

Sistemi Distribuiti

HTML e XHTML

Questo lavoro è concesso in uso secondo i termini di una licenza Creative Commons
(vedi ultima pagina)

Sommario

- HTML: semantica di visualizzazione
- Storia di HTML
- Principali problemi tecnologici
- Standard rilevanti
 - HTML 1.0 e CSS
 - Modularizzazione e HTML 1.1
- XHTML Basic
 - principali moduli
 - convalida del codice
 - uso degli standard

Introduzione

- HTML
 - un linguaggio per descrivere documenti ipermediali
 - documenti fatti di titoli, paragrafi, testo ecc.
 - collegamenti (link) ad altri documenti
 - contenuti di altri media (es: immagini)
- Linguaggio di marcatura
 - obbedisce ad un DTD di SGML

Introduzione

- SGML
 - sintassi per descrivere la struttura gerarchica di documenti
 - il “padre” di XML (anni '70); più complesso
 - DTD: schemi per i documenti
- Meno restrittivo di XML
 - i tag di chiusura possono mancare
 - maiuscole e minuscole
 - apici per gli attributi

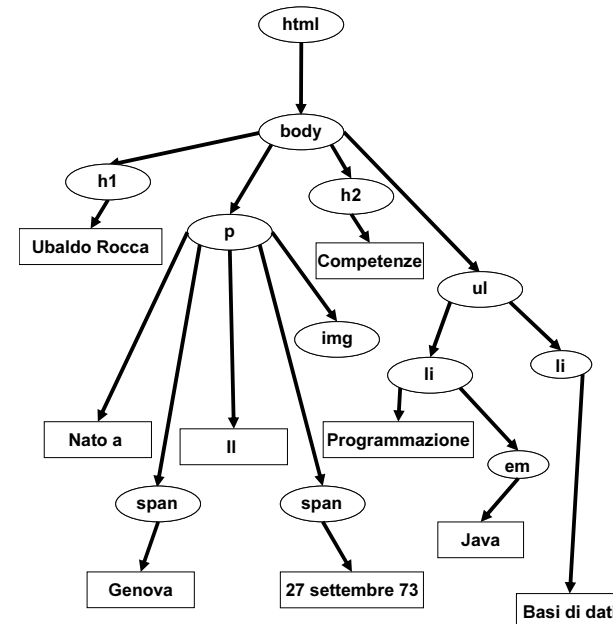
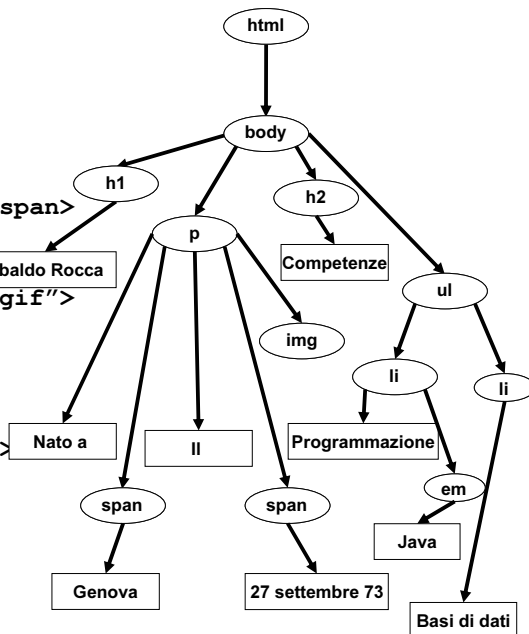
Introduzione

```
<html>
<head>
  <title> Il portale universale </title>
</head>
<body>
  <h1>Portale Universale</h1>
  <p>Questa &egrave; la pagina principale del
    <b>Portale Universale</b>.
    
    <a href="materiale.html">Materiale Interessante</a>
    ...
    <br />
    <i>Ultima modifica 23 Febbraio 2004</i>
  </p>
</body>
</html>
```

