

F. 2. Open data, big data: ¿hacia dónde nos dirigimos?

Antonia Ferrer-Sapena y Enrique A. Sánchez-Pérez

1 febrero 2012

Ferrer-Sapena, Antonia; Sánchez-Pérez, Enrique A. (2013). "Open data, big data: ¿hacia dónde nos dirigimos?". *Anuario ThinkEPI*, v. 7, pp. 150-156.



Resumen: En unos momentos en que se está popularizando el movimiento de "datos abiertos" en distintos contextos: administración pública, ciencia y empresas, es importante conocer cuál es el alcance que tiene. El término ha generado muchas expectativas relacionadas con la transparencia a todos los niveles, la reutilización y el libre acceso. Aquí se hace una revisión de alguna de las expectativas generadas por este movimiento.

Palabras clave: Open data, Datos abiertos, Transparencia, Administración pública, Gobierno.

Title: Open data, big data: where do we go?

Abstract: The "open data" movement is becoming more popular in different contexts: public administration, science and companies; therefore, it is important to know its scope. The term has generated a lot of expectations related to transparency at all levels, data reuse and free access. This review discusses some of the expectations generated by this movement.

Keywords: Open data, Transparency, Public administration, Government.

Está de moda todo aquello que lleve la palabra "open", la palabra "data", o ambas. Recurrentemente encontramos los términos *big data*, *open data*, *data mining*... Ya no se trata sólo de compartir información o de permitir el libre acceso a las publicaciones, ahora queremos que sean accesibles también los datos, de manera libre y gratuita. Los datos nos invaden en todos los contextos de la sociedad, generados por las aplicaciones de la web 2.0, los gobiernos, las investigaciones científicas, las publicaciones o las organizaciones.

Todos estos datos se encuentran cada vez más ligados a los movimientos "open". Aunque quizás convendría aclarar las diferencias, implicaciones y relaciones que pueden existir entre *big data*, *open data*, *data mining*, *open science*, aquí únicamente abordaremos los dos primeros.

Se entiende como *open data* la apertura (o puesta a disposición de cualquiera) de datos digitales, pero en este concepto también tiene cabida la de datos en documentos impresos. Hay una enorme cantidad de datos depositados en archivos y bibliotecas que se han construido gracias a los impuestos de los ciudadanos que están sin digitalizar. La promoción del acceso, por tanto,

está vinculada a la transparencia de los procesos. En función de quien sea el productor de los datos se hablará de transparencia en el gobierno, en la investigación o en las empresas.

"Open data son todos los datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona (OKF, 2012)"

En el caso de la Administración la apertura suele ir unida a la creación de servicios de valor añadido, ya sea por parte de la propia Administración o de empresas que quieran crear nuevas utilidades basadas en los datos. Pero el concepto *open data* va más allá de la mera reutilización. Su filosofía pretende que los datos se encuentren disponibles de manera libre para todo el mundo, sin restricciones de *copyright*. Actualmente este tipo de datos abiertos son, mayoritariamente geográficos, estadísticos, genómicos, médicos y biológicos.

Según el manual de la *Open Knowledge Foundation* (2012), *open data* son los datos que pueden ser utilizados libremente, reutilizados y redistribuidos por cualquier persona, permitiendo la creación de obras derivadas que reconozcan la misma licencia. Sus características fundamentales son:

- Disponibilidad y acceso. Los datos deben encontrarse disponibles, preferiblemente mediante la descarga a través de internet, y permitir su modificación. Su reproducción debe ser a un coste razonable.

- Reutilización y redistribución en condiciones óptimas pudiendo combinarse con otros conjuntos de datos.

- Participación universal: todos deben ser capaces de utilizarlos, reutilizarlos y redistribuirlos. No deben existir restricciones para su comercialización.

- En el contexto empresarial *big data* se refiere a grandes conjuntos de datos que tienen algún valor para las empresas. En algunos casos han sido generados por los científicos, aunque los más conocidos proceden de movimientos de *Facebook*, *Twitter* y *Google*. El tratamiento suele realizarse en tiempo real utilizando tecnologías cada vez más potentes para la gestión de bases de datos, la visualización y el análisis. En muchas ocasiones los datos se almacenan en la nube y son manipulados desde allí.

