



Revista de Comunicación
ISSN: 1684-0933
ISSN: 2227-1465
rcom@udep.edu.pe
Universidad de Piura
Perú

Algoritmos, aplicaciones y Big data, nuevos paradigmas en el proceso de comunicación y de enseñanza-aprendizaje del periodismo de datos

Flores Vivar, Jesús Miguel

Algoritmos, aplicaciones y Big data, nuevos paradigmas en el proceso de comunicación y de enseñanza-aprendizaje del periodismo de datos

Revista de Comunicación, vol. 17, núm. 2, 2018

Universidad de Piura, Perú

Disponible en: <https://www.redalyc.org/articulo.oa?id=589466337013>


DOI: <https://doi.org/10.26441/RC17.2-2018-A12>



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial 4.0 Internacional.

Algoritmos, aplicaciones y Big data, nuevos paradigmas en el proceso de comunicación y de enseñanza-aprendizaje del periodismo de datos

Algorithms, applications and Big Data, new paradigms in the process of communication and teaching-learning of data journalism

Jesús Miguel Flores Vivar jmflores@ucm.es
Universidad Complutense de Madrid, España
 <https://orcid.org/0000-0003-1849-5315>

Revista de Comunicación, vol. 17, núm. 2, 2018

Universidad de Piura, Perú

Recepción: 14 Abril 2018
Aprobación: 10 Julio 2018

DOI: <https://doi.org/10.26441/RC17.2-2018-A12>

Redalyc: <https://www.redalyc.org/articulo.oa?id=589466337013>

Resumen: Las tecnologías disruptivas y su impacto en el periodismo y comunicación obligan a asumir retos en el aprendizaje de nuevas técnicas para el procesamiento de datos e información. El conocimiento interdisciplinar se pone de manifiesto en la formación de nuevos perfiles profesionales. El periodismo de datos es una muestra de ello, por lo que la inmersión hacia una cultura de datos debe estar precedida por una concienciación en el aprendizaje de aplicaciones de noticias, algoritmos o el tratamiento del Big data, elementos que configuran nuevos paradigmas entre los periodistas de los medios en Internet.

Con la revisión de textos, observación directa de aplicaciones seleccionadas y estudio de caso, se establecen unas conclusiones que encierran una creciente demanda en el conocimiento de nuevas técnicas. Los resultados arrojan el uso de recursos tecnológicos y la propuesta de cambios en la malla curricular de las facultades de comunicación.

Palabras clave: periodismo, datos, investigación, aplicaciones, algoritmos.

Abstract: Disruptive technologies and their impact on journalism and communication force us to assume challenges in learning new techniques for data and information processing. Interdisciplinary knowledge is evident in the teaching of new professional profiles. Data journalism is an example of this, so the immersion into a data culture must be preceded by awareness in the learning of news applications, algorithms or the treatment of Big Data, elements that configure new paradigms among journalists of the media on the Internet.

With the revision of texts, direct observation of selected applications and case study, some conclusions are established that contain a growing demand in the knowledge of new techniques. The results show the use of technological resources and the proposal of changes in the curricula of the communication faculties.

Keywords: journalism, data, investigative, apps, algorithms.

1. Introducción

Los cambios y transformaciones tecnológicas, fagocitadas por Internet, implican algunas adaptaciones de la profesión periodística y de los medios de comunicación al ser los primeros afectados por estos cambios vertiginosos, incluyendo, el propio concepto de periodismo. Es obvio que el periodismo es algo más -bastante más- que tener una página web, colgar

un vídeo en YouTube o narrar la vida del vecino en Facebook. Por lo que el hecho de que estemos viviendo un proceso de cambio y adaptación a las nuevas tecnologías, no exime al profesional de tener que saber y conocer cómo se redacta una noticia, cómo se hace un editorial, de qué manera se estructura una información, cual es el origen y la historia de los medios de comunicación y un largo etcétera que ni se improvisa ni se adquiere de forma automática. Sin embargo, la forma como la noticia debe ser presentada está cambiando más rápido de lo que imaginamos. Y gran parte de la causa-efecto se encuentra en el uso de aplicaciones tecnológicas o tecnologías disruptivas, como también se les conoce, que directa e indirectamente afectan al periodismo, a los profesionales de los medios y a los estudiantes de Periodismo o Comunicación.

Entre 2009 y 2012, empezamos a percibir una forma novedosa de contar historias periodísticas denominado periodismo de datos. Desde 2014, prácticamente cada semana ha venido apareciendo un nuevo sitio sobre este entorno, en donde aplicaciones algorítmicas y otras tecnologías para la visualización de datos, juegan un papel cada vez más importante. Al margen de las iniciativas que algunos profesionales -individual o colectivamente-, empezaron a diseñar y crear modelos informativos basados en los datos masivos (Big data), los medios de comunicación, principalmente, los más emblemáticos, se han volcado en el desarrollo de algoritmos y uso de herramientas y recursos dentro de sus sitios web con el objetivo no solo de explorar y expandir esta vertiente informativa, sino, también, como una estrategia de imagen vanguardista del medio. En esta expansión del fenómeno del periodismo de datos se suma, por un lado, el tratamiento de los datos masivos y por otro, el concepto de acceso abierto de la información en la que se encuentran involucrados diversos actores sociales.

A nivel internacional, ésta es la tónica que pervive en los grandes periódicos del mundo. Un ejemplo de ello lo protagoniza *The New York Times* con una iniciativa lanzada en 2014 y que lleva por nombre, *The Upshot*, cuyo responsable es *David Leonhardt*, periodista ganador de un Pulitzer que también fuera redactor jefe del *Washington Post*.

En España, los periódicos generalistas como El Mundo, El País, ABC o La Vanguardia, por citar algunos, vienen explorando, aunque no en la magnitud deseada, esta novedosa manera de incursionar en la información que rompe los esquemas en cuanto a las formas de visualización de la información a la que los lectores han estado –o están– acostumbrados. En este escenario, tanto desde una perspectiva profesional como académica, asistimos a ver iniciativas en el uso práctico de este campo del periodismo.

En el contexto actual, nos encontramos con un abanico de herramientas y recursos online que, en algunos casos, causa cierto desasosiego a periodistas y otros profesionales de la comunicación, sobre todo, aquellos cuya incursión en el campo tecnológico data de pocos años atrás. Así, este estudio muestra cómo emerge y se realiza el periodismo de datos, su uso por los principales medios de comunicación españoles, las distintas leyes y reglamentos que hacen posible la expansión, el auge

de proyectos e iniciativas, el tipo de tecnologías utilizadas, el proceso de comunicación generado por algoritmos y el rol de la universidad en la formación y transferencia de nuevos conocimientos para la consolidación del nuevo perfil profesional.

En este escenario, desde hace algunos años, cada vez más voces, tanto académicas como profesionales, se alzan para poner en entredicho el futuro del periodismo como profesión, motivado por las consecuencias del impacto tecnológico que catapulta un tipo de periodismo de datos masivos y de bases de datos. Pero, ¿Cómo adaptamos las características del paradigma de bases de datos a la nueva realidad mediática inmersa en una cultura de red? En la actualidad, nos encontramos con un abanico de herramientas y recursos online que es necesario conocer y aprender por lo que este trabajo aborda los antecedentes del periodismo de datos, su uso por los principales medios españoles, las distintas leyes y reglamentos que hacen posible la expansión, el auge de proyectos e iniciativas, la proliferación y tipos de tecnologías utilizadas, el proceso de comunicación que genera el algoritmo y el rol de la universidad en la formación y transferencia de nuevos conocimientos para la consolidación del perfil profesional de la era del Big data. Este último aspecto es de vital importancia ya que vemos como, en contrapartida, emergen profesionales con nuevos conocimientos, adquiridos de forma autodidacta, quienes 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, una primera hipótesis se deriva de las tendencias tecnológicas en las redacciones, en donde las aplicaciones y los algoritmos se constituyen en piezas fundamentales en el desarrollo del periodismo de datos, tanto de medios tradicionales como de medios únicamente digitales. Una segunda hipótesis es que las facultades de comunicación asumen nuevos roles en la formación de periodistas que demanda, a escala global, una sociedad hiper-conectada por lo que estos centros tienen un papel fundamental y de vital importancia en la formación de periodistas de nuevo perfil, con un tipo de conocimientos que requiere de una revisión permanente de los planes de estudio. Pero, ¿Cuáles son los contenidos de esos planes de estudio para la formación de nuevos perfiles profesionales? ¿Es suficiente el 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? ¿Cuáles son las herramientas multimedia que permiten una visualización óptima de los datos?

