

Periodismo de datos y transparencia al margen de los grandes medios: Un estudio comparativo de *Civio* y *ProPublica*

*Data journalism and transparency in nonprofit
newsrooms: a comparative study of Civio and
ProPublica*

Helena M. Cortés del Álamo

*Profesora de Periodismo
(Universidad Carlos III de Madrid)*

María Luengo Cruz

*Profesora titular interina de Periodismo
(Universidad Carlos III de Madrid)*

Carlos Elías

*Catedrático de Periodismo
(Universidad Carlos III de Madrid)*

Fecha de recepción: 15 de abril de 2018

Fecha de revisión: 12 de junio de 2018

Fecha de publicación: 1 de julio de 2018

Para citar este artículo: Cortés del Alamo, H. M., Luengo Cruz, M. y Elías, C. (2018).

Periodismo de datos y transparencia al margen de los grandes medios: Un estudio comparativo de Civio y ProPublica, *Icono 14*, 16 (2), 66-87. doi: 10.7195/ri14.v16i2.1177

Resumen

El auge del “big data” ha colocado al periodismo de datos en el centro del debate. Mediante un análisis comparativo de entrevistas con profesionales dedicados a esta especialidad en la americana ProPublica y la española Civio, este artículo busca esclarecer el modo en el que dichos informadores perciben la relación entre periodismo de datos y transparencia (tanto institucional como periodística) a ambos lados del Atlántico. El estudio destaca que, aunque la percepción del trabajo y los procedimientos son muy similares en ambas redacciones, el mayor acceso a los datos públicos y las leyes de mecenazgo americanas hacen que Estados Unidos vaya un paso por delante en la incorporación del periodismo de datos.

Palabras clave

Periodismo de datos - Transparencia - Democracia - ProPublica - Civio - Análisis comparativo

Abstract

The rise of “big data” has brought data journalism at the centre of the public debate. Drawing from in-depth interviews with data journalists working at the American ProPublica and the Spanish Civio, this article aims to explore journalists’ perceptions about the relationship between data journalism and transparency (both institutional and journalistic) in the two countries. The article shows that, although interviewees from both media organisations agree in the way they perceive transparency in relation to their work, the implementation of transparency laws and sponsorship in the US allows this country to go ahead in adopting data journalism.

Key Words

Data journalism - Transparency - Democracy - ProPublica - Civio - Comparative analysis

1. Introducción

El periodismo, al menos en su vertiente más tradicional, no pasa por su mejor momento: cierre de medios, precariedad laboral y una transición a la era digital que aún no ha logrado encontrar un modelo económicamente solvente capaz de sostener un cuarto poder fuerte e independiente. Tal y como explica *The State of Media*, el último informe anual del Pew Research Center sobre el estado de la industria periodística en Estados Unidos, el número de empleos en el sector ha bajado hasta los 33.000, el mayor descenso desde 2009 (Pew Research Center, 2016). En España, la situación es similar. Tras alcanzar máximos superiores a los 10.000 periodistas en paro durante la crisis de 2008, las cifras han mejorado ligeramente hasta situarse en los 7.890 parados en 2016, niveles del segundo año de la crisis (APM, 2017). Pero esta, según las mismas fuentes, también ha provocado problemas como el aumento de la carga de trabajo y dificultades para ejecutarlo correctamente. El conjunto de estos datos revela un mercado de los medios digitales en España expansivo, muy diversificado y con un notable grado de consolidación. A pesar de que el modelo de negocio sigue siendo el principal caballo de batalla, los cybermedios han sabido nadar contracorriente y, en un contexto económico adverso, se han expandido en número, al tiempo que han diversificado su tipología. Autores como McChesney y Pickard (2011) creen que los medios están sometidos a cambios financieros, legales y tecnológicos que, en el contexto actual de crisis económica, ofrecen también una oportunidad para refugiarse de las malas rachas económicas, porque la demanda de información seguirá existiendo.

La digitalización también está modificando todas las facetas de la profesión y propiciando la aparición de nuevos roles tecnológicos y nuevas especialidades como editor gráfico, editor de vídeo, portadista, especialista en la experiencia de usuario o periodismo de datos, especialidad no muy extendida que gozará de mayor aceptación en los nativos digitales (APM, 2017). De hecho, en los últimos años se han desarrollado varias iniciativas sin ánimo de lucro, como *Civio* y *ProPublica*, que buscan recuperar ese lugar central del periodismo en la vida social a través de reportajes de investigación útiles e interesantes que aprovechan todas las herramientas que brindan las nuevas tecnologías y el “big data”. Tras la crisis de 2007, los medios pasaron

de un panorama de incertidumbre económica a un modelo de cambio sin precedentes con consecuencias en lo económico, industrial y cultural (Franklin, 2016). De hecho, estos medios son herederos de la explosión del periodismo de datos.

No obstante, para poder trabajar con datos es muy importante que estos sean accesibles y estén expresados en estándares abiertos que faciliten su uso. En Estados Unidos, la ley que garantiza el acceso de los periodistas a la información pública es la Freedom of Information Act (1966). También la Unión Europea publicó la Directiva 2003/98/EC de noviembre de 2003, que anima a los distintos gobiernos a facilitar el uso de información pública. A raíz de esta norma comunitaria nació, en España, la Ley de Transparencia, acceso a la Información Pública y buen gobierno (2013), que tiene como objetivo “garantizar el derecho de los ciudadanos a acceder a la información, obligar a las distintas administraciones a ser transparentes y fiscalizar la actividad pública”.

