E.M. nº 45 Septiembre-Diciembre 2013

EVOLUCIÓN DE LA INFORMACIÓN Y SU CAPTACIÓN EN LA WEB DEL FUTURO

Mariano Rico Almodóvar
Departamento de Ingeniería Informática. Universidad Autónoma de Madrid

RESUMEN

Este artículo muestra la evolución de las investigaciones en Web Semántica, la que presumimos será la Web en un futuro cercano. Se describe la manera en la que un usuario común utilizará esta nueva Web, así como los retos que deben superar estas tecnologías para llegar a popularizarse; entre otros, la reticencia de muchas organizaciones a poner sus datos en abierto, el equilibrio entre la privacidad del individuo y el derecho de los demás a conocer sus datos. También se muestran ciertos detalles de la llamada "Ley de Transparencia" que pueden facilitar mucho la pronta emergencia de esa nueva forma de generación y búsqueda de información.

1. YO GOOGLEO, TÚ GOOGLEAS... (Cómo funcionamos con las búsquedas)

Hace unos días mi hija de 8 años me pidió "buscar en Internet" algo que su profesora les había pedido en clase. Al parecer, un científico les iba a dar una charla sobre la Antártida y la profesora les pidió que preparasen en casa algunas preguntas para plantearle. Abrimos el navegador, y mi hija tecleó "preguntas sobre la Antártida".

Tenemos tan integrado el uso de Google, el afamado "motor de búsqueda" en jerga técnica, que no paramos a pensar cómo hacemos las búsquedas o qué limitaciones tienen. Quizás los más veteranos en tecnologías web recuerden AltaVista, el primer buscador. Corría el año 1995, y resultó bastante útil esta herramienta capaz de ayudarnos a encontrar qué sitios web alojaban páginas con el texto indicado. Entonces, la incipiente Web no estaba "indexada" y sólo el boca a boca o algunos servicios de directorios nos ayudaban a descubrir otros sitios web. Pocos años más tarde aparecería Google y la continuación de esta historia es bien conocida: desde entonces ocupa un lugar privilegiado en el mundo, despachando dos de cada tres consultas en Estados Unidos, y un porcentaje aun mayor en España (1).

Tanto en AltaVista como en Google la idea del buscador ha sido la misma: a partir de unas palabras, proporcionar la lista de páginas web que tienen "mayor relación" con estas palabras. Y no se trata de una tarea sencilla en absoluto. Requiere tener indexado (esto es, catalogado y analizado por programas informáticos) un número de páginas que aumenta a velocidad de vértigo, atender decenas de miles de búsquedas cada segundo, y darles una respuesta en las décimas de segundo que dura un parpadeo.

La parte más compleja de esta tarea, la que dota a Google de su magia, es calcular esa "mayor relación". La mayor parte de nosotros coincidiremos en que Google calcula de manera aceptable esa "mayor relación" entre las palabras que le proporcionamos y los resultados obtenidos, y por eso ocupa el primer puesto en el ranking de buscadores. Pero, también coincidiremos en que Google no

“entiende” las preguntas. Por ejemplo, si tecleamos en Google “libros que citen a libros de García Márquez” veremos que no saca el resultado esperado. Sin embargo, todos tenemos amargas experiencias de búsquedas infructuosas por contener palabras demasiado comunes o que tienen gran contenido semántico. Qué abogado no ha soñado con poder pedir “sentencias de primera instancia sobre apropiación indebida que llegaron hasta el Tribunal Supremo”. O qué historiador no estaría encantado con poder obtener la lista de “reyes que llegaron al trono siendo los cuartos en la línea sucesoria en el momento de su nacimiento”.

Los humanos entendemos sin apenas ambigüedad el significado de cada palabra porque podemos enmarcarlas en un contexto. El más próximo es el contexto de la frase, pero ésta tiene un significado en el contexto de un párrafo, y éste en el contexto de un documento. Pero no paramos aquí, un documento se contextualiza en una cultura y ésta en un marco temporal. Esta contextualización nos permite minimizar las ambigüedades del lenguaje humano. Pretender realizar esta tarea mediante programas informáticos es una labor extremadamente compleja que requiere de ingentes cantidades de cálculo. Una tarea en la que los científicos de Ciencias de la Computación han estado trabajando durante décadas.

