

Posicionamiento Semántico

RDFa, Microdatos, etc.

Jose Emilio Labra Gayo

Departamento de Informática

Universidad de Oviedo

POSICIONAMIENTO SEMÁNTICO

Hipótesis:

Cuanta más semántica tiene una página, más fácil de localizar

Objetivo:

Añadir semántica a las páginas Web

¿Cómo añadir semántica a HTML?

Incluir RDF como comentarios en HTML

Enlazar a ficheros RDF externos

GRDDL

Microformatos

RDFa

Microdatos

Proyecto Schema.org

Herramientas Google & JSON-LD

RDF como comentarios en HTML

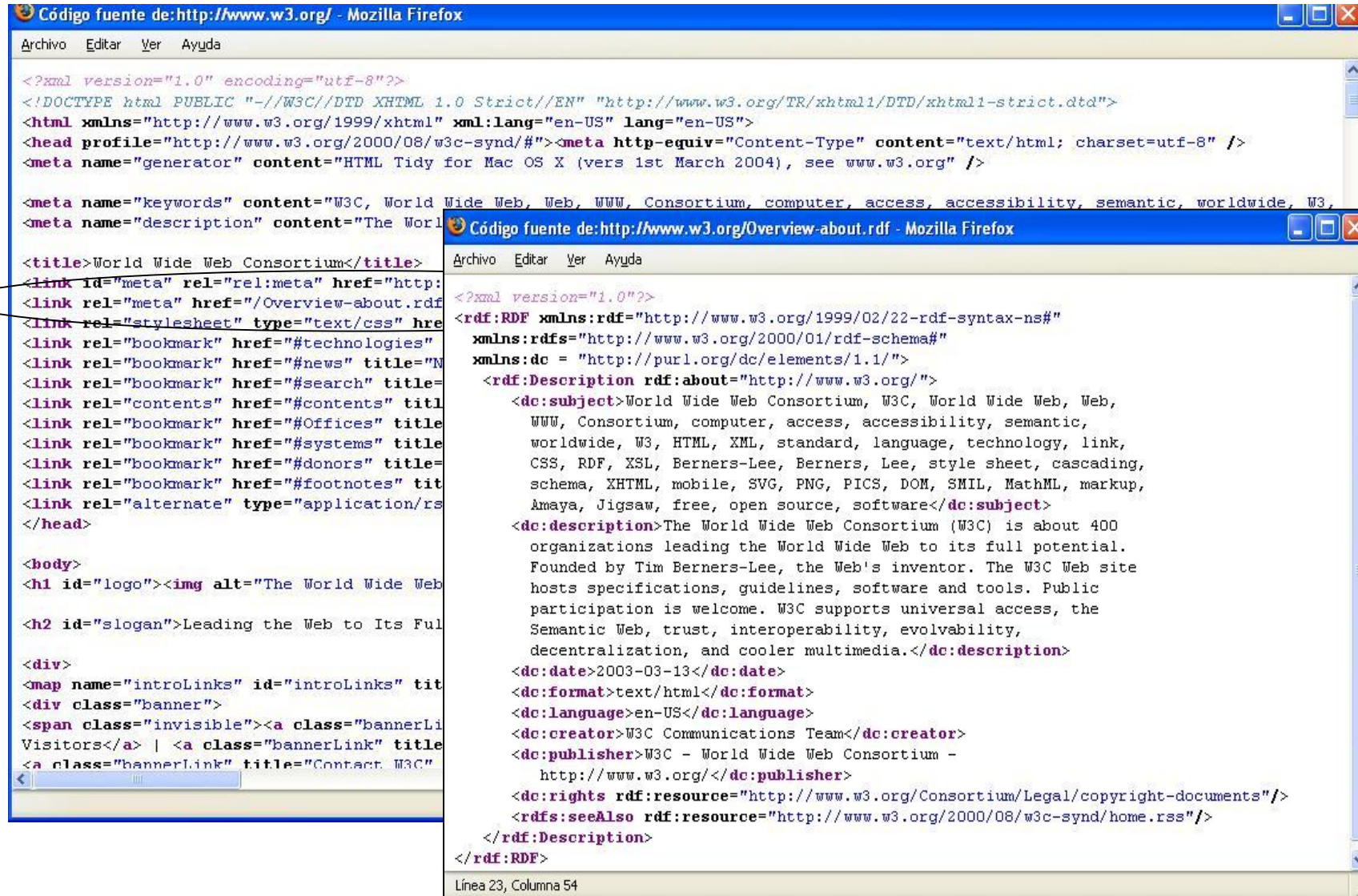
Problema: no es posible incluir RDF sin más

