

OPENDATA, OPORTUNIDAD ESCONDIDA Y SEMILLA DE LA WEB SEMÁNTICA. OPENDATA, SEMANTIC WEB SEED AND A HIDDEN OPPORTUNITY.

Autores:

Agustina Piedrabuena Moraleda. Telefónica.

agustina.piedrabuenamoraleda@telefonica.es

Luis Criado Fernández. luis@criado.org

Resumen:

Todo lo que dábamos por establecido está cambiando de manera vertiginosa. Un tema que parece estar especialmente agitado en estos días es el vínculo entre las personas y los datos: "the data". Los ciudadanos empiezan a ser conscientes de que el conocimiento y tratamiento de datos públicos no es un privilegio limitado al uso exclusivo de las administraciones sino un derecho de todos, que puede permitir transformar el modelo de gobernanza pública y además, generar todo un nuevo sector económico en torno a él. Necesitamos verdaderas plataformas de OpenData en la Web para mirar al futuro. OpenData implica que se conecten los datos a nivel semántico y por lo tanto será la semilla del nuevo paradigma de la Web, la Web Semántica. Si somos capaces de aprovechar la oportunidad que las iniciativas OpenData y los desarrollos de web semántica nos brindan, se convertirán en una auténtica oportunidad de transformación social y económica.

Abstract:

All that we took for granted is changing dramatically. One issue that seems to be especially interesting in these days is the link between people and data, "the data". Citizens are becoming aware that the knowledge and treatment of public data is not a privilege limited to the exclusive use of the governments but a right of everyone, which may allow transforming public governance model and also create a whole new industry around it. We need effective OpenData platforms to face the future. OpenData enables connection the data at the semantic level and therefore it will be the seed of the new paradigm of the Web, the Semantic Web. If we can get the chance that OpenData initiatives and semantic web developments give us, it will become a huge opportunity for genuine social and economic transformation.

Palabras clave: OpenData, web semántica.

Keywords: OpenData, semantic web.

Todo lo que dábamos por establecido está cambiando de manera vertiginosa. Y un tema que parece estar especialmente agitado, o al menos cuestionado por todos, es el vínculo entre las personas y los datos: “the data”. Los ciudadanos empiezan a ser conscientes de que el conocimiento y tratamiento de datos públicos no es un privilegio limitado al uso exclusivo de las administraciones sino un derecho de todos, que puede permitir transformar el modelo de gobernanza pública y además, generar todo un nuevo sector económico en torno a él.

El movimiento OpenData está en un punto de inflexión. Hemos pasado de una primera fase mantenida esencialmente por el activismo a una segunda fase en que empieza a estar ya en la agenda de muchos gobiernos locales, regionales y estatales. A pesar de estos éxitos, la apertura de datos apenas está en el inicio de su largo camino porque aunque su potencial es inmenso, su impacto en la sociedad sigue siendo muy limitado. Uno de los grandes impedimentos es el desconocimiento de qué es y qué implica la apertura de datos. En este artículo vamos a intentar despejar algunas de estas incógnitas.

¿QUÉ ES LA APERTURA DE DATOS PÚBLICOS (OpenData)?

La apertura de datos públicos **-también conocido como apertura de datos, Datos Abiertos u OpenData-** es una filosofía y práctica que persigue que determinados datos e informaciones pertenecientes a las Administraciones Públicas sean accesibles y estén disponibles para todo el mundo, sin restricciones técnicas ni legales, en formatos digitales, estandarizados y abiertos, siguiendo una estructura clara que permita su manejo y comprensión. El fin de la iniciativa Open Data es que la información pueda ser redistribuida y reutilizada tanto por los ciudadanos como por empresas para conseguir un beneficio para todas las partes. De esta forma, la sociedad —ciudadanía, empresas y cualquier institución— puede acceder fácilmente a estos para informarse o para crear nuevos servicios aumentando el valor social y también el valor comercial.