Este artículo compara la relación entre periodismo de datos y transparencia en dos organizaciones sin ánimo de lucro que emplean el “big data” en sus prácticas y productos informativos: la americana *ProPublica* y la española *Civio*. En el primer apartado, se esboza el marco teórico para explicar qué se entiende por transparencia en el contexto del trabajo periodístico con datos. La revisión bibliográfica conduce a una doble reflexión de la transparencia como condición necesaria de las instituciones sociales para poder desarrollar este tipo de periodismo, y como uno de los rasgos más destacados que lo definen. Después, se desarrolla el estudio exploratorio de *ProPublica* y *Civio* a través de entrevistas en profundidad con el equipo de profesionales dedicados al periodismo de datos en ambos medios. La discusión se plantea en términos comparativos y con base a estos objetivos específicos:

1. Esclarecer cómo los encuestados perciben la transparencia de las instituciones públicas que han de suministrar los datos.
2. Examinar las opiniones de los entrevistados sobre la propia claridad y transparencia en la visualización de sus informaciones.
3. Analizar el modo en que los periodistas describen la influencia de su trabajo en la creación de una sociedad más participativa y democrática.

1.1. Desde la objetividad a la transparencia en el periodismo de datos

El periodismo de datos no es un invento moderno. A principios de siglo, Max Weber en Alemania y Walter Lippmann y Robert E. Park en Estados Unidos ya plantearon las primeras sinergias entre el trabajo del periodista y el científico social (Dader y Gómez, 1993). En su libro *Liberty and the news*, publicado en 1920, Lippmann ya hablaba de una crisis en el periodismo asociada a la intención de los medios de no trabajar con datos objetivos, sino con opiniones. Con el objetivo de acercarse lo más posible al ideal de libertad, este teórico proponía “asegurar una investigación de los hechos tan imparcial como sea posible” (Lippmann, 1920:71).

En los años treinta, la revista *Fortune* publicó la primera encuesta científica realizada directamente por una empresa periodística y en 1939 *Reader's Digest* comenzó a realizar experimentos de campo (Dader, 1997). Más tarde, periodistas como Meyer comenzaron a sentirse fascinados por las técnicas metodológicas empleadas en el análisis de las ciencias sociales y empezaron a plantearse si estas podrían aplicarse al campo periodístico para intentar acercarse a ese ideal de objetividad que guiaba la investigación sociológica (Meyer, 1973). El periodismo de precisión, en palabras de su ideólogo, “aplica los métodos científicos de investigación social y comportamental a la práctica del periodismo” (Meyer, 1973:6). Los teóricos Weaver y McCombs analizaron el fenómeno y concluyeron que había dos fases en la implementación del periodismo de precisión en las redacciones. En la primera, los periodistas comenzaron a convertir en noticias los informes elaborados por sociólogos. En una segunda fase de consolidación, realizaban los experimentos con métodos sociológicos o científicos o empleaban esas mismas técnicas para evaluar el rigor de los trabajos que les llegaban de especialistas (Dader, 1997).

Las teorías de Meyer encontraron un buen campo para su desarrollo en lo que los anglosajones llamaron CAR (“computer-assisted reporting” o reportero asistido por un ordenador). En este caso, los ordenadores se emplean como una herramienta clave para gestionar, almacenar y extraer historias de enormes bases de datos: “La combinación de potentes ordenadores con las posibilidades de la World Wide Web ofrece la posibilidad de almacenar y analizar enormes plataformas de datos

para descubrir ideas para una historia o añadir hechos y pruebas a una historia que acaba de ser investigada por métodos clásicos de periodismo” (Wormer, 2007:394).

Las primeras investigaciones con ordenadores primigenios tuvieron lugar a finales de los setenta, cuando se fundó la Organización de Periodistas de Investigación de Estados Unidos (IRE). A partir de los años ochenta, comenzaron a llegar también los primeros reconocimientos en forma de prestigiosos premios. En 1985, el *Dallas Morning News* logró el primer Pulitzer por un artículo elaborado con las técnicas del periodismo de datos en el que denunciaba la segregación racial en la adjudicación de viviendas de protección pública en Texas.

El ordenador, y más tarde el surgimiento de internet ya en los noventa, ha conseguido facilitar varias tareas que sin él sería imposible realizar o al menos poco práctico aventurarse en ellas, como hacer preguntas concretas a una base de datos, explorar tendencias creando rápidas visualizaciones y encontrar lo imposible, ya que en muchas ocasiones las notas de prensa evitan airear ciertas informaciones (Leonard, 1992).

A principios de los 2000, el término CAR cayó en desuso en favor del concepto periodismo de datos. Mientras el primero se veía como una herramienta auxiliar para ayudar a las investigaciones a largo plazo, el segundo, más empleado en el mundo periodístico, se refiere a colocar estas técnicas informáticas en el centro de la investigación como forma de trabajo principal (Coddington, 2015). Además, el periodismo de datos pone mayor énfasis en la visualización.

Ya en el ámbito de la investigación académica actual, Martinisi (2013) afirma que el periodismo de datos se basa en tres pilares: el auge de la tecnología, la aparición de nuevas prácticas periodísticas (destrezas matemáticas, habilidades como buscar en la web profunda, saber gestionar los formularios de petición de la información, organizar y limpiar los datos...) y un tercer motivo relacionado con la implicación ciudadana. La sociedad civil, mantiene Martinisi (2013:60), “es cada vez más receptiva a la importancia de tener datos abiertos”.

No obstante, el periodismo de datos tiene un problema: muchas veces los datos solo son datos (estadística) y no existe un relato para comprenderlos (Elías, 2015).

