

# El periodismo de datos como especialización de las organizaciones de noticias en Internet

Jesús Flores Vivar  
jmflores@ucm.es

Cecilia Salinas Aguilar  
saliflor@movistar.es

Recibido: 03 de diciembre de 2012

Aceptado: el 28 de junio de 2013

Universidad Complutense de Madrid (España)

**Resumen:** El presente trabajo corresponde al análisis sobre una forma de especialización periodística basado en los datos y que las organizaciones de noticias (antiguas empresas de comunicación) vienen incorporando esta especialidad como innovación del periodismo. El periodismo de datos (*Data Journalism*) consiste en usar estadísticas y herramientas de visualización para crear y contar mejor las historias de forma novedosa y atractiva para la audiencia de Internet. Es la evolución de lo que ya hace algunos años se conocía como periodismo de precisión o periodismo de escritorio (*Desktop Journalism*).

El artículo, partiendo desde una contextualización analítica, pretende reflexionar cómo el periodismo de datos, cuyo uso se desprende de las tecnologías existentes unidas al tratamiento de datos, se complementan las diferentes formas y contenidos dando paso a una nueva especialidad del periodismo, concretamente, del periodismo de investigación.

**Palabras clave:** Datos, Internet, periodismo de datos, investigación, perfiles.

**Abstract:** This work relates to the analysis of a form of specialization based on *Data Journalism* and news organizations (old Media companies) have been incorporating this specialty as innovation journalism. *Data Journalism* involves using statistical and visualization tools to create and tell stories better in new and attractive to the Internet audience. It is the evolution of what some years ago was known as Precision Journalism or Desktop Journalism.

The article, starting from an analytical contextualization, aims to demonstrate how *Data Journalism*, whose use is clear from existing technologies linked to data processing,

*supplementing the various forms and contents, giving way to a new specialty of Journalism and, specifically, Investigative Journalism.*

**Key words:** Data, Internet, Data Journalism, Investigation, Profiles.

## 1. Introducción

Los procesos evolutivos de la sociedad están exigiendo a los medios informativos más rigor y profesionalidad en sus publicaciones. Las Tecnologías de la Información y Comunicación (TIC) y, concretamente, Internet y las bases de datos proporcionan gran capacidad para procesar, comparar y analizar críticamente todo el maremágnum de información, es decir, los contenidos. Surge así el llamado “periodismo de datos” o *Data Journalism*.

Internet, como gran base de datos universal y plataforma de contenidos y recursos digitales es la causa de que la sociedad, el periodismo y los medios estén viviendo la “era del dato”. El periodismo de investigación y su vertiente, el periodismo de datos, se ha constituido en una disciplina de gran éxito a nivel internacional. En este contexto, hablar de periodismo de datos es referirnos a técnicas nuevas en la creación de información periodística que requiere de conocimientos específicos y destrezas en el uso de bases de datos y presentación (visualización) de la información en nuevas interfaces. Diarios como *The New York Times* o el británico *The Guardian* vienen incorporando este modelo periodístico basado en los datos y las visualizaciones como diferenciador de sus noticias y elemento clave para atraer lectores en un momento tan difícil para el periodismo. Incluso, el gigante Google apoya este modelo periodístico con la convocatoria, en 2012, del internacional Data Journalism Awards.

Por ello, desde una perspectiva más actual, en el saber hacer del periodismo profesional (el de toda la vida y que sumado al conocimiento y habilidades tecnológicas) nos encontramos con el periodista de datos como un perfil profesional emergente, cuyo entorno de trabajo se basa fundamentalmente en la construcción de visualizaciones cada vez más novedosas que combinan contenidos (de información) más tecnología y que tienen múltiples aplicaciones en las diversas formas del periodismo especializado.

### 1.1. Hipótesis

Los periodistas comienzan a diseñar su propio futuro, ilusionados con proyectos que permiten que la profesión siga existiendo más allá de los grandes medios. En este contexto, la universidad juega un papel fundamental y de vital importancia en la formación de los periodistas de nuevo perfil, con un tipo de conocimientos que obliga a una revisión permanente de los planes de estudio.

Pero, ¿cuáles son los contenidos de esos planes de estudio? Como referencia cabe citar a instituciones de prestigio como la Universidad de Columbia de Nueva York, la City University Nueva York (CUNY), el Poynter Institute de Florida, la City University de

Londres o la Universidad de Birmingham (Reino Unido), instituciones que han incorporado una asignatura o seminario sobre *Data Journalism* (periodismo de datos) en sus cursos de grado y postgrado. Más aún, algunas escuelas, han convertido el entorno del periodismo de datos en una titulación de máster.

No obstante, surgen algunas incógnitas como: ¿Es suficiente este cambio estructural en los planes de formación de los estudiantes de periodismo? ¿Qué tipo de conocimientos se han de considerar en la nueva estructura de los estudios periodísticos? ¿Cómo adaptamos las características de la cultura de datos a la nueva realidad mediática? ¿Cómo transformamos los datos en formatos reutilizables y que nos permitan hacer investigaciones poderosas? ¿Cómo visualizamos los datos con herramientas multimedia? ¿Cómo podemos utilizar las tablas, gráficos y mapas o aplicaciones como Google Fusion Tables, Maps, Hotmaps? De ahí que nuestra primera hipótesis es que las Facultades de Comunicación deben cumplir y asumir nuevos roles en la formación de periodistas que demanda, a escala global, una audiencia hiperconectada.

Por otro lado, en la aplicación práctica del periodismo de datos vemos que su desarrollo se da en países que tienen una buena legislación en cuanto a leyes de acceso a la información pública se refiere, es decir, leyes de transparencia. Consideramos que es importante una ley de transparencia para el desarrollo normal de este modelo periodístico basado en los datos. Así, nuestra segunda hipótesis plantea la necesidad de profundizar en la consolidación de una buena ley de acceso y transparencia, tan necesitada en los tiempos actuales.

## 1.2. Objetivos

El objetivo que persigue este artículo contempla no solo estudiar y describir el funcionamiento del nuevo perfil profesional que se viene expandiendo al alimón de la red; sino, también, analizamos su área de influencia en el mundo de la información y sus futuras aplicaciones en la profesión periodística (y consecuente efervescencia del periodismo de datos); y planteamos la necesidad de una formación adecuada y el tipo de conocimientos que requieren los periodistas, así como el rol de las Facultades de Periodismo en la configuración de este nuevo perfil profesional (Flores, 2011: 28).

Por lo demás, el estudio y reflexiones que aquí planteamos es parte del análisis de una investigación más amplia sobre técnicas del periodismo de datos y acceso abierto que los autores de este artículo desarrollan para una posterior publicación<sup>1</sup>.

