

Tema 4. Tecnologías para la web 3.0

Segunda Parte: Mercado Semántico

J. Javier Samper Zapater
José Fco. García Calderaro
Departamento de Informática

VNIVERSITAT  ID VALÈNCIA

Marcado semántico

Introducción

Microdatos

Microformatos

RDFa: Introducir semántica en los documentos XHTML.

GRDDL: Obtención de RDF desde XML, XHTML documentos.

Introducción

- La web 3.0 o web semántica, en esencia, consiste en conseguir que las máquinas entiendan qué significan los contenidos de tal manera que luego sea más fácil localizarlos e interpretarlos.
- Así, por ejemplo, si Flickr reconoce el lugar dónde se ha sacado una fotografía, puede mostrarla acompañada de un mapa que aporte información complementaria al usuario o, por poner otro caso, si un buscador como Google reconoce al autor de una web, puede mostrar esa información durante una búsqueda.

Introducción (II)

- Para conseguir eso, se han incorporado algunas etiquetas especiales a la última versión del lenguaje HTML, como la etiqueta `<header>` que indica que un conjunto de elementos forman la cabecera de una web. Sin embargo, no son suficientes para describir la tremenda variedad de contenidos que hay en la red, así que se están lanzando otras propuestas para el etiquetado semántico de las webs.
- Las etiquetas semánticas son unos tags especiales que definen significados. Por ejemplo, este podría ser un párrafo normal en HTML, marcado tan solo por la etiqueta `<p>` que define los párrafos:

`<p>Nero Wolf experto en cultivo de orquídeas</p>`

- Pero se le podrían añadir etiquetas semánticas para diferenciar el nombre de la profesión:

`<p><nombre>Nero Wolf</nombre>,</p>`

`<profesión>experto en cultivo de orquídeas</profesión></p>`

Introducción (III)

- En Google lo llaman "fragmentos enriquecidos", aunque en inglés queda más elegante: rich snippets. Nosotros lo conocemos como "etiquetado semántico", aún cuando en muchas ocasiones no aportan semántica. Poner una etiqueta de fecha, url, nombre de autor o título no tiene mucho que ver con semántica y sí con estructuración.
- La idea global surge de la necesidad de encontrar un mecanismo que permita al HTML estructurar la información. En esencia, estos fragmentos de código suponen un mecanismo sencillo para agregar significado (semántico o no) a un documento que, sin dejar de ser legible para una persona, aporta estructura al sistema de información que sea capaz de procesar su código fuente.
- Lo curioso del caso es que, aunque este sistema existe desde hace tiempo, no es hasta hace muy poco que se consideran claves para el posicionamiento web. Algunos autores afirman que su uso incrementa hasta un 30% la visibilidad de un sitio web

Fuente: <http://jsenso.es/node/184>

Introducción (IV)

- Hay tres tipos principales de marcado semántico, los microdatos, los microformatos y las extensiones RDFa.
- Aunque la filosofía es la misma, la forma de llevar a cabo el proceso varía, y no es lo mismo hacerlo con cada uno de ellos.

Microdatos (en HTML5)

- Los **microdatos** son una de las principales aportaciones de la **versión 5 del HTML**
- Los microdatos consisten en grupos de pares nombre-valor. Estos grupos se denominan **items** y cada par nombre-valor es una **propiedad**». Para entender qué significa, imaginemos que tenemos el siguiente bloque de información sobre una película:

Título: Blade Runner

Dirección: Ridley Scott

Reparto: Harrison Ford, Rutger Hauer, Sean Young, Edward James Olmos, Daryl Hannah

Año: 1982.

Microdatos (en HTML5)

- Ese bloque de información forma un grupo, es decir, un ítem y se define como tal incluyendo el **atributo itemscope** (que vendría a significar, «el ámbito de un ítem») en un div (u otro tag) que lo contenga:

```
<div itemscope>
```

Título: Blade Runner

Dirección: Ridley Scott

Reparto: Harrison Ford, Rutger Hauer, Sean Young, Edward James Olmos, Daryl Hannah

Año: 1982.

```
</div>
```

Microdatos (en HTML5)

- Ese grupo está formado por pares de elementos que forman *propiedades*: la propiedad título, la propiedad director, la propiedad reparto y la propiedad año. Para definir las se usa el **atributo itemprop** (propiedad del ítem) dándole un valor entre comillas en el que se explica qué significa cada una.
- Si es información que va a ser visualizada, el atributo se puede declarar en cualquier etiqueta contenedora —como <div>, <p> o —, de lo contrario, podemos emplear una etiqueta <meta>, como en el siguiente ejemplo, en el cual se indica la ruta de la imagen que deberían mostrar los buscadores en una página de resultados . Cuando se utiliza una etiqueta meta, el valor se declara en el **atributo content**.

```
<div itemscope>
```

```
Título: <span itemprop="título"> Blade Runner </span>
```

```
Dirección: <span itemprop="autor">Ridley Scott </span>
```

```
Reparto: <span itemprop="reparto"> Harrison Ford, Rutger Hauer, Sean Young, Edward James Olmos, Daryl Hannah </span>
```

```
Año: <span itemprop="años">1982.</span>
```

```
<meta itemprop="imagen" content="imagen_representativa.jpg">
```

```
</div>
```

Microdatos : Schema.org

i Soportado por la mayoría de los motores de búsqueda !

- Para que estas etiquetas semánticas funcionen deben referenciarse en algún lugar, es decir, en algún sitio tiene que haber una explicación de que la propiedad <nombre> significa el nombre y los apellidos de alguien. Y, además, más o menos todo el mundo debería usar las mismas, pues si uno emplea *autor*, otro *autoría* y el de más allá *responsable*, sería un caos tremendo.
- Por suerte, para contribuir a la web semántica, Google, Microsoft y Yahoo se pusieron de acuerdo en concentrar esas explicaciones en un solo sitio y así nació el proyecto schema.org, donde poco a poco se añaden nuevas especificaciones o *schema* para microdatos. Su uso aporta cierto **nivel de normalización** .
- Para usar uno de estos *schemas* basta con añadir en el div (o la etiqueta contenedora que sea) donde se especifica el itemscope el **atributo itemtype** y como valor la URL del esquema que estemos usando. Por ejemplo, para usar la especificación de las películas, que se encuentra en esta URL: <http://www.schema.org/Movie> habría que usar este itemtype:

```
<div itemscope itemtype="http://schema.org/Movie">
```

```
Aquí irían los datos de la peli </div>
```



Microdatos : Creación de Esquemas

Schema Creator

The easiest way to create HTML with schema.org microdata

<http://schema-creator.org/>

Schemas

- [Person](#)
- [Product](#)
- [Event](#)
- [Organization](#)
- [Movie](#)
- [Book](#)
- [Review](#)

Plugins

- [Schema Creator for WordPress](#)

Getting Started

Structured data is a way for search engine machines to make sense of content in your [HTML](#). Google and other search engines created a structured data standard called [schema.org](#).

If you're like most web designers or [SEO](#)s, you might find [schema.org](#) a little difficult to grasp at first. That's why we created [Schema Creator](#)—to help you quickly build and get started with [schema.org](#) microdata.

We've included some of the most popular schemas for you to create. After you create a schema, you can copy and paste the code to your site or add additional [item properties](#) to it.

Plugin para Joomla : J4Schema



Brought to you by the creators of [Google Analytics Configuration Tool](#), [Raven Internet Marketing Tools](#).

© 2012 Sitening LLC | [Give us Feedback](#)

Microdatos: Combinación de Esquemas

- Los schema se pueden combinar, por lo que las posibilidades son infinitas.
- Imagina que tienes una web que se llama «Los libros de Sandra» y en un itemscope has definido qué eres su autora:

```
<div itemscope itemtype="http://schema.org/WebPage">  
<span itemprop="author">Sandra</span>  
</div>
```

Pero quieres añadir más datos personales que no vienen en el esquema de WebPage sino en el de Person. Como este esquema está vinculado con *author*, en esa propiedad añades un itemscope:

```
<div itemscope itemtype="http://schema.org/WebPage">  
<span itemprop="author" itemscope  
itemtype="http://schema.org/Person">Sandra</span>  
</div>
```

Microdatos (en HTML5)

- Y ya dentro de esa propiedad puedes añadir los datos de Person que necesites, como el correo electrónico:

```
<div itemscope itemtype="http://schema.org/WebPage">  
<span itemprop="author" itemscope itemtype="http://schema.org/Person">Sandra  
<span itemprop="email"> elcorreoquesea@talycual.com</span>  
</span>  
</div>
```

Para comprobar si hemos definido bien los microdatos de una web se puede utilizar una aplicación de las herramientas para webmaster de Google que te indica qué microdatos ha comprendido este buscador: [Rich Snippets Testing Tool](#).

Microdatos (en HTML5)

En resumen:

```
<div itemscope itemtype="URL del schema empleado">  
<span itemprop="Propiedad a definir">valor<span>  
<meta itemprop="Propiedad a definir" content="valor">  
</div>
```

Donde:

- **itemscope** sirve para indicar qué comienza y termina un bloque de información,
- **itemtype** para definir el esquema que se está utilizando,
- **itemprop** para indicar una propiedad del bloque,
- **content** para indicar el valor de una propiedad cuando se define en un tag meta (que no se visualiza en pantalla).

Incorporar microdatos en Wordpress usando las especificaciones de Schmea.org:
<http://www.mmfilesi.com/blog/microdatos-en-wordpress/>

Microformatos

- Los **microformatos** son porciones de código HTML (o XHTML) que permiten estructurar información aprovechando los atributos "id" o "class" empleadas por algunas etiquetas del HTML.
- Permiten aplicar una nomenclatura estandarizada para los atributos de HTML «class», «rel» y «rev» con fines semánticos.
- Por ejemplo, un microformato muy extendido (gracias a WordPress) es **hCard**, que sirve para definir la identidad de una persona.
- Así, en este fragmento de código HTML normal donde se indican algunos datos de alguien:

```
<div>  
<span>Italo</span>  
<span>Calvino</span>  
</div>
```

- Se pueden añadir microformatos para indicar que denotan la identidad de una persona, como su nombre y su apellido.

```
<div class="vcard">  
<span class="given-name">Italo</span>  
<span class="additional-name">Calvino</span>  
</div>
```

Microformatos



El sistema de trabajo con microformatos se basa en lo que llaman POSH, acrónimo de Plain Old Semantic HTML, «el viejo, plano y semántico HTML», que se rige por los siguientes principios:

- los datos tienen que estar estructurados
- hay que separar el estilo de los datos
- todo el código debe estar validado
- el nombre de cada clase tiene que cumplir una función semántica y no para la presentación del documento.

Microformatos: Los más usados

- **XFN** (<http://gmpg.org/xfn/>) for human relationships (using the rel attribute; rel is also a way to extend Creative Commons metadata)
`Jane`
- **Geo** (<http://microformats.org/wiki/geo>) for location (using simple `<meta>` tags)
`<div class="geo">GEO: 37.386013, -22.082932 </div>`
- **hCalendar** (<http://microformats.org/wiki/hcalendar>) for calendar events (mapping the commonly used iCalendar format to XHTML).
`<div class="vevent">
 http://www.web2con.com/
 Web 2.0 Conference:
 <abbr class="dtstart" title="2007-10-05">October 5</abbr>-
 <abbr class="dtend" title="2007-10-20">19</abbr>,
 at the Argent Hotel, San Francisco, CA
</div>`
- **XOXO** (<http://microformats.org/wiki/xoxo>) for outlines and blogroll-like subscriptions.
- **hCard** (<http://microformats.org/wiki/hcard>) for address books (mapping the commonly used vCard format to XHTML).
- Others:
 - **Reviews** (<http://microformats.org/wiki/hreview>)
 - **Resumes** (<http://microformats.org/wiki/hresume>)
 - Or even very domain specific like wines – **hwine**

Microformatos: Ejemplo

```
<dl class="dublincore">
  <dt>Título:</dt><dd class="title">Microdatos, microformatos y RDFa: parecen lo mismo, pero no</dd>
  <dt>Dirección:</dt><dd><a href="http://jsenso.es/node/184" class="identifier">http://jsenso.es/node/184</a></dd>
  <dt>Descripción:</dt><dd class="description">Se analizan los tres métodos para enriquecer documentos HTML: microformatos, microdatos y RDFa.</dd>
  <dt>Palabras clave:</dt><dd class="subject">Microdatos, Microformatos, RDFa</dd>
  <dt>Código de idioma:</dt><dd class="language">es</dd>
  <dt>Autor/a:</dt><dd><a href="http://jsenso.es" class="creator">Jose A. Senso</a></dd>
</dl>
```

Fuente: <http://jsenso.es/node/184>

Microformatos: cont. Ejemplo

Con lo que podría identificar sin problema el título, resumen, palabras clave... sin que tratara estos datos como una simple cadena de caracteres. Para el usuario todo este proceso sería transparente, y vería sólo esto:

Título:

Microdatos, microformatos y RDFa: parecen lo mismo, pero no

Dirección:

<http://jsenso.es/node/184>

Descripción:

Se analizan los tres métodos para enriquecer documentos HTML: microformatos, microdatos y RDFa.

Palabras clave:

Microdatos, Microformatos, RDFa

Código de idioma:

es

Autor/a:

[Jose A. Senso](#)

Fuente: <http://jsenso.es/node/184>

hCard Microformat Example

- **hCard** (<http://microformats.org/wiki/hcard>) para agendas (mapeando el formato común vCard a XHTML).
- Considerar el siguiente código HTML con marcado de microformatos embebido:

```
<div class="vcard">
  <a class="url fn"
    href="http://paginaspersonales.deusto.es/dipina">Diego López de Ipiña</a>
  <div class="org">Universidad de Deusto</div>
  <div class="adr">
    <div class="street-address">Avda. Universidades
    24</div>
    <span class="locality">Bilbao</span>
    <span class="postal-code">48007</span>
    <span class="country-name">Spain</span>
  </div>
  <div class="tel">944139000 ext.2548</div>
</div>
```

Microformat Wizards

- hCard Creator: <http://tantek.com/microformats/hcard-creator.html>
- <http://microformats.org/code/hcard/creator> (más completo)
- Aprender [hCard authoring](#)

hCard Creator

Last updated: 2005-153

This is here mostly for historical purposes.