Para solucionarlo, hubo un intento en *The Guardian* en el que usaron el periodismo de datos para explicar las revueltas que hubo en Gran Bretaña en 2011, pero se basaron en la Teoría Fundamentada para buscar un relato. La Teoría Fundamentada es útil para generar una hipótesis a partir de los datos obtenidos en una investigación. Es decir, nos puede ayudar a obtener una explicación –un titular- cuando no tenemos una hipótesis predeterminada (Elías, 2015). La Teoría Fundamentada apareció en 1967 de la mano de dos experimentados sociólogos estadounidense, Barney Glaser y Anselm Strauss, en su ya clásico “The Discovery of Grounded Theory”.

El pionero del periodismo de precisión, Phillip Meyer, escribió una columna en *The Guardian* titulada “Riot theory is relative (Meyer, 2011)” donde analizaba la aplicación de la Teoría Fundamentada a los reportajes de periodismo de datos sobre las revueltas inglesas de 2011. Meyer sostiene que los periodistas que cubren historias importantes necesitan más y mejores herramientas y recuerda que si su éxito, en la cobertura de los disturbios de Detroit de 1967, fue el uso de encuestas y los emergentes programas informáticos, el éxito del reportaje de *The Guardian* “Reading the riots” radica no tanto en lo que obtuvieron, sino en la exploración de si la Teoría Fundamentada, aplicada por los científicos sociales de la London School of Economics, puede considerarse una herramienta para construir una narrativa –un “storytelling”- que explique por qué los ingleses se manifestaban en 2011. En este sentido, señala que el proyecto de *The Guardian* ha sido “más audaz y difícil” de lo que fue el suyo (Meyer, 2011).

Muchos informáticos, atraídos por la idea romántica de que las habilidades digitales también pueden dedicarse al bien común, se han unido a estos equipos multidisciplinarios (Cohen, Li, Yang y Yu, 2011). En estudios recientes, Anderson (2012) argumenta que el uso de estas nuevas tecnologías depende tanto de las herramientas digitales disponibles como de las instituciones sociales que dan acceso a los datos.

En su mejor versión, el periodismo de datos, según los profesores Parasie y Dagiral (2013), fortalece la democracia principalmente de tres formas: estrecha el concepto de objetividad periodística (permitiendo contrastar las distintas versiones), ofrece nuevas herramientas para vigilar el trabajo del gobierno y ayuda a los ciudadanos a participar en política con sus propios análisis de datos.

Sin embargo, tal y como apuntan académicos como Schudson (2010), no hay que olvidar que los avances tecnológicos en sí mismos, como ya ocurrió con la aparición de la imprenta y el telégrafo, no contribuyen a crear una sociedad más participativa y democrática: “La tecnología no produjo esto sola. Es difícil imaginar la historia sin movimientos civiles” (Schudson, 2010: 104).

1.2. El “big data” y la ética de los datos

Si hablamos de periodismo, debemos tener en cuenta algunos aspectos éticos del “big data” analizados entre otros por Boyd y Crawford (2012). El primero de todos es que el “big data” cambia la definición de conocimiento. Igual que el telescopio modificó el conocimiento sobre la posición de la Tierra, el acceso a grandes cantidades de datos cambia la percepción de la realidad que teníamos hasta ahora.

Por otro lado, los periodistas deben tener cuidado con la maliciosa tendencia a ver patrones, conexiones o ambos en sucesos aleatorios o datos sin sentido. Leinweber (2007) demostró que usando técnicas de minería de datos podemos demostrar una correlación espuria entre la producción de mantequilla de Bangladesh y los cambios del índice Standard & Poor’s 500. Que parezca que hay relación no quiere decir que exista. El periodismo de datos y de grandes datos es relevante, pero no puede eclipsar la importancia de otras formas de obtener datos: entrevistas u opiniones de personas relevantes que nos explican mejor la realidad.

Otro problema es que “acceso fácil” no implica que sea ético o legal. El artículo 18.4 de la Constitución Española (1978), establece que “la ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos”.

El “big data” genera nuevas divisiones sociales. Como ha demostrado el caso de Cambridge Analytics, compañías como Google, Twitter o Facebook son propietarias de los datos de sus usuarios y deciden qué datos comparten libremente, cuáles venden y cuáles se reservan. Como sostiene Manovich (2011), un antropólogo que trabaje para Facebook o un sociólogo de Google tiene acceso a datos que el resto de los académicos del mundo no tendrán. En el periodismo de datos y en el “big

data” también puede existir un sesgo de género. Esta labor exige conocimientos de ingeniería informática, matemáticas y estadística, disciplinas donde las mujeres aún son minoritarias, mientras que en periodismo ya existe una paridad. Habría que potenciar el acceso de las periodistas a esta especialidad periodística.

1.3. Dificultades en la implantación

Pese a sus ventajas, el periodismo de datos no ha logrado imponerse como método de trabajo en las redacciones. Según explican Solop y Wonders (2016), hay varios motivos por los que esta forma de trabajar aún no es mayoritaria: la organización de las propias redacciones, el mercado y motivos estructurales relacionados con el poder de los gobiernos.

Otras críticas reconocen el valor de estas formas de “periodismo administrativo” (Kreiss, 2015), pero consideran que hablar de los datos y los modelos técnicos y ha limitado la conversación actual sobre qué es el periodismo. “La información es solo uno de las siete u ocho contribuciones del periodismo a la sociedad, a la que también ofrece investigación, análisis, empatía social, foros públicos y formas de movilizar al público (Kreiss, 2015:64)”. Kreiss afirma que, aunque estos nuevos modelos apuesten por la innovación, no está claro que lleven a una mayor pluralidad. De hecho, explica, la audiencia se siente más involucrada en las noticias cuando estas van más allá de la mera información y se apuesta por el análisis y la crítica (Kreiss, 2015).