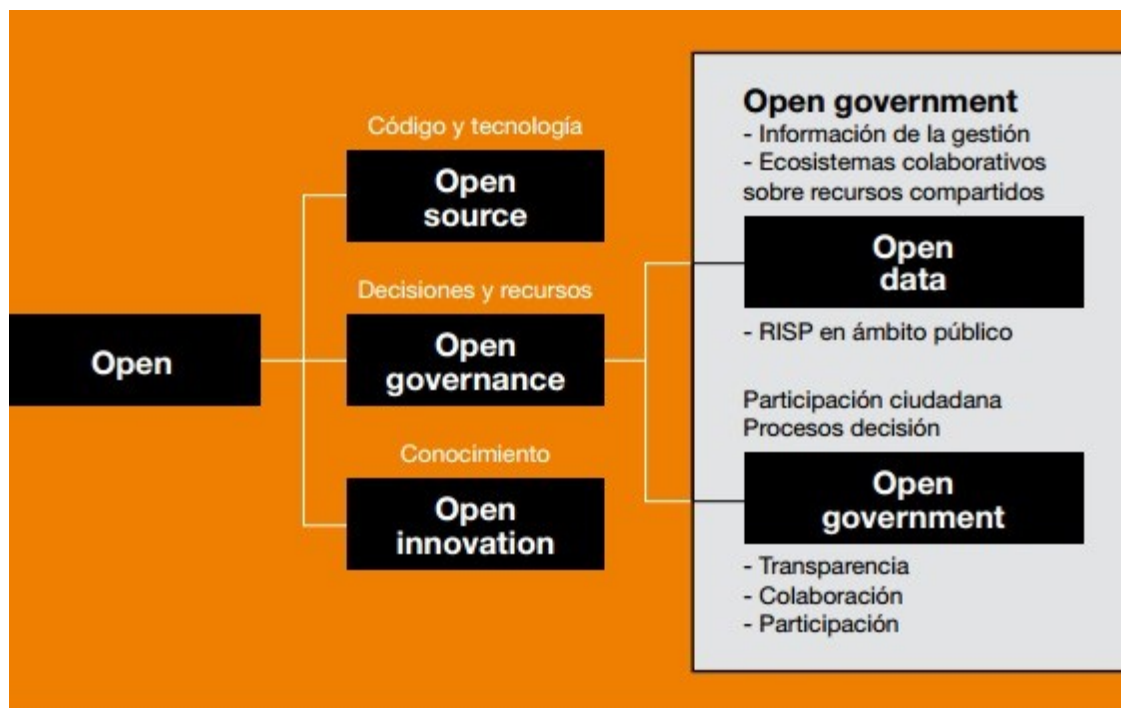


Figura 1. Contexto del OpenData dentro de lo 'open'.

¿QUÉ ES RISP?

La Reutilización de la Información del Sector Público es el objetivo principal de la iniciativa OpenData. Consiste en poner la información del sector público disponible, en bruto y en formatos estándar abiertos, facilitando su acceso y permitiendo su reutilización tanto a particulares como a empresas para distintos fines.

El sector público produce una gran variedad de información que es potencialmente reutilizable, como puede ser la información social, económica, geográfica, estadística, etc. Esta información es muy atractiva para su reutilización ya que es completa, fiable y de calidad.

Se quiere lograr que la reutilización y redistribución de la información pública repercuta en beneficio tanto de las administraciones públicas que sacan a la luz sus datos, como de los ciudadanos y empresas que ven cubiertas algunas de sus necesidades.

OpenData EN ESPAÑA.

La adhesión a la iniciativa OpenData se está haciendo de forma progresiva en todo el mundo. Comenzó con el proyecto data.gov en EEUU, el sitio donde se encuentra el catálogo de datos abiertos del gobierno federal, y poco a poco se ha ido extendiendo a otros países. Cada vez son más los gobiernos que anuncian y comienzan iniciativas de transparencia y gobierno abierto y que publican sus catálogos de datos. Actualmente existen proyectos de OpenData desarrollados además de en EEUU, en países como Canadá, Reino Unido, Australia, Nueva Zelanda, Suecia o Noruega y también en España.

En España, el Ministerio de Industria y Comercio está llevando a cabo el [Proyecto Aporta](#), que pretende colocar a España en la vanguardia de la reutilización de la información del sector público en Europa. Este proyecto está enmarcado dentro del [Plan Avanza2](#), que afronta el reto de conseguir que la utilización de las Tecnologías de la Información y la Comunicación (TIC) aporten cada vez más ventajas a los ciudadanos.

El proyecto Aporta (ahora datos.gob.es, como portal de carácter nacional que organiza y gestiona el Catálogo de Información Pública de la Administración General del Estado) promueve una cultura de reutilización de la información en el ámbito de las Administraciones, concienciando de la importancia y el valor que tiene esta información y su posterior reutilización. También quiere facilitar la puesta a disposición de toda la información posible para aprovechar el mercado de reutilización de la información en España a través de acciones tanto de formación y concienciación, como con el desarrollo de una sobre reutilización de información generada por el sector público, llamada Guía Aporta.

Existen ya algunos proyectos que han sido desarrollados por diversas Administraciones Públicas, como por ejemplo: [Open Data Euskadi](#), [Catálogo de Datos de Asturias](#), [Dades Obertes Generalitat de Catalunya](#), [Dades Obertes Caib](#) (Islas Baleares), [Datos Abiertos Andalucía](#), [Datos Abiertos Castilla la Mancha](#), [Datos Abiertos Castilla y León](#) y [Open Data Xunta de Galicia](#) a nivel de Comunidades Autónomas. También se están llevando a cabo proyectos de Open

Data a nivel local, como en el [Open Data de Barcelona](#), [Ayuntamiento de Badalona](#), [Dades Obertes de Lleida](#), el [Ayuntamiento de Zaragoza](#) y el [Ayuntamiento de Gijón](#).