Pero ha sido en esta última década en la que se ha tenido disponible un binomio singular que augura buenos resultados. De un lado, se han desarrollado las tecnologías de lo que se ha denominado “Web Semántica”. Por otro lado, se ha alcanzado una potencia de cálculo sin precedentes gracias al uso simultáneo de cientos, miles, de ordenadores, lo que denominamos en jerga “computación distribuida” y que últimamente se vende comercialmente bajo la palabra de moda “computación en la nube”.

2. DATOS, DATOS, DATOS (Datos estructurados, Wikipedia y esDBpedia)

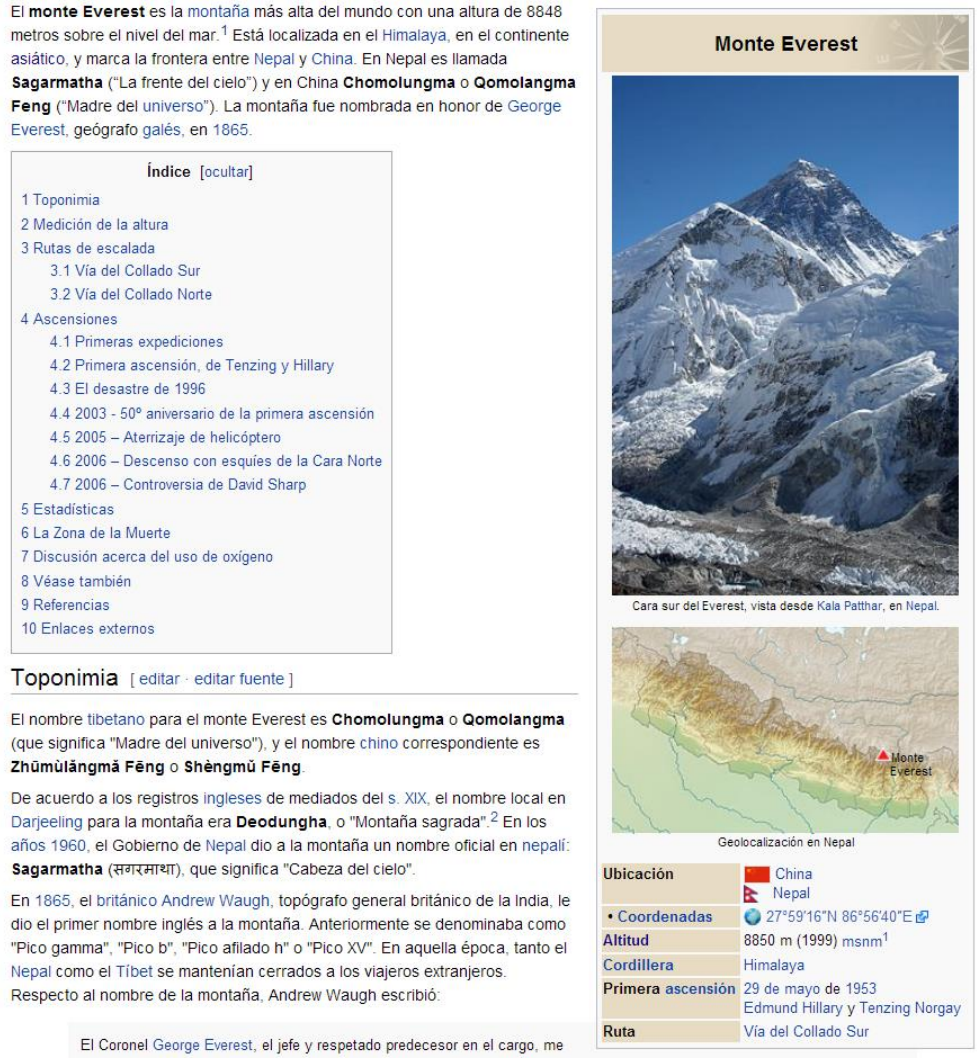
El primer paso para poder dotar de semántica a los datos es disponer de “datos estructurados”, esto es, datos contextualizados y con un valor concreto. Un ejemplo muy simplificado de dato estructurado podría tener esta representación textual “Everest, montaña de Asia, altura 8.848 metros”. Las técnicas actuales de análisis de textos por ordenador no permiten eliminar por completo las ambigüedades del lenguaje humano escrito. Si intentásemos obtener el dato estructurado del ejemplo mediante el análisis de textos, veríamos que hay textos referidos a diversas empresas “Everest” (en mi ciudad hay pastelerías, restaurantes y colegios), un club de fútbol con ese nombre en Guayaquil, un municipio de Dakota del Norte (EE.UU.); y que, en el caso de la montaña, no hay consenso en cuanto a la altura.