En lo que sigue, el artículo lleva a cabo un análisis exploratorio de los principales argumentos sobre transparencia y periodismo de datos esgrimidos por Martinisi (2013), Schudson (2010), LaFleur (2014) y Anderson (2012). La comparación permitirá observar las diferencias y similitudes en la percepción de periodistas de datos en *ProPublica* y *Civio*, dos organizaciones periodísticas diferentes. En concreto, el estudio trata de responder a estas preguntas: ¿Qué grado de importancia tiene la transparencia de las instituciones en el desarrollo de este tipo de periodismo? ¿Cómo aplican los profesionales del periodismo este concepto sus investigaciones? ¿Consideran realmente que los datos son objetivos y no manipulables? ¿En qué medida el papel del público es relevante para la reconstrucción de sus historias? ¿En qué grado estas formas de periodismo fomentan el espíritu crítico en la audiencia?

2. Material y métodos

ProPublica ha sido seleccionado por su posición como medio de referencia dentro del índice del Global Investigative Journalism Network¹. Su popularidad, la máxima alcanzada por un medio sin ánimo de lucro según el índice Alexa², viene también avalada por dos premios Pulitzer (2010, 2011). En el caso español, se ha seleccionado el único de esta lista de medios de investigación: la fundación ciudadana Civio, distinguida con el Premio iRedes Categoría Institucional (2015).

Para el estudio de *ProPublica*, se ha entrevistado a miembros todos los perfiles profesionales que, en 2016, formaban parte del equipo de creación de aplicaciones de noticias “news applications”, el departamento de *ProPublica* encargado del periodismo de datos: Jeff Larson (entonces director adjunto), Ryann Grochowski Jones (periodista ahora responsable del área de datos), Derek Willis y Sisi Wei (desarrolladores, esta última hoy responsable también del área de “News applications”). Por su parte, en Civio, una organización mucho más pequeña, se ha ampliado la diversidad muestra y se han realizado entrevistas a David Cabo, fundador, Eva Belmonte, responsable de proyectos, Miguel Ángel Gavilanes, periodista, y Raúl Díaz, desarrollador. Por último, se ha entrevistado a analistas destacados como Mar Cabra, impulsora del periodismo de datos en nuestro país y editora en el ICIJ (International Consortium of Investigative Journalists), antigua directora de Civio, y José Luis Dader, investigador del ramo en la Universidad Complutense de Madrid.

Se han empleado las entrevistas en profundidad propias de una metodología cuantitativa porque se buscaba ver, con exhaustividad, todas las perspectivas de los trabajadores dedicados a esta labor. De esta forma, es posible indagar en el trasfondo de los sujetos investigados, se obtiene información pormenorizada sobre los valores y experiencias de quienes responden y se puede acomodar el cuestionario al entrevistado al encuestado particular (Taylor y Bogdan, 1992).

3. Resultados

3.1. El caso de *ProPublica*

ProPublica fue fundado en 2007 por Paul Steiger, aunque comenzó a publicar sus primeras investigaciones con datos en 2008. Se declara no partidista ni afín a ninguna ideología, sino estrictamente apegado a los ideales de imparcialidad periodística. Su misión, explican, es promover el periodismo de investigación que interese al público. De hecho, la mayoría de sus investigaciones están relacionadas con cuestiones de sanidad, educación, seguridad y economía y fomentan activamente la participación del lector, a quien dan acceso directamente a los datos. En cuanto a su modelo de financiación, destacan los subsidios, grandes donaciones y las aportaciones de los lectores (en total, más de quince millones de euros). Además, explican, también aceptan publicidad y exploran constantemente nuevos sistemas de obtención de ingresos, como la venta de datos y libros. “Una organización como *ProPublica* no sería viable en España, porque es lo que se denomina *non profit organizations*, organizaciones sin ánimo de lucro cuyos donantes se desgravan el 100%. En nuestro país esta cantidad es el 25% o el 35%. Además, en Estados Unidos tienen la cultura de la transparencia más interiorizada (*accountability*). En España no somos conscientes de que esto también puede ayudar a cambiar las cosas”, apunta Mar Cabra (entrevista 1). Dader (entrevista 2) también afirma que en Estados Unidos hay una mayor cultura de la transparencia, lo que anima al desarrollo de este tipo de proyectos. “Las leyes de transparencia ayudan, por eso en Estados Unidos es un gran fenómeno y en España casi un tabú. Aquí no hemos tenido casi hasta 2013 una ley de este tipo, y hay que ver hasta dónde llega su reglamento”, afirma este experto.

Conscientes de la importancia del periodismo de datos, *ProPublica* tiene un equipo (*news applications team*) volcado con el periodismo de datos. Scott Klein, entonces director de este equipo y director adjunto, coordinaba a periodistas con distintos perfiles: editores de datos, desarrolladores y reporteros. Sin embargo, según los entrevistados, todos realizan labores tanto de limpieza y adecuación de los datos como análisis y redacción: “El tiempo que pasamos en cada tarea depende del momento del proceso en el que estemos: yo puedo estar escribiendo un artículo un

día o programando y diseñando la web del proyecto otro”, explica la desarrolladora Sisi Wei (entrevista 3), responsable adjunta del área. También los reporteros, como Larson (entrevista 4), son trabajadores multitarea.

La periodista Ryann Grochowski Jones, reportera y responsable adjunta del área de datos, afirma que una de las grandes diferencias del medio respecto a otros es el tiempo que dedican a cada proyecto, que va desde varios meses hasta incluso años: “Esto nos permite realizar un análisis completo, bien pensado e investigado. A veces no es factible para una organización de noticias formar un grupo similar, así que al menos hay que asegurarse de que todos tus reporteros son capaces de entender y trabajar con datos”, afirma Grochowski (entrevista 5). Tal y como explica el profesor Dader, *ProPublica* es “la continuación de ese gran periodismo de precisión de los ochenta y noventa. Tienen la concepción que había antes del periodismo de datos”, añade.