Por todo ello, el objetivo que persigue este análisis contempla no solo estudiar y describir los recursos tecnológicos, el uso de algoritmos y el funcionamiento del esquema formativo y los roles asumidos por las facultades de comunicación, sino, también, proponer contenidos que implique nuevos conocimientos que los estudiantes de periodismo han de poseer para crecer profesionalmente en las organizaciones de noticias (Flores y Cebrián, 2012, p. 28). Estos contenidos vienen precedidos

por los análisis de distintas herramientas y aplicaciones utilizadas en las diferentes iniciativas sobre periodismo de datos que se han estudiado.

Este trabajo forma parte de los resultados parciales del proyecto de investigación “Innovación y experimentación de nuevas narrativas y medios con tecnologías disruptivas emergentes”, proyecto competitivo de la Universidad Complutense de Madrid, que analiza el estado de la cuestión acerca de la formación, el desarrollo profesional, el tipo de aplicaciones que es necesario conocer y el reto que supone para las facultades de comunicación.

2. Marco referencial. Elementos del periodismo de datos: algoritmos, aplicaciones y datos masivos y su enfoque formativo.

Los antecedentes en el estudio del fenómeno acerca del Periodismo de datos se remontan a 2005, cuando *Adrian Holovaty* creó el sitio web *ChicagoCrime.org*, un *mashup* en donde incluía una serie de datos públicos extraídos de la base de datos de la policía de Chicago que visualizaba utilizando aplicaciones cartográficas como Google maps, dando origen a lo que inicialmente se conocía como el *Data Delivery Editor* y, posteriormente, *Data Driven Journalist*. Hoy, simplemente, *Data Journalist*.

Con este antecedente, los propulsores del periodismo de datos sitúan a la creación de *Holovaty* en el punto de partida. No obstante, algunos investigadores consideran los inicios en otros proyectos, desarrollados ese mismo año, con la creación de otros mashups (era el nombre que adquiriría el producto creado) como el caso de *Housingmaps* (<http://www.housingmaps.com/>), desarrollado por *Paul Rademacher* para su uso en páginas de clasificados como *Craigslist*. De modo comparativo, *Chicago Crime* muestra un enfoque más periodístico, más de servicio público, mientras que *Craigslist+Googlemaps* (como así se le denominó al perfil de *Rademacher*) es más un *mashup* que busca en el sitio web de páginas clasificadas. En este sentido, la creación de *Holovaty* tiene valor periodístico dado que su perfil es de periodista con conocimientos de programación informática.

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 (*software*). De ahí que el periodismo de datos consiste en usar herramientas estadísticas y de visualización para contar mejor las viejas historias y descubrir nuevas historias que contar. Es, en opinión de muchos, la nueva veta del periodismo de investigación. Pero, el desarrollo y puesta en práctica de este tipo de periodismo requiere del conocimiento de otras disciplinas cuyos aprendizajes deben darse transversalmente en los planes de estudios de periodismo. Los antecedentes de la transversalidad de conocimientos, no sólo de contenido y de tecnología (con Internet a la cabeza), sino del anclaje con otras disciplinas de las Ciencias Sociales, como la Estadística

o la Sociología, viene dado por la opinión de diversos expertos y de tendencias de este nuevo ecosistema.

En este contexto, Internet, sus recursos subyacentes y otras tecnologías del conocimiento fagocitan un tipo de periodismo que, sin ser nuevo, ayuda al periodismo en la construcción de una nueva forma de acceder a la noticia. Así, para Alain Joannès (2011, p.24), periodista francés experto en nuevos medios y autor de *Data Journalism. Bases de données et visualisation de l'information*, no se trata ni de una revolución ni de un nuevo periodismo, sino, simplemente, de una búsqueda y tratamiento de información mediante las herramientas que disponemos actualmente y para una audiencia, también, actual. El *modus operandi* es una metodología de investigación a partir de cifras, estadísticas, fondos cartográficos, etc., una herramienta para poner de relieve hechos, aportar pruebas y presentar resultados.

Liisa Rohumaa, profesora de la *Bournemouth University* y Paul Bradshaw (2011, p. 18), profesor de la *Birmingham University*, ambas del reino Unido y autores de *The Online Journalism Handbook: Skills to survive and thrive in the digital age*, afirman que “el periodismo de datos (*data journalism*, en inglés) es aquel en el que se usa el poder de las computadoras (ordenadores) para encontrar, contrastar, visualizar y combinar información proveniente de varias fuentes”. Para Andrea Fama (2011, p.45), 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.

Bill Kovach y Tom Rosenstiel (2012, p.34), en su libro, *Los elementos del Periodismo*, afirman que, “en el periodismo de investigación moderno el poder de análisis que permite la informática sustituye a menudo a la observación personal del reportero”. La serie de reportajes titulada “The Color of Money”, con la que el redactor Bill Dedman, del *Atlanta Journal and Constitution*, que obtuvo un premio *Pulitzer* en 1989, constituye uno de los primeros ejemplos de esta modalidad. Para los autores, la informática empleada (denominada CAR, acrónimo de *Computer Assisted Reporting*) del modo adecuado puede modificar la capacidad del periodismo de investigación porque tiene el potencial de trascender las tradicionales anécdotas y entrevistas y amasar una cantidad abrumadora de pruebas documentales.

Con estos antecedentes, podemos afirmar que la evolución de la transformación por la que atraviesa el periodismo y la comunicación, en general, ha traído como consecuencia una consolidación del Periodismo de datos, modelo periodístico que utiliza una serie de variantes tecnológicas, las mismas que, en muchos casos, es difícil seguir la secuencia de su evolución. De hecho, se dan circunstancias en los que mientras una aplicación empieza su despliegue (*deployment*), ya se han diseñado y creado otras *apps* (*development*) que rápidamente se posicionan en el mercado a través de los recursos sociales de internet, quedando algunas de las herramientas en el camino del aprendizaje (Flores y Cebrián, 2012).

A todo esto, se suma que las técnicas del periodismo de datos pueden ser utilizadas 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. Por ejemplo, en el periodismo deportivo se aplican estas técnicas y se logran muchas innovaciones, dada la gran cantidad de datos involucrados y muchos profesionales con conocimiento en esta área. Pero, es en el periodismo de investigación donde se han logrado grandes innovaciones usando ordenadores para crear enlaces entre informaciones o encontrar historias entre grandes conjuntos de datos mediante técnicas de cruce de bases de datos. De hecho, donde más está creciendo el periodismo de datos es en la información política puesto que cada vez más interesan a los ciudadanos los datos acerca de la gestión que hacen los políticos en el Gobierno, máxime si se encuentra en juego, por ejemplo, los impuestos que cada ciudadano aporta a las arcas del estado.

En este contexto, en una unidad de profesionales en un medio dedicada al periodismo de datos, el perfil de los periodistas debe ser el de personas con habilidades para la investigación tradicional, incluyendo el periodismo asistido por computador (CAR); periodistas con conocimientos de estadística, y diseñadores de información junto a programadores informáticos que puedan crear secuencias de comandos para facilitar a periodistas y usuarios a obtener y aprovechar mejor la información.

Actualmente, el modelo CAR está siendo utilizado por casi todas las redacciones del mundo. Independientemente de que su uso se restrinja no sólo al periodismo de datos sino, además, al desarrollo de cualquier modelo o especialidad periodística. Por ello, se destaca que el uso de las técnicas actuales del CAR junto a las técnicas de la estadística, forman un binomio necesario en el desarrollo y expansión del periodismo de datos.

