



# ¿DATOS ABIERTOS? SÍ, PERO DE FORMA SOSTENIBLE



**Marc Garriga-Portolà**



**Marc Garriga-Portolà** es ingeniero en informática y licenciado en marketing. Experto (y activista) en procesos de apertura de datos públicos, es miembro del grupo impulsor del proyecto de datos abiertos del *Ayuntamiento de Barcelona: Open Data BCN*.

*Ayuntamiento de Barcelona. Tánger, 98. 08018 Barcelona*

<http://www.bcn.cat/opendata>

<http://mgarrigap.info>

[mgarrigap@gmail.com](mailto:mgarrigap@gmail.com)

## Resumen

Cada vez hay más gente que opina que estamos transformando el modelo de sociedad, que estamos cambiando de la sociedad industrial a la sociedad del conocimiento, una sociedad en la que los datos, la información, son su materia prima. En este contexto es donde se enmarcan las iniciativas de apertura de datos que, en muy poco tiempo, se están extendiendo por las administraciones públicas de todo el mundo. Se define qué es un proceso de apertura de datos, cuáles son los motivos que justifican su realización y se proporcionan unas pautas de cómo hacerlo, desde una óptica muy pragmática y de forma sostenible.

## Palabras clave

Apertura de datos, Gobierno abierto, Transparencia, Administración pública, Reutilización de la información, Reutilización de información del sector público, RISP.

**Title: Open data? Yes, but in a sustainable way**

## Abstract

More and more people think that we are transforming the model of society, changing from an industrial to a knowledge model, where data and information are its raw material. In this context many open data initiatives are being developed by public administrations around the world in a short time period. The open data process is defined, the reasons to carry out it are explained, and some guidelines on how to do it in a very pragmatic and sustainable manner are given.

## Keywords

Open data, Open government, Transparency, Public administration, Information reuse, Public sector information reuse, PSI reuse.

**Garriga-Portolà, Marc.** "¿Datos abiertos? Sí, pero de forma sostenible". *El profesional de la información*, 2011, mayo-junio, v. 20, n. 3, pp. 298-303.

DOI: 10.3145/epi.2011.may.08

## Introducción

El 20 de enero de 2009 Barack Hussein Obama fue investido como cuadragésimo cuarto presidente de los Estados Unidos de América; su discurso de investidura tuvo por título *Un nuevo nacimiento de la libertad* en conmemoración del segundo centenario del nacimiento de Abraham Lincoln.

Es significativo que, en su primer día en el cargo de presidente, lo primero que firmó Barack Obama fue un memorándum destinado a todas las agencias gubernamentales con el objetivo de fomentar la transparencia, la participación y la colaboración entre el gobierno y la ciudadanía.

Artículo recibido el 07-04-11

Aceptación definitiva: 10-04-11

La *Open government directive* marca las pautas que deben seguir las agencias gubernamentales americanas para conseguir más transparencia, participación y colaboración. Entre otros objetivos, obliga a abrir los datos públicos y a publicarlos en el portal *Data.gov* creado para potenciar la apertura de la información pública:

[http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda\\_2010/m10-06.pdf](http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf)

## Definición

Para acometer el ejercicio de sus servicios, las administraciones públicas disponen de muchos conjuntos de datos que,

en buena parte, son o pueden ser abiertos a la sociedad sin que implique ningún problema de privacidad.

Estos datos pueden ser reutilizados para otras finalidades, lo cual constituye el principal objetivo de los procesos de apertura de la información pública.

Se define la apertura de datos públicos –también conocido por las palabras inglesas *open data*– como el proceso que pone al alcance de la sociedad los datos públicos de los que dispone la Administración, en formatos digitales, estandarizados y abiertos, siguiendo una estructura clara que permita su comprensión y reutilización.

Gracias a los procesos *open data*, los ciudadanos, las empresas y, en general, cualquier organización pueden acceder cuando quieran y fácilmente a estos datos para informarse o para crear nuevos servicios, y aumentar el valor social y, si es el caso, también el valor comercial de esta información.

“ *Open data* es el proceso que pone al alcance de la sociedad los datos públicos de los que dispone la Administración, en formatos digitales, estandarizados y abiertos ”

### ¿Por qué es importante el *open data*?

Son varios los motivos que impulsan a las administraciones públicas a realizar un proceso de apertura de sus datos no confidenciales.

El primero es, si se me permite, de matiz ético: todo lo que se ha generado con dinero público debería poder ser accesible para toda la sociedad sin discriminación (excepto las lógicas restricciones legislativas); así de claro, así de sencillo.