Otra posible método para obtener datos estructurados, distinto del análisis automático de textos, consiste en aprovechar el trabajo colaborativo de proyectos como Wikipedia. Esta enciclopedia online es una fuente gratuita de información bastante reputada, que contiene errores, pero no muchos más que los que tienen otras fuentes de pago. Además, esta fuente se mantiene permanentemente actualizada gracias al esfuerzo desinteresado de miles de personas. Si usamos esta fuente podemos estar seguros de que la altura asignada al Everest es la más consensuada ya que, en caso de disputa, Wikipedia dispone de una jerarquía de reputados revisores para resolver estos conflictos.

Una de las claves de Wikipedia es que elimina las ambigüedades los términos, esto es, tiene entradas distintas para la montaña, el club de fútbol ecuatoriano y el municipio estadounidense. Además, y esta es otra clave fundamental, muchas de las entradas disponen, además del texto, de una “caja-resumen” (infobox) que aporta uniformidad. Por ejemplo, en la entrada de la Wikipedia española (del idioma español) del monte Everest (véase la Figura 1) podemos ver la caja-resumen común a todas las montañas en la parte derecha, dentro de un recuadro que tiene en su cabecera el texto “Monte Everest”. En ese recuadro hay fotos que describen la entrada y, más abajo, una tabla de atributos y sus

valores. En este caso se pueden ver, entre otros, los atributos “Ubicación”, “Coordenadas” y “Altitud”, así como sus respectivos valores.

Figura 1. Monte Everest en Wikipedia. A la derecha se observa la caja-resumen (infobox) con los datos de esta montaña.



El **monte Everest** es la montaña más alta del mundo con una altura de 8848 metros sobre el nivel del mar.[1] Está localizada en el Himalaya, en el continente asiático, y marca la frontera entre Nepal y China. En Nepal es llamada **Sagarmatha** (“La frente del cielo”) y en China **Chomolungma** o **Qomolangma Feng** (“Madre del universo”). La montaña fue nombrada en honor de George Everest, geógrafo galés, en 1865.

Índice [ocultar]

1 Toponimia
2 Medición de la altura
3 Rutas de escalada
3.1 Vía del Collado Sur
3.2 Vía del Collado Norte
4 Ascensiones
4.1 Primeras expediciones
4.2 Primera ascensión, de Tenzing y Hillary
4.3 El desastre de 1996
4.4 2003 - 50º aniversario de la primera ascensión
4.5 2005 – Aterrizaje de helicóptero
4.6 2006 – Descenso con esquíes de la Cara Norte
4.7 2006 – Controversia de David Sharp
5 Estadísticas
6 La Zona de la Muerte
7 Discusión acerca del uso de oxígeno
8 Véase también
9 Referencias
10 Enlaces externos

Toponimia [editar · editar fuente]

El nombre tibetano para el monte Everest es **Chomolungma** o **Qomolangma** (que significa "Madre del universo"), y el nombre chino correspondiente es **Zhūmùlǎngmǎ Fēng** o **Shèngmǔ Fēng**.

De acuerdo a los registros ingleses de mediados del s. XIX, el nombre local en Darjeeling para la montaña era **Deodungha**, o "Montaña sagrada".[2] En los años 1960, el Gobierno de Nepal dio a la montaña un nombre oficial en nepalí: **Sagarmatha** (सगरमाथा), que significa "Cabeza del cielo".