Please use the official microformats.org hCard Creator.

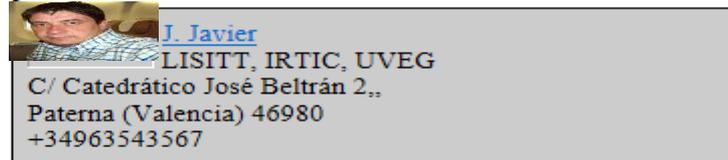
hCard-o-matic

name	J. Javier		
company	LISITT, IRTIC, UVEG		
street	C/ Catedrático José Beltrán 2..		
city	erna (Valencia)	state	<input type="text"/>
zip	46980		
phone	+34963543567		
url	http://robotica.uv.es/~jsamper		
photo url	/robotica.uv.es/~jsamper/images/foto.jpg		

code

```
<div class="vcard">
  
  <a class="url fn" href="http://robotica.uv.es/~jsamper">J. Javier</a>
  <div class="org">LISITT, IRTIC, UVEG</div>
  <div class="adr">
    <div class="street-address">C/ Catedrático José Beltrán 2., </div>
    <span class="locality">Paterna (Valencia)</span>
    <span class="postal-code"> 46980</span>
  </div>
  <div class="tel">+34963543567</div>
```

preview



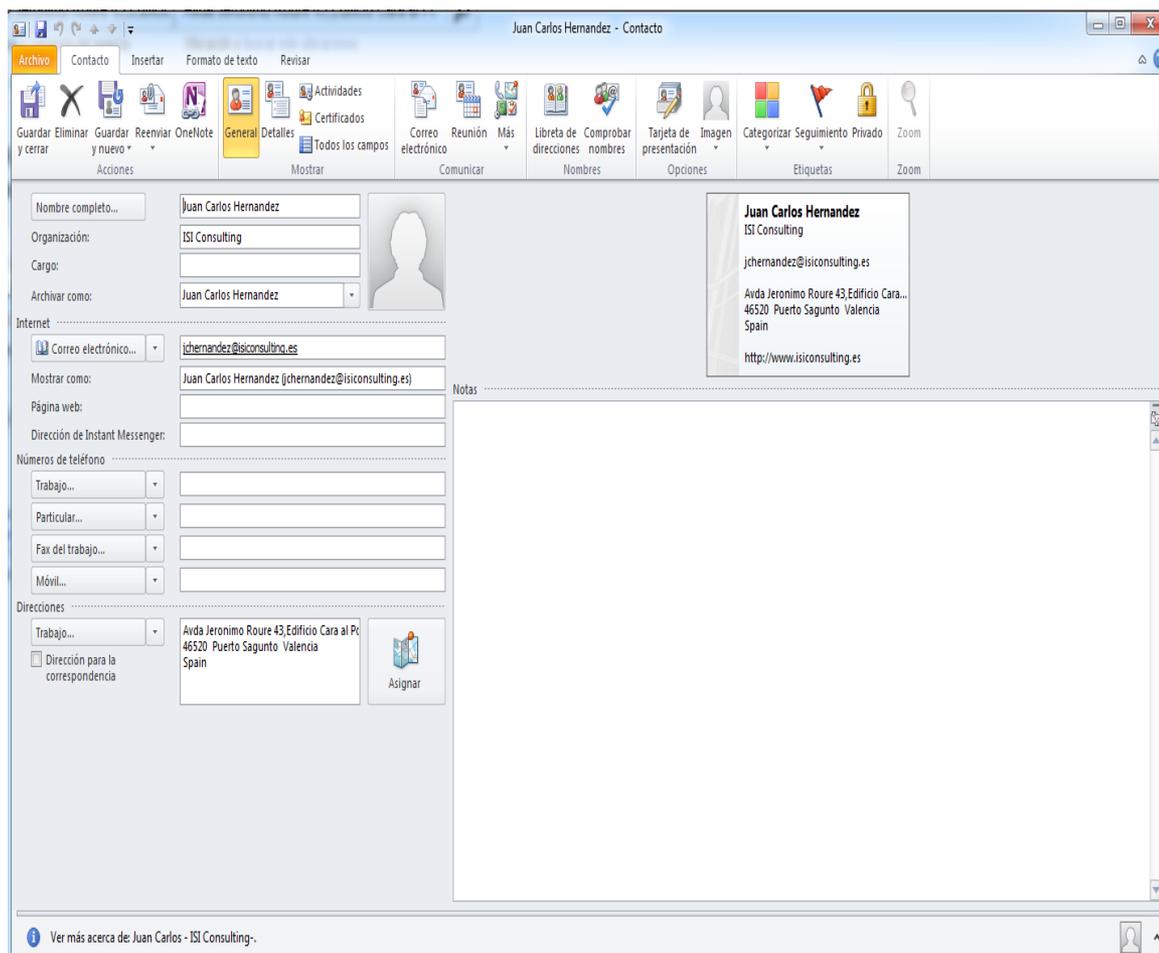
 [J. Javier](#)
LISITT, IRTIC, UVEG
C/ Catedrático José Beltrán 2.,
Paterna (Valencia) 46980
+34963543567

Uso de enlace con Vcard para agregar dirección en agenda (ej. en Outlook)

Add us to your address book

Es un formato estándar para el intercambio de información personal, específicamente tarjetas personales electrónicas (electronic business cards). Las vCards son usualmente adjuntadas a mensajes de e-mail, pero pueden ser intercambiadas en muchas otras formas, como en la World Wide Web o a través de códigos QR etc.

BEGIN:VCARD
VERSION;TYPE=WORK:3.0
FN:Juan Carlos Hernandez
N:Juan Carlos Hernandez;;;;
PROFILE:VCARD
ADR;;;Avda Jeronimo Roure
43,Edificio Cara al Port;Puerto
Sagunto;Valencia;
46520;Spain
EMAIL:jchernandez@isiconsulting.es
ORG:ISI Consulting
URL:http://www.isiconsulting.es
END:VCARD



Convertir vCard en hCard

<http://www.andypemberton.com/sandbox/hcardconvert/index.php>

vCard-2-hCard

[vCard](#) to [hCard](#) Converter

vCard file to convert

C:\Users\javi\Downloads\PAFS.vcf

Examinar...

[reset](#)

Preview



Juan Carlos Hernandez
ISI Consulting
jchernandez@isiconsulting.es
Avda Jeronimo Roure 43
Puerto Sagunto, Valencia 46520 Spain

```
<div id="hcard-Juan-Carlos-Hernandez" class="vcard">
  <span class="fn n">
    Juan Carlos Hernandez      </span>
  <ul>
    <li class="org">ISI Consulting</li>

  </ul>
  <ul>
    <li class="email">
      <a class="value"
href="mailto:jchernandez@isiconsulting.es">jchernandez@isiconsulting.es</a>
    </li>

  </ul>
  <ul>
    <li class="adr">
      <div>
      </div>
      <div class="street-address">Avda Jeronimo Roure 43</div>
      <span class="locality">Puerto Sagunto</span>,
      <span class="region">Valencia</span>
      <span class="postal-code">46520</span>
      <span class="country-name">Spain</span>
    </li>
  </ul>
</div>
```

Otros Wizard

XFN 1.1 Creator

Name

URL another web address of mine

friendship contact acquaintance friend none

physical met

professional co-worker colleague

geographical co-resident neighbor none

family child parent sibling spouse kin none

romantic muse crush date sweetheart

`Javier Samper`

hCalendar Creator

hCalendar-o-matic

summary

event title

location

url

start :

end :

TZ

description

tags

(comma separated)

code

```
<div class="vevent" id="hcalendar-event-
title"><time class="dtstart" datetime="2013-09-
03">September 3, 2013</time>--<time class="dtend"
datetime="2013-09-03"></time> <span
class="summary">event title</span>

<p style="font-size: smaller;">This <a
href="http://microformats.org/wiki/hcalendar">hCal
endar event</a> brought to you by the <a
href="http://microformats.org/code/hcalendar/creat
or">hCalendar Creator</a>.</p></div>
```

compact code

```
<div class="vevent" id="hcalendar-event-
title"><time class="dtstart" datetime="2013-09-
03">September 3, 2013</time>--<time class="dtend"
datetime="2013-09-03"></time> <span
class="summary">event title</span> <p style="font-
size: smaller;">This <a
```

preview

September 3, 2013– event title

This [hCalendar event](#) brought to you by the [hCalendar Creator](#).

Microformatos cambiarán la interacción Web

- **Sitios de agregación**

- blogging + microformats son una buena combinación
 - Si quieres vender algo, usa un blog con microformat hlisting (<http://microformats.org/wiki/hlisting>) y un site como <http://www.edgeio.com/> lo encontrará cuando agregue anuncios clasificados a través de la Web

- **Compartir información con una comunidad específica**

- *geocast* ([RSS](#) con una carga útil de [geo](#)), para las ubicaciones de las rutas de bici de montaña, y otra gente se suscribirá al feed usando Google Earth

- **Búsqueda Temática**

- Al publicar tu comic con un [microformat agreed upon by the web comic community](#), el resto de la comunidad encontrará fácilmente tu trabajo con un buscador.

- **El navegador como un “Information Broker”**

- Los navegadores futuros asociarán marcado semántico de datos en la Web con aplicaciones específicas.
- Ejemplos:
 - Información en el sitio Web de contacto estará asociado con la aplicación de contactos
 - Los eventos se asocian con la aplicación de calendario favorito
 - Ubicaciones estarán asociados con su aplicación de mapas favoritos
 - Los números de teléfono se asociarán con tu favorita aplicación de VOIP

Websites usando Microformats

- Marcado de su contenido utilizando microformatos como un medio de hacer el sitio en sí una API.
 - [Flickr](#), permitiendo a los usuarios [geotag](#) fotos
 - [Yahoo! Local](#), la cual codifica cada resultado de búsqueda con un [hCard](#)
 - [Upcoming.org](#), la cual codifica eventos con [hCalendar](#)
 - Revyu.com – permite revisar y establecer ratios de cosas
 - <http://eventful.com/>
 - <http://www.linkedin.com/in/harryhalpin>
 - <http://yedda.com/>
 - <http://local.yahoo.com>

Principales Microformatos

- Firefox 3 y posteriores usan Microformatos
 - http://wiki.mozilla.org/Firefox/Feature_Brainstorming:Microformat_Handling
- Internet Explorer 8 y superiores los soportan
 - <http://factoryjoe.com/blog/2006/10/29/internet-explorer-80-will-support-microformats/>

Interrogative Word	Action	Tool	Microformat
Who?	Meet a person	Contacts book	hCard
Where?	Find a location	Map	adr, geo
When?	Schedule	Calendar	hCalendar
What? How? Why?	Domain Specific		



Feed



Contact



Location



Event

El navegador como un Broker de Información

- Nuevo rol de los navegadores:
 - **Detectar esta información en las páginas web y manejar esta información fuera para otras aplicaciones.**
 - Se pasa de un simple renderizador de HTML a un *information broker*.
- La detección de Microformatos debería ser diseñado como una completa plataforma extensible y abierta.
 - Aplicaciones de gestión de contactos, calendario, y mapeo usan la API de los navegadores para integrar con el sistema de detección de microformatos del navegador.



Operator Firefox Add-on

- Operator aprovecha microformatos ya están disponibles en muchas páginas web para proporcionar nuevas maneras de interactuar con servicios web.

Videos ejemplo:

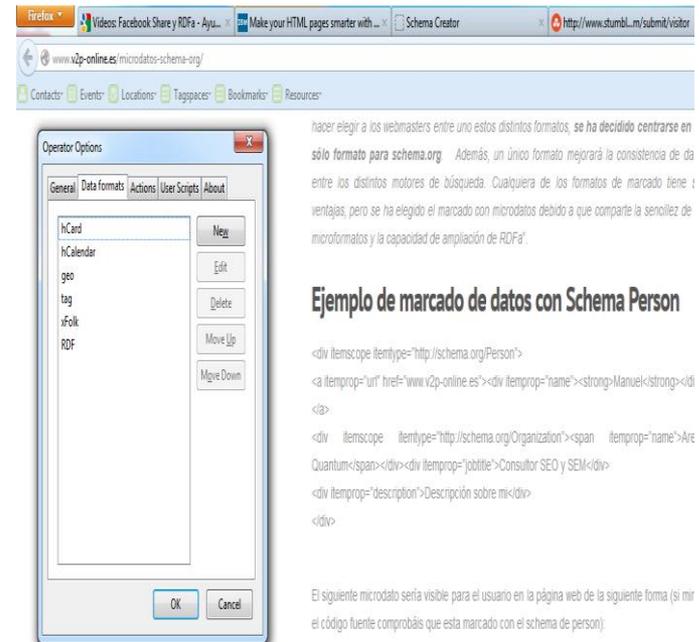
- <http://www.youtube.com/watch?v=Y9A6aypBNw8>
- <http://www.youtube.com/watch?v=BSBKVYUIUml>

- Ejemplo: Flickr + GoogleMaps

- <http://www.flickr.com/photos/dbaron/299291151>

- Ejemplo:

- <http://en.wikipedia.org/wiki/HCard> + Outlook



Firefox - Videos Facebook Share y RDFa - Ayu... - Make your HTML pages smarter with ... - Schema Creator - http://www.stumb.../submit/visitor

www.v2p-online.es/microdatos-schema-org/

Contactor - Evento - Location - Tagpaces - Bookmark - Resources

Operator Options

General Data formats Actions User Scripts About

hCard
hCalendar
geo
tag
vFolk
RDF

New
Edit
Delete
Move Up
Move Down

OK Cancel

haber elegido a los webmasters entre uno de estos distintos formatos, se ha decidido centrarse en sólo formato para schema.org. Además, un único formato mejorará la consistencia de datos entre los distintos motores de búsqueda. Cualquiera de los formatos de marcado tiene ventajas, pero se ha elegido el marcado con microdatos debido a que comparte la sencillez de microformatos y la capacidad de ampliación de RDFa.