---

1 La investigación más amplia se desarrolla en el Internet Media Lab (UCM: 970620) y, que uno de los autores, viene abordando a través del Proyecto *Cybermedia II: Desarrollos e innovaciones del periodismo en las redes sociales en Internet y telefonía móvil. Convergencias, modelos de negocios, servicio y formación* (MINECO-Plan Nacional de I+D+i- Referencia: CSO2011-25235).

## 2. Periodismo de datos: teoría, contexto y evolución

El periodismo de datos (o *Data Journalism*) se ha convertido en un campo del periodismo de investigación de gran éxito a nivel internacional. Para ello, partimos del hecho que las tecnologías de la información han cambiado por completo el concepto que teníamos del mundo. Se han creado miles de herramientas nuevas que han diversificado los usos de las que ya conocíamos. En esta vorágine de novedades, se comenzó a “mezclar” distintas herramientas para nuevos usos. La ciencia informática y la telemática (junto a la abundancia de información y contenidos que se genera cada vez más en la red y al uso de Google que facilita la búsqueda y localización de datos) ha dado pie al surgimiento de un entorno profesional que se explota cada vez más en los medios.

Por ello, destacamos que el factor más importante del periodismo de datos es, sin duda, la posibilidad técnica. La digitalización de documentos y hechos acaecidos en el mundo, en general, ha producido una increíble cantidad de material para buscar, seleccionar, procesar, analizar, comparar y finalmente publicar por parte del periodista. Este proceso requiere de la disponibilidad de herramientas y programas especiales, así como las habilidades técnicas y matemáticas para usarlos (Fama, 2010: 7).

Para Giannina Segnini<sup>2</sup> (*ápur* Ramírez, 2012), el periodismo de datos no remplaza las formas tradicionales del oficio. Sólo es una herramienta para hacer mejor el trabajo por lo que plantea que en el periodismo de datos deben contemplarse cinco pasos básicos:

- *Obtención de los datos*: Es la primera fase, que para muchos parece muy fácil, pero no lo es y cada vez se vuelve más compleja, porque hay servidores públicos que no entienden por qué se desea acceder a una base de datos, para qué le puede servir a un periodista. Aquí es clave no ceder e insistir con todos los protocolos que permiten el acceso a la información.
- *Limpieza de los datos*: Estos normalmente vienen repetidos, códigos que no están correctos, información que no es de interés para el fin planteado, entonces hay que revisarlos y normalizarlos de acuerdo con nuestra necesidad. Esta es la parte más difícil para el periodista. Vale la pena hacerse un curso sobre hoja de cálculo, CSV.
- *Análisis*: “Es el paso más importante en periodismo de base de datos, no se trata de publicar datos, es hacer lo que hacemos los periodistas desde tiempos inmemoriales, sacar una conclusión”, asevera Segnini. Para esto es necesario comparar la información encontrada, correlacionarla con otra, cruzar datos para descubrir tendencias, patrones, comportamientos atípicos.
- *Verificación de la información*: Las bases de datos no son la verdad absoluta. Verificamos cada resultado que encontramos, “esto se llama calle”. Hay periodistas

---

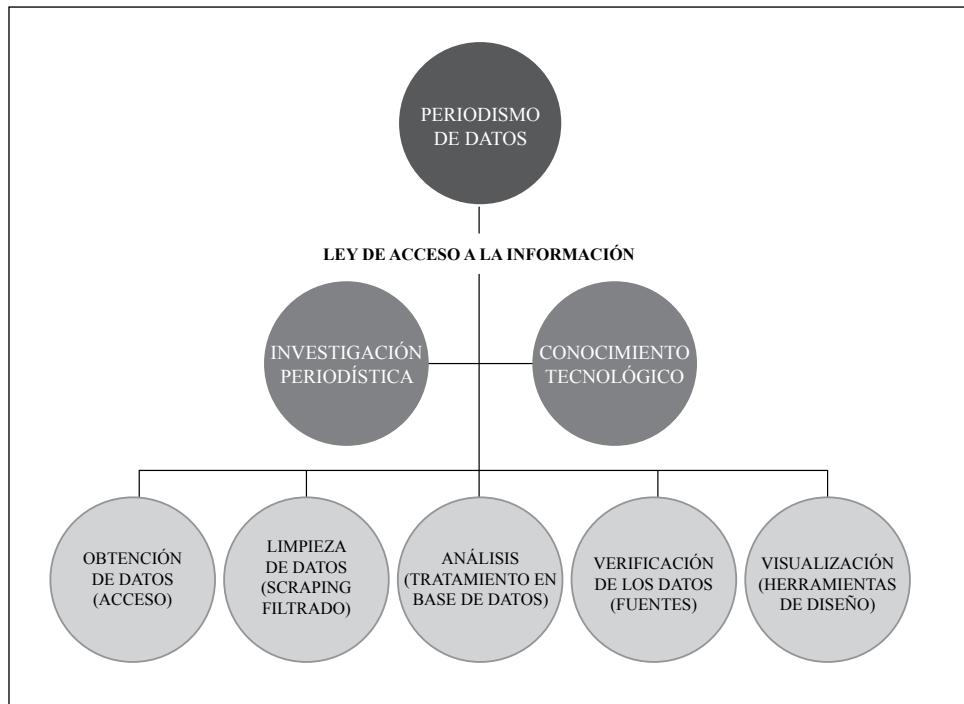
2 Periodista e investigadora del periódico *La Nación* de Costa Rica. Es editora de la Unidad de Investigación de la mencionada publicación. Con su equipo reveló la corrupción en licitaciones públicas de equipos médicos que involucró a la estatal Caja Costarricense de Seguro Social (CCSS), al consorcio finlandés Fischel y a tres ex presidentes de Costa Rica, dos de los cuales terminaron en la cárcel; así como los sobornos de Alcatel para resolver cargos en su contra en Estados Unidos.

que creen que todo está en Google y que el buscador es la verdad revelada. La mala noticia para esta clase de periodistas es que el rigor seguirá siendo la mejor arma para el periodismo de calidad.

- **Visualización:** Pensar como diseñador es hoy una obligación de todo periodista. Es necesario encontrar la forma de explicar de la mejor manera posible a la audiencia algún dato. Por eso, resulta relevante sugerir que si se piensa la información, antes de escribirla, en términos de cómo presentarla puede ayudarnos a hacer informaciones más agradables y claras. Cómo presentarles a los lectores esta información, que en no pocas ocasiones resulta confusa, es una pregunta que deberíamos sumar a las cinco W.