Es difícil prever hasta dónde pueden llegar las operaciones *big data*. Algunos ejemplos llamativos son el seguimiento de la campaña de marketing de la película *Amanecer 2* y la campaña del candidato Barak Obama en las elecciones presidenciales de EUA de 2012:

1) La película *Amanecer 2* es un ejemplo de cómo interpretar los sentimientos a través del análisis de datos, analizando los tweets generados en la promoción de la película. La campaña de marketing fue a nivel mundial y estaba basada en las redes sociales. Los movimientos en *Twitter* se analizaron con *Social sentiment index* de *IBM*. <http://www.ibm.com/analytics/us/en/conversations/social-sentiment.html>

Esta herramienta permite prever los comportamientos del mercado mediante el análisis de grandes cantidades de datos aportados por las redes sociales. En el caso de *Amanecer 2*, se estudiaron los tweets publicados once días antes de su estreno en EUA. Los compararon con los de las películas competidoras en ese momento para determinar los sentimientos que cada una de ellas generaba y así poder predecir el comportamiento



<http://open-data.europa.eu/es>

en taquilla. El estudio de estos tweets dictaminó que el sentimiento positivo generado en el pre lanzamiento (un 90%), pasó a ser de un 75% el día del estreno. El motivo del cambio no fue por las críticas negativas, sino por los lamentos por finalizar la saga (*IBM*, 2012; *Pico*, 2012a).

2) En las pasadas elecciones americanas los dos partidos políticos rastrearon a sus potenciales votantes en la Red a través del cruce de varias bases de datos, por ejemplo, la derivada de la venta de tazas publicitarias. Así, los *big data* han contribuido a predecir quién iba a ser el vencedor y han ayudado a Obama a conseguir el éxito, facilitándole sobre todo la búsqueda de fondos para su campaña (*Ticbeat*, 2012).

Todos estos datos en bruto suelen ser difíciles de entender, de ahí que se hayan elaborado distintas herramientas de visualización que facilitan su interpretación. Quizá una de las empresas más potentes en España para realizar proyectos de visualización y análisis de datos sea *Quadrigram*. <http://www.quadrigram.com>

Una recopilación de herramientas gratuitas de visualización de datos se puede encontrar en los blogs *DoctorMetrics* y *Periodismo ciudadano*, en los que se hace una breve descripción de sus utilidades. En estos blogs se hace referencia a las siguientes herramientas: *Tableau (Public)*, *Weave*,

Gephi, Many eyes, NodeXL, Data-Driven Documents (d3), Axiis, Google fusion tables, Wordle, Dipity, Batchgeo, Twitter widgets, Google forms, Trendistic.

<http://www.doctormetrics.com/2012/08/29/herramientas-visualizacion-datos>

<http://www.periodismociudadano.com/2011/02/25/herramientas-para-el-periodismo-ciudadano-visualizacion-de-datos>

“La empresa tiene claras las ventajas que produce el *big data*, pero no las que puede proporcionar el *open data*”

Otra herramienta interesante es la que ha implementado *LinkedIn* para facilitar la visualización de las redes de colaboración a través del mapeo de contactos, o la aplicación gratuita de tratamiento de datos *Google refine*.

<http://inmaps.linkedinlabs.com/network>

<http://code.google.com/pl/google-refine>

En el contexto empresarial, ¿qué es prioritario, *open data* o *big data*?

Por supuesto, ambas tecnologías se pueden implementar al mismo tiempo, pero puestos a priorizar se puede seguir un estudio de *Gartner* (**Bernal**, 2012a), según el cual exponer los datos en abierto proporcionará más beneficios a las empresas que el *big data*. La empresa tiene claras

las ventajas que produce el *big data*, pero no las que puede proporcionar el *open data* todavía. El estudio señala que la transparencia en las organizaciones será vital para acercarse a los clientes, y esto será posible con una buena estrategia para compartir y controlar los datos que se ofrezcan en abierto. Este proceso de apertura permitirá a las empresas conquistar a los clientes y generar más beneficios. La empresa debe tener una estrategia clara en este sentido. La clave se encuentra en conocer con quién se debe compartir los datos para conseguir el mayor beneficio. Los datos hacen las organizaciones sean más inteligentes, pero abiertos las hacen más ricas (**Newman**, 2012). Según *Gartner*, la apertura de datos es el primer paso para conseguir la monetarización de la información empresarial.