En 1865, el británico Andrew Waugh, topógrafo general británico de la India, le dio el primer nombre inglés a la montaña. Anteriormente se denominaba como "Pico gamma", "Pico b", "Pico afilado h" o "Pico XV". En aquella época, tanto el Nepal como el Tíbet se mantenían cerrados a los viajeros extranjeros. Respecto al nombre de la montaña, Andrew Waugh escribió:

El Coronel George Everest, el jefe y respetado predecesor en el cargo, me

Monte Everest

Cara sur del Everest, vista desde Kala Patthar, en Nepal.

Monte Everest

Geolocalización en Nepal

Ubicación	China Nepal
• Coordenadas	27°59'16"N 86°56'40"E
Altitud	8850 m (1999) msnm[1]
Cordillera	Himalaya
Primera ascensión	29 de mayo de 1953 Edmund Hillary y Tenzing Norgay
Ruta	Vía del Collado Sur

En el ejemplo de la figura, Everest (el monte) usa la “caja-resumen montaña”, una plantilla que se definió para almacenar datos como altitud, ubicación o cordillera. En la Wikipedia del idioma inglés, a los usuarios interesados en dar información de una montaña se les recomienda el uso de la plantilla “caja-resumen montaña”, y así se han descrito más de 13.000 montañas. La Wikipedia del idioma inglés ofrece casi 6.000 de estas plantillas, y casi 900 la Wikipedia del español. Para hacernos una idea de la cantidad de información estructurada contenida en la Wikipedia podemos suponer que cada plantilla aporta datos de un único “concepto”; en el ejemplo anterior, el concepto es “Montaña”. En la Wikipedia del idioma inglés, el concepto más usado después del de “Persona” (con más de un millón de entradas), es “Localidad”, con algo más de 300.000 entradas. De cada localidad descrita en Wikipedia usando esta plantilla, los programas informáticos pueden extraer datos estructurados como la ubicación (latitud y longitud), el número de habitantes, o el área (en kilómetros cuadrados)… y así hasta 400 datos distintos de cada una de las 300.000 localidades. En el caso de la Wikipedia del idioma español, la plantilla más usada es “Taxón”, usado para describir especies biológicas, con casi 130.000 entradas (descripciones de especies biológicas que lo usan), con datos como reino, familia, orden… hasta 150 datos distintos para cada especie.

Pero estos datos estructurados necesitan algo más para ser datos semánticos, necesitan un soporte matemático que les dote de un significado preciso. En el ejemplo anterior vimos que de la caja-resumen de la entrada de Wikipedia del monte Everest podíamos saber que tiene 8.850 metros y que se encuentra en la cordillera del Himalaya, pero ¿qué es una cordillera?, y ¿qué es un metro?. Aparte de traducciones a otros idiomas, y aunque los términos "Cordillera" y "Metro" tengan su propia entrada en Wikipedia, debemos tener una descripción matemática precisa que permita, por ejemplo, saber en qué país se encuentra esa cordillera, o convertir pies en metros.

Heredado de la metafísica, los ingenieros informáticos denominan con el término "ontología" a la descripción matemática de los conceptos (también denominados "clases") que pertenecen a un dominio concreto del conocimiento, y de las relaciones entre estos conceptos. La capacidad que tienen las ontologías de describir con precisión los conceptos y relaciones está fuera del ámbito de este artículo, pero un ejemplo puede resultar ilustrativo. Supongamos un conjunto de datos que describen que una persona concreta, digamos "Mariano", tiene dos hijos, digamos "Ángela" y "José". También se indica en ese conjunto de datos que "Mariano" tiene un hermano "Ángel", y que "Ángel" tiene dos hijos "Daniel" y "Santiago". Todos estos individuos pertenecen a la clase "Persona", y las relaciones entre ellos (*padre_de*, *hijo_de*, *hermano_de*) están definidas en una hipotética ontología llamada "Familia". Esta ontología describe con tal precisión las relaciones de parentesco, por ejemplo, las relaciones *primo_de* o *sobrino_de*, que si añadimos un nuevo dato al conjunto, por ejemplo, que "Mariano" es *padre_de* "Calaf", un programa informático será capaz de inferir nuevos datos como que "José" es *hermano_de* "Calaf", "José" es *primo_de* "Daniel", "Calaf" es *primo_de* "Daniel", etc. A estos programas informáticos se les denomina en jerga técnica, y por motivos obvios, "razonadores".