2. 1. El algoritmo como elemento clave en el ecosistema del ciberperiodismo y los nuevos medios

Kevin Kelly (2012) investigador y experto en tecnologías (ex editor de la revista *Wired* y autor del libro *Fuera de Control* y el bestseller *New Rules for a New Economy*), afirma que las tecnologías buscan para nosotros. Pero ¿Qué buscan? De esta afirmación y posterior interrogante nos valemos para indicar que el uso de aplicaciones no se han creado *per se* para el periodismo o comunicación, sino que corresponde adaptar, desarrollar y diseñar un entorno más específico de las tecnologías para su uso en el entorno ciberperiodístico. Un desarrollo *ad-hoc* de un todo que debe funcionar según los flujos de trabajo o parámetros que los medios tengan trazados. Es decir, un flujograma de funciones trazadas que genera un algoritmo.

Actualmente, la técnica de los algoritmos juega un papel fundamental en la búsqueda de datos e información digital para la investigación periodística. Los expertos afirman que sólo hace falta saber un poco de lógica para crear nuestro propio algoritmo cuyo fin sea procesar

la información. Por ejemplo, la mayoría de agregadores de noticias están hechas mediante algoritmos que permiten seleccionar y encontrar información según el interés de cada usuario. Google es, por excelencia, el rey de los algoritmos.

2. 2. 1. Los contornos de la vida algorítmica

El algoritmo se está convirtiendo en la unidad esencial en la creación de la nueva información y comunicación. El grupo de investigación *The Mellon Initiative In Digital Cultures* asociado al *Center for Science and Innovation Studies* (CSIS), dirigido por el Prof. Dr. Mario Biagioli, de la Universidad de California Davis, organizó en mayo de 2014, el congreso “Los Contornos de la Vida Algorítmica” (*The Contours of Algorithmic Life*). La iniciativa del grupo de investigación pone de relieve los desafíos de la vida algorítmica que ya se da de forma global y está cada vez más presente en la vida de las personas y sobre todo, como es nuestro caso, en la inclusión de disciplinas y profesiones como el periodismo y la comunicación.

Por ello, la propuesta del congreso realizado por el referido grupo de investigación asociado al CSIS de la UC Davis, ha tenido una visión interdisciplinaria, corroborada por la cantidad de asistentes (profesionales, académicos e investigadores) pertenecientes a distintas disciplinas académicas. Y, es que cada vez más, los profesionales e investigadores de distintas disciplinas, entre los que se cuentan, principalmente, aquellos relacionados con la comunicación, medios digitales y las nuevas formas periodísticas, tienen más claro el cómo los algoritmos permean nuestra experiencia de vida, bordeando los límites y las fronteras de lo que puede y no puede ser adaptado, traducido o incorporado en el pensamiento algorítmico, convirtiendo todo esto en un ecosistema de contención (algorítmica).

En este contexto, las conclusiones del congreso fueron que el principio del algoritmo se basa en la especificación del espacio potencial de una acción, generando una noción de un modo universal lo que lleva a los discursos sobre el empoderamiento, la eficiencia, la transparencia y la inclusión. Pero en última instancia, los algoritmos sólo son capaces de hacer inteligible y valioso todo aquello que puede ser separado, cuantificado, manejado, procesado y *gamificado* y este dominio limitado hace que los algoritmos necesariamente sean excluyentes.

Los algoritmos configuran cada vez más nuestro mundo, nuestro pensamiento, nuestra economía, nuestra vida política, y nuestros cuerpos. Vemos esto, por ejemplo, en la respuesta de las redes de la NSA u otros servicios de vigilancia sobre la actividad de la red que amenaza cada vez más la privacidad de las personas al conllevar una vigilancia política bajo control algorítmico. De hecho, al menos, el 30% de las acciones de comercio electrónico que realizamos es actualmente algorítmico y automático, situación que, en algunos casos, nos ha llevado a varios colapsos inexplicables. Con *Amazon* . *Netflix* a la cabeza, las decisiones de compra a la que son sometidos los usuarios están hechas por un algoritmo.

Es importante, por tanto, explorar tanto los usos específicos de los algoritmos y la cultura algorítmica de una manera más amplia, incluyendo otros temas como la programación, minería de datos y visualización, política de los algoritmos, vigilancia digital, tecnología móvil y geo localizada. Hoy, cualquier disciplina es proclive a tener una vida algorítmica que afecta a su entorno profesional.

En este sentido, los debates que se circunscriben a la vida algorítmica deben ser concebidos abiertamente, incidiendo particularmente en los estudios de software, estudios de códigos críticos, estudios de rendimiento crítico, estudios culturales y de medios de comunicación, antropología social, ingeniería inversa, humanidades y ciencias sociales, así como las artes visuales, la música y los estudios de sonido. Es importante avanzar en propuestas experimentales o híbridas en donde el arte de lo multimedia y los recursos en Internet juegan un papel preponderante.

2. 1. 2. Concepto y alcance del algoritmo.

Pero ¿Qué es un algoritmo? ¿Qué conocimientos son necesarios para crearlos? ¿Por qué se han convertido en una especie de piedra filosofal en la construcción de la información? ¿Qué relación guarda con el periodismo de datos? Según los manuales, un algoritmo es el proceso de una acción. Un proceso o flujo de trabajo que está supeditado a ordenes (parametrizadas) que el creador (del algoritmo) previamente ha configurado.

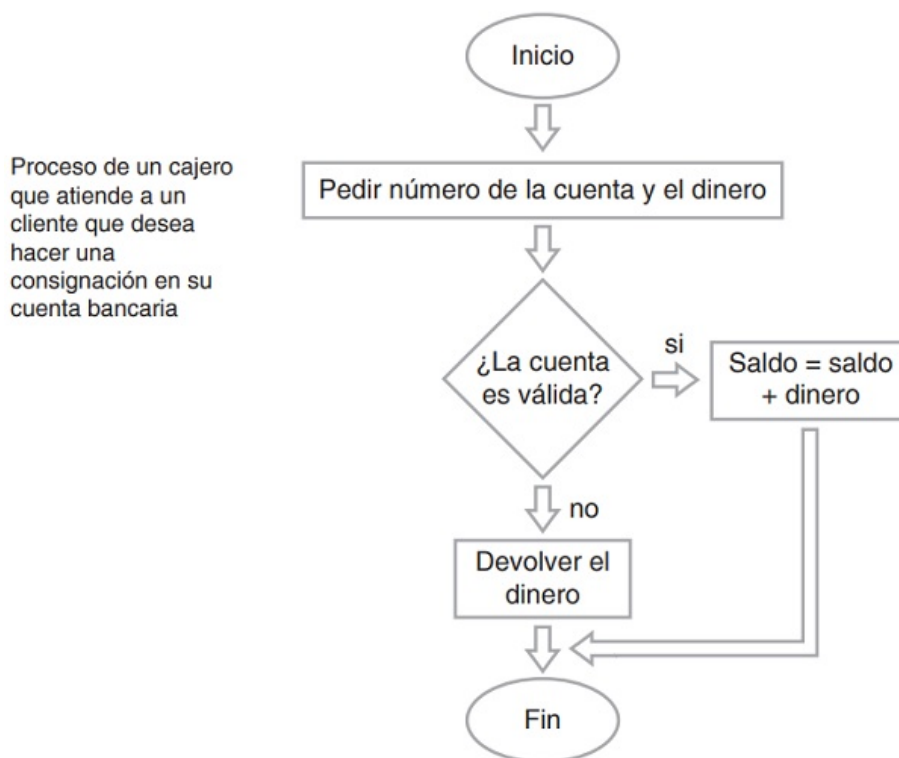


Figura 1
Ejemplo de proceso de algoritmo simple.

Fuente: Elaboración propia

El empleo del algoritmo en matemáticas, lógica, ciencias de la computación y disciplinas relacionadas es más frecuente. Pero, también en las disciplinas del periodismo, potenciado por el proceso y tratamiento de los datos masivos y el libre acceso a la información.

