



Vol 4, Nº 11 (diciembre/december 2011)

## **ANÁLISIS DE MÉTODOS DE ETIQUETADO GEOGRÁFICO DE SERVICIOS TURÍSTICOS EN DOCUMENTOS HTML EN INTERNET**

**José Luis García Balboa**

Grupo de Investigación Ingeniería Cartográfica. Universidad de Jaén  
[jlbalboa@ujaen.es](mailto:jlbalboa@ujaen.es)

**Francisco Javier Ariza López**

Grupo de Investigación Ingeniería Cartográfica. Universidad de Jaén  
[fjariza@ujaen.es](mailto:fjariza@ujaen.es)

**Manuel Antonio Ureña Cámara**

Grupo de Investigación Ingeniería Cartográfica. Universidad de Jaén  
[maurena@ujaen.es](mailto:maurena@ujaen.es)

**Alfonso Ureña-López**

Grupo de Investigación Sistemas Inteligentes de Acceso a la Información. Universidad de Jaén  
Campus Las Lagunillas s/n. 23071, Jaén, España  
[laurena@ujaen.es](mailto:laurena@ujaen.es)

### **RESUMEN**

En el sector turístico la posición es una información muy relevante. En los últimos años el etiquetado geográfico o geoetiquetado está cobrando una especial relevancia, sobre todo con la generalización del uso de dispositivos móviles con conexión a internet. En esta misma línea, está en pleno auge el desarrollo y uso de servicios basados en localización o LBS (Location Based Services). Por ello es importante que los documentos html vinculados a una determinada localización, sean geoetiquetados, y así facilitar la labor de los agentes automáticos de búsqueda. Existen varias alternativas para realizar este geoetiquetado las cuales se presentan y analizan para observar ventajas, inconvenientes y poder enjuiciarlas de manera aplicada y realizar una propuesta de mejora. Además, se analiza el grado de difusión del geoetiquetado de páginas html de actividades y negocios del sector turístico.

**Palabras clave:** etiquetado geográfico, geoetiquetado, html, web

## ABSTRACT

In the tourist industry the location is a very relevant data. Geographic tagging, or geotagging, is becoming more important nowadays, especially because of the increased use of mobile devices with wireless internet access and the development and use of Location Based Services. Therefore, it is important that the html documents pointing to a location are geotagged, thus facilitating the task of automatic search agents. The alternatives available are presented in this paper, their main characteristics being summarized in a comparative analysis. Finally, the extent to which geotagging is used in web pages related to the tourist industry is also analysed.

**Keywords:** geographic tagging, geotagging, html, web

## 1. INTRODUCCIÓN

Desde la democratización de la información geográfica y de los sistemas de navegación (p.e. TomTom), es usual que algunas páginas web incluyan información relativa a la posición geográfica de la organización o actividad que presentan (Figura 1). Sin embargo, esta forma de incluir las coordenadas, en el contenido y poco estructurada, ha sido superada por las geoetiquetas.

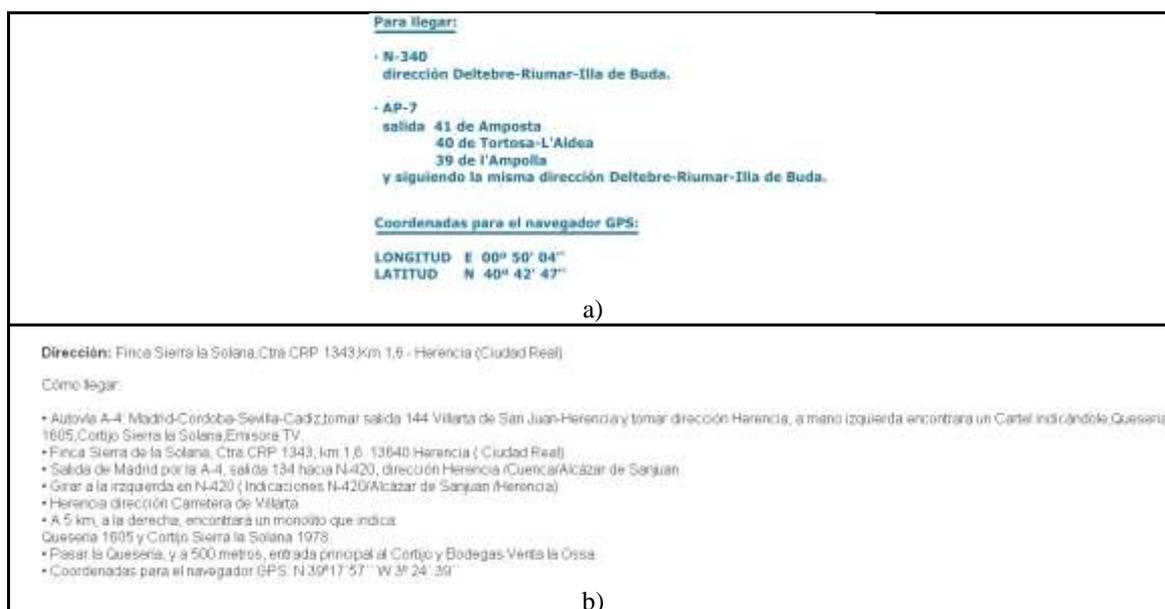


Figura 1. Ejemplo de inclusión de las coordenadas en dos páginas web.

Otra forma de incluir la posición, de una manera más indirecta, pero también más visual, es mediante la inclusión de un mapa, por ejemplo de Google Maps (Figura 2). Esta opción está disponible en dispositivos móviles y permite la visualización por parte del usuario.



Figura 2. Dispositivo móvil con información de un establecimiento hostelero. La posición de indica mediante un mapa de Google Maps (derecha de la pantalla).

Volviendo al etiquetado geográfico, o geoetiquetado, éste es el resultado de incluir información geográfica (etiquetas geográficas o geoetiquetas) junto a alguna información publicada en cualquier formato (fotografías, vídeos, páginas web, etc.). Por tanto estas geoetiquetas son metadatos, cuyo objetivo es enriquecer la información al asociarle una localización espacial. Su empleo está en pleno auge, conforme se está haciendo habitual el uso de dispositivos móviles (smartphones o teléfonos inteligentes) con cámara de fotos digital y receptor GNSS integrados, con conexión a internet, lo que permite tomar una fotografía, etiquetarla con las coordenadas del punto de disparo y subirla a internet. Dado que se pretende que estas geoetiquetas sean utilizadas por el gran público, sin necesidad de disponer de formación específica, suelen incluir información bastante sencilla, siendo el caso más simple el incluir las coordenadas de latitud y longitud.

Actualmente, el principal motivo de geoetiquetar la información es el de permitir el desarrollo de los servicios basados en localización o LBS (Location Based Services). Los LBS son aplicaciones que integran la localización geográfica con el concepto general de servicio y que pueden ser definidos como (Schiller y Voisard, 2004): servicios que integran la localización o posición de un dispositivo móvil con otra información para proporcionar un valor añadido al usuario.

Muchos de los documentos html existentes actualmente en internet contienen información que está vinculada a una localización (una página web de un hotel, por ejemplo). Para facilitar que esta información pueda ser integrada en servicios LBS es importante que estos documentos estén geoetiquetados. Sin estas geoetiquetas, los agentes automáticos de búsqueda no encontrarán la información sobre la posición y por tanto no pasarán a formar parte de las bases de datos que utilizan los servicios LBS.

El presente trabajo pretende realizar una revisión de los diferentes métodos de geoetiquetado que se vienen utilizando para enriquecer los documentos html publicados en la red, resumiendo sus principales características, explicitando sus ventajas e inconvenientes y realizando algunas propuestas de mejora.