HTML: semantica di visualizzazione della pagina

- I documenti HTML corrispondono ad alberi
- In fase di visualizzazione, la nidificazione dei nodi corrisponde ad una nidificazione di riquadri
- Distinguiamo riquadri
 - *Inline*: il flusso non viene interrotto
 - *Block*: il flusso viene interrotto

```
<html>
<body>
  <h1>Ubaldo Rocca</h1>
  <p>
    Nato a <span>Genova</span>
    il <span>27 settembre
    1973</span>
    ... 
  </p>
  <h2>Competenze</h2>
  <ul>
    <li>Programmazione
      <em>Java</em></li>
    <li>Basi di Dati</li>
  </ul>
</body>
</html>
```



Principali Problemi del Linguaggio

- Problema 1: Evoluzione molto rapida
 - mancanza di metodologie consolidate
 - struttura e presentazione sono indipendenti
- Indicazione del Consorzio
 - separare struttura e presentazione
 - CSS livello 1 e CSS livello 2

Principali Problemi del Linguaggio

- Problema 2: Codice scorretto
 - i browser sono “onnivori”
 - i DTD dello standard non sono rispettati
 - problemi per le applicazioni
- Indicazione del Consorzio
 - adottare la sintassi XML
 - XHTML 1.0: una riformulazione di HTML 4.0 secondo lo standard di XML

Principali Problemi del Linguaggio

- Problema 3: Diverse capacità dei browser
 - diverse versioni di HTML, diversi DTD
 - diverse piattaforme (palmari, WebTV, ...)
- Indicazione del Consorzio
 - modularizzazione di XHTML e negoziazione
 - principio di “accessibilità”
 - estensibilità
 - “Modularization of XHTML”, XHTML1.1

Standard Rilevanti

- XHTML 1.0
 - gennaio 2000
- CSS livello 1, CSS livello 2
 - dicembre '96, maggio '98
- Modularization of XHTML, XHTML 1.1
 - aprile 2001, maggio 2001
- XHTML Basic
 - dicembre 2000

XHTML 1.0

- Riformulazione di HTML 4.0
 - formato e DTD XML e non SGML
- In altri termini
 - i documenti XHTML corretti sono alberi validi rispetto ad un DTD XML fissato dal consorzio
 - il DTD descrive quali elementi e quali attributi possono comparire in una pagina XHTML

XHTML 1.0

- DTD di riferimento
 - DTD strict XHTML1-strict.dtd
è quello di riferimento; presentazione CSS
 - DTD transitional XHTML1-transitional.dtd
è da intendersi come temporaneo
 - DTD frameset XHTML1-frameset.dtd
è di fatto “deprecato”
- Ci occuperemo solo di XHTML1-strict

XHTML 1.0

- Vincoli sulla sintassi
 - preambolo XML
 - dichiarazione del DTD es:


```
<!DOCTYPE html PUBLIC
  "-//W3C//DTD XHTML 1.0 Strict//EN"
  "xhtml1-strict.dtd">
```
 - l'elemento radice è <html>
 - il namespace del documento è
http://www.w3.org/1999/xhtml
 - valgono le regole della sintassi di XML

XHTML 1.0

- Vincoli sulla sintassi
 - il documento deve essere ben formato
 - elementi vuoti; es:

 - valori degli attributi tra apici;
es: xmlns="http://www.w3.org/1999/xhtml"
 - codifica delle entità; es: &
 - numerose entità predefinite: Latin-1
characters, Special characters, Symbols

XHTML 1.0 e CSS

- Separazione tra struttura e presentazione
 - i tag descrivono solo la struttura
 - fogli di stile CSS: associano caratteristiche di visualizzazione agli elementi (>>)
 - foglio di stile standard del browser; es: testo in carattere Times 10 punti tondo

Modularizzazione

- Idea
 - divisione del linguaggio in moduli
 - per consentire di implementare frammenti diversi del linguaggio in client diversi
 - meccanismo di estensibilità: è possibile per lo sviluppatore definire nuovi moduli
- Modulo
 - frammento del DTD di XHTML1-strict
 - collezione di elementi e relativi attributi

Elenco dei Moduli

- Nucleo ("Core")
 - Structure, Text, Hypertext, List
- Estensioni del testo
 - Presentation, Edit, Bidirectional Text
- Maschere
 - Basic Forms, Forms
- Tabelle
 - Basic Table, Table
- Immagini
 - Image
- Mappe
 - Client Side Map, Server Side Map
- Frame
 - Frame, Target, Iframe
- Fogli di stile
 - Stylesheet, Link
- Altri
 - Intrinsic Events, Metainformation, Scripting, Object, Base, Name, Legacy

