

Historia y bases del periodismo de datos

[\[1.a\] La historia del periodismo de datos y su desarrollo reciente](#)

[\[1.b\] Las nuevas posibilidades de la tecnología para el uso de datos](#)

[1.a] *La historia del periodismo de datos y su desarrollo reciente*

Los datos siempre existieron y el periodismo siempre los usó, pero la facilidad con la que actualmente se pueden conseguir y compartir, las innumerables posibilidades de procesamiento que ofrecen las nuevas tecnologías y la ayuda que pueden brindar cientos de ciudadanos comprometidos, cambió radicalmente la producción de información y la forma de contar la realidad.

En términos teóricos, hemos asistido a un cambio de paradigma cuyo quiebre puede encontrarse en los inicios del siglo XIX, y su máximo desarrollo en nuestros días, intensificado y facilitado por la aparición de nuevas tecnologías. Un conjunto de técnicas que modificó el modo de ejercer una gran cantidad de profesiones, entre ellas, y quizás sustancialmente, el periodismo. Desde un aspecto práctico, esto ha modificado completamente las condiciones de posibilidad del ejercicio de la práctica periodística, signando la aparición del así llamado “periodismo de datos”. ¿En qué consiste esta rama del periodismo? ¿Es una rama específica del periodismo, un complemento, o una parte sustancial de la profesión? Intentaremos responder estos interrogantes a lo largo de esta clase, para lo cual será necesario hacer un breve repaso diacrónico.

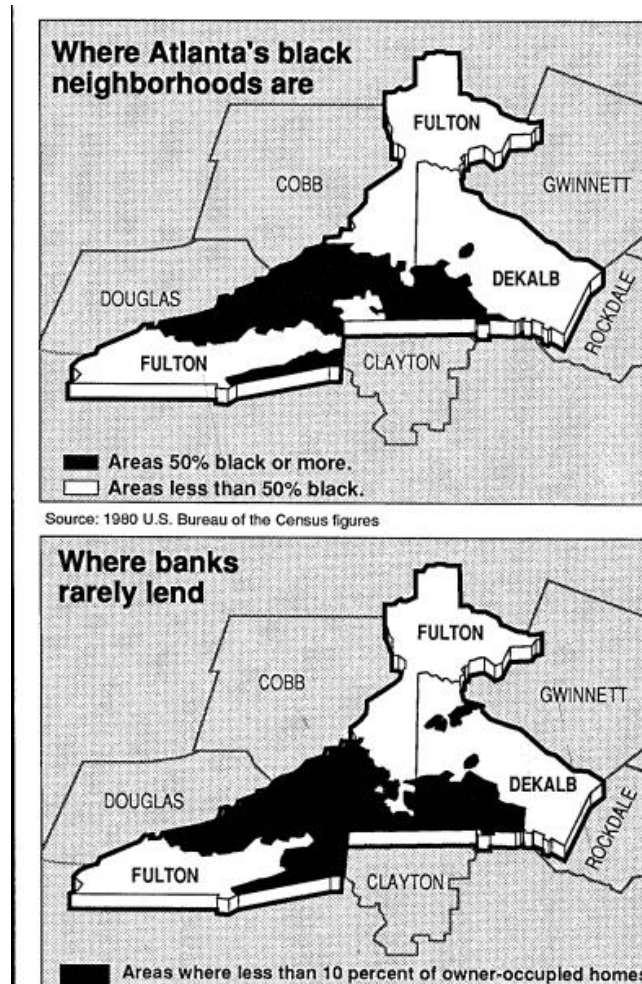
Podemos remontar los orígenes del periodismo de datos a la primera edición de The Guardian,

en 1821, cuando se publicó un tabla con la cantidad de alumnos que iban a escuelas públicas en Manchester. La tabla mostraba que los datos oficiales, de 8 mil niños, estaban muy lejos de los 25 mil que realmente estaban asistiendo y que necesitaban educación gratuita.