## 2. Métodos de etiquetado geográfico en páginas web (html)

A la hora de incluir etiquetas geográficas en cualquier documento html, no existe un estándar al uso que determine unívocamente cómo deben incluirse. Al contrario, el panorama es bastante confuso, al existir diferentes alternativas, y poca información al respecto. En muchas ocasiones la información que fluye en la red se refiere más a las costumbres en el uso de este etiquetado, a través de blogs, foros etc. Estos diferentes métodos, a los que denominaremos *estilos*, se resumen a continuación.

### 2.1. Etiquetados en la cabecera del html: ICBM y GT

El estilo de geoetiquetado *ICBM* es uno de los más sencillos. Su nombre procede de las siglas InterContinental Ballistic Missile (misil balístico intercontinental). Este estilo de geoetiquetado está ligado al concepto de *dirección ICBM* (ICBM address), que se utilizó en el antiguo proyecto *Usenet mapping project* (Raymond, 1996). Usenet es uno de los sistemas más antiguos de comunicación en red, y aún hoy está en uso, ya que es el sistema de discusión distribuida más conocido como grupos de noticias.

El estilo proporciona el par de coordenadas latitud, longitud, en formato decimal y en el sistema WGS84. La información se incluye en la cabecera del archivo html mediante un *metatag* o *etiqueta html*, cuyo nombre es "ICBM" y cuyo contenido es el par de coordenadas latitud, longitud, por este orden. A modo de ejemplo, las coordenadas de la entrada a la catedral de Jaén (España) son latitud igual a 37.76491°N y longitud igual a -3.79049°W), y la etiqueta se introduciría según el siguiente formato:

```
<meta name="ICBM" content="37.76491, -3.79049">
```

Hoy día éste es un método bastante habitual para geoetiquetar cualquier material en la red. Prueba de ello son dos proyectos que permiten poner en marcha servicios LBS a partir de las páginas web recogidas en sus bases de datos: GeoURL (<http://geourl.org>) y A2B (<http://a2b.cc/>). En ambos sitios web se recogen directorios que relacionan una gran cantidad de direcciones URL con localizaciones geográficas. Es decir, si los propietarios de las páginas web dan de alta su página web en A2B, los usuarios pueden realizar búsquedas basadas en la posición, de forma que accedan a las páginas web de lugares o establecimientos próximos a su posición actual.

El estilo de geoetiquetado *GeoTag* (en adelante GT), es también bastante sencillo. Al igual que el estilo ICBM, se basa en la inclusión de información en la cabecera del archivo html, en este caso mediante una o varias etiquetas. Como ejemplo de servicio LBS, A2B también las utiliza.

El estilo GT ha sido propuesto por la IETF (Internet Engineering Task Force), que es una comunidad abierta de la ASOC (Internet Society), que proporciona documentación técnica para el desarrollo de internet. El estilo GT es actualmente un borrador o "internet-draft", cuya última versión es la 8, documentada en Daviel y Kaegi (2007), que en caso de ser aprobado pasaría a ser un RFC (Request For Comments), que constituiría una propuesta formal de estándar. Aunque aún se trate de un borrador, este estilo se ha venido utilizando frecuentemente. La primera versión del borrador es de 1999, lo que da idea del tiempo transcurrido desde que se planteó su uso por primera vez.

Como se indica en Daviel y Kaegi (2007), este geoetiquetado pretende ser una forma concisa, no ambigua, simple de usar y compatible con herramientas de edición existentes, para

proporcionar información sobre localización a los robots web que revisitan las páginas cada varias semanas. Los tipos de etiquetas son los siguientes<sup>1</sup>:

- Posición (`geo.position`). Se utiliza para proporcionar la latitud y longitud de una localización en el sistema WGS84, y opcionalmente la elevación, por este orden.
- País (`geo.country`). Se corresponde con un identificador del país, según la norma ISO 3166-1 (ISO, 2006). En el caso de España el identificador es “ES”.
- Dirección (`geo.a1`, `geo.a2`, etc.). Para la etiqueta `geo.a1` se propone utilizar una lista controlada de subregiones, como la proporcionada por la norma 3166-2 (ISO, 2007). Por ejemplo, la etiqueta `geo.a1` para el caso de la provincia de Jaén sería “J”. El resto de etiquetas serían libres, pero su uso sería más descriptivo que de indexación, por la posible ambigüedad en su utilización, dadas las diferentes convenciones a la hora de nombrar un lugar. En resumen, puede considerarse como una estructura jerárquica para ir presentando la dirección postal de una ubicación.
- Hito (`geo.lmk`). Se trata de una etiqueta que se propone para incluir el nombre común del objeto que se está describiendo.

En cuanto al uso de estas etiquetas, en principio, como mínimo hay que utilizar la etiqueta de posición (`geo.position` con la latitud y longitud), lo que garantiza una indexación con exactitud de un punto que representa la posición de lo que se está documentando. No obstante, esto tendría poco sentido para se está haciendo referencia a un país o región. En ese caso se puede omitir la etiqueta de posición, y en su lugar se incluirían las etiquetas de país (`geo.country`) y dirección (`geo.a1`, etc.)

Siguiendo con el ejemplo de la catedral de Jaén, las etiquetas a incluir serían las siguientes:

```
<meta name="geo.position" content="37.76491; -3.79049">
<meta name="geo.country" content="ES">
<meta name="geo.a1" content="J">
<meta name="geo.a2" content="Jaén">
<meta name="geo.lmk" content="Catedral de Jaén">
```

## 2.2. Estilo de microformato Geo (MFG)

Este estilo se basa en la utilización de los conocidos como microformatos, término que fue acuñado en 2004. Los microformatos son un conjunto de formatos de datos simples y abiertos, elaborados a partir de estándares existentes y ampliamente adoptados (Microformats, 2011a). La idea es la de reutilizar lo que ya está en uso, de forma que mediante pequeñas adaptaciones, se puedan resolver problemas específicos. La información que contiene un microformato se añade mediante dos elementos del código html el primero de ellos cualquier elemento html o xhtml que soporte la creación de una clase (p.e. *li*, *p*, *div*, etc.) junto con la clase *geo* y el segundo siempre debe ser un elemento *span* (o *abbr*, como se verá más adelante) más el atributo *class* que identifica el atributo de latitud y longitud. Así se puede crear una estructura para el documento siguiendo un mecanismo genérico (WC3, 1999).

Todos los microformatos (aprobados y en borrador) se encuentran en Microformats (2011a). Uno de los microformatos, que actualmente se clasifica como borrador, aunque está en uso, es el *Geo* (Microformats, 2011b) (en adelante se denominará MFG). Este microformato permite etiquetar la latitud y longitud en grados decimales y en el sistema WGS84 (actualmente está

---

<sup>1</sup> Con anterioridad a la versión 8 del borrador del estilo GT, las etiquetas definidas eran: *geo.position*, *geo.placename* y *geo.region*. La primera ha permanecido, la segunda ha sido sustituida por la etiqueta *geo.lmk* y la tercera ha sido sustituida por las etiquetas *geo.country* y *geo.a1*. Aún puede encontrarse un gran número de documentos html que usan esta estructura más antigua.

bajo estudio la inclusión de una extensión del microformato para incluir información sobre elevación). A diferencia de los estilos ICBM y GT, en este caso se pueden incluir en un archivo html tantas localizaciones como se desee, ya que el microformato no se incluye en la cabecera, sino que forma parte del cuerpo del documento.