RDF/XML tiene una sintaxis que impide la validación

Solución: ¿incluir RDF como comentarios?

```
<div id="f-lastmod"> This page was last modified 12:26, 11 July 2006.</div>
<div id="f-copyright">This wiki is licensed to the public under a
  <a href="http://creativecommons.org/licenses/by/3.0/"
    class='external '
    title="http://creativecommons.org/licenses/by/3.0/"
    rel="nofollow">Creative Commons Attribution 3.0</a> license.<br/>
  Your use of this wiki is governed by the
  <a href="/CWiki:Terms_of_Use">Terms of Use</a>.
<!-- <rdf:RDF xmlns="http://web.resource.org/cc/"
  xmlns:dc="http://purl.org/dc/elements/11/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <Work rdf:about="">
    <license rdf:resource="http://creativecommons.org/licenses/by/2.5/" />
    <dc:type rdf:resource="http://purl.org/dc/dcmitype/Text" />
  </Work>
  <License rdf:about="http://creativecommons.org/licenses/by/2.5/">
    <permits rdf:resource="http://web.resource.org/cc/Reproduction"/>
    <permits rdf:resource="http://web.resource.org/cc/Distribution"/>
    <requires rdf:resource="http://web.resource.org/cc/Notice"/>
    <requires rdf:resource="http://web.resource.org/cc/Attribution"/>
    <permits rdf:resource="http://web.resource.org/cc/DerivativeWorks"/>
  </License>
</rdf:RDF> -->
</div>
<div id="f-about">
<a href="/CWiki:About" title="CWiki:About">About CWiki</a>
```

Enlazar a un fichero mediante <link>



```
Código fuente de:http://www.w3.org/ - Mozilla Firefox
Archivo  Editar  Ver  Ayuda

<?xml version="1.0" encoding="utf-8"?>
<!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" xml:lang="en-US" lang="en-US">
<head profile="http://www.w3.org/2000/08/w3c-synd/#"><meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<meta name="generator" content="HTML Tidy for Mac OS X (vers 1st March 2004), see www.w3.org" />

<meta name="keywords" content="W3C, World Wide Web, Web, WWW, Consortium, computer, access, accessibility, semantic, worldwide, W3,
<meta name="description" content="The World

<title>World Wide Web Consortium</title>
<link id="meta" rel="rel:meta" href="http:
<link rel="meta" href="/Overview-about.rdf
<link rel="stylesheet" type="text/css" hre
<link rel="bookmark" href="#technologies"
<link rel="bookmark" href="#news" title="N
<link rel="bookmark" href="#search" title=
<link rel="contents" href="#contents" titl
<link rel="bookmark" href="#Offices" title
<link rel="bookmark" href="#systems" title=
<link rel="bookmark" href="#donors" title=
<link rel="bookmark" href="#footnotes" tit
<link rel="alternate" type="application/rs
</head>

<body>
<h1 id="logo"><img alt="The World Wide Web
<h2 id="slogan">Leading the Web to Its Ful

<div>
<map name="introLinks" id="introLinks" tit
<div class="banner">
<span class="invisible"><a class="bannerLi
Visitors</a> | <a class="bannerLink" title
<a class="bannerLink" title="Contact W3C"

Código fuente de:http://www.w3.org/Overview-about.rdf - Mozilla Firefox
Archivo  Editar  Ver  Ayuda

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:dc = "http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about="http://www.w3.org/">
<dc:subject>World Wide Web Consortium, W3C, World Wide Web, Web,
WWW, Consortium, computer, access, accessibility, semantic,
worldwide, W3, HTML, XML, standard, language, technology, link,
CSS, RDF, XSL, Berners-Lee, Berners, Lee, style sheet, cascading,
schema, XHTML, mobile, SVG, PNG, PICS, DOM, SMIL, MathML, markup,
Amaya, Jigsaw, free, open source, software</dc:subject>
<dc:description>The World Wide Web Consortium (W3C) is about 400
organizations leading the World Wide Web to its full potential.
Founded by Tim Berners-Lee, the Web's inventor. The W3C Web site
hosts specifications, guidelines, software and tools. Public
participation is welcome. W3C supports universal access, the
Semantic Web, trust, interoperability, evolvability,
decentralization, and cooler multimedia.</dc:description>
<dc:date>2003-03-13</dc:date>
<dc:format>text/html</dc:format>
<dc:language>en-US</dc:language>
<dc:creator>W3C Communications Team</dc:creator>
<dc:publisher>W3C - World Wide Web Consortium -
http://www.w3.org/</dc:publisher>
<dc:rights rdf:resource="http://www.w3.org/Consortium/Legal/copyright-documents"/>
<rdfs:seeAlso rdf:resource="http://www.w3.org/2000/08/w3c-synd/home.rss"/>
</rdf:Description>
</rdf:RDF>

Línea 23, Columna 54
```