En coincidencia con los pasos mencionados, proponemos que el periodismo de datos se basa: *a)* en la investigación periodística; *b)* en el conocimiento tecnológico; y *c)* en una buena ley de acceso a la información. Sin una ley de transparencia o acceso a la información pública, difícilmente se puede llevar a cabo el desarrollo del periodismo de datos en toda regla. A su vez, el conocimiento tecnológico debe darse, en primer lugar, en la forma como obtenemos o accedemos a los datos; en la selección de los datos mediante técnicas de filtrado; en el análisis y tratamiento de bases de datos; en la verificación mediante el cruce de datos estadísticos con el uso de la informática que permita llegar a las fuentes; y, en la visualización de la información tratada.

**Figura 1:** Mapa contextual para el desarrollo del periodismo de datos.



**Fuente:** Elaboración propia.

En este contexto, vemos que el periodismo de datos es posible gracias a la evolución tecnológica, que ha dado lugar a la digitalización de los hechos que suceden en el mundo y el uso de herramientas de intercambio de conocimientos unida a la conciencia cívica y la necesidad de transparencia de los periodistas, activistas, administradores y personas que encuentran representación en las filas de la cultura informal del *Open Data* (“datos abiertos”). El crecimiento del periodismo de datos, derivado, entre otros, de las nuevas oportunidades que ofrece el desarrollo tecnológico para impulsar la cultura cívica de los datos abiertos, lo encontramos en el ejemplo de Wikileaks, que desde su aparición se convirtió en un fenómeno de información. Al respecto, diversos expertos del nuevo periodismo en plataformas tecnológicas hacen sus aportaciones sobre el nuevo entorno profesional.

En 2012, Rosental Calmon Alves<sup>3</sup>, en la conferencia de apertura del XIII Congreso de Periodismo Digital, en la ciudad de Huesca (España), afirmaba que “en un mundo basado en bases de datos, crece la importancia del periodismo de datos y de visualización de datos”. Para Fama (2011), el periodismo de datos es una técnica de acoplamiento fascinante de tecnologías y datos de acceso público, cuyo potencial de desarrollo es inmenso porque se basa en interminables campos de aplicación del periodismo.

Para Ramonet<sup>4</sup> (2011: 18), el periodismo de [bases de] datos es otro tipo de periodismo que está prosperando entre los nuevos nichos de Internet. Se le considera un gran logro de la democracia moderna porque permite buscar y acceder, de forma inmediata, a bases de información creadas por instituciones públicas o privadas. No obstante, el periodismo de datos se basa en un principio fundamental del periodismo de siempre: “los hechos son sagrados”.

Paul Bradshaw (2011)<sup>5</sup> publicó en el periódico *The Guardian* un artículo titulado “Cómo hacer periodismo de datos”, en donde afirma que el periodismo de datos es aquel “en el que se usa el poder de las computadoras para encontrar, contrastar, visualizar y combinar información proveniente de varias fuentes”. Esta afirmación es corroborada mediante su propuesta de “pirámide invertida de periodismo de datos”.

En concordancia con el mencionado artículo, vemos que las técnicas del periodismo de datos pueden utilizarse en todas las áreas del periodismo. Política, Deporte, Nacional, Sucesos, Economía, Medio Ambiente, Cultura, etc., dando por sentado que este tipo de periodismo tenga más repercusión en unas áreas que en otras.

Así, en el periodismo deportivo se aplican estas técnicas y se logran muchas innovaciones, dada la gran cantidad de datos involucrados y muchas personas con conocimiento profundo en interés en esta área. Por supuesto, en el periodismo de investigación es donde

---

3 Rosental Calmon Alves es catedrático de la Journalism School of Texas University, Estados Unidos.

4 Ignacio Ramonet es periodista y director de *Le Monde Diplomatique*.

5 Paul Bradshaw es periodista, profesor e investigador de la Universidad de Birmingham, Reino Unido.

más se han producido grandes innovaciones utilizando ordenadores (computadoras) para crear enlaces entre informaciones o encontrar historias entre grandes conjuntos de datos. Asimismo, el periodismo de datos es creciente en la información política, puesto que cada vez más interesan a los ciudadanos los datos acerca del gasto de los gobiernos. Diversas organizaciones de noticias han empezado a crear unidades de trabajo desarrollando periodismo de datos para cualquier área del medio. Según Bradshaw (2011):

“Muchas publicaciones ya no cuentan con una unidad investigativa, pero podemos decir que las unidades de periodismo de datos deberían adoptar objetivos similares para hacer que la información pública tenga significado y sea más fácil de digerir”.

Respecto al tipo de formación que debe tener un periodista de datos, algunos expertos no dudan en afirmar que las Escuelas y Facultades de Periodismo no han formado a profesionales para desarrollar este tipo de tareas. En una unidad de profesionales dedicada al periodismo de datos, el perfil de los periodistas debe ser personas con habilidades para la investigación tradicional, incluyendo el periodismo asistido por computador (CAR); periodistas con conocimientos de estadísticas, diseñadores de información y programadores que puedan crear secuencias de comandos para facilitarle a los periodistas y a los usuarios obtener y aprovechar la información. Sobre estos aspectos formativos, profundizaremos en los siguientes apartados.

En cualquier caso, es importante saber que los profesionales que deseen enfocar su trayectoria hacia el periodismo de datos deben asumir unos conocimientos básicos de programación, diseño y manejo de bases de datos, sin que esto sea una especie de *kit* en la formación. Tanto Fama (2011) como Bradshaw (2011) coinciden en que hacen falta habilidades para practicar el periodismo de datos. No obstante, lo más importante es que el periodista debe ser capaz de encontrar información, entenderla y hacerla inteligible para el público. Sin ese factor, la programación, el diseño y otros conocimientos no tendrán ninguna utilidad. Esto nos lleva a pensar que antes de empezar a trabajar con aplicaciones y describir habilidades concretas, el periodista debe partir desde una amplia comprensión de lo que puede hacer y, en segundo término, adquirir aquellos conocimientos que sean necesarios para lograr un proyecto informativo en particular.

Otro tema de grandes debates es sobre los géneros y los formatos narrativos que surgen a partir del periodismo de datos a través de plataformas digitales. Las macroempresas del mundo de las tecnologías ven en el periodismo de datos una oportunidad para desarrollar y crear aplicaciones que puedan ser utilizadas por profesionales y usuarios. El ejemplo lo tenemos en Google Corp. Los mapas se han convertido en una forma de periodismo nativo de la web, éxito derivado en parte por el lanzamiento de la aplicación de Google Maps en 2005.