Siguiendo con el ejemplo de la catedral de Jaén, la información a incluir sería la siguiente:

```
<div class="geo">
  <span class="latitude">37.76491</span>;
  <span class="longitude">-3.79049</span>
</div>
```

En ejemplo anterior, la información que lee el usuario y la que lee cualquier agente automático de búsqueda es la misma. Si se desea presentar la información de latitud y longitud en otro formato, para que sea más entendible por el ser humano (por ejemplo, como grados, minutos y segundos), existe un patrón de diseño que permite hacerlo. En concreto, el patrón de diseño *abbr-design-pattern* (<http://microformats.org/wiki/abbr-design-pattern>) permite utilizar el elemento *abbr* en lugar de *span*, y añadir el atributo *title*, que es el que contiene la información estandarizada y compatible con el microformato. Así, siguiendo con el ejemplo que se está manejando se podría incluir la siguiente información:

```
<div class="geo">
  <abbr class="latitude" title="37.76491">N 37° 45' 53.7''</abbr>
  <abbr class="longitude" title="-3.79049">W 3° 47' 25.7''</abbr>
</div>
```

Es muy importante aclarar que en caso de desear incluir el nombre de la localización, además de la latitud y longitud, debe utilizarse el microformato *hCard* (Microformats, 2011c), mucho más rico al tener a disposición del usuario muchos más atributos de tipo *class*, entre ellos los de atributos *geo*. En realidad el MFG es una representación del atributo *geo* del microformato *hCard*. Lo mismo ocurrirá si se desea incluir la dirección postal, en cuyo caso debe utilizarse el microformato *hCard* con el atributo *adr*. La Universidad de Jaén podría incluirse de la siguiente forma si se quiere incluir el nombre y la dirección postal de la misma:

```
<div class="vcard">
  <span class="fn org">Catedral de Jaén</span>
  <span class="adr">
    <span class="street-address">Plaza de Santa María</span>
    <span class="locality">Jaén</span>,
    <span class="region">J</span>
    <span class="postal-code">23001</span>
    <span class="country-name">Spain</div>
  </span>
  <span class="geo">
    <abbr class="latitude" title="37.76491">N 37° 45' 53.7''</abbr>
    <abbr class="longitude" title="-3.79049">W 3° 47' 25.7''</abbr>
  </span>
</div>
```

### 2.3. Estilo RDF Geo (RDFaG)

El RDF (Resource Description Framework, Marco de Descripción de Recursos) es un modelo de datos que fue creado por el W3C (World Wide Web Consortium), dentro del concepto de Web semántica. El objetivo de este modelo es añadir información adicional a las páginas web, en forma de metadatos y de una forma determinada, para que pueda ser procesada de forma automática por cualquier agente automático de búsqueda de información en la red.

El RDF Interest Group, que a partir de 2004 pasó a denominarse Semantic Web Interest Group (SWIG), seleccionó varias áreas temáticas (ontologías) cuyos conceptos son de amplia difusión, en el sentido de que podrían ser utilizados por una gran cantidad de aplicaciones. Entre estas ontologías seleccionaron la referente a la información geoespacial. No obstante hay que aclarar que no forma parte de las estandarizaciones elaboradas por el W3C. La información geoespacial es representada mediante un vocabulario básico, denominado *SWIG Basic Geo Vocabulary* (W3C, 2006). Este vocabulario básico permite describir puntos con la latitud, longitud y altitud.

Para introducir los metadatos ajustados al modelo RDF en una página web se puede utilizar RDFa, que es una sintaxis para incorporar la información RDF en un documento html a través de atributos (véase W3C, 2008). En el caso del vocabulario básico Geo, los atributos de RDFa que se han de utilizar para la información geográfica son los siguientes:

- `typeof`. Para indicar de qué tipo es la información descrita. Aquí se indicaría que se trata de una información de tipo *geo:Point*, que quiere decir que se va a describir la localización de un punto.
- `property`. Se refiere a la propiedad (o atributo) que se va a describir. Serían tres las propiedades que se pueden incluir: *geo:lat* para la latitud, *geo:long* para la longitud y *geo:alt* para la altitud (ésta es opcional)
- `content`. Se incluye a continuación de la propiedad y contiene el valor del atributo, que sería el valor de latitud, longitud o altitud, en cada caso. Las dos primeras se introducen en grados decimales sobre WGS84 y la tercera en metros sobre el elipsoide de referencia local (así se indica en la documentación).

A la utilización de RDFa para representar información ajustada al vocabulario básico Geo en RDF se le denominará en adelante RDFaG. Siguiendo con el ejemplo de la catedral de Jaén, la información a incluir sería la siguiente:

```
<div typeof="geo:Point">
  <span property="geo:lat" content="37.76491"> N 37° 45' 53.7''</span>
  <span property="geo:long" content="-3.79049"> W 3° 47' 25.7''</span>
</div>
```

Como se puede comprobar, la información que lee el usuario y la contenida en los metadatos están en formatos diferentes. La primera estaría en formato sexagesimal y la segunda en formato decimal, que es el que se ajusta al vocabulario básico Geo y es capaz de ser entendida por cualquier agente automático de búsqueda de información. Es lo mismo que se ha indicado en el estilo MFG en cuanto al patrón *abbr-design-pattern*.

Con posterioridad a las iniciativas indicada en el subapartado anterior se crearon en el W3C los denominados Incubator Groups (XGs) con la idea de acelerar el desarrollo de nuevas propuestas relacionadas con la web, ya que el desarrollo de estándares suele ser lento. De esta forma se consigue obtener un informe (XG Report) en poco tiempo, que con posterioridad puede incluirse en el proceso de estandarización habitual. Uno de estos XGs fue el W3C Geospatial Incubator Group (GeoXG, ya desaparecido). Uno de los primeros informes del GeoXG fue la elaboración del *W3C Geospatial Vocabulary* (W3C, 2007), que puede considerarse como una propuesta de actualización del vocabulario básico. El GeoXG planteó inicialmente varios ejemplos prácticos para mostrar la necesidad de fortalecer la web en la componente geoespacial y decidió adoptar el modelo de datos GeoRSS (<http://www.georss.org/gml>) de geosindicación de contenidos, que permite describir no sólo puntos, sino también rectángulos, líneas y polígonos. Se trata de un modelo ligeramente diferente y reducido respecto al modelo de entidades simples de OGC (<http://www.opengeospatial.org/standards/sfs>). Será, por tanto, interesante, comprobar si en el futuro esta propuesta se convierte en un estándar de W3C.

### 3. Análisis comparativo de estilos

En el apartado 2 se han resumido cuáles son las principales formas de incluir información geográfica en una página web: etiquetado en la cabecera (ICBM y GT), MFG y RDFaG. Hasta la fecha, no se ha declarado ninguna de las anteriores propuestas como estándar que haga decantarse por uno u otro. Incluso en muchas ocasiones lo que se realiza es incluir la información geográfica en todos estos *estilos*, de forma que se maximice la posibilidad de que los agentes automáticos de búsqueda encuentren la información relativa al geoetiquetado. Realmente es algo que no requiere un esfuerzo especial. No obstante es conveniente realizar un análisis comparativo que permita vislumbrar las ventajas, inconvenientes, semejanzas y diferencias que hay entre ellos.

#### 3.1. Resumen de principales características

Para facilitar el repaso a las principales características de los estilos de geoetiquetado descritos en el apartado 2, en la tabla 1 se presenta un resumen. Estas características se irán analizando en los siguientes apartados

**Tabla 1. Principales características de los estilos de geoetiquetado analizados**

Característica	ICBM	GT	MFG	RDFaG
Organismo que lo formaliza	ninguno	IETF	Microformats	SWIG, W3C
Documentación sobre el estilo	No	sí	sí	Sí
Posición en el documento html	cabecera	cabecera	cuerpo	Cuerpo
Tipo de localización	puntual	puntual	puntual	puntual <sup>3</sup>
Datos obligatorios	latitud, longitud	latitud, longitud <sup>1</sup>	latitud, longitud	latitud, longitud
Datos opcionales	-	altitud, país, región, dirección postal, nombre común	- <sup>2</sup>	altitud
Formato de las coordenadas geográficas	grados decimales	grados decimales	grados decimales	grados decimales
CRS horizontal	WGS84	WGS84	WGS84	WGS84
CRS vertical	-	WGS84	-	“elipsoide de referencia local”
Otros CRS	No	no	no	no
Incertidumbre	No	número de decimales	no	no
Posicionamiento indirecto	No	sí	no	no
Recomendaciones de uso	No	algunas directrices	no	no

<sup>1</sup> Si las coordenadas hacen referencia a un país o región, el estilo GT permite obviar las coordenadas latitud, longitud e incluir las etiqueta referente al país y región.  
<sup>2</sup> El estilo MFG es una parte del microformato *h-card*, que incluye muchos más datos (por ejemplo las direcciones a través de la clase *adr*).  
<sup>3</sup> Aquí se ha considerado el vocabulario básico Geo, no la propuesta del GeoXG basada en GeoRSS.

