

Área Abierta. Revista de comunicación
 audiovisual y publicitaria
 ISSN: 1578-8393 / ISSNe: 1578-8393
<http://dx.doi.org/10.5209/ARAB.59521>

 EDICIONES
 COMPLUTENSE

Big Data: crisis y nuevos planteamientos en los flujos de comunicación de la cuarta revolución industrial

María José Revuelta Bayod¹

Recibido: 13 de marzo de 2018/ Aceptado: 19 de mayo de 2018

Resumen. El artículo trata de establecer unas nociones generales sobre la aplicación del *Big Data* a diversos ámbitos que interactúan con la disciplina de la comunicación en el contexto de la cuarta revolución industrial en la que estamos inmersos, de forma que podamos vislumbrar líneas de investigación necesarias tanto para la práctica profesional como investigadora. Nuestro cometido en el presente artículo es acotar este tema entre la multitud de aplicaciones y problemas que suscita e iniciar la formulación de hipótesis sobre el panorama actual. Nos centramos en la descripción y estudio de casos relacionados con el *Big Data* y su aplicación al periodismo de datos y al marketing electoral, para vislumbrar los cambios y las líneas de continuidad con tendencias que se habían gestado en el siglo pasado. Finalmente entendemos la necesidad de un enfoque interdisciplinar, más imprescindible que en anteriores periodos, además del apremio por adquirir nuevos conocimientos y nuevas prácticas profesionales.

Palabras clave: *Big Data*, cuarta revolución industrial, periodismo de datos, neuromarketing electoral

[en] Big Data: Crisis and New Approaches in the Communication Flows of the Fourth Industrial Revolution

Abstract. The article tries to establish some general notions about the application of big data to various fields that interact with the discipline of communication in the context of the fourth industrial revolution in which we are immersed, so that we can glimpse lines of research necessary for both professional practice as a researcher. Our task in this article is to limit this issue among the multitude of applications and problems that arise and start the formulation of hypotheses on the current scene. We focus on the description and study of cases related to big data and its application to data journalism and electoral marketing, to glimpse the changes and lines of continuity with trends that had developed in the last century. Finally, we understand the need for an interdisciplinary approach, more essential than in previous periods, in addition to the pressure to acquire new knowledge and new professional practices.

Key words: Big Data, Fourth Industrial Revolution, Data Journalism, Electoral Neuromarketing.

Sumario. 1. Introducción. 2. Axiomas, hipótesis y metodología. 3. *Big Data*, qué son, cómo se gestionan. 4. *Big Data* en el Periodismo. El Periodismo de datos. 5. Pronósticos y campañas electorales. Neuromarketing político. 6. Conclusiones. 7. Bibliografía

¹ Centro Universitario Villanueva (España)
 E-mail: mjrevuelta@gmail.com

Cómo citar: Revuelta Bayod, M.J. (2018) *Big Data: crisis y nuevos planteamientos en los flujos de comunicación de la cuarta revolución industrial*, en *Área Abierta. Revista de comunicación audiovisual publicitaria* 18 (2), 309-324. <http://dx.doi.org/10.5209/ARAB.59521>

1. Introducción

En el año 2016, el Foro de Davos, principal evento anual del Foro Económico Mundial (*World Economic Forum*), tuvo como tema central la cuarta revolución industrial. Klaus Schwab, fundador y director general de esta organización, señalaba en un artículo:

We stand on the brink of a technological revolution that will fundamentally alter the way we live, work, and relate to one another. In its scale, scope, and complexity, the transformation will be unlike anything humankind has experienced before (Schwab, 2016). [Nos encontramos al borde de una revolución tecnológica que alterará fundamentalmente la forma en que vivimos, trabajamos y nos relacionamos unos con otros. Es su escala, alcance y complejidad, la transformación será diferente a todo lo que la humanidad haya experimentado hasta ahora].

La primera revolución industrial, aparecida entre 1750 y 1830, se basó en la energía de vapor y el comienzo de la mecanización. La segunda se basó en las revoluciones en el campo de la electricidad y el motor de combustión (1870). La tercera se basó en los ordenadores, la tecnología de la comunicación e internet; comenzó en los años sesenta y alcanzó su punto álgido en el año 2000. Actualmente estamos en un momento de transición hacia la cuarta revolución industrial, que supone una combinación de sistemas digitales, biológicos y físicos, con importantes avances en robótica, genética y biotecnología, nanotecnología, inteligencia artificial e impresión 3D, entre otros. Se habla del fin de las energías fósiles pues existen las tecnologías necesarias para utilizar energías limpias. En círculos académicos, económicos y políticos se tratan temas como la urgencia de modificar el modelo económico partiendo de un cambio de valores donde prime la satisfacción de las necesidades básicas de todos los habitantes del planeta, de la necesidad de preservar la integridad del planeta, de maximizar el bienestar humano, de un modelo económico cíclico que reintegre los recursos en el sistema, de ciudades inteligentes donde se gestione una convivencia más humana y se prescindiera de las grandes cadenas de suministro... (WEF, 2016). Y una gran diferencia con las anteriores revoluciones es la rapidez de los cambios: la repercusión económica de la primera revolución industrial se percibió 100 años después, pero la celeridad de los avances en la actualidad hace que revolución económica y tecnológica vayan casi sincronizadas.

Las nuevas tecnologías, entendidas como la convergencia de sistemas digitales, físicos y biológicos, las usamos en multitud de aspectos de nuestra vida, pero también nos usan, puesto que en esta interacción se recopila gran cantidad de información que se encuentra en objetos que utilizamos en nuestras actividades cotidianas: tarjetas de crédito, GPS, teléfonos móviles, antenas de telefonía, contadores de luz y agua, internet... Las inversiones mundiales están dirigidas hacia proyectos de innovación tecnológica y en las mismas lo que prima es la recopilación masiva de datos, *big data*, procedentes de las interacciones de nuestro yo tecnológico, nuestra identidad digital, y que sobrepasa el tratamiento habitual de bases de datos.