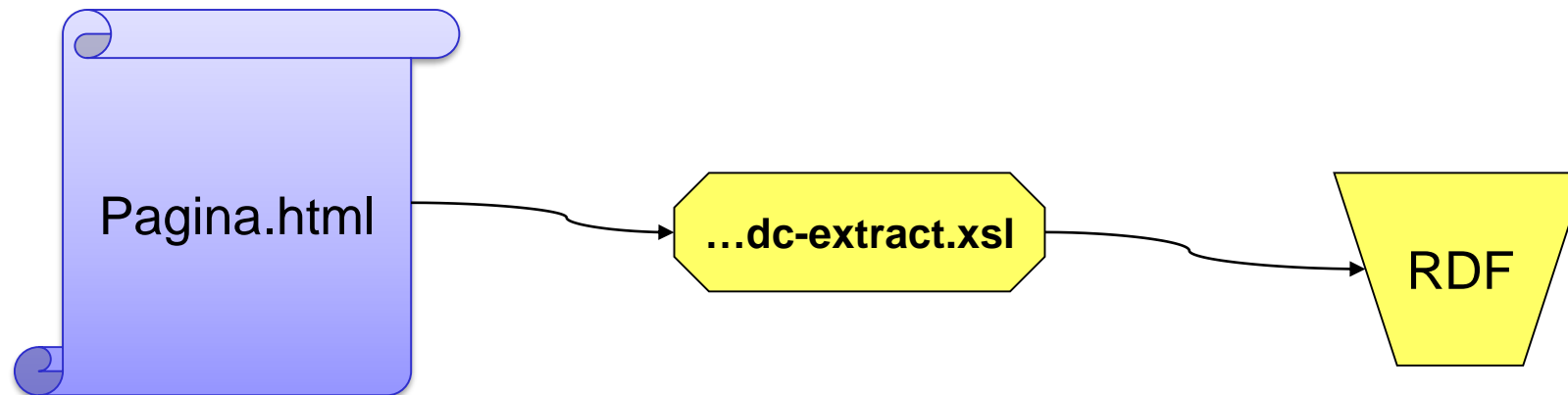
GRDDL

Gleaning Resource Descriptions over Dialects of Languages

Genera RDF a partir de ficheros XML/HTML

Utiliza XSLT para transformar XML/HTML a RDF

Asocia transformaciones XSLT a tipos de documentos



Más información

<https://www.w3.org/wiki/GrddlImplementations>

Microformatos

Usar atributos de HTML para añadir *algo* de semántica

Valores predefinidos en atributo `class`

A veces también: `id`, `title`, `rel` y `rev`

Posibilidades: eventos, contactos, lugares, etc.