Fte. Elaboración propia

### 3.2. Existencia de especificaciones y herramientas.

Los estilos GT, MFG y RDFaG sí están establecidos formalmente, respectivamente por IETF (aunque actualmente se trata de un borrador o “internet-draft”), *microformats community* y por W3C. Esto conlleva la existencia de páginas web que describen estos estilos, indicando las sucesivas revisiones realizadas sobre el estilo, señalando si existe alguna discusión abierta, problemas pendientes de resolver, etc.

Por el contrario, el geoetiquetado ICBM ha venido utilizándose conforme se difundía su uso por la red, pero sin que ninguna organización lo describiera formalmente. En cualquier caso, se trata del estilo de geoetiquetado más sencillo con lo cual tampoco supone un problema grave. Eso también permite suponer que no va a evolucionar para introducir mejoras.

Sería cómodo que los propios editores html facilitarán la inclusión del geoetiquetado, y que tuvieran en cuenta los documentos que definen los distintos estilos. No obstante, dado que las geoetiquetas son estructuras de código bastante simples, se introducen editando el código html con las herramientas habituales que permiten los distintos editores.

No obstante, en la red existen algunas herramientas sencillas que facilitan la incorporación de los estilos de geoetiquetado, principalmente el ICBM y GT, como es el caso de <http://www.geotag.de/generator/en.html>. Esta herramienta facilita el fragmento de código html a insertar en la cabecera del documento, solicitando al usuario la dirección postal de interés. Otras herramientas son: <http://www.mygeoposition.com/>, <http://www.willamowius.de/geo-tags.html>, <http://www.addressfix.com/>. También existen herramientas de validación (p.e. <http://www.geotag.de/validator/en.php>).

### 3.3. Información contenida

La mayoría de los estilos contemplan únicamente el etiquetado de localizaciones de carácter puntual. Por tanto, cualquier fenómeno del mundo real, independientemente de tu tamaño o extensión, sólo puede ser georreferenciado mediante un punto. Por el contrario, las opciones basadas en el vocabulario del GeoXG, que podría usarse en el RDFaG, admitiría líneas, superficies y rectángulos.

Tres estilos, ICBM, MFG y RDFaG sólo proporcionan información sobre las coordenadas geográficas del punto o puntos de interés (aunque hay que recordar que el MFG es una parte del microformato *h-card*, que permite añadir mucha más información). Por el contrario, el estilo GT es más rico, ya que permite incluir, además, información sobre el país, región, dirección postal y nombre común del objeto. Incluso en este caso la información sobre el país y región está estandarizada. Esto puede facilitar la recuperación de información no basada exclusivamente en la coordenada o realizar búsquedas basadas en la dirección postal. En otras palabras, el estilo GT permite tanto el posicionamiento por coordenadas como el posicionamiento denominado “indirecto” o por identificadores geográficos, es decir, basado en una relación con una localización dada por uno o varios fenómenos geográficos (el país, la región, etc.). No obstante, este posicionamiento indirecto no sigue formalmente ningún sistema de referencia espacial por identificadores geográficos, en conformidad con ISO 19112 (ISO, 2003). En este sentido, sería muy positivo fijar un sistema de referencia global en conformidad con esta norma.

Merece un comentario aparte la información sobre la coordenada de altitud, ya que dos estilos, ICBM y MFG, no la consideran, aunque sobre éste último se está considerando introducir una extensión del microformato para poder incluir esta información.

### **3.4. Ubicación de la información en el archivo html**

Dos estilos consideran la inclusión de la información en la cabecera del archivo html: el ICBM y GT. Por tanto son útiles para páginas web cuyo contenido aluda a una única localización. Es decir, si un documento contiene información sobre varias ubicaciones, sería más adecuado dividirlo en varios html, de forma que se incluyera la información sobre la posición en cada una de las cabeceras de los nuevos documentos (por ejemplo, una página por cada restaurante que se quiere georreferenciar). Al ubicarse en las cabeceras, esta información no es directamente accesible por el usuario que lee la página, aunque puede ser encontrada sin problema por agentes de búsqueda. Si se quisiera que esta información sobre la localización geográfica fuera explícita para el usuario, debería también incluirse como texto simple en el cuerpo del documento html.

Dos estilos consideran incluir la información sobre el geoetiquetado en el propio cuerpo del documento, el MFG y el RDFaG. Es decir, permiten ir incluyendo atributos al texto conforme se va redactando el documento html, de forma que la información sobre la localización que se pone de forma explícita a disposición del usuario (por ejemplo, para que pueda introducirla en su navegador GNSS) queda enriquecida y puede ser interpretada por un agente automático. De este modo en una misma página web se pueden incluir geoetiquetas sobre diferentes ubicaciones (por ejemplo, todos los restaurantes de una ciudad en el mismo documento html). Sin embargo, en el estilo MFG se tiene como inconveniente la necesidad de asignar la información geográfica a través de clases, por lo que si es necesario incluir formatos html a los elementos que contienen este tipo de información, deben declararse dentro de las estructuras CSS pertinentes o asignar los parámetros de formato individualmente a cada campo. Por el contrario, el RDFaG no tiene este inconveniente al no estar limitado a elementos class y su desarrollo dentro del estándar RDF admite su inclusión en bases de datos XML o en metadatos sobre información como el Dublín Core.

Acerca de la redacción explícita de información geográfica en el cuerpo del documento, es interesante consultar el Anexo D de la norma ISO 6709 (ISO, 2008), en la que se dan una serie de sugerencias para representar la latitud y la longitud de cara al usuario.

### **3.5. Formato de las coordenadas geográficas y sistema de referencia.**

Todos los estilos requieren introducir la información sobre la latitud y la longitud en un mismo formato, que es el de grados decimales. No obstante, el usuario puede estar habituado a otros formatos, como el de grados en formato sexagesimal o coordenadas UTM. En el caso de los estilos ICBM y MFG no se trata de un problema, ya que la información reside en la cabecera del archivo html, la cual no es directamente accesible por el usuario. En el caso de los estilos MFG y RDFaG, dado que trabajan mediante atributos de la información vertida en el cuerpo del documento, supondría una limitación a la hora de presentar las coordenadas geográficas al usuario. En ambos casos esto se ha solventado de forma que se pueda presentar al usuario una información explícita, en un formato totalmente libre, aunque es recomendable seguir la propuesta del Anexo D de ISO 6709 (ISO 2008), y un valor para el atributo en el formato requerido, grados sexagesimales.

En cuanto al sistema de referencia, se propone siempre la utilización de WGS84 en lugar de otros sistemas de referencia más antiguos y locales (como ED50 en España o NAD27 en EE.UU.) es lógica. Junto con la utilización del formato de coordenadas en grados decimales, facilita la estandarización y el intercambio de información.

No obstante es importante que el usuario sepa qué supone tomar información georreferenciada en antiguos sistemas locales. Es habitual que un usuario normal, sin formación específica, considere que la coordenada de latitud y de longitud es única, desconociendo que existen

diferentes sistemas de referencia. Por ello es conveniente que siempre que se documente algún tipo de estilo (unas especificaciones, un manual, una guía de uso, etc.) se remarque la importancia de este aspecto.