El término inteligencia artificial surgió de una reunión de jóvenes investigadores organizada por John McCarthy en Dartmouth en el año 1956, donde se planteó el reto de simular mediante algoritmos determinados aspectos de la inteligencia. Desde este momento se perfila una nueva especialidad interdisciplinar donde debían converger física, biología, economía, psicología cognitiva y matemáticas. Unos años antes, en 1950, Alan Turing, matemático, informático teórico, criptógrafo y filósofo inglés, había propuesto en su ensayo *Computing Machinery and Intelligences*, el test de Turing, cuya finalidad era comprobar que si una máquina poseía inteligencia debía poder comunicarse de la misma forma que lo hacemos los humanos. Desde los años 50 hasta finales de los 80 la IA (Berlanga, 2016), tuvo su aplicación en sistemas donde la información y el conocimiento estaban estructurados y organizados con un sistema de reglas. Es el periodo de la IA simbólica con un modelo de razonamiento por deducción. Pero a finales de este periodo se vuelve la mirada hacia los años 60 para rescatar el aprendizaje automático (*Machine Learning*), un subcampo de las ciencias de la computación que estudia la formulación de algoritmos para explicar las regularidades que aparecen en un conjunto de datos, aplicando métodos inductivos y extrayendo de esta forma modelos generales. El avance de las nuevas tecnologías durante los años 90 hace surgir situaciones en las que el volumen de los datos sobrepasa las capacidades de almacenamiento en disco, lo que dificulta la extracción de conocimiento de esos datos. Será a partir del año 2010 cuando el interés sobre el *Big Data* se incrementa de forma exponencial, por la disminución del coste de almacenamiento y el surgimiento de nuevos campos de aplicación (*Smart Cities*, internet de las cosas, sistemas de reconocimiento facial y de voz, robots, mecanismos de algunos drones, videojuegos, *trending* de alta frecuencia para predecir las fluctuaciones de los mercados). Laboratorios de alta tecnología experimentan sobre sus aplicaciones en base a sistemas de redes neurales y microprocesadores cuánticos.

La IA se define como la “rama de las ciencias computacionales encargada de estudiar modelos cómputo capaces de realizar actividades propias de los seres humanos en base a dos de sus características primordiales: el razonamiento y la conducta” (López, 2015). Y un nuevo ámbito que se está desarrollando paralelamente con las ciencias cognitivas, es la relación de la IA con la conciencia y las emociones (Ortega, 2016).

2. Axiomas, hipótesis y metodología

La desregulación comunicativa iniciada en los años setenta en EEUU y en los ochenta en Europa, junto con el incremento en la velocidad en las transmisiones de datos, inician una carrera de concentraciones y sinergias que da lugar en el terreno de los mass media a la formación de grandes grupos multimedia con alcance mundial y que abarcan todos los medios de información tradicionales, internet y la cultura de masas. Así se creó el grupo Murdoch o la Fox TV. Pero además otras empresas comienzan a involucrarse en el sector de la comunicación, como compañías eléctricas- General Electric- compañías telefónicas- Telefónica española- pero también del sector bancario, seguros, etc. El avance de las nuevas tecnologías también fomentó la aparición de medios alternativos desvinculados de los intereses de estas grandes corporaciones, de forma que hay quien preconizaba cierto debilitamiento de esos

grandes emporios multimedia. ¿Es esto cierto? ¿Cómo ha influido el desarrollo de tecnologías que soportan *Big Data* en ello? ¿Cómo han reaccionado estos protagonistas de la esfera mediática mundial ante el avance tecnológico de la cuarta revolución industrial?

El proceso de toma de decisiones en el marketing ha ido incrementando exponencialmente sus interacciones con el consumidor y también con el no consumidor, en base a las informaciones proporcionadas por el yo tecnológico de sus públicos. Según la clasificación realizada por IBM (Mysore, Khupat y Jain, 2013), el problema del marketing es recoger los sentimientos del consumidor, datos éstos que son recogidos en la web y en las redes sociales. En este artículo expondremos casos de marketing político que ha utilizado como materia prima para sus estrategias comunicativas el *Big Data*.

La ciencia cognitiva estudia cómo la información es procesada y representada en la mente/cerebro. Es un ámbito interdisciplinar en el que convergen la psicobiología cognitiva, la lingüística, la inteligencia artificial, la neurociencia, la antropología cognitiva y la sociología cognitiva. En el ámbito de la comunicación de masas la alusión a la esfera emocional del receptor, fue adquiriendo mayor peso por su efectividad en el mensaje desde los inicios del siglo XX; pero la ciencia cognitiva ha demostrado con sus teorías las razones de esa efectividad y se ha profundizado la técnicas aplicables a diferentes ámbitos, demostrando que en nuestras emociones hay implícita una tendencia a actuar. Es el caso del neuromarketing.

En la aplicación de la metodología se ha recurrido al método inductivo a partir del estudio de casos escogidos por su relevancia mediática, además del recurso a fuentes bibliográficas y recursos web multimedia.

3. *Big Data*, qué son, cómo se gestionan

Esta ingente cantidad de datos que son almacenados, ha sido denominada *Big Data* y supone un valor añadido para la nueva economía. Los métodos tradicionales de bases de datos resultan ineficientes para gestionarlos y aparecen nuevas tecnologías de gestión. El término *Big Data* no sólo hace referencia al volumen de datos, sino a las tecnologías asociadas a la captación, administración y visibilidad de los mismos; y no sólo se tiene en cuenta el volumen sino la variedad de los mismos y la velocidad de acceso y procesamiento. En la recogida de datos en el análisis clásico de gestión de bases de datos, el proceso era plantear unas preguntas y realizar una recogida y análisis para responder a las cuestiones planteadas. Pero el *Big Data* recoge todo tipo de información, enormes cantidades de datos estructurados, semi-estructurados y no estructurados (señales de móvil, *emails*, *logs*, sensores, mensajes en redes sociales...) y en tiempo real, lo que lleva a realizar nuevas preguntas diferentes a las iniciales y poder explorar otros campos.

Comenzó a hablarse de las cuatro V del *Big Data* – volumen, variedad, velocidad y veracidad- pero se han establecido tres más: viabilidad, visualización y valor de los datos (IIC, 2016):

Volumen: se refiere a la cantidad de datos que son generados y almacenados y que tiene como finalidad transformarlos en acciones útiles para las diferentes entidades que los manejan. Cuando se habla de *Big Data*, no se refiere a una cantidad específica pero se suele utilizar cuando se trata

con términos de petabytes y exabytes.

Gigabyte = 10^9 Terabyte = 10^{12} Petabyte = 10^{15} Exabyte = 10^{18}

Variación: Se refiere a la naturaleza de los datos, esto es, a las formas, tipos y fuentes en las que se registran los datos. Las bases de datos tradicionales gestionaban datos estructurados (en filas y columnas), pero el *Big Data* añade a éstos, los semi-estructurados (*software*, informes...) y no estructurados, donde caben todo tipo de documentos de texto, vídeos e imágenes generados por las redes sociales, mails, foros y chats... Los mismos son generados en todos nuestros dispositivos: teléfonos móviles, vídeo, sistemas GPS, sensores digitales, medidores eléctricos...

