

# The future of search and SEO

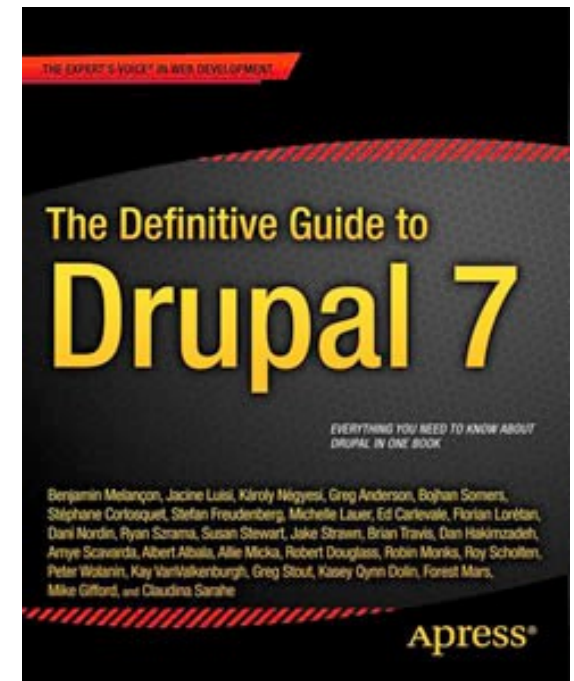


Stéphane Corlosquet & Nick Veenhof  
BADCamp, October 2013  
Berkeley, CA

**Acquia**™

# About the speakers

- Stéphane “scor” Corlosquet
  - 7 years with Drupal
  - Software engineer @ Acquia
  - Drupal 7 RDF core maintainer
  - Drupal Security Team member
  - Co-authored the **Definitive Guide to Drupal 7**
  - Contrib modules: RDF Extensions, SPARQL, schema.org, WebID
  - Member of the RDFa WG at W3C