La Real Academia de la Lengua, recoge dos acepciones del algoritmo. Por un lado es un Conjunto ordenado y finito de operaciones que permite hallar la solución de un problema. Y por otro, es un Método y notación en las distintas formas del cálculo. Para Harry Guinness (2017), un algoritmo es un conjunto prescrito de instrucciones o reglas bien definidas, ordenadas y finitas que permite realizar una actividad mediante pasos sucesivos que no generen dudas a quien deba realizar dicha actividad. En un escenario con un estado inicial y una entrada, siguiendo los pasos sucesivos se llega a un estado final y se obtiene una solución.

En este escenario, en la configuración de los nuevos medios el uso de algoritmos para la optimización de los contenidos informativos se ha convertido en una tendencia creciente. De ahí que, desde hace pocos años, algunos medios han empezado a hacer exploraciones sobre este entorno. En la actualidad, el uso de algoritmos se ha visto fagocitado por su desarrollo cada vez más simple, lo que ha hecho posible una accesibilidad mayor a las tecnologías digitales por parte de usuarios que no han estado acostumbrados a ellos, por ejemplo, los periodistas. En paralelo, su uso también se debe a la incorporación de profesionales de otras disciplinas más técnicas en las salas de redacción, formando un binomio de periodista y programador, como anteriormente se hacía entre el periodista y el reportero gráfico o el periodista y camarógrafo.

Por ello, cada vez más periodistas se atreven a crear contenidos aprovechando los recursos y herramientas que proporciona internet. Algunos, más osados, se atreven con el código sincretizando de esta forma el conocimiento periodístico con las tecnologías más avanzadas, pero, al mismo tiempo, más asequibles.

2. 1. 3. Estudio de caso: Quakebot. El algoritmo en la primicia de las noticias.

Diversos expertos del periodismo en internet, afirman que los algoritmos están en todas partes. Al día de hoy ya dominan los mercados de valores, la composición musical, la conducción de coches y la redacción de noticias, y son artífices de extensas pruebas matemáticas; además, sus poderes para la autoría creativa sólo están empezando a cobrar forma (Finn, 2018, p.36). En esta línea, para Valle (2014) “son la base de los buscadores de noticias, de las redes sociales o de los servidores de publicidad que muestran los anuncios en los medios digitales”. La relación existente entre un algoritmo y el periodismo de datos es cada vez más pertinente. Para Juan Varela (2008), el átomo de la información ya no es la noticia y sus moléculas no son los formatos periodísticos.

La oferta de los medios es cada vez menos un producto -o una colección de ellos- acabado para públicos objetivos. Los nuevos medios son desarrollos informativos donde el hiperenlace, la estructura en red de la información y su carácter abierto, así como la interacción de las herramientas tecnológicas y multimedia

permiten desarrollar los contenidos como un proceso. Los medios se convierten en plataformas donde los usuarios, llegados a través de diferentes soportes y vías (acceso distribuido) podrán ejecutar programas informativos a través de algoritmos que aún tenemos por descubrir (Varela, 2008, p.25).

En este contexto, parafraseando la teoría macluhaniana de “el medio es el mensaje” es probable que estamos ante nuevo un paradigma en donde “el medio es el algoritmo”. Un ejemplo de algoritmo para las noticias, se desprende de una nota del *Knight Center* de Periodismo para las Américas que en un post da referencia de un terremoto de magnitud 4.4 (Escala de *Richter*) que despertó a los residentes de California el lunes 17 de marzo de 2014. El primer reportero en cubrir la noticia fue un algoritmo llamado *Quakebot*.

Quakebot fue diseñado, en 2011, por *Ken Schwencke*, periodista y programador del diario estadounidense *Los Angeles Times*, para recopilar datos de interés periodístico de alertas publicadas por el Servicio Geológico de los Estados Unidos (*USGS*, siglas en inglés), introduciéndolos a una plantilla de noticias escrita con anterioridad. La nota preliminar se guarda en el sistema de manejo de contenido del periódico mientras que se envía una alerta por correo a un editor humano, quien revisará la nota y decidirá si la publica o no, en línea. Cuando *Quakebot* recibió una alerta del *USGS* señalando un terremoto de magnitud 3.0 o mayor (algo menor no causaría mucho impacto) a las 6:25 a.m., hora del Pacífico, envió un correo a *Schwencke* quien publicó la nota pocos minutos después, asegurando de esta forma que *Los Angeles Times* fuera el primer periódico en reportar la noticia.

Siguiendo el apunte de Varela (2008) “el desafío del periodismo es reinventarse para mantener los viejos criterios de calidad y sumar los nuevos y las oportunidades de la tecnología y la acción de los ciberciudadanos”. Así, desarrollar códigos informativos cuyos contenidos propios y ajenos se ofrecen en una plataforma en donde la conexión entre información y usuarios y su portabilidad permite el desarrollo de algoritmos procesables en común o individualmente, a semejanza de lo desarrollado por *Schwencke*, con el objetivo de programar resultados en forma de una oferta de información inteligible y de valor para el conocimiento de una sociedad cada vez más hiperconectada.

Este es, por tanto, el verdadero desafío de los futuros medios. Aunque siempre habrá quien mantendrá que su selección y sus criterios -a menudo cerrados y no explícitos- son la identidad a la que muchos deben adherirse. Y es más que probable que los periodistas seguirán esta tendencia. Por ello, es también un reto de las universidades en la formación periodística, como se trata en otros apartados de este artículo.

Por tanto, vemos que la tecnología no sólo ha cambiado el modo de cómo se escriben las noticias sino, también, quién las está escribiendo. Los algoritmos inteligentes, con acceso a conjuntos de datos y dirigidos por periodistas, ahora son capaces de escribir rápidamente noticias básicas sobre temas deportivos, económicos y judiciales.

2. 2. Herramientas y aplicaciones en el tratamiento de datos

En el seguimiento realizado sobre la evolución del periodismo de investigación y en su forma derivada en periodismo de datos, detectamos que el ritmo de evolución y cambio en el mundo tecnológico y, concretamente, en el informático, es frenético, acelerado y un tanto delirante. A la vez que emergen aplicaciones que se crean cada vez con más frecuencia y cuyos nombres y usos no terminamos de conocer, hay profesionales del mundo de la red que se han hecho un hueco por su afán de superación y su constante trabajo en este campo; nos han demostrado que el cambio es constante.

Por otro lado, se da la circunstancia que en la medida que los medios españoles empezaron a utilizar los recursos tecnológicos disponibles, la mayoría gratuitos, su uso estaba constreñido a aquellos cuya habilidad o conocimiento tecnológico iba más allá de la exclusividad en la redacción de contenidos. Hoy, ese panorama está cambiando dado que cada vez son más los profesionales, incluso desde su etapa de estudiantes que dominan una serie de aplicaciones y recursos online, situación que les permite hacer una inmersión en el entorno del *Data Journalism* (periodismo de datos). Conocimientos como la ingeniería inversa, la codificación o programación de lenguajes, aunque no sea de uso común por parte de los periodistas, pero, si de nociones básicas, va ganando más adeptos entre este colectivo que demanda en conocer y aprender este tipo de tecnología desde la universidad.

Por tanto, consideramos que esta evolución del periodismo augura un futuro prometedor de los reporteros. Aunque para ello tendrán que pasar por incluir en su tablero de conocimientos, temas como la experiencia en análisis de datos, la búsqueda y selección de información, visualización de datos, aplicación de datos en historias periodísticas, codificación algorítmica, manejo de herramientas y recursos digitales, entre otros, que serán las claves para un tipo de periodismo que muchos empiezan a explorar.

Detectamos que las grandes empresas tecnológicas ven en el periodismo de datos una oportunidad para desarrollar y crear aplicaciones que puedan ser utilizadas por periodistas, profesionales de medios, investigadores y otros usuarios. Las aplicaciones (descritas) constituyen ejemplos de cómo la factoría Google muestra un especial interés en este entorno profesional de la comunicación. Uno de estos ejemplos, lo encontramos en los mapas que se han convertido en una forma de periodismo nativa de la web (“mapeador de información”), éxito derivado, en parte, de la aplicación Google maps, lanzada en 2005 (Flores y Cebrián, 2012).