Otro ejemplo lo encontramos en el tipo de periodismo que permite a los usuarios profundizar en los detalles de algún asunto utilizando bases de datos tipo Access. A todo esto se suman emergentes formatos narrativos que permiten explorar los asuntos públicos a través de juegos (por ejemplo: retar a un usuario a que maneje un presupuesto para que tenga una

visión de los retos que asumen un gobierno y las implicaciones de sus decisiones); el *crowdsourcing* (formato de trabajo colectivo) ahora más viables con el poder de los ordenadores para trabajar en red (computadores para hacer conexiones). También están los *mashups*, que permiten combinar distintas fuentes de información y proporcionan al usuario la opción de usar a su antojo el producto informativo.

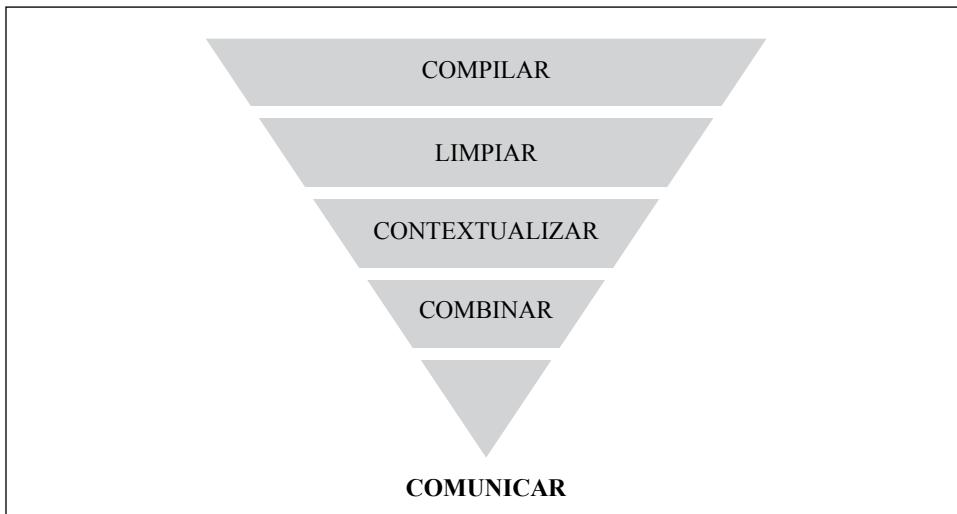
**Figura 2:** Modelo de periodismo de datos con mapas, desarrollado por *The New York Times* en Google Maps sobre las elecciones presidenciales en Estados Unidos.



Fuente: *The New York Times* (2007).

Es evidente que todo el periodismo tiene como base la información, pero al usar la palabra “datos” o “data”, implícitamente nos referimos a un tipo de información particular que puede ser procesada por sistemas informáticos y potentes herramientas de computación. La pirámide invertida del periodismo de datos (figura 3) muestra, de forma simple, como se avanza desde una gran cantidad de información incrementada exponencialmente y que la misma se va concretando hasta llegar al punto de proporcionar a la audiencia, unos resultados más claramente posibles. La apuesta de Bradshaw en el modelo de pirámide invertida, atraviesa cuatro etapas en el proceso del periodismo de datos: compilar, limpiar, contextualizar, combinar, para al final, comunicar.

**Figura 3:** Pirámide invertida del periodismo de datos.



**Fuente:** Bradshaw (2011).

**a) Compilar**

Bradshaw considera que el periodismo de datos empieza en una de las dos formas: o tienes una pregunta que necesita datos o tienes un set de datos que necesitan ser interrogados. En este sentido, la compilación de información es lo que lo define como un acto de periodismo de datos y puede tener varias formas:

- Datos propuestos directamente por una organización.
- Datos encontrados usando técnicas de búsqueda avanzada para surcar en las profundidades de sitios web (por ejemplo, de gobiernos).
- Al hacer *scraping* (rascar) de bases de datos escondidas online usando herramientas como OutWit Hub y Scraperwiki.
- Convertir documentos en algo que pueda ser analizado usando herramientas como DocumentCloud.
- Tomando información de API.
- Recolectando los datos por el propio periodista a través de la observación, encuestas, formularios online o *crowdsourcing*.

**b) Limpiar**

Tener información es sólo el comienzo. Estar confiado en las historias escondidas dentro de las bases de datos significa poder confiar en la calidad de los datos y eso significa limpiarla. Limpiar en general tiene dos formas: remover el error humano y convertir el dato en un formato que es consistente con otro dato que el periodista esté usando. Hay formas simples, como buscar y reemplazar, de limpiar los datos en Excel o Google Docs, clasificando los datos y usando filtros para que sólo se muestren entradas duplicadas.

**c) Contextualizar**

Como cualquier fuente, la información no puede siempre ser confiable. Viene con sus propias historias, prejuicios y objetivos. Por ello, como con cualquier fuente, se necesita hacer preguntas, tales como: ¿quien recolectó los datos?, ¿cuando y por qué motivos?, ¿cómo fue recolectada (la metodología)? También es necesario entender la jerga como códigos que representan categorías, clasificaciones o ubicaciones y terminología de especialistas. Todo esto, probablemente, nos lleve a compilar más información. Por ejemplo, conocer el número de crímenes en una ciudad es interesante, pero sólo se vuelve relevante cuando se contextualiza junto con la población, número de policías, niveles de los crímenes de los últimos cinco años, percepción del crimen, niveles de desempleo y demás. El alfabetismo estadístico es una obligación en este nivel.

**d) Combinar**

Las buenas historias se pueden encontrar en un solo set de datos, pero a menudo es necesario combinar dos juntos. La combinación clásica en este apartado es el *mashup* de mapas: tomar un set de datos y combinarlo con datos de mapas para proveer una visualización instantánea de cómo algo está distribuido en el espacio. Esto se ha convertido en una acción común (sobre todo, gracias a la API de Google Maps que ha sido de gran utilidad para el periodismo). Una combinación más trivial es fusionar dos o más sets de datos con un punto de información común (por ejemplo, el nombre de un político, una escuela o una ubicación). Eso usualmente significa asegurarse que el punto particular del dato está formateado con el mismo nombre a través de cada uno de los sets de datos.

**e) Comunicar**

En el periodismo de datos necesitamos visualizar los resultados. Esta visualización puede hacerse en un mapa, cuadro, infografía o animación. Sin embargo, a esto debemos sumar que hablamos de periodismo por lo que hay mucho más a considerar (desde la clásica narrativa, las aplicaciones de noticias, casos de estudio y personalización).