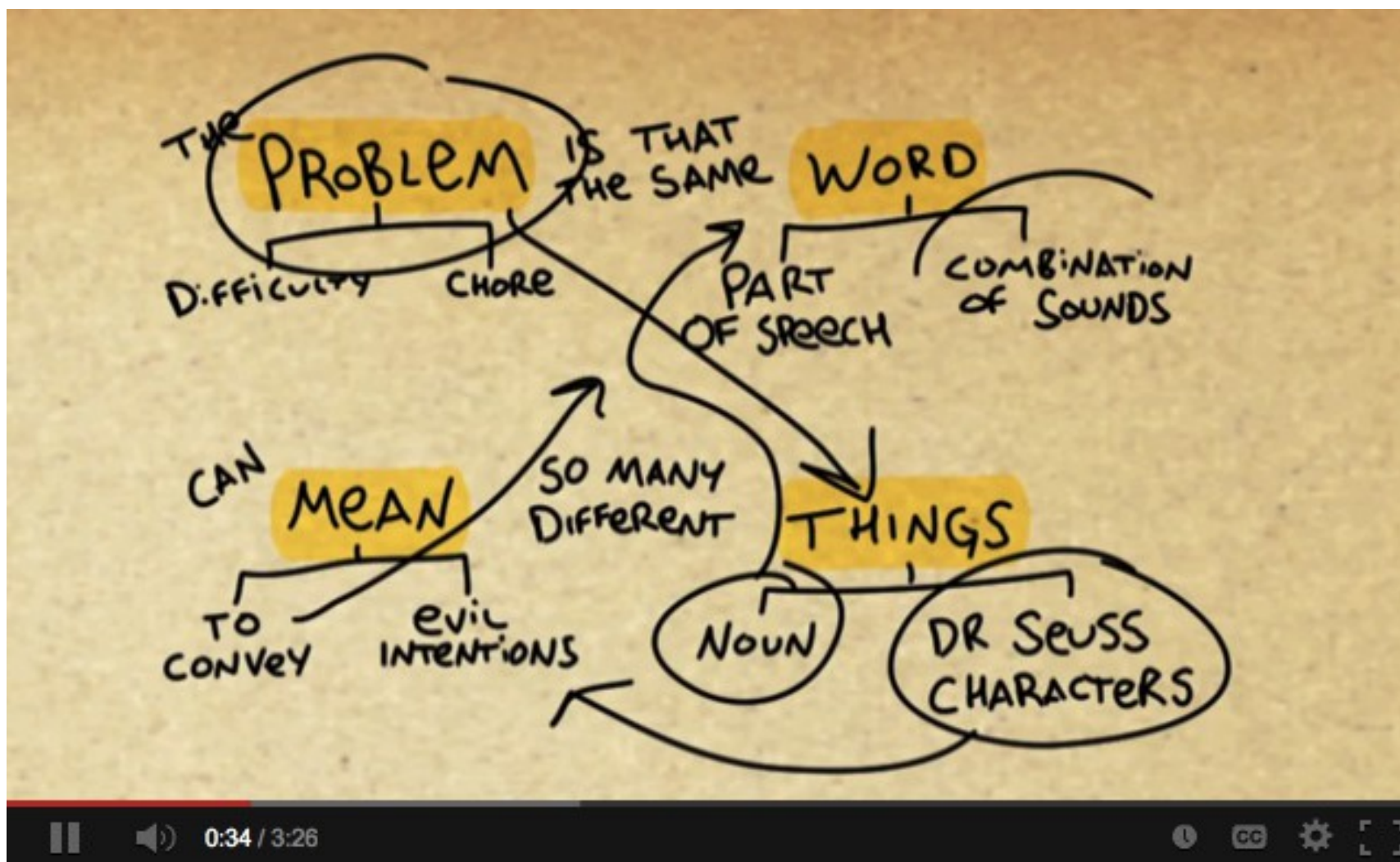
# About the speakers

- Nick “Nick\_vh” Veenhof
  - 6 years with Drupal
  - Search engineer @ Acquia
  - Drupal 7 Apache Solr maintainer
  - Very much involved in the community across North America and Europe

Acquia™



# The problem



Speak now

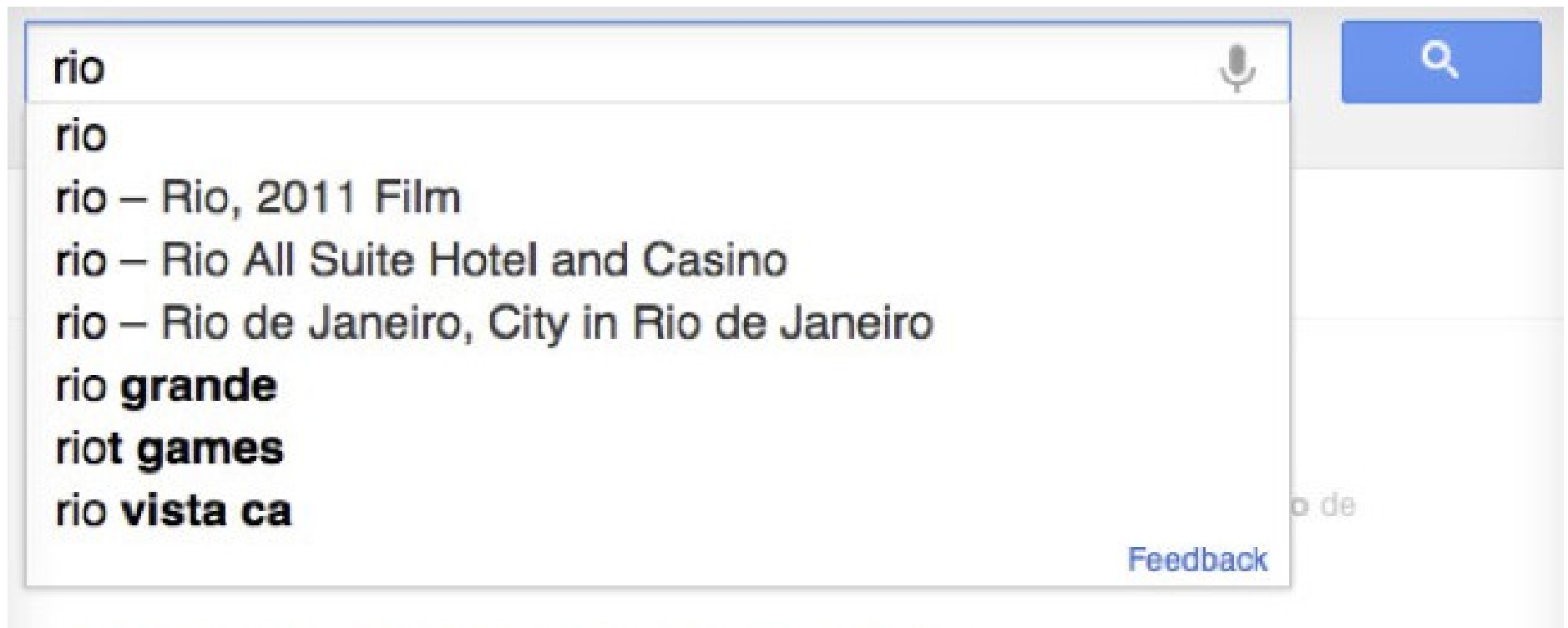


# Knowledge Graph

- Google's internal knowledge base
- Helps user find information quickly + easily
- Knows about people, places, cities, celebrities, movies, work or art...
- “Things, not strings” - no keyword matching
- Hummingbird algorithm update (largest update since 2001)