También hay motivos de tinte social. Abrir los datos públicos es hacer universal el acceso a esta información ya que se dan las mismas oportunidades a todos de disponer de ella. Consecuentemente se evita la discriminación, se lucha contra la opacidad informativa y se erradican los grupos de poder basados en información reservada.

Por otro lado, la apertura de los datos económicos permite que los ciudadanos estén informados en todo momento de dónde invierten el dinero las administraciones públicas. Es un proceso claro de transparencia que facilita que la Administración pueda rendir cuentas a los ciudadanos de forma continua.

Dicho de otra forma, es una herramienta muy potente para luchar contra la corrupción y un primer paso para que la ciudadanía pueda participar de las decisiones, esté más implicada en el sector público y colabore en el diseño, prestación y evaluación de los servicios. Es el primer paso hacia un *gobierno abierto*.

De hecho, la apertura de los datos públicos también genera un impacto positivo en la economía, no sólo por los motivos ya presentados que, de por sí, ya repercuten directamente en la economía sino que, además, un proceso de apertura

de los datos públicos es una manera de generar riqueza, especialmente de aquellas empresas y profesionales de la información y sus tecnologías.

Desde un punto de vista interno de la Administración, poner datos al alcance de la sociedad le permite a ésta validar y controlar la calidad de los mismos, repercutiendo en una mejora de la eficiencia del sector público.

También desde el punto de vista de la Administración, un proceso de apertura de datos permite que la ciudadanía conozca más las organizaciones públicas, que perciba mejor cuáles son sus límites así como que se dé cuenta de que los servicios públicos no son gratuitos; dicho de otra forma, que conozca más al sector público redundando, a la larga, en una mejor percepción de los servicios públicos.

Teniendo en cuenta la actual situación de crisis económica, es importante destacar que mediante un proceso de apertura de datos públicos se incrementa el volumen de servicios públicos ya que las empresas, o en general cualquier organización, o incluso los ciudadanos, pueden utilizar la información abierta para desarrollar servicios que complementarán aquellos proporcionados por la Administración resultando en un mejor servicio al ciudadano (con independencia de quién preste los servicios).

### ¿Cómo lo hacemos? (visión teórica)

Una vez vistos los motivos que justifican la necesidad de abrir los datos públicos conviene indicar cómo realizar dicha apertura; hay que recordar que el objetivo último es la reutilización de la información pública.

“ Todo lo que se ha generado con dinero público debería poder ser accesible para toda la sociedad ”

Lo primero a tener en cuenta son los ocho principios de los datos públicos que definió el *Open Government Working Group* a finales de 2007:

- 1) Abrir todos los datos públicos\*.
- 2) Publicar los datos directamente desde la fuente, con un nivel de granularidad tan detallado como se pueda, sin agregarlos ni modificarlos, sólo *raw data*.
- 3) Actualizarlos en los períodos temporales adecuados para preservar su valor. Por ejemplo, no tiene sentido abrir la información del tráfico si ésta no es actualizada muy a menudo.
- 4) Facilitar su uso proporcionando herramientas de accesibilidad, catalogación, búsqueda y control de la calidad de la información para toda la sociedad. Además, hay que abrir los datos teniendo en cuenta la más amplia gama de propósitos fruto de su reutilización.
- 5) Facilitar la reutilización automática mediante el uso de formatos de archivo adecuados para este fin.
- 6) Proporcionarlos a cualquier persona, sin ninguna discriminación ni limitación.

7) Usar formatos no propietarios para que nadie tenga un control exclusivo.

8) Usar una licencia libre que fomente la reutilización, sin sujeción a derechos de autor que la limiten, o incluso la impidan\*.

\*Excepto en aquellos casos en que se permiten restricciones por razones de seguridad, de privacidad o que estén regulados por una ley específica o por un procedimiento administrativo.

<http://www.opengovdata.org/home/8principles>

**David Eaves**, uno de los expertos más reconocidos en el campo de la apertura de datos, resumió estos principios en sólo tres grandes leyes de los datos abiertos:

- Si no puede ser indexado, no existe.
- Si no está disponible en un formato abierto que pueda ser leído y procesado con éxito por un ordenador, no se motiva a (re)utilizar el dato.
- Si el marco legal no permite a alguien readaptarlo, no se motiva a utilizarlo.

<http://eaves.ca/2009/12/11/las-tres-leyes-de-los-datos-abiertos>

El objetivo último del *open data* es la reutilización de la información pública

Estas leyes persiguen un único objetivo: cómo abrir los datos para que puedan ser reutilizados.