### **3. El periodismo de datos como especialización periodística en los medios digitales**

Nos encontramos en una situación de expectación e incertidumbre. Los medios no saben como desarrollar los nuevos modelos informativos. Los cambios en los que se encuentran inmersos son sólo tentativas de viabilidad. Por ello, algunas voces auguran un futuro incierto de los medios y, en consecuencia, de los profesionales. También, se pone en entredicho la profesión periodística, la publicitaria, la empresarial, etc. Es, por tanto, una crisis estructural del periodismo y la comunicación que afecta a todos los sectores relacionados con los medios de comunicación.

Como antecedente del periodismo de datos, nos situamos en los años 60 con la consolidación del periodismo de investigación mediante la implantación del *New Journalism* o “Nuevo periodismo” en los Estados Unidos. Esta modalidad permitió desarrollar con fuerza la figura

del *muckraker* (ahora *New Muckraker*) y se produjo por la toma de conciencia de la libertad y derechos del pueblo frente a los gobernantes, influenciados por la revolución de Mayo del 68. Los periodistas no se conforman con las versiones oficiales, se vuelven desafiantes y escépticos frente al poder y se convierten el valedores de los intereses ciudadanos.

Desde los revolucionarios años 60, la situación del periodismo ha cambiado mucho. Tras el escándalo del Caso Watergate<sup>6</sup>, recogido por Carl Bernstein y Bob Woodward, en *The Washington Post* (por el que recibieron el premio Pulitzer en 1973) se han sucedido muchos trabajos de investigación basándose en la recuperación y análisis de datos e información pública como la que hicieron los citados periodistas del *Post* en los archivos de la Biblioteca del Congreso estadounidense.

Sin embargo, en la actualidad, so pretexto de la crisis que afecta también a las empresas periodísticas, estas han encontrado la excusa para despedir a media plantilla con lo que la información en los medios es más superficial y de menor calidad. A esto se suma que el desarrollo del periodismo de investigación (y por consiguiente, del periodismo de datos) es demasiado oneroso, puesto que liberar a un periodista para que investigue un tema no está a alcance de cualquier medio de comunicación. Como consecuencia de todo esto se puede decir que el periodismo de datos, aunque es un valor en alza, no suele estar entre los planes de trabajo diario de los medios.

En el seguimiento que hemos realizado sobre la evolución del periodismo de investigación y su derivación en periodismo de datos, detectamos que el ritmo de evolución y cambio en el mundo tecnológico y, más concretamente, en el informático, es frenético, acelerado y un tanto delirante. Nombres de aplicaciones que hasta hace poco desconocíamos, así como personalidades del mundo de la red que se han hecho un lugar, por su afán de superación y su constante trabajo en este campo, nos han demostrado que el cambio es constante.

Para entender el importante e innovador perfil del *Data Journalism*, ha sido preciso conocer y analizar la figura del *Data Delivery Editor*, profesional que crea los *mashups*, como nuevo entorno de información, en donde se cuenta una historia, se proporciona información utilizando diferentes elementos tecnológicos (gráficos, mapas, estadísticas).

#### **4. Aptitudes y actitudes en el perfil del periodista de datos**

Aunque el análisis de la información es una labor intrínseca a la profesión periodística, en los últimos años estamos asistiendo al desarrollo y auge de un nuevo método periodístico basado en el análisis de grandes volúmenes de datos y que involucra a programadores, estadísticos o diseñadores. Se trata de contar historias a partir de los datos.

---

<sup>6</sup> Como se sabe, la difusión de este acto de espionaje provocó la dimisión del presidente Richard Nixon en 1974.

Los periodistas nos enfrentamos a nuevos desafíos y problemas en esta era de abundancia informativa, en donde la escala de la producción de datos y la diversidad de formatos demandan nuevos estándares, recursos y herramientas pero, sobre todo, una visión renovada (visión global) sobre nuestro trabajo. Un nuevo perfil.

Los periodistas que quieran trabajar en los nuevos medios, deben aprender y conocer en profundidad la gran cantidad de recursos y herramientas tecnológicas que les permitirá desarrollar su trabajo. Pero, ¿cuáles son las habilidades que necesita un periodista de datos? ¿Qué destrezas? ¿Dónde se forman esos nuevos perfiles? ¿Qué contenidos formativos se requieren?

Alan Grimm, colaborador del área de reclutamiento del Poynter Institute, ofrece algunos consejos a los periodistas que quieran aprender las más importantes tecnologías digitales. En primer lugar, sugiere aprender de varias áreas del periodismo y mostrarse menos preocupado sobre el aprendizaje de programas específicos. El periodista no se debe fijar tanto en el uso del software (programa) ya que cuando un reportero aprende, por ejemplo, a crear y producir videos usando algún programa conocido o de moda, bien puede transferir esas habilidades a otros programas.

En línea con lo expuesto por Grimm, otros expertos coinciden en que el profesional debe saber como hacer una grabación y edición de audio y video, edición de fotos (por ejemplo, con el uso de Photoshop como norma general) y utilizar bases de datos (Access o Excel son buenos programas de software para empezar). En un contexto más avanzado y más vanguardista, la aparición constante de recursos tecnológicos y, sobre todo, teniéndose en cuenta el grado de aceptación y adaptación por parte de los medios a los nuevos formatos de publicación, debe ser una tarea prioritaria el conocimiento y aprendizaje de los diferentes Sistemas de Gestión de Contenidos (CMS, por sus siglas en inglés).

El perfil del periodista de datos necesita de conocimientos que le ayuden a gestionar bases de datos y cruzar la información que contienen con otros elementos o conocer herramientas y recursos que le ayude a crear híbridos de información, conocidos como *mashup*. En este contexto, el periodista de datos (*Data Journalist*) se define como un profesional que puede estar formado en documentación, periodismo o cualquier otra disciplina, que posee un aditivo de conocimiento necesario para la gestión datos y clasificación de información, así como para la búsqueda, procesamiento, interpretaciones concluyentes, análisis de datos y capacidad para crear y publicar información con una narrativa propia que se desprende del propio género utilizado. Un buen periodista de datos debe manejar grandes cantidades de información, debe clasificarlo y después, introduce esos datos en las aplicaciones tecnológicas de uso, por donde considere que su información tendrá mas impacto.

El periodista de datos que crea o gestiona un *mashup* (en un mapa de Google) deberá, previamente, haber analizado diferentes datos, contrastado fuentes, antes de su publicación final que deberá hacerlo de forma ordenada y coherente. Como hemos indicado, este tipo de perfil profesional aun no está muy presente en los medios de comunicación, pues la importancia de los datos o *mashups* es todavía prácticamente testimonial. Sin embargo,

puede que en un futuro no muy lejano se popularicen estos perfiles en los medios digitales, tal y como ya se vienen consolidando en los medios estadounidenses. El siguiente paso, en la función del periodista de datos, será controlar las novedades que puedan surgir, buscar nueva información, actualizar la existente y, en último término, investigar nuevos servicios para crear nuevas herramientas o mejorar las que ya tienen.