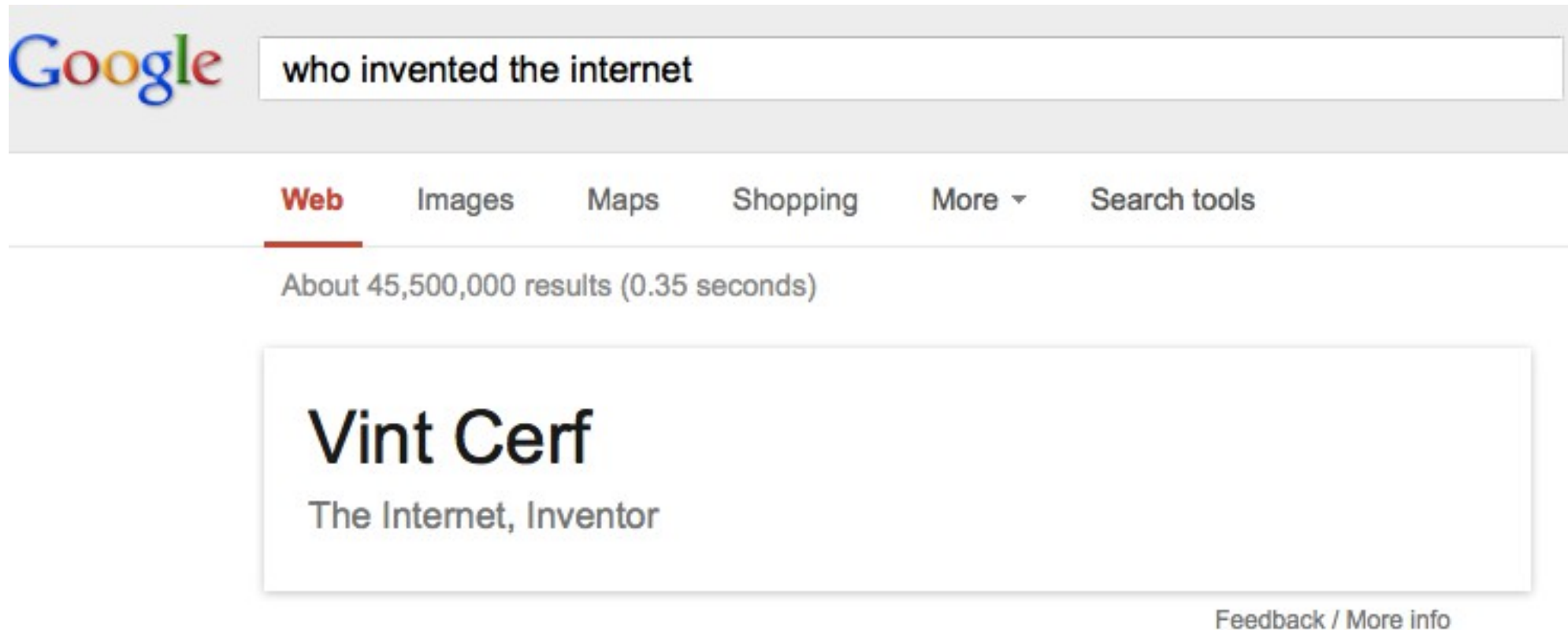
# Knowledge Graph

- Live Entity disambiguation



# Knowledge Graph

- Direct answers



The image is a screenshot of a Google search interface. At the top left is the Google logo. To its right is a search bar containing the text "who invented the internet". Below the search bar is a horizontal menu with tabs: "Web" (which is underlined in red), "Images", "Maps", "Shopping", "More" (with a downward arrow), and "Search tools". Below the menu, it says "About 45,500,000 results (0.35 seconds)". The main content area features a large white box with a light gray border. Inside this box, the name "Vint Cerf" is displayed in a large, bold, black font. Below the name, the text "The Internet, Inventor" is shown in a smaller, gray font. At the bottom right of the white box, there is a link that says "Feedback / More info".