18 Catalogs filtered from 189 originally ([Reset All Filters](#))

Stars	Catalog	Location	Country	
★★★★★	Asturias Public Data	Asturias	Spain	edit
★★★★★	Zaragoza Public Data Catalogue	Saragossa	Spain	edit
★★★★★	Abert@s (Galicia Open Data)	Galicia	Spain	edit
★★★★★	Balearic Islands Open Data	Balearic Islands	Spain	edit
★★★★★	Catalonia Open Data Catalog	Catalonia	Spain	edit
★★★★★	Gijón Public Data Catalogue	Gijón	Spain	edit
★★★★★	Open Data Euskadi	Basque Country	Spain	edit
★★★★★	Open Data Junta de Castilla y León	Castile and León	Spain	edit
★★★★★	OpenData BCN	Barcelona	Spain	edit
★★★★★	Portal de Datos Abiertos de JCCM	Castilla La Mancha	Spain	edit
★★★★★	Andalusia Open Data Catalog	Andalusia	Spain	edit
★★★★	Badalona Open Data	Badalona	Spain	edit
★★★★	Lleida Open Data	Lleida	Spain	edit
★★★★	Open Data Catalog Spain	España	Spain	edit
★★★★	Open Data Navarra	Navarra	Spain	edit
★★★★	Opengov.es	Spain	Spain	edit
★★★	Open Data Córdoba	Cordoba	Spain	edit
★	Extremadura Reutiliza	Extremadura	Spain	edit

Figura 2. Iniciativas OpenData en España (Fundación CTIC).

Contrariamente a lo que nos pueda llevar a creer el tradicional pesimismo español que nos empuja a creer que España siempre está a la cola de prácticamente todo y nos hace mirar a nuestros vecinos con envidia, **en materia de OpenData España es el segundo país más avanzado de Europa**. España puede considerarse el segundo país más desarrollado en OpenData público (RISP) de la UE tras el Reino Unido, que junto con los Estados Unidos pueden considerarse líderes mundiales. España muestra una buena receptividad de las administraciones y una elevada descentralización de las iniciativas. Dentro del ámbito público, la administración del País Vasco puede considerarse la más destacada a nivel autonómico mientras que el Instituto Geográfico Nacional es la entidad pública más avanzada, aunque el lanzamiento de proyectos se está multiplicando en los últimos meses.

OpenData Y GOBIERNO ABIERTO: LA TRANSFORMACIÓN DE LA ADMINISTRACIÓN PÚBLICA.

El Gobierno Abierto surge como un nuevo paradigma y modelo de relación entre los gobernantes, las administraciones y la sociedad: transparente, multidireccional, colaborativo y orientado a la participación de los ciudadanos tanto en el seguimiento como en la toma de decisiones públicas.

Para algunos autores el concepto trae aparejado un quiebre de envergadura y un cambio radical del paradigma sobre lo que significa gobernar en sociedades complejas: “Muchas personas piensan que la idea de un gobierno abierto tiene que ver con la libertad de información, pero su impacto es mucho más profundo. De hecho, es parte del cambio más grande en la estructura y arquitectura del sector público desde el siglo pasado. Hay una enorme cantidad de datos que posee el gobierno, tales como datos sobre el cambio climático, el éxito de los empresarios, las emisiones de gas, los accidentes de bicicleta, etc. Con gobiernos que empiezan a generar datos y los ponen a disposición en bruto para los ciudadanos, la gente se podrá auto-organizar para usar/reutilizar éstos datos creando valor. No se trata de externalización o privatización. Se trata de una nueva división del trabajo en la sociedad acerca de cómo crear valor público. El resultado es una mejora de los servicios y un gobierno que cueste menos” (Tapscott, 2011).

Visto de este modo, “Se trata de una nueva cultura de la comunicación, un nuevo modelo organizativo y la liberación del talento creativo dentro y fuera de los perímetros de la función pública. No hablamos solo de tecnología, sino de una tecnología social y relacional que impulsa y estimula una cultura de cambio en la concepción, gestión y prestación del servicio público” (Gutiérrez-Rubí, 2010).

Distintos gobiernos en todo el mundo han iniciado un proceso progresivo para impulsar y llevar a cabo estrategias ligadas al concepto de “Gobierno Abierto” (Open Government), cuyo común denominador está ligado a mejorar los niveles de transparencia mediante la apertura de datos públicos (para ejercer control social y rendición de cuentas, *accountability*) y la reutilización de la información del sector

público (RISP); facilitar la participación de la ciudadanía en el diseño e implementación de las políticas públicas (e incidir en la toma de decisiones); y favorecer la generación de espacios de colaboración entre los diversos actores, particularmente entre las administraciones públicas y entre éstas y la sociedad civil y el sector privado. Dichos esfuerzos se enmarcan en el propósito de fortalecer los sistemas democráticos, incrementar los niveles de confianza de la ciudadanía en las instituciones políticas, potenciar la participación y el compromiso cívico, y mejorar la calidad, eficacia y eficiencia de los gobiernos y sus aparatos administrativos. Todo ello a través de la generación de nuevos espacios para la innovación institucional, el desarrollo de mecanismos ligados a la co-creación en la provisión de servicios públicos y al trabajo colaborativo, y la puesta en marcha de nuevos entramados organizativos y sistemas de gestión orientados a fomentar la apertura y una nueva forma de gobernar en el contexto de la sociedad red (Castells, 2010).