Ante este panorama, no es de extrañar que surjan iniciativas de instituciones y empresas cuyo objetivo sea promover el entorno investigador y práctico del periodismo de datos como el *Global Editors Network* (GEN) y Google, que desde 2012, han lanzado la convocatoria del Premio Periodismo de Datos (datajournalismawards.org). Los Premios de Periodismo de Datos (DJA) constituyen el primer concurso

internacional que reconoce la destacada labor en el campo del periodismo de datos en todo el mundo.

2. 3. El Big data en la creación de contenidos informativos de los medios

Durante mucho tiempo los tecnólogos han sido los “guardianes” y creadores de repositorios que ha proporcionado grandes cantidades de datos a las instituciones y empresas almacenando la información de la mejor manera posible. Ahora estamos ante un escenario que produce cierto vértigo ¿qué hacemos o qué hacen las empresas e instituciones con toda la información almacenada? Más aún: es evidente que cada vez más, la información y los datos se extienden dentro del mundo “*social media*” de los usuarios participativos y está saliendo del círculo cerrado de las organizaciones, instituciones y empresas. Todo ello, debe hacernos pensar que ya no se trata solo de datos ordenados y perfectamente clasificados como se ha venido haciendo, sino que ahora hay que tratar datos heterogéneos, de multitud de fuentes, con estructuras y formatos muy diversos. La Web en general, y las redes sociales en particular, producen cantidades ingentes de datos impensables hace años.

Según Viktor Mayer-Schönberger y Kelly Cukier (2013, p.35) “*Big data*” es un término aplicado a conjuntos de datos que superan la capacidad del software habitual para ser capturados, gestionados y procesados en un tiempo razonable. Los tamaños del “*Big data*” se encuentran constantemente en movimiento creciente, de esta forma en 2012 se encontraba dimensionada en un tamaño de una docena de *terabytes* hasta varios *petabytes* de datos en un único *data set*.

Si bien es cierto que las tecnologías de la información han catapultado el concepto, los datos antes de la revolución digital han existido siempre, por lo que su análisis no es completamente nuevo. Uno de los ejemplos más antiguos es el caso de del análisis de las muertes por cólera realizada por el doctor John Snow, uno de los precursores de la epidemiología en el Londres de 1854. Mediante la visualización geográfica de las muertes que se habían producido por cólera en el mes de septiembre de ese año, el doctor Snow verificó que la variable que mejor explicaba las muertes era la existencia de un séptico de aguas contaminadas en *Broad Street*. Gracias al análisis explico que la gente que más enfermaba se encontraba viviendo cerca del pozo, mientras que los menos eran aquellos que vivían lejos (Ontiveros, 2017, p. 38).

El concepto de datos masivos, macro datos o *Big data* (en inglés) no es tanto una cuestión de cantidad de datos como, sobre todo, de capacidad de buscar, agregar y cruzar unos datos para obtener otros. El término no se refiere solo a los datos en bruto sino a las técnicas por las cuales se pueden obtener conclusiones interesantes de esos datos que están, en algunos casos, en abierto (Elías, 2015, p. 34). Los motivos por los cuales se ha acuñado el término Big data para hacer referencia a las diversas tecnologías para su procesamiento y cuyo entorno se ha convertido en tendencia en todos los sectores económicos, incluido el de los medios de comunicación, son muy variados.

Para Emilio Ontiveros (2017, p. 39):

El principal motivo es concepto de la transformación digital que ha hecho que se haya empezado a recopilar información a gran escala de diferentes procesos empresariales. El segundo motivo es la aparición de tecnologías que nos permiten almacenar y procesar volúmenes de información de una manera barata y escalable.

El tercero es la adquisición de conocimientos en el mundo empresarial de que existen una gran variedad de técnicas estadísticas, matemáticas e informáticas, tradicionalmente conocidas desde los años cincuenta en el sector académico y científico. El cuarto y último motivo es, precisamente, lo que se ha dado en llamar economía de los datos, la consciencia de que tanto los datos en sí mismos como lo que podemos derivar de ellos (información, conclusiones, predicciones, etc.) pueden ayudar a las empresas a generar más ventas, disminuir costes e incluso, generar nuevos negocios por sí mismos.

No obstante, no hay que olvidar que los datos masivos son avances tecnológicos y como cualquier tecnología, puede entenderse desde una aproximación utópica o distópica (Elías, 2015, p.36). Sus defensores sostienen que el *Big data* posee un potencial enorme de riqueza y conocimiento, permitiendo resolver problemas sociales y económicos y de áreas tan complejas y diversas como las investigaciones en cura del cáncer, biotecnológica o investigación aeroespacial. Pero, también hay detractores, atribuyendo a este concepto los problemas de seguridad de la información o la violación de los derechos personales, por lo que un importante sector de la sociedad considere que el Big data encierra más peligros que beneficios. Para Cathy O'Neil (2017, p. 11), las aplicaciones fundamentadas en las matemáticas (que crean algoritmos para el procesamiento de datos masivos) que alimentaban la economía de los datos se basaban en decisiones tomadas por seres humanos que no eran infalibles. Seguro que algunos de esas decisiones se tomaban con la mejor de las intenciones, pero muchos de los modelos algorítmicos se programaban en base a los prejuicios, equivocaciones y sesgos humanos en unos sistemas informáticos que dirigen cada vez más nuestras vidas.

En este contexto, vemos que el mundo del *Big data* es muy interesante para las empresas. Y con más frecuencia para las organizaciones de noticias, habituadas a procesar toda la información social publicada por sus potenciales usuarios (antes conocida como audiencia) y aprovecharla para crear un perfil de los mismos y adelantarse a sus peticiones y necesidades.

En este sentido, muchos tecnólogos empiezan a adaptarse en este nuevo eco- sistema, considerando que ya no es suficiente con ser un “analista de datos”, sino en convertirse en un valor en alza llamado “científico de datos”. Diversos investigadores coinciden en que el científico de datos debe tener una formación similar a la del analista de datos, pero, además, de los conocimientos avanzados de informática (sobre todo en herramientas y lenguajes como Hadoop, Pig, Python y Java), estadística, análisis y matemáticas, también debe tener un sólido conocimiento del negocio (en nuestro caso, de los medios) y, sobre todo, debe tener la capacidad de comunicarse tanto con las áreas de tecnología como las de negocio, dentro de la organización. Este es uno de los perfiles que viene

demandándose como un elemento fundamental en el mundo de los datos masivos.

Según un informe de la Universidad de Harvard (<https://datascience.harvard.edu/>) “El científico de datos...es el trabajo del siglo XXI”. Diversos expertos coinciden en algunas de las funciones de esta nueva forma de gestión de la información:

- - Garantizar el flujo de la información (la información adecuada, a la persona adecuada, en el momento adecuado)
 - Extender la gestión de la información (ampliar la cobertura de la información que somos capaces de gestionar)
 - Reducir la complejidad y reducir los silos de información (accesible desde cualquier dispositivo, en cualquier momento y en cualquier lugar con la capacidad de agregar fuentes, formatos y localizaciones en una visión única).

Para *Michael Coley* (1987), los datos, organizados y empleados debidamente, pueden convertirse en información. La información, absorbida, comprendida y aplicada por las personas, puede convertirse en conocimientos. Los conocimientos aplicados frecuentemente en un campo pueden convertirse en sabiduría, y la sabiduría es la base de la acción positiva.

En base a este precepto, las organizaciones de noticias están abocadas en crear equipos multidisciplinarios (científicos de datos, programadores y periodistas) para el tratamiento de los datos masivos proveniente de los grandes archivos de contenidos y de los “rastros” (navegación y participación) dejados por sus lectores cuando acceden a sus sitios de noticias.

Vemos pues, que el término Big data abarca cada vez más campos del conocimiento debido a su amplitud de desarrollo. Los datos masivos se han empleado con acierto en las ciencias naturales y desde hace poco tiempo, ha empezado a introducirse en el ámbito de las ciencias sociales, en áreas tales como economía, publicidad y política. Y también en el periodismo, aunque su aplicación data de unos pocos años atrás, pero con una tendencia creciente.