Otro de los puntos analizados en esta investigación es la presentación de los datos. Todos los profesionales coinciden en que estos, por sí mismos, son objetivos, “aunque es posible cogerlos para que digan lo que quieres que digan”, mantiene Wei. Esto se puede conseguir alterando la forma de presentarlos. “Siempre hay cierto peligro, incluso limpiando los datos, pero hay que tener cierto compromiso con lo que haces. Si estoy buscando una tendencia soy más proclive a ver esa que otras”, apunta Wei. *ProPublica*, con el objetivo de evitar estos errores, realiza test estadísticos para afirmar que lo que cuentan es real e incluso cambian la historia si los datos no la prueban. Además, siempre realizan sus comprobaciones a posteriori: aplican un protocolo de comprobación básico (<https://github.com/propublica/guides/blob/master/data-bulletproofing.md>), comparten su metodología y datos y piden a un compañero que reproduzca su investigación.

Los profesionales de *ProPublica* también puntualizan que la transparencia está directamente relacionada con la accesibilidad a los datos. Wei observa lo siguiente: “Gran parte del trabajo que hacemos, incluso con datos que ya son accesibles, es presentarlos de forma que sean fáciles de usar. La transparencia no solo quiere decir poner a disponibilidad del público nuestras aplicaciones de noticias, sino también ponerlos de tal forma que la búsqueda en ellos sea sencilla”.

Los periodistas insisten en un aspecto que muchas veces se olvida en relación con los datos: su depuración. “No recuerdo ninguna base de datos que no haya necesitado ningún proceso de limpieza. Tienes que entrevistar a los datos como a las fuentes”, dice Grochowski.

Este pensamiento común respecto al trabajo con los datos no es tan unánime cuando la conversación se centra en la relación con las instituciones. Jeff Larson insiste en las ventajas de tener una ley de transparencia como ocurre en Estados Unidos: “Hay leyes (Freedom of Information Act) muy efectivas. Normalmente podemos solicitar los datos que buscamos, las agencias tienen que evaluar al menos nuestra solicitud y lo más normal es que usemos esta ley para obtenerlos”.

No obstante, sus compañeras señalan los agujeros negros de esta normativa, que, por ejemplo, no se aplica a empresas privadas. “Además, si los organismos no la siguen no hay alternativa. Tienes que demandar a un alto tribunal y no merece la pena”, sostiene. Su compañera Grochowski puntualiza que también es importante el manejo de los tiempos: “A veces una petición se convierte en un largo toma y daca entre el reportero y la agencia estatal”.

Todos los periodistas entrevistados integrantes de *ProPublica* creen firmemente que el periodismo de datos contribuye a crear una sociedad más transparente, ya que da al público las herramientas necesarias para que sean ellos mismos quienes prueben la historia. En palabras de Grochowski, y en línea con las tesis de Meyer: “Las aplicaciones de noticias son la mejor forma de ayudar a las organizaciones periodísticas a mejorar la transparencia, por la mera razón de que estás permitiendo a la audiencia examinar los datos cuidadosamente contigo. No estás contando algo, estás dando al público el poder de investigar”.

Otra de las ventajas de trabajar con datos, según el desarrollador Derek Willis, es que estos permiten también ir al detalle: “Pueden mostrar tanto el dibujo y las conclusiones generales como las pequeñas situaciones individuales. Además, nuestros *datasets*, que también publicamos, pueden ser usados por otros medios para crear sus propias historias. Cerca de cien diarios han escrito historias basándose en nuestros datos”, subraya Larson (entrevista 6).

Respecto a la relación del periodismo de datos y la participación, los profesionales de *ProPublica* defienden que un mayor acceso a los datos anima a la sociedad a involucrarse en los temas que preocupan. “El objetivo es poner en manos del público datos que les ayuden a tomar decisiones más adecuadas”, dice Wei. Los entrevistados emplean como ejemplo uno de sus proyectos, que tenía como objetivo descubrir en qué centros de emergencias atendían antes. Willis también subraya el papel didáctico de sus trabajos: “Contribuyen a crear una sociedad más democrática en la medida de que ayuda a la gente a entender el papel de las instituciones esenciales en sus vidas”.

Finalmente, los entrevistados también consideran que el periodismo de datos es esencial en para el futuro del periodismo de investigación: “Uno de los mayores problemas del periodismo de investigación es obtener fuentes que quieren hablar, cuando en la mayoría de los casos se pueden usar datos como fuente”, insiste Larson. Por otro lado, sostiene Wei, tener este tipo de habilidades permite a los periodistas no depender de expertos: “Si no sabes cómo usarlos no podrás comprobar que lo que te dicen las fuentes oficiales es cierto”.

3.2. El caso de *Civio*

Civio es una organización sin ánimo de lucro creada en 2011 por Jacobo Elosua y David Cabo, que decidieron acercar a nuestro país iniciativas de periodismo ciudadano como *mySociety* en Reino Unido, *Ciudadano Inteligente* en Chile o la *Sunlight Foundation* estadounidense. En 2012 pusieron en marcha sus primeros proyectos: “¿Dónde van mis impuestos?” y “Tuderechoasaber.es”. Después llegaron el “Indultómetro” (febrero 2013), “España en llamas” (2012 y 2013) y “El BOE nuestro de cada día”, que analiza el Boletín Oficial del Estado.