Respecto al sistema de referencia empleado en altitud, para el estilo GT se solicita que se utilice el sistema WGS84. Aunque no es común en cartografía el utilizar altitudes elipsoidales, la difusión de la utilización de navegadores GNSS, y la necesidad de utilizar un sistema de referencia único, favorece su uso. No obstante, es importante informar al usuario sobre qué supone el uso de la altitud sobre WGS84. En este sentido, en Daviel y Kaegi (2007) se advierte de la diferencia que puede provocar respecto a un sistema de referencia altimétrico local (por ejemplo, referido al nivel medio mal mar y utilizando altitudes ortométricas). En relación al RDFaG, este asunto no está tan claro; la documentación indica que la altitud se refiere “sobre el elipsoide de referencia local”. Esto choca con la idea de utilizar el sistema WGS84, que utiliza el elipsoide global GRS80. Podría pensarse que en realidad quería referirse al sistema de referencia altimétrico local (por ejemplo, el sistema de altitudes ortométricas sobre el geoido, utilizado en España). Esto último provocaría problemas a la hora de que los usuarios utilicen navegadores GNSS que suelen ofrecer la altitud elipsoidal sobre GRS80, sin aplicar ningún modelo de geoido para obtener altitudes ortométricas.

En cualquier caso, la concreción del sistema de referencia considerado sólo es importante cuando las coordenadas proporcionadas tienen una incertidumbre de carácter métrico (como la que puede ofrecer un navegador GNSS) y se refieren a objetos del mundo real de poca extensión (una pequeño edificio por ejemplo). A modo de ejemplo, tiene poca o ninguna importancia conocer el sistema de referencia si se van a proporcionar las coordenadas de una ciudad; igual sucedería si se quieren dar las coordenadas de un edificio de poca extensión pero con una incertidumbre de varias centenas de metros. Esto es así porque el cambio en el valor de las coordenadas que supone cambiar de sistema de referencia es inferior a la incertidumbre de las coordenadas y/o a la extensión del objeto del mundo real que se está geoetiquetando. En este sentido, en Daviel y Kaegi (2007), se indica que siempre que la incertidumbre sea inferior a 1 km, las coordenadas deben convertirse al sistema WGS84. Esto es consistente con la norma ISO 6709 (ISO, 2008), que indica que hay que especificar el CRS (Coordinate Reference System, Sistema de Referencia de Coordenadas) para aplicaciones que requieren una exactitud mayor a 1 km.

Es interesante mencionar la compatibilidad de los estilos de geoetiquetado con la Norma ISO 6709 (ISO, 2008), cuyo objetivo es la normalización de la representación de localizaciones geográficas puntuales mediante coordenadas. Esta norma establece que una localización siempre ha de constar de una tupla de coordenadas y una identificación del CRS. Será el CRS el que especifique las direcciones positivas de cada eje de coordenadas, el orden de las coordenadas y las unidades. En caso de no indicar ningún CRS, se considera que la posición del punto adquiere un mayor grado de incertidumbre y que la tupla consta de la latitud, antecediendo a la longitud, expresándolas en grados decimales, e indicando a continuación la altitud (o profundidad) si se trata de una localización tridimensional. El respetar este orden se considera crítico para el caso de emergencias en marina y navegación, ya que tradicionalmente éste ha sido el orden de estas coordenadas; utilizar otro orden podría generar situaciones de riesgo.

Todos los estilos considerados siguen este orden. Esto es importante para los estilos ICBM y GT, que presentan las coordenadas en una única etiqueta. Por el contrario, en los estilos MFG y RDFaG la latitud y longitud se presentan en etiquetas independientes, y por ello el orden no tiene importancia para agentes automáticos de búsqueda. No obstante, dado que estos últimos estilos pueden ir ligados a la presentación al usuario de la información geográfica de forma explícita en el cuerpo del documento, se recomienda seguir la regla general de anticipar la latitud a la longitud.

En cuanto al CRS, ya se ha comentado que en todos los estilos se ha convenido utilizar WGS84, aunque no se explicita en ninguna etiqueta. Esto es consistente con el RFC 5870 (Mayrhofer y Spanring, 2010), en relación a la estructura del URI (Uniform Resource Identifier, o Identificador Uniforme de Recurso) para localizaciones geográficas, denominado ‘geo’ URI. Este URI permite incluir el parámetro opcional “crs”, pero si no se incluye, por defecto se ha de entender que el CRS es el WGS84. Si en el futuro se viera conveniente abrir los estilos de geoetiquetado a otros CRS, sería conveniente incluir una etiqueta a tal efecto, de forma similar al parámetro “crs” del RFC 5870, antes mencionado. En este sentido, y en la línea de los adoptado en las Infraestructuras de Datos Espaciales, parece lógico que se utilizara el sistema de códigos EPSG (European Petroleum Survey Group) que barre todas las opciones posibles a nivel mundial, tanto en datums, elipsoides como sistemas de proyección. Por ejemplo, el sistema WGS84 se codificaría como EPSG 4326 (sólo componente horizontal: latitud y longitud) o 4979 (componentes horizontal y vertical: latitud, longitud y altitud elipsoidal).

### 3.6. Incertidumbre de las coordenadas geográficas

Todo conjunto de coordenadas geográficas (latitud, longitud, altitud) que represente la posición de un punto relativo a la ubicación de un fenómeno del mundo real, lleva asociado un *error*, que viene a ser la diferencia entre estas coordenadas y las coordenadas verdaderas o de referencia. Este error es teóricamente desconocido, al no poder conocer las coordenadas verdaderas.

Dado lo anterior, se recurre a la denominada *incertidumbre de medida*, que sí puede ser evaluada. La incertidumbre de medida caracteriza la dispersión de los valores, y viene dada por un parámetro, habitualmente la desviación típica (pasando a denominarse “incertidumbre típica de medida”). Existen otros términos relacionados, como *exactitud de medida* o *precisión de medida*, que presentan matices diferentes y por tanto no deben confundirse entre sí (para profundizar, véase Ruiz y otros, 2010).

En el caso del geoetiquetado, ningún estilo considera necesario el incluir información sobre la incertidumbre de las coordenadas proporcionadas. Por tanto el usuario no tiene a su disposición ningún parámetro, cuantitativo o cualitativo, sobre la calidad de estas coordenadas y en consecuencia las limitaciones de uso que conlleva. En realidad esto sólo supone un problema cuando se desea proporcionar la localización de un objeto del mundo real de poco tamaño o extensión (un pequeño edificio por ejemplo). No obstante, se puede considerar que esto último es actualmente es muy frecuente, dado que la difusión del uso de navegadores GNSS conlleva que cada vez sea más frecuente geoetiquetar objetos de todo tipo para facilitar que distintos usuarios utilicen su navegador para encontrar el objeto en cuestión (por ejemplo, que un cliente encuentre un restaurante). Esto conlleva que las coordenadas geográficas proporcionadas sean de una incertidumbre reducida, del orden de varios metros, y aquí sí sería conveniente indicar que se cumple esta premisa. También conlleva, como se ha dicho anteriormente, que esté claro el sistema de referencia en el que se ofrecen las coordenadas (en principio WGS84).

Parece interesante la propuesta de Daviel y Kaegi (2007) en relación al estilo GT, donde se sugiere que si se está geoetiquetando un país o una parte de él (como un estado o una provincia), no se indiquen las coordenadas geográficas, prescindiendo de la etiqueta *geo.position*, y sólo se cumplimenten las etiquetas *geo.country* y *geo.a1*.

