Drupal 7 and the Semantic Web

Stéphane Corlosquet, <scorlosquet@gmail.com>
MIND Informatics
ESIP Federation Webinar
May 16th, 2012
About the speaker

- 6 years with Drupal
- Drupal core maintainer (RDF)
- Drupal Security Team member
- Co-authored the
  *Definitive Guide to Drupal 7*
- Co-maintain RDF Extensions, SPARQL, schema.org
- Member of the RDFa WG
- Software engineer at
  MIND Informatics, Cambridge, MA
Drupal

- Dries Buytaert - small news site in 2000
- Open Source - 2001
- Content Management System
- LAMP stack
- Non-developers can build sites and publish content
- Control panels instead of code

http://www.flickr.com/photos/funkyah/2400889778
Drupal

- Open & modular architecture
- Extensible by modules
- Standards-based
- Low resource hosting
- Scalable
Building a Drupal site

See what's coming up next month in the LEGO Store Monthly Mini Build!

April 2009

Built by: ____________________

Ask a Store Associate for more details or visit www.LEGO.com/LEGOstores.

http://www.flickr.com/photos/toomuchdew/3792159077/
Building a Drupal site

- Create the content types you need:
  Blog, article, wiki, forum, polls, image, video, podcast, e-commerce... (be creative)

http://www.flickr.com/photos/georgivar/4795856532/
Building a Drupal site

• Enable the features you want:

  Comments, tags, voting/rating, location, translations, revisions, search...

http://www.flickr.com/photos/skip/42288941/
Building a Drupal site

6000 free contributed modules

- Google Analytics
- Wysiwyg
- Captcha
- Calendar
- XML sitemap
- Five stars
- Twitter
- ...

http://www.flickr.com/photos/kaptainkobold/1422600992/
The Drupal Community

http://www.flickr.com/photos/x-foto/4923221504/
“It’s really the Drupal community and not so much the software that makes the Drupal project what it is. So fostering the Drupal community is actually more important than just managing the code base.” - Dries Buytaert
Who uses Drupal?
Who uses Drupal?
Who uses Drupal?
Who uses Drupal?

http://buytaert.net/tag/drupal-sites
Try Drupal 7

• Download and Install Drupal 7
  • Grab latest release http://drupal.org/project/drupal
  • Mac OS: MAMP http://www.mamp.info/
  • Acquia Stack http://acquia.com/downloads

• Drupal Gardens: free Drupal 7 site
  http://www.drupalgardens.com/
Drupal 7 and RDF
History of RDF in Drupal

- rdf.php (2000, Dries)
- FOAF, vCard (2004, walkah)
- Relationship (2005, dman)
- Semantic Search (2006, hendler)
- RDF (2007, Arto)
- OpenCalais (febbraro, 2008)
- RDF CCK (2008, scor)
- and more...
Experiment prior Drupal 7

- Expose Drupal's data in RDF
- Automatically
- No RDF expertise required
- Module based

http://www.flickr.com/photos/neys/238686518
Architecture

- User driven data model
- Content type = RDF class
- Field = RDF property
- Node = RDF resource
Content types and Fields

Add fields and groups to the content type, and arrange them on content display.
You can add a field to a group by dragging it below and to the right of the group.
Note: Installing the Advanced help module will let you access more and better

<table>
<thead>
<tr>
<th>Label</th>
<th>Name</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Node module form.</td>
<td>Text</td>
</tr>
<tr>
<td>First name</td>
<td>field_fn</td>
<td>Text</td>
</tr>
<tr>
<td>Last name</td>
<td>field_ln</td>
<td>Text</td>
</tr>
<tr>
<td>Picture</td>
<td>field_picture</td>
<td>Image</td>
</tr>
<tr>
<td>Colleagues</td>
<td>field_friend</td>
<td>Node reference</td>
</tr>
<tr>
<td>Current projects</td>
<td>field_current_project</td>
<td>Node reference</td>
</tr>
</tbody>
</table>
Content types and Fields

Global settings
These settings apply to the Gender field in every content type in which it appears.

- Required

Number of values:
1
Maximum number of values users can enter for this field.
Warning! Changing this setting after data has been created could result in the loss of data!

Allowed values
Create a list of options as a list in Allowed values list or as an array in PHP code. These values will be the same for Gender in all content types.

Allowed values list:
- Male
- Female
Node

Stephane Corlosquet

First name: Stephane
Last name: Corlosquet
Homepage: http://openspring.net/
Email: stephane.corlosquet@deri.org
        scorlosquet@gmail.com
Colleagues: Aidan Hogan
          Andreas Harth
          Axel Polleres
          Tim Clark
Current projects: Drupal
                Drupal Semantics
                Science Collaboration Framework
Mapping to RDF

Specify the RDF class of this content type. You can also map the CCK fields to RDF properties.