DAY SCHOOLS.—Establishments	Boys	Girls	Total	Ann. Exp.	Remarks.
Grammar School	155	155	1800	} Taught, clothed and boarded. } } Taught and clothed. } } And offertory money: do. do. } } (Suppose)—Taught and clothed. } } Funds arising from Sacramental } } Offerings. } } (Suppose)—Expences raised by vo- } } luntary Subscription. } } Taught, clothed and boarded. by } } voluntary Subscription. } } (Suppose)—Taught and partly } } clothed. This School is supported } } by the benevolence of a single } } individual. } } Voluntary Subscription, and Col- } } lections at Churches. }
Blue Coat ditto.....	80	80	2000	
Green Coat ditto.....	50	50	200	
Collegiate Church ditto.....	50	50	40	
Strangeways ditto.....	10	10	100	
St. Mary's ditto.....	12	12	24	40	
St. John's ditto.....	9	9	} 40	
St. Paul's ditto.....	20	20		
Ladies' Jubilee	30	30	250	
Back King-street.....	21	21	40	
NATIONAL SCHOOLS, Granby-row.....	194	119	313	} 600	
Bolton-street, Salford	300	170	470		
	851	381	1232	£5110	
<i>Dissenters.</i>					
LANCASTRIAN SCHOOL, Marshall-st.	692	225	917	400	Voluntary Subscription.
UNITARIAN, Mosley-street.....	35	35	50	Ditto ditto
CATHOLIC.....	198	121	319	104	Ditto ditto
SUNDAY SCHOOLS.					
<i>Establishment.</i>					
Collegiate Church, Shude Hill.	291	205	496	} £1023	
St. Ann's, Back King-street.....	50	56	106		
St. Mary's, Back South Parade	130	110	240		
St. Paul's, Green-street.	170	183	353		
Turner-street.....	68	71	139		
Jersey-street.....	314	281	595		
St. George's, St. George's.....	141	112	253		
St. John's, St. John's-street.....	118	163	281		
St. James's, St. James's-street.....	162	198	360		
St. Michael's, Miller-street.....	234	352	586		
St. Peter's, Jackson's-row.....	..	120	120		
Alport Town.....	90	90		
St. Clement's and St. Luke's, Bennet-street	335	1071	1406		
St. Stephen's, Blooms-street.....	181	297	478		
Oldfield-road	139	204	343		
Trinity, King's Head Yard.....	220	300	520		
Hulme, Duke-street.....	185	189	374		
All-Saints, Oxford-road	196	191	387	30	
Arwick.....	60	110	170	25	
} This is, perhaps, the largest School in the Kingdom. It cost about £2,200, of which £512 0 10½ was contributed in small sums by the Teachers and Scholars.					

Más cerca en el tiempo y de las técnicas conocidas, en 1989, el periodista estadounidense Bill Dedman se ganó el Premio Pulitzer gracias a una investigación hecha con una base de datos, cuando el método todavía se llamaba "Computer Assisted Research" (Investigación asistida por

una computadora). Demostró, gracias a un análisis de bases de datos, que existía discriminación en la forma en la que se otorgaban los créditos hipotecarios en los Estados Unidos: un blanco con bajos ingresos tenía más posibilidad de obtenerlo que un negro con ingresos medios.



Su desarrollo más reciente y en la forma actual fue gracias a los avances de la tecnología y empezaron en paralelo tanto en medios tradicionales (The Guardian en el Reino Unido), como en medios nativos de Internet (Propublica en los Estados Unidos). A partir de entonces se ha desarrollado en distintos formatos y tipos de medios. En América Latina y en el país existen interesantes ejemplos de este de periodismo como La Nación Data de la Argentina o La Nación

de Costa Rica. ¿Cómo se llegó hasta este punto?

El periodismo de precisión, precursor del periodismo de datos

El periodismo de datos como hoy lo conocemos no surgió de un momento a otro. Fue más bien el resultado de la evolución de distintas técnicas que se fueron desarrollando a lo largo del siglo XX, y cuya cristalización es lo que hoy identificamos como “periodismo de datos”. Si bien es difícil encontrar la marca de origen de manera clara, el “periodismo de precisión” aparece como el antecedente más inmediato.

El periodismo de precisión tiene en sus raíces, a diferencia del periodismo tradicional, la construcción de historias a partir de tres pilares: la informática, la estadística y las ciencias sociales. La primera nutre al periodismo de precisión de una regimentación de la información en una estructura de almacenamiento. La estadística, por su parte, constituye una materia prima que brinda datos y diversos tipos de análisis a su ejercicio. Y las ciencias sociales tales como la sociología, la antropología, la economía y la ciencia política contribuyen, dentro del amplio espectro que abarcan, a nutrir a este tipo de periodismo de herramientas de contextualización para la construcción de historias.

Los orígenes del “periodismo de precisión” pueden encontrarse en los Estados Unidos, a partir de la investigación del periodista Philip Meyer, del *Miami Herald*, en la cual reveló por qué los seguros escolares tenían un costo tan elevado para los norteamericanos. La razón por la cual se lo considera el padre del periodismo de precisión a Meyer es que cruzó una serie de datos que nadie había tenido en mente, para llegar a la conclusión de que las compañías de seguros financiaban la campaña presidencial de los candidatos estadounidenses.