Ejemplo de marcado de datos con Schema Person

```
<div itemscope itemtype="http://schema.org/Person">  
<a itemprop="url" href="http://www.v2p-online.es"><div itemprop="name"><strong>Manuel</strong></div>  
</a>  
<div itemscope itemtype="http://schema.org/Organization"><span itemprop="name">Arena Quantum</span></div><div itemprop="jobtitle">Consultor SEO y SEM</div>  
<div itemprop="description">Descripción sobre mi</div>  
</div>
```

El siguiente microdato sería visible para el usuario en la página web de la siguiente forma (si miras el código fuente comprobáis que está marcado con el schema de person)

Manuel
Arena Quantum
Consultor SEO y SEM

Operator Firefox Add-on

The screenshot shows a Firefox browser window with the address bar at www.mattcutts.com/blog/thoughts-on-google-instant/. The browser's toolbar includes a search bar with the text "Thoughts on Google Instant" and a "+" button. Below the address bar is a navigation bar with links for Contacts (268), Events, Locations, Tagspaces (1), Bookmarks, and Resources (1). The main content area features a navigation menu on the left with names like Mrinmay Bhattacharjee, Kathy, Adam Jusko, Gary Ashton, Rodrigo Silveira, Nuttakorn Rattanachaisit, Mr Cakey, the one who blogs, Steve St. Germain, Brian Ussey, Quadrille, Brett, Halvorsen, Ron Tumbokon, Ben Joven, Ozer Erke, George Charalampakis, Jeff Hancock, Kim Clinkunbroomer, Buzzillionsim, Jason Cavett, gordon, Josh L, Andy Symonds, Ann Donnelly, Daniel J. Mount, Frank Reed, and don. The main article is titled "Matt Cutts: Gadgets, Google, and SEO" and is categorized under "Thoughts on Google Instant", posted on September 8, 2010. The article text discusses Google's search infrastructure and the announcement of Google Instant. A right sidebar contains a "FOLLOW ME" section with links to Twitter and Google+, a search box, and a "CATEGORIES" section listing various topics with their respective counts.

Firefox Thoughts on Google Instant

www.mattcutts.com/blog/thoughts-on-google-instant/

Contacts (268) Events Locations Tagspaces (1) Bookmarks Resources (1)

Mrinmay Bhattacharjee
Kathy
Adam Jusko
Gary Ashton
Rodrigo Silveira
Nuttakorn Rattanachaisit
Mr Cakey
the one who blogs
Steve St. Germain
Brian Ussey
Quadrille
Brett
Halvorsen
Ron Tumbokon
Ben Joven
Ozer Erke
George Charalampakis
Jeff Hancock
Kim Clinkunbroomer
Buzzillionsim
Jason Cavett
gordon
Josh L
Andy Symonds
Ann Donnelly
Daniel J. Mount
Frank Reed
don

HOME ABOUT ME GOOGLE/SEO DISCLAIMER DISCLOSURE

Matt Cutts: Gadgets, Google, and SEO

Thoughts on Google Instant

Posted September 8, 2010 in Google/SEO

The blogosphere is [absorbing](#) today's announcement of [Google Instant](#). I wanted to give some context and some thoughts.

Google's web search (and web search infrastructure) team tries to do several things well:

- We want the most comprehensive index of the web. We explore ways to crawl the web deeper, faster, and better, from increasing our index size or indexing speed to crawling web forms to discovering links in JavaScript.
- We try to return relevant, useful results. Hundreds of people work on lots of improvements to our ranking algorithms.
- We try to return your search results really fast.
- We try to improve our search user interface (UI).

The first three things aren't highly visible. Average users might not notice changes like Caffeine (improved indexing) or a better algorithm to detect hacked sites—although we have seen effects like users searching more when we deploy a fresher index. A bunch of people at Google have come up with amazing ways to make your search results faster. We've shared [many of those insights](#) to help make the web faster as a whole.

FOLLOW ME

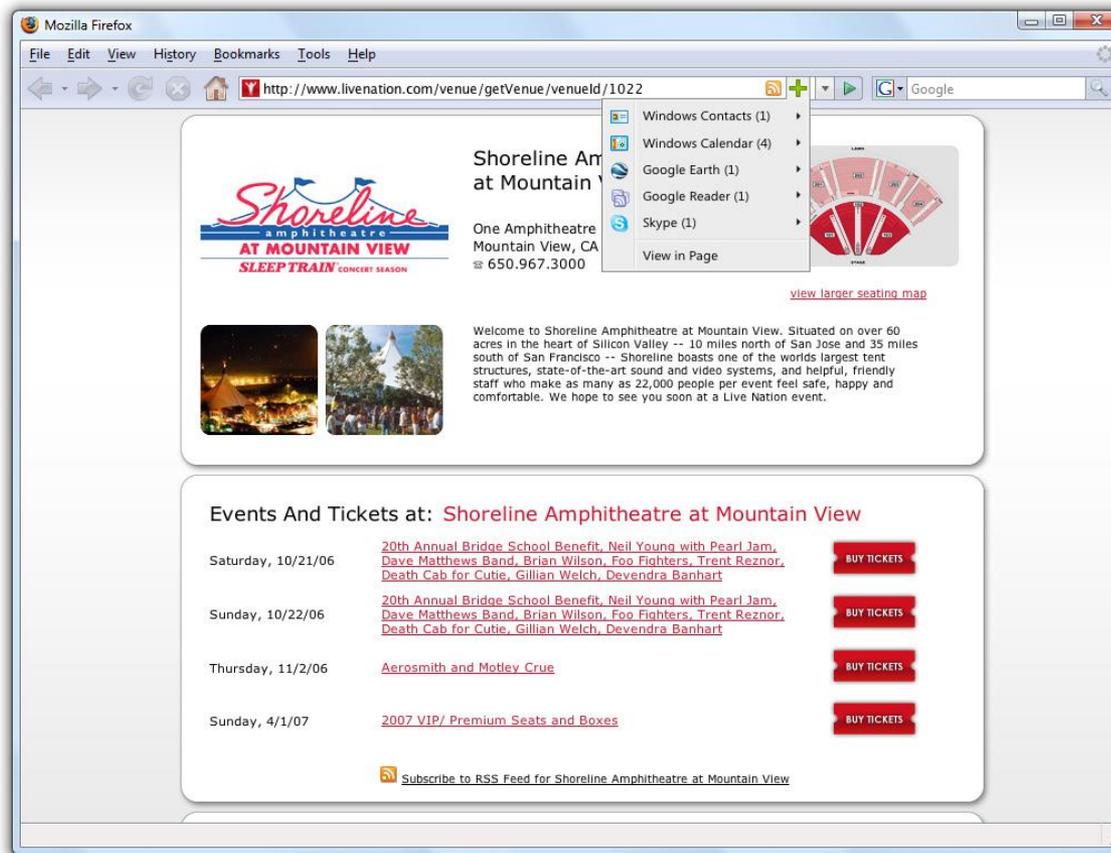
- Twitter
- Google+

Search this website...
Search

CATEGORIES

- 30 days (33)
- Android (5)
- Books/Magazines (24)
- Chrome (23)
- Food (6)
- Fun (44)
- Gadgets/Hack (94)
- Games (5)
- Glossary/Definitions (2)
- Google/SEO (548)
- How to (54)
- iPhone (17)
- Leffovers (5)

Firefox 3: The Web Information Broker



Microformatos vs Microdatos

Desde el punto de vista práctico, las razones que hacen decantarse por el uso de microdatos y no por microformatos son:

- mucho **más cómodo y limpio** separar las etiquetas semánticas (aparte de las nativas de HTML5) de las que se usan para maquetar (diseño).
- Microdatos son **más flexibles, escalables e intuitivos**. Como permiten combinar diferentes esquemas o definiciones («schema»), se puede etiquetar semánticamente cualquier tipo de formato de forma mucho más precisa y adaptable.
- Microdatos alcanzan un mayor nivel de **normalización** mediante los esquemas.
- **Es muy probable que los microdatos terminen por sustituir a los microformatos**, porque los demiurgos de Internet y la informática doméstica —Google, Microsoft, Yahoo y el buscador ruso Yandex— se han puesto de acuerdo en apostar por este tipo de etiquetado semántico

RDFa



- **RDFa** es un conjunto de extensiones de XHTML propuestas por W3C para introducir semántica en los documentos. RDFa aprovecha atributos de los elementos meta y *link* de XHTML y los generaliza de forma que puedan ser utilizados en otros elementos. Además se ha definido una correspondencia simple que permite extraer tripletes RDF.
- Las empresas identifican la necesidad de establecer y administrar un marco referencial de contenidos, sobre el que se apoyaran diferentes publicaciones, tanto en la red como en otros medios de comunicación.
- RDFa está diseñado para ser compatible con HTML 4, HTML 5, o XHTML.

<http://www.ibm.com/developerworks/library/wa-rdfalite/>

RDFa 1.1 Lite

- Resource Description Framework (RDF) se ha convertido en un formato cada vez más pragmático en el tiempo. RDF anotación (RDFa) ha tenido mucho éxito como sistema de anotaciones en documentos HTML en línea en la web.
- Es soportado por los motores de búsqueda entre ellos Google , en la forma de Rich Snippets.
- La aparición de los microdatos y la iniciativa Schema.org han ejercido presión para simplificar RDFa aún más. El W3C actuó y produjo una versión muy simplificada: **RDFa 1.1 Lite**. W3C Recommendation el 07 de Junio de 2012 (<http://www.w3.org/TR/rdfa-lite/>).
- Nosotros vamos a tratar en las siguientes diapositivas con RDFa Lite.

Atributos de RDFa

Los atributos que se utilizan son:

- **Vocab:** que permite que el autor declare un vocabulario único para un fragmento de código HTML
- **Prefix:** puede asignar prefijos para representar URLs y, utilizando dichos prefijos, pueden abreviarse los elementos del vocabulario.
- **typeof:** indica de que tipo es la instancia descrita.
- **about:** una URI que indica el recurso que describen los metadatos y que remite al documento actual por defecto
- **rel, rev, href y resource:** atributos que establecen una relación o relación inversa con otro recurso
- **property:** aporta una propiedad para el contenido de un elemento
- **content:** atributo opcional que se sobrepone al contenido del elemento cuando se usa el atributo *property*
- **datatype:** atributo opcional que indica el tipo de datos del contenido

Ventajas de RDFa

Hay cinco principios de los metadatos interoperables que cumple RDFa

- Independencia del editor: cada sede web puede usar sus propios estándares.
- Reutilización de datos: se debe tratar de no duplicar los datos;
- RDFa hace innecesario separar las secciones XML y HTML de los mismos contenidos.
- Autocontención: las secciones de XML y HTML pueden mantenerse separadas.
- Modularidad del esquema: Los atributos son reusables
- Escalabilidad: se pueden añadir campos adicionales con la única condición de que se mantenga la capacidad de extraer semántica de los datos del archivo XHTML.

RDFa Elements

- RDFa-aware software can pull useful metadata from **span**, **link** and **meta** elements with only minor modifications to these elements.
 - **span** elements are most popular for RDFa because you can insert them anywhere in the body of an HTML document, you can add the same RDFa attributes to any elements you like.
 - **link** and **metadata** are useful to insert RDFa in the **head** of an HTML document
 - The **a** linking element is also popular for storing RDFa metadata
 - It expresses a relationship between one resource (the document where it's stored) and another (the resource it links to).
 - **rel** attribute indicates the relationship in the triple

3 Categories of RDFa Use Cases

- **Inline metadata about document components**

`<p>This photo was taken by`

- El elemento `span` y sus valores de atributo permiten que una herramienta RDFa-aware consiga las siguientes tripletas del documento (en RDF/XML):

```
<rdf:Description
  rdf:about="file:///C:/dat/xml/rdf/rdfa/photo1.jpg">
  <dc:creator>Mark Birbeck</dc:creator>
</rdf:Description>
```

`<p>Last revision of document:`

- La tripleta resultado usa el valor `content` de la fecha:

```
<rdf:Description
  rdf:about="http://www.snee.com/docs/mydoc1.html">
  <dc:date>20070315T15:32:00</dc:date> </rdf:Description>
```

- Metadata about the containing document
- Out-of-line metadata

More Triples, Fewer Subjects

- RDFa permite crear múltiples tripletas desde el mismo tema sin “ensuciar” el documento demasiado.
- Los elementos hijo heredan el atributo **about** del elemento padre (resource <http://www.snee.com/img/myfile.jpg>)

- Si ningún atributo **about** es encontrado entonces él es el actual documento:

```

  <span property="dc:subject" content="Niagra
  Falls"/>
  <span property="dc:creator" content="Richard
  Mutt"/>
  <span property="dc:format" content="img/jpeg"/>
</img>
```

3 Categories of RDFa Use Cases

- Inline metadata about document components
- **Metadata about the containing document**
 - Los metadatos sobre el documento con ningún contenido visualizable se puede almacenar en el elemento **head** del documento

```
<html xmlns:fm="http://www.foomagazine.com/ns/prod/">
  <head>
    <title>Is Black the New Black?</title>
    <meta property="fm:newsstandDate" content="2006-04-03"/>
    <meta property="fm:copyEditor" content="RSeIavy"/>
    <meta property="fm:copyEdited" content="2006-03-28T10:33:00"/>
  </head>
  <body>
```
 - Un extractor RDFa obtiene el siguiente RDF / XML de este:

```
<rdf:Description rdf:about="">
  <fm:newsstandDate>2006-04-03</fm:newsstandDate>
  <fm:copyEditor>RSeIavy</fm:copyEditor>
  <fm:copyEdited>2006-03-28T10:33:00</fm:copyEdited>
</rdf:Description>
```

- Out-of-line metadata

3 Categories of RDFa Use Cases

- Inline metadata about document components
- Metadata about the containing document
- **Out-of-line metadata**

```
<html xmlns:fm="http://www.foomagazine.com/ns/prod/">
<head><meta about="#recipe13941"><meta property="fm:ComponentID">XZ3214</meta><meta
  property="fm:ComponentType">Recipe</meta><meta
  property="fm:RecipeID">r003423</meta></meta>
</head>
<body>
  <h>Add Some Tex Mex Sizzle to Your Kid's Lunch</h>
  <section id="recipe22143">
    <h>Amigo Corn Dogs</h>
  </section>
  <section id="recipe13941">
    <h>EZ Bean Tacos</h>
  </section>
</body>
</html>
```