Modularizzazione

- Evoluzione dello standard
 - lo standard stabilisce come definire altri moduli
 - le successive versioni di XHTML saranno ottenute combinando moduli di XHTML 1.0 e nuovi moduli
 - devono contenere almeno i moduli "core" (Structure, Text, Hypertext, List)
- XHTML 1.1: "Module Based XHTML"
 - selezione di moduli di XHTML 1.0 + modulo Ruby
- Attualmente: XHTML 2.0

Modularizzazione

- Obiettivo di questa lezione
 - presentare i moduli fondamentali del linguaggio XHTML 1.0
- Impostazione
 - cominciamo da XHTML Basic di cui verrà presentata una panoramica completa
 - questo dovrebbe fornire gli strumenti per utilizzare XHTML 1.0 Strict nel complesso

XHTML Basic

- Insieme minimale di moduli
 - per tutti i browser (cHTML e WML)
 - ragionevole capacità di formattazione
- Organizzazione
 - moduli del nucleo
 - fogli di stile esterni
 - immagini
 - semplici maschere
 - semplici tabelle

XHTML Basic

- Insieme minimale di moduli
 - ma ampia capacità di formattazione
- Funzionalità di XHTML Basic
 - moduli del nucleo (struttura, testo, ipertesto, liste)
 - modulo per fogli di stile esterni
 - modulo per le immagini
 - modulo per semplici maschere
 - modulo per semplici tabelle

XHTML Basic

- | | |
|---|--|
| <ul style="list-style-type: none"> ➤ Elenco completo <ul style="list-style-type: none"> – Structure, Text, Hypertext, List – Image – Basic Forms – Basic Tables – Link – Object – Meta, Base | <ul style="list-style-type: none"> ➤ Mancano <ul style="list-style-type: none"> – script (es: JavaScript) – presentazione (es: b) – mappe – tabelle nidificate – frame – “legacy” tag (es: font) |
|---|--|

XHTML Basic

➤ Intestazione di una pagina XHTML-basic

```
<?xml version="1.0" ?>
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
    "DTD-xhtmlbasic/xhtml-basic10.dtd">
```

➤ Il DTD di XHTML Basic

- xhtml-basic10.dtd
- numerosi file, uno per ciascun modulo

XHTML Basic

➤ Dove reperire i DTD

- tutti i DTD sono reperibili dal sito del W3C

➤ www.w3.org

- sezione dedicata a HTML e XHTML
- sono in linea tutti gli standard
- i DTD sono distribuiti con i documenti relativi agli standard (XHTML Basic e XHTML 1.0)

XHTML Basic

➤ Nella lezione

- Structure, Text, Hypertext, List
- Image
- Basic Tables
- Meta

➤ Nelle lezioni successive

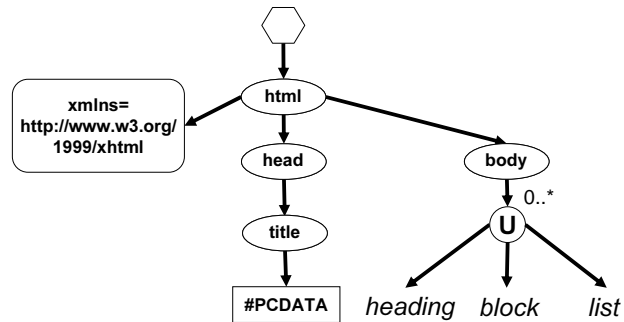
- Basic Forms
- Link
- Object
- Base

XHTML Basic

➤ Vincoli sulla sintassi di XHTML 1.0

```
<?xml version="1.0" ?>
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
    "DTD-xhtmlbasic/xhtml-basic10.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
...
</html>
```

Modulo Structure



Modulo Structure

elemento	attributo	modello di contenuto
html	xmlns ...	(head, body)
head	...	(title)
title	...	#PCDATA
body	...	(heading block list)*

