right corner, partially enclosing the text.

Crawler & Spider

Sommario

- Introduzione ai Crawler & gli Spider.

Le pagine web ed i crawler

Si fanno diverse attività per rendere interessanti i contenuti di un sito ai propri visitatori e spesso si trascurano altri tipi di ottimizzazioni rivolte ad altri visitatori: i crawler dei motori di ricerca.

Le pagine web ed i metadati

I crawler scandagliano il web per favorire l'indicizzazione dei siti dei motori di ricerca.

Spesso le sole pagine web non bastano e si ricorre all'uso dei metadati.

I metadati sono dati che descrivono altri dati.

Indicizzazione e conoscenza

Con riferimento al motore di ricerca più famoso al mondo si può affermare che:

Un conto è permettere a Googlebot di analizzare la struttura del sito web attraverso la navigazione dei link. Cosa del tutto diversa è aggiungere a questo conoscenza delle informazioni distribuite mediante il sito web.

Microdata

I Microdata rappresentano un modo standardizzato per **aggiungere conoscenza** nelle pagina web, **estendendo** così il markup HTML5.

Al fine di inserire queste notazioni semantiche si è scelto di inserire 5 attributi globali, che assolvono a questo compito, ad ogni TAG del linguaggio HTML5.

Microdata: itemscope

Identifica un'entità, un oggetto (per esempio un film, un articolo di un blog).

```
<div>  
    <div itemscope >  
    </div>  
</div>
```


Microdata: itemtype

Esprime la tipologia cui l'entità appartiene, idealmente per mezzo di un vocabolario condiviso.

```
<div>  
    <div itemscope itemtype = "http://schema.org/Movie" >  
    </div>  
</div>
```

Vocabolari per itemtype

L'attributo itemtype può riferirsi a definizioni di entità nei seguenti siti:

data-vocabulary.org

schema.org

Microdata: itemid

Un identificativo univoco globale ove disponibile (per esempio il codice ISBN di un libro).

```
<div>  
  <div itemscope itemtype =http://schema.org/Movie  
  id="numb1" itemid="http://www.imdb.com/title/tt0816692">  
  </div>  
</div>
```

Microdata: itemprop

Ciascuna proprietà dell'entità sarà contrassegnata da questo attributo.

```
<div>  
  <div itemscope itemtype =http://schema.org/Movie >  
    <div>1. <em  
itemprop="name">Interstellar</em></div>  
  </div>  
</div>
```

Microdata: itemref

Esplicita la relazione tra proprietà ed entità di appartenenza ove questa non possa essere inferita dal contesto (l'entità non è innestata nell'entità di riferimento).

```
<div>  
  <div itemscope itemtype =http://schema.org/Movie >  
    <div>1. <em itemprop="name">Interstellar</em>  
    <strong itemprop="director" itemscope="http://  
      schema.org/Person">  
      <span itemprop="name"> Christopher Nolan </span>  
    </strong>  
  </div>  
</div>  
</div>
```

Microdata: Correttezza sintattica

Non esistono modi per validare un microdata, migliore sarà la marcatura delle informazioni e più possibilità si hanno di far capire al crawler cosa trattano le pagine web (al fine di ottenere un rich snippet). Il seguente link verifica la correttezza sintattica:

<https://developers.google.com/structured-data/testing-tool/>

Significato di un'immagine

Sembra che non esista un algoritmo in grado di interpretare il significato di immagine!

Microformat

I Microformat hanno molte analogie con microdata.

La principale differenza è che non usano dei TAG specifici per marcare la pagina, ma usano dei nomi di classe specifici. Ne sono esempi i seguenti:

hCard

hCalendar

hReview

Microformat: correttezza sintattica

Anche in tal caso non esistono veri e propri validatori, ma solo un sito di riferimento:

<http://microformats.org/>

Inoltre, esso mette a disposizione un utile strumento per creare hCard, hCalendar, ecc:

<http://microformats.org/code-tools>

Domande & approfondimenti

MOBILE DEVELOPER