Por otro lado, también en Daviel y Kaegi (2007) se sugiere cuidar la localización del punto que va a representar la ubicación de lo que se pretende geoetiquetar, cuando se está tratando con pequeños objetos del mundo real y una incertidumbre reducida (del orden de pocos metros). En este caso puede ser más interesante capturar las coordenadas de la puerta de entrada al objeto que el centro del objeto, como puede ser el caso de un edificio. Si se captura el centro del edificio en lugar de la puerta, aunque el punto capturado tuviera una incertidumbre reducida, el usuario no tendría la certeza de por dónde se entra al edificio. Esto último conlleva, en cierto sentido, un aumento de la incertidumbre de la información que se quiere proporcionar al

usuario, que ante todo debe ser útil. Evidentemente, toda regla del geoetiquetado, como la anterior respecto a la entrada a edificios, debería quedar reflejada en unas especificaciones del estilo de geoetiquetado, que actualmente se centran en el formato de las geoetiquetas, pero comentan poco o nada sobre su uso.

En relación a la referida ausencia de información expresa sobre la incertidumbre, es interesante observar la propuesta de Daviel, Kaegi y Kofahl (2007), que no se refiere al geoetiquetado, sino a la inclusión de una extensión de tipo geográfico para las transacciones http (es decir, las peticiones que se realizan desde un navegador web a un servidor para que le devuelva el contenido de una página web). En esta propuesta se sugiere un identificador denominado *geo-position*, que incluye la posibilidad de informar sobre la incertidumbre de la posición mediante la clave *epu* (estimated position uncertainty). Se propone que esta incertidumbre se corresponda con el radio de un círculo (o esfera) con una probabilidad del 95%. Aunque no se indica expresamente, esto se corresponde con la medida CE95 definida en la norma ISO 19138 (ISO, 2006). También hay que citar en este sentido el RFC 5870 (Mayrhofer y Spanring, 2010), que permite incorporar un parámetro “u” opcional sobre la incertidumbre. Se trata de un parámetro único, ya se trate de una localización bidimensional o tridimensional. Podría ser interesante incorporar algo similar en los estilos de geoetiquetado.

Existe la posibilidad de relacionar el número de cifras decimales de las coordenadas proporcionadas con su incertidumbre, tal y como se sugiere en Daviel y Kaegi (2007) y en ISO (2008). Una menor incertidumbre, se correspondería con un mayor número de cifras decimales. A modo de ejemplo, si se han capturado con un navegador GNSS las coordenadas de la entrada de un edificio, éstas deberían darse con 4 decimales (asumiendo, que 1 grado son unos 111 km, 1 diezmilésima de grado supone unos 10 metros) y en WGS84. En cambio si sólo se sabe que el edificio está en una gran ciudad, se podrían tomar las coordenadas del centro de la ciudad a partir de cualquier fuente cartográfica y ofrecer éstas coordenadas para el edificio, pero sólo con dos decimales (la centésima de grado equivaldría a 1 km aproximadamente) y por supuesto sin necesidad de que sean WGS84.

Obsérvese, por otro lado, que el RFC 5870 indica que el número de cifras decimales no debe ser asociado a la incertidumbre, recomendando el anteriormente referido parámetro “u” para informar sobre esta última. Esto saca a la luz que el tratamiento de la incertidumbre es un asunto pendiente en la búsqueda de la interoperabilidad de la información geográfica en internet. La introducción de un parámetro específico para la incertidumbre sería lo más oportuno.

#### **4. Análisis de la situación actual en España del uso de geoetiquetas en el sector turístico**

Como se ha ido indicando son diversas las opciones de geoetiquetado, sin embargo no existe un estudio ni análisis del grado de aplicación de estas geoetiquetas, ni a nivel mundial, ni a nivel nacional, ni mucho menos centrado en el sector turístico. El único dato general que se ha podido encontrar procede de *Whois DataBase* (<http://reviews.gcoupon.com>), donde se indica que se ha explorado el millón de sitios web más relevantes del mundo y se presenta el resultado del conteo de las etiquetas ICBM en dichos sitios. El dato que ofrece este sitio es la aparición de 3267 casos en ese millón de sitios analizados, es decir, sólo un 0.0033% (un tres por mil). Desgraciadamente, no se sabe la fecha del trabajo, los criterios de selección, ni se dispone de ninguna información adicional sobre estos datos.

En este trabajo se pretende realizar una primera aproximación general a la situación que se da en España. Para ello se ha desarrollado el siguiente proceso:

- Selección de una muestra de sitios web españoles.
- Descarga automatizada de los sitios por medio de una araña o robot.

- Búsqueda de geoetiquetas (ICBM, GT, MFG, DRFaG) en las páginas descargadas. También se han considerado geoetiquetas del tipo GeoRSS, por el ya comentado interés para el futuro desarrollo de un estándar de W3C. Junto a las geoetiquetas también se ha incluido en la búsqueda la detección de enlaces a Google Maps como forma de obtener una medida de comparación con otra forma de georreferenciación de elementos de la web y que se considera bastante extendida.
- Análisis de los resultados.

A continuación se comentan los aspectos más relevantes del proceso y de los resultados obtenidos.

El aspecto que más condiciona el estudio es la selección de la muestra. En nuestro caso no se puede realizar un muestreo estadístico y de tipo aleatorio por la imposibilidad de conocer el marco o población de manera previa, sino que se ha desarrollado un muestreo intencional u opinático. Por tanto se ha adoptado un criterio de selección guiada por criterios. Básicamente se han adoptado tres criterios:

- Sectorial: Se han considerado diversos sectores de actividad pensando en las necesidades cotidianas de un ciudadano normal. En total se han considerado 12 sectores, algunos de ellos son: agencias viaje, cadenas de hoteles, Campings, etc.
- Liderazgo: Se han incluido los sitios de las organizaciones líderes (por sector, por zona, etc.).
- Espacial: Se ha pretendido que existiera una distribución espacial que cubriera, en la medida de lo posible, todo el territorio nacional (islas, península, regiones).

En total se seleccionaron manualmente 136 direcciones de sitios web con el reparto sectorial de casos indicado en la Tabla 2.

**Tabla 2. Distribución de la muestra de sitios web analizada en España**

Tipo de actividad	Casos
Agencias Viajes	4
Banca	6
Cadenas de Hoteles	4
Campings	21
Casas rurales	24
Comercio (General)	12
Comercio (Grandes Almacenes, Superficies y Supermercados)	9
Ocio cultura (museos, cines, teatros, óperas)	18
Puertos, Aeropuertos	4
Restaurantes	20
Servicios religiosos	5
Transportes	9

El proceso de visita y descarga automática mediante robot se realizó a principios del mes de junio de 2011. Se limitó la profundidad de los enlaces en el sitio a cuatro niveles. Sólo se descargaron páginas HTML. En total se descargaron 730 Mbytes y 1952 páginas distintas correspondientes a 136 sitios.

El proceso de búsqueda de las geoetiquetas consiste en la localización del patrón formado por las palabras clave de cada tipología de etiquetas. Este proceso se puede realizar de manera elemental con cualquier editor de textos, sin embargo, en nuestro caso la complejidad está en la automatización del proceso de visita de las 1952 páginas.

Los resultados obtenidos son bien pobres en cuanto a cantidad y a posibilidades de análisis. Además, el tipo de muestreo aplicado sólo permite informar sobre sus resultados, pero restringidos a este caso, sin la posibilidad de extrapolar. La Tabla 3 presenta los datos. Los resultados se reducen a la localización de un total de 27 páginas etiquetadas (todas los formatos), correspondientes a 136 sitios.