La plantilla está formada por diez personas de distintos perfiles (periodistas, desarrolladores, diseñadores web, gestores de márketing, captadores de donantes...). Haciendo gala de la transparencia que exigen a los poderes públicos, esta fundación publica anualmente sus cuentas. En 2016, *Civio* cerró con un déficit de 122 euros, resultado de restar los 336.578 euros de gastos (259.000 en personal) a los 336.456 euros de ingresos que generó la empresa, principalmente por su actividad

propia. Actualmente, sus dos grandes proyectos en desarrollo son “Medicamentalia” y “¿Quién cobra la obra?”. Prácticamente todos sus trabajos son muy visuales e interactivos, y siempre se preocupan por detallar la metodología empleada y dejar un repositorio con los datos empleados para que los usuarios o medios puedan bucear en busca de nuevas historias.

Lo primero que aclaran los periodistas y programadores que trabajan en *Civio* sobre la organización es que no es un medio al uso, sino “una fundación ciudadana que recurre a dos herramientas –periodismo y tecnología– para facilitar la transparencia, el acceso a la información y la participación ciudadana”, apunta su principal impulsor, David Cabo, programador (entrevista 7). “Somos proveedores de información sobre nuestras áreas de actividad, pero no somos un medio ni competimos con ellos”.

Dados los recursos disponibles, gran parte del trabajo de *Civio* con los datos se centra en actualizar los proyectos que ya tienen, como “Quienmanda.es...” y trabajar en nuevas iniciativas, como “Medicamentalia” o “El indultómetro”. Para seleccionar los proyectos es importante “que afecte a una problemática o sensibilidad social, que sea un tema que demande un debate con más datos de los disponibles y que podamos garantizar la integridad de nuestros datos”, subraya Eva Belmonte, responsable de proyectos de *Civio* (entrevista 8).

Al contrario que otros periodistas del ramo, Miguel Ángel Gavilanes no considera que sea necesario tener un conocimiento especializado para trabajar con datos. “Lo más importante es entender bien qué están contando esos datos, cómo se han recogido, qué variables representan, cuál es el nivel de detalle y las lagunas que dejan [...]. Sí es cierto que conocimientos matemáticos o estadísticos para hacer cálculos, como medianas, ayudan”, explica el periodista (entrevista 9). Todos los profesionales de *Civio*, de hecho, coinciden en la necesidad de que el equipo de trabajo sea multidisciplinar.

Civio considera la transparencia casi como un derecho, o al menos como un valor cultural y administrativo. “Implica reconocer el derecho de cualquier ciudadano u organización a recibir la información que consideren relevante, y también la obligación de proporcionarla. Debe formar parte del contrato social entre representantes y representados, y sobre eso tratamos de sensibilizar tanto a los cargos públicos

como a cualquier ciudadano”, afirma Cabo. Una idea que comparten periodistas del medio como Gavilanes, que considera que buscar historias en bases de datos con un propósito informativo puede llevarnos a lograr una sociedad más transparente.

Además, no solo hablan de la publicación en datos, sino también de su acceso en formatos con los que se pueda trabajar: “Si las instituciones facilitan datos en crudo y en los formatos adecuados que hacen posible la reutilización y la comprensión, sería un gran paso adelante”, puntualiza Belmonte. Para Gavilanes, hay formatos cerrados (PDF e imágenes escaneadas) que dicen muy poco sobre la voluntad de transparencia de las instituciones que los publican.

También hay unanimidad entre los trabajadores de *Civio* sobre la objetividad de los datos: dependen del método empleado para analizarlos. “Un dato es capaz de reflejar una visión y la contraria. No es que sean objetivos, es que son mudos. Sin un trabajo riguroso de análisis, de recopilación, investigación y sin contexto, los datos no llevan a ningún lado”, concluye David Cabo. También las visualizaciones deben estar al servicio de las historias que cuentan, subraya el desarrollador Raúl Díaz (entrevista 10), porque “de otro modo habría sesgos y no hablaría de objetividad”. Y por si algo falla, también *Civio* tiene su propia forma de comprobar una y otra vez los datos y volver a la historia (*fact-check*) antes de publicarla.

La normativa española sobre la transparencia dificulta mucho el trabajo a estos profesionales. “Aquí no tuvimos ley de transparencia hasta 2013 y muchos proyectos que en Estados Unidos pueden hacerse a golpe de *click* requieren mucho más tiempo y esfuerzo”, reconoce Cabra. En relación con las leyes de transparencia, Dader ve una paradoja en el desarrollo del periodismo de datos en nuestro país: “Paradójicamente, con leyes menos avanzadas, como la Ley de procedimientos administrativos de 1992, se hicieron grandes reportajes en los años noventa”.

En cuanto al futuro del periodismo de datos y su contribución a la investigación, los trabajadores de *Civio* apuntan tres vías distintas de evolución directamente relacionadas entre sí. David Cabo apuesta por crear más unidades especializadas en contar historias de este tipo. “Quizá, el siguiente paso debería ser que estos equipos se dedicasen 100% a la investigación, con plazos más extensos que los que demanda el día a día informa-

tivo”, reconoce Cabo. En este sentido, Eva Belmonte y Raúl Díaz creen que, aunque la tendencia es creciente, hay que esperar a que se consolide en las redacciones para hablar sobre cómo afectará a la forma de investigar en la profesión. Finalmente, Miguel Ángel Gavilanes apunta que puede ayudar a crear “metodologías de investigación más precisas”.