La primera ley tiene en cuenta una realidad en internet: si *Google* y los otros motores de búsqueda no pueden encontrar una información entonces ésta *no existe*. Consecuentemente, si queremos que los datos sean reutilizados por la sociedad, necesitamos codificarlos convenientemente para que sean encontrados por los motores de búsqueda y lleguen a los usuarios finales.

La segunda ley dice que hay que ofrecer los datos en los formatos que faciliten su reutilización, por ejemplo su integración en un *mashup* con *Google Maps*, o en formatos estándar que permitan reaprovechar aplicaciones realizadas para datos abiertos de otras administraciones, etc. Estos formatos dependerán de las características de cada dato.

Se necesita un marco legal que permita compartir lo que se ha creado para motivar a la sociedad a ofrecer nuevos servicios con la información abierta o simplemente mostrar un hecho interesante que se deduzca de dicha información. Dicho de otra forma, los datos abiertos tienen que ofrecerse con unas condiciones de uso que fomenten al máximo su reutilización, aunque sea para fines comerciales.

Si se asigna una condición de uso restrictiva, como por ejemplo, el cobro de tasas, se limitará enormemente la reutilización de la información resultando en una clara pérdida en los beneficios de la apertura de los datos públicos.

### Cómo hacerlo de forma práctica

Estas leyes y principios son adecuadas... para un mundo ideal.

Para un mundo real hay que tener en cuenta otras cuestiones que modelarán nuestro proceso de apertura de datos públicos.

### Respaldo político

Lo primero que hay que hacer es convencer a los dirigentes políticos de la administración dónde se quieren abrir los datos; sin un claro *soporte político* es imposible seguir. Conviene aclarar que abrir los datos es, ante todo, un proceso político. Es cierto que también tiene una vertiente técnica importante, pero en esencia es un proceso político con consecuencias políticas.

Este soporte se plasmará en un acuerdo escrito que sirva para dar cobertura legal al proceso y como un elemento importante en la tarea de convencer a los responsables de los datos para que los abran a la sociedad.

### Inventario

Una vez tenemos el respaldo político debemos realizar un inventario de los datos públicos que existen especificando una meta-información que los enriquezca: su origen (en qué sistema de información residen), su descripción, su responsable, frecuencia de actualización, en qué formato o formatos se encuentran disponibles, etc.

Teniendo en cuenta este inventario, el presupuesto de que se dispone y el calendario previsto, se planificará un plan de trabajo que, además de detallar qué conjuntos de datos se quieren abrir (y en qué momento), también tendrá en cuenta un concepto importante: la sostenibilidad del servicio.

### Sostenibilidad en actualizaciones

Sostenibilidad en el sentido de que no sólo hay que abrir datos públicos de forma estática; la mayoría de los datos están "vivos", evolucionan con el tiempo, y consecuentemente es necesario habilitar procesos automáticos que actualicen los cambios en los datos con la frecuencia que sea necesaria (tal como hemos visto en el tercer principio del *Open Government Working Group*).

No hay que infravalorar estos procesos automáticos, pues entrañan una importante complejidad especialmente en su mantenimiento, debido a la heterogeneidad de las fuentes de datos y a las diferentes frecuencias de actualización.

No obstante, tales procesos automáticos son indispensables para poder ofrecer un servicio de calidad.

### Sostenibilidad presupuestaria

Y sostenibilidad teniendo en cuenta las fuertes restricciones presupuestarias que actualmente están sufriendo las administraciones públicas.

Esta situación conlleva que se ofrezcan, de entrada, aquellos conjuntos de datos que son más fáciles (y más baratos) de abrir; es decir, la facilidad de abrirlos está en función de cuán "bien ordenados" estén los datos en los sistemas de información.

Conviene aclarar que un proceso de apertura de datos no es caro. Habitualmente el problema proviene del hecho de que abrir los datos hace aflorar las carencias de la gestión interna de (algunos de) los datos públicos en las administraciones: generación y/o actualización de los datos de forma manual,

calidad de los datos y su evaluación, formatos en los que está disponible la información (generalmente formatos poco proclives para la reutilización), entre otras dificultades.

### Sostenibilidad en formatos

Sostenibilidad en cuanto a los formatos en que se ofrecen los datos. En este punto hay que encontrar un equilibrio entre proporcionar los datos de la mejor manera para poder ser reutilizados con facilidad (o sea, en formato RDF<sup>2</sup> o incluso en *linked data*) o en los formatos en los que ya están codificados actualmente, habitualmente propietarios como *Microsoft Excel*, entre otros.