En su obra “Precision Journalism and the 1988 US Elections”, Meyer define a este tipo de práctica periodística: “Es la aplicación de métodos científicos de investigación social y comportamental a la práctica del periodismo”¹. Vemos, en el cuño de la definición misma, que

¹ MEYER, P., *Precision Journalism and the 1988 US Elections*, en *International Journal of Public Opinion Research*, 1989, volumen 1, número 3, Pág. 196.

la metodología científica forma parte del núcleo de esta práctica, podemos decir, en dos sentidos: por un lado, porque se nutre de lo que de la ciencia resulta, y por el otro, porque imita o toma como modelo el tipo de construcción científica. La idea de que cualquier persona, aplicando los mismos criterios, pasos y métodos que utilizó el periodista, llegará a los mismos resultados no es más que una réplica de cómo opera el método científico.

Fundamentado desde la modernidad por el método cartesiano, la idea es que las conclusiones que de la ciencia se desprenden son universales: valen para todos los casos previos, los actuales y los futuros. El periodismo de precisión imita, como podemos ver desde su definición, esta idea de universalidad de la ciencia: si todos tomamos los datos que se utilizaron para construir una historia y hacemos el camino (siempre explícito) que hizo el periodista, entonces llegaremos a las mismas conclusiones.

Ahora bien, el contexto histórico es fundamental para entender no sólo a qué se dedicaba el periodismo de precisión, sino también a qué se contraponía. La década de 60 fue sin dudas, a nivel literario, la década del así llamado “Nuevo periodismo”. Encarnado en figuras como Tom Wolfe, Truman Capote y Gay Talese, esta corriente periodística supo combinar de manera perfecta literatura y periodismo: en contra de la clásica pirámide invertida que rezaba que el cómo, cuándo, dónde, qué y por qué debían estar a la cabeza de cualquier historia, estos autores quebraron las reglas para demostrar que había otras formas de contar historia sin salirse del género.

Sin embargo, como una reacción a este movimiento, surgió el periodismo de precisión. Como explica María Ángeles Chaparro,

El primero que incorporó por primera vez el concepto de “periodismo de precisión” a la universidad fue Everette Dennis, quien en 1971 utilizó este término en un curso de la Universidad de Oregon para referirse a algunos reportajes de difícil clasificación donde se aplicaba el método científico, lo que se contraponía al enfoque literario del Nuevo Periodismo que estaba tan en boga en aquella época (CHAPARRO, M.A., *La evolución del periodismo de precisión: el blog de The Guardian sobre el periodismo de datos*, Universidad Internacional de La Rioja, 2013, en Actas del I Congreso Internacional Comunicación y Sociedad, Pág. 3)

Esto llevó a movimientos como el “Periodismo asistido por computadora”, que cuenta con

ciertas investigaciones que dan cuenta del uso de esta técnica. Se trata de “cualquier proceso que utiliza la ayuda de las computadoras durante la recolección de insumos para noticias”². Las hojas de cálculo se volvieron el centro de atención de estos investigadores. Por ejemplo, en 1969, Clarence Jones, del *Miami Herald*, hizo un trabajo buscando patrones sobre justicia penal. En 1972, David Burnham del *New York Times* expuso las diferencias entre las tasas de delitos denunciados por la policía. Elliot Jaspin (*The providence Journal*), utilizó en 1986 distintas bases de datos para denunciar malos historiales de conducción y antecedentes penales en el transporte escolar. Y Bill Dedman, como mencionamos antes, de *Atlanta Journal-Constitution* ganó el premio Pulitzer en 1988 por “El color del dinero”, donde estudió la cuestión de la discriminación racial en la otorgación de préstamos hipotecarios³.

El caso de Dedman tiene sus raíces, a su vez, en un sistema llamado *Computer Assisted Research*. Esta metodología irrumpió como un tipo de investigación legal que utilizaba documentación del ámbito legal (tales como sentencias, jurisprudencia y demás) en el formato de bases de datos electrónicas. Esto supuso un gran avance en relación con la organización y jerarquización de la información, ya que la metodología de búsqueda permitía organizar los tópicos, proveer información relacionada y abría a un trabajo exhaustivo que al cual el periodismo no estaba del todo acostumbrado.

Como explica Giannina Segnini, “los beneficios del Periodismo Asistido por Computadora irradian particularmente el ejercicio del periodismo investigativo porque multiplican las posibilidades para generar ideas para proyectos de investigación, minimizan la dependencia en las fuentes tradicionales y permiten al reportero obtener y entrevistar universos de datos para descubrir hechos ocultos”⁴. Es esto justamente lo que se abrió con este tipo de herramientas como *Computer Assisted Research*.