Tal y como afirman sus creadores, expertos como Mar Cabra inciden en que la misión de *Civio* no es idéntica la de *ProPublica*, sino que se parece más a asociaciones como *My society* en Reino Unido: “*Civio* es un 50% *ProPublica* y otro 50 % la *Sunlight Foundation*. Es más una organización que lucha por la transparencia que una productora de contenidos. Si no creábamos contenidos para animar a la gente a participar, no volvían”, recuerda Cabra.

José Luis Dader plantea que este tipo de investigaciones que requieren muchos recursos se están refugiando fuera de los medios, en organizaciones como estas. Como argumenta Cabra, con la situación fiscal actual, estos medios subsisten en gran parte de la asesoría en cuestiones de transparencia a otros organismos como ayuntamientos. “Ahora está más de moda el periodismo de precisión o datos pero en ocasiones es muy superficial, solo se vuelcan los datos y se les da dinamismo”, apunta Dader.

4. Discusión

Al comparar el trabajo de *Civio* y *ProPublica*, lo primero que hay que destacar son las diferencias en cuanto a los recursos que disponen (22,7 millones de euros de ingresos de *ProPublica* frente a los 336.578 euros de *Civio*), algo que Cabra y Dader atribuyen a las distintas deducciones que tiene cada sistema fiscal. Esta menor cantidad de presupuesto tiene su reflejo en el número de trabajadores que tiene cada medio y en los proyectos que puede sacar adelante. Mientras que *ProPublica* tiene cerca de cuarenta empleados que trabajan en unas cincuenta investigaciones, en *Civio* la redacción apenas alcanza la docena.

Sin embargo, retomando los objetivos, la ética de trabajo es similar en ambos medios. Tanto los trabajadores de *Civio* como los de *ProPublica*, en línea con los planteamientos de Martinisi (2013), valoran muy positivamente leyes de transparencia como la FOIA, pese a que también tenga sus puntos negros como los tiempos de espera o

las cuestiones relacionadas con la privacidad. Asimismo, ambos insisten en la importancia de la accesibilidad y el formato de publicación. Otra cuestión importante está relacionada con la publicación de la metodología y los datos en bruto (y cumplen estos principios) para que el público los explore o puedan ser replicados por otros diarios.

En cuanto al segundo objetivo, relacionado con su propio trabajo con los datos y su transparencia, todos coinciden en que los datos pueden ser manipulados para mostrar una y otra cosa. En sí mismos, apunta Cabo, son “mudos”. Por ello, su objetividad depende del método empleado para analizarlos y la honestidad del profesional. Siempre hay cierto peligro derivado de los prejuicios del periodista o incluso del método elegido para visualizar los datos. Sin embargo, ahí entra en juego un buen filtrado y procesamiento de los datos, así como una serie de comprobaciones que deben hacerse a posteriori. Todos los periodistas interrogados afirman que cuentan con esta comprobación o “fact-checking”. De ahí también que sea tan importante poner a disposición de los usuarios el método, para que los propios lectores puedan reconstruir el camino realizado por el periodista para llegar a las conclusiones publicadas.

Los profesionales de uno y otro lado del Atlántico divergen cuando se trata de analizar la preparación de estos profesionales. Desde *ProPublica* abogan por considerar estas técnicas casi como una especialidad y formar a todos los periodistas de la redacción en ellas. En cambio, desde *Civivo* periodistas como Gavilanes no creen que haga falta una preparación específica, sino “entender bien qué están contando esos datos, cómo se han recogido, qué variables representan, cuál es el nivel de detalle y las lagunas que dejan”, aunque reconocen que ciertos conocimientos mínimos de estadística y matemáticas ayudan.

Asimismo, todos insisten en que poner estas herramientas en manos del público ayuda a crear una sociedad más transparente y permite a los lectores, e incluso a grupos de trabajo afectados por las informaciones publicadas, investigar por sí mismos. Este tipo de periodismo, según expertos como Mar Cabra, podría ser una herramienta clave en el futuro para conseguir recuperar la credibilidad en los medios, porque ya no se insta a los lectores a que se creen lo que el medio publica, sino que se les permite que lo comprueben por ellos mismos.

En cuanto a la contribución de estas informaciones a mejorar la democracia, hay dudas entre los entrevistados. En general, las respuestas apuntan a que el periodismo de datos dota a los ciudadanos de herramientas que antes no tenía, pero no está comprobado que esto genere una sociedad más democrática y participativa, ni que impulse directamente a la movilización social.

Finalmente, ambos medios divergen sobre el papel que jugarán los datos en el futuro del periodismo de investigación. En *ProPublica* consideran el trabajo con el “big data” como una herramienta fundamental, la más importante en el desarrollo de esta disciplina en el futuro, ya que permite lograr informaciones a las que es difícil acceder si uno no tiene buenas fuentes. En cambio, desde *Civio* son más cautelosos, y creen que antes de hablar del papel de los datos hay que esperar a que se consoliden en las redacciones.

Para concluir, si analizamos las principales diferencias entre el periodismo de datos en España y en Estados Unidos, dos países que corresponden a dos sistemas de medios distintos según la clasificación de Hallin y Mancini (2004) -el sistema polarizado y pluralista, propio de sistemas mediterráneos como España, y el anglosajón o liberal, propio de Estados Unidos, Canadá, Reino Unido e Irlanda-, vemos que, aunque la filosofía de trabajo es la misma, factores legislativos e industriales separan el trabajo que realizan estos profesionales a uno y otro lado del Atlántico.

El periodismo de datos es más fuerte en Estados Unidos gracias a una legislación más favorable, mejores medios de financiación y una mayor tradición de este tipo de investigaciones. En este ecosistema, los medios son más populares (tienen más tirada o, en este caso, aceptación popular), independientes y con profesionales más formados en tareas más especializadas (Hallin y Mancini, 2004).