Por ello, consideramos que para hacer periodismo de datos son necesarias personas con habilidades en la investigación tradicional, incluyendo el periodismo asistido por computador, periodistas con conocimientos de estadísticas, buenos diseñadores de información y programadores que puedan crear secuencias de comandos para facilitarle a los periodistas y a los usuarios obtener y aprovechar la información. Es decir, un equipo de trabajo multidisciplinar que aun no se asimilan en las redacciones de los medios.

En este sentido, los profesionales tienen que asumir nuevos roles que van surgiendo, como ya lo vienen haciendo diversos periodistas de la talla de Aron Pilhofer, editor del área interactiva del *The New York Times*, quien dirige un equipo de periodistas y desarrolladores que crean aplicaciones basadas en datos para mejorar los tiempos de presentación de los informes online; o el caso de Simon Rogers, responsable del Datablog del periódico británico *The Guardian*. Sin embargo, la opinión generalizada es que todavía es difícil convencer a los periodistas de que el periodismo de datos es una forma eficaz para contar una historia.

Se trata, pues, de un asunto de mentalidad y de habilidades. Para muchos periodistas aún es difícil comprender las posibilidades que ofrece esta forma de periodismo. En palabras de Bradshaw, “los periodistas están acostumbrados a contar historias desde una perspectiva particular para una gran masa. Permitirle al usuario contar su propia historia, haciendo clic en un mapa o profundizando en un tema específico, significa renunciar un poco al poder como periodista, y eso requiere algo de humildad”. Para Pilhofer el valor del periodismo de datos radica en el hecho de permitir generar historias que antes era muy difícil de hacer.

La evolución de los contenidos y la tecnología lleva a transitar por caminos profesionales que empiezan a encontrar un punto de convergencia. Una convergencia que empieza a ser real y necesaria. La idea de juntar periodistas y programadores de software proviene de mediados de 2000. Empezó a llevarse a cabo cuando Richard Gordon y Aron Pilhofer, del TNYT, que decidieron promover encuentros entre periodistas y programadores de software, coincidieron en San Francisco con Burt Herman, un periodista emprendedor fundador de Storify. Entre los tres, no sólo tuvieron la misma idea, sino que coincidieron en el nombre a esos encuentros denominándoles “Hacks/Hackers”. *Hacks* (“recortar”) que es lo que hacen los periodistas y *hacker* que es, en definitiva, “programar”. Hoy, esos tipos de encuentros vienen replicándose en diversas ciudades estadounidenses, latinoamericanas y europeas, incluida España.

Pese a estas iniciativas, la cultura del periodista tradicional sumado a otras razones, como la falta de recursos para la capacitación, hace que existan (aún) muy pocas personas que dediquen tiempo a desarrollar sus habilidades para este tipo de periodismo. De ahí que muchas de las innovaciones en periodismo de datos se están realizando o se hayan iniciado fuera de los medios, como por ejemplo el realizado por Adrian Holovaty en Estados

Unidos, periodista y programador y uno de los primeros creadores de *mashups* denominado ChigagoCrime.org, un sitio web creado en 2005, en donde se informaba puntualmente sobre los índices de criminalidad que se cometían en la ciudad de Chicago.

Por todo esto, con el fin promover esta reinvestigación del periodismo, diferentes organizaciones como Global Editors Network (GEN) y Google lanzaron, en 2012, el Premio Periodismo de Datos (datajournalismawards.org). Los Premios de Periodismo de datos (DJA) constituye el primer concurso internacional que reconoce la destacada labor en el campo del periodismo de datos en todo el mundo. Con estos premios, en donde el DJA reconoce la excelencia editorial en el campo del periodismo de datos, se busca contribuir al establecimiento de normas y poner de relieve las mejores prácticas del periodismo de datos, así como demostrar el valor de datos entre los editores y ejecutivos de los medios.

## **5. La formación sobre el periodismo de datos en las Facultades de Comunicación. Imbricación del periodismo de investigación con el conocimiento tecnológico**

Una vez lograda la convergencia tecnológica y de formatos, se debe lograr la convergencia en la mente de las personas. No cabe duda que el propio concepto de periodismo digital, de periodismo online o de ciberperiodismo (Flores y Arruti, 2001: 17) se ha consolidado ya como una forma de realizar un trabajo profesional de la información, basado en una cultura de red o cibercultura en donde subyacen, primero, el conocimiento y segundo, las destrezas y habilidades que permitan utilizar las diversas herramientas y recursos que se encuentran en Internet y que trae consigo desarrollar un nuevo formato informativo. Los medios de comunicación que apuestan por los cambios y los medios que han nacido sólo en la red, demandan de unos profesionales que son escasos, que no se encuentran fácilmente, pero que,矛盾oramente, no se forman en las Facultades de Comunicación.

El periodismo de datos debería ser hoy una materia más en los nuevos currículos de las Facultades de Periodismo. Conocer las bases de datos públicas a las que se puede acceder, las que se pueden pedir a través de un simple derecho de petición o incluso las que van a necesitar ser “hackeadas” porque son inexpugnables, pero contienen información relevante para la opinión pública.

Las Facultades de Comunicación, tanto españolas como iberoamericanas, no contemplan aun dentro de sus planes de estudio la formación para el periodismo de datos, como si lo hacen ya, desde hace algunos años, las Escuelas de Periodismo de universidades anglosajonas.

### **5.1. La formación del *Data Journalist* en centros especializados internacionales**

El periodista de datos que ya ejerce esta práctica en los medios ha tenido que formarse de forma autodidacta o, en el mejor de los casos, asistiendo a cursos de postgrado en universidades internacionales, sobre todo de Estados Unidos, Reino Unido, Francia, Suecia

o Alemania. Algunas organizaciones que brindan formación sobre periodismo de datos son el Centro para el Periodismo Investigativo (The Centre for Investigative Journalism, CIJ) que tiene sede en el Reino Unido y la Asociación de Editores y Reporteros de Investigación (Investigative Reporters & Editors) en su siglas, IRE.