- La tripleta resultante :

```
<rdf:Description rdf:about="file:///C:/dat/xml/rdf/rdfa/test5.html#recipe13941">
  <fm:ComponentType>Recipe</fm:ComponentType>
  <fm:RecipeID>r003423</fm:RecipeID>
  <fm:ComponentID>XZ3214</fm:ComponentID>
</rdf:Description>
```

Ejemplo sencillo de RDFa

```
<div xmlns:dc="http://purl.org/dc/elements/1.1/"  
  about="http://www.example.com/books/wikinomics">  
  <span property="dc:title">Wikinomics</span>  
  <span property="dc:creator">Don Tapscott</span>  
  <span property="dc:date">2006-10-01</span>  
</div>
```

Veámoslo paso a paso

Alice publica un blog y le gustaría proporcionar información estructurada adicional en sus páginas, como la fecha de publicación o el título. A ella le gustaría utilizar los términos definidos en el vocabulario Dublin Core [[DC11](#)], un conjunto de términos ampliamente utilizados, por ejemplo, en la industria editorial o bibliotecas. Su blog ya contiene dicha información:

```
<html>
<head> ... </head>
<body>
...
<h2>The Trouble with Bob</h2> <p>Date: 2011-09-10
<p>Date: 2011-09-10</p>
...
</body>
```

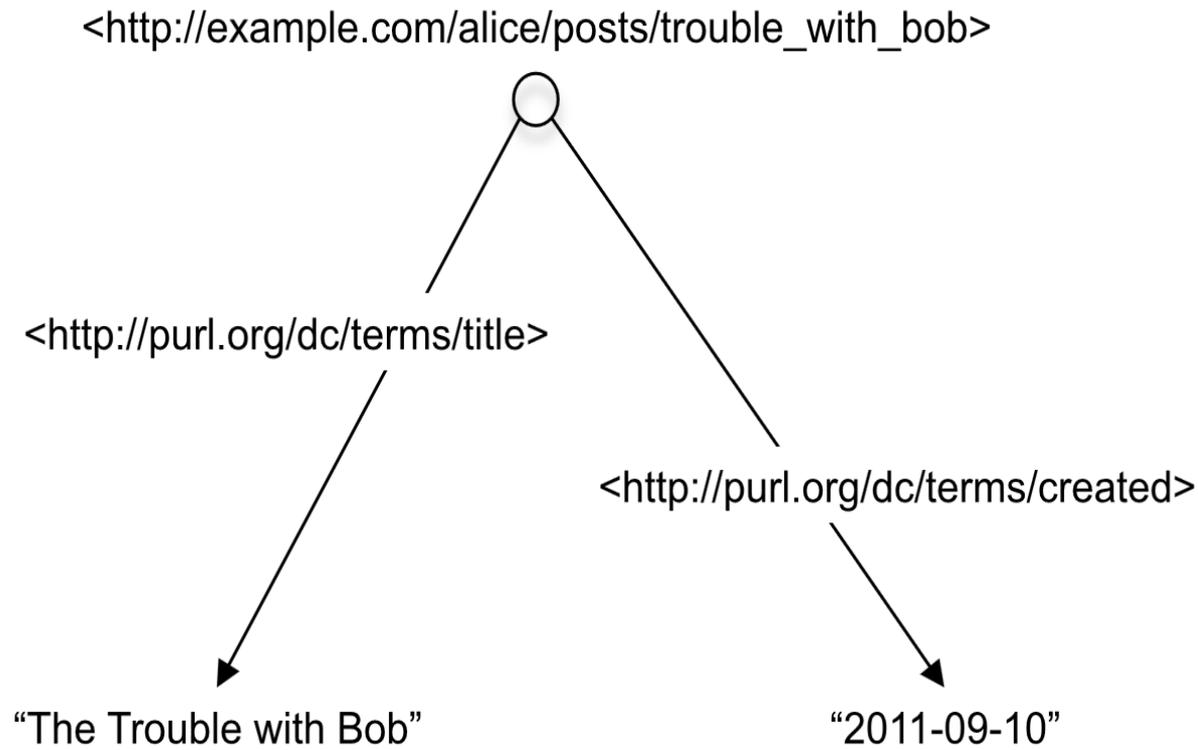
Veámoslo paso a paso (II)

Sin embargo, esta información está pensada únicamente para las personas; los ordenadores necesitan algunos métodos sofisticados para extraerla. Sin embargo, utilizando RDFa, ella puede marcar su página para expresar *datos estructurados*:

```
<html>
<head> ... </head>
<body>
...
<h2 property="http://purl.org/dc/terms/title">The Trouble with Bob</h2>
<p>Date: <span property="http://purl.org/dc/terms/created">2011-09-
10</span></p>
...
</body>
```

Veámoslo paso a paso (III)

Un modo útil de visualizar los datos estructurados sería:



Veámoslo paso a paso (IV)

- Vale la pena destacar que RDFa utiliza URLs para identificar casi todo. Por esta razón, en lugar de utilizar propiedades como title o created, se usa <http://purl.org/dc/terms/title> y <http://purl.org/dc/terms/created>. La razón tras esta decisión de diseño se basa en la **portabilidad de los datos, la consistencia y el intercambio de información**.
- El uso de URLs **elimina la posibilidad de ambigüedades** en la terminología. Sin asegurar que no hay ambigüedad, el término "title" puede significar "el título de una obra", "el título de un puesto de trabajo", o "la escritura de una propiedad inmobiliaria". Cuando cada término del vocabulario es una URL, está disponible a un clic de distancia una detallada explicación del significado del mismo. Esto permite que tanto personas como máquinas puedan seguir el enlace para encontrar el significado de un término concreto de un vocabulario. Mediante el uso de una URL para identificar un tipo de título en particular, por ejemplo <http://purl.org/dc/terms/created>, tanto las personas como las máquinas pueden entender que dicha URL se refiere de forma inequívoca a la "Fecha de creación de un recurso", tal como una página web.
- Si las direcciones URL son las mismas, los términos del vocabulario significan lo mismo si son <> entonces no. También se hace muy fácil crear nuevos términos y documentos de vocabularios. Si se puede publicar un documento en la Web, automáticamente se tiene la capacidad para crear un documento que contiene un nuevo vocabulario con nuevos términos.

Veámoslo paso a paso (V)

- El ejemplo anterior muestra como Alice puede marcar el texto para hacerlo legible por máquina. A ella también le gustaría marcar los enlaces de forma que sean legible por máquina, para expresar el tipo de enlace que está siendo descrito. RDFa permite al editor añadir un "sabor", por ejemplo una etiqueta, a un enlace sobre el que se puede hacer clic, y que puedan entender las aplicaciones informáticas. Esto hace que el mismo marcado sea útil tanto para las personas como para las máquinas.
- En el pie de página de su blog, Alice ya ha declarado que su contenido puede ser reutilizado libremente, siempre y cuando ella reciba la debida acreditación de su autoría cuando sus artículos sean citados. El código HTML incluye un enlace a una licencia Creative Commons [[CC-ABOUT](#)]:
- `<p>All content on this site is licensed under a Creative Commons License. ©2011 Alice Birpemschwick.</p>`

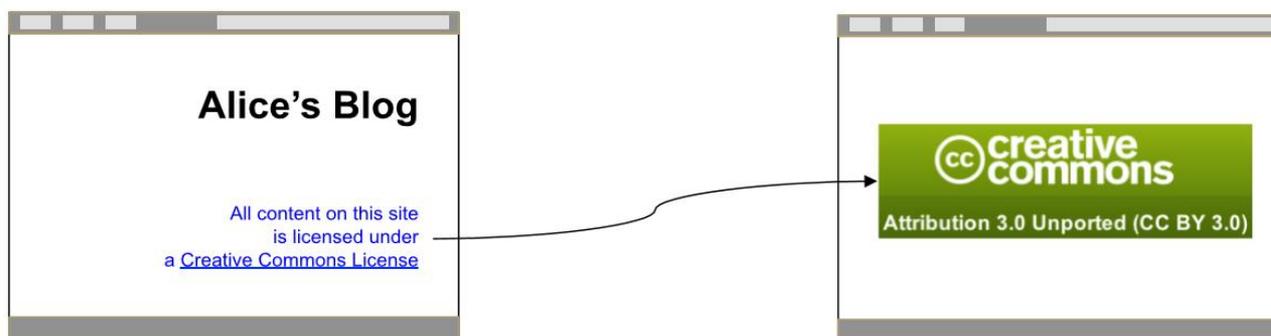
Veámoslo paso a paso (VI)

- Una persona entiende claramente esta frase, en particular el *significado* del enlace con respecto al documento actual: indica la licencia del documento, las condiciones bajo las que puede distribuirse el contenido de la página. Desafortunadamente, cuando Bob visita el blog de Alice su navegador únicamente ve un enlace plano que podría apuntar tanto a uno de los amigos de Alice como a su currículum vitae. Para que el navegador de Bob entienda que dicho enlace actualmente apunta a una página con los términos de la licencia del documento, Alice necesita añadir algún *sabor*, alguna indicación al respecto del tipo de enlace del que se trata.
- Alice puede añadir este sabor usando otra vez el atributo property. En efecto, cuando el elemento contiene el atributo href (o src), property se asocia automáticamente con el valor de este atributo en lugar del contenido textual del elemento a. El valor del atributo es el identificador URL <http://creativecommons.org/ns#license>, definido por [Creative Commons](#):

```
<p>All content on this site is licensed under  
<a property="http://creativecommons.org/ns#license"  
href="http://creativecommons.org/licenses/by/3.0/"> a Creative Commons License</a>. ©2011  
Alice Birpemschwick.</p>Con este pequeño cambio, el navegador de Bob comprenderá ahora que este  
enlace tiene un sabor: indica la licencia del blog:
```

Veámoslo paso a paso (VII)

- Con este pequeño cambio, el navegador de Bob comprenderá ahora que este enlace tiene un sabor: indica la licencia del blog:



RDFa: Atributo Vocab

- En una serie de casos de uso sencillos, como en el ejemplo del blog de Alice, los autores HTML normalmente utilizarán únicamente un vocabulario. Sin embargo, mientras que la tarea para generar direcciones URL completas a través de sistemas de gestión de contenidos (Content Management Systems, CMS) no reviste problema alguno, hacerlo a mano puede ser tedioso y propenso a errores por parte de las personas. Para paliar este problema RDFa introduce el atributo vocab que permite que el autor declare un vocabulario único para un fragmento de código HTML.

Así, en lugar de:

- `<html> <head> ... </head> <body> ... <h2 property="http://purl.org/dc/terms/title">The Trouble with Bob</h2> <p>Date: 2011-09-10</p> ... </body> Alice`

puede escribir:

- `<html> <head> ... </head> <body vocab="http://purl.org/dc/terms/"> ... <h2 property="title">The Trouble with Bob</h2> <p>Date: 2011-09-10</p> ... </body>`

RDFa: Atributo Vocab II

- Los valores de las propiedades individuales son ahora "términos", los cuales simplemente se concatenan a la URL definida mediante el atributo vocab. El atributo se puede colocar en cualquier elemento HTML (es decir, no sólo en el elemento body como en el ejemplo) y su efecto es válido para todos los elementos por debajo de ese punto.
- Los vocabularios por defecto y las URIs completas pueden combinarse en cualquier momento. Es decir, Alice podría escribir:

```
<html> <head> ... </head>
<body vocab="http://purl.org/dc/terms/"> ...
<h2 property="title">The Trouble with Bob</h2>
<p>Date: <span property="http://purl.org/dc/terms/created">2011-09-
10</span></p>
...
</body>
```

RDFa: Atributo Vocab III

Uso de dos vocabularios:

Nota:

el atributo vocab en el párrafo de la licencia anula la definición heredada del cuerpo del documento

```
<html>
<head> ... </head>
<body vocab="http://purl.org/dc/terms/">
...
<h2 property="title">The Trouble with Bob</h2>
<p>Date: <span property="created">2011-09-10</span></p>
...
<p vocab="http://creativecommons.org/ns#">All content on this site is licensed under <a
property="license" href="http://creativecommons.org/licenses/by/3.0/"> a Creative Commons
License</a>. ©2011 Alice Birpemschwick.</p>
</body>
</html>
```

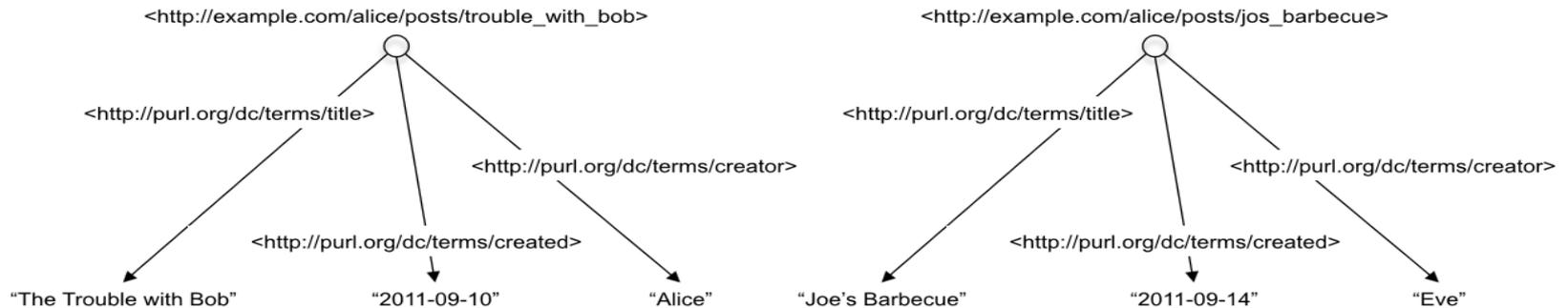
RDFa: Atributo Resource

Por supuesto, la página del blog de Alice puede contener múltiples entradas. A veces Eve, la hermana de Alice, también introduce alguna entrada en el blog. En la página principal del blog se enumeran las 10 entradas más recientes, cada una con su propio título, autor y párrafo introductorio. ¿Cómo debería marcar Alice el título de cada una de las entradas de forma individual pese a que todas aparecen en la misma página? RDFa proporciona el atributo resource para especificar el "contexto", es decir, la dirección exacta a la que se aplica el marcado RDFa:

```
<body vocab="http://purl.org/dc/terms/">
...
  <div resource="/alice/posts/trouble_with_bob">
    <h2 property="title">The trouble with Bob</h2>
    <p>Date: <span property="created">2011-09-10</span></p>
    <h3 property="creator">Alice</h3>
    ...
  </div>
...
  <div resource="/alice/posts/jos_barbecue">
    <h2 property="title">Jo's Barbecue</h2>
    <p>Date: <span property="created">2011-09-14</span></p>
    <h3 property="creator">Eve</h3>
    ...
  </div>
</body>
```

RDFa: Representación de múltiples items

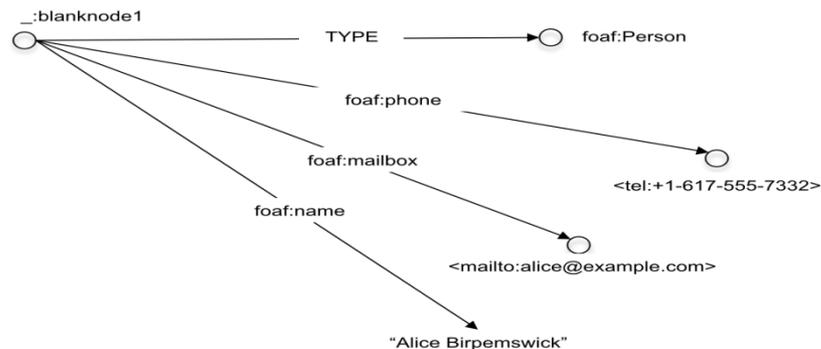
Representación de Múltiples items por página, dónde cada entrada del blog se representa mediante un nodo, con propiedades asociadas al mismo.