**RDF class:**

- person

- foaf:Person

- <http://xmlns.com/foaf/0.1/Person> Person* A person.
- <http://ebiquity.umbc.edu/ontology/person.owl#Person> Person
- <http://rs.tdwg.org/ontology/voc/Person#Person> Person* A representation of a human being, living or dead.
- <http://www.cyc.com/2003/04/01/cyc#Person> persons* The collection of all human beings. Something is an instance of #$Person just in case it is a member of the species Homo Sapiens, and thus #$Person is an instance of #$BiologicalSpecies
## Mapping to RDF

<table>
<thead>
<tr>
<th>Label</th>
<th>Name</th>
<th>Type</th>
<th>RDF property</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last name</td>
<td>field_last</td>
<td>Text</td>
<td>foaf:surname</td>
</tr>
<tr>
<td>Picture</td>
<td>field_picture</td>
<td>Image</td>
<td>foaf:depiction</td>
</tr>
<tr>
<td>URI</td>
<td>field_myuri</td>
<td>Link</td>
<td>owl:sameAs</td>
</tr>
<tr>
<td>Homepage</td>
<td>field_homepage</td>
<td>Link</td>
<td>foaf:homepage</td>
</tr>
<tr>
<td>Email</td>
<td>field_email</td>
<td>Text</td>
<td>foaf:mbox</td>
</tr>
<tr>
<td>Colleagues</td>
<td>field_friend</td>
<td>Node reference</td>
<td>relationship works</td>
</tr>
<tr>
<td>Resume</td>
<td>field_resume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blog</td>
<td>field_blog</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current projects</td>
<td>field_current</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past projects</td>
<td>field_past</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>field_gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organizations</td>
<td>field_organization</td>
<td>Node reference</td>
<td>foaf:member</td>
</tr>
<tr>
<td>Posts</td>
<td>field_post</td>
<td>Node reference</td>
<td>foaf:made</td>
</tr>
</tbody>
</table>
Experiment continued

- 1. Automatic site vocabulary generation
- 2. Mapping Content Models to existing ontologies
- 3. Data endpoint for SPARQL querying
- 4. Lazy loading of external data (data import)
Drupal 7 and RDF

- Drupal 7 core is RDFa enabled!
- RDFa output by default on blogs, forums, comments, etc. using FOAF, SIOC, DC, SKOS
Rich Snippets
Google search results for 'banana bread recipes'.

1. Banana Banana Bread Recipe - Allrecipes.com
   - 6501 reviews - 1 hr 20 mins - 229 cal
   - Why compromise the banana flavor? This banana bread is moist and delicious with loads of banana flavor! Friends and family love my recipe and say it's by far...

2. Banana Bread Recipes - Allrecipes.com
   - 1327 reviews - 1 hr 10 mins - 218 cal
   - Looking for the perfect banana bread recipe? Allrecipes has more than 120...

   - 1 hr 5 mins
   - Jun 15, 2006 – Moist and delicious classic banana bread recipe. Easy to make, no need for a mixer. Ripe bananas, butter, sugar, egg, vanilla, baking soda, ...

4. Banana Bread Recipe | Simply Recipes
   - 487 reviews - 3 hrs 10 mins
   - 3 days ago – Food Network invites you to try this banana bread recipe.
Avatar - Rotten Tomatoes
It might be more impressive on a technical level than as a piece of storytelling, but Avatar reaffirms James Cameron's singular gift for imaginative, absorbing ... www.rottentomatoes.com/m/avatar

Avatar :: rogerebert.com :: Reviews

Avatar Review
Avatar Review Poetry :: Prose :: Reviews :: Interviews :: Art. Home; About; Submissions; Index of Contributors; Current Issue; AV13: Summer 2011; Past Issues. AV12: Summer 2010 www.avatarreview.net

Avatar Review :: Movies :: Reviews :: Paste
Never before have so many cynics wanted to hate a movie, and never have they been so thoroughly thwarted, as with James Cameron's Avatar. I had more than my share of ... www.pastemagazine.com/articles/2009/12/avatar-review.html

Avatar Reviews, Ratings, Credits, and More at Metacritic
Expert score: 83/100 · PG-13 · 162 min · 35 reviews
Jake Sully is a former Marine confined to a wheelchair. But despite his broken body, Jake is still a warrior at heart. He is recruited to travel light years to the ... www.metacritic.com/redirrectcritic?m=avatar
Why Structured Data in HTML

• Help machines extract relevant data from HTML
• Can make use of this data in amazing ways (e.g. enhanced search results)
Structured Data in HTML

- Add or alter HTML attributes
- Syntaxes
  - Microformats (@class, @rel)
  - RDFa (@property, @about, @typeof, ...)
  - Microdata (@itemscope, @itemtype, @itemprop, ...)
  - RDFa Lite
What is Schema.org?