Es oportuno recordar que lo que estamos ofreciendo son datos, no documentos; esta afirmación implica que no tiene sentido abrir archivos en formatos pdf, Word o similares. Un servicio *open data* no es un servicio *open access*, sino un sitio donde se ofrecen datos en bruto o en crudo (*raw data*) para poder ser reutilizados y crear otros servicios públicos. Un documento, tal como define **Alberto Ortiz**<sup>3</sup>, “es una colección de datos reunida de forma temporal para un fin determinado” y yo añadiría “donde se mezclan datos e información de presentación de dichos datos, información que no es interesante cuando el objetivo es la reutilización de los datos”.

Lo que estamos ofreciendo son datos, no documentos

Por este motivo los documentos no pueden ser ofrecidos como información en un portal *open data*.

Por otra parte, **Tim Berners-Lee** propuso una clasificación de los portales de datos abiertos en 5 niveles en función de cómo ofrecen la información y qué formato utilizan. Cuanto mayor sea el nivel más fácil será reutilizar la información de forma automática:

- 1) en cualquier formato;
- 2) en formatos estructurados;
- 3) en formatos estructurados no propietarios;
- 4) en formato RDF;
- 5) enlazados (*linked data*).

<http://inkdroid.org/journal/2010/06/04/the-5-stars-of-open-linked-data>

Ésta será una de las decisiones que habrá que tomar en el momento del diseño de una iniciativa de apertura de datos: a qué nivel queremos llegar.

La decisión no es banal ya que ofrecer información en los niveles altos (4 y 5) implica dedicar esfuerzos en la transformación de los datos a formatos RDF o *linked data*, en lugar



Iniciativas *open data* en España, <http://datos.fundacionctic.org/sandbox/catalog/faceted>

de destinarlos a abrir más conjuntos de datos (en formatos de menor calidad, desde un punto de vista de la reutilización).

En este sentido conviene destacar que actualmente España es una potencia mundial en la apertura de datos públicos no sólo por la cantidad de iniciativas en este sentido sino también por su calidad.

Según la *Fundación CTIC (Centro Tecnológico de la Información y la Comunicación)* a principios de mayo de 2011 en España había 12 servicios de apertura de datos en funcionamiento de nivel 2 o superior (el primer nivel de la clasificación que propuso **Tim Berners-Lee** es tan básico que personalmente no lo considero relevante).

La mayoría de estas iniciativas españolas están clasificadas en los mejores peldaños: 5 son de categoría 4 y 2 están en el máximo nivel. Es oportuno tener en cuenta que en todo el mundo sólo hay 7 servicios de apertura de datos con un nivel 4 (de los cuales 5 son españoles), y 4 de la máxima distinción (2 de España).

Un servicio de datos abiertos no sirve de nada si no se mantiene constantemente actualizado

### Sostenibilidad en las condiciones de uso

Es muy importante decidir qué condiciones de uso tienen los datos abiertos dado que éste es el elemento clave que permitirá, o no, la reutilización de la información, que es el objetivo principal.

Lo ideal es no poner ninguna barrera. Esencialmente esto implica no cobrar por el uso de los datos abiertos ni mermar la posible ganancia del servicio resultante, y no poner limitaciones a la creatividad a la hora de diseñar servicios. Una buena manera es utilizar una licencia *Creative Com-*

mons CC-BY que sólo obliga a especificar la autoría de los datos, de dónde proceden, no pone ninguna restricción adicional. Además, utilizar este tipo de licencias nos ahorra las “tradicionales” que conllevan que primero tenga que realizarse una petición, que ésta sea procesada y que se tengan que gestionar y controlar las licencias en uso.

Hay que tener en cuenta que algunas administraciones públicas obtienen ingresos de determinada información, y es evidente que será difícil abrir dicha información con unas condiciones de uso que no conlleven alguna restricción económica. Hay que dejar bien claro que éstos son casos muy concretos y, en general, de uso específico.

Otro caso de restricciones en el uso procede de aquellos conjuntos de datos que tiene la administración pública pero que, por el motivo que sea, no son de su propiedad. Por tanto la administración no puede proporcionar derechos a terceros sobre ellos. Este

es un caso también muy concreto donde habitualmente se licencian los datos con unas condiciones de uso que permiten su consulta (dado que son públicos) pero no crear productos a partir de ellos.