**Tabla 3. Resultados de uso de las geoetiquetas en la muestra de sitios seleccionada.**

	ICBM	GT	MFG	RDFaG	GeoRSS	Google Maps
Patrón buscado	name="ICBM"	"geo.	class="geo"	"geo:	"georss"	maps.google.com
Nº Sitios encontrados	2	3	1		0	21
% de Sitios encontrados	1.47%	2.21%	0.74%	0.00%	0.00%	15.44%
Nº de Páginas encontradas	9	52	2		0	48
%º de Páginas encontradas	0.46%	2.66%	0.10%	0.00%	0.00%	2.46%
Total de sitios analizados: 116. Total de páginas analizadas: 873						

Fte. Elaboración propia

Mediante técnicas de remuestreo, conocidas como bootstrap (Efron, 1979), se han obtenido los resultados que se presentan en la Tabla 4 relativos a los valores de presencia de etiquetas en sitios con una estructura poblacional como la indicada en la Tabla 2. De esta forma, se puede establecer un valor medio para la estimación y el error asociado a la misma (desviación).

**Tabla 4. Valores estimados por bootstrap.**

	Media (%)	Varianza	Desviación
Caso ICBM	2.09	2.24	1.50
Caso GT	3.01	3.04	1.74
Caso MFG	1.00	1.02	1.01
Caso Google Maps	20.95	17.79	4.22

Como se ha indicado, los resultados son pobres y no permiten hacer un análisis ni por tipo de actividad, ni por distribución espacial, ni por ningún otro aspecto, pero son rotundamente significativos en cuanto a demostrar un uso prácticamente nulo o testimonial de estas etiquetas en nuestro país, y en un orden bastante inferior a la inclusión de llamadas a Google Maps. Sin embargo queremos resaltar también varios aspectos que nos resultan especialmente llamativos:

- Respecto a la información procedente de de Whois DataBase (<http://reviews.goupon.com>), el grado uso en España de geoetiquetas ICBM en los sitios web es mayor.
- Por tipos de actividad los sitios que poseen mayor número de geoetiquetas son los campings.
- Aquellas páginas que disponen de etiquetas ICBM también tienen etiquetas GT.
- Las etiquetas del tipo RDFaG y GeoRSS no tienen difusión alguna.

- La difusión de los mapas de Google en la muestra seleccionada es mayor que la de las técnicas de geoetiquetado, quizá por simplicidad (inclusión de un elemento con el código incrustado de página de Google Maps) o por su valor añadido visual.

## 5. Propuesta de uso

Basadas en el análisis de las opciones de geoetiquetado, tanto desde una perspectiva más conceptual como aplicada (por medio de las páginas analizadas), a continuación se realiza una propuesta de uso de estas técnicas. El objetivo es favorecer un uso lo más consistente posible con la expresión correcta de la posición y la identificación de los objetos, así como de la elección de la opción y la mejor forma de inclusión de las geoetiquetas en los propios documentos html que se vayan a publicar.

### 5.1. Elección del estilo

Dado que se tienen varios estilos de geoetiquetado cuyo objetivo viene a ser el mismo, una de las primeras cuestiones a plantearse es si hay que decantarse siempre por uno de ellos, o si habría que tomar la decisión según algún criterio.

La verdad es que no puede decirse que uno de ellos sea un estándar dominante sobre los demás. Como ya se ha referido, dos estilos introducen la información geográfica en la cabecera del documento html (ICBM y GT) y dos la introducen en el cuerpo (MFG y RDFaG). Ambos planteamientos han de considerarse complementarios, no excluyentes. Véanse los siguientes ejemplos que vienen a ilustrar esta relación:

- *Página web de un hotel.* En este caso lo que interesa es proporcionar las coordenadas del hotel, es decir, una única ubicación. Bastaría con incluir el geoetiquetado en la cabecera. No obstante, esta información no es visible para el usuario de la página (sí para agentes automáticos de búsqueda). Si se quisiera hacerla visible, puede incluirse también la información sobre la posición en el cuerpo del documento. Esta última información podría o no geoetiquetarse, ya que la información útil para los agentes automáticos de búsqueda ya está en la cabecera. En caso de decidir geoetiquetarla, las coordenadas han de ser las mismas que las incluidas en la cabecera.
- *Página web con todos los hoteles de una provincia.* Aquí ya hay más de una ubicación de interés, que puede ser del orden de varias decenas o centenas de hoteles. Esto se traduce en la necesidad de realizar el geoetiquetado en el cuerpo del documento, al ir incluyendo la información de cada uno de los hoteles. No obstante, también sería adecuado incluir el geoetiquetado en la cabecera, en este caso aludiendo a la provincia en cuestión, a la que pertenecen todos los hoteles. Para esto último es muy útil el estilo GT, que permite utilizar etiquetas relativas al país y provincia de forma estandarizada.

En relación a lo anterior, puede decirse que es conveniente que se utilice un estilo relativo a la cabecera en todo documento html que pretenda ser geoetiquetado. De este modo los agentes automáticos de búsqueda pueden comenzar por la cabecera y detectar que se trata de un documento con geoetiquetas y luego rastrear el cuerpo del documento buscando más información. A la hora de elegir entre el estilo ICBM y el estilo GT, la decisión se decanta hacia este último, que contempla la posibilidad de incluir más información, y además está siendo formalizado por el IETF. No obstante, no cuesta ningún trabajo incluir por ahora una línea más de código con la geoetiqueta del estilo ICBM, que prácticamente coincide con la línea de código de la etiqueta *geo.position*, del estilo GT.

En relación a la utilización de MFG o RDFaG, la decisión está menos clara. Suele considerarse que el uso de los microformatos es más simple que el RDF y que por tanto acerca al gran público la posibilidad de introducir metadatos en los documentos html. Ambos son soportados

hoy en día por los principales buscadores, como puede leerse en la ayuda de herramientas para webmaster de Google, en el apartado dedicado a fragmentos enriquecidos y marcado estructural.

## 5.2. Resumen de reglas de uso

Por último, es útil resumir en una serie de reglas todo lo visto anteriormente, con un objetivo meramente práctico, para facilitar un correcto uso de los estilos de geoetiquetado:

- *Elección de estilos.* Si sólo se va a geoetiquetar una ubicación, es aconsejable utilizar un estilo de cabecera GT (el ICBM también se puede añadir, pero no sustituir al GT). Opcionalmente, si se las coordenadas se ofrecen al usuario en el cuerpo del documento, éstas pueden ser etiquetadas mediante MFG o RDFaG. Si se van a geoetiquetar varias ubicaciones en un mismo documento, se recomienda utilizar el estilo GT haciendo referencia a la zona o región que englobe a todas las ubicaciones y el estilo MFG o RDFaG (cualquiera de ellos) para geoetiquetar cada una de las ubicaciones.
- *Información a incluir en el geoetiquetado.* La información esencial es la de las coordenadas planimétricas, latitud y longitud, por este orden y en grados sexagesimales. La inclusión de la altitud, en metros, es opcional en función de los requisitos de uso de la información suministrada (por ejemplo, no parece esencial incluirla para geoetiquetar un documento html de un hotel, pero sí sería interesante para objetos relativos a actividades de senderismo, como los manantiales de un parque natural). En el caso del estilo GT, se recomienda añadir las etiquetas *geo.country*, *geo.al* y *geo.lmk*, si la información a geoetiquetar puede considerarse relativa a un país, provincia y objeto, respectivamente.
- *Sistema de coordenadas planimétrico.* Sólo cobra relevancia para georreferenciar ubicaciones con una incertidumbre reducida (inferior al km). En este caso la norma es utilizar WGS84, para lo cual es necesario realizar la conversión oportuna en caso de ser necesario.
- *Sistema de coordenadas altimétrico.* Sólo cobra relevancia para georreferenciar ubicaciones con una incertidumbre reducida (inferior a 100 m). Aunque el sistema de referencia oficial suele utilizar altitudes ortométricas, y por tanto la cartografía oficial, la generalización en el uso de navegadores GNSS trae consigo el uso de un sistema de referencia global, utilizando la altitud elipsoidal sobre WGS84, que es el recomendado.
- *Ubicación del punto que representa la ubicación de un objeto.* Para objetos de gran extensión como un parque público, una ciudad, una provincia, un parque natural, un país, la ubicación puede quedar aproximadamente en el centro, o a un punto característico (ayuntamiento de una ciudad, capital de un país). Para objetos de menor extensión, como edificios, o parte de ellos, es más conveniente que el punto se sitúe en la entrada que da acceso al interior del recinto. Hay que recordar que el estilo GT permite obviar las coordenadas si se está etiquetando un país, estado o provincia.
- *Número de decimales e incertidumbre.* El número de cifras decimales de las coordenadas de latitud y longitud (y elevación) ha de estar relacionado con el tamaño del objeto que se está geoetiquetando y con la incertidumbre de las coordenadas. Se aumenta el número de decimales conforme el objeto es más pequeño y las coordenadas se conocen con menor incertidumbre. Se reduce el número de decimales si el objeto es más grande o las coordenadas presentan mayor incertidumbre.