En este escenario, cabe destacar que dicho concepto no es nada nuevo. A fines de los años 70 del siglo XX, apareció por primera vez en el espacio político británico y en su concepción original trataba diversas cuestiones relacionadas con el secreto de gobierno e iniciativas para “abrir las ventanas” del sector público hacia el escrutinio ciudadano, con el objeto de reducir la opacidad burocrática (Chapman y Hunt, 1987). En la actualidad, se ha posicionado como un nuevo eje articulador de los esfuerzos por mejorar las capacidades de gobierno y modernizar las administraciones públicas (OECD, 2010), todo ello potenciado por las posibilidades que alberga la Web 2.0, que facilitan el intercambio de información interactiva y en tiempo real, interoperabilidad, diseño centrado en el usuario y espacios genuinos de colaboración y promoción de la participación. Ello se refleja en la identificación de una serie de potenciales beneficios (OECD, 2010), tales como:

- (R)establecer una mayor confianza en el gobierno.
- Garantizar mejores resultados al menor coste. El co-diseño y ejecución de políticas, programas y (provisión de) servicios con los ciudadanos, las empresas y la sociedad civil ofrece el potencial para explotar un depósito más amplio de ideas y recursos.

- Elevar los niveles de cumplimiento de las políticas y asegurar que las decisiones adoptadas sean percibidas como legítimas por la ciudadanía.
- Asegurar la equidad de acceso a la formulación de políticas públicas mediante la reducción del umbral para el acceso a los procesos de decisión que enfrentan las personas como barreras para la participación.
- Fomentar la innovación y nuevas actividades económicas.
- Mejora de la eficacia mediante el aprovechamiento de los conocimientos y los recursos de los ciudadanos que de otra forma se enfrentan a barreras para participar. La participación ciudadana puede garantizar que las políticas sean más específicas y atiendan a sus necesidades, eliminando desgastes potenciales.

Los principios básicos en los que se sustentan las cada vez más numerosas iniciativas ligadas al Gobierno Abierto tienen su origen en el Memorando sobre Transparencia y Gobierno Abierto que el gobierno americano basó en los siguientes ejes: transparencia, participación y colaboración (Obama, 2009). Observando la evidencia reciente y la evolución de las estrategias de Gobierno Abierto en el mundo, podemos sintetizar los contenidos de base de cada uno de estos tres pilares:

- **Transparencia.** Un gobierno transparente proporciona información sobre lo que está haciendo, pone a disposición sus fuentes y bases de datos, y publica los planes de actuación por los que puede ser considerado responsable frente a la sociedad. Ello fomenta y promueve la rendición de cuentas ante la ciudadanía y un permanente control social, al mismo tiempo que refuerza la confianza, la cohesión social y el fortalecimiento institucional.
- **Participación.** Un gobierno participativo promueve el derecho de la ciudadanía a participar activamente en la formulación de las políticas públicas y a facilitar el camino para que las administraciones públicas se beneficien del conocimiento, ideas y experiencia de los ciudadanos mejorando la calidad y eficacia de la acción pública. Promueve la creación de nuevos espacios de encuentro y dialogo que favorezcan el protagonismo, implicación y deliberación de los ciudadanos en los asuntos públicos, todo lo cual obedece a responder a la complejidad creciente de los problemas que son el origen de

las políticas públicas en un contexto en el que la tecnocracia y las jerarquías dan paso a las plataformas deliberativas, la gestión mediante redes y la sabiduría de las multitudes (o “WikiGovernment”).

- **Colaboración.** Un gobierno colaborativo compromete e implica a los ciudadanos y demás agentes sociales en el esfuerzo por trabajar conjuntamente para resolver los (cada vez más complejos) problemas públicos, aprovechando el potencial y energías disponibles en vastos sectores de la sociedad. Ello supone la colaboración, cooperación y el trabajo coordinado no sólo con la ciudadanía, sino con las empresas, asociaciones, organizaciones de la sociedad civil y demás actores, y permite el esfuerzo conjunto dentro de las propias administraciones, entre ellas y sus funcionarios transversalmente.

En términos prácticos y para concretar dichos principios, las estrategias de Gobierno Abierto se han venido cristalizando a través de dos ejes fundamentales:

1. La apertura de datos públicos (open data) y “portales de transparencia” que promueven, por un lado, la innovación usando (o reutilizando) los datos públicos como catalizador de nuevas aplicaciones y servicios y, por tanto, transforma a los gobiernos de proveedores de servicios a gestores de plataformas, permitiendo que otras entidades y/o usuarios, utilizando los datos públicos liberados, puedan crear nuevas prestaciones en la Web, generar nuevas actividades económicas y agregar valor público. Por otro lado, devolviendo a los ciudadanos los datos que son financiados por ellos mismos, incrementando así la transparencia, la rendición de cuentas y el escrutinio público permanente.
2. La apertura de procesos (open process), las plataformas de colaboración y el uso de la Web 2.0 y redes sociales, que persigue facilitar la comunicación y la participación, aprovechar el conocimiento y la experiencia de los ciudadanos para ayudar en el diseño de políticas y provisión de servicios públicos, y la colaboración (en red) dentro y entre las administraciones públicas y más allá.

En síntesis, el Gobierno Abierto fluye desde la convicción de repensar gobiernos y administraciones, de transformar la sociedad y contribuir a desarrollar democracias más sanas, de pasar de las jerarquías a las redes, de asumir compromisos transversales y generar en conjunto con otros actores sociales y económicos, valor público y social. Hoy en día, representa un emergente movimiento en todo el mundo, más allá de la divergencia en cuanto a su definición y la aplicación de las ideas que le dan sustento.