Google

who invented the internet

**Web** Images Maps Shopping More ▾ Search tools

About 45,500,000 results (0.35 seconds)

**Vint Cerf**

The Internet, Inventor

[Feedback / More info](#)










# Knowledge Graph

- List and collections

events in Boston

Web Images Maps Shopping News More Search tools

## Boston events

						
<b>Paul McCartney</b> Jul 9	<b>CMA Songwriters Series Featurin...</b> Jul 11	<b>Harborwalk Sounds 2013...</b> Jul 11	<b>The Brooklyn Brewery Mash...</b> Jul 9	<b>Belle and Sebastian</b> Jul 9	<b>Fort Point Fitness: Sunset...</b> Jul 16	<b>A Midsummer's Night With the...</b> Jul 16

# Knowledge Graph

- Entity summary

+Dave Search Images Maps Play YouTube News Gmail Documents Calendar More

Google

marie curie

8 personal results, 17,300,000 other results (0.32 seconds)

Everything

Images

Maps

Videos

News

Shopping

Books

More

Mountain View, CA

Any time

Past hour

Past 24 hours

Past 2 days

Past week

Past month

Past year

Custom range...

Marie Curie - Wikipedia, the free encyclopedia

en.wikipedia.org/wiki/Marie\_Curie

Marie Skłodowska-Curie (7 November 1867 – 4 July 1934) was a French-Polish physicist and chemist famous for her pioneering research on radioactivity.

→ Marie Curie Cancer Care - Pierre Curie - Irène Joliot-Curie - Radioactive decay

Marie Curie - Biography - Nobelprize.org

www.nobelprize.org/nobel\_prizes/physics/.../marie-curie-bio.html


Short profile from the foundation that awards the Nobel Prize.

Marie Curie - Biography - Nobelprize.org

www.nobelprize.org/nobel\_prizes/chemistry/.../marie-curie-bio.html

Marie Curie, née Maria Skłodowska, was born in Warsaw on November 7, 1867, the daughter of a secondary-school teacher. She received a general education ...

Images for marie curie - Report images




Marie Curie and The Science of Radioactivity

www.aip.org/history/curie/

The life of Marie Curie, from the AIP Center for History of Physics. Text by Naomi Pasachoff and many illustrations describe Curie's contributions to the science of ...

Marie Curie



es.wikipedia.org

Marie Skłodowska-Curie was a French-Polish physicist and chemist famous for her pioneering research on radioactivity. She was the first person honored with two Nobel Prizes—in physics and chemistry. Wikipedia

Born: November 7, 1867, Warsaw

Died: July 4, 1934, Sancellemoz


Spouse: Pierre Curie (m. 1895–1906)

Children: Irène Joliot-Curie, Ève Curie


Discovered: Radium, Polonium

Education: École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris, University of Paris


People also search for




Albert Einstein




Pierre Curie



Ernest Rutherford



Louis Pasteur






John Dalton

Report a problem

# Knowledge Graph

- Rich Snippets

[Advanced search](#)

SearchAbout 5,210,000 results (0.40 seconds)

Everything

Images

Maps

Videos


News

Shopping

Recipes

More

San Francisco, CA  
Change location



★★★★★

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 ...

Banana Bread I Recipe · Brown Sugar Banana Nut Bread I · Nutty Banana Bread

Banana Bread Recipes - Allrecipes.com

allrecipes.com > Recipes > Bread - Cached

Looking for the perfect **banana bread recipe**? Allrecipes has more than 120 ...



★★★★★

1327 reviews · 1 hr 10 mins · 218 cal


Sour cream guarantees a moist and tender **loaf**. And **bananas** are sliced ...

+

Show more results from allrecipes.com

Banana Bread Recipe | Simply Recipes

simplyrecipes.com/recipes/banana\_bread/ - Cached



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, ...

Banana Nut Muffins · Orange Bread · Glazed Lemon Bread

Ingredients

	Yes	No
bananas	<input type="checkbox"/>	<input type="checkbox"/>
baking soda	<input type="checkbox"/>	<input type="checkbox"/>
walnuts	<input type="checkbox"/>	<input type="checkbox"/>
pecans	<input type="checkbox"/>	<input type="checkbox"/>
vanilla extract	<input type="checkbox"/>	<input type="checkbox"/>

# Knowledge Graph

- 570 million real-world objects
- 18 billion facts and relationships
- Built from content gathered from the Web
- **Freebase**
  - Acquired by Google in 2010
  - 39 Million topics
  - 1,9 Billion facts and relationships
- Wikipedia
- More? → **Video** and **video**



# Other search engines similar efforts

- Yahoo!