Los tecnólogos (informáticos, ingenieros, analistas) saben manejar los datos, pero son los periodistas los llamados a interpretar y contar una historia a través de los datos. Para Elías (2015, p. 34), los ‘periodistas son los que en contacto con el poder político y económico, usan su olfato para ordenar los datos y descubrir información que quizás los citados poderes prefieran mantener ocultos. Una de las primeras aplicaciones del tratamiento de los datos masivos a las ciencias sociales se evidenció en un artículo de la revista *Nature*, publicada en 2009, en donde se demostraba cómo, en función de las búsquedas de Google, podían determinarse con mayor exactitud, los brotes de gripe en Estados Unidos, que con las habituales encuestas y estadísticas de los centros médicos y el sistema nacional de salud. (Mayer-Schönberger y Cukier, 2013, p. 12).

En los últimos años, ha cambiado de manera notable el tratamiento de los datos masivos por parte de los periodistas, quienes empiezan a

especializarse en estos entornos. Los grandes medios de comunicación, europeos y americanos, han abierto secciones de *Open data*, con el deseo expreso de integrar el periodismo de datos y el trabajo con grandes volúmenes de información en sus prácticas profesionales habituales. Para Carlos Elías (2015, p. 35), éste y otros cambios están contribuyendo a la aparición de una audiencia interesada en el acceso interactivo a información de calidad, y en algunos casos, también a recuperar el maltrecho negocio de la información.

2. 4. Enfoque formativo de los elementos del periodismo de datos.

En un post publicado en 2014, en el blog Cultura de Red (del Sistema de divulgación científica Madrid), respecto a la importancia del periodismo como profesión y disciplina del conocimiento, se hace referencia al análisis (de otro artículo) sobre el estado de la formación periodística en las facultades de comunicación. En dicho análisis se sugiere una reestructuración de los planes de formación con un giro radical de 180 grados. El tan denostado Plan Bolonia (proyecto que consiste en la adaptación y unificación de criterios educativos en todos los centros europeos, divide la enseñanza superior en tres niveles. El grado tiene una duración de cuatro años con 240 créditos, y sustituye a las diplomaturas y licenciaturas. El máster supone una especialización en un ámbito concreto o multidisciplinar y puede tener una duración de un año con 60 créditos o dos con 120. Al doctorado se accede a través de un máster específico, y por término medio los estudios se prolongan durante cuatro años), puede ser el recurso a utilizar para lograr esos cambios. Algunas universidades ya lo vienen haciendo, pero no en la magnitud ni con la rapidez deseada. Cabe recordar que en dichos cambios subyacen dos factores de vital importancia. Uno, el futuro de los estudiantes de periodismo y comunicación, cuya formación idónea depende de lo que reciba en su facultad. Y dos, la pervivencia de la propia disciplina del periodismo, como garante y pilar fundamental de toda democracia.

Urge, por tanto, un cambio estructural de los planes formativos en donde se aborde, por ejemplo, la enseñanza de las aplicaciones descritas, tanto desde una perspectiva crítica, reflexiva o teórica, como práctica. La currícula de estudios debe incluir, al menos de forma introductoria, el diseño, creación y desarrollo del algoritmo aplicado a los nuevos medios y al nuevo periodismo basado en los datos. La razón de todo este cambio es elocuente: en el desarrollo del periodismo de datos se vienen utilizando una serie de recursos online, unas veces gratuitos, otros de pago y otros que se crean de forma compartida.

Esta situación, que implica los cambios propuestos, no sólo pasa en las universidades españolas, europeas o latinoamericanas, sino, también, en las escuelas estadounidenses. En junio de 2014, un post publicado por Jon Marcus en el *Nieman Journalism Lab* y extraído del *Nieman Reports* (ambas, iniciativas de la Fundación Nieman, centro de la Universidad de Harvard, EE.UU.), hacía referencia a este tema con el título “*Como las escuelas de periodismo están tratando de conectar con las salas de redacción*”.

Jon Marcus (2014), editor del área de Educación Superior del Informe Hechinger (*The Hechinger Report*) que examina el futuro de la enseñanza del periodismo, sostiene que “La enseñanza del Periodismo ha llegado al mismo punto de inflexión de incertidumbre que el periodismo en sí se ha alcanzado”. Y, quienes piensan así, son al parecer, muchos más de los que nos pudiésemos imaginar.

En esta línea, el panorama de las universidades estadounidenses con escuelas de periodismo no es nada halagüeño. Si bien es cierto que hay escuelas innovadoras en la formación, como es el caso del nuevo programa de postgrado en innovación de medios en la *Northeastern University*, en donde los estudiantes pueden aprender codificación, visualización de información, videografía, gestión de base de datos e incluso, diseño de video-juegos. O, el Centro *Knight* para el Desarrollo Empresarial del Periodismo de la *City University de Nueva York .CUNY*) cuyos planes tienen por objetivo incubar a los empresarios periodísticos quienes pueden ganar, además, becas de maestrías y certificados avanzados en enfoques innovadores para el negocio de los medios. O, la Universidad de Columbia, en donde su Escuela de Periodismo y el Departamento de Ciencias de la Computación se han unido (a través de un equipo trabajo) para preparar un programa dirigido a periodistas y otros interesados en las tecnologías de bases de datos y el uso práctico de las mismas. Estos ejemplos, son sólo una pequeña muestra en relación a la cantidad de universidades con estudios en periodismo y que requieren de un gran cambio.

No hacer esto, conllevaría entrar en un bucle difícil de salir. Según Marcus (2014), las universidades están cerrando o proponen cerrar las escuelas de periodismo, o, en algunos casos, realizar la fusión con otros departamentos. La inscripción de estudiantes que optaban por periodismo, está cayendo - de manera espectacular, para los programas de postgrado -. En contrapartida, está aumentando las inscripciones en nuevos centros de formación u otras instituciones que ponen el énfasis en los medios digitales. Y es que las nuevas formas de enseñanza en línea y nuevas credenciales amenazan todo del monopolio de la educación superior que otorga el crédito académico de las universidades. En España, véase el caso de escuelas de periodismo, creadas por los propios medios de comunicación.

Nos guste o no, esta es una situación irreversible. Como la tecnología avanza, los profesionales quieren más formación (al igual como ya tienen, por ejemplo, los programas de periodismo de la *Northeastern University*). En Estados Unidos, las fundaciones, que llenaron el vacío dejado cuando las empresas de medios de comunicación dejaron de subvencionar/ patrocinar las escuelas de periodismo, están presionando cada vez más agresivamente por la reforma en los estudios. En respuesta a ello, algunas escuelas de periodismo están transformando abruptamente sus planes para enseñar nuevas formas de medios y nuevos métodos periodísticos en la entrega de las noticias.

3. Metodología

La metodología empleada en el desarrollo de este trabajo, se basa, primero, en la revisión bibliográfica sobre los conceptos fundamentales de cada aspecto, tratados desde la óptica del periodismo y comunicación en donde impera lo multimedia y digital. Se ha escogido esta técnica puesto que la investigación bibliográfica es la etapa de la investigación científica donde se explora qué se ha escrito en la comunidad científica sobre un determinado tema o problema. Es decir, ¿Qué hay que consultar y cómo hacerlo? Segundo, en la selección de aplicaciones tecnológicas necesarias para el desarrollo del periodismo de datos y, por consiguiente, desarrollos algorítmicos que emergen en los principales medios digitales, verificando y contrastando las diversas utilidades de las herramientas seleccionadas en el ámbito de la investigación periodística.