OpenData Y SU DIMENSIÓN INDUSTRIAL.

Como decíamos, OpenData es el contexto económico en el que las Administraciones Públicas ponen a disposición de los ciudadanos los datos en su poder y permiten su combinación con datos privados para la creación de nuevos servicios o la mejora de los existentes. OpenData pone a disposición de terceros juegos de datos sin tratar, raw data, con objeto de que se usen para consulta, análisis, seguimiento de la gestión, generar nuevos productos y servicios, o en general, cualquier actividad a la que estos terceros los quieran destinar.

Una estrategia de Datos Abiertos se vertebra en torno a un portal que organiza y permite acceder al catálogo de datos publicado por una organización, pero abarca muchos más elementos como son:

- Tecnología que permita extraer y publicar la información, enriquecer de la información, usar las capacidades que aporta la web semántica, usar elementos que aprovechen los datos como aplicaciones y sistemas de análisis y control de Gestión.
- Organización que tenga interiorizado el que el acceso y gestión de los datos, e información, genera riqueza dentro y fuera de esta. Y que dentro de sus parámetros de desarrollo esté el derribar las barreras a la información y establecer mecanismos estándar de consumo.
- Objetivo que identifique para que se publiquen los datos y que se pretende conseguir.



Figura 3. Cadena de valor del OpenData.

El primer componente de la cadena son las fuentes de datos, que pueden tener origen en organizaciones públicas o privadas, dividiéndose estas últimas en organizaciones con ánimo de lucro (empresas) o sin él (fundaciones y ONG).

El segundo elemento de esta cadena de valor son los esquemas legales y los mecanismos técnicos que permitan la reutilización, que incluye una evaluación legal de las fuentes, la definición de licencias de reutilización, etc. Así mismo hay que implantar la infraestructura tecnológica para la publicación y puesta a disposición del público de acuerdo con las condiciones legales establecidas.

Los infomediarios, son los creadores de los productos y servicios basados en las fuentes. De nuevo pueden ser entidades de carácter público, privado o del tercer sector y, por tanto, sus objetivos pueden variar desde el mero interés económico, hasta el desarrollo de la propia sociedad, o la promoción de la transparencia o la participación, entre otros.

Basados en estos datos los infomediarios pueden generar, bien nuevos conjuntos de datos mediante el tratamiento y la combinación, aplicaciones para dispositivos móviles, tablets u otros dispositivos, o puramente servicios, bien sean disponibles en línea o como asesoría, que constituyen el cuarto elemento de esta cadena de valor.

El último eslabón de esta cadena son los usuarios, bien de ámbito ciudadano (con modelos de negocio gratuitos o sufragados por publicidad) o desde el ámbito profesional, que utilizan los productos como una parte de su propio negocio.

Los actores de esta cadena de valor se ven condicionados por los reguladores existentes, tanto dentro del marco legal, con una profusa normativa pública, como dentro del marco tecnológico, donde la elaboración, publicación y difusión de estándares, son elementos claves para que el sector pueda desarrollarse hasta su máximo potencial.

El estudio PIRA (Oct-2000) para la Unión Europea cifraba en torno al 1,4% del PIB el potencial económico de la UE (en un margen de 26.000 a 134.000 millones de euros), de la reutilización de información solo del sector público. Posteriormente el informe MEPSIR70 (2006) estima el valor económico de la reutilización de las informaciones generadas en el ámbito público en un margen de 26.000 a 47.000 millones de euros para la UE, lo que representaría para España un valor mínimo de unos 2.080 millones de euros, lo que a tasas estándar de ocupación de empleo podría suponer unos 45.000 empleos en los próximos 10 años.

Según datos del "**Estudio de Caracterización del Sector Infomediario**" realizado por el Observatorio Nacional de las Telecomunicaciones y para la Sociedad de la Información (ONTSI), el volumen de negocio asociado directamente a la actividad de las empresas que generan aplicaciones, productos y/o servicios de valor añadido destinados a terceros, a partir de la información del Sector Público, se sitúa entre 332 y 550 millones de euros. Este mismo informe señala que la información reutilizada procede mayoritariamente de organismos nacionales y que la mitad de las empresas reutilizan también información internacional, esencialmente la producida desde las instancias europeas.

Aunque se dispone de algunos informes sobre el valor económico de la reutilización de datos del sector público, es evidente que la actualidad, tanto general como particular, vendrá no sólo del tratamiento diferenciado de estos datos sino de su mezcla con datos privados, con licencias tanto abiertas como cerradas. Uno de los puntos a resolver es la complejidad jurídica de la mezcla de datos públicos (que

deberían ser de libre acceso) con datos privados que puedan estar restringidos con licencias propietarias.

Otra gran oportunidad –y otra gran fuente de datos públicos- será la progresiva gestión de las ciudades y de otras infraestructuras de uso público y privado a través de comunicaciones Machine-to-Machine (M2M). Es una faceta de las denominadas ‘**smart cities**’. Muchos de esos datos podrían ponerse a disposición de reutilizadores, tanto públicos como privados, para la generación de nuevas aplicaciones.