**YAHOO!**

18,600 results


WEB IMAGES VIDEO BLOGS MORE ▾

**FILTER BY TIME**

- Anytime
- Past day
- Past week
- Past month

**Cascal**  
[cascalrestaurant.com](http://cascalrestaurant.com)  
(650) 940-9500  
400 Castro St, Mountain View, CA ▾  
[Menu](#) | [More Info »](#)

★★★★★ (65)  
65 Reviews - I am so sick of people suggesting Cascal who have never been there. Don't let the decor fool you - that was done by someone with taste... [more](#)



**Cascal - Welcome to the Journey**  
[www.cascalrestaurant.com](http://www.cascalrestaurant.com)

[Tapas](#)      [Special](#)  
[Lunch](#)      [Contact](#)  
[Music And](#)

[More results from cascalrestaurant.com »](#)

**Cascal - Mountain View, CA**  
★★★★★ (1237 Reviews) - (650) 940-9500 - Mountain View, CA  
(650) 940-9500 · 400 Castro St · "The prosciutto wrapped dates and mushroom empanadas are amazing."  
[www.yelp.com/biz/cascal-mountain-view](http://www.yelp.com/biz/cascal-mountain-view) - [Cached](#)  
[More results from yelp.com »](#)

**Cascal Restaurant - Contact - Cascal - Welcome to the Journey**  
Cascal is centrally located at the corner of California and Castro streets in the heart of downtown Mountain View, California.  
[www.cascalrestaurant.com/about\\_contact.asp](http://www.cascalrestaurant.com/about_contact.asp) - [Cached](#)

# Other search engines similar efforts

- Bing - Satori

The image is a screenshot of a Bing search results page for the query "avatar reviews". The Bing logo is in the top left, and the search bar contains "avatar reviews". Below the search bar are tabs for "Web", "Videos", "Movies", and "More". The "Web" tab is selected. On the left side, there are sections for "RELATED SEARCHES" and "SEARCH HISTORY". The "RELATED SEARCHES" section lists various search terms related to Avatar, including "Avatar the Last Airbender Review", "YouTube Avatar Review", "Movie Avatar Reviews", "Lego Avatar Review", "Avatar Review Film", "Avatar Review 3D", "Avatar Review Yahoo!", "Avatar Review Christian", and "Avatar Review". The "SEARCH HISTORY" section lists "dirt devil vacuum", "avatar", "avatar movie", "imdb", and "Similar searches avatar". The main content area shows "ALL RESULTS" for "1-10 of 37,900,000 results". The first result is "Avatar - Rotten Tomatoes", which includes a snippet about the movie's technical level and a link to the Rotten Tomatoes page. The second result is "Avatar :: rogerebert.com :: Reviews", which includes a snippet about the movie's cast and a link to the Roger Ebert review. The third result is "Avatar Review", which includes a snippet about the movie's poetry and a link to the Avatar Review website. The fourth result is "Avatar Review :: Movies :: Reviews :: Paste", which includes a snippet about the movie's cynics and a link to the Paste magazine review. The fifth result is "Avatar Reviews, Ratings, Credits, and More at Metacritic".

**bing**  
Web

avatar reviews

Web Videos Movies More▼

RELATED SEARCHES

- Avatar the Last Airbender Review
- YouTube Avatar Review
- Movie Avatar Reviews
- Lego Avatar Review
- Avatar Review Film
- Avatar Review 3D
- Avatar Review Yahoo!
- Avatar Review Christian

SEARCH HISTORY

- dirt devil vacuum
- avatar
- avatar movie
- imdb
- Similar searches avatar

ALL RESULTS

1-10 of 37,900,000 results · [Advanced](#)

**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](http://www.rottentomatoes.com/m/avatar)

**Avatar :: rogerebert.com :: Reviews**  
Jake Sully Sam Worthington Neytiri Zoe Saldana Grace Sigourney Weaver Col. Miles Quaritch Stephen Lang Trudy Chacon Michelle Rodriguez  
[rogerebert.suntimes.com/apps/pbcs.dll/article?AID=/20091211/REVIEWS/912119998](http://rogerebert.suntimes.com/apps/pbcs.dll/article?AID=/20091211/REVIEWS/912119998)

**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](http://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](http://www.pastemagazine.com/articles/2009/12/avatar-review.html)

**Avatar Reviews, Ratings, Credits, and More at Metacritic**

**Rotten Tomatoes**  
83%  
279 reviews

# What about Drupal?

- Drupal sites hold a lot of the data on the Web
  - > 1 Million sites
  - 2% of the web
  - Variety of content: news, gov, corporate, media, commerce
- Drupal's data is structured inside
  - Entities: Node, User, Comment, Taxonomy term
  - Properties: fields (text, entity reference...)