Modulo Structure

```
<?xml version="1.0" ?>
  <!DOCTYPE html
    PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
    "DTD-xhtmlbasic/xhtml-basic10.dtd">
  <html xmlns="http://www.w3.org/1999/xhtml">
    <head>
      <title>Esempio</title>
    </head>
    <body>
      ...
    </body>
  </html>
```

Modulo Text

- Contiene elementi detti di flusso ("*flow*")
 - Heading
 - titoli delle sezioni del documento
 - Block
 - blocchi identificabili del documento
 - Possono contenere altri blocchi, inline e testo
 - Inline
 - porzioni di testo, contengono solo inline o testo

Modulo Text

➤ Heading

`<h1>`, `<h2>`, `<h3>`,
`<h4>`, `<h5>`, `<h6>`

➤ Block

- `<p>`: paragrafi
- `<pre>`: testo preformattato
- `<address>`: indirizzi
- `<blockquote>`: citazioni (attributo `cite`)
- `<div>`: raggruppamento di elem. in blocco

➤ Inline

- ``: enfaticizzato
- ``: rafforzato
- `
`: interruzione di linea
- `<q>`: citazioni brevi (attributo `cite`)
- ``: raggruppamenti
- ...

Modulo Text

```
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>Esempio</title>
  </head>
  <body>
    <h1>Titolo principale</h1>

    <p>Questo &grave; un documento di esempio
      relativo ad <em>XHTML Basic</em> <br />
      Sono andato a capo</p>
  </body>
</html>
```

Modulo Hypertext

➤ Elemento `<a>` (ancora)

- collegamenti ad altri URI
- attributo `href` con riferimento ad URI
- si aggiunge agli elementi *inline*

elemento	attributo	modello di contenuto
a	href ...	(#PCDATA <i>inline</i>)

Modulo Hypertext

```
<body>
  <h1>Titolo principale</h1>
  <p>Questo &grave; un documento di esempio
    relativo ad <em>XHTML Basic</em> <br />
    Sono andato a capo</p>
  <a href="http://www.dia.uniroma3.it">
    <strong>Sito del DIA</strong>
  </a>
  <a href="documenti/prova.html">
    Una pagina di prova
  </a>
</body>
```

Modulo *List*

➤ Liste di elementi

- Numerate (``) e non numerate (``)
- Elementi ``

elemento	attributo	modello di contenuto
ol	...	(li)+
ul	...	(li)+
li	...	(#PCDATA flow)*

Modulo *List*

```
<body> ...
  <a href="http://www.dia.uniroma3.it">
    <strong>Sito del DIA</strong>
  </a>
  <a href="documenti/prova.html">
    Una pagina di prova
  </a>
  <ul>
    <li>primo elemento di una lista</li>
    <li>secondo elemento di una lista</li>
  </ul>
</body>
```

Modulo *Image*

➤ Elemento ``

- esprime il collegamento ad un'immagine
- formati ordinari: gif, jpg, etc.
- si aggiunge agli elementi *inline*
- Attributi fondamentali: *src*, *alt*, *width*, *height*

elemento	attributo	modello di contenuto
img	src alt width, height, ...	EMPTY

Modulo *Image*

```
<body>
  ...
  <ul>
    <li>primo elemento di una lista</li>
    <li>secondo elemento di una lista</li>
  </ul>
  
</body>
```

Modulo *Table*

➤ Tabelle

- elemento radice: **<table>**
attributo summary per descrivere il contenuto
- didascalia dopo la radice: **<caption>**
- una o più righe (table row): **<tr>**
- fatte di una o più celle
 - celle di intestazione (table header): **<th>**
 - celle dei valori (table data): **<td>**

Modulo *Table*

elemento	attributo	modello di contenuto
table	summary ...	(caption?, tr+)
caption	...	(#PCDATA inline)*
tr	...	(th td)+
th	colspan, rowspan, scope (row column) ...	(#PCDATA flow)* meno altri elementi table
td	colspan, rowspan ...	(#PCDATA flow)* meno altri elementi table