Velocidad: es la rapidez que exige el procesamiento de esa información por las constantes interconexiones que se generan. Hay procesos que exigen que la monitorización y el tratamiento de la información se realice en tiempo real.

Veracidad: hace referencia al grado de credibilidad, a la fiabilidad de la información.

Viabilidad: las entidades deben ser realistas en las necesidades técnicas y humanas necesarias para proporcionar un uso eficaz de los datos, estableciendo recursos personalizados en cada proyecto.

Visualización de los datos: es imprescindible la utilización de herramientas de visualización, para poder llegar a una comprensión de los datos.

Valor de los datos: los datos deben transformarse en información que sea útil para la toma de decisiones de una entidad. El procesamiento de estos datos por lo tanto, debe estar adecuadamente dirigido.

Un proceso *Big Data* se reparte en las fases de recopilación, almacenamiento, procesamiento y análisis de los datos. Lo primero que se debe tener claro para emprender este tipo de proyecto, no es qué información debo analizar, sino ¿qué problema estoy tratando de resolver? Y para ello se puede comenzar por establecer una clasificación de los tipos de datos de *big data* (Barranco, 2012):

1. *Web and Social Media:* se refiere a los contenidos de las redes sociales, Facebook, Twitter, LinkedIn, etc.
2. *Machine to Machine:* son las tecnologías que conectan diferentes dispositivos, para lo que se utilizan sensores o medidores que captan alguna variable (presión, velocidad, temperatura...). Estos datos se transmiten a otras aplicaciones para transformarlos en información.
3. *Big Transaction Data:* se trata de datos estructurados y semi-estructurados referentes al almacenamiento de los registros de facturación y llamadas (CDR).
4. *Biometrics:* Información biométrica como escaneo de la retina, huellas digitales, reconocimiento facial, genética, etc. Muy importantes para las agencias de investigación.
5. *Human Generated:* son los contenidos generados por el ser humano, tales como archivos y correos electrónicos, presentaciones, documentos de texto, archivos de audio...; pero también los metadatos sobre ese archivo: quién lo ha leído, tipo de archivo, quién tiene acceso a él, en qué carpeta está...

En función del tipo de entidad y de los intereses generales y específicos que hayan motivado un proceso *Big Data*, se utilizarán diferentes tecnologías para almacenar,

procesar y gestionar. Las tecnologías son el *software* que captura, transforma, procesa y analiza los datos y que pueden ser tanto libres como propietarias. Algunas de estas son (Hernández, Duque y Moreno, 2017):

Hadoop: librería de Apache.

MapReduce: utilizado inicialmente por Google para solucionar el problema de *Page Rank*.

HBase: también desarrollada por Apache y que se utiliza para acceder a lectura y escritura de datos en tiempo real.

Cassandra: también de Apache y utilizada para tratar problemas de datos críticos en el hardware o en la nube. Twitter es una de las empresas que la utiliza dentro de su plataforma.

Mahout: también de Apache, para la creación de aplicaciones de *Machine Learning*.

Jaql: donado por IBM a la comunidad de software libre

Lucene: proyecto de Apache que realiza búsquedas sobre textos

La funcionalidad de estos datos viene cuando se extrae valor de los mismos y es donde entra el concepto de ciencia de los datos o *Data Science*. Este campo interdisciplinar, que no se aplica únicamente a grandes cantidades de datos, involucra métodos científicos, procesos y modelos para realizar una extracción del valor de los datos. Incluye técnicas como la analítica descriptiva, la estadística, la minería de datos o *Data Mining* y el aprendizaje automático o *Machine Learning* (Hernández, Duque y Moreno, 2017).

En los siguientes epígrafes especificaremos en algunos de los casos las técnicas utilizadas para el tratamiento de *Big Data*.

4. Big Data en el Periodismo. El Periodismo de datos

En 1988 el periodista Bill Deadman recibió el premio *Pulitzer* por demostrar criterios racistas en la concesión de créditos para la adquisición de viviendas en Atlanta. Para recopilar datos utilizó el análisis de bases de datos, iniciando el periodismo asistido por ordenadores (*CAR Computer Assisted Reporting*); éste sería el antecedente de lo que se ha denominado periodismo de datos, que es el que se realiza con las herramientas del *Big Data* y *Open Data*.

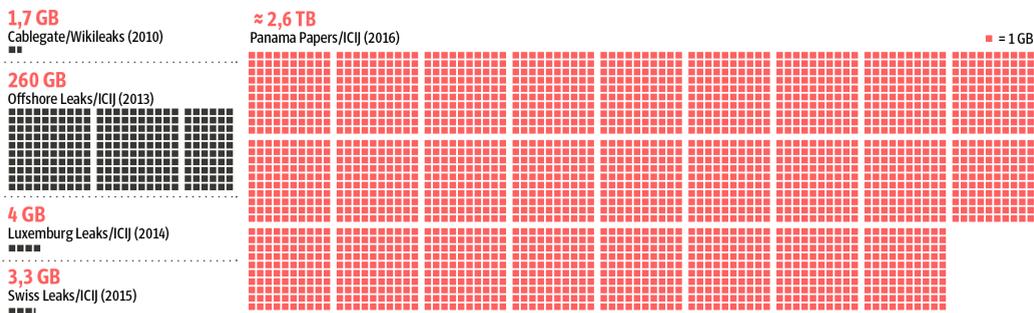
El periódico *The Guardian* fue pionero en la introducción del *Big Data*, con su *Datablog*, una sección que aparecía en su versión online y donde se publicaban únicamente noticias y reportajes basados en datos. El portal *La Nación* de Argentina, el *New York Times* con su iniciativa *The Upshot* o *Le Monde*, también se encuentra entre los primeros en ofrecer datos e investigaciones basados en grandes visualizaciones de datos. En España existe el grupo de *Periodismo de Datos Meetup* que lleva en funcionamiento desde octubre de 2011, creado para promover esta disciplina en nuestro país. Uno de los objetivos del grupo es la formación de periodistas en este campo, pues es necesaria la adquisición de “nuevas capacidades para ser capaces de leer, combinar y extraer información de las bases de datos” (MediaLabPrado, 2011)

Los casos más importantes del periodismo de datos han ido sucediendo: la filtración de Hervé Falciani (2008), conocida como *la lista Falciani*, de los nombres de unos 130.000 potenciales evasores fiscales de varios países de la Unión Europea. Julian Assange con *Wikileaks*, en el año 2010, con la publicación de los documentos secretos del Pentágono y del Departamento de Estado, que fueron obtenidos por la soldado Chelsea Manning, hoy en prisión, y que trataba sobre las muertes de civiles en la guerra de Irak. Edward Snowden que denunció el programa de cibervigilancia masiva *Prism* que realizaban las agencias de seguridad de inteligencia de Estados Unidos en colaboración con otros países aliados².