Desmitificando el *open data*

Entre las ventajas del *open data* se han dado por sobreentendidos determinados supuestos que no son verdaderos. Una visión clara acerca de las posibilidades de este movimiento se ofrece en una reciente selección de opiniones recopiladas en la web del *Open data seminar* (**Farrell**, 2012a), que recoge varias reflexiones de interés:

En relación con la reutilización

Combinar distintos tipos de datos produce siempre un valor superior. Por ejemplo, los datos del horario de autobuses con un mapa. Según señalan **Scassa y Campbell** (2009), la combinación de elementos puede plantear problemas de protección de datos ya que la información recopilada para unos fines específicos no debe ser utilizada para otros sin consentimiento. Lo ilustra con un ejemplo de la ciudad de Ottawa, donde se encuentran relacionados los crímenes y las llamadas a la policía a través de un mapa. Esta información podría ser utilizada por las compañías de seguros y hacer corresponder las tasas de seguro con los índices de criminalidad, o utilizarlos para campañas de comercialización en función del número de robos. Según **Campbell (Farrell, 2012a)** esta utilización viola las circunstancias por las que se recopilaron estos datos. Entonces, ¿hasta qué punto existe un conflicto entre las condiciones del mercado y las iniciativas ciudadanas de datos abiertos?

Jo Bates (2012), más crítica, considera que hay que analizar hasta qué punto estas iniciativas

The screenshot shows the GreatSchools website interface. At the top, there's a navigation bar with 'School Officials | Join Us | Sign In' and social media icons. Below that is a search bar for schools by district or city. The main content area is divided into sections: 'Find a School' with tabs for different school levels (Preschool, Elementary, Middle, High, College), 'Hot Topics' with various educational categories, and a 'Start here: Find a school' section with a search form. Below the search form is an advertisement for 'OXFORD ROYALE SUMMER SCHOOL' and a 'Most Popular' list of articles.

<http://greatschools.org>

no deben ser utilizadas para la privatización y la desregulación en un intento de crear nuevos mercados de información en el sector público en lugar de prestar servicios públicos. **Bates** pone como ejemplo *Palantir Technologies*, empresa líder en el espacio de datos abiertos y que tiene a *In-Q-Tel* como segunda inversora. *In-Q-Tel* es el nombre de un fondo de inversiones creado originalmente por la *CIA* con el fin de favorecer la formación de empresas que puedan tener potencial como colaboradoras para los servicios de inteligencia.

A la pregunta de si hay alguna manera de controlar los datos abiertos en la economía de libre mercado, **Campbell** esboza un paralelismo entre la economía de los bienes de información y la de bienes culturales. Reflexiona sobre cómo unas políticas gubernamentales que apoyen a los productores nacionales pueden hacer frente a las grandes empresas que están formando oligopolios alrededor de los datos abiertos. Por desgracia, el movimiento *open data* exige que los datos se proporcionen sin ningún tipo de restricción, exigiendo la no discriminación en la concesión de licencias sean ya ricos o pobres los que las utilicen.

En España la legislación marca que para poder reutilizar la información del sector público se debe solicitar al órgano competente. En la solicitud deben identificarse los documentos reutilizables y especificar los fines, comerciales o no comerciales, de la reutilización. La resolución para la reutilización debe ser expresa, en caso contrario, se entiende desestimada (*Real decreto 1495/2011*, de 24 de octubre), por lo que el contexto es semejante al señalado por **Campbell**.

En relación con la aplicación del modelo científico a los gobiernos

Otra de las aportaciones compara el método de trabajo de los científicos con el de los políticos. En el campo científico, determinados métodos están comenzando a ser anticuados. Con el fin de hacer afirmaciones verificables, los científicos deben dar acceso a más información, por ejemplo a las medidas de cálculo que se tomaron para generar los resultados. La pregunta que nos deberíamos hacer es si este método puede ser trasladado al contexto político y así comprender qué motivó la toma de decisiones en el gobierno. Si en el mundo científico la procedencia del conocimiento y la justificación son vitales, en el contexto político puede implicar aportar los razonamientos sobre por qué se han tomado determinadas decisiones e incluir las pruebas que se han utilizado para apoyarlas. Esta es una visión idealista de lo que pueden significar los datos abiertos: cómo la práctica científica puede proporcionar un marco útil para ayudar a los principios del gobierno y ser una guía de la comunicación del mismo y promover la transparencia.