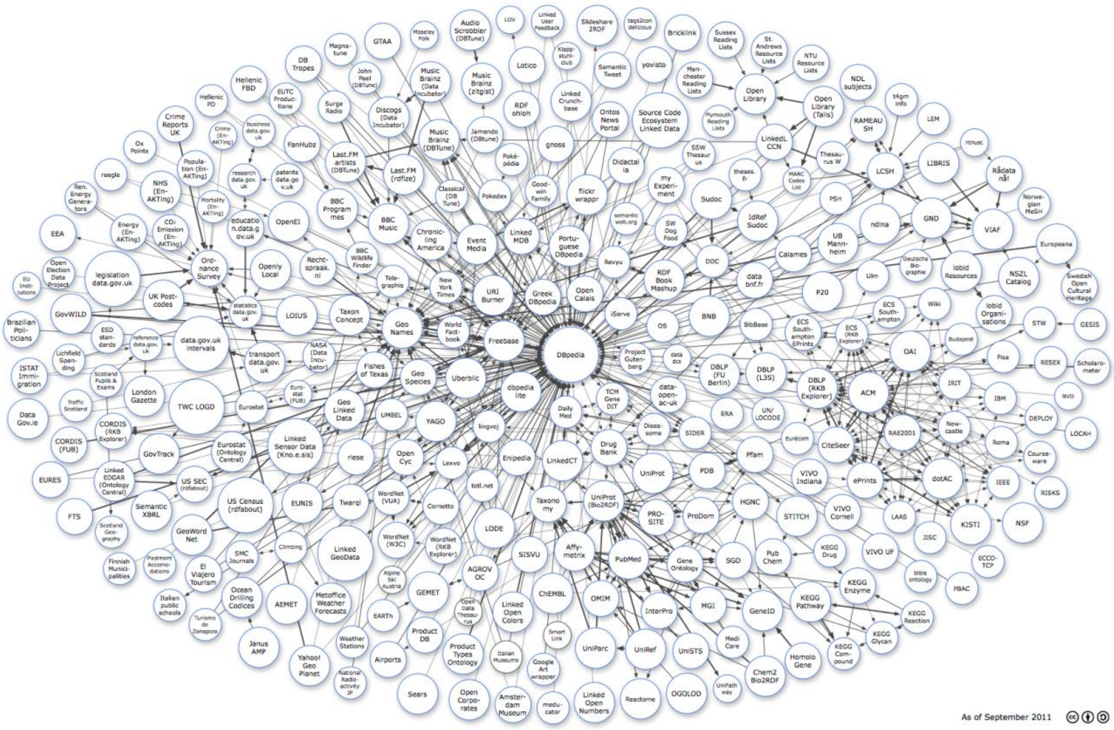
Durante muchos años, la capacidad de descripción (expresividad) de las ontologías ha estado reñida con la capacidad de capacidad de cálculo, de forma que, cuando una ontología estaba descrita con mucho detalle (era muy expresiva), los razonadores tardaban mucho tiempo en inferir los nuevos datos o, incluso, había casos en que se podía demostrar que nunca podrían inferir todos los nuevos datos aun disponiendo de tiempo infinito. En la última década se ha logrado un equilibrio entre la capacidad de descripción y la capacidad de cálculo, de forma que las ontologías actuales garantizan una capacidad descriptiva alta y unos tiempos de cálculo razonables.