Desarrollo mediante wiki: www.microformats.org

Ejemplo:

Oficina situada en coordenadas 23.4, -1.8

```
Oficina situada en coordenadas
<span class="geo">
  <span class="latitude">23.4</span> y
  <span class="longitude">-1.8</span>
</span>
```

Utiliza la especificación geo

Microformatos

Otro ejemplo. Describir información de contacto (hCard)

Sin microformatos:

```
<div>  
  <p>Jose Labra</p>  
  <p>Universidad de Oviedo</p>  
  <p>+34-985103394</p>  
</div>
```

Con microformatos: hCard

```
<div class="vcard">  
  <p class="fn">Jose Labra</p>  
  <p class="org">Universidad de Oviedo</p>  
  <p class="tel">+34-985103394</p>  
</div>
```


Microformatos

Diversas propuestas:

XFN (XHTML Friends Network): relaciones de amistad

hCard: personas y organizaciones (basado en vCard)

Incluye **adr** para direcciones postales y **geo** para lugares geográficos

hCalendar: eventos. Se basa en iCalendar

hAtom: ficheros Atom

hProduct: describir productos

hResume: curriculum vitae de una persona

hReview: revisiones y valoraciones

XOXO: listas

Rel-license: valores de licencias

www.microformats.org

RDFa

RDFa (2004) añade semántica a documentos HTML

Inspirado en microformatos

Se codifican tripletas RDF mediante atributos de HTML

RDFa 1.1 (2011)

Intenta facilitar la creación de documentos

Admite prefijos de espacios de nombres

Permite crear perfiles y vocabularios

Ejemplo RDFa

El libro *La Primavera* de Juan Torre fue publicado el sábado

```
<body vocab="http://schema.org/">
  <div typeof="Book"
    about="http://Leer.com/Libro123">
    <p>El libro <span property="name">La Primavera</span> de
    <span property="author">Juan Torre</span> fue publicado
    <span property="datePublished" content="2014-05-04">
      el sábado
    </span>
  </div>
</body>
```

En Turtle @prefix s: <http://schema.org/> .

```
<http://Leer.com/libro123> a schema:Book;
  s:name "La Primavera";
  s:author "Juan Torre";
  s:datePublished "2014-05-04" .
```

RDFa

Sujeto: se especifica mediante atributo **about**

Predicado: mediante **property**, **rel**, **rev**

Objetos (URIs) mediante **href**, **resource** ó **src**

Tipos: mediante **typeof**

Literales: mediante **content** el contenido propio del elemento.

Atributo opcional **datatype** para tipo de datos

Declarar prefijos de espacios de nombres (**prefix**)

Declarar vocabularios a utilizar (**vocab**)

Herramientas RDFa

RDFa Play: <http://rdfa.info/play/>

Google Rich Snippets Tool:

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

RDFa Developer (Extensión de Firefox)

RDF-Translator: <http://rdf-translator.appspot.com/>

Sindice Inspector: <http://sindice.com/developers/inspector>

StructuredData.org <http://linter.structured-data.org/>

RDFa Live loop <http://rdfa.digitalbazaar.com/live-loop/>

Microdatos



Propuestas para HTML5

Enriquecer documentos con valores de vocabularios

- Grupos de parejas nombre/valor

- Nuevo grupo mediante `itemscope`

- Nombres mediante `itemprop`

- Valores mediante `itemtype`

Genera JSON

Ejemplo



El libro *La Primavera* de Juan Torre fue publicado el sábado

```
<p itemscope
  itemid="http://leer.com/libro123"
  itemtype="http://schema.org/Book">
  El libro
  <span itemprop="name">La Primavera</span> de
  <span itemprop="author">Juan Torre</span> fue publicado el
  <time itemprop="datePublished" content="2014-05-04">sábado</time>.
</p>
```

```
{ "items": [ { "type": [ "http://schema.org/Book" ],
  "id": "http://leer.com/libro123",
  "properties": {
    "author": [ "Juan Torre" ],
    "name": [ "La Primavera" ],
    "datePublished": [ "2014-05-04" ]
  }
} ] }
```

