

“Cuando la información era escasa, la mayor parte de nuestros esfuerzos estaban dedicados a buscarla y recogerla. Ahora que hay información abundante, el procesamiento es más importante”.

Philip Meyer. Profesor emérito y profesor titular de la cátedra Knight de Periodismo en la Universidad de Carolina del Norte

Periodista de Datos, nace un nuevo reportero

La reportería tradicional a la que nos hemos acostumbrado comienza a cambiar desde el momento en que los periodistas:

- Deciden comenzar a procesar grandes cantidades de datos.
- Consideran los datos una fuente potencialmente poderosa.
- Analizan grandes cantidades de datos y llegan a hallazgos sorprendentes.

Sandra Crucianelli, especialista en periodismo de precisión, dice que el periodismo de datos no es nuevo, solo que comienza a conocerse con ese nombre a partir de hace un par de años, pero “ya se postuló en los 80 que el método científico podía aplicarse a la investigación periodística y 20 años antes Philip Meyer, docente y especialista en temas digitales, había postulado que los métodos de la investigación social podían aplicarse al periodismo; en todos los casos, los periodistas reportamos los métodos, compartimos los documentos procesados y sin procesar, de modo que puedan ser descargados a una PC del usuario (cultura de open data). Luego, el usuario puede recorrer el mismo camino que el lector y llegar a la misma conclusión, refutarla o encontrar otras conclusiones.

Es entonces el periodismo de datos, investigación en profundidad, la cual utiliza las herramientas ofrecidas por la tecnología para hacer un trabajo mucho más completo. En palabras de Gianina Segnini, experta en el tema y Premio a la Excelencia de la Fundación Gabriel García Márquez: “La gente piensa que es complicado y que hay que ser ingeniero, pero no. Se trata de adoptar las herramientas para generar ideas y citas, buscar informaciones de referencia, hacer reportería, analizar datos y presentarlos”.

El periodismo de datos es el periodismo de siempre que incluye:

1. Métodos del periodismo en profundidad.
2. Precisión.
3. Análisis de los datos.

¿Qué es lo que hace que el Periodismo de Datos sea diferente?

Que combina el tradicional “olfato para las noticias” y la capacidad de narrar una historia convincente, con la escala y alcance de la información digital disponible.

El Periodismo de Datos acerca la información incomprensible que producen los gobiernos y las empresas que manejan nuestros intereses y recursos a los ciudadanos, para despertar en ellos el interés y así tengan los suficientes elementos de juicio para tomar las propias decisiones y exigir cumplimiento por lo que intentan ocultar.

El Periodismo de Datos plantea una nueva forma de narrar las historias. Antes las formas discursivas tradicionales eran las reinas: explicación, descripción, narración y argumentación. Ahora se complementan con las herramientas que proporcionan los medios digitales para analizar en profundidad y corroborar o comprobarlo casi todo. Esa interacción ofrece territorios ilimitados en una nueva narración digital que permite alimentarse a diario y construirse de la mano de la ciudadanía, que es asume un papel activo en este nuevo escenario periodístico.

¿Con qué trabaja el PD?

Con un volumen enorme de datos abiertos, muchos de los cuales están en formatos cerrados, por tanto, hay que hacer una tarea previa de apertura de datos.

¿Con qué fin se hace?

- Ayuda al periodista a contar una historia convincente con una carga de datos alta.
- Ayuda al periodista a explicar cómo se relaciona una historia con un individuo, esto es con el ciudadano. Y quizá este es el aspecto más importante de este tipo de periodismo. La investigación periodística como tal ha ido cambiando en la medida en que el ecosistema tecnológico pone a disposición de todos, nuevas herramientas.

Ahora bien, el criterio de usabilidad de las mismas es válido en la medida en que se explican los fenómenos con mayor coherencia y credibilidad. En ese sentido, el periodismo de investigación es uno, que se nutre de nuevas herramientas tecnológicas para entender y explicar el universo de datos que contienen historias hasta este momento difícilmente digeribles e inalcanzables para el ciudadano, e incluso hasta hace poco para el periodista.

¿Qué se requiere para hacer Periodismo de Datos?

Lo primero y más importante es armar un equipo apasionado. Un grupo de personas que de manera constante se estén formulando preguntas a partir de lo que no comprende la ciudadanía y frente a lo cual esta requiere respuestas. Esas, se las debe dar el periodista. Una primera y buena forma de comenzar, es queriendo entender cómo funciona el Estado. Y eso se puede estudiándolo en la Universidad o en la vida práctica a partir de todos los datos públicos desperdigados que hay en la red.

Tres aspectos a tener en cuenta:

Si en el periodismo de investigación somos curiosos, en el de Datos auscultamos de manera extrema.

Si en el periodismo de investigación consumimos enormes cantidades de documentos en lo que Daniel Santoro llama actividad 'glúteo cerebral' en el de datos todo lo que vamos revisando lo vamos procesando administrándole una categoría de búsqueda que permita hacer usable ese contenido para el trabajo que estemos haciendo y para los venideros.

Si en el periodismo de investigación se debe dudar de todo, en el de datos debe dudar hasta de ti mismo cuando se trata de revisar y autoeditar la historia. No hay que perder de vista que el periodismo de datos es el de precisión, y cualquier error no solo tumba la historia, también su