En relación con la democracia cognitiva

Johnson (2012) argumenta que las plataformas de datos abiertos pueden atraer a personas interesadas en la solución de problemas complejos. La participación en la resolución de problemas puede provenir de ciudadanos voluntarios que estén dispuestos a trabajar. Los beneficios, argumenta, son simples:

a) Dentro del gobierno se obtendrán mejores ideas y una comprensión más clara de los problemas si se centran en ellos más personas, y más si éstas dedican parte de su tiempo libre a resolverlos. Este es un principio básico ya anunciado por **Henry Farrell** (*George Washington University*) y **Cosma Rohilla-Shalizi** (*Carnegie Mellon University/The Santa Fe Institute*) sobre la democracia cognitiva, quienes señalan cómo se encuentran mejores soluciones cuando se amplía la densidad y diversidad de las mentes dedicadas a la resolución de problemas y se cuenta con las herramientas adecuadas.

b) Si se amplía el número de voluntarios que tratan de mejorar el Gobierno, más allá del hecho de ir a votar, se diluye la línea que separa a los políticos-burócratas de los ciudadanos de a pie. A medida que se amplía el número de voluntarios, se ensancha y atrae a más personas que tienen talento para compartir. El resultado final es el aumento del compromiso, de la participación ciudadana y una mayor conciencia acerca de los servicios que proporcionan los estados y los desafíos a los que se enfrentan.

“La organización que recopila y mantiene la información no siempre se encuentra en oposición de utilizarla”

En relación con la prensa

La apertura de datos plantea una nueva manera de proporcionar información a los periódicos y ha revolucionado el periodismo. Hasta el momento los periodistas de investigación eran los que desenterraban la corrupción; al abrir los datos, la información se encuentra más accesible y cualquier ciudadano puede investigar, con lo que el número de investigaciones puede ser mayor. **Matthew Yglesias** (2012) señala que esto ha revolucionado el modelo de negocio de los periódicos tradicionales y ha hecho que los periodistas incrementen su productividad. En EUA cuentan con distintas herramientas para el periodismo económico. Por ejemplo, la base de datos *Fred* (compilada por el *Federal Reserve Bank of St.*

Louis), asociada a Alfred (*Archival federal reserve economic data*) permite comparar series de datos económicos.

<http://research.stlouisfed.org/fred2>

<http://alfred.stlouisfed.org>

Ese tipo de bancos de datos se utilizan normalmente para los escritos políticos de **Joe Weisenthal** en *Business insider*, o en el blog de **Paul Krugman** en el *New York times*.

En este contexto se ve cómo el costo marginal de la transmisión de información se encuentra cada vez más cerca de cero. Así se producen dos fenómenos paralelos:

- cada vez es más difícil internalizar el valor de producción de la información debido a que los hechos una vez publicados tienden a extenderse más allá del control del productor;

- la información se vuelve socialmente más valiosa.

Expectativas generadas por el *open data*

Tras estas visiones relativamente optimistas del movimiento *open data*, **Clay Shirky** (2012), profesor de la *University of New York* y experto en redes sociales, aclara los puntos que hay que conocer para no decepcionarnos por las expectativas generadas. Para este autor, sirve para mejorar el servicio, pero no para acabar con la corrupción. Señala que debemos ser conscientes de que:

a) No es un movimiento político o cultural. En sus fundamentos este movimiento no tiene objetivos políticos. Trata de mejorar los servicios que se ofrecen al ciudadano.

b) No se está haciendo nada por la transparencia y la rendición de cuentas del gobierno.

c) No existe una clara diferenciación entre lo que es sin ánimo de lucro y lo que no.

d) Aquellos que son buenos con los datos no son brillantes a la hora de contar historias. Pero las personas no consumen hechos, consumen historias. Quienes comprenden la importancia de los datos, por lo general, son las personas más entusiastas acerca de su interpretación y, como resultado, sistemáticamente sobrestiman el interés ciudadano por los datos. Un gráfico sobre el cambio climático no es suficiente para que todos los ciudadanos lo entiendan.

e) Las instituciones que cuentan buenas historias no siempre tienen capacidad para el análisis de los datos. A las agencias de noticias se les paga por contar historias que son verdad, pero no siempre están informadas o tienen la capacidad suficiente para la interpretación de los datos.

f) La transparencia en ocasiones es sólo transparencia. Existen muchos datos disponibles con los que crear una ilusión de apertura de datos, pero cuya utilidad resulta escasa y baja calidad, frenando la exposición de los datos significativos.