# Challenge

How can you make Drupal data easily understandable by search engines?



# Schema.org standard

- General web schema for everyone to use
- 400+ types (book, recipe, event, person...)
- Sponsored and supported by major search engines: Google, Yahoo!, Bing, Yandex
- Multiple formats: RDFa, Microdata, JSON-LD

## Thing > Event

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

Property	Expected Type	Description
Properties from <a href="#">Thing</a>		
<a href="#">description</a>	Text	A short description of the item.
<a href="#">image</a>	URL	URL of an image of the item.
<a href="#">name</a>	Text	The name of the item.
<a href="#">url</a>	URL	URL of the item.
Properties from <a href="#">Event</a>		
<a href="#">attendees</a>	<a href="#">Person</a> or <a href="#">Organization</a>	A person attending the event.
<a href="#">duration</a>	Duration	The duration of the item (movie, audio recording, event, etc.) in <a href="#">ISO 8601 date format</a> .
<a href="#">endDate</a>	Date	The end date and time of the event (in <a href="#">ISO 8601 date format</a> ).
<a href="#">location</a>	<a href="#">Place</a> or <a href="#">PostalAddress</a>	The location of the event or organization.
<a href="#">offers</a>	<a href="#">Offer</a>	An offer to sell this item—for example, an offer to sell a product, the DVD of a movie, or
<a href="#">performers</a>	<a href="#">Person</a> or <a href="#">Organization</a>	The main performer or performers of the event—for example, a presenter, musician, or
<a href="#">startDate</a>	Date	The start date and time of the event (in <a href="#">ISO 8601 date format</a> ).

# How does schema.org apply to Drupal?



Breaking News Story

## Content Type

---

## Fields

### Article

---

- Title
- Author
- Published

### Event

---

- Name
- Date
- Location

### Related Stories

- Title
- Author
- Link

# Schema.org module for Drupal

- Map your content types and fields to the schema.org terms

The screenshot shows the configuration page for a content type in the Schema.org module. On the left is a sidebar with settings tabs: Submission form settings, Publishing options, Display settings, Comment settings, Menu settings, and Schema.org settings. The main area is titled 'Type' and contains a text input with 'art' typed in. Below the input is a list of schema.org types: ApartmentComplex, ArtGallery, Article, AutoPartsStore (highlighted with a mouse cursor), DepartmentStore, NewsArticle, PerformingArtsTheater, ScholarlyArticle, and VisualArtsEvent. At the bottom are two buttons: 'Save content type' and 'Delete content type'.

Settings	Type
<b>Submission form settings</b> Title	art
<b>Publishing options</b> Published , Promoted to front page	ApartmentComplex
<b>Display settings</b> Display author and date information.	ArtGallery
<b>Comment settings</b> Open, Threading , 50 comments per page	Article
<b>Menu settings</b>	AutoPartsStore
<b>Schema.org settings</b>	DepartmentStore
	NewsArticle
	PerformingArtsTheater
	ScholarlyArticle
	VisualArtsEvent

Save content type    Delete content type

# Schema.org module for Drupal

- Map your content types and fields to the schema.org terms

The screenshot shows the Drupal administration interface for the Schema.org module. The breadcrumb trail is: Home » Administration » Structure » Content types » Event » Manage fields. The page title is 'Location'. There are four buttons: 'EDIT', 'FIELD SETTINGS', 'WIDGET TYPE', and 'DELETE'. The 'FIELD SETTINGS' tab is active, showing 'EVENT SETTINGS' for the 'Location' field. The settings include: 'Label' (Location), 'Required field' (unchecked), 'Help text' (empty), 'Instructions to present to the user below this field on the editing form.' (empty), 'Allowed HTML tags' (a list of tags), 'Size of textfield' (60), 'Text processing' (Plain text selected), and 'DEFAULT VALUE' (The default value for this field, used when creating new content).