This site provides a collection of schemas, i.e., html tags, that webmasters can use to markup their pages in ways recognized by major search providers. Search engines including Bing, Google and Yahoo! rely on this markup to improve the display of search results, making it easier for people to find the right web pages.

Many sites are generated from structured data, which is often stored in databases. When this data is formatted into HTML, it becomes very difficult to recover the original structured data. Many applications, especially search engines, can benefit greatly from direct access to this structured data. On-page markup enables search engines to understand the information on web pages and provide richer search results in order to make it easier for users to find relevant information on the web. Markup can also enable new tools and applications that make use of the structure.

A shared markup vocabulary makes easier for webmasters to decide on a markup schema and get the maximum benefit for their efforts. So, in the spirit of sitemaps.org, Bing, Google and Yahoo! have come together to provide a shared collection of schemas that webmasters can use.

We invite you to get started!

New! View our blog at blog.schema.org.
Schema.org

• Describe the type of your content (Person, Event, Recipe, Product, Book, Movie, etc.)
  – 290 types and counting

• Each type has a set of properties
  – Common properties: name, description, image, url
  – Specific properties depending on the type (see type page on schema.org)
  – 240 properties and counting
## Thing > Event

An event happening at a certain time at a certain location.

<table>
<thead>
<tr>
<th>Property</th>
<th>Expected Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Properties from</strong> Thing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>description</td>
<td>Text</td>
<td>A short description of the item.</td>
</tr>
<tr>
<td>image</td>
<td>URL</td>
<td>URL of an image of the item.</td>
</tr>
<tr>
<td>name</td>
<td>Text</td>
<td>The name of the item.</td>
</tr>
<tr>
<td>url</td>
<td>URL</td>
<td>URL of the item.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Property</th>
<th>Expected Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Properties from</strong> Event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>attendees</td>
<td>Person or Organization</td>
<td>A person attending the event.</td>
</tr>
<tr>
<td>duration</td>
<td>Duration</td>
<td>The duration of the item (movie, audio recording, event, etc.) in <strong>ISO 8601 date format</strong>.</td>
</tr>
<tr>
<td>endDate</td>
<td>Date</td>
<td>The end date and time of the event (in <strong>ISO 8601 date format</strong>).</td>
</tr>
<tr>
<td>location</td>
<td>Place or PostalAddress</td>
<td>The location of the event or organization.</td>
</tr>
<tr>
<td>offers</td>
<td>Offer</td>
<td>An offer to sell this item—for example, an offer to sell a product, the DVD of a movie, or a CD of a music album.</td>
</tr>
<tr>
<td>performers</td>
<td>Person or Organization</td>
<td>The main performer or performers of the event—for example, a presenter, musician, or actors.</td>
</tr>
<tr>
<td>startDate</td>
<td>Date</td>
<td>The start date and time of the event (in <strong>ISO 8601 date format</strong>).</td>
</tr>
</tbody>
</table>
Schema.org module for Drupal

- UI instead of code
- Map your content types and fields to the schema.org terms

http://drupal.org/project/schemaorg
Example: Event

BAD CAMP 2011
Bay Area Drupal Camp

October 21 to 23, 2011
UC Berkeley, California
A culmination of the brightest Drupal minds in the technology hub of the world

BADcamp 2011
The Bay Area Drupal camp is the largest FREE Drupal event in the world.

Come join us October 21st - 23rd at UC Berkeley, in California.

Schedule of events

OCTOBER 21
Free Drupal Training
Core Developer Summit
Higher Education Summit
Non-Profit Summit

OCTOBER 22
BADcamp Sessions & BoFs
An Evening of Drupal

Register
Build your schedule
Rich Snippet testing tool

- [http://www.google.com/webmasters/tools/richsnippets](http://www.google.com/webmasters/tools/richsnippets)
Examples in the wild

• schema.org sandbox
  – Dries: http://goo.gl/Ztx90
  – Apple pie recipe: http://goo.gl/yc9qv

• definitivedrupal.org
  – DrupalCon Munich: http://goo.gl/jgMvw
  – List of events: http://goo.gl/SJBXS
Examples in the wild

• University of Waterloo
  - School of Public Health and Health Systems launch:
    http://goo.gl/Df9hp

• Curling tournament calendar
  - European Curling Championships 2012:
    http://goo.gl/YXgXl
  - World Women’s Curling Championships 2013:
    http://goo.gl/BDNZW
Thanks!

- [http://drupal.org/project/schemaorg](http://drupal.org/project/schemaorg)
  - Download module (beta)
  - Documentation on drupal.org
  - Screencast + examples
- Make use of existing HTML structured data
  - sindice.com
  - [www.commoncrawl.org](http://www.commoncrawl.org) (video)