En relación con la transparencia

Aaron Swartz (2012) indicaba que encontramos distintos tipos de personas que se entremezclan en este contexto:

- los curiosos que quieren conocer dónde van a parar sus impuestos;

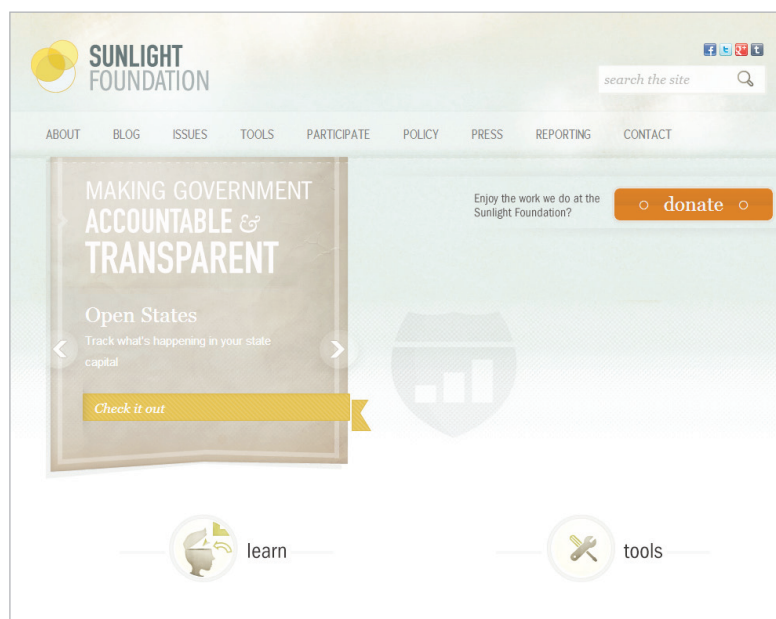
- los activistas que creen que al revelar los datos sobre lo que hace el Gobierno en realidad pueden detener la corrupción; y

- las empresas que pueden aprovecharlo para generar aplicaciones de valor añadido.

Considera que nos debemos preguntar si realmente el *open data* pretende alertarnos de las infracciones cuando todo el mundo intenta ocultarlas. Considera que las organizaciones creadas específicamente para perseguir la transparencia gubernamental serán las únicas capaces de conseguirlo, tal y como hace por ejemplo la *Sunlight Foundation*.

<http://sunlightfoundation.com>

En este contexto también se presentan dos visiones antagónicas pero a la vez complementarias: las optimistas acerca de los beneficios sociales de la tecnología y la parte política donde se encuentra una lucha entre facciones con diferentes necesidades e intereses. Así, **Henry Farrell** (2012) sostiene que la democracia tiene que reducir al mínimo las disparidades de poder y aprovechar la diversidad cognitiva. En este contexto los datos



<http://sunlightfoundation.com>

abiertos pueden aportar tres ventajas:

1. La principal ventaja de los grupos de presión es a menudo el control selectivo de la información más que el de la financiación. Hacer que los datos relevantes se encuentren disponibles puede ser aprovechado por los infrarrepresentados.

2. Los datos abiertos de calidad pueden ayudar a limitar la tendencia de los partidos para crear su propia información y su interpretación.

3. La apertura de algunos datos puede acercarnos a lo que se denominan procesos de cognición colectiva.

“Los propietarios de datos han de liberar los datos a pesar de los desafíos políticos y técnicos”

Innovación abierta y democracia

Beth Noveck (2012) ve los datos abiertos como fundamento de una democracia compleja y como fuente y base de innovación en el gobierno:

- conllevan un gobierno más transparente y responsable;
- son más significativos para el rediseño de las instituciones que para hacer visibles malas prácticas;
- proporcionan la base para el desarrollo de soluciones en conjunto;
- posibilitan que participen terceros y es lo que hace que éstos tengan una opción transformadora.