El caso de los *Papeles de Panamá*, supone un hito en el periodismo de datos por considerarse la mayor filtración periodística de la historia. La filtración comenzó en la primavera de 2015 con una llamada anónima a un periodista de investigación del diario alemán *Süddeutsche Zeitung* y, siguiendo el ejemplo de los dos periodistas de *Le Monde* con el caso de la lista Falciani, decidieron compartirlo con el Consorcio Internacional de Periodistas de Investigación (ICIJ, *International Consortium of Investigative Journalist*). Esta organización es una red internacional con más de 190 periodistas de investigación de más de 95 países y fue fundada en 1997 por el periodista estadounidense Chuck Lewis. Cuenta con experiencia en manejar *big data* y realizar investigaciones transnacionales, además de experiencia y conocimiento sobre paraísos fiscales por haber realizado desde 2013 “cuatro investigaciones vinculadas a ese tema: *Secrecy for Sale: Inside the Global Offshore Money Maze*, *China Leaks*, *Luxembourg Leaks* y *Swiss Leaks*” (García Rey, 2016)

The scale of the leak

Volume of data compared to previous leaks



Escala de los “Papeles de Panamá”. [Fuente: http://www.alternet.org/files/screen_shot_2016-04-04_at_12.01.06_pm.png]

Se manejaron 2.6 terabytes y 11.5 millones de documentos y fueron presentados a partir de mayo de 2016. Durante un año más de 370 periodistas de 107 medios de comunicación de 76 países se asociaron para descifrar todo el material. Para la

2 Este caso quedaría reflejado en el documental ganador de un Oscar en 2014, *Citizenfour*, dirigido por Laura Poitras

digitalización de esos millones de documentos se utilizó el reconocimiento óptico de caracteres (OCR, *optical carácter recognition*). En la realización de búsquedas a gran escala se utilizó *Apache Solr*, a través de la interface *Blacklight Project*. Posteriormente estos documentos “fueron estructurados después en un esquema de relaciones (tipo nodo-arista) para crear una base de datos orientada a grafos usando la tecnología *Neo4j*, lo que finalmente permitió hacer uso de técnicas de análisis de *big data* para encontrar las relaciones entre individuos y datos financieros ...” (Arcilla, Barbosa y Cabezuelo, 2016).

La investigación supuso un duro golpe al sistema financiero mundial y mostraba la actividad del despacho de abogados panameño Mossak, desvelando que miles de personas ocultaban su patrimonio en sociedades situadas en paraísos fiscales, cómo se utilizan empresas *offshore* para la financiación de grupos terroristas, y las repercusiones políticas y sociales que suponen estas acciones. Se abrió entonces el debate sobre la necesidad de adoptar nuevas medidas de control financiero coordinadas internacionalmente. Hasta aquí una victoria del cuarto poder ejerciendo su función, pero existen datos que dieron lugar a otras interpretaciones.

Una de las críticas recibidas se debía a la aparición en las portadas de muchos periódicos de la imagen del presidente ruso, Vladimir Putin, cuando su nombre no aparecía en la documentación filtrada (Colás, 2016; Bernstein, Blum, Zihlmann y Thompson, 2016; Roderick, 2016; Harding, 2016)). Por esta razón, se atribuyeron intereses geopolíticos en la investigación, debido a la postura del dirigente en Ucrania o en Siria, donde son más que conocidas las tensiones entre Rusia y la OTAN (Murray³, 2016; Zegada, 2016 periódico *El País*).

El hecho de que no aparecieran nombres estadounidenses en la lista, introdujo la sospecha de que el escándalo financiero tenía como objetivo redirigir los grandes flujos financieros desde paraísos fiscales hacia EEUU. Los ciudadanos estadounidenses ricos que se sirven de paraísos fiscales, eligen las Bermudas, las Islas Caimán o las Islas Vírgenes Británicas, o dentro de su propio país en estados como Delaware y Nevada (*RT*, 2016; Zegada, 2016).

Al contrario que hace *Wikileaks* con sus filtraciones, la información no se publicó completa y únicamente fue accesible a grandes medios de comunicación de investigación. Los periodistas del diario alemán *Süddeutsche Zeitung*, comparten su filtración con el ICIJ, por su experiencia en manejar *big data* y su metodología de trabajo:

Jerarquía mínima para garantizar una colaboración eficaz y construir confianza [...].

La investigación necesita del periodismo de datos [...].

Plataforma de colaboración para generar conocimiento y compartirlo [...].

Seguridad [...].

Memorias intermedias y planificación de la publicación [...].

Diligencia informativa obligada [...].

Cuatro filtros de calidad de la información [...].

Publicación colegiada [...].

Transparencia. El ICIJ tiene una vocación periodística de servicio público. ...La publicación de todo el conjunto de archivos podría dañar la honorabilidad y violar la privacidad de mucha gente que no tiene un interés público, además de provocar muchos conflictos legales a los medios socios de la

3 Craig Murray, ex diplomático británico y activista de derechos humanos.

investigación. Como dijo el director del ICJI, Gerard Ryle, a *Wired*: ‘No somos *WikiLeaks*. Estamos intentando mostrar que el periodismo no puede ejercerse de forma responsable’. Y por último y no menos importante, esa base de datos liberada es un bonus terapéutico para conspiranoicos. (García Rey, 2016)

A esa última alusión de Gerard Ryle, en la revista *Wire*, respondió *Wikileaks* al día siguiente en su cuenta de twitter: “Washington DC based Ford, Soros funded soft-power tax-dodge, “ICIJ” has a WikiLeaks problema # PanamaPapers”⁴ acusando a ICJI de estar al servicio de los intereses estadounidenses en todo el mundo. El Consorcio Internacional de Periodistas de Investigación (ICIJ, *International Consortium of Investigative Journalist*⁵), está financiado por el *Center for Public Integrity* de EEUU, la *Ford Foundation*, el *Carnegie Endowment*, el *Rockefeller Family Fund*, la *WK Kellogg Foundation* y la *Open Society Foundation* (financiada por George Soros). Itimo y no menos importante, esa base de datos liberada a periodismo no puede ejercerse de forma responsable “praciones del president

La organización Attac⁶, afirma sobre este acontecimiento que fue una operación orquestada por intereses de los gobiernos para mantener la presión fiscal sobre la ciudadanía mediante la exposición de casos de personalidades y organizaciones, intereses geoestratégicos y económicos concretos (redirigir el flujo de capitales hacia EEUU y deslegitimar al dirigente ruso en el caso de Siria y Ucrania) y una nueva reconfiguración del Nuevo Orden Mundial con el establecimiento de nuevas instituciones globales que monitoreen con facilidad el dinero electrónico en detrimento del físico (Nadal, Piketty y Roberts, 2016).