Dashboard Content **Structure** Appearance People Modules Configuration Reports Help Hello admin Log out

Home » Administration » Structure » Content types » Event » Manage fields

Location o EDIT FIELD SETTINGS WIDGET TYPE DELETE

**EVENT SETTINGS**

These settings apply only to the Location field when used in the Event type.

**Label \***

Location

☐ Required field

**Help text**

Instructions to present to the user below this field on the editing form.  
Allowed HTML tags: <a> <b> <big> <code> <del> <em> <i> <ins> <pre> <q> <small> <span> <strong> <sub> <sup> <tt> <ol> <ul> <li> <p> <br> <img>

**Size of textfield \***

60

**Text processing**

☒ Plain text

☐ Filtered text (user selects text format)

**DEFAULT VALUE**

The default value for this field, used when creating new content

<http://drupal.org/project/schemaorg>

# Schema.org module for Drupal

- Map your content types and fields to the schema.org terms

The screenshot shows the Drupal administration interface for mapping a field to a Schema.org term. The breadcrumb trail is: Home » Administration » Structure » Content types » Event » Manage fields. The field being edited is 'Location'. The main heading is 'LOCATION SCHEMA.ORG MAPPING'. Under the 'Property' section, the text 'locatio' is entered in a text box. A dropdown menu is open, showing the following options: 'contentLocation', 'homeLocation', 'jobLocation', 'location', and 'workLocation'. To the right of the dropdown, the text 'Id.' is visible. Below the dropdown, there is a text area with the placeholder text '<p> <br> <img>'. At the bottom, there is a 'DEFAULT VALUE' section with the text 'The default value for this field, used when creating new content'.

<http://drupal.org/project/schemaorg>

# Schema.org mappings in HTML

```
<div property="schema:location">Wellington, New Zealand </div>
```

# Preview Rich Snippet

- <http://www.google.com/webmasters/tools/richsnippets>

## Structured Data Testing Tool

URL

HTML

<http://rdfsolr13test.devcloud.acquia-sites.com/node/6>

PREVIEW

Sign in to view the retrieved HTML.

Google search results

Google Custom Search

## Preview

[DrupalSouth New Zealand | rdfsolr13dev.devcloud.acquia-sites.com](#)

[rdfsolr13test.devcloud.acquia-sites.com/node/6](#)

Feb 14, 2014 - Wellington, New Zealand

The excerpt from the page will show up here. The reason we can't show text from your webpage is because the text depends on the query the user types.



# Examples

- Recipe preview
- Event preview
- Person preview

**Can I haz Rich  
Snippetz...**

**... on my Drupal site  
too?**



# rich\_snippets.module

Yes!

## Belgian Fries

[rdfsolr13test.devcloud.acquia-sites.com/node/26](http://rdfsolr13test.devcloud.acquia-sites.com/node/26)



