

SISTEMAS E SERVIZOS DE INTERNET

Practica 1 - Web Semántica

Jacobo Martínez Gómez

1. CONCEPTOS FUNDAMENTALES

1.1.WEB SEMANTICA. Definición sobre este concepto. Qué implica y cuál es el objetivo. Opinión personal acerca de este término y sus consecuencias.

Es una extensión de la web que permite que el contenido de las paginas lo puedan interpretar las maquinas y algoritmos. Constituye una forma de marcado de metadatos sobre el contenido de las paginas web.

En mi opinión, es algo esencial y de gran utilidad para poder analizar y examinar el contenido de paginas web principalmente cuando se utilizan técnicas como web scraping que se centra en recuperar mediante programas información de un sitio web simulando la navegación de un humano.

1.2.MICROFORMATOS. Definición de microformato. Aplicación y utilidades. Descripción breve de los microformatos disponibles.

Un microformato es una forma sencilla de añadir significado semántico al contenido de una pagina web con la finalidad de que sea comprensible para los humanos sin embargo para una maquina puede ser texto plano.

Hay 4 tipos de microformatos:

- Microformatos Personales: también llamados microformatos de autor, permiten marcar al autor de la pagina utilizando rel="autor" y los enlaces a redes sociales con rel="me".
- Microformatos de Empresa: sirven para descubrir los datos de la empresa, utilizando el microformato hCard, este tipo de marcado tiene gran utilidad para mejorar el posicionamiento SEO de una pagina web.
- Microformatos de Eventos: sirve para empresas que gestionan eventos, usando el microformato hCalendar, vinculando la información necesaria en correspondencia con los eventos.
- Microformatos de Reviews: este marcado se utiliza para ofrecer información sobre el análisis de un producto o servicio.

Microformatos disponibles

- hAtom: para marcar feeds Atom dentro de HTML.
- hCalendar: para eventos
- hCard: para informacion de contacto.
- hReview: para criticas de artículos, servicios, productos.
- hResume: para curriuculums.
- rel-directory: para apuntar enlaces a una página de un directorio.
- rel-nofollow: advierte a los buscadores web que no deben asignar importancia a la página al que apunta el enlace.
- rel-tag: indica que la pagina que contiene el enlace esta relacionada con la etiqueta especificada.
- xFolk: para especificar enlaces guardados en un marcador social.

- XFN: para relaciones sociales mediante enlaces.
- XOXO: para hacer esquemas con listas.

1.3. MICRODATOS. Definición de microdatos. Aplicación y utilidades.

Los microdatos son un conjunto de etiquetas incluidas en html que proporcionan contenido semántico al contenido de la pagina web. Este contenido puede ser interpretado por las maquinas utilizando un script genérico y de forma consistente.

2. CÓMO CONVERTIR/ENRIQUECER UN DESARROLLO BASADO EN HTML5 A NIVEL SEMÁNTICO UTILIZANDO LOS CONCPETOS ANTERIORES

2.1. Adaptar la estructura del desarrollo que se adjunta con la práctica (web_mei.zip) a un etiquetado semántico basado en HTML5. Identificar que secciones y estructuras son susceptibles de este etiquetado y convertir los archivos y c digo necesario para que todo funcione exactamente igual, pero con etiquetado semántico.

```
<body>
  <!-- cabecera -->
  <header id="header">
    <div id="contentHeader">
      <figure id="logo"><a href="index.html"></a></figure>
      <nav id="menu">
        <ul>
          <li><a href="#principal"
class="btnSeccion">PRINCIPAL</a></li>
          <li><a href="#informacion"
class="btnSeccion">INFORMACIÓN</a></li>
          <li><a href="#plan" class="btnSeccion">PLAN DE
ESTUDIOS</a></li>
          <li><a href="#horarios"
class="btnSeccion">HORARIOS</a></li>
          <li><a href="#docentes"
class="btnSeccion">DOCENTES</a></li>
          <li><a href="#faq" class="btnSeccion">FAQ</a></li>
          <li><a href="#contacto"
class="btnSeccion">CONTACTO</a></li>
        </ul>
      </nav>
      <figure id="logoUVigo"><a href="http://www.uvigo.es"
target="_blank"></a></figure>
    </div>
```

```
</header>
<!-- principal -->
<section id="principal">
  <div class="seccion">

  </div>
</section>

<div class="clear"></div>
<!-- informacion -->
<section id="informacion">
  <div class="seccion">
  </div>
</section>
<div class="clear"></div>
<!-- plan de estudios -->
<section id="plan">
  <div class="seccion">
  </div>
</section>
<div class="clear"></div>
<!-- horarios -->
<section id="horarios">
  <div class="seccion">
  </div>
</section>
<div class="clear"></div>
<!-- docentes -->
<section id="docentes">
  <div class="seccion">
  </div>
</section>
<div class="clear"></div>
<!-- faq -->
<section id="faq">
  <div class="seccion">
  </div>
</section>
<div class="clear"></div>
<!-- contacto -->
<footer id="contacto">
  <div class="seccion">
    <div class="infoContacto">
    </div>
  </div>
</footer>
</body>

</html>
```

2.2. Analizar entidades de información de la URL

<http://master.enxeñeriainformatica.es> que no sean estructurales del lenguaje de marcado para poder etiquetarlas semánticamente.

2.3. ¿Qué usarías para la implementación? ¿microformatos o microdatos? ¿Nos llegaría con los tipos del vocabulario shchema.org si usamos microdatos?