Aunque las leyes de transparencia son esenciales, lo principal en España es la falta de una ley de mecenazgo como la americana que anime a los ciudadanos a contribuir con estas asociaciones. Esto, sumado a la crisis de los medios tradicionales, hace que no haya una abundancia de recursos que fomente el desarrollo de estos medios. En este sistema de medios, los productos periodísticos no son tan exitosos, los poderes públicos son más intervencionistas y los periodistas no están tan especializados (Hallin y Mancini, 2004).

Sin embargo, esta investigación solo toma en consideración dos sistemas de medios. En este sentido, en el futuro las investigaciones podrían continuar comprobando la situación de este tipo de periodismo de datos y la relación entre transparencia y democracia en países representativos de otros sistemas de medios, como el germánico. O incluso se podría extender la mirada con las posteriores revisiones que han hecho Hallin y Mancini (2004) de sus propias teorías incluyendo nuevas áreas como el periodismo latinoamericano, donde periódicos como *Clarín* han demostrado su buen hacer dentro del periodismo de datos.

Referencias

- Anderson, C.W. (2012). Towards a sociology of computational and algorithmic journalism. *New Media & Society*, 15(7),1005 - 1021 Doi: 10.1177/1461444812465137
- Asociación de la Prensa de Madrid (2017). *Informe anual de la profesión periodística en España 2016*. Madrid: APM goo.gl/cUyRTb
- Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information, Communication, & Society*, 15, 662-679. Doi: 10.1080/1369118X.2012.678878
- Coddington, M. (2015). Clarifying Journalism's Quantitative Turn, *Digital Journalism*, 3:3, 331-348, Doi: 10.1080/21670811.2014.976400
- Cohen, S. ; Li, Ch.; Yang, J. y Yu, C. (2011). Computational journalism. A call to arms to database researchers. *Proceedings of the Conference on Innovative Data Systems Research*, 148-151.
- Dader, J.L. (1997). *Periodismo de precisión. La vía socioinformática de construir noticias*. Madrid: Síntesis.
- Elías, C. (2015). *Big data y periodismo en la sociedad red*. Madrid: Síntesis
- Franklin, B. (2014) The Future of Journalism. *Journalism Studies*, 15:5, 481-499, Doi: 10.1080/1461670X.2014.930254
- Glaser B y Strauss A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine
- Hallin, D y Mancini, (2004). *Sistemas mediáticos comparados: Tres modelos de relación entre los medios de comunicación y la política*. Barcelona: Hacer.

- Kreiss, D. (2016). Beyond administrative journalism: Civic skepticism and the crisis in journalism. In J. Alexander, E. Breese, & M. Luengo (Eds.), *The Crisis of Journalism Reconsidered: Democratic Culture, Professional Codes, Digital Future*, 59-76. *Cambridge University Press*. doi:10.1017/CB09781316050774.005
- LaFleur, J. (2014). Data and transparency: perils and progress in Transparency. *Politics and the media: Accountability and open government*. Londres: Reuters Institute
- Leinweber, D (2007). Stupid Data Miner Tricks: Overfitting the S&P 500. *The Journal of Investing*, 16(1),15-22 Doi: 10.3905/joi.2007.681820
- Leonard, T. (1992). Databases in the newsroom: Computer-assisted reporting. *Online*, 16 (3), 62-65
- Lippman, W. (1920). *Liberty and the news*. Nueva York: Harcourt, Brace and Howe. goo.gl/k4Vq3r
- McChesney, R, Pickard, V (2011). *Will the Last Reporter Please Turn Out the Lights: The Collapse of Journalism and What Can Be Done to Fix It*. New York: The New York Press.
- Manovich, L. (2011). Trending: The promises and the challenges of big social data. *Debates in the digital humanities*, 2, 460-475. goo.gl/piKvnd
- Martinisi, A. (2013). Data journalism and its role in open government. *Selected papers of the Vth international scientific conference e-governance*, 58-68
- Meyer, P. (1973). *Precision Journalism: A Reporter's Introduction to Social Science Methods*. Indianapolis: Indiana University Press
- Meyer, P. (2011). "Riot theory is relative" *The Guardian*. (Recuperado el 6 de junio de 2018)en <http://www.theguardian.com/commentisfree/2011/dec/09/riot-theory-relative-detroit-england>
- Parasie, S. y Dagiral, E. (2013). Data-driven journalism and the public good: computer-assisted reporters and programmer-journalists in Chicago. *New media & society*, 15 (6), 853-871. Doi: 10.1177/1461444812463345
- Pew Research Center (2016). *State of the News Media 2016 United States: Journalism*. org goo.gl/G38uhp
- Schudson, M. (2010). Political Observatories, Databases& News in the Emerging Ecology of Public Information. *Daedalus*, 139 (2), 100-109. Doi: 10.1162/daed.2010.139.2.100

Solop, F, Wonders, A (2016). Data Journalism Versus Traditional Journalism in Election Reporting: An Analysis of Competing Narratives in the 2012 Presidential Election. *Electronic News*, 10 (4), 203-223. Doi: 10.1177/1931243116656717

Taylor, S.J. y Bogdan. R (1992): *Introducción a los métodos cualitativos de investigación*. Paidós :Madrid

Wormer, H. (2007). Figures, statistics and the journalist: an affair between love and fear. Some perspectives of statistical consulting in journalism. *Springer*, 91 (4) 391-397. Doi: 10.1007/s10182-007-0041-2

Notas

[1] Global Investigative Journalism Network es una red internacional de medios sin ánimo de lucro dedicados a la investigación.

[2] El índice Alexa clasifica el rendimiento y la popularidad de una página web en relación con el millón de páginas más popular de la web.