Building Interoperable Web Metadata

Hal Abelson, Ben Adida

MIT
Computer Science & Artificial Intelligence Laboratory
6 April 2006
Customers
IPTC

- NewsML
- Standards-compliant language
- strong metadata support
Creative Commons

• 50M licensed pages
  (not including Flickr photos)
• Self-contained HTML chunk
• cc:license
• (Similar use case to FOAF, DC, GeoURL)
Demo
Web Publishers

Blogr™

shutr™

calendar
Requirements of Interoperable Metadata
Requirement:
Publisher Independence

- Trackbacks
- URL commenting on

- Camera and Lens Info
- Resolution, Size, etc..
<h2>Great New Sushi Place</h2>
<p>it was amazing, I tell you. The freshest fish I've seen in years. And the toro...</p>

licensed under a <a href="http://creativecommons.org/licenses/by-nc-nd/2.5/">CC License</a>.
Requirement: Self-Containment

Great new Sushi place

it was amazing, I tell you. The freshest fish I've seen in years. And the toro...

Posted by Ulrich User 12 days ago

Posted in sushi, moderate price

License: some rights reserved

Author: Ulrich
Requirement: Modularity

DeskBlog
All Your Blogs

Blogr

shutr™

Tags
- food
- sushi
- salmon
- prawn
- ginger
- e-kagen
- japanese
- brighton

Posted in sushi, moderate price
Requirement: Evolvability

- Blogr
  - Posted in sushi, moderate price

- shutr™
  - Tags: food, sushi, salmon, prawn, ginger, e-kagen, japanese, brighton

- DeskBlog
  - All Your Blogs

- calendar
  - Tags: work, conference, talk
    - Suggestions: work conference talk

- Equivalence
Publisher Independence
Data Reuse
Self-Containment
Schema Modularity
Evolvability
RDF/A
(Primer WD #1 published 4 weeks ago)
licensed under a
<a href="http://cc.org/licenses/by/2.5/">Creative Commons License</a>.
licensed under a
<a rel="cc:license"
    href="http://cc.org/licenses/by/2.5/">
    Creative Commons License
</a>.
This document was written by Ben Adida and is licensed under a Creative Commons License, and a translation of ‘metadata’.
This document was written by <span property="dc:creator">Ben Adida</span> and is licensed under a <a rel="cc:license" href="http://cc.org/licenses/by/2.5/">Creative Commons License</a>, and a <a rel="cc:citeURL" property="cc:derivativeDescription" href="http://mit.edu/hal/metadata">translation of ‘metadata’</a>, originally by <span property="cc:attributedTo">Hal</span>.
This photo was taken by <span property="dc:creator">Ben Adida</span> and is licensed under a <a rel="cc:license" href="http://cc.org/licenses/by/2.5/">Creative Commons License</a>.
A Few Details

• Metadata about internal chunks:
  about="#feed"

• Existing semantics of HTML reused:
  ul, ol, li
## RDF/A by Requirement

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publisher Independence</td>
<td>everyone decides their own schema</td>
</tr>
<tr>
<td>Data Reuse</td>
<td>rendered data = structured data</td>
</tr>
<tr>
<td>Self-Containment</td>
<td>chunk of HTML is self-contained</td>
</tr>
<tr>
<td>Schema Modularity</td>
<td>attributes are reusable</td>
</tr>
<tr>
<td>Evolvability</td>
<td>RDF/OWL equivalence statements</td>
</tr>
</tbody>
</table>
RDF/A
Elegant Degradation

- Fully marked up XHTML
- Client software interprets as much as it can understand, structuring it for the user to see.
- Anything the client doesn’t understand is just XHTML-rendered.
<h2>Events</h2>
<a>Calendar</a>.
<p>Events.</p>
<link rel="rss:items" href="#items" />
<ul id="items">
  <li href="#item_56">Lessig - Free Culture</li>
  <li href="#item_57">Zittrain - E2E is Dead</li>
</ul>

<div id="item_56" about="#item_56" role="cal:Vevent">
  <h3>Lessig - Free Culture</h3>
  <p>
    Larry will be giving his talk on Free Culture at <a>CSAIL</a> on <span>2006-03-08 14:00-05</span> for about <span>1</span> hour
  </p>
</div>
XHTML 1.1 Module

- RDF/A as an XHTML 1.1 Module
- Main differences:
  - `<link>` and `<meta>` appear in BODY.
  - `about` and `property` attributes.
  - CURIES for compact URIs.
- SemWeb Deployment Working Group.
“RDF/A's balance of XML rigour and XHTML 1.1 pragmatism (placing it between XHTML2 and microformats) makes it something that we'll be quite interested in following, and we'd like to make sure that Firefox is a suitable host for RDF/A development and deployment.”

Mike Shaver, Mozilla
Back to the Demo
Questions?
Backup Slides
Technology for Web Metadata

- XML
- RDF
- Microformats
- GRDDL
- RDF/A
Beautiful Sushi.

Taken on January 15th, 2006.

licensed under a

<a href="http://creativecommons.org/licenses/by/2.5/">CC License</a>. 
<xml version="1.0"?>
<Photos
 xmlns="http://shutr.net/schema.xsd">
 ...
 <Photo id="123">
  <title>Beautiful Sushi</title>
  <creator>Ulrich</creator>
  <date>2006-01-15</date>
  <license
   url="http://cc.org/licenses/by/2.5/">
   ...
 </license>
 </Photo>
</Photos>
<Photos xmlns="http://shutr.net/schema.xsd">
  ...
  <Photo id="123">
    <title>Sushi</title>
    <creator>Ulrich</creator>
    <date>2006-01-15</date>
    <license url="http://cc.org/licenses/by/2.5/">
      ...
    </license>
  </Photo>
  ...
</Photos>

<Photos xmlns="http://flickr.net/schema.xsd">
  ...
  <Photo id="456">
    <title>Sushi 2</title>
    <creator>Ursula</creator>
    <date>2006-01-16</date>
    <license url="http://cc.org/licenses/by-nc/2.5/">
      ...
    </license>
  </Photo>
  ...
</Photos>
RDF

Some Rights Reserved

This work is licensed under a Creative Commons License.
RDF - A Refresher

- Everything is a triple: subject, property, object.
- Every item is either a URI or a literal (e.g. 3, “Ben”)

<> cc:license <http://cc.org/licenses/by/2.5/>.

<> dc:creator “Ben Adida”.

</images/sunset.jpg> dc:title “Sunset on the Beach”.
Beautiful Sushi.
Taken on January 15th, 2006.
licensed under a
<a href="http://cc.org/licenses/by/2.5/">CC License</a>.

<!--
-->

dc:title "Beautiful Sushi";
dc:creator "Ulrich";
dc:date "2006-01-15";
cc:license <http://cc.org/licenses/by/2.5/>.
-->
Beautiful Sushi.

Taken on January 15th, 2006.

licensed under a

<a href="http://creativecommons.org/licenses/by/2.5/">CC License</a>.
“If some usage of the web is too nascent to have converged, we can’t easily codify it, so we choose to pass on the problem.”

Ryan King, Microformats.org

http://microformats.org/blog/2006/01/17/microformats-are-semantic-markup/
<div class="photo">
<img src="/ulrich/photos/123.jpg" />
<span class="phototitle">
    Beautiful Sushi.
</span>
Taken on <span class="photodate">January 15th, 2006.</span>
licensed under a <a rel="license" href="http://cc.org/licenses/by/2.5/">CC License</a>.