Por otra parte, el gobierno de Panamá convocó a un comité de expertos liderado por el premio nobel de economía Joseph Stiglitz⁷, pero él y su compañero suizo Mark Pieth renunciaron meses después de su nombramiento por falta de transparencia y “poca voluntad para acometer los cambios que fueran necesarios” (Efe, 2016).

Independientemente de estos casos, y del debate que pueda suscitar el papel del periodista como *gatekeeper*, está claro que el perfil del profesional está cambiando, pues es necesaria una formación, además de en periodismo, en tecnología de la información. Se ha constituido así, un método de trabajo con fases de extracción, limpieza, procesamiento y visualización de los datos, para lo cual es también necesario el conocimiento de portales de *Open Data*, que son la apertura de fuentes públicas para el acceso libre a esa información: el *Portal de datos de Naciones Unidas*, *Portal Open Data de la UE*, *Ask The EU*, *Datos.gob.es* (datos del gobierno español), *Data Eukadi*, *Open Corporate*... un conocimiento de herramientas de extracción de datos como *Scraperwiki* o *Pentaho Data* y el manejo de herramientas de visualización como *Linkurious*. No se trata de que los periodistas aprendan a programar, pero sí que entiendan que los datos masivos contienen historias; “y que puedan tomar datos y una idea y con la ayuda de un experto (de la misma manera que trabajan con fotógrafos y artistas gráficos ahora) puedan contar esa historia” (Gonzalo, 2013).

4 Cuenta de Twitter de WikiLeaks. Recuperado de <https://twitter.com/wikileaks/status/717454569861611520>

5 <http://www.icij.org>

6 Attac (Asociación por la Tasación de las Transacciones financieras y por la Acción Ciudadana) es un movimiento independiente internacional y autofinanciado. Promueve entre otros objetivos, el control democrático de los mercados financieros. www.attac.org

7 Economista y profesor estadounidense, premio nobel de Economía (2001). Desempeñó labores en la administración Clinton, en el Banco Mundial como vicepresidente y economista jefe.

Empresas como Google, Facebook, Amazon, han cimentado su estrategia empresarial sobre el *Big Data*, y también han estado bajo el punta de mira del periodismo:

- El periodista francés, Jean-Baptiste Malet publicó en el año 2013 *En los dominios de Amazon* (Trama). Inicialmente comenzó con un reportaje infiltrándose como trabajador en los almacenes de Montelimar, pero terminó convirtiéndose en un libro donde relata la explotación laboral de los trabajadores de *Amazon* (Corroto, 2016). El libro se vendía por *Amazon*. En agosto de 2015 *The New York Times* (Kantor y Streitfeld, 2015) publicó un artículo donde relata las duras condiciones de trabajo de los trabajadores de cuello blanco, esta vez en las oficinas de la compañía de Seattle, aplicando lo que denominan darwinismo intencional. En el año 2013 Jeff Bezos, el dueño de Amazon había comprado el *Washington Post*, y el resultado cinco años después es una reestructuración del periódico poniendo el mayor énfasis en las nuevas tecnologías, con unos ingresos publicitarios en internet en más de cien millones de dólares y toda la financiación necesaria para sus investigaciones periodísticas. Así lo declara Bob Woodward, uno de los periodistas que destapó el *Watergate* junto a Carl Bernstein y que, tras 45 años continúa en la plantilla del periódico (Cano, 2017).
- Sobre Google han planeado sospechas de haber pactado con el gobierno chino para restringir los servicios y contenidos en ese país, cediendo así a la censura que el gobierno ejerce sobre los internautas chinos, limitando el acceso a determinados contenidos. Igualmente, Microsoft, Yahoo y Cisco parece que cedieron a las presiones del gobierno chino (EFE, 2009; Agencias, 2006; Agencias, 2006). También fue señalada, junto a otras compañías de nuevas tecnologías, como una de las involucradas en el programa *Prism*, orquestado desde la Agencia de Seguridad Nacional (NSA) de EEUU.

5. Pronósticos y campañas electorales. Neuromarketing político

Desde hace años las predicciones electorales clásicas que publican los grandes medios de comunicación parecen fallar. En el 2012 fallaron los sondeos que pronosticaban la igualdad entre Obama y el candidato republicano Matt Romney, el *Brexit* o la victoria de Trump.

Durante la campaña de 2008, Obama comienza a desarrollar parte de su estrategia electoral basándose en el *Big Data*, hecho que terminaría de consolidar en 2012. En la primera campaña presidencial empleó a ocho ingenieros que trabajaban con datos almacenados en servidores y en algún momento con la nube de Amazon. Se rodeó de un equipo de científicos de datos capitaneados por Dan Wagner, hoy CEO de *Civis*. Se había diseñado una *app* para la campaña de Obama y se cruzaban datos con *Open Data*, y, a grandes rasgos, se consiguió de esta forma obtener información sobre los indecisos (revistas que leían, amigos de Facebook...) con la colaboración de 300.000

voluntarios que realizaban preguntas previamente diseñadas hacia este público que vacilaba en la decisión sobre su voto. En el año 2012 el manager de campaña, Jim Messina, vuelve a incidir en la necesidad de apoyarse en el *Big Data*. Para ello se valieron de acciones offline, como encuestas telefónicas y campañas puerta a puerta y acciones online. Toda esta información les llevó a realizar mejoras en sus propuestas, basándose en las opiniones de la gente y personalizando los argumentos para los diferentes tipos de votantes, centrándose nuevamente en los indecisos. Esta misma estrategia basada en *Big Data* es la que siguió la consultora de Jim Mesina en las elecciones del año 2016 en España para hacer remontar al PP, pasando de 123 a 137 escaños. El presupuesto era ajustado y se reducen las partidas de publicidad para tener más recursos para Facebook y focalizarse en los votantes indecisos del partido de Ciudadanos; desde TMG lo conocen como “lograr el voto táctico” mediante una estrategia de *microtargeting* para conectar directamente con los potenciales votantes y Facebook era la red social en la que focalizarse, puesto que desde enero de 2016, seis millones de personas habían generado 90 millones de interacciones sobre temas políticos (Mucha y Negre, 2016).