Cuando los datos estructurados se relacionan con conceptos o relaciones de una o varias ontologías, diremos que disponemos de datos semánticos. En el caso de la Figura 1, los datos estructurados que se pueden extraer de cada caja-resumen se pueden relacionar con los conceptos y relaciones de las ontologías. Por ejemplo, podemos ligar "Monte Everest" con la clase (concepto) "Montaña" de una hipotética ontología "Accidentes geográficos" en la que se describen fielmente clases como "Cordillera" o relaciones como "en_país" o "altura". La tarea de ligar los datos estructurados con los conceptos y relaciones de las ontologías, lo que en definitiva es dotar de significado semántico, se realiza de forma manual. Pero una vez establecidos estos enlaces, los razonadores realizan de forma automática el resto. Continuando con el ejemplo, una vez establecido que las entradas que usan la "caja-resumen montaña" pertenecen a la clase "Montaña", y establecido el significado semántico de los atributos que componen la plantilla, las 13.000 páginas de la Wikipedia inglesa que describen montañas pueden ser convertidas por programas informáticos en datos semánticos.

Hay miles de ontologías que describen fielmente conceptos, y las relaciones entre estos conceptos, en dominios tan dispares como lo la genética o el patrimonio cultural. Incluso hay ontologías de "dominio amplio" que relacionan entre sí los conceptos y relaciones de las ontologías de dominios específicos. Una de las más importantes es DBpedia (dbpedia.org), el equivalente semántico de la Wikipedia. Los idiomas más importantes de la Wikipedia tienen su propia DBpedia, y en el caso del idioma español se encuentran en es.dbpedia.org, donde se almacenan 100 millones de datos semánticos disponibles para su descarga o consulta.

Pero DBpedia no es el único silo de datos semánticos. Como se puede ver en la Figura 2, DBpedia ocupa un lugar central en la web de datos, no tanto por el número de datos semánticos de que dispones, como por el número de vínculos (relaciones) que tiene con otros silos de datos.

Figura 2. La Web de datos. DBpedia ocupa un lugar central por el alto número de vínculos que tiene con el resto de silos de datos semánticos.



Toda esta ingente cantidad de datos semánticos de alta calidad, permanentemente actualizados (como es el caso de DBpedia, que se alimenta de Wikipedia), puede ser utilizada directamente por usuarios humanos. Por ejemplo, los datos disponibles en esDBpedia sobre “Madrid” se pueden “ver” (se proporciona una visualización para los datos semánticos) en http://es.dbpedia.org/resource/Madrid). Pero, sobre todo, están disponibles para ser usados por programas informáticos de todo el mundo porque están accesibles a través de Internet.

Para mitigar la sobredosis de conceptos y terminología, animo al lector a que visite YouTube y busque “IBM Watson Jeopardy”. En esos vídeos podrá ver el nivel actual de desarrollo de tecnologías como el reconocimiento de voz, el análisis del significado y la recuperación de la información. En este ejemplo, IBM puso a sus científicos a trabajar en un objetivo muy ambicioso: que su sistema (Watson) fuese capaz de vencer a los mejores jugadores de Jeopardy, un juego en el que se hacen preguntas de cultura general a tres concursantes. El sistema de IBM tenía que oír la pregunta (al igual que sus dos contrincantes humanos), entenderla, buscar una respuesta y, si sabía la respuesta, tocar un pulsador para detener el tiempo y responder con voz humana. El resultado de este experimento fue demoledor: Watson venció con absoluta solvencia a los dos mejores jugadores humanos de los últimos años. Esto sucedió en 2011. Quince años antes, Deep Blue, un superordenador específicamente creado para jugar al ajedrez venció al campeón del mundo Garry Kasparov.

Aunque Watson es un prototipo experimental y costoso, permite hacernos una idea del tipo de tareas que podrán resolver las máquinas en el futuro cercano. Como ejemplo de producto para el gran público podemos citar a Siri, el asistente personal incluido en los iPad y iPhone de Apple. Sin llegar a los niveles de Watson, Siri permite tener conversaciones con el propietario del dispositivo del estilo

"cambia la cita con el médico a mañana". En España, la compañía Sherpa ha lanzado el llamado "Siri español" para dispositivos Android.