Para ello, además del método de la revisión de textos, hemos optado por utilizar los métodos de la observación directa y el estudio de caso, unas metodologías de investigación cualitativa mediante la cual se verifica o construye una teoría con base en la observación de muestras o casos seleccionados de la realidad (Sierra y Álvarez, 1998). La observación directa ha de lograr una comprensión objetiva de la realidad al garantizar la recogida de información de todos y cada uno de los indicadores como pueden ser las diversas aplicaciones tecnológicas utilizadas en este campo del periodismo que se basa en los datos. En cuanto al método del estudio del caso que permite al analista comprender el fenómeno estudiado, crear una base sobre la cual desarrollar investigaciones más extensas y estudiar sujetos o situaciones únicas, su relación e importancia en su contexto, el estudio de caso utilizado se basa en el algoritmo de noticias Quakebot de Los Angeles Times, lo que nos permitirá establecer un marco teórico y referencial que emana de las definiciones de las aplicaciones encontradas en sus sitios web. Estos métodos nos ayudarán a definir el estado de la cuestión y sus posibles extrapolaciones en la mayoría de los medios periodísticos. Finalmente, se reflexiona sobre el aspecto formativo de las universidades proponiéndose nuevos contenidos para la potenciación y consolidación del nuevo perfil profesional.

4. Análisis de los resultados.

El uso de aplicaciones es cada vez más común en las salas de redacción de las organizaciones de noticias por parte de los periodistas. Aunque algunos expertos en la exploración y desarrollo del periodismo de datos, afirman que no es bueno encasillarse con una determinada aplicación, recomiendan, no obstante, que lo importante es seguir el método utilizado en el desarrollo de cada fase (Figura 3). Lo que si queda claro es que la innovación tecnológica que realizan las empresas (de tecnología) es real y tienen el foco puesto en esta disciplina profesional y del conocimiento. Recursos tecnológicos de fácil acceso como son los que proporciona la factoría de Google dan fe del interés que el Periodismo de datos genera. Aplicaciones como *Google Refine*, *Maps*,

Google Fusion Tables, Earth, Hojas de cálculo, *Docs, Dipity*, Mapas cartográficos, etc., son sólo algunas de las herramientas que se necesitan para hacer periodismo de datos.

De ahí que los profesionales de medios y otros interesados, creadores de sitios web con visos de periodismo de datos en España, utilizan estas y otras aplicaciones que indistintamente sirven para realizar cada una de las fases del periodismo de datos. En la siguiente tabla, se detallan algunas de las herramientas utilizadas, según los *inputs* recogidos en los Cursos de Periodismo de Datos realizados en el marco de los Cursos de Verano El Escorial UCM, en 2012, 2015 y 2017.

Tabla 1

Aplicaciones de libre acceso utilizadas por profesionales y medios en el desarrollo del periodismo de datos.

Nombre de la aplicación o recurso online	Dirección web o descarga	Descripción
JUNAR	http://www.junar.com/	Captura, procesa, publica y comparte colecciones de datos de interés social, tales como presupuestos públicos, auditorías, censos, reportes financieros, o mediciones de monitoreo ambiental.
GOOGLE REFINE	https://code.google.com/p/google-refine/	Aplicación de filtrado de documentos desestructurados de Google. Inicialmente, se tenía que descargar e instalar la aplicación dentro del PC. Hay que disponer de una cuenta de Google.
TABLEAU (PUBLIC)	http://www.tableausoftware.com/public/	Es una herramienta gratuita de visualización de datos. Ayuda a crear modelos de contenido interactivo, que luego se puede incrustar en un sitio web o compartirlo. Transforma bases de datos y hojas de cálculo en fuentes de investigación.
FUSION TABLES	https://support.google.com/fusiontables/answer/2571232	Servicio de Google Drive que permite subir y compartir archivos en la nube; colaborar con otras personas; hacer mapas de manera rápida y fácil a partir de los datos. Fusion Tables es una herramienta web, enfocada a trabajar de forma colaborativa en la nube.
TREEMAPS	http://www.treemap.com/	Recurso de pago que desarrolla soluciones de visualización interactivas.
DRIVE/DOCS	https://www.google.com	Es el concepto sobre documentos en la nube (Cloud computing) al igual que otros recursos como DropBox. Inicialmente, Google Drive proporcionaba 5 GB gratis. Su escalabilidad actual supera los 25 GB. Su complemento con la aplicación DOC y SPREADSHEETS. Así se añade la función de compartir documentos, hojas de cálculos, imágenes y carpetas. Permite realizar encuestas y escribir textos en línea con varios colaboradores a la vez.

Fuente: información recogida de los sitios web de aplicaciones (2017). Elaboración propia.

Tabla 1 (Cont.)

Aplicaciones de libre acceso utilizadas por profesionales y medios en el desarrollo del periodismo de datos.

MAPS COORDINATE	http://www.google.com/enterprise/mapsearth/products/coordinate.html	Google Maps Coordinate es una herramienta de gestión de empleados que mejora la eficiencia de los equipos móviles. Al obtener visibilidad en tiempo real de la ubicación de los equipos y de las tareas que están realizando, el trabajo se puede programar de forma más inteligente y eficaz.
SPREADSHEETS	http://www.google.com/sheets/about	Son hojas de cálculo de Google. Otro recurso que se ha incorporado al pack de Drive, junto con Docs. Tiene fórmulas incorporadas, tablas dinámicas y opciones de formato que ahorran tiempo y simplifican las tareas habituales con las hojas de cálculo.
GOOGLE MAPS	http://www.google.com/maps/about/	Uno de los potentes recursos de Google para crear historias geolocalizadas, imprescindible también para la creación de mashups
COMETDOCS	http://www.cometdocs.com/	Convertidor de archivos PDF a Word, Excel, PowerPoint, etc. Convierte varios formatos a PDF. Es un Sistema Online de Gestión de Documentos
PDEFTOEXCELONLINE (actual NITRO)	https://www.pdfexcelonline.com/	Convertidor de PDFs a Excel. Convierte archivos PDF desde y hacia cualquier formato de Microsoft Office. Edita cualquier documento, incluyendo imágenes, párrafos y páginas.
ZAMZAR	http://www.zamzar.com/	Convertidor Online de formatos a ficheros PDF.
OCROPUS	https://code.google.com/p/ocropus/	Es un sistema OCR escrito en Python, centrándose en el uso de la máquina de aprendizaje a gran escala para abordar los problemas en el análisis de documentos. Pertenece a Google.

Fuente: información recogida de los sitios web de aplicaciones (2017). Elaboración propia.

Respecto al diseño y creación de algoritmos, no pasa lo mismo que con las aplicaciones, recurriendo en muchos casos, a expertos informáticos. Este es, por tanto, el desafío de los medios y de los profesionales. La transversalidad de conocimientos de los periodistas da lugar al algoritmo de conocimientos que estos y otros profesionales deben asumir cuando se incorporen a los medios u organizaciones de noticias. Aunque siempre habrá quien mantendrá que su selección y sus criterios -a menudo cerrados y no explícitos- es la identidad a la que muchos deben adherirse.

Confiamos en que los periodistas seguirán esta tendencia. Por ello, es también un reto de las universidades en la formación periodística (como se trata en otros apartados). En la siguiente Figura 2, se propone un boceto de algoritmo informativo de los medios de comunicación.

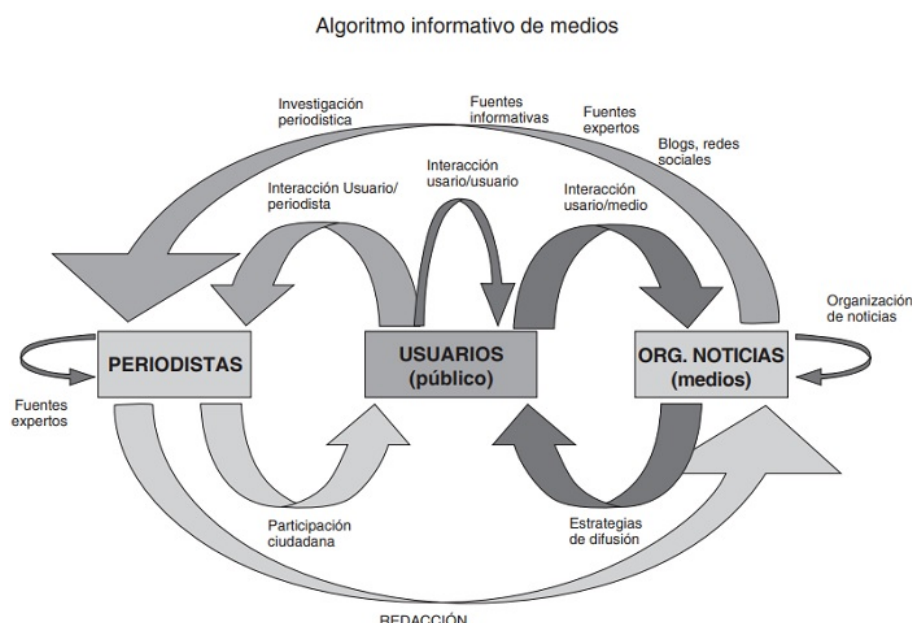


Figura 2
Modelo de flujograma algorítmico aplicado a los medios
Fuente: Elaboración propia

Distintas aplicaciones hacen que el periodismo de datos funcione. Así, este tipo de periodismo, ha de pasar por distintas fases y en cada fase, ha de utilizarse un tipo de recurso o aplicación informática que nos lleve hasta el objetivo de la publicación de la información.

