

Sistemas de Información para la Web

Práctica 8 - Información semántica: Microdatos

Esteban Montes Morales - UO246305

Ejercicio 1: Cuestiones

¿Deberían tener valor para itemid todos los items?

No tiene por qué, sintácticamente hablando, pero es altamente recomendable, ya que normalmente sirven para identificar de forma única las entidades, y de esta forma podrá obtenerse la información necesaria de ella

¿Qué valor debería asignarse?

Una URI/URN que identifique de forma concisa dicha entidad y que sirva de fuente de información para la misma. Dependiendo del contexto, es muy posible que la decisión del itemid sea crucial.

¿Si hubiera varios candidatos cuál escogerías?¿Por qué?

Tal y como comentaba antes, depende de la situación. Tal vez si el texto que estoy "tagueando" se situa en un contexto bibliográfico, prefiera utilizar identificadores que tengan por ejemplo como fuente la wikipedia, mientras que en caso de estar "tagueando" un texto meramente informativo, como por ejemplo, una guía turística, prefiera utilizar los enlaces a los sitios oficiales, o a páginas relevantes dedicadas al turismo.

¿Cuál crees que es la solución de compromiso que podría resultar menos polémica en la mayoría de los casos?

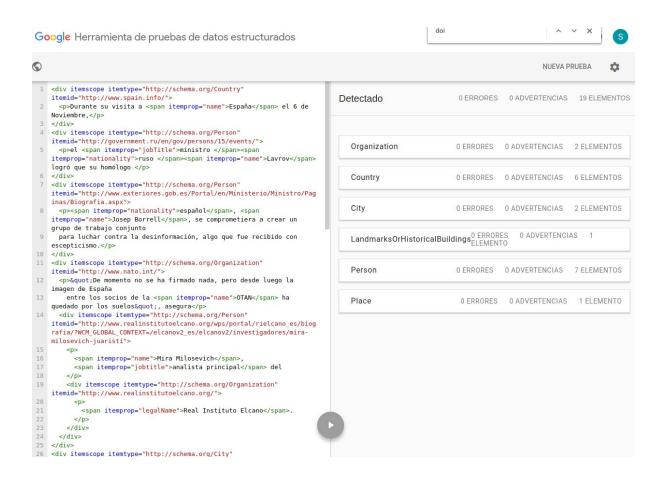
Creo que fuentes como wikipedia serían posiblemente las más correctas genéricamente, debido al contenido variopinto que hay para cada entrada.

¿Qué inconvenientes concretos te ha supuesto la obligación de incrustar los metadatos allí donde te "forzaba" la estructura del texto?

- Falta de expresividad
 - En mi opinión se hace muy verboso y es muy complejo anidar y relacionar información y entidades.
- Ambigüedad a la hora de identificar el tipo de la entidad
 - Aunque esto también es un problema común a RDF o a JSON-LD, y es que es muy importante elegir correctamente el tipo, ya que vas a definir información del texto en base a las propiedades disponibles para ese tipo concreto.
- Duplicado de entidades
 - Una misma entidad puede aparecer repetida en múltiples ocasiones, y tener que definir una entidad con el mismo tipo y el mismo identificador, en lugar de poder reutilizarlo, me parece un punto desfavorable.

- No se puede añadir información adicional no visible (metainformación)
 - Debes ceñirte a la estructura del texto, y no puedes añadir información que no esté presente como tal en el mismo. Esto es una desventaja clara, ya que tal vez al lector humano, no le interese toda esa información adicional no visible, pero una máquina puede utilizarlo para relacionar información y generar nueva.

Ejercicio 1: Herramienta de pruebas de Google



Ejercicio 2: Herramienta de pruebas de Google