Modulo *Table*

```

<body>
...
<table>
  <caption>Lezioni</caption>
  <tr>
    <th>Data</th><th>Argomento</th><th>Tipo</th>
  </tr>
  <tr>
    <td>Lun 17 feb 03</td><td>HTTP</td><td>Lezione</td>
  </tr>
  <tr>
    <td>Mar 18 feb 03</td><td>XML</td><td>Lezione</td>
  </tr>
  <tr>
    <td>Mer 19 feb 03</td><td>XML</td><td>Esercizi</td>
  </tr>
...
</table>
...
</body>

```

XHTML >> Principali Moduli >> XHTML Basic

Esempio

HTML code snippets corresponding to the table structure:

```

<table>
<tr>
  <th>A</th>
  <th>B</th>
  <th>C</th>
  <th>C</th>
</tr>
<tr>
  <td>xxx</td>
  <td>yyy</td>
  <td>zzz</td>
  <td>ww</td>
</tr>
<tr>
  <td colspan="2">due righe</td>
  <td>123</td>
  <td rowspan="3">tre col.</td>
</tr>
</table>

```

Modulo Meta

➤ Si aggiunge al contenuto di **head**

– **head** (**title**, **meta**+)

– Insieme di coppie name, content

ES: `<meta name="author" content="Pinco" />`

– **http-equiv**: intestazioni HTTP

ES: `<meta http-equiv="Expires" content="<data>" />`

elemento	attributo	modello di contenuto
meta	name, content, http-equiv ...	EMPTY

Modulo Meta

```
<?xml version="1.0" ?>
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
    "DTD-xhtmlbasic/xhtml-basic10.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
    <title>Esempio</title>
    <meta name="author" content="Pippo" />
    <meta name="keywords" content="esempio, HTML" />
  </head>
  <body>
    ...
  </body>
</html>
```

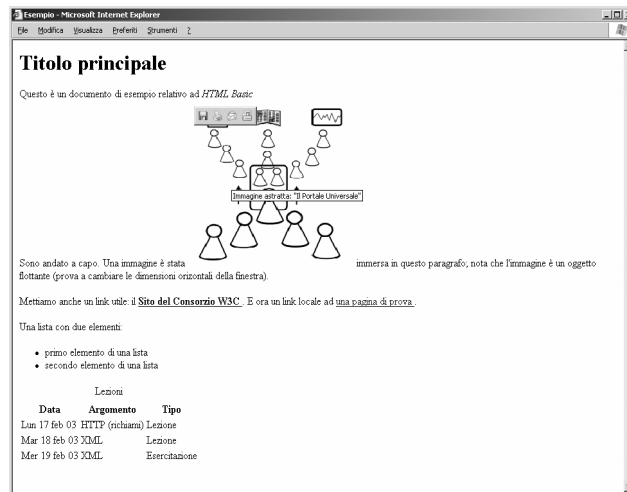
Esempio Completo

```
<?xml version="1.0" ?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
  "DTD-xhtmlbasic/xhtml-basic10.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
  <title>Esempio</title>
  <meta name="author" content="Pippo" />
  <meta name="keywords" content="HTML" />
</head>
<body>
  <h1>Titolo principale</h1>
  <p>Questo &grave; un documento
    di esempio relativo ad <em>HTML
    Basic</em> <br />
    Sono andato a capo. Una immagine
    &grave; stata  immersa in questo paragrafo;
    nota che l'immagine &grave; un oggetto
    flottante (prova a cambiare le dimensioni
    orizzontali della finestra).
  </p>
  Mettiamo anche un link utile: il
  <a href="http://www.w3.org">
    <strong>Sito del Consorzio W3C</strong>
  </a>. E ora un link locale ad
  <a href="prova.html">
    una pagina di prova
  </a>.
</body>
</html>
```

Esempio Completo (cont)

```
<p>
  Una lista con due elementi:
  <ul>
    <li>primo elemento di una lista</li>
    <li>secondo elemento di una lista</li>
  </ul>
</p>
<table>
  <caption>Lezioni</caption>
  <tr>
    <th>Data</th><th>Argomento</th><th>Tipo</th>
  </tr>
  <tr>
    <td>Lun 17 feb 03</td><td>HTTP</td><td>Lezione</td>
  </tr>
  <tr>
    <td>Mar 18 feb 03</td><td>XML</td><td>Lezione</td>
  </tr>
  <tr>
    <td>Mer 19 feb 03</td><td>XML</td><td>Eser.</td>
  </tr>
</table>
</body>
</html>
```