Usaría microdatos ya que son los mas utilizados hoy en día. Los microformatos están pensados para facilitar la interpretación del contenido para una persona, sin embargo, para una maquina es simple texto plano. En ese caso es más apropiado utilizar microdatos.

2.4. Realizar extensión de tipos si es necesario.

2.5. Adjuntar el etiquetado para el desarrollo. Crear una lista de las estructuras enriquecidas con metadatos, y mostrar como seria el etiquetado para cada una de ellas (tiempos, duraciones, entidad de información, etc).

- Duracion:

```
<div itemscope itemtype="https://schema.org/Course">
  <h2 class="color-primary mb40 mb-xs-24">Máster Universitario
  en Enxeñaría Informática</h2>
  <p class="lead">
    <span itemprop="description">
      O <span itemprop="name"> Máster Universitario en
      Enxeñaría Informática </span> pola <span itemprop="provider">
      Universidade de Vigo </span> ten como <b>obxectivo primordial</b>
      <span itemprop="educationalCredentialAwarded">
        a preparación de profesionais altamente
        cualificados, capaces de planificar,
        deseñar, dirixir, coordinar e xestionar
        proxectos, produtos, procesos e instalacións en todos os
        ámbitos da Enxeñaría Informática. </span>
      </p>
      <p class="lead">Este Máster ten unha carga docente de <b>
      <span itemprop="numberOfCredits"> 90</span>
      ECTS</b>, impártese durante
      <b>
        <time itemprop="duration"
        itemtype="https://schema.org/Schedule" datetime="R3/PM4">
        tres cuatrimestres </time>
        </b> (un curso e medio), e ten <b>tres
        orientacións</b> ou intensificacións cun <b>elevado nivel de
        empleabilidade</b> na contorna socioeconómica galega e nacional.
```

```

        </span>
    </p>
</div>

```

- Docentes:

```

<div itemscope itemtype="https://schema.org/Person">
    <span itemprop="name">Barreiro Alonso, Enrique</span>
    <p><span itemprop="hasCredential">Licenciado en Ciencias Económicas
e Empresariais</span>, e <span itemprop="hasCredential">Doutor Enxeñeiro
en Informática</span> pola <span itemprop="affiliation"> Universidade de
Vigo</span>. <span itemprop="disambiguatingDescription"> Ten participado
en diversos proxectos con empresas, relacionados coa implantación de
sistemas de información empresarial. Colabora como
<span itemprop="jobTitle">experto técnico na avaliación de proxectos de
I+D+i</span> para a <span itemprop="affiliation">Agencia para la
Certificación de la Innovación Española</span>, e dende 2013 é <span
itemprop="jobTitle">docente internacional</span> da <span
itemprop="affiliation"> Universidade EAN de Bogotá (Colombia)</span>, na
que imparte a materia de Sistemas de Información
Organizacional.</span></p>
    <a href="mailto:enrique@uvigo.es" class="email" itemprop="email">
enrique@uvigo.es</a><br>
    <a href="tel:988368911" class="phone" itemprop="telephone"> 988
368 911</a><br><br>
</div>

```

- Pie de página:

```

<div itemscope itemtype="https://schema.org/Person">
    <p class="lead mb32 mb-xs-16">
        <span
itemprop="occupationLocation"><b>Coordinador</b><br></span>
        <span itemprop="name"> Francisco Javier Rodríguez
Martínez<br></span>
        <span itemprop="jobTitle"> Doctor en Ingeniería
Informática</span> ·n<span itemprop="affiliation"> Universidade de
Vigo<br></span>
        <span itemprop="telephone"> +34 988 387 022</span> - <span
itemprop="email">coordinador.mei.esei@uvigo.es</span>
    </p>
    <p>O Máster conta coa colaboración e apoio do
        <span itemprop="affiliation"><b> CPEIG (Colexio
Profesional de Enxeñaría Informática de
Galicia)</b></span>.
    </p>
</div>

```

2.6. Probar con una herramienta de prueba de datos estructurados.

Detalles

Elementos detectados	
✓ Máster Universitario en Enxeñaría Informática ^	
type	Course
description	O Máster Universitario en Enxeñaría Informática pola Universidade de Vigo ten como obxectivo primordial a preparación d e profesionais altamente cualificados, capaces de planificar, deseñar, dirixir, coordinar e xestionar proxectos, produtos, pr ocesos e instalacións en todos os ámbitos da Enxeñaría Info rmática.
name	Máster Universitario en Enxeñaría Informática
provider	
type	Thing
name	Universidade de Vigo
educationalCredentialAwarded	a preparación de profesionais altamente cualificados, capac es de planificar, deseñar, dirixir, coordinar e xestionar proxect os, produtos, procesos e instalacións en todos os ámbitos d a Enxeñaría Informática.
numberOfCredits	90
duration	R3/PM4

2.7. Microdatos en realidad. Un caso real. Adjuntar un caso real que esté usando microdatos y explica cual crees que es el motivo para ello.

Un ejemplo real de una web que utiliza microdatos es la web <https://www.xataka.com/>. Pienso que el motivo principal por lo que webs como esta utilicen microdatos es debido a la intención de subir el SEO de la pagina dentro de buscadores como Google. Gracias a esto los buscadores son capaces de entender la pagina y descifrar de que trata su contenido y como lo muestra.