Debemos ser conscientes que Internet afecta y cambia la dicotomía entre sociedad e información e impacta de lleno en el periodismo y la comunicación (Flores, 2011: 75). Cambia la forma y los contenidos y el modo de acceso a estos, por lo que medios y profesionales son los más afectados por estos cambios. De ahí que, una primera cuestión consiste en superar las barreras tecnológicas en su uso y concepto, principalmente este último, ya que muchos medios están impregnándose de una cibercultura en sus redacciones. Cada vez más empresas exigen periodistas digitales. Eso significa que los profesionales deben entrar en una dinámica de “aprender a aprender”, aunque antes, quizás, deban aprender a desaprender. Por tanto, el reciclaje se convierte en algo esencial y prioritario.

La demanda de profesionales para los medios digitales exige una formación con nuevos perfiles para la realidad actual de la industria periodística en la que está imbricado el desarrollo del periodismo de datos. Tal necesidad es tan apremiante que algunas empresas periodísticas elaboran planes de formación interna. Más aún, los medios de comunicación más importantes, en connivencia con centros oficiales de estudios, crean Escuelas de Periodismo en donde imparten cursos de postgrado formando profesionales para los nuevos medios. En España, iniciativas de periódicos como *El País*, *ABC* y *El Mundo* son solo una pequeña muestra de ello.

La incursión, primero, de los medios sociales (basados, principalmente, en las redes sociales y el *microblogging* y representados por Facebook, MySpace, Twitter, Flickr, etc.) y ahora, de aplicaciones para la creación de *mashups* y el avance y desarrollo del periodismo de datos, implica una reconfiguración de los modelos formativos que se han venido dando hasta la fecha para el periodismo tradicional. Ante el avance de las tecnologías que crean nuevos perfiles, algunos medios se resisten, pero nada puede parar el cambio de paradigma en la comunicación. El uso de mapas, de herramientas de análisis, de estadísticas, de optimización en la localización de datos, etc., entornos necesarios para el periodismo de datos requieren de un cúmulo de conocimientos que la universidad debe implementar en sus planes de formación para periodistas de nuevo cuño.

Instituciones como la Universidad de Columbia de Nueva York, la City University New York (CUNY), el Poynter Institute de Florida, la City University de Londres o la Universidad de Birmingham han incorporado una asignatura o seminario sobre *Data Journalism* en sus cursos de grado y postgrado. Más aún, algunas escuelas, han convertido el entorno del periodismo de datos en una titulación de máster. Los contenidos en los planes de estudios que imparten estos centros cubren todos los aspectos del periodismo de datos, que rebasan, principalmente, en la visualización de datos, uso de aplicaciones y el tratamiento estadístico, además de las propias técnicas del periodismo de investigación.

Las tecnologías (móvil, Internet, blogs, medios sociales, etc.) se han convertido en herramientas inseparables del ciudadano del siglo XXI. Esto trae como consecuencia que

la formación e investigación en las Ciencias de la Información y Comunicación sea cada vez más constante como apoyo al desarrollo y expansión de los nuevos medios.

El periodista de datos debe adaptar su mentalidad y cultura hacia un entorno de red, lo que implica una concienciación de una voluntad de experimentación continua como la ruta más probable para el éxito en los nuevos medios. Esto incluye no sólo probar las nuevas tecnologías, sino aprender a valorar el compromiso de que los medios no sean vistos como simples páginas de un sitio web.

De ahí que el rol de la universidad debe basarse en una formación en TIC avanzada desde sus inicios. En este sentido, las Facultades de Ciencias de la Información y de la Comunicación, deben adaptar sus planes de estudio en donde la tecnología sea un conocimiento transversal (en todas las etapas formativas) y forme a nuevos perfiles profesionales. La estrategia universitaria debe comprender la creación de laboratorios (*labs*) en sus Facultades de Comunicación (para investigación y transferencia de conocimiento) y cuyo objetivo sea desarrollar proyectos pilotos sobre periodismo de datos y la creación de *mashups*.

## **5.2. El proceso formativo del periodismo de datos en la universidad española**

La web supone una era de cambios y transformaciones. La profesión periodística y los medios de comunicación son los primeros afectados por estos cambios vertiginosos, incluyendo el propio concepto de periodismo. Estamos ante una auténtica globalización de la información.

En este contexto, preocupados por una perspectiva de constante innovación tecnológica y adquisición de una “cibercultura de red” que debe implantarse en los medios y empresas de comunicación, grupos de universidades e instituciones trabajan con el objetivo de poner en marcha proyectos que redunden en beneficio no sólo de la profesión periodística sino también de los profesionales que trabajen en ella, que de manera directa o indirecta se sienten afectados, así como de estudiantes del periodismo y la comunicación.

El periodista de datos es un perfil en alza que requiere una formación básica, al menos, en varios campos del conocimiento, como la Estadística, Programación, Diseño, Sociología, Psicología, además del Periodismo. Un perfil que pueda gestionar de forma eficiente grandes cantidades de información y tenga la disciplina de organizar esos datos, creando un servicio útil que ayude a comprender noticias, encontrar recursos o, simplemente, obtener información adicional y con un amplio nivel de detalles.

Desde la universidad se viene trabajando en estos aspectos que inciden en la forma como se va a desarrollar el nuevo periodismo basado en los datos en conjunción con las tecnologías emergentes. Diversas investigaciones, como el proyecto Cybermedia<sup>7</sup>,

---

<sup>7</sup> *Cybermedia II: Desarrollos e innovaciones del periodismo en las redes sociales en Internet y telefonía móvil. Convergencias, modelos de negocios, servicio y formación*, del Plan Nacional de I+D. 2012-2014.

dan fe de la preocupación por profundizar en los aspectos de la formación de perfiles para los nuevos medios. La ingente cantidad de datos, Internet, telefonía móvil y otras tecnologías emergentes han abierto un gran campo en el entorno cambiante de los medios de comunicación y, por consiguiente, en la demanda de un nuevo tipo de perfil profesional.

Una de las conclusiones del XIII Congreso de Periodismo Digital de Huesca destaca que los periodistas comienzan a diseñar su propio futuro, ilusionados con proyectos que permiten que la profesión siga existiendo más allá de los grandes medios. En este contexto, la universidad juega un papel fundamental y de vital importancia en la formación de los periodistas de nuevo cuño, con un tipo de conocimientos que obliga a una revisión permanente de los planes de estudio.

Vemos pues que las proliferaciones de materias y cursos referidos al periodismo de datos se dan, en mayor medida, en las universidades anglosajonas (Estados Unidos y Reino Unido) y en menor medida en universidades europeas y escasa presencia en los centros españoles. No obstante, en España empiezan a generarse algunas iniciativas en esta materia. Como referencia mencionamos el propio Congreso de Periodismo Digital de Huesca, de 2012, en donde se desarrolló una mesa de trabajo titulada “Periodismo de datos, nuevas historia y nuevas formas de contar historias”.