Es evidente que no poder crear servicios implica eliminar el principal motivo por el que se abren los datos, sin embargo, en casos muy específicos puede tener sentido.

La principal restricción en la reutilización de información abierta puede venir de la tentación que tengan las administraciones públicas si alguno de los servicios creados a partir de dicha información tiene un importante éxito. Es muy probable que quieran compartir el éxito y, obviamente, las ganancias económicas obtenidas. No obstante, aunque pueda parecer sencillo, no es fácil cobrar por el uso de datos públicos, y se puede llegar a la paradoja de tener una gestión de cobro más cara que la cantidad que se cobra.

También es probable que quieran compartir los gastos del servicio, aunque, como decíamos anteriormente, no acostumbra a salir caros. Lo que sí puede resultar caro es solucionar las carencias en la generación y gestión de los datos públicos.

“ A más información abierta más potencia tendrán las consecuencias (positivas) del servicio ”

### Sostenibilidad en el fomento de la reutilización

Una buena manera de espolear el uso del servicio de datos abiertos es un concurso de aplicaciones/servicios.

En 2007 la ciudad de Washington decidió poner al alcance de los ciudadanos información pública de manera libre

The screenshot shows the homepage of 'Apps for Democracy Community Edition'. At the top, there's a header with the title and logos, including the District of Columbia Office of the Chief Technology Officer. Below the header, there are two call-to-action buttons: a red one saying 'Got what it takes to mash-up DC's data to win some cash and tons of street cred?' and a blue one saying 'Then sign-up for Apps for Democracy and start coding here!'. A prominent yellow banner reads 'Interested in launching your own Apps for Democracy inspired initiative? You're in luck! We've prepared a detailed guide showing you how to create one. Download the PDF here.' Below this, there's a 'What:' section with text about the first edition's success and a link to 'Submit your ideas and problems here!'. A 'Ways to Engage' section lists four steps: 'Add Your Neighborhood's Insights & Vote', 'Signup to Build Apps for Democracy', 'Tweet this!', and 'Share!'. The URL 'http://www.appsfordemocracy.org' is at the bottom.

<http://www.appsfordemocracy.org>

y gratuita. Fue una de las primeras administraciones en dar este paso y se encontraron que tenía un uso muy bajo. Por este motivo decidieron realizar el concurso *Apps For Democracy* para fomentar la creación de servicios basados en estos datos públicos. Según informan en su web, los resultados fueron espectaculares: -"Nuestra solución fue convocar el concurso *Apps for Democracy*, que costó 50.000 US\$ a Washington DC pero creó 47 [...] aplicaciones web con un valor estimado de más de 2.600.000 US\$".

Se demostró que si se fomenta adecuadamente el uso de portales *open data*, el resultado es rentable para la Administración, no sólo a nivel económico sino también a nivel social.

<http://www.appsfordemocracy.org>

### Resumen

Recapitulando, los pasos a seguir para abrir los datos públicos de una administración son los siguientes:

- Apoyo político plasmado en un acuerdo del más alto nivel posible.
- Realización de un inventario de los datos disponibles.
- Decidir qué datos se quieren abrir, en función de cómo estén gestionados internamente, del presupuesto que se disponga, de cuestiones políticas así como de tener muy en cuenta la sostenibilidad del servicio.
- Decidir qué nivel de calidad queremos que tenga nuestro portal de datos abiertos: un nivel más alto conllevará más tareas para “cocinar” los datos antes de ofrecerlos a la sociedad, pero ésta lo tendrá mucho más fácil para poder reutilizarlos.
- Abrir los conjuntos de datos (ponerlos a disposición de la ciudadanía) y automatizar la actualización de esta información.

# iralis®

International Registry for Authors:  
Links to Identify Scientists

¡Ni un día más sin decidir tu firma!

Inscríbete en <http://iralis.org>

es:

- una **guía** para los autores hispanos para que firmen sus trabajos en el formato internacional usual
- una **base de datos** que registra las variantes de firma usadas por cada autor en diferentes épocas
- un **buscador** que usa automáticamente todas las variantes registradas

***El formato de  
firma internacional***

Con el apoyo de:

El profesional de la

**información**



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE CIENCIA  
E INNOVACIÓN



**FECYT**

FUNDACIÓN ESPAÑOLA PARA LA CIENCIA Y LA TECNOLOGÍA

**CIEPI**

CENTRO INTERNACIONAL PARA LA INVESTIGACIÓN  
EN ESTRATEGIA Y PROSPECTIVA DE LA INFORMACIÓN