La organización que recopila y mantiene la información no siempre se encuentra en posición de utilizarla. Cuando los datos abiertos son procesables por máquinas, es posible crear herramientas, visualizaciones, modelos y análisis que pueden hacer más valiosa la información.

Para resolver problemas complejos es necesario que trabajen de manera conjunta personas con distintas habilidades. Actualmente los gobiernos no suelen aprovechar el conocimiento colectivo para mejorar la gobernabilidad y resolver problemas. En este contexto los *hackathons*¹ se convierten en un modelo para la acción, para la gobernanza participativa. Avanzar hacia la innovación abierta como una forma predeterminada de trabajar en el Gobierno no es fácil. Se necesita persuadir a los propietarios de datos para que transformen el papel al formato digital, y luego liberar los datos a pesar de los desafíos políticos y técnicos.

La innovación participativa es mucho más que la simple publicación de datos conjuntos. Las instituciones tienen que estar dispuestas a colaborar con los innovadores, con los activistas de la

sociedad civil. Actualmente son probablemente las grandes empresas las únicas entidades con los recursos suficientes para hacer algo con la información en bruto. Si los datos se encuentran en formato abierto, los innovadores pueden crear nuevas visualizaciones para ayudar a los consumidores a entender los costos y tomar decisiones informadas. La web *GreatSchools.org* ayuda a los padres a elegir entre las escuelas públicas recopilando los datos de rendimiento escolar, en cooperación con el *Departamento de Educación* de EUA.

En definitiva, tal y como expresa **Tom Lee** (2012), pensamos en los datos abiertos con tendencia a un excesivo optimismo, pero realmente aportan beneficios a la democracia.

Conclusiones

Hasta ahora trabajar con datos era bastante difícil por el coste que suponía trasladarlos a soportes informáticos. Actualmente las estadísticas y los estudios econométricos están llegando de manera masiva al mundo *open data*; ejemplo de ello son los datos abiertos del *Banco Mundial* y de los distintos gobiernos. De hecho, en diciembre de 2012 se lanzó la versión beta del portal de datos abiertos de la *Comisión Europea*: <http://open-data.europa.eu/es>

Los distintos softwares en abierto permiten realizar análisis cuantitativos de manera fácil desde cualquier ordenador. Ello está contribuyendo a que el número de trabajos se incremente de manera considerable. Sus ventajas son cuantiosas: eliminan redundancias, evitan tener que normalizar datos, permiten fusionarlos y localizarlos. Si se comparte todo este trabajo habrá más tiempo para la investigación.

Los datos abiertos aseguran una asignación más eficiente de los escasos recursos con los que contamos actualmente. Hasta ahora han sido tratados, cuidados y protegido su acceso ya que ellos han sido y son la base para nuestras investigaciones. Muchas veces no les sacamos más partido por no tener la suficiente capacidad o tiempo, desaprovechando los conocimientos de otros.

Esta apertura de datos también puede mejorar la transparencia, no sólo en las políticas gubernamentales, sino también en las investigaciones, fomentando así el rigor científico.

Junto a las ventajas señaladas y las prevenciones acerca de lo que se puede esperar actualmente del *open data*, hay que señalar las reflexiones de **Fritjof Capra** y **Evgeny Morozov**.

Fritjof Capra (2003) señala cómo las redes cambian la manera de entender la sociedad actual, además de que en nuestro escenario proliferan y crecen los medios sociales de información y de redes. Cada vez existe un mayor número de

tecnologías para la vida diaria y todo ello contribuye a que haya un volumen de datos y de conexiones entre personas, redes e informaciones con un crecimiento cuyos límites aún no se vislumbran. Este crecimiento y la complejidad que van generando las relaciones que se establecen son un verdadero desafío.