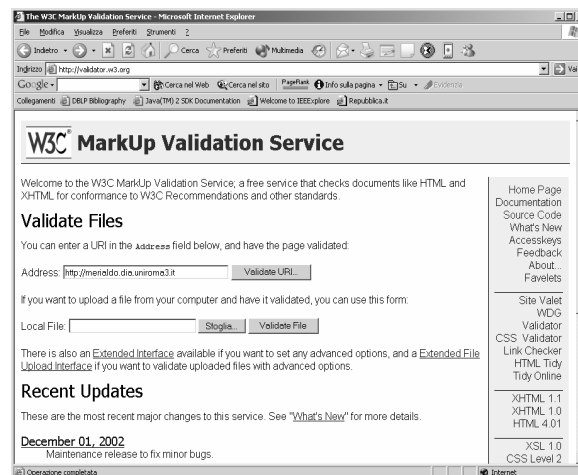
Esempio Completo



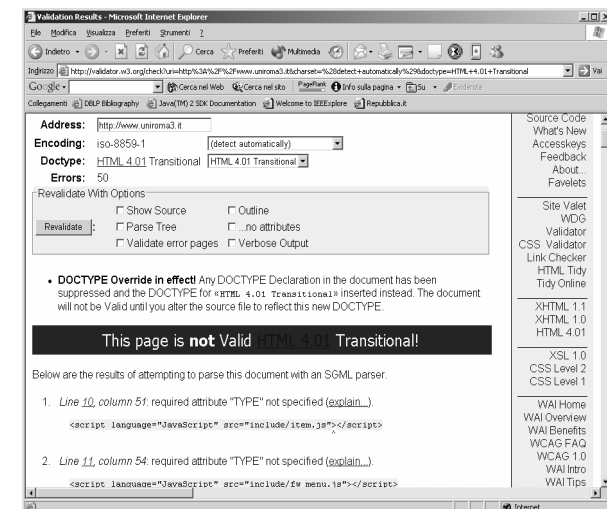
Convalida del Codice

- Il browser non effettua convalide
- E' consigliato verificare la correttezza
 - operazione noiosa ma utile
- Validatori
 - <http://validator.w3.org>, validatore in linea
 - ...

Validatore



Validatore



Da XHTML Basic a XHTML 1.0 Strict

- Per la maggior parte delle applicazioni
 - XHTML Basic e CSS consentono di sviluppare la maggior parte dei siti
 - vantaggio: è praticamente visualizzabile da qualsiasi browser e dispositivo
- In alcuni casi
 - se i moduli sono insufficienti, allora è possibile utilizzare XHTML 1.0 Strict

Da XHTML Basic a XHTML 1.0 Strict

- Limiti tipici di XHTML Basic
 - non consente di utilizzare tabelle nidificate >> modulo Tables di XHTML 1.0 Strict
 - non consente di utilizzare script nella pagina (es: JavaScript): >> modulo Script
 - non consente di utilizzare alcuni controlli nelle maschere: >> modulo Form
 - non consente di specificare stili CSS per i singoli elementi: >> modulo Style

Utilizzo degli Standard

- In questi casi
 - è necessario consultare lo standard di XHTML 1.0 Strict: documenti del consorzio
 - purtroppo la consultazione non è semplicissima
- Esempio
 - supponiamo di voler verificare la sintassi e la semantica di un elemento di XHTML Basic oppure di XHTML 1.0 Strict

Termini della Licenza

Termini della Licenza

- Questo lavoro:
 - si basa su materiale originale preparato dal prof. G. Mecca della Università della Basilicata nell'ambito del progetto Diogene (<http://www.difa.unibas.it/gmecca/diogene>), e successivamente adattato dall'ing. Paolo Merialdo della Università Roma Tre.
 - viene concesso in uso secondo i termini della licenza "Attribution-ShareAlike" di Creative Commons.
- Per ottenere una copia della licenza, è possibile visitare <http://creativecommons.org/licenses/by-sa/1.0/> oppure inviare una lettera all'indirizzo Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.