También se destaca la materia de Periodismo de Datos en los Cursos de Verano El Escorial de la Universidad Complutense, que se realizó en julio de 2012<sup>8</sup>, cuyo objetivo del curso era que los periodistas (y estudiantes de periodismo) adquieran los principios básicos de cómo transformar datos en poderosos recursos visuales que implican la comunicación y narrativa multimedia.

La Universidad Pompeu Fabra ha ofrecido para el presente curso 2012-2013 el Programa de postgrado de Visualización de la Información. Dicha materia pretende dotar de los conocimientos conceptuales, tanto básicos como avanzados, ligados a la disciplina del periodismo de datos en tres áreas principales: minería de datos (*Data Mining*), visualización de la información y la utilización de gráficos.

A estas iniciativas, se suma el Máster de Periodismo Profesional ABC-UCM, que imparte una asignatura relacionada con el periodismo de datos y el Máster de Periodismo Multimedia Profesional de la Universidad Complutense de Madrid, que ha incorporado, para el curso 2012-13, una asignatura sobre periodismo de datos y redes sociales.

Estas innovaciones que impulsan algunas Facultades de Comunicación evidencian que los contenidos de cualquier curso sobre periodismo de datos deben ayudar a estudiantes carentes de conocimientos de ingeniería informática (o de programación) a construir las competencias

---

8 La materia “Periodismo de datos”, que se realizó en julio de 2012 en los Cursos de Verano El Escorial-UCM, fue dirigida por el profesor Jesús Flores Vivar, uno de los autores de este artículo.

necesarias para diseñar, desarrollar y evaluar visualizaciones complejas. En este sentido, un curso de estas características debe cumplir con los objetivos de despejar las dudas sobre ¿cómo adaptar las características de la cultura de datos a la nueva realidad mediática? ¿cómo transformar los datos en formatos reutilizables y que permitan hacer investigaciones poderosas? y ¿cómo visualizar los datos, incluyendo tablas, gráficos y mapas?

Por todo esto, no es descabellado incluir dentro de los nuevos planes de formación periodística el aprendizaje de técnicas para el manejo de herramientas informáticas y de métodos de acceso al gran laboratorio de Internet.

## 6. Conclusiones

Una primera conclusión que se desprende de este estudio demuestra cómo emerge y se realiza el periodismo de datos en los principales medios de comunicación. Los periodistas de datos son una fuerza creciente del periodismo de investigación que se ha venido practicando desde tiempos remotos. La historia del periodismo de investigación nos hace ver cómo estos periodistas, en muchos casos, han realizado trabajos que han hecho tambalear a un gobierno o han desentrañado casos de corrupción.

Ahora los nuevos periodistas de investigación escudriñan en datos públicos, cuentan cómo les ayudan las tecnologías digitales y *online*, por lo que es necesario hacer una reflexión, en el actual contexto de crisis de los medios, sobre el futuro del periódico de papel y las posibilidades que brinda la web con la proliferación de esta forma especializada del periodismo.

Tanto este tipo de periodismo como el desarrollado con el uso de datos, utilizan diferentes recursos y aplicaciones tecnológicas que se traduce en una sinergia de conocimiento de ambos entornos, lo que conlleva un nuevo planteamiento en la formación de los periodistas. La formación sobre el perfil del periodista de datos en las Facultades de Comunicación viene dándose de forma residual. Es necesario hacer una revisión en los planes de formación periodística de 180 grados e intentar ir en paralelo con el avance y evolución de las tecnologías digitales que la sociedad ya ha tomado por asalto.

## Fuentes consultadas

Baker, S. (2008). *The Numerati*. Nueva York: G Library Agency.

Bradshaw, P. (2010). “Data Journalism Blog”. Extraída el 27/II/2012 desde <http://www.datajournalismblog.com/tag/paul-bradshaw/>

Bradshaw, P. y Rohumaa, L. (2011). *The Online Journalism Handbook: Skills to Survive and Thrive in the Digital Age*. Londres: Longman Practical Journalism Series.

- Fama, A. (2011). *Open data. Data Journalism. Transparenza e informazione al servizio delle società nell'era digitale*. Roma: Narccisa.
- Flores, J. y Cebrián, M. (2012). “El data journalism en la construcción de mashups para medios digitales”. XIII Congreso Periodismo Digital. Zaragoza.
- Flores, J.  
\_(2012). “Ecosistema del periodismo de datos”. *Comunicaçao & Sociedade*, vol. 34, núm. 1, pp. 7-35.  
\_(ed.) (2011). *Reinventar el Periodismo y los Medios*. Madrid: Fragua.  
\_(2009). “Nuevos modelos de comunicación, perfiles y tendencias en las redes sociales”. *omunicar*, núm. 33, pp. 73-81.
- Flores, J. y Arruti, A. (2001). *Ciberperiodismo: Nuevos enfoques, conceptos y profesiones emergentes en el mundo infodigital*. Madrid: Ediciones 2010.
- Fuentes, A. (1998). “Periodismo de Precisión: La vía socio informática de descubrir noticias”. Extraída el 19/II/2012 desde <http://www.saladeprensa.org/art04.htm>
- Kovach, B. y Rosenstiel, T. (2003). *Los elementos del periodismo*. Madrid: Aguilar.  
Levinson, P. (2009). *New New Media*. Nueva York: Person Education.
- Lohr, S. (2012). “The Age of Big Data”. Extraída el 20/II/2012 desde [http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?\\_r=1&ref=stevelohr](http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?_r=1&ref=stevelohr)
- Mancini, P. (2011). *Hackear el periodismo. Manual de laboratorio*. Buenos Aires: La Crujía.
- Meyer, P. (1991). *Precision Journalism*. Indianápolis: Indiana University Press.
- Open Knowledge Foundation (2010). *The Data Journalism Handbook*. Londres: European Journalism Centre.
- Ramírez, F. (2012, enero 11). “Periodismo de datos, periodismo de siempre”. Extraída el 10/IX/2013 desde <http://www.lapatria.com/blogs/periodismo-de-datos-periodismo-de-siempre>
- Ramonet, I. (2011). *La explosión del periodismo*. Buenos Aires: Clave Intelectual.
- Rogers, S. (2011). *Facts are Sacred: The power of data* [Kindle Edition]. Londres: Guardian Books.

## Páginas electrónicas

Hacks-Hackers Journalism & Techonology  
<http://hackshackers.com/>

Poynter Institute de Florida  
<http://www.poynter.org/>

Premio Datajournalism de Google  
<http://datajournalismawards.org/>

Storify  
<http://storify.com>