² GARRISON, B. (1998). *Computer-assisted reporting*, (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

³ CRUCIANELI, S. Introducción al periodismo de datos.

<https://knightcenter.utexas.edu/mooc/datos/semana1/periodismodedatosintroduccion.pdf>

⁴ SEGNINI, G., *Introducción al Periodismo Investigativo y al Periodismo Asistido por Computadora*, en Agencia de los Estados Unidos para el Desarrollo Internacional (USAID), Guatemala, 2008, Pág. 46.

Lo que se puede apreciar es que, en los casos que someramente citamos, el periodista trabaja por su cuenta con datos que recolecta para confirmar, refutar o contrastar sus hipótesis (basta pensar en casos de gran trascendencia nacional como el *Watergate*, paradigma de este tipo de trabajo).

¿Dónde estuvo, entonces, el cambio radical dentro de la profesión? En sustituir las fuentes orales, de las cuales el Nuevo Periodismo no sólo se servía sino hacía de ellas su fundamento, por fuentes precisas, escritas y cuantificables. En este sentido, el periodismo de precisión les exigía por esa época a los periodistas que manejasen técnicas que, en un principio, no son sustantivas del periodismo, sino de otras disciplinas asociadas a las ciencias exactas, tales como la estadística.

Ahora bien, ¿cómo llegamos del periodismo de precisión al periodismo de datos? ¿Cuál es el vínculo entre uno y otro concepto? Como explica Segnini, no se trata más que de una evolución de un término al otro:

Según Flores y Salinas, este nuevo tipo de periodismo “consiste en usar herramientas estadísticas y de visualización para contar mejor las viejas historias y descubrir nuevas historias que contar. Es la nueva veta del periodismo de investigación” (, 2012: 7). Mar Cabra, una de las mejores periodistas españolas en esta especialidad, define este concepto en pocas palabras: “Se trata de recabar y analizar grandes cantidades de información y datos detallados para después hacerlos comprensibles a la audiencia a través de artículos, visualizaciones o aplicaciones” (, 2012) (SEGNINI, G., *Ibid.*)

La especificidad del periodismo de datos, a través de su evolución desde el periodismo de precisión, radica entonces en ser una rama del periodismo de investigación que se nutre de fuentes directas que surgen a partir de volúmenes de información significativos.

Quien mejor lo define sea, quizás, Paul Bradshaw:

¿Qué hace que el periodismo de datos sea diferente del resto? Quizás son estas nuevas posibilidades que aparecen cuando se combina el tradicional “olfato” para las noticias y la habilidad para contar una buena historia con la magnitud y el rango que abre la información digital disponible en la actualidad. [...] El periodismo de datos puede ayudar a los periodistas a contar una historia compleja a través de infografías interactivas” (GRAY, J., BOUNEGRU, y CHAMBERS, L. (2012): *Data Journalism Handbook. How journalists can use data to improve the news.* Londres: OpenKnowledge Foundation. Disponible en: <http://www.datajournalismhandbook.org/1.0/en/index.html> La traducción es nuestra

Con lo visto hasta aquí, entonces, podemos dar con la especificidad del periodismo de datos: en tanto evolución del periodismo de precisión, y en tanto es una rama particular del periodismo en general, se trata de un ejercicio en el cual el periodista extrae sus historias de bases documentales cuyo lenguaje son los datos. Los testimonios, trascendidos, rumores y demás fuentes pueden ayudar las “intuiciones” del periodista que, como tales, no constituyen una parte del trabajo terminado sino una fracción que puede disparar a distintas fuentes documentales, pero nunca sustituirlas. El tratamiento informático de los datos es el núcleo duro del periodismo de datos, sin el cual se cae su propia definición.

Con estas caracterizaciones hechas, podemos darnos cuenta que los avances tecnológicos que ocurrieron durante los últimos años facilitaron la aparición de medios que practican el periodismo de datos, e incluso de periodistas independientes que abrieron sus propias líneas investigativas. En este sentido, la dependencia tecnológica que el periodismo de datos tiene con estos avances mencionados hace necesario detenernos en estas nuevas posibilidades, que abrieron a un verdadero giro copernicano dentro de la profesión.

[1.b] Las nuevas posibilidades de la tecnología para el uso de datos

Como mencionamos anteriormente, los desarrollos previos al periodismo de datos ya contenían en potencia lo que, años después, explotaría con las novedades que el uso masivo de Internet trajo consigo. Con el paso de los años, el desarrollo del periodismo cambió tanto que sus consecuencias, si bien pueden verse en la actualidad, son aún de dimensiones desconocidas.