Civis Analytics, *Resonate*, *Nationbuilder* y *Quid*, son algunas de las empresas de análisis de datos y que venden a partidos políticos (Kuchler, H. 2016). Se segmenta a los votantes y luego se les lanzan mensajes digitales y por tv para solicitar su voto. Se realizan miles de encuestas al mes y luego se completan sus datos a través del seguimiento de las *cookies* de los participantes en las encuestas.

Trump realizó una inversión mínima en publicidad pero millonaria en *Big Data* y *Data Science*. El equipo estaba dirigido por Steve Bannon, miembro del consejo de administración de *Cambridge Analytica*, y llegaron a manejar 250 millones de perfiles con 5.000 datos por perfil (López, 2017), basado en el desarrollo del *text mining*, el *análisis del sentimiento* y el NLP (*natural language processin*). Al igual que en los anteriores casos expuestos el objetivo es extraer información subjetiva del *microtarget*, identificando el sentimiento hacia el candidato, su polaridad e intensidad, de forma que puedan construirse los mensajes pertinentes para que sea efectiva la persuasión.

En otro sentido, se constata el transvase bidireccional de recursos humanos e información entre grandes compañías y partidos políticos: después de la victoria de Obama en 2012, su ex directora de análisis, Carol Davidsen, fue contratada como vicepresidenta de tecnología en la empresa *Rendtrak*, un servicio de medición de medios. El director de análisis de la campaña, Dan Wagner, creó *Civis Analytics*, uno de cuyos principales inversores es Eric Schmidt, ex presidente ejecutivo de Google y desde agosto de 2015, presidente de *Alphabet Inc.* después de la reorganización de Google. Este último se incorporó a la compañía política de Hillary Clinton (Huchler H. , 2016).

6. Conclusiones

Sin pretender caer en la distopía, hemos percibido que el hecho de que las implicaciones de los avances en tecnología sobre la economía se produzcan en un breve periodo de tiempo, presenta una serie de consecuencias:

El periodismo de datos necesita más recursos económicos para financiarse cuando se trata de *Big Data*, y en el caso del Consorcio Internacional de Periodistas

de Investigación (ICIT, *International Consortium of Investigative Journalist*) su vinculación con fundaciones de grandes corporaciones y el filtro informativo a la hora de publicar sus informaciones, levanta sospechas y da lugar a interpretaciones sobre la finalidad de su periodismo. Recibió el premio *Pulitzer* en el año 2017. Por otra parte, Julien Assange, programador, ciberactivista, periodista, fundador y miembro de *Wikileaks*, que también ha sido galardonado con varios premios y condecoraciones⁸, permanece aislado en la embajada de Ecuador en Londres desde el año 2012. La razón de ello fue la acusación por parte de la justicia sueca de dos presuntos delitos sexuales⁹, pero lo que temían los defensores de Assange es que desde Suecia se le extraditase a los Estados Unidos, donde ha sido acusado de delitos de espionaje y traición y podría condenársele a la pena de muerte.

Aparecen nuevos actores en la escena. Empresas como Amazon, Google o Facebook no habrían alcanzado el éxito sin el *Big Data*. Han estado en el punto de mira del periodismo, pero este hecho no han tenido consecuencias en su actividad empresarial:

- Amazon ha levantado económicamente a uno de los diarios más prestigiosos del planeta, y las reacciones ante las condiciones laborales en sus almacenes y sus oficinas no se han percibido. Los ciudadanos parece que están dispuestos a pagar el “coste social” a cambio de un servicio competente.
- Los internautas continuamos utilizando el servicio de Google, a pesar de la filtración sobre su colaboración en el programa *Prism* de control global de las multitudes que denunció Snowden.

Los medios alternativos que surgieron con las nuevas tecnologías en el siglo pasado y que podían ejercer un contrapeso, utilizan las redes de Google, Facebook y Twitter.

Las grandes multinacionales gestadas en el pasado siglo se sienten amenazados por los nuevos actores digitales. De hecho, se habla en Europa y en EEUU de controlar a estos nuevos gigantes digitales de Silicon Valley, que controlan *Big Data* y concentran una gran proporción de la publicidad en las redes sociales, mediante leyes antimonopolio (Smith E. 2018). Mientras en Estados Unidos se está negociando la posible regulación del duopolio digital de Google y Facebook, los grandes emporios mediáticos se han metido en una carrera de adquisición de *star-up* digitales para acelerar la innovación tecnológica en sus corporaciones. Y las más activas son “...las empresas financieras, de medios de comunicación e industria, puesto que acaparan el 77% de las inversiones” (Prieto M. 2017).

8 Premios Amnistía Internacional de los Medios Británicos en 2009; en 2008 premio del *Index on Censorship* de la revista *The Economist* (fuente: *Wikipedia*)

9 La Fiscalía de Suecia cerró la causa en mayo de 2017 y el ministerio público sueco retiró los cargos por acoso sexual en 2015. La policía británica mantiene vigente la orden de detención por no presentarse ante la corte de Westminster en 2012 y se teme su extradición a EEUU. El grupo de Trabajo de Naciones Unidas sobre Detenciones Arbitrarias lo ha calificado en el año 2016 de “detención arbitraria”. La justicia británica mantiene la orden de detención. Desde EEUU se le ha declarado “enemigo público número 1” por las últimas filtraciones sobre la CIA. Fuente: El País. Recuperado de https://elpais.com/tag/julian_assange/a

No es una novedad que en las campañas electorales se aplican técnicas procedentes de la comunicación audiovisual, la publicidad y el marketing y que se apela sobre todo al ámbito emocional de los potenciales electores para inducir un determinado comportamiento a la hora de decidir su voto. Con la aplicación de las ciencias cognitivas al marketing se consigue ofrecer una explicación sobre la efectividad de estos mensajes, dando lugar al neuromarketing; pero además, con la incorporación del *Big Data*, el diseño de una campaña electoral se convierte en una sofisticada herramienta tecnológica capaz de personalizar cada vez más los mensajes, analizando sentimientos y actitudes de los potenciales electores a través de las redes sociales que suministran datos estructurados y no estructurados. Datos suministrados por nuestro yo tecnológico, a los que se les aplican algoritmos que desentrañan aquello que parece que ocultamos en las encuestas oficiales. Los resultados muestran la victoria de Obama en dos campañas electorales, la de Rajoy en las últimas elecciones o el triunfo de Trump. Los cuatro acontecimientos en contra de los pronósticos publicados en los periódicos.