Microdatos



Atributo	Definición	Valores
itemscope	Contenedor de microdatos Crea un nuevo grupo de pares nombre/valor	Valor booleano
itemtype	Define el tipo de los ítems	Lista de URIs
itemid	Identificador del ítem	URI
itemref	Referencia donde se pueden obtener más datos	Identificador
itemprop	Propiedad a definir	Nombre de la propiedad a declarar

Microdatos



Los valores de la propiedad definida en `itemprop` dependen del tipo de elemento:

Elemento	Valor
<code>meta</code>	Valor de <code>content</code>
<code>a</code> , <code>area</code> , <code>link</code>	Valor de <code>href</code>
<code>audio</code> , <code>embed</code> , <code>video</code> , <code>iframe</code> , <code>img</code> , <code>source</code> , <code>track</code>	Valor de <code>src</code>
<code>object</code>	Valor de <code>data</code>
<code>data</code>	Valor de <code>value</code>
Resto de elementos	Contenido textual

PROYECTO SCHEMA.ORG

Proyecto apoyado por: Google, Yahoo, Bing, YandeX, W3c

Sintaxis: Microdatos, RDFa Lite, JSON-LD

Semántica: Vocabulario jerárquico

- Herencia universal de Thing

- Clases con una serie de propiedades

- Las subclases heredan las propiedades de las clases

Información oficial: <http://schema.org>

Más información: <http://www.w3.org/wiki/WebSchemas>

Ejemplo de oferta

```
<div itemscope
  itemtype="http://schema.org/Offer"
  itemid="http://ejemplo.org/Oferta24">
  <h1 itemprop="name">Batidora</h1>
  <p>Precio: <span itemprop="price">19.95€</span>
  <p itemscope itemprop="reviews"
    itemtype="http://schema.org/AggregateRating">
  Puntuación:
  <span itemprop="ratingValue">3</span> sobre
  <span itemprop="bestRating">5</span>.
  Basado en <span itemprop="ratingCount">25</span> usuarios
  </p>
</div>
```

```
@prefix s: <http://schema.org/> .
```

```
<http://ejemplo.org/Oferta24> a s:Offer;
  s:name "Batidora";
  s:price "19.95€";
  s:reviews [ a s:AggregateRating;
    s:bestRating "5";
    s:ratingCount "25";
    s:ratingValue "3" ] .
```

Jerarquía (<http://schema.org/docs/full.html>)

Datatype

Boolean, Number (float,integer), Date, Text (url)

Thing (propiedades: name, description, image, url)

CreativeWork (Book, Movie, MusicRecording, Recipe, TVSeries, ...)