“El periodismo está siendo sometido a una reevaluación fundamental del núcleo de sus principios e ideas. Las nociones de quienes hablan como periodistas, y las mismas bases sobre las cuales el periodismo ha operado como un negocio viable están siendo reevaluadas, a medida que la institución del periodismo se encuentra con la tecnología digital”, explican Lucy

Morieson y Nikki Usher⁵. El punto es que, según estos autores, el uso masivo de la web trajo consigo innumerables cambios, y que estos afectaron sensiblemente el modo de ejercer la profesión.

Estos cambios, si bien bastante heterogéneos como para hacer una teoría acabada, pueden dividirse en dos grandes grupos. Por un lado, por el lado de la audiencia. Rosen habla de “las personas antes conocidas como audiencia”⁶, en lugar de “audiencia” o “lectores”. El autor pone el foco de la investigación en cómo cambió la relación entre el periodista y el consumidor de noticias, a través del fenómeno del *blogging*, las herramientas de *social media*, y las distintas formas del así llamado “periodismo ciudadano”. Todos estos cambios fueron posibilitados por los avances en el terreno digital.

Sin embargo, lo que nos interesa es el segundo aspecto, en relación a la práctica desde el periodista y no el lector. Este segundo grupo de cambios hizo que el periodista adquiriera herramientas que antes le eran esquivas, o bien imposibles de conseguir.

Nos referimos a las herramientas que facilitan las búsquedas de datos, la extracción, el refinamiento, la puesta en contexto, la visualización y la comunicación. Cuando hablamos de búsqueda, apuntamos un poco más allá de los motores clásicos: portales de *Open Government*, bases de datos públicas, ONGs, catálogos académicos, bibliotecas digitales, y fuentes de *papers*. A estas herramientas se le suman las de refinamiento, para su posterior trabajo en detalle: conversoras de formatos, indexadoras, limpiadoras de datos, creación y conversión de tablas, etc.

Todo este tipo de aplicaciones digitales son deudoras de los cambios que Internet, mirado desde el lado del periodista, trajo consigo. Debido a los grandes volúmenes de datos, herramientas técnicas y coeficientes en juego, el periodismo incorpora al programador como parte del equipo de trabajo, en lo que suele conocerse como “periodismo asistido por hackers”.

Lo interesante de los cambios que la tecnología le trajo a la profesión es, podemos decir, su

⁵ MORIESON, L. y USHER, N., *Mapping the Future of news in the Digital World*, en “Nexus: New Intersections in Internet Research”, Peter Lang Publishing, Nueva York, 2011, Pág. 78. La traducción es nuestra.

⁶ ROSEN, J., *The people formerly known as the audience*, Press Think, 2006, en <http://journalism.nyu.edu>

ligazón con el concepto de “democracia” y, en rigor, con una democracia de más alcance y accesible a cualquier ciudadano. El hecho de que chequear información esté al alcance de cualquier ciudadano hace que el sistema democrático abra más aristas, antes quizás obturadas por la falta de acceso a estas herramientas: “Estos servicios de orientación pública [se refiere a las herramientas de las que estamos hablando] pueden ser vistos como las funciones que custodian a la democracia [...], dándole representación a grupos que, de otro modo, no podían ser escuchados”⁷.

Ahora bien, si esto ha abierto el juego a que cualquier ciudadano pueda chequear la información, aparecen nuevos desafíos para la profesión: “A pesar de las expectativas en torno a los potenciales democráticos de la comunicación digital [...], hay una necesidad vital de preservar las funciones de los servicios públicos consagrados en las noticias impresas y de asegurar que esto sea traducido a los nuevos espacios *online*”⁸. Es decir, trasladar el modo en el que la confianza pública anclada en los medios de comunicación tradicionales a los nuevos medios, que se sirven de herramientas que están al alcance de cualquier ciudadano.

Con el desafío planteado, algunos medios han logrado aceptar las nuevas reglas del juego, elevando la calidad del periodismo y fundamentando de modo cada vez más transparente y riguroso su trabajo.

⁷ MORIESON, L. y USHER, N., *Op. Cit.*, Pág. 90. La traducción es nuestra.

⁸ *Ibid.*

Fuentes:

<http://desocultar.files.wordpress.com/2010/04/segnini-introduccion-al-periodismo-investigativo-y-al-periodismo-asistido-por-computadora.pdf>

https://www.academia.edu/3616384/La_evolucion_del_periodismo_de_precision_el_blog_de_The_Guardian_sobre_periodismo_de_datos

<https://knightcenter.utexas.edu/mooc/datos/semana1/periodismodedatosintroduccion.pdf>

https://www.academia.edu/6187227/Mapping_the_Future_of_News_in_the_Digital_Era