No pretendemos caer en una visión apocalíptica, puesto que esta cuarta revolución industrial también está mostrando importantes avances positivos, pero a las anteriores conclusiones podemos añadir que existe un preocupante aumento de las desigualdades en los países ricos según el informe de la OCDE¹⁰. En EEUU, *Axiom* empresa líder en captación y análisis de datos, clasifica a los consumidores en 70 grupos socioeconómicos distintos y aquí aparecen los *waste* (basura); son los offline, los que no existen porque carecen de yo tecnológico (Singer, N. 2012; Arnaz, 2016).

Estas circunstancias nos hacen considerar que necesitamos una interpretación humanista que establezca una transversalidad entre ciencias y tecnología y ciencias sociales y por supuesto filosofía y ética¹¹. Y en nuestro ámbito de conocimiento, una actualización de conocimientos que no vuelvan a hacer perder en nuestro país, 350.000 puestos de trabajo por falta de formación:

En España hay más de 350.000 puestos de trabajo sin cubrir debido a la falta de competencias digitales. Dice César Alierta que nuestro país en 2020 necesitará 3 millones de jóvenes con formación digital, mientras que las universidades actualmente no se adaptan a estas exigencias. Los perfiles demandantes están relacionados con el análisis de big data, la famosa ciberseguridad, el comercio electrónico, internet de las cosas, aplicaciones móviles, comercio electrónico y el marketing digital (*20 Minutos*, 2017: 21/05/2017)

Este texto surge dentro del marco del Proyecto de Investigación UCM-Santander. Referencia: PR26/16-6B-1, titulado: PERLAD, *Pensamiento y representación literaria y artística digital ante la crisis de Europa y el Mediterráneo*.

10 *Understanding The socio-economic divide in Europe*. Background report. 26 January 2017. Recuperado de <https://www.oecd.org/els/soc/cope-divide-europe-2017-background-report.pdf>

11 El Gobierno de España ha convocado en 2017, un Grupo de Sabios con expertos del mundo académico, institucional y empresarial para que elabore un Libro Blanco sobre IA y *big data*. El objetivo es potenciar estos dos ámbitos en España. El informe abarcará aspectos económicos, sociales y éticos y está prevista su publicación en primavera de 2018 (EuropaPress, 2017).

7. Bibliografía

- Agencias. (25/06/2009). “El gobierno chino bloque Google”. *ABC*. Recuperado de http://www.abc.es/hemeroteca/historico-25-06-2009/abc/Internacional/el-gobierno-chino-bloquea-google_922024713781.html
- Agencias. (02/02/2006). “Críticas a Google, Yahoo, Microsoft y Cisco por ceder a la censura de Pekín”. *El País*. Recuperado de https://elpais.com/tecnologia/2006/02/02/actualidad/1138872479_850215.html
- Arcilla-Calderón Carlos, Barbosa-Caro Eduar y Cabezuelo-Lorenzo Francisco. “Técnicas de Big Data: análisis de textos a gran escala para la investigación científica y periodística”. *El profesional de la información*, julio-agosto 2016, v.25, no.4 pp 623-631
- Arnaz, R. (21/02/2016). “Descubre Axiom, la empresa de la que nunca has oído hablar pero que lo sabe todo sobre ti y vende la información a multinacionales”. *La Información*. Recuperado de: https://www.lainformacion.com/tecnologia/economia-negocios-y-finanzas/descubre-axiom-la-empresa-de-la-que-nunca-has-oido-hablar-pero-que-lo-sabe-todo-sobre-ti-y-vende-la-informacion-a-multinacionales_paekjpcchoapsj3ukfzomy5
- Barranco Frangoso, Ricardo. “¿Qué es el Big Data?”. (2012). *IBM developerWorks*. Recuperado de <https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/> (Fecha de acceso: 15/12/2017)
- Berlanga, Antonio. (2016). “El camino desde la Inteligencia Artificial al Big Data”. *Revista Índice*, julio 2016, n.68. Recuperado de: <http://www.revistaindice.com/numero68/p9.pdf> (Fecha de acceso: 30/12/2017)
- Bernsteni, J. Blum P. Zihlmann O. Thompson D. (03/04/2016). “Todos los hombres de Putin. La red de dinero ‘offshore’ conectada al presidente ruso”. *El Confidencial*. Recuperado de https://www.elconfidencial.com/economia/papeles-panama/2016-04-03/papeles-panama-paper-red-dinero-putin-lider-ruso-mossack-fonseca_1177787/
- Cano, F. (18/06/2017). “¿Ha salvado Jeff Brezos al Washington Post?”. *El Español*. Recuperado de https://www.elespanol.com/economia/empresas/20170616/224228288_0.html
- Colás, X. (03/04/2016). “Vladimir Putin, implicado en los Papeles de Panamá”. *El Mundo*. Recuperado de <http://www.elmundo.es/internacional/2016/04/03/57017a22e2704ee6748b45e5.html>
- Corroto, P. (14/10/2013). “Jean- Baptiste Malet: ‘Amazon no respeta las reglas del capitalismo’”. *El diario.es*. Recuperado de: https://www.eldiario.es/turing/Jean-Baptiste-Malet-Amazon-capitalismo_0_185831480.html
- EFE, Panamá (06/08/2016). “Stiglitz renuncia al comité panameño de expertos creado tras los papeles de Panamá”. *La Vanguardia*. Recuperado de: <http://www.lavanguardia.com/economia/20160806/403726860215/stiglitz-renuncia-comite-papeles-panama.html>
- EFE. (25/06/2009). “El gobierno chino bloque Google”. *ABC*. Recuperado de http://www.abc.es/hemeroteca/historico-25-06-2009/abc/Internacional/el-gobierno-chino-bloquea-google_922024713781.html
- García Rey, Marcos. (2016). “Los papeles de Panamá: su intrahistoria periodística”. *Cuadernos de periodistas: revista de la Asociación de la Prensa de Madrid*, no.32, pp. 50-60
- Gonzalo, M. (05/08/2013). “Los datos masivos (o big data) son el nuevo oro” Entrevista a Viktor Mayer-Schönberger. *Diarioturing*. Recuperado de https://www.eldiario.es/turing/Big-data_0_161334397.html
- Harding, L. (03/04/2016). “Revealed: the 2bn dolars offshore trail that leads to Vladimir