NIVELES DE PARTICIPACIÓN OpenData

Lo importante es estar dispuesto a compartir datos, y una vez que se tiene claro este compromiso, sería un buen camino proporcionar datos abiertos con el objetivo de lograr cinco estrellas que supone situarse en la excelencia tecnológica en el ámbito de compartir datos siguiendo los estándares de la Web Semántica, esto es, usando anotación en RDF u OWL, para poder ser explotado mediante SPARQL.

RDF es un lenguaje que se puede representar en XML para definir recursos en la Web. El mecanismo que utiliza RDF para la definición de recursos es la tripleta, este modelo proporciona un ámbito semántico simple pero eficaz, que se basa en tres términos; sujeto o recurso, propiedad o predicado y objeto o literal. Con el lenguaje de ontologías web (OWL) se permite representar ontologías a partir de un vocabulario más amplio y una sintaxis más fuerte que RDF. El lenguaje OWL proporciona la funcionalidad de aplicar lógica descriptiva. Resulta especialmente interesante, el hecho de que OWL, puede embeberse en RDF/RDFS y expresarse como tripletas.

Cuando tenemos muchos datos semánticos, en cualquiera de los estándares que hemos comentado, surge la necesidad del acceso a los datos de una forma sencilla. SPARQL se diseñó como lenguaje de consultas para tripletas y viene a ser el equivalente a SQL con las bases de datos relacionales.

Y ¿qué es esto de las cinco estrellas? Es una propuesta de clasificación (Berners-Lee, 2006) para determinar cómo de abiertos y reutilizables son los datos que puede ofrecer una entidad en su sitio web. Dicha clasificación se basa en cinco niveles, desde una estrella hasta cinco. De manera, que actualmente, las webs que proporcionan datos abiertos procuran alcanzar el objetivo de cinco estrellas.

Una estrella, se logra fácilmente, basta ofrecer los datos en cualquier formato (pdf, una imagen, etc). Para lograr una estrella es suficiente especificar en la web de la entidad que los datos publicados son abiertos.

Para conseguir una segunda estrella, se debe facilitar datos de manera estructurada, como puede ser un archivo "Excel". En caso de que se utilice un formato estructurado no propietario, por ejemplo, en lugar de Excel utilizar CSV, se obtienen tres estrellas.

La cuarta estrella implica utilizar los estándares de la Web Semántica. Por lo tanto, ya hay que utilizar URIs para identificar conceptos (clases y propiedades), lo que implica el uso de ontologías RDF/OWL. Esto significa que un tercero podrá enlazar con nuestros datos. Es decir, con cuatro estrellas publicamos datos formalizados con los estándares de la Web Semántica y permitimos que terceros establezcan enlaces (Linked Data) con nuestros datos. Cuatro estrellas implica implementar las tres primeras reglas de Linked Data también propuestas por Tim Berners-Lee.

Finalmente, para obtener la quinta y última estrella, tendremos también que establecer enlaces (Linked Data) con terceros (cuarta regla de Linked Data)

Todo esto, sirve para dotar a la Web actual de datos formalizados, datos semánticos para los sistemas informáticos. Es obvio que Berners-Lee, desde el Consorcio de la W3C, estimula que las webs evolucionen a cinco estrellas con el propósito de disponer de datos formalizados lo más actuales posibles, favoreciendo que la Web actual evolucione a una Web de datos semánticos, a partir de la cual se puede pensar en aplicaciones semánticas a nivel global, construyendo, entre todos la Web Semántica.

EJEMPLOS PRÁCTICOS

Durante todo el artículo hemos procurado dar una visión de lo que es y para qué sirve la iniciativa OpenData. Es ahora el momento de explicar cómo se pueden utilizar.

El ritmo de proyectos en este ámbito es tremendo y resulta imposible relatar todas las referencias de publicación de datos abiertos. Pero para avanzar siempre es útil estudiar con ejemplos, de manera que en este epígrafe vamos a repasar algún ejemplo que sea lo más cercano posible.

Una forma de participar, puede ser realizando un planteamiento gradual, es decir, si formalizar datos en RDF supone retrasar o no publicar datos abiertos, entonces pongamos metas más cercanas; dos o tres estrellas están al alcance de cualquier organización con ánimo de transparencia, participación, colaboración, etc... Un buen ejemplo de esto es la Empresa Municipal de Transportes de Madrid (EMT). Esta empresa ha creado una web concreta para OpenData. El acceso a la plataforma OpenData de EMT por parte de terceros queda regulado a través de un acuerdo firmado por ambas partes, en el que se describen la finalidad y condiciones de uso de los datos. Cualquier persona física o jurídica puede solicitar el acceso. Obsérvese que OpenData no significa publicar datos de manera descontrolada, se puede regular.

Un ejemplo con más estrellas puede ser la Web de Datos Abiertos de Castilla-La Mancha. Si analizamos este portal, nos llama la atención el recuadro granate con el título de “últimos datos abiertos” (bajo la licencia Creative Commons-Reconocimiento). Seleccionando cualquiera de los enlaces nos vamos a una pantalla que nos permite diferentes formas de acceder a la información. Observe que a la derecha dispone de diferentes opciones de descarga, del estilo de la siguiente captura de pantalla



Figura 4. Opciones de descarga de datos (opendata.jccm.es).