“La apertura de datos puede mejorar la transparencia tanto en las políticas gubernamentales como en la investigación, fomentando así el rigor científico”

Por su parte **Evgeny Morozov** ofrece una visión crítica de la Red, considerándola un elemento favorecedor de las dictaduras totalitarias capaces de controlar la vida de los ciudadanos y monitorizar su vida diaria. La conferencia en la que participó en Madrid ha sido comentada ya por nuestros compañeros del grupo *Durga* en su blog (*Big data*, s. f.). En ella abordaron distintos aspectos de cómo la Red afecta a nuestras vidas, en especial *Google* y las redes sociales, sin olvidar el creciente control que supone la era de los *big data* y las *smart cities*.

Estamos pues en un contexto donde en función de las utilidades que les den o los datos que se nos ofrezcan pueden ser elementos que favorezcan o no a los ciudadanos.

Nota

1. Un *hackathon* es un encuentro de programadores para desarrollar software.

Bibliografía

Bates, Jo (2012). *Open data movement redux: tribes and contradictions*. <http://goo.gl/Wqbbk>

Bernal, Natasha (2012). “Open data, mejor que *big data* para generar beneficios a empresas”. *TICbeat*, 24 agosto. <http://innovacion.ticbeat.com/open-data-mejor-big-data>

Capra, Fritjof (2003). *Las conexiones ocultas*. Anagrama. ISBN: 978 84 339 6188 4

Farrell, Henry (2012a). “Open data seminar”. *Crooked timber*, 17 July. <http://crookedtimber.org/2012/07/17/open-data-seminar-2>

Farrell, Henry (2012b). “Trish, Reiner and the politics of open data”. *Crooked timber*, 4 July. <http://crookedtimber.org/2012/07/04/trish-reiner-and-the-politics-of-open-data>

Federal Reserve Bank of St. Louis. Economic research - St. Louis Fed. <http://research.stlouisfed.org>

IBM (2012). “IBM analytics – conversations: social sentiment index – United States”. *IBM.com*. <http://www.ibm.com/analytics/us/en/conversations/social-sentiment.html>

Johnson, Steven B. (2012). “Searching for John Snows”. *Crooked timber*, 27 June. <http://crookedtimber.org/2012/06/27/searching-for-john-snows>

Lee, Tom (2012). “Open data: better politics, winning politics... but still politics”. *Crooked timber*, 6 July. <http://crookedtimber.org/2012/07/06/open-data-better-politics-winning-politics-but-still-politics>

Morozov, Evgeny (2012). *El desengaño de internet. Los mitos de la libertad en la red*. Destino. ISBN: 978 84 233 2779 9

Newman, David (2012). “Open for business: learn to profit by open data”. *Gartner*. <http://www.gartner.com/id=1947015>

Noveck, Beth (2012). “Open data – The democratic imperative”. *Crooked timber*, 5 July. <http://crookedtimber.org/2012/07/05/open-data-the-democratic-imperative>

Open knowledge foundation (2012). *The open data handbook*. Open Knowledge Foundation. <http://opendatahandbook.org>

Pico, Raquel C. (2012). “Análisis de datos para predecir el éxito (o el fracaso) de *Amanecer 2*”. *TICbeat*, 30 nov. <http://innovacion.ticbeat.com/analisis-datos-predecir-exito-fracaso-amanecer-2>

Scassa, Teresa; Campbell, Lisa M. (2009). “Data protection, privacy and spatial data”. En: Goodchild, Helen (ed.). *Spatial data quality*. CRC Press, pp. 211-219. ISBN: 978 1 4398 1012 5 <http://www.crcnetbase.com/doi/abs/10.1201/b10305-24>

Shirky, Clay (2012). “Cooperation and corruption”. *Crooked timber*, 2 July. <http://crookedtimber.org/2012/07/02/cooperation-and-corruption>

Swartz, Aaron (2012). “A database of folly”. *Crooked timber*, 3 July. <http://crookedtimber.org/2012/07/03/a-database-of-folly>

Ticbeat (2012). “La campaña de *big data* que dio la victoria a Obama”. *TICbeat*, 8 nov. <http://bigdata.ticbeat.com/campana-big-data-dio-victoria-obama>

Wikipedia (2012). *Datos abiertos*. http://es.wikipedia.org/w/index.php?title=Datos_abiertos

Yglesias, Matthew (2012). “Open data journalism”. *Crooked timber*, 28 June. <http://crookedtimber.org/2012/06/28/open-data-journalism>