- Putin”. *The Guardian*. Recuperado de: <https://www.theguardian.com/news/2016/apr/03/panama-papers-money-hidden-offshore>
- Hernández-Leal Emiley, Duque-Méndez Néstor y Moreno-Cadavid, Julián. “Big Data: una exploración de investigaciones, tecnologías y casos de aplicación”, *Tecnológicas*, vol.20, no.39, mayo-agosto, 2017.
- Huchler, H (Financial Times). (08/06/2016). “El big data entre en campaña”. *Expansión*. Recuperado de <http://www.expansion.com/economia-digital/innovacion/2016/06/08/575566f9ca4741b91e8b4670.html>
- Kantor J. y Streitfeld D. (15/08/2015). “Inside Amazon: Wrestling Big Ideas in a Bruising Workplace”. *The New York Times*. Recuperado de: <https://www.nytimes.com/2015/08/16/technology/inside-amazon-wrestling-big-ideas-in-a-bruising-workplace.html>
- “Las 7 V del Big Data: Características más importantes”.(2016). *Instituto de Ingeniería del conocimiento*. Recuperado de: <http://www.iic.uam.es/innovacion/big-data-caracteristicas-mas-importantes-7-v/> (Fecha de acceso: 15/12/2017)
- “Infografía Big Data: las 7 V”. (2016). *Instituto de Ingeniería del conocimiento*. Recuperado de: <http://www.iic.uam.es/innovacion/big-data-infografia-7-v/>(Fecha de acceso: 15/12/2017)
- López Takeyas, Bruno. (2015). “Introducción a la inteligencia artificial”. *Instituto Tecnológico de Nuevo Laredo*. Recuperado de <http://www.itnuevolaredo.edu.mx/takeyas> (Fecha de acceso: 15/01/2018)
- López Zafra, J.M. (22/02/2017). “Política 2.0: El rastro digital es la clave”. *El Confidencial*. Recuperado de: https://blogs.elconfidencial.com/economia/big-data/2017-02-22/politica-demoscopia-encuestas-rastro-digital_1336020/
- “El gobierno crea un Grupo de Sabios para elaborar un libro blanco sobre Inteligencia Artificial y Big Data” *EuropaPress*. (14/11/2017). Recuperado de: <http://www.europapress.es/economia/noticia-gobierno-crea-grupo-sabios-elaborar-libro-blanco-inteligencia-artificial-big-data-20171114133920.html>. (Fecha de acceso: 15/01/2018)
- MediaLab Prado. “Seminario sobre Periodismo de Datos” 2011. Recuperado de: <https://www.medialab-prado.es/actividades/seminario-sobre-periodismo-de-datos> (Fecha de acceso: 01/02/2018)
- Murray, Craig. “Corporate Media Gatekeepers Protect Western 1% from Panama Leak” .(03/04/2016). *Blog Craig Murray*. Recuperado de <https://www.craigmurray.org.uk/archives/2016/04/corporate-media-gatekeepers-protect-western-1-from-panama-leak/>
- Muncha M. y Negre, J. (03/07/2016). “La gurú de San Francisco que hizo ganar las elecciones a Mariano Rajoy”. *El Mundo*. Recuperado de <http://www.elmundo.es/cronica/2016/07/03/57779fc0ca4741301d8b4609.html>
- Mysore Divakar, Khupat Shirkant, y Jain Shweta. “Big data y analytics”. (2013). *IBM DevelopersWorks*. Recuperado de: <https://www.ibm.com/developerworks/ssa/library/bd-archpatterns1/index.html>. (Fecha de acceso: 20/01/2018)
- Nadal A., Piketty T. y Roberts, M. “Los ‘Papeles de Panamá’: el papel de los paraísos fiscales en el capitalismo actual. Dossier”. (28/04/2016). *Attac España*. Recuperado de <https://www.attac.es/2016/04/28/los-papeles-de-panama-el-papel-de-los-para-isos-fiscales-en-el-capitalismo-actual-dossier/>
- Ortega, Andrés. (2016). “Eros, máquinas y emociones”. *La imparable marcha de los robots*. Madrid: Alianza, pp. 33-57
- ‘Papeles de Panamá: “El escándalo de las ‘offshores’ redirigirá billones de dólares a EEUU’’. (05/04/2016). *RT Noticias*. Recuperado de <https://actualidad.rt.com/economia/203895-papeles-panama-billones-eeuu>

- Prieto, M. (05/10/2017). “¿Por qué WalMart, BBVA o Ford compran tecnológicas?” *Expansión*. Recuperado de: <http://www.expansion.com/economia-digital/companias/2017/10/05/59c94bdf46163f770f8b45d3.html>
- Roderick, P. (05/04/2016). “Putin Caught In Huge Panama Papers Scandal”. *Forbes*. Recuperado de <https://www.forbes.com/sites/paulroderickgregory/2016/04/05/putin-caught-in-huge-panama-papers-scandal/#272176987d31>
- Schwab, Klaus. (2016). “The Fourth Industrial Revolution: what it means, how to respond”. Recuperado de <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/> (Fecha de acceso: 30/01/2018)
- Singer, N. (22/07/2012). “Consumer Data, but Not for Consumers”. *The New York Times*. Recuperado de <http://www.nytimes.com/2012/07/22/business/acxiom-consumer-data-often-unavailable-to-consumers.html>
- Smith E. (20/01/2018). “The techlash against Amazon, Facebook and Google- and what they can do”. *The Economist*. Recuperado de <https://www.economist.com/news/briefing/21735026-which-antitrust-remedies-welcome-which-fight-techlash-against-amazon-facebook-and>
- World Economic World. (2016). “La Cuarta Revolución Industrial”. Recuperado de <http://weforum.org>. (Fecha de acceso: 29/01/2018)
- Zegada, A. (17/04/2016). “ICIJ protege a los más ricos de los Papeles de Panamá”. *El País en*. Recuperado de: <http://www.elpaisonline.com/index.php/blogs/la-billetera/item/212561-icij-protege-a-los-mas-ricos-de-los-papeles-de-panama>
- “350.000 puestos de trabajo están sin cubrir en España por falta de formación digital” (21/05/2017). *20 Minutos*. Recuperado de <https://www.20minutos.es/noticia/3042400/0/350000-empleos-espana-sin-cubrir-falta-formacion-digital/>