3. CONVERSIÓN Y TRATAMIENTO DE LOS DATOS (Necesidad de algo distinto: Caso de Adopta un…)

En las oficinas de las empresas o de las administraciones, apenas se trabaja con datos estructurados. A excepción de las bases de datos o las hojas de cálculo, la mayor parte de las aplicaciones ofimática está orientadas a crear documentos que van a ser modificados y leídos por humanos. Un contraejemplo sería la capacidad los ficheros pdf de almacenar lo que Adobe denomina "meta-datos", pero todavía está poco extendido su uso. El uso de meta-datos (datos estructurados) permitirá, por ejemplo, que el pdf de un documento científico almacene datos de los métodos experimentales utilizados para generar los datos mostrados en el documento, y permita garantizar que es reproducible (véase http://www.force11.org/beyondthepdf2).

Para poder añadir semántica a nuestras tareas rutinarias, para poder incluir la semántica en nuestro día a día, primero tenemos que facilitar la creación de datos estructurados. Para ilustrar la dificultad de esta tarea describiré a continuación unos casos reales de necesidad de datos estructurados en los que la falta de estructura de los documentos fue la clave del problema, y la colaboración desinteresada de personas fue la clave de la solución.

En junio de 2009, el parlamento de Reino Unido publicó más de un millón de documentos con los gastos de sus parlamentarios (2) en un ejercicio de "transparencia". Sin embargo, la mayor parte de estos documentos eran archivos pdf en los que las tablas de datos eran imágenes de poca calidad, por lo que su conversión de manera automática a texto era muy difícil o generaba muchos errores. El periódico The Guardian proporcionó una aplicación web para que cualquiera pudiera ver esos pdf y pudiese pasar a texto la información contenida de forma manual. La iniciativa fue un gran éxito gracias a la participación ciudadana, y se logró convertir toda esa información textual en datos estructurados con los que se pudieron hacer análisis, gráficos y comparativas.

En septiembre de 2011, el congreso y el senado españoles publicaron las declaraciones patrimoniales de sus miembros. De nuevo, el formato de los datos era pdf con imágenes incrustadas, y fue la iniciativa de David Cabo, de la asociación Pro Bono Público, la que permitió que una comunidad de voluntarios realizase la conversión de los datos. Este tipo de movimientos se han denominado "crowdsourcing", y están emparentados con movimientos que aspiran a una e-democracia de representación directa.

Recientemente, la iniciativa "Adopta un corrupto" ha permitido obtener datos estructurados, una hoja Excel (3), a partir de los "papeles de Bárcenas" proporcionados por El País (4).

4. TAREAS PENDIENTES (Ley de transparencia y Legislación)

España es el único de los países de Europa con más de un millón de habitantes que no dispone de una legislación que permita a sus ciudadanos pedir datos a sus gobernantes. Este hecho es particularmente significativo puesto que Europa ha sido pionera en el mundo al crear el primer tratado sobre el derecho de sus ciudadanos a la información. Según este tratado, vigente desde 2008 para los 47 países miembros del Consejo Europeo, cualquier particular puede solicitar documentos a cualquier entidad pública de la Unión Europea, esto es, Parlamento Europeo, Consejo Europeo o Comisión Europea (pero no a instituciones de cada país), sin necesidad de justificar el uso que vaya a hacer de esa información, y sin coste para el solicitante. Ya en 2001, la Unión Europea se comprometió a responder en 15 días laborables a las solicitudes de información de sus ciudadanos siempre que la respuesta fuese en formato electrónico u ocupase menos de 20 páginas de papel.

En España, la conocida como "Ley de Transparencia", recientemente aprobada, pero rodeada de controversia, *no* permite solicitar esta información a la Administración, ni impone la obligación de dar una respuesta en un plazo máximo y sin coste para el interesado. El trámite parlamentario fue muy costoso pero, afortunadamente, los detalles relativos al formato de los datos quedaron recogidos en la lista de enmiendas aprobadas el 2 de julio de 2013. En particular, una de estas enmiendas indica que los datos publicados por la Administración deben seguir la llamada "Normativa de Interoperabilidad", donde se especifica el formato de los datos, y donde se obliga a seguir los estándares de la Web de Datos.