Pero lo que nos interesa son datos con estructura semántica, de manera que podemos trabajar directamente con datos formalizados en RDF utilizando el lenguaje SPARQL. Por lo tanto, lo primero es disponer de un programa cliente que permita realizar consultas SPARQL. Una buena opción es instalar *twinkle*. Esta herramienta es gratuita y está escrita en Java.

Un vez instalado y arrancado el programa, para usar los datos de la Web de Castilla-La Mancha sólo tiene que copiar la ruta del enlace a los datos RDF en “Data URL”



Figura 5. Uso de datos (opendata.jccm.es).

Con la consulta anterior obtenemos todas las tripletas contenidas en el fichero RDF, de octubre del 2012 del portal de OpenData de Castilla-La Mancha. A partir de aquí, es cuestión de ir depurando las consultas SPARQL para obtener información más útil. El problema es que normalmente hay que estudiar el modelo, lo mismo que ocurre cuando intentamos realizar consultas SQL en un modelo relacional.

Supongamos que queremos obtener una lista de las ofertas de trabajos que se ofrecen en Castilla La Mancha, podemos entonces escribir algo del estilo siguiente:

PREFIX jccm:

<http://www.jccm.es/cs/Satellite/OpenData/RDF/schema/1212729415646#>

SELECT ?oferta ?masinformacion ?fecha

WHERE {

 ?subjccm:especialidad ?oferta .

 ?subjccm:link ?masinformacion .

 ?subjccm:fecha_alta ?fecha

} ORDER BY ?fecha

Esta consulta, como se aprecia en la captura de pantalla, nos muestra la lista completa de ofertas, con el enlace para ampliar información, así como la fecha de publicación.

data URL	http://opendata.jccm.es/contenidos/portal/ccurl/1008/844/empleopublico_10_2012,0.rdf
----------	--

```

PREFIX jccm: <http://www.jccm.es/cs/Satellite/OpenData/RDF/schema/1212729415646#>
SELECT ?oferta ?masinformacion ?fecha
WHERE {
    ?subjccm:especialidad ?oferta .
    ?subjccm:link ?masinformacion .
    ?subjccm:fecha_alta ?fecha
}
ORDER BY ?fecha
    
```

oferta	masinformacion	fecha
Cocinero	http://www.jccm.es/empleo/cuerpo.php?par=3\$1\$10&tipo=51	2012-10-01 08:02:00
Ayudante cocina	http://www.jccm.es/empleo/cuerpo.php?par=4\$15\$10&tipo=51	2012-10-01 08:02:00
Personal de Servicios Domésticos	http://www.jccm.es/empleo/cuerpo.php?par=5\$3\$10&tipo=51	2012-10-01 08:02:00
GESTION ADMINISTRATIVA	http://www.jccm.es/empleo/cuerpo.php?par=2\$1\$10&tipo=50	2012-10-01 08:30:00
JURÍDICA	http://www.jccm.es/empleo/cuerpo.php?par=1\$3\$10&tipo=50	2012-10-01 08:49:00
ADMINISTRACIÓN GENERAL	http://www.jccm.es/empleo/cuerpo.php?par=1\$1\$10&tipo=50	2012-10-01 08:49:00
JURÍDICA	http://www.jccm.es/empleo/cuerpo.php?par=1\$3\$10&tipo=50	2012-10-01 08:50:00
GESTION ADMINISTRATIVA	http://www.jccm.es/empleo/cuerpo.php?par=2\$1\$10&tipo=50	2012-10-01 08:50:00
Administrativo	http://www.jccm.es/empleo/cuerpo.php?par=3\$1\$10&tipo=50	2012-10-01 08:50:00
ADMINISTRACIÓN GENERAL	http://www.jccm.es/empleo/cuerpo.php?par=1\$1\$10&tipo=50	2012-10-01 08:53:00
ADMINISTRACIÓN GENERAL	http://www.jccm.es/empleo/cuerpo.php?par=1\$1\$10&tipo=50	2012-10-01 08:54:00
ADMINISTRACIÓN GENERAL	http://www.jccm.es/empleo/cuerpo.php?par=1\$1\$10&tipo=50	2012-10-01 08:55:00
ADMINISTRACIÓN GENERAL	http://www.jccm.es/empleo/cuerpo.php?par=1\$1\$10&tipo=50	2012-10-01 08:56:00
JURÍDICA	http://www.jccm.es/empleo/cuerpo.php?par=1\$3\$10&tipo=50	2012-10-01 08:56:00
GESTION ADMINISTRATIVA	http://www.jccm.es/empleo/cuerpo.php?par=2\$1\$10&tipo=50	2012-10-01 08:57:00

Figura 6. Consulta de datos (opendata.jccm.es).

Estamos directamente reutilizando los datos publicados, de hecho estamos tratando los datos como si fuera una base de datos, explotándolos directamente con

SPARQL. Pero además, estos datos se pueden mezclar fácilmente con otros OpenData que están en otras webs.

REFLEXIONES FINALES.