Total cook time: 25 mins

Summary Yield Servings Source <http://allreci>

15 minutes Cooking time 10 minutes Total tin

Russet Potatoes (Peeled, rinsed and dried) 1 t Salt (Ad

# RDF support in Rich Snippet



**Rich Snippets**



**Rich Snippets Event**

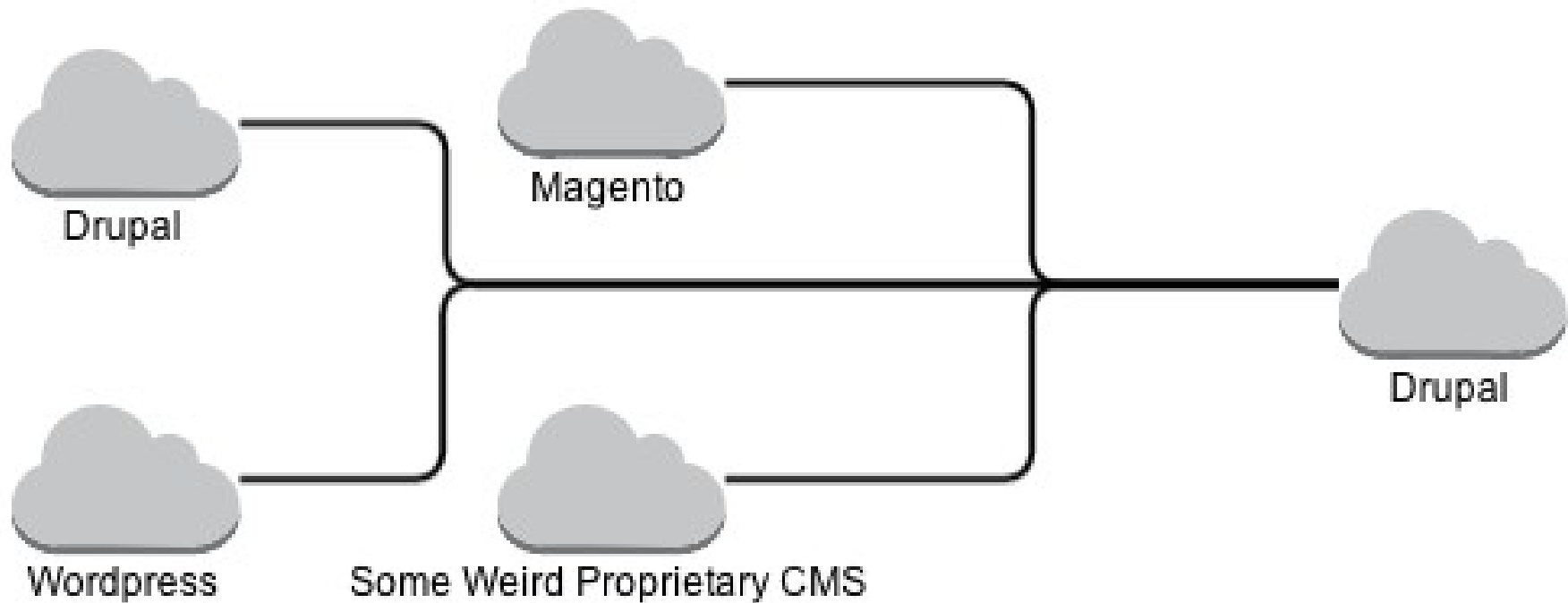


**Rich Snippets Person**



**Rich Snippets Recipe**

# Multisite search + Rich Snippets



# Without Rich Snippet module

---

## Belgian Fries

Summary Yield Servings Source <http://allrecipes.com/recipe/apple-pie-by-grandma-ople/> Prep time 15 minutes Cooking time 10 minutes Total time 25 minutes Description Belgian Fries Ingredients 2 lb Russet Potatoes (Peeled, rinsed and dried) 1 t Salt (Add to ...

admin - 10/23/2013 - 13:20 - This site (rdfsolr13dev.devcloud.acquia-sites.com)

---

## BADCamp 2013

Saturday, October 26, 2013 Berkeley, CA The format of BADCamp is a mix of presentations, workshops, and networking. The focus is on community knowledge, sharing and ...

admin - 10/22/2013 - 09:38 - rdfsolr13dev.devcloud.acquia-sites.com

**These are really from two different sites, but they are a copy of each other. Not very clear is it?**

# With Rich Snippet module

---

## [Belgian Fries](#)

[rdfsolr13test.devcloud.acquia-sites.com/node/26](#)



Total cook time: 25 mins

Summary Yield Servings Source <http://allrecipes.com/recipe/apple-pie-by-grandma-ople/> Prep time

15 minutes Cooking time 10 minutes Total time 25 minutes Description Belgian Fries Ingredients 2 lb

Russet Potatoes (Peeled, rinsed and dried) 1 t Salt (Add to ...

**Same results, rich\_snippets enabled**

## [BADCamp 2013](#)

[rdfsolr13dev.devcloud.acquia-sites.com/node/1](#)

Sat, Oct 26, 2013 - Berkeley, CA

Saturday, October 26, 2013 Berkeley, CA The format of BADCamp is an "un-conference" not unlike

DrupalCampNY, DrupalCampLA, DrupalCampSeattle and BarCamp (the granddaddy of them all). The emphasis is on community knowledge, sharing and ...

# Benefits of Rich Snippets

- Same look and feel as Google search for private content
- Schema.org mappings used for Google search and internal Solr search
- Mix Drupal and non-Drupal data



# Modules used

- [Apache Solr Search](#)
- [Apache Solr Multisite Search](#)
- [Schema.org](#)
- [Rich Snippets](#)
- Blog post: [Bridge Drupal and Non Drupal content with RDF and Solr](#)

# Thanks!

Stéphane Corlosquet:

- [scorlosquet@gmail.com](mailto:scorlosquet@gmail.com)
- [@scorlosquet](#)
- <http://openspring.net/>

• Nick Veenhof:

- [nick.veenhof@gmail.com](mailto:nick.veenhof@gmail.com)
- [@nick\\_vh](#)
- <http://veenhof.be/>