Hacking Climate Change with Drupal and the Semantic Web

David Peterson & Nicholas Roberts

Drupalcon 2008, Szeged Hungary
“the value of your information grows exponentially with your ability to combine it with new information”

– Henry Story, Sun Microsystems
In its present form, the Web is a loose collection of data islands called websites.
Clear Picture?
Structured Data = Information

• Data and information are not the same
• Data is raw, non-structured text and numbers
• Information is taking that raw data and turning it into something meaningful
Steinhausen Flying Back Automatic Watch - Rose Gold-TW499RG

Price: $500.00

Availability: In Stock. Ships from and sold by Timecove.
Steinhausen Flying Back Automatic Watch-Rose Gold-TW499RG<br>Other products by Steinhausen<br><br><b class="h1">Special Offers and Product Promotions</b><br><div class="content"
<ul class="qoUL">
Get free shipping on this item when you spend $125.00 or more on Qualifying Items offered by Timecove. <a href="/gp/promotions/detailspopup/A19RHN75E613JQ" target="AmazonHelp" onclick="return amz_is_PopWin(this.href,'AmazonHelp',width=450,height=600,resizable=1,scrollbars=1,toolbar=1,status=1');">Here's how</a> (restrictions apply)<br></ul>
<script type="text/javascript">
showHideElement('specialOffersHidden', 'specialOffersDiv');
showHideElement('productPromosHidden', 'heroQuickPromoDiv');
</script>
<a name="moreAboutThisProduct" id="moreAboutThisProduct"></a>
<hr class="bucketDivider" noshade="noshade" size="1">
<br><b class="h1" align="left">Product Specifications</b><br><div class="bucket" align="left">
<table class="techSpecTable" cellspacing="0">
<tbody><tr class="techSpecRowHeader">
<td class="techSpecHeader" colspan="2">Watch Information</td></tr>
<tr class="techSpecRow"><td class="techSpecTD1">Brand Name:</td><td class="techSpecTD2">Steinhausen</td></tr>
<tr class="techSpecRow"><td class="techSpecTD1">Part Number:</td><td class="techSpecTD2">TW499RG</td></tr>
</tbody></table>
</div>
Towards Understanding

• What if the computer knew that this was *something*?
• And not just anything but a watch
Connections

• Definition of watch is linked to Wikipedia entry

• Now things become really interesting
  – You now have created a shared vocabulary
  – Use within your enterprise to overcome integration difficulties
  – Use on the global Web
Linked Data

LastFM: Tracks

Geonames: Sydney

Geo:lat: -33.87
Geo:long: 151.20

Jamendo: Sydney Bands
Embedded Metadata

- The metadata is carried along with the data
- Standards
  - Microformats
  - RDFa
- New forms of documents
Query the Web

- Extract structured information via microformats & RDFa
- Create rich data feeds on the fly

Sun JSP  PHP  ASP.Net
SELECT ?episode, ?chalkboard_gag
WHERE {
?episode skos:subject <Simpsons_episodes_season_12>.
}
## Chalkboard Gags – Episode 12

<table>
<thead>
<tr>
<th>episode</th>
<th>chalkboard_gag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homer vs. Dignity</td>
<td>&quot;I will not surprise the incontinent&quot;</td>
</tr>
<tr>
<td>A Tale of Two Springfields</td>
<td>&quot;I will not plant subliminAL messaGOREs&quot;</td>
</tr>
<tr>
<td>Bye Bye Nerdie</td>
<td>&quot;I will not scare the vice president&quot;</td>
</tr>
<tr>
<td>Children of a Lesser Clod</td>
<td>&quot;Today is not Mothra's Day&quot;</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
GLOBAL
“It's not the documents, it is the things they are about which are important. Obvious, really.”

Tim Berners-Lee
Meaning
What do you **Really** Mean?

- Concept tagging
- Paris
  - City?
  - Person?
- What happens when someone tags their photo with “paris”
Concept Tagging

• “par” into tags box
  – person?
  – city?

- Paris is the capital city of France. It is situated on the River Seine, in northern France, at the heart of the Île-de-France region ("Région parisienne"). Paris has an estimated population of 2,153,600 within its administrative limits. The Paris Unité urbaine (or "urban unit") is an area of unbroken...
Clarity

• Image is tagged with the concept of “Paris the city”

• A lot of things come with that for free
  – Disambiguates the item being tagged, meaning is explicit
  – Adds a wealth of metadata
  – Searches are much more powerful
    • “give me all images that are about cities in European with over 1 million people”
More Powerful, Automatically

• Not a radical re-wiring of existing infrastructure
• *Existing* data, and *existing* connections can be made far more powerful
• A richer and more intuitive way to express oneself
Semantic Web

- There will be many Semantic Webs
- Lives alongside current Web
- Integrating data by giving it more structure
- Flexible data model
Semantic Web Toolkit

- RDF [data]
- OWL [schema]
- SPARQL [query]
- GRDDL [legacy data]
Hacking Climate Change with Drupal and the Semantic Web

Drupalcon 2008, Szeged Hungary

Attribution-Share Alike 2.5 Australia
Image Attribution

- Prawn – http://www.flickr.com/people/hotmeteoor/
- Blender – Intel
- Birds – http://www.flickr.com/people/caitlinburke/
- Islands – http://www.flickr.com/people/mureena/
- Globe - http://www.flickr.com/people/19877537@N00/
- Lamp, Sharks, Spiral – http://www.flickr.com/people/tscarlisle/
- Sharks – http://www.flickr.com/people/psmithson/
- Social Graph 1 – http://www.fidgt.com/
- Social Graph 2 – Cognitive Science Research Unit - Universite Libre de Bruxelles
- Bart – FOX Entertainment
- Drug Discovery 1 & 2 – John Wilbanks, Science Commons