En todo el mundo, los gobiernos y los ciudadanos se están dando cuenta de que el intercambio de conocimiento ayuda a hacer las cosas mejor, y están utilizando las plataformas de información para llevar a la gente a encontrar nuevas soluciones. El desafío es enorme. Para los gobiernos la apertura de datos y la nueva gobernanza proporcionan una oportunidad real para refundar nuestros espacios democráticos. Así, un gobierno abierto es la respuesta más adecuada para el impulso democrático a participar, a involucrarse, a ser parte, a comprometerse. La apertura hace que sea más fácil generar confianza, legitimidad y responsabilidad. Para las empresas, la apertura de datos, unida a otras técnicas como minería de datos, bigdata, movilidad de dispositivos móviles, sensorización urbana, etc, puede llegar una cantidad ingente de nuevas oportunidades que den lugar a nuevos modelos de negocio y nuevos modelos productivos que supongan un auténtico despegue de la actividad económica asociada.

Existen cuatro grandes retos que amenazan el futuro de la apertura de datos:

- Abrir los datos no es un mero trámite. Muchos gobiernos ven los procesos de apertura de datos como una cuestión estrictamente de cumplimiento de transparencia; es mucho más que eso. Las administraciones deben entender que son organizaciones que gestionan datos y que éstos deben estar gestionados con el mismo rigor que se gestionan otros bienes materiales.
- Armonización. Como decíamos al principio, actualmente hay muchos portales OpenData, y más que habrá. Sin embargo, el número de portales abiertos no es lo más importante, lo que realmente cuenta es que estos portales ofrezcan datos armonizados con el resto de portales. Dicho de otra forma, que no convirtamos cada portal en un silo independiente del resto. Hay que establecer acuerdos para abrir un mínimo conjunto de datos y en los mismos formatos. Se deben dejar de hacer mapas con los portales OpenData y empezar a hacer mapas indicando datasets armonizados entre diferentes

iniciativas de apertura de datos. Un camino puede ser armonizar estandarizando ontologías comunes.

- La inseguridad jurídica es otro de los obstáculos a la reutilización. La mayoría de los conjuntos de datos ofrecidos carecen de condiciones de licenciamiento claras. Unas veces por aparecer en el aviso legal un mensaje críptico al respecto y otras veces por ni tan siquiera aparecer nada.
- Ampliación del movimiento OpenData. A pesar de los logros conseguidos, necesitamos más masa crítica, más datos formalizados de acuerdo a los estándares (RDF, RDFa u OWL)

OpenData conviene por todas las razones que se han expuesto a lo largo de este artículo. Es una oportunidad social y económica de riqueza. Necesitamos una verdadera plataforma de OpenData en la Web para mirar al futuro, OpenData implica que se conecten los datos (Linked Data) a nivel semántico y por lo tanto es la semilla del nuevo paradigma de la Web, la Web Semántica. De manera que en gran medida, el futuro de la Web Semántica dependerá del éxito de OpenData.

REFERENCIAS

- ABELLA, Alberto. *Reutilización de información pública y privada en España*. [Fecha de la consulta: 13/10/2012] Disponible en <http://www.rooter.es/publicaciones>

- BERNERS-LEE, Tim. *Linked Data*. [Fecha de la consulta: 13/10/2012] Disponible en <http://www.w3.org/DesignIssues/LinkedData.html>

- *The Danish Government eGovernment Strategy 2011-2015. Good basic data for everyone. A driver for growth and efficiency*. 13/10/2012. <http://www.digst.dk/Home/ServiceMenu/English/Digitisation/~media/Files/English/BasicDataUKweb20121008.ashx>

- *Digital Government. Building a 21st Century Platform to Better Serve the American People*. [Fecha de la consulta: 13/10/2012] Disponible en

<http://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government.html>

- EAVES, David. *Eaves.ca*. [Fecha de la consulta: 13/10/2012] Disponible en <http://eaves.ca/>

- *EMT Movilidad 2.0*. [Fecha de la consulta: 13/10/2012] Disponible en <http://www.emtmadrid.es/movilidad20/opendata.aspx>

- *Fundación CTIC*. [Fecha de la consulta: 13/10/2012] Disponible en <http://www.fundacionctic.org/>

- GARRIGA, Marc. *Caldo Casero*. [Fecha de la consulta: 13/10/2012] Disponible en <http://www.caldocasero.es/>

- NIN, Andrés. *Apuntes electrónicos*. [Fecha de la consulta: 13/10/2012] Disponible en <http://apunteselectronicos.blogspot.com.es/>

- *The Obama's Administration commitment to Open Government: a status report*. [Fecha de la consulta: 13/10/2012] Disponible en <http://www.whitehouse.gov/open>

- *The Open Data Handbook*. [Fecha de la consulta: 13/10/2012] Disponible en <http://opendatahandbook.org>

- *Open Knowledge Foundation*. [Fecha de la consulta: 13/10/2012] Disponible en <http://okfn.org/opendata/>

- *Portal de Datos Abiertos JCCM*. [Fecha de la consulta: 13/10/2012] Disponible en <http://opendata.jccm.es>

- RAMÍREZ-ALUJAS, Álvaro. En *GIGAPP Estudios Working Papers*. ISSN 2174-9515. [Fecha de la consulta: 13/10/2012] Disponible en <http://www.gigapp.org/es/working-papers-gigapp>

- *Twinkle: a SPARQL Query Tool.* [Fecha de la consulta: 13/10/2012]

Disponible en <http://www.ldodds.com/projects/twinkle/>