## 6. Conclusiones

Este trabajo se centra en el análisis comparativo de los métodos, o estilos, más difundidos para incorporar información geográfica en los archivos html, lo que viene a denominarse etiquetado geográfico o *geoetiquetado*. La incorporación de este geoetiquetado es cada día más importante, conforme se van desarrollando servicios basados en la posición o LBS (location-based services). En la red existe información sobre las alternativas de geoetiquetado, pero se trata de una información bastante dispersa, a veces poco clara o contradictoria, ya que se fundamenta en experiencias de uso de distintos usuarios. Este trabajo ha pretendido presentar y analizar los principales estilos en uso, haciendo referencia a las fuentes que los definen.

Los estilos analizados han sido dos que incluyen la información en la cabecera mediante etiquetas: *ICBM* y *GeoTag* (denominados aquí ICBM y GT, respectivamente), y dos que permiten incluir la información en el cuerpo del documento: mediante el uso del microformato *Geo*, y mediante el uso del vocabulario RDF *Geo* (MFG y RDFaG, respectivamente).

Este trabajo ha incluido una parte experimental centrada en el estudio de la adopción de las geoetiquetas en nuestro país. El análisis se ha realizado sobre un total de 136 sitios web relativos a sectores de uso cotidiano de los ciudadanos. No se trata de un estudio con significación estadística determinada pero sí de gran rotundidad en los resultados que indican la adopción prácticamente nula de las geoetiquetas en nuestro país. Las geoetiquetas no las usan las grandes cadenas (comercios, hoteles, etc.) ni los pequeños comercios. Su adopción parece muy aleatoria y escasa.

Aún así, convencidos del interés de esta tecnología, se ha realizado una propuesta de uso, en la que se plantea la complementariedad entre los estilos relativos a la cabecera y al cuerpo del documento. Se sugiere un planteamiento de uso en esta línea, según la cantidad de ubicaciones que se pretendan geoetiquetar en un mismo documento html (una o varias). También se han resumido algunas directrices generales atendiendo a la información a incluir, los sistemas de referencia, la ubicación del punto respecto al objeto etiquetado y el número de cifras decimales.

## 7. Agradecimientos

Este trabajo ha sido cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER) y por el proyecto P08-TIC-4199 de Excelencia de la Junta de Andalucía. Asimismo, agradecemos a la Junta de Andalucía la financiación económica del Grupo de Investigación Ingeniería Cartográfica (PAIDI-TEP-164) desde 1997 hasta la fecha.

## 8. Referencias

**Daviel, A.; Kaegi, F. A. (2007).** *Geographic registration of HTML documents*. The Internet Engineering Task Force (IETF). Disponible en línea en: <http://tools.ietf.org/html/draft-daviel-html-geo-tag-08> (último acceso el 23 de marzo de 2011).

**Daviel, A., Kaegi, F. A. y Kofahl, M. (2007).** *Geographic extensions for HTTP transactions*. The Internet Engineering Task Force (IETF). Disponible en línea en: <http://tools.ietf.org/html/draft-daviel-http-geo-header-05> (último acceso el 23 de marzo de 2011).

**ISO (2003).** ISO 19112:2003. Geographic information -- Spatial referencing by geographic identifiers. International Organization for Standardization.

**ISO (2006).** ISO 3166-1:2006. Codes for the representation of names of countries and their subdivisions -- Part 1: Country codes. International Organization for Standardization.

**ISO (2006).** *ISO 19138:2006. Geographic information -- Data quality measures.* International Organization for Standardization.

**ISO (2007).** ISO 3166-2:2007. Codes for the representation of names of countries and their subdivisions -- Part 1: Country subdivision code. International Organization for Standardization.

**ISO (2009).** ISO 6709:2008. Standard representation of geographic point location by coordinates. International Organization for Standardization.

**Mayrhofer, A. y Spanring, C. (2010).** *A Uniform Resource Identifier for Geographic Locations ('geo' URI).* The Internet Engineering Task Force (IETF). Disponible en línea en: <http://datatracker.ietf.org/doc/rfc5870/> (último acceso el 20 de mayo de 2011).

**Microformats (2011a).** *Microformats wiki.* Disponible en línea en: [http://microformats.org/wiki/Main\\_Page](http://microformats.org/wiki/Main_Page) (último acceso el 23 de marzo de 2011).

**Microformats (2011b).** *Geo.* Microformats wiki. Disponible en línea en: <http://microformats.org/wiki/geo> (último acceso el 23 de marzo de 2011).

**Microformats (2011c).** *hCard 1.0.* Microformats wiki. Disponible en línea en: <http://microformats.org/wiki/hcard> (último acceso el 23 de marzo de 2011).

**Ruiz, A.M., García, J.L. y Mesa, J.L. (2010).** Error, incertidumbre, precisión y exactitud, términos asociados a la calidad espacial del dato geográfico. En: Alcázar y otros (Eds.), *Catastro: formación, investigación y empresa. Selección de ponencias del I Congreso Internacional de Catastro Unificado y Multipropósito*, Servicio de Publicaciones, Universidad de Jaén, pp.95-102.  
([http://coello.ujen.es/congresos/cicum/ponencias/Cicum2010.2.02\\_Ruiz\\_y\\_otros\\_Error\\_incertidumbre\\_precision.pdf](http://coello.ujen.es/congresos/cicum/ponencias/Cicum2010.2.02_Ruiz_y_otros_Error_incertidumbre_precision.pdf), último acceso: mayo 2011).

**Raymond, E. S. (1996).** *The new hacker's dictionary.* The MIT Press. Massachusetts Institute of Technology.

**Schiller, J. y Voisard, A. (2004).** *Location-Based Services.* San Francisco, Morgan Kaufmann, Elsevier.

**W3C (1999).** *HTML 4.01 Specification.* W3C Recommendation 24 December 1999. Disponible en línea en: <http://www.w3.org/TR/1999/REC-html401-19991224/html40.pdf.gz> (último acceso el 23 de marzo de 2011).

**W3C (2006).** *Basic Geo (WGS84 lat/long) Vocabulary.* Revision 1.21. W3C Semantic Web Interest Group. Disponible en línea en: <http://www.w3.org/2003/01/geo> (último acceso: 29 de mayo de 2011).

**W3C (2007).** *W3C Geospatial Vocabulary.* W3C Incubator Group Report 23 October 2007. Disponible en línea en: <http://www.w3.org/2005/Incubator/geo/XGR-geo-20071023/> (último acceso: 29 de mayo de 2011).

**W3C (2008).** *RDFa Primer. Bridging the Human and Data Webs.* W3C Working Group Note 14 October 2008. Disponible en línea en: <http://www.w3.org/TR/xhtml-rdfa-primer> (último acceso el 23 de marzo de 2011).