Aparte de los aspectos políticos, los nuevos hábitos de uso de los ciudadanos con los contenidos digitales, entre los que se encuentran nuestros datos, producen situaciones complejas que chocan con la normativa vigente o que ni siquiera quedan amparadas por la misma. Podemos citar el "caso Bruce Willis" (5), en el que el actor demandó a Apple por no poder dejar en herencia a sus hijos sus archivos digitales (libros, música, vídeos) almacenados en iTunes. En este caso, chocan un derecho tan básico como el de propiedad privada con el derecho de las empresas sobre los materiales intangibles que venden.

Otro caso de conflicto, este en relación con los datos de la esfera pública, es el planteado por la Agencia Española de Meteorología (AEMET) que, en 2012 dejó de proporcionar (6) sus datos meteorológicos de forma gratuita, para convertirlo en un servicio de pago. En este caso, los datos de una institución pagada con fondos públicos no quedan en la esfera pública, desoyendo las iniciativas Open Data que promueven la puesta en público de los datos de las instituciones públicas.

5. CONCLUSIONES

El sueño de Tim Berners-Lee, inventor de la Web y responsable de la institución que establece los estándares tecnológicos usados por la Web, iba mucho más allá de la actual Web. Él bosquejó y participó en la creación de los estándares tecnológicos de la Web de Datos, y es uno de los principales impulsores del movimiento Open Data. En España, cada vez más gobiernos e instituciones se suman a esta iniciativa y proporcionan datos "en abierto". Destaca el proyecto Aporta (http://datos.gob.es/datos), de la Administración central, que es parte del Plan Avanza 2 (http://www.planavanza.es) para el fomento de la Sociedad de la Información, y que tiene como objetivo "promover una cultura de reutilización de la información del sector público".

De forma análoga a la evolución que han tenido los datos personales en las redes sociales, donde el usuario decide la visibilidad que tienen sus datos, es razonable suponer que pronto surja un movimiento en el que los ciudadanos deseen poner en práctica, y de manera sencilla, los derechos que contempla la LOPD (Ley Orgánica de Protección de Datos de Carácter Personal) para saber qué uso hacen las empresas de los datos personales de sus clientes. Tecnológicamente es posible hacer que, en el futuro cercano, un ciudadano pueda saber qué empresas usan alguno de sus datos, o que sea capaz de revocar el permiso para usar alguno de sus datos de cliente a una o varias empresas. Este ciudadano podrá cambiar de domicilio, modificar este dato una sola vez, y esa modificación será comunicada de forma automática a todas las empresas a las que el ciudadano ha autorizado a usar ese dato.

Tenemos la tecnología, una legislación prometedora, y un volumen razonable de datos semánticos, ¿qué nos falta para tener un buscador semántico?. Mi opinión es que sólo falta algo de tiempo. Probablemente, antes de que mi hija acabe secundaria, todos seremos capaces de preguntarle a Google (o al equivalente de la época) como pregunta mi hija, es decir, como si le preguntásemos a una bola de cristal.

REFERENCIAS

1. *comScore.* [Online] Febrero 2013: http://www.comscore.com/Insights/Press_Releases/2013/3/comScore_Releases_February_2013_U.S._Search_Engine_Rankings.
2. *Unido, Parlamento de Reino.* [Online]: http://mpsallowances.parliament.uk/mpslordsandoffices/hocallowances/allowances-by-mp/.
3. *País, Proporcionados por El.* Google Drive. [Online] [Cited: 8 11, 2013.]: https://docs.google.com/spreadsheet/ccc?key=0Al4XVjh6YN-RdERVWGYyUjYxUkJ4Ulg4aU9OQkR6NkE#gid=0.
4. *País, El.* Todos los papeles de Bárcenas. [Online] [Cited: 8 11, 2013.]: http://elpais.com/especiales/2013/caso_barcenas/todos_los_papeles.html.
5. *Osborne, Hilary.* [Online] 09 3, 2012. [Cited: 10 14, 2013.]: http://www.theguardian.com/money/2012/sep/03/do-you-own-your-digital-content.
6. *Change.org.* change.org. [Online]: http://www.change.org/es/peticiones/agencia-estatal-de-meteorolog%C3%ADa-aemet-que-no-se-cierre-el-acceso-a-los-datos-por-ftp.