RDFa: Atributo typeof

Uso de `typeof` (para declarar un nuevo item de datos de un tipo concreto) y vocabulario FOAF: Alice indica el contenido de la página que representa su nombre completo, su dirección de correo electrónico y su número de teléfono

```
<div vocab="http://xmlns.com/foaf/0.1/" typeof="Person"><p>  
<p>  
<span property="name">Alice Birpemschwick</span>,  
Email: <a property="mbox" href="mailto:alice@example.com">alice@example.com</a>,  
Phone: <a property="phone" href="tel:+1-617-555-7332">+1 617.555.7332</a>  
</p>  
</div>
```



RDFa: atributo typeof y la Red Social

Alice es miembro de 5 sitios diferentes de redes sociales. Está cansada de introducir información repetida sobre sus amigos en cada red social nueva, por lo que decide enumerar a sus amigos en un único lugar de su sitio web, combinándolo con sus propios datos FOAF. Con RDFa, puede indicar esto a sus amistades en su propia página web y permitir que los sitios de las redes sociales lean automáticamente esta información. Hasta el momento, Alice ha enumerado tres personas pero no ha especificado su relación con ellos; podrían ser sus amigos o sus poetas favoritos del siglo XVII. Para indicar que los conoce utiliza la propiedad **FOAF foaf:knows**:

```
<div vocab="http://xmlns.com/foaf/0.1/" typeof="Person">
  <p>
    <span property="name">Alice Birpemswick</span>,
    Email: <a property="mbox" href="mailto:alice@example.com">alice@example.com</a>,
    Phone: <a property="phone" href="tel:+1-617-555-7332">+1 617.555.7332</a>
  </p>
  <ul>
    <li property="knows" typeof="Person">
      <a property="homepage" href="http://example.com/bob/"><span property="name">Bob</span></a>
    </li>
    <li property="knows" typeof="Person">
      <a property="homepage" href="http://example.com/eve/"><span property="name">Eve</span></a>
    </li>
    <li property="knows" typeof="Person">
      <a property="homepage" href="http://example.com/manu/"><span property="name">Manu</span></a>
    </li>
  </ul>
</div>
```

RDFa: atributo Prefix

Uso de Prefix y vocab juntos

RDFa ofrece la posibilidad de utilizar los términos con prefijo: un atributo especial `prefix` puede asignar prefijos para representar URLs y, utilizando dichos prefijos, pueden abreviarse los elementos del vocabulario. Se utiliza la sintaxis `prefix:reference`: simplemente se concatena la dirección URL asociada con `prefix` a `reference` para crear una URL completa.

```
<html> <head>
...
</head>
<body vocab="http://purl.org/dc/terms/" prefix="schema:
http://schema.org/">
  <div resource="/alice/posts/trouble_with_bob" typeof="schema:BlogPosting">
    <h2 property="title">The trouble with Bob</h2>
    ...
    <h3 property="creator" resource="#me">Alice</h3>
    <div property="schema:articleBody">
      <p>The trouble with Bob is that he takes much better photos than I do:</p>
    </div>
    ...
  </div>
</body></html>
```

RDFa : atributo Content

La presencia del atributo *content* ordena al procesador RDFa anular el comportamiento por defecto que utiliza el contenido textual, pasando a utilizar el valor del atributo *content*. Mediante este atributo Alice podría proporcionar una fecha más legible, al tiempo que mantiene un contenido inequívoco para las máquinas al utilizar datos estructurados.

```
<html>
<head>
...
</head>
<body> ... <h2 property="http://purl.org/dc/terms/title">The Trouble with Bob</h2> <p>Date: <span
property="http://purl.org/dc/terms/created" content="2011-09-10">10th of September, 2011</span></p>
...
</body>
</html>
```

Aproximación "tradicional" de añadir metadatos simples a una página web ha utilizado la cabecera del documento a través de los elementos y meta.

```
<html>
<head prefix="og: http://ogp.me/ns#" >
...
<meta property="og:title" content="The Trouble with Bob" />
<meta property="og:type" content="text" />
<meta property="og:image" content="http://example.com/alice/bob-ugly.jpg" />
...
</head>
<body>...
</body> </html>
```

RDFa: atributo para Tipo de Datos

El valor utilizado para indicar una fecha puede ser ambigua para las máquinas. Por supuesto, si un programa "sabe" que <http://purl.org/dc/terms/date> se refiere a una fecha, entonces por supuesto que es posible deducir que la cadena "2011" se refiere a un año. Pero puede que los procesadores, por ejemplo, proporcionen una presentación visual de todos los datos estructurados de una página específica, y podría utilizar un componente de software para interpretarlo y mostrarlo como un año y otro para hacer lo mismo pero como un número entero. Para evitar ambigüedad hay que establecer información adicional. Ésta puede ser transmitida al procesador RDFa mediante el atributo RDFa datatype:

```
<p>All content on this site is licensed under  
  <a property="http://creativecommons.org/ns#license"  
href="http://creativecommons.org/licenses/by/3.0/">  
  a Creative Commons License</a>. ©<span property="dc:date" datatype="xsd:gYear">2011</span>  
Alice Birpemschwick.</p>
```

donde `xsd:gYear` se refiere a <http://www.w3.org/2001/XMLSchema#gYear>, y es uno de los tipos de datos estándar definidos por la especificación de tipos de datos del W3C [XSD] que contiene tipos como el booleano, entero, fecha o coma flotante. (xsd es uno de los prefijos por defecto de RDFa.)

Rdfa:atributo About

RDFa introduce un atributo separado, denominado about, que puede usarse como una alternativa a resource en el establecimiento del contexto. Supongamos que se desea crear una página índice separada para todas sus entradas, y la única información que nos gustaría incluir en ella, como datos estructurados, son las referencias a los títulos.

```
<ul>
  <li about="/alice/posts/trouble_with_bob" property="title">The trouble with Bob</li>
  <li about="/alice/posts/jos_barbecue" property="title">Jo's Barbecue</li>
  ...
</li>
```

La diferencia fundamental entre about y resource es que el primero solo se utiliza para establecer el contexto, ya sea combinándolo con el atributo property en el mismo elemento o no. Esto también significa que, para dicho uso, about y resource son intercambiables

RDFa: atributo rel

En contraste con el atributo property, **rel** *nunca* considera el contenido textual de un elemento (o el valor del atributo content). En su lugar, si no se ha especificado claramente el objeto de un enlace mediante, por ejemplo, el atributo resource o el atributo href, el procesador supone que debe continuar y encontrar uno o más objetos en la jerarquía y utilizarlos. Esto es lo que sucede en este caso: el atributo knows del elemento ul no incluye ningún objeto evidente; sin embargo, el procesador lo encuentra en los elementos individuales li y los utiliza. El uso de este tipo de patrón está muy extendido en RDFa.

```
<div vocab="http://xmlns.com/foaf/0.1/" resource="#me">
  <ul rel="knows">
    <li resource="http://example.com/bob/#me" typeof="Person">
      <a property="homepage" href="http://example.com/bob/"><span property="name">Bob</span></a>
    </li>
    <li resource="http://example.com/eve/#me" typeof="Person">
      <a property="homepage" href="http://example.com/eve/"><span property="name">Eve</span></a>
    </li>
    <li resource="http://example.com/manu/#me" typeof="Person">
      <a property="homepage" href="http://example.com/manu/"><span property="name">Manu</span></a>
    </li>
  </ul>
</div>
```

Herramientas RDFa

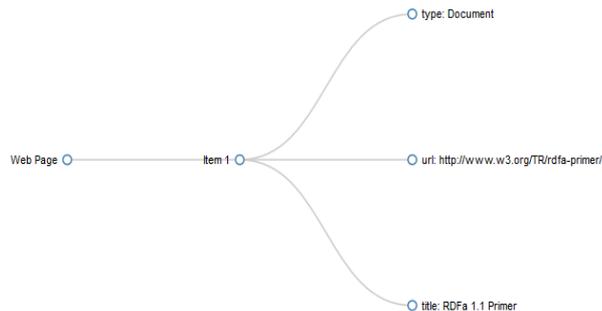
- Hay una amplia variedad de herramientas que se pueden utilizar para generar o procesar datos RDFa. Una buena fuente que recopila referencias de estas herramientas es la [página sobre RDFa de la wiki de la Web Semántica del W3C](#). Pese a todo se debe tener cuidado ya que algunas de estas herramientas puedan estar relacionadas con versiones anteriores de RDFa.
- Otra fuente puede ser la [página web de implementación de la comunidad RDFa](#). Ambas fuentes están en constante cambio. Por cierto, la segunda fuente forma parte de un [sitio web de una comunidad](#) que contiene otros ejemplos de uso de RDFa, información general, así como información sobre como participar en la comunidad.
- En particular, los ejemplos de fragmentos RDFa se pueden probar usando el [editor en tiempo real RDFa 1.1](#) que también puede mostrar una representación visual de los datos estructurales subyacentes.

Editor en tiempo real de RDFa

The screenshot shows a web browser window with the URL `rdfa.info/play/`. The page has a navigation bar with links for `RDFa`, `Play`, `Documentation`, `Tools`, and `Developers`. A yellow warning banner at the top states: "WARNING: This is beta software, there are bugs. This tool is not yet ready for production use. That said, it's pretty useful already, and the bugs preventing production use will be sorted out soon." Below the warning, there are tabs for "Examples" with sub-tabs for `Person`, `Social Network`, `Event`, `Place`, `Product`, and `SVG`. The main content area is split into two columns. The left column contains instructions: "Paste your HTML+RDFa code in here, or click on an example above. A preview of the page will appear to the right with a data visualization below. You can also see the raw data by clicking on the 'Raw Data' tab." It also includes a link to "You can learn more about RDFa by reading the" followed by a code block:

```
<span vocab="http://schema.org/" typeof="Document">
  <a property="url" href="http://www.w3.org/TR/rdfa-primer/">
    <span property="title">RDFa 1.1 Primer</span></a>
</span>
```

 The right column contains similar instructions: "Paste your HTML+RDFa code in here, or click on an example above. A preview of the page will appear to the right with a data visualization below. You can also see the raw data by clicking on the 'Raw Data' tab. You can learn more about RDFa by reading the [RDFa 1.1 Primer](#)." At the bottom of the main area, there are two tabs: `Visualization` (selected) and `Raw Data`.



Software to Pull RDFa from XHTML

- Usage pattern of RDFa:
 - Extract embedded RDFa triples
 - Load them into a triplestore in memory or on disk
 - Develop an application around your extracted data
- RDFa2RDFXML.xsl XSL StyleSheet
 - <http://www-sop.inria.fr/acacia/soft/RDFa2RDFXML.xsl>
- Web Service which receives an XHTML and returns RDFa
 - Use <http://ben.adida.net/card> to test service at <http://torrez.us/services/rdfa/>
 - If you want to use it in your semantic web application, simply point your SPARQL query to: <http://torrez.us/services/rdfa/>[URL of HTML page]

¿Qué dice GOOGLE?

- Existen argumentos para preferir cualquiera de los estándares existentes, pero creemos que los microdatos ofrecen el equilibrio necesario entre la posibilidad de ampliación de RDFa y la sencillez de los microformatos, así que ese es el formato por el que hemos optado.
- Google recomienda utilizar schema.org para marcar los vídeos. schema.org es un nuevo esquema de marcado en páginas que utiliza un vocabulario reconocido por Google, Yahoo! y Microsoft. Para obtener más información sobre schema.org, consulta [esta página](#) (Google sigue admitiendo el contenido existente marcado con Facebook Share y RDFa).
- Cuando la información de vídeo se ha marcado en el cuerpo de una página web, Google puede identificarla y utilizarla para mejorar los resultados de búsqueda.

<https://support.google.com/webmasters/answer/162163?hl=es>
<https://support.google.com/webmasters/answer/1211158>

Google Knowledge Graph

- **Knowledge Graph es una base de conocimiento usada por Google para mejorar los resultados obtenidos con su motor de búsqueda** mediante información de búsqueda semántica recolectada de una amplia gama de recursos.
 - Se añadió al motor de búsqueda de Google en 2012
- Provee información estructurada y detallada acerca de un tema además de una lista de enlaces a otros sitios.
- Esta información se deriva de muchos recursos, que incluyen el CIA World Factbook, Freebase y Wikipedia
 - Su red semántica contenía más de **570 millones de objetos y más de 18 mil millones de sucesos acerca de –y relaciones entre– esos diferentes objetos** que se usan para entender el significado del término índice ingresado en la búsqueda
- URL:
 - Knowledge Graph
<http://www.google.com/insidesearch/features/search/knowledge.html>
 - Freebase: http://wiki.freebase.com/wiki/Main_Page

GOOGLE, facebook y yahoo

Actualmente, Google usa las tecnologías semánticas para mostrar **snippets enriquecidos** e incorporar información a su **Knowledge graph**. Por otro lado, las redes sociales **Facebook** y **Twitter** las usan para que, a la hora de publicar páginas, éstas aparezcan con información personalizada o adicional haciendo uso de los vocabularios **Open Graph** y **Twitter Cards**.

Google reconoce tanto el formato de marcado **Facebook Share** como el formato de marcado **Yahoo! SearchMonkey RDFa**. El uso de uno o de ambos formatos para marcar vídeos directamente en HTML ayuda a Google a interpretar y a presentar mejor el contenido de vídeo. Asegúrate de que el marcado aparezca en el contenido HTML sin ejecutar JavaScript ni Flash.

Google Vídeo soporta Facebook Share y RDFa

artículo de Javier Casares publicado el 15 de septiembre de 2009...

Aunque ya se estaba trabajando duro para hacer crecer la cantidad de vídeos en los resultados de búsqueda a través de sistemas como el media RSS, parece que Google ha decidido que ha de utilizar las herramientas libres de terceros para poder seguir aumentando.

Es por esto que **han anunciando la disponibilidad** de leer el sistema de vídeos de Facebook y el RDFa que utiliza Yahoo! Searchmonkey. De esta forma, podemos incluir códigos como los siguientes en nuestro sitio (personalmente recomiendo el RDFa al ser estándar):

RDFa (Yahoo! SearchMonkey):

```
<object width="512" height="296" rel="media:video" resource="http://example.com/video_object.swf?id=12345" xmlns:media="http://search.yahoo.com/searchmonkey/media/" xmlns:dc="http://purl.org/dc/terms/">
  <param name="movie" value="http://example.com/video_object.swf?id=12345" />
  <embed src="http://example.com/video_object.swf?id=12345" type="application/x-shockwave-flash" width="512" height="296"></embed>
  <a rel="media:thumbnail" href="http://example.com/thumbnail_preview.jpg" />
  <a rel="dc:license" href="http://example.com/terms_of_service.html" />
  <span property="dc:description" content="Cute Overload defines Baroo? as: Dogspeak for 'Whut the...?' Frequently accompanied by the Canine Tilt and/or wrinkled brow for enhanced effect." />
  <span property="media:title" content="Baroo? - cute puppies" />
  <span property="media:width" content="512" />
  <span property="media:height" content="296" />
  <span property="media:type" content="application/x-shockwave-flash" />
  <span property="media:region" content="us" />
  <span property="media:region" content="uk" />
  <span property="media:duration" content="63" /></object>
```

Facebook Share:

```
<meta name="title" content="Baroo? - cute puppies" />
<meta name="description" content="The cutest canine head tilts on the Internet!" />
<link rel="image_src" href="http://example.com/thumbnail_preview.jpg" />
<link rel="video_src" href="http://example.com/video_object.swf?id=12345"/>
<meta name="video_height" content="296" />
<meta name="video_width" content="512" />
<meta name="video_type" content="application/x-shockwave-flash" />
```

GOOGLE , Facebook y Schema.org

- El protocolo Open Graph de Facebook cumple correctamente su función, pero no proporciona la información detallada que necesitan los motores de búsqueda para mejorar la experiencia del usuario. Una única página web puede tener muchos componentes y tratar de diversos temas. Aunque marques tu contenido con el protocolo Open Graph de Facebook, schema.org ofrece una forma adicional de proporcionar información más detallada sobre entidades particulares de la página.
- Por ejemplo, una página sobre un grupo de música puede incluir algunos de los siguientes elementos (o todos ellos):
 - una lista de álbumes,
 - el precio de cada álbum,
 - una lista de las canciones de cada álbum junto con un enlace para escuchar fragmentos de cada canción,
 - una lista de próximos conciertos,
 - las biografías de los miembros del grupo.
- Cada uno de estos fragmentos de información se puede representar como un tipo de contenido de schema.org. Si marcas estos fragmentos con el esquema de schema.org, los motores de búsqueda podrán entender mejor el contenido de tu página.
- Google no utiliza el marcado para mejorar la posición del sitio en los resultados de búsqueda en este momento. No obstante, los fragmentos enriquecidos pueden destacar tus páginas en los resultados de búsqueda, por lo que es posible que se note un aumento en el tráfico.

<https://support.google.com/webmasters/answer/162163?hl=es>
<https://support.google.com/webmasters/answer/1211158>

GRDDL (pronunciado “griddle”)

- Gleaning Resource Descriptions from Dialects of Languages. Recomendación del W3C el 11 de septiembre de 2007.
- Objetivo: enlazar un documento XML con una transformación que extraiga su semántica.
- Una técnica (mecanismo sencillo) que permite obtener datos en RDF desde documentos (dialectos) XML y en particular desde XHTML. Ej: Extraer microformatos, extraer RDFa.
- GRDDL nos ayuda a asociar los algoritmos de transformación necesarios para cada dialecto, permitiendo la combinación de conceptos expresados en diferentes formatos.
- Los autores pueden asociar explícitamente procedimientos de transformaciones XSLT, usando el elemento “link” en la cabecera (head) del documento. También pueden usarse funciones de Javascript.

Alternativamente, la información necesitada para obtener la transformación puede ser mantenida en perfiles de metadatos asociados o en el espacio de nombres del documento.

GRDDL: Marcado

GRDDL es un marcado para :

- Declarar que un documento XML incluye datos “gleanable” (PROFILE) o
- Enlazar (linking) a un algoritmo (típicamente XSLT) para “gleaning” los datos RDF del documento (TRANSFORMATION).

El marcado incluye:

1. Un atributo namespace-qualified para usar en documentos XML de propósito general.

```
xmlns:grddl='http://www.w3.org/2003/g/data-view#'  
grddl:transformation="glean_title.xsl"
```

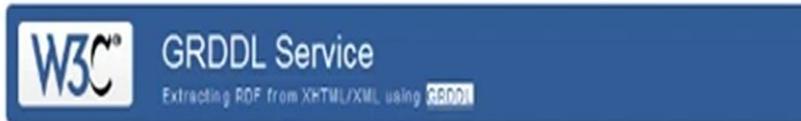
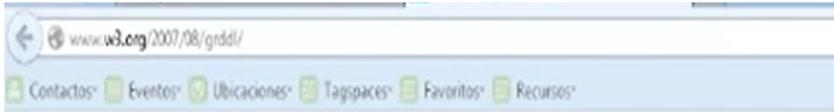
2. Una relación “profile-qualified link” para usar en documentos XHTML válidos.

```
<head profile="http://www.w3.org/2003/g/data-view">  
<link rel="transformation" href="glean_title.xsl" />
```

Implementaciones GRDDL

- Lista de implementaciones:
 - <http://esw.w3.org/topic/GrddlImplementations>
 - Ejemplo: La framework Jena ofrece una implementación de GRDDL disponible en la página <http://jena.sourceforge.net/grddl/> , desde donde pueden descargarse los .jar que implementan la librería ofrecida, denominada **Jena GRDDL Reader**
- Transformaciones GRDDL actualmente disponibles:
 - <http://www.w3.org/wiki/CustomRdfDialects>
 - Ejemplo: RDFa Profile (<http://ns.inria.fr/grddl/rdfa/>)
 - RDFa2RDFXML.xsl
 - Ver VIDEO : <http://www.youtube.com/watch?v=XT3DJJn4tRk>

GRDDL Service



Validate by URI Validate by Direct Input

Validate by URI

Address of the document to extract RDF from:

Output format: RDF/XML as application/rdf+xml

get GRDDL results

Stuff used to build this service

Python script based on
python-librdf
python, apache, etc.
script \$Revision: 1.46 \$ of \$Date: 2011/01/12 13:47:38 \$
by Dominique Hazaël-Massieux

Este fichero XML no parece tener ninguna información de estilo asociada. Se muestra debajo el árbol del documento.

```
<?xml:lang es >
<rdf:RDF>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <rdf:type rdf:resource="http://www.w3.org/2006/vcard/vCard"/>
    <rdf:Description>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <us:photo rdf:resource="http://www.vlbe.net/SemanticKipu/vtatars/anatax_perfil_1.jpg"/>
    <rdf:Description>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <us:fn>Alejandro</us:fn>
    <rdf:Description>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <us:nickname>Alejandro</us:nickname>
    <rdf:Description>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <us:url rdf:resource="http://www.enet-tech.com"/>
    <rdf:Description>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <us:title>Desarrollador Senior</us:title>
    <rdf:Description>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <us:email rdf:resource="mailto:fvnm@gmail.com"/>
    <rdf:Description>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <us:bday>1985-12-14</us:bday>
    <rdf:Description>
  <rdf:Description rdf:nodeID="r13593460038351z2">
    <rdf:type rdf:resource="http://www.w3.org/2006/vcard/vAddress"/>
    <rdf:Description>
```

Aquí se puede observar como se a obtenido un archivo RDF de hCard con GRDDL

Transformaciones GRDDL

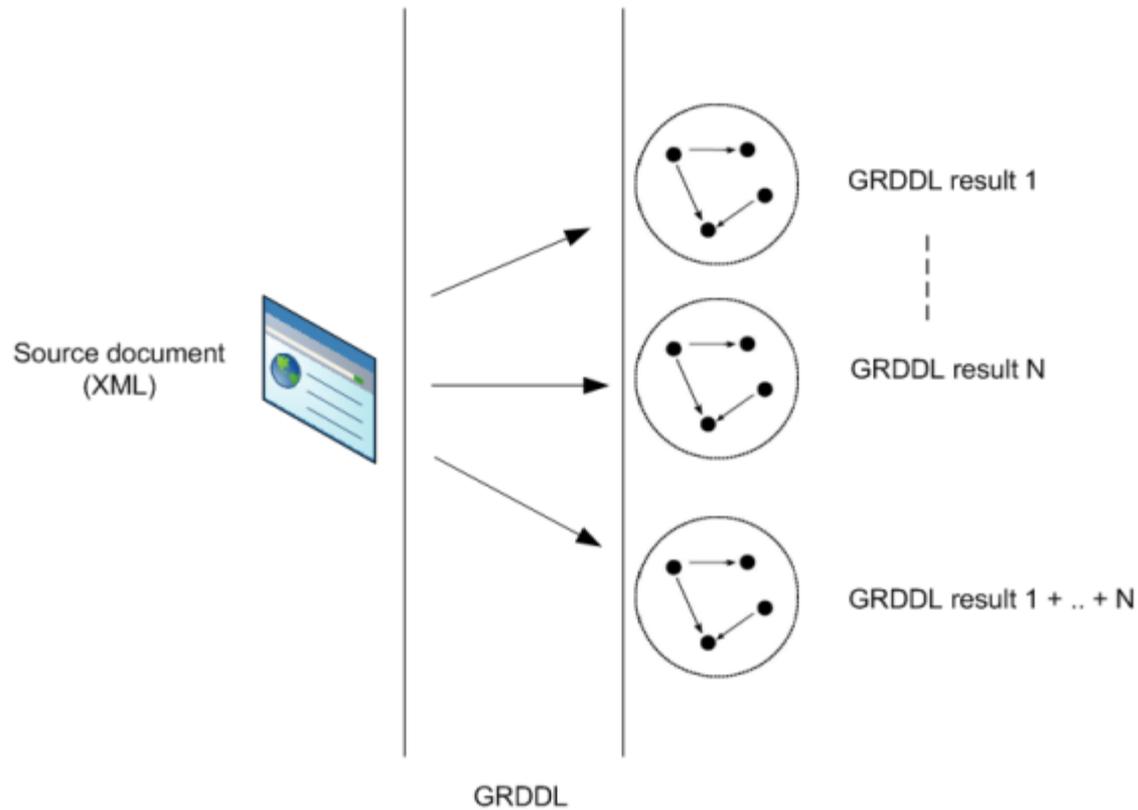
HTML-based Dialects

Dialect	Domain	Schema	Mapping	Status
Dublin Core in HTML (legacy)	Bibliographic Metadata	Dublin Core	dc-extract	Valid 2003-2003, now superseded
Dublin Core in HTML (current)	Bibliographic Metadata	Dublin Core	dc-html2rdfxml.xml	Maintained by DCMI
eRDF	generic	RDF	extract-rdf.xml , profile	by IanDavis
GeoURL markup	Geo (lat/long)	GeoInfo	grokGeoURL.xml	e.g. Joe Lambda's homepage
glink	HTML links/link@rel	#html, DC	grolinks.xml	maintained by DannyAyers
hCalendar	Calendars and Events	RDF Calendar	glean-hcal.xml , hcal_profile	maintained by DanConnolly since Dec 2002
hCard	Contact Information	vcard ontologies (2006)	hcard2rdf.xml , hcard_profile	maintained by Walsh, Halpin, and Suda since Nov 2006; used in xtech schedule
hDoap	Software project description	DOAP	hDoap	maintained by DannyAyers
hReview	Opinions, Ratings and Reviews	RDF Review	hreview2rdfxml.xml	See RDF Review
hTurtle	generic	RDF	XSLT service , profile	by SeanPalmer
metaprof	meta/link@rel + generic	various	xh2rdf.xml	maintained by MasahideKanzaki
RDFa	generic	RDF	RDFa profile	by Fabien Gandon, collaborating with the GRDDL and RDFa WGs
Rel-License	Licenses	Creative Commons	grokCC.xml	e.g. Joe Lambda's homepage
Rel-Tag	Tags, Keywords, Categories	Tag Ontology , SKOS , DC	?	(todo DannyAyers)
Site Summaries	Syndication	RSS	html2rss.xml	in production on the W3C homepage
XFN	Social Graph	XFN + # as namespace	grokXFN.xml	maintained(?) by DanConnolly
Site Summaries	Syndication	RSS	html2rss.xml	in production on the W3C homepage
XFN	Social Graph	XFN+hCard as FoaF	hFoaF.xml	Maintained by MartinMcEvoy
hAudio	Audio Syndication	hAudio as Music Ontology	Mo-hAudio.xml	Maintained by MartinMcEvoy
RDF-3T	generic	RDF	RDF-3T	Maintained by MartinMcEvoy

Non-HTML based dialects

Dialect	Domain	Schema	Mapping
RSS	Syndication	RSS 1.0	feed-rss1.0.xml
Atom	Syndication	Atom/OWL	atom2rdf_xslt-1.0.xml
Atom	Syndication	RSS 1.0	atom2rss1.0.xml
OPML	Feedlists	FOAF/RSS 1.0	opml2blogroll.xml
OPML	Feedlists	FOAF/RSS 1.0/SKOS	opml2skosroll.xml
Metalink	File sharing	metalink	(under revision 2007-10)
APML	Attention Profiling	apml	-
SAWSDL	web service description	SAWSDL	profile, SAWSDL_GRDDL.xml
XTM	topic maps	XTM	profile, XTM_GRDDL.xml

GRDDL: Múltiples resultados



GRDDL, Ejemplo I: de XHTML a RDF

Incluir un perfil (profile) URI para GRDDL en el documento para declarar que el marcado puede ser interpretado usando GRDDL.

```
<html xmlns="http://www.w3.org/1999/xhtml">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Some Document</title>
    <link rel="transformation" href="http://www.w3.org/2000/06/dc-extract/dc-extract.xsl" />
  <meta name="DC.Subject" content="ADAM; Simple Search; Index+; prototype" />
  ...
  </head>
  ...
</html>
```



```
<rdf:RDF xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about="">
    <dc:subject>ADAM; Simple Search; Index+;
  prototype</dc:subject>
  </rdf:Description>
</rdf:RDF>
```

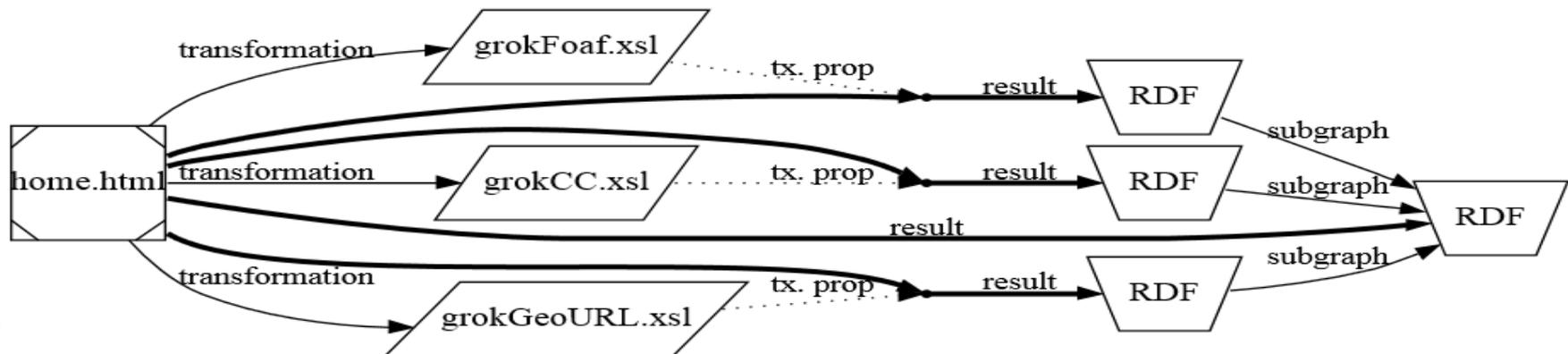
GRDDL, Ejemplo II:

Múltiples transformaciones en XHTML

Un documento XHTML puede ajustarse a una serie de dialectos simultáneamente y vincularse a más de una transformación GRDDL. Sin embargo, dado que el atributo “href” de los elementos “link” y “a” aceptan únicamente una sola referencia URI, deben ser utilizados varias instancias de estos elementos para establecer múltiples enlaces:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
<head profile="http://www.w3.org/2003/g/data-view">
  <title>Joe Lambda's Home page [an example of RDF in XHTML]</title>

  <link rel="transformation" href="http://www.w3.org/2003/12/rdf-in-xhtml-xslts/grokFOAF.xsl" />
  <link rel="transformation" href="http://www.w3.org/2003/12/rdf-in-xhtml-xslts/grokCC.xsl" />
  <link rel="transformation" href="http://www.w3.org/2003/12/rdf-in-xhtml-xslts/grokGeoURL.xsl" />
  ...
```



GRDDL: Ejemplo III

Añadir un elemento “link” conteniendo la referencia a la transformación específica GRDDL (aquí HTML en hCalendar es transformada a RDF).

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
<head profile="http://www.w3.org/2003/g/data-view">
  <title>Robin's Schedule</title>
  <link rel="transformation"
href=http://www.w3.org/2002/12/cal/glean-hcal/>
</head>
<body> .....
```

GRDDL: Usando GRDDL con XML namespace

Asociar una transformación GRDDL no solo con un documento individual , sino con un conjunto de “dialectos” , que comparten un espacio de nombres XML. Para esto, se incluye un *grddl:namespaceTransformation* property en un resultado GRDDL del documento del espacio de nombres.

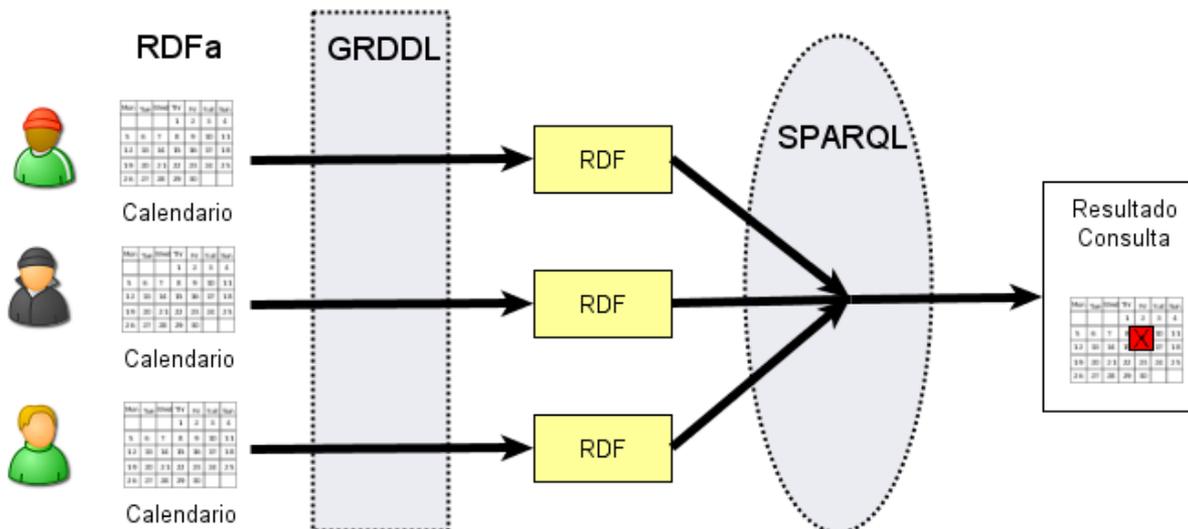
```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dataview="http://www.w3.org/2003/g/data-view#">
  <rdf:Description rdf:about="http://www.w3.org/2004/01/rdxh/p3q-nsexample">
  <dataview:namespaceTransformation
    rdf:resource="http://www.w3.org/2004/01/rdxh/grokP3Q.xsl"/>
  </rdf:Description>
</rdf:RDF>
```



“Todas en conjunto”: Ejemplo de extracción de datos usando RDFa, GRDDL y SPARQL:

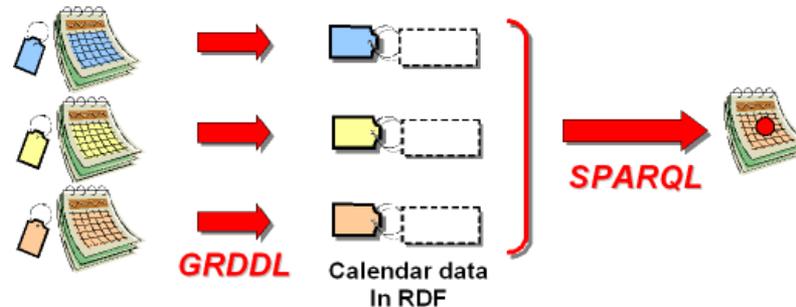
Se desea establecer una reunión entre tres personas, que tienen publicados en sus sitios Web los calendarios de sus citas y eventos.

1. Estos datos están expuestos en páginas XHTML de forma gráfica, que además incluye información en RDFa.
2. Una herramienta nos permite extraer, mediante GRDDL, los datos de sus calendarios en un formato homogéneo y fácil de tratar (RDF), para poder procesarlo posteriormente.
3. Se realiza una consulta sobre la disponibilidad de las personas para un cierto día a una hora concreta. Los datos consultados están en formato RDF y la consulta se podría realizar mediante SPARQL.
4. La herramienta procesa y analiza el resultado obtenido, concluyendo si las personas están disponibles en el instante que se había elegido previamente.



GRDDL and XHTML: Scheduling a Meeting

- Example 1:
 - “Jane is trying to see if at any point next year she can schedule a meeting with all three of her friends, despite the fact that all of her friends publish their calendar data in different ways.”
- GRDDL provides a number of ways for GRDDL transformations to be associated with content, each of which is appropriate in different situations.
 - The simplest method for authors of HTML content is to embed a reference to the transformations directly using a `link` element in the `head` of the document.



Robin amigo de Jane usa hCalendar

MicroFormat

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
xml:lang="en" lang="en">
<head>
<title>Robin's Schedule</title>
</head>
<body>
<ol class="schedule">
<li>2006
<ol>
<li class="vevent">
<strong class="summary">Fashion
Expo</strong> in
<span class="location">Paris,
France</span>:
<abbr class="dtstart" title="2006-10-
20">Oct 20</abbr> to
<abbr class="dtend" title="2006-10-
23">22</abbr>
</li>
...
```

```
</ol>
</li>
<li>2007
<ol>
<li class="vevent">
<strong class="summary">Web
Design Conference</strong> in
<span class="location">Edinburgh,
UK</span>:
<abbr class="dtstart"
title="2007-01-08">Jan 8</abbr> to
<abbr class="dtend" title="2007-
01-11">10</abbr>
</li>
...
</ol>
</li>
</ol>
</body>
</html>
}
```

Robin amigo de Jane usa hCalendar MicroFormat

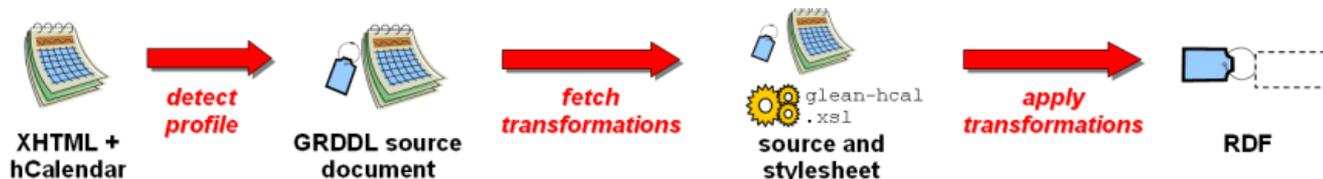
- To explicitly relate the data in this document to the RDF data model the author needs to make two changes.

1. Add a profile attribute to the head element to denote that her document contains GRDDL metadata.

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
  "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Robin's Schedule</title>
  </head>
  <body>
    ...
```

2. Add a **link** element containing the reference to the specific GRDDL transformation for converting HTML containing **hcalendar** patterns into RDF.

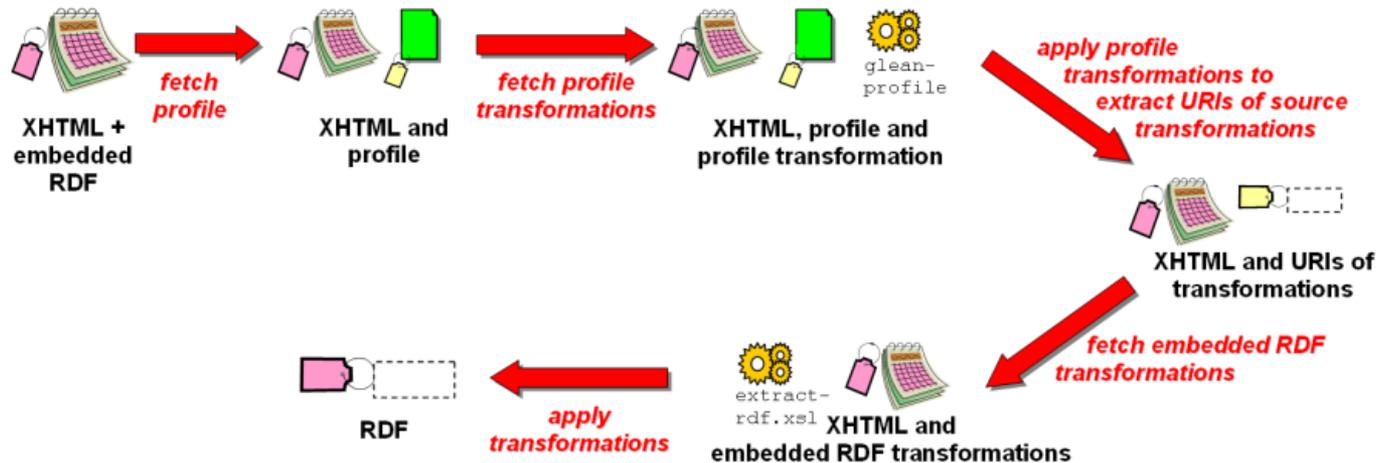
```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
  "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Robin's Schedule</title>
    <link rel="transformation" href="http://www.w3.org/2002/12/cal/glean-hcal"/>
  </head>
  <body>
    ...
```



David, amigo de Jane usa embedded RDF

- Embedded RDF has a link to a GRDDL transformation in its profile document.

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN" "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
  <head profile="http://purl.org/NET/erdf/profile">
    <title>Where Am I</title>
    <link rel="schema.cal" href="http://www.w3.org/2002/12/cal#" />
  </head>
  <body>
    <p class="-cal-vevent" id="tiddlywinks">
      From <span class="cal-dtstart" title="2006-10-07">7 October, 2006</span>
      to <span class="cal-dtend" title="2006-10-13">12 October, 2006</span>
      I will be attending the <span class="cal-summary">National Tiddlywinks
      Championship</span> in
      <span class="cal-location">Bognor Regis, UK</span>.
    </p>
    ...
  </body>
</html>
```



Jane usa RDFa para su planificación

- Jane stores her calendar directly in RDFa (<http://www.w3.org/TR/grddl-primer/janeschedule.html>)
 - GRDDL Transformation for RDFa to convert RDFa to RDF/XML is available at:
 - <http://www.w3.org/TR/2007/NOTE-grddl-primer-20070628/RDFa2RDFXML.xsl>

- Example:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML+RDFa 1.0//EN"
    "http://www.w3.org/Markup/DTD/xhtml+rdfa-1.dtd">
<html xmlns:cal="http://www.w3.org/2002/12/cal/icaltzd#"
xmlns:xs="http://www.w3.org/2001/XMLSchema#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns="http://www.w3.org/1999/xhtml">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Jane's Blog</title>
    <link rel="transformation" href="RDFa2RDFXML.xsl"/>
  </head>
  <body>
    <p about="#event1" class="cal:Vevent">
      <b property="cal:summary">Weekend off in Iona</b>:
      <span property="cal:dtstart" content="2006-10-21" datatype="xs:date">Oct 21st</span>
      to <span property="cal:dtend" content="2006-10-21" datatype="xs:date">Oct 23rd</span>.
      See <a rel="cal:url" href="http://freetime.example.org/">FreeTime.Example.org</a> for
      info on <span property="cal:location">Iona, UK</span>.
    </p>
    ...
  </body>
</html>
```

Fusionando las tres agendas

- One of the advantages of the RDF data model is that RDF data can be easily merged by adding it to a RDF store
 - Jane can merge and query all the calendars together once they are transformed into RDF through GRDDL
 - Jane uses SPARQL to query her data, which automatically merges the calendar data sources before running the query.
- Next slide SPARQL query explanation:
 - The **SELECT** line determines which variables will appear in the results, here one of the start dates, one of the stop dates, a location and a summary.
 - The **FROM** lines identify the data sources to use in the query, in this case the RDF/XML derived from Jane, David and Robin's original documents.
 - The **WHERE** section provides a pattern which can match three events.
- Results of next slide query:

start1	stop1	loc1	summ1
"2007-01-08	"2007-01-11"	Edinburgh, UK"	Web Design Conference"

Ejemplo 1 SPARQL

```
PREFIX ical: <http://www.w3.org/2002/12/cal/icaltzd#>
PREFIX xs: <http://www.w3.org/2001/XMLSchema#>
SELECT ?start1 ?stop1 ?loc1 ?summ1 ?summ2 ?summ3
FROM <http://www.w3.org/TR/grddl-primer/janeschedule.rdf>
FROM <http://www.w3.org/TR/grddl-primer/robin-hcal-grddl.rdf>
FROM <http://www.w3.org/TR/grddl-primer/david-erdf.rdf>
WHERE
{
  ?event1 a ical:Vevent;
           ical:summary ?summ1 ;
           ical:dtstart ?start1 ;
           ical:dtend ?stop1 ;
           ical:location ?loc1.
  ?event2 a ical:Vevent;
           ical:summary ?summ2 ;
           ical:dtstart ?start2;
           ical:dtend ?stop2;
           ical:location ?loc2.
  ?event3 a ical:Vevent;
           ical:summary ?summ3 ;
           ical:dtstart ?start3;
           ical:dtend ?stop3;
           ical:location ?loc3.
}
```

```
FILTER ( ?event1 != ?event2 && ?event2 !=
?event3 && ?event1 != ?event3 ) .
  FILTER ( xs:string(?start1) =
xs:string(?start2) ) .
  FILTER ( xs:string(?stop1) =
xs:string(?stop2) ) .
  FILTER ( xs:string(?loc1) = xs:string(?loc2)
) .
  FILTER ( xs:string(?start1) =
xs:string(?start3) ) .
  FILTER ( xs:string(?stop1) =
xs:string(?stop3) ) .
  FILTER ( xs:string(?loc1) = xs:string(?loc3)
) .
  FILTER ( xs:string(?start3) =
xs:string(?start2) ) .
  FILTER ( xs:string(?stop3) =
xs:string(?stop2) ) .
  FILTER ( xs:string(?loc3) = xs:string(?loc2)
) .
  FILTER ( xs:string(?summ1) <=
xs:string(?summ2) ) .
  FILTER ( xs:string(?summ2) <=
xs:string(?summ3) ) .
}
```

Mashing-Up Microformatos: Reservando un Hotel

- Example 2:
 - “Jane wants to book a hotel in Edinburgh”
 - Combines data dialects as different as reviews and social networks in order to guarantee the booking a hotel with a high review from a trusted friend.
- This example highlights the role of GRDDL in aggregating data from a variety of different formats and of using RDF as a common format to "mashup" all sorts of data, not just calendar data.
 - XFN file can be converted to RDF with the use of another GRDDL Transform for XFN
 - <http://www.w3.org/TR/2007/NOTE-grddl-primer-20070628/grokXFN.xsl>
 - The hotel review file uses **hReview** that can be converted to RDF with the following transformation:
 - <http://www.w3.org/TR/2007/NOTE-grddl-primer-20070628/hreview2rdfxml.xsl>
- Using GRDDL we can glean information, including the ratings about the hotels and the reviews given by Jane’s friends.

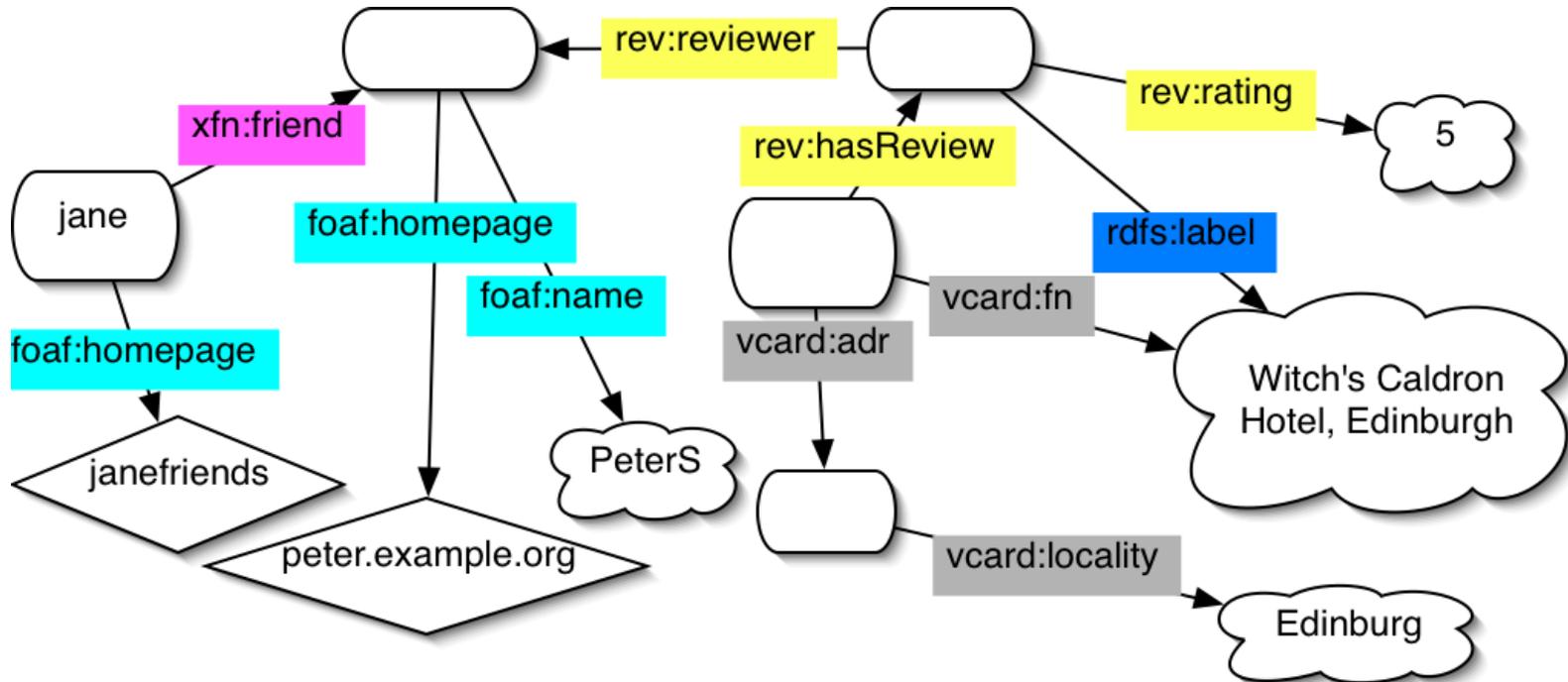
Microformato XFN

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
  "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd" >
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">
  <head profile="http://www.w3.org/2003/g/data-view
    http://gmpg.org/xfn/11>
    <link rel="transformation" href="http://www.w3.org/TR/grddl-
      primer/grokXFN.xsl" />
    <title>Jane's XFN List</title>
  </head>
  <body>
    <h1>Jane's <abbr title="XHTML Friends Network">XFN</abbr> List</h1>
    <ul class="xoxo">
      <li class="vcard"><a href="http://peter.example.org/" class="url
fn" rel="met colleague friend">Peter Smith</a></li>
      <li class="vcard"><a href="http://john.example.org/" class="url
fn" rel="met">John Doe</a></li>
      <li class="vcard"><a href="http://paul.example.org/" class="url
fn" rel="met">Paul Revere</a></li>
    </ul>
  </body>
</html>
```

Microformato hReview

```
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Hotel Reviews from Example.com</title>
    <link rel="transformation"
      href="http://www.w3.org/TR/grddl-primer/hreview2rdfxml.xsl"/>
  </head>
  <div class="hreview" id="_665">
    <div class="item vcard">
      <b class="fn org">Witch's Caldron Hotel, Edinburgh</b>
      <ul>
        <li>
          <a class="url" href="http://witches.example.com/">Homepage</a>
        </li>
      </ul>
      <span><span class="rating">5</span> out of 5 stars</span>
      <ul>
        <li class="adr">
          <div class="type">intl postal parcel work</div>
          <div class="street-address">313 Cannongate</div>
          <div><span class="locality">Edinburgh</span>, <span class="postal-code">EH8 8DD</span>
            </div> <span class="country-name">United Kingdom</span></div>
        </li>
      </ul>
      <div class="tel"><abbr class="type" title="work msg">Homepage</abbr>: <a
        class="value" href="tel:+44 1317861235">+44 1317862235</a></div>
```

Fusionando “reviews” y gente



Fuente: <http://www.w3.org/TR/grddl-primer/>

Ejemplo 2 SPARQL Query

- Find hotels with specific ratings or higher from a group of her trusted friends”

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rev: <http://www.purl.org/stuff/rev#>
PREFIX vcard: <http://www.w3.org/2006/vcard/ns#>
PREFIX xfn: <http://gmpg.org/xfn/11#>
SELECT DISTINCT ?rating ?name ?region ?homepage ?xfnhomepage ?hotelname
FROM <http://www.w3.org/TR/grddl-primer/janefriends.rdf>
FROM <http://www.w3.org/TR/grddl-primer/hotel-data.rdf>
WHERE {
  ?x rev:hasReview ?review;
    vcard:ADR ?address;
    vcard:FN ?hotelname.
  ?review rev:rating ?rating .
  ?address vcard:Locality ?region.
  FILTER (?rating > "2" && ?region = "Edinburgh").
  ?review rev:reviewer ?reviewer.
  ?reviewer foaf:name ?name;
    foaf:homepage ?homepage.
  ?y xfn:friend ?xfnfriend.
  ?xfnfriend foaf:homepage ?xfnhomepage.
  FILTER (?xfnhomepage = ?homepage).
}
```

rating	name	region	homepage	xfnhomepage	hotelname
"5"	"PeterS"	"Edinburgh"	<http://peter.example.org/>	<http://peter.example.org/>	"Witch's Caldron Hotel, Edinburgh"

Referencias

Microdatos: <http://www.mmfilesi.com/blog/que-son-los-microdatos/>

Microformats

<http://microformats.org/>

<http://www.mmfilesi.com/blog/microformatos-o-microdatos/>

Alex Faaborg content on Microformats

<http://blog.mozilla.com/faaborg/2006/12/11/microformats-part-0-introduction/>

<http://people.mozilla.com/~faaborg/files/20070416-web20Expo/faaborg-Web20Expo.pdf>

RDFa

Introducing RDFa

<http://www.xml.com/lpt/a/1691>

Introducing RDFa, Part Two

<http://www.xml.com/lpt/a/1698>

RDFa Primer 1.0

<http://www.w3.org/TR/xhtml-rdfa-primer/>

RDF/A Syntax

<http://www.w3.org/2001/sw/BestPractices/HTML/2005-rdfa-syntax>

GRDDL:

Primer: Using GRDDL & Microformats to Aggregating data

<http://suda.co.uk/sandbox/GRDDL/Primer.htm>

GRDDL Primer

<http://www.w3.org/TR/2007/NOTE-grddl-primer-20070628/>

GRDDL Digital Library Example

<http://www.sop.inria.fr/acacia/personnel/Fabien.Gandon/tmp/grddl/rdfaprimer/PrimerRDFaSection.html>

• Appendix B – Semantic Web. Dr. Diego Lz. De Ipiña GZ. De Artaza.

<http://paginaspersonales.deusto.es/dipina> . Deusto Tech. Universidad de Deusto. Deustuko Unibertsitatea