Dichas fases son: la obtención de datos; limpieza de datos; análisis; verificación y visualización.

En la Figura 3, se proponen las fases del Periodismo de datos que podrían desarrollarse con diversas aplicaciones y cuyo aprendizaje, en las facultades y escuelas de periodismo, puede verse complementado, como en el caso de España, por la Ley de Acceso a la Transparencia. En este sentido, los contenidos formativos en la enseñanza de aplicaciones para este tipo de periodismo deben modularse en función de estas fases.



Figura 3

Esquema sobre las fases del Periodismo de datos

Fuente: Esquema formativo en la asignatura de Periodismo de datos en la UCM (2017). Elaboración propia.

Según Flores y Salinas (2012, p.32) los cambios formativos pasan por buscar mayor acercamiento entre las organizaciones de noticias (empresas) y la universidad. En este emergente ecosistema formativo, no se trata de quien busca a quien sino que ambas se necesitan, por lo que una estrategia puede ser la creación de unidades de trabajo (Task force) encargados de diseñar el nuevo orden curricular. Muchas veces los cambios en la currícula de estudios están realizados única y exclusivamente por académicos, sin tenerse en cuenta lo que realmente demanda el mercado de los medios para el ejercicio del nuevo periodismo. Debemos tener claro que el nuevo diseño curricular de la formación periodística debe darse entre los medios y las facultades de comunicación.

5. Resultados y Conclusiones

El análisis realizado, arroja unos resultados que dejan entrever, en primer lugar, el escaso conocimiento de los medios para desarrollar los nuevos modelos informativos fagocitados por los océanos de datos (Big data) o, en el mejor de los casos, una exploración lenta de los nuevos entornos del periodismo de investigación. No obstante, en contrapartida, otras voces expertas lanzan mensajes acerca de un nuevo modelo informativo, con grandes oportunidades de desarrollo y expansión basado en el uso de los datos.

El algoritmo se ha convertido en la nueva esencia de la comunicación e información. Esta afirmación se deriva de las reflexiones de expertos y del análisis de medios digitales que dentro de sus equipos de trabajo cuentan con un especialista, de perfil técnico, encargado de apoyar las historias periodísticas con el desarrollo de algoritmos. En esta misma línea, se perfila una tendencia en el desarrollo de aplicaciones periodísticas cuyo objetivo es llegar a un público que interactúa cada vez más.

Desde la perspectiva de la formación, aprendizaje y diseño curricular, en España y otros países, las universidades están empezando a hacer algunos cambios, aunque también con bastante lentitud. La clave frente a esta revolución del periodismo de datos está en la rapidez para adaptarse a los cambios, algo de lo que carecen gran parte las universidades españolas en donde se imparte esta disciplina.

Las tecnologías digitales deben tratarse como un modelo de enseñanza-aprendizaje transversal y específica en las escuelas de periodismo, impartándose durante todo el tiempo de duración de estudios de esta disciplina. De esta forma, se podría profundizar en el uso práctico de herramientas y aplicaciones más avanzadas. El binomio Periodismo + Programación, y con ello, la inmersión y aprendizaje del mundo algorítmico, debe pasar de ser un paradigma a un estado natural de las ciencias de la comunicación y, por consiguiente, de los estudios de periodismo.

Una conclusión en la que parecen coincidir estudiantes de periodismo, profesionales y organizaciones de noticias es que se ven buenas perspectivas de desarrollo profesional en el ámbito del periodismo de datos. La formación, mediante unos planes de estudio acordes con la demanda de la industria, serán elementos diferenciadores de la calidad.

Bibliografía

- Bradshaw, P. & Rohumaa, L. (2011). *The Online Journalism Handbook: Skills to Survive and Thrive in the Digital Age*. London, UK: Longman Practical Journalism Series.
- Cooley, M. (1987). *¿Arquitecto o abeja? el precio humano de la tecnología*. London: The Hogarth Press.
- Elías, C. (2015). *Big data y periodismo en la sociedad red*. Madrid: Editorial Síntesis.
- Fama, A. (2011). *Open Data e Data Journalism*. Roma: Edt.Narccisa. Lsdi.
- Finn, E. (2018). *La búsqueda del algoritmo. Imaginación en la era de la informática*. Barcelona: Alph Decay Ed.
- Flores, J.M. y Cebrián, M. (2012). El data journalism en la construcción de mashups para medios digitales. Zaragoza, *XIII Congreso Periodismo Digital*. Edit. APA.
- Flores, J. M. y Salinas, C. (2012). Retos de las Facultades de Comunicación: Formación de nuevos perfiles de periodistas con base en la innovación y Tecnologías del conocimiento. En *Revista de Tecnología Conocimiento y Sociedad*. Illinois: USA Common Ground Publisher.
- Guinness, H. (2017). How Facebook's news feed sorting algorithm works. Recuperado de: <https://www.howtogeek.com/290919/how-facebook-news-feedsorting-algorithm-works/>
- Joannès, A. (2011). *Data journalism: bases de données et visualisation de l'information*. Paris: CFPJ Éditions.
- Kelly, K. (2011). *What technology wants?* New York: Penguin.
- Kovach, B y Rosenstiel, T. (2012). *Los elementos del Periodismo*. Barcelona: Aguilar.

- Mayer-Schönberger, V. & Cukier, K. (2013). *Big data. La revolución de los datos masivos*. Madrid: Turner publicaciones.
- O'Neil, K. (2017). *Armas de destrucción matemática. Como el big data aumenta la desigualdad y amenaza la democracia*. Madrid: Capitan Swing libros, S.L.
- Ontiveros, E. y Lopez, V. (2017). *Economía de los datos. Riquezas 4.0*. Madrid: Fundacion Telefonica.
- Open Knowledge Foundation (2010). *The Data Journalism Handbook*. European Journalism Centre and the (R.U.). O'Reilly Media.
- Sierra, V. M. y Álvarez de Zayas, C. (1998). *La investigación científica en la sociedad del conocimiento*. México: Willey.
- Valle, A. (2014). Periodismo y algoritmos. *El Mundo*. (recuperado el 03-02-2018). Disponible en: <https://www.larazon.es/opinion/tribuna/periodismo-y-algoritmos-MX8202428>
- Varela, J. (2009). Redes, plataformas y algoritmos. En: *La revolución de la prensa digital. Cuaderno de Comunicación Evoca* (Dir. Jose Cerezo), pp: 23-28. Madrid: Evoca Comunicación e Imagen.

Recursos web

- Premio Datajournalism de Google (2012): <http://datajournalismawards.org/>
- Hechinger Report dirigida por Jon Marcus. <http://hechingerreport.org/>
- Knight Center de la Universidad de Texas (EE.UU.) <https://knightcenter.utexas.edu/es/blog/00-15307-periodista-utiliza-algoritmo-para-escribir-nota-sobre-terremoto-en-solo-minutos>
- Blog Cultura de red del Sistema de divulgación científica Madrid. <http://www.madrimasd.org/blogs/culturadered>