Event

Organization

Person

Place

Product

Review

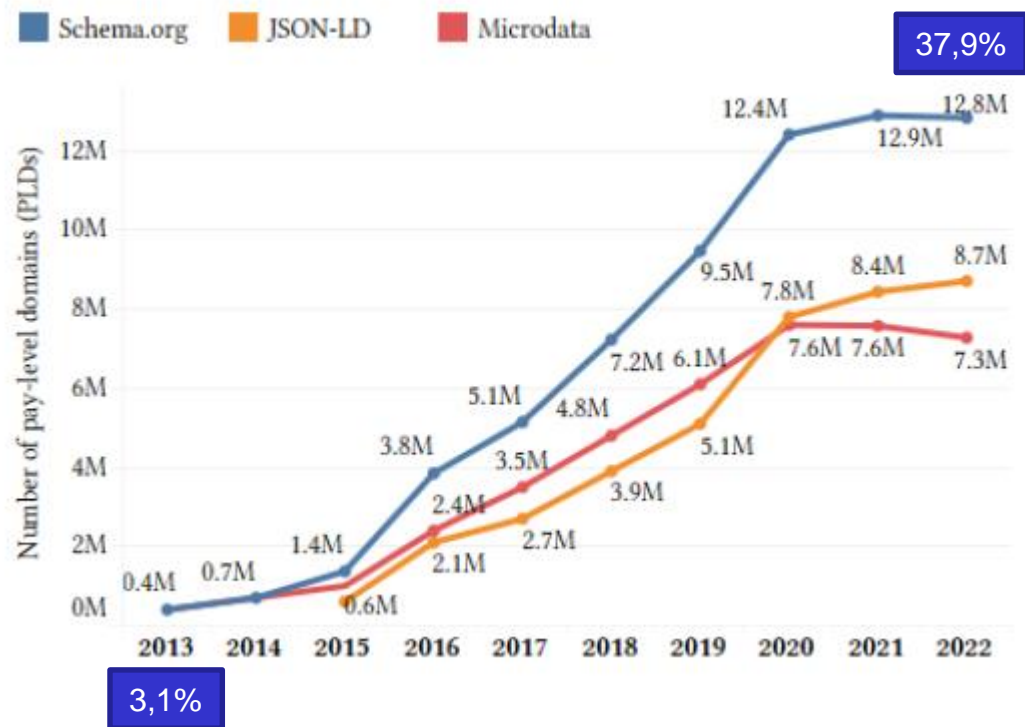
Action

...

Adopción de schema.org

Estudio sobre datos de páginas web (Common Crawl)

<https://webdatacommons.org/structureddata/schemaorg/>



Release	Dominios	Declaraciones	Tamaño
2022-12	12,838,680	105,579,841,305	1,831 GB
2021-12	12,897,837	94,445,547,365	1,691 GB
2020-12	12,414,851	90,416,246,904	1,339 GB
2019-12	9,470,865	25,366,269,937	542 GB
2018-12	7,226,784	20,135,070,334	469 GB
2017-12	5,144,983	14,424,702,204	431 GB
2016-10	3,846,805	26,719,043,004	707 GB
2015-11	1,370,192	12,910,049,866	267 GB
2014-12	700,000	11,067,728,953	262 GB
2013-11	400,000	10,395,924,438	227 GB

Adopción de Schema.org

Noticias: Nytimes, guardian.com, bbc.co.uk,

Películas: imdb, rottentomatoes, movies.com

Trabajos: careerjet.com, monster.com, indeed.com

Personas: linkedin.com,

Productos: ebay.com, alibaba.com, sears.com, cafepress.com, sulit.com, fotolia.com

Vídeos: youtube, dailymotion, frequency.com, vinebox.com

Medicina: cvs.com, drugs.com

Eventos: wherevent.com, meetup.com, zillow.com, eventful

Música: last.fm, myspace.com, soundcloud.com

Otros: yelp.com, allmenus.com, urbanspoon.com

Grupo comunitario: <https://www.w3.org/community/schemaorg/>

Structured data tool (Google)

Herramientas que promueve Google

Structured Data tool:

<https://search.google.com/structured-data/>

Utilización de JSON-LD

JSON-LD permite representar datos RDF

```
<script type="application/ld+json">
{
  "@context": "http://schema.org",
  "@type": "Event",
  "name": "Reunión",
  "startDate" : "2016-0-14T21:30",
  "location" : {
    "@type" : "Place",
    "name" : "Escuela de Ingeniería Informática",
    "address" : "C/Valdés Salas S/N, Oviedo"
  }
}
</script>
```


Otras herramientas

RDF-Translator: <http://rdf-translator.appspot.com/>

Structured data Tool: <http://www.google.com/webmasters/tools/richsnippets>

LiveMicrodata: <http://foolip.org/microdatajs/live/>

StructuredData.org: <http://linter.structured-data.org/>

Any23: <http://any23.org/>

Microdata generator: <http://microdatagenerator.com/>

Otras Aplicaciones

Facebook Open Graph Protocol (RDFa)

Drupal 7: Soporte para RDFa

GoodRelations: <http://www.heppnetz.de/projects/goodrelations/>

Posicionamiento semántico (Semantic SEO)

Originalmente RDFa, ahora admite Microdatos

Adoptado por Overstock, BestBuy, ...

LinkedOpenCommerce

<http://linkedopencommerce.com/>

Google

Data Highlighter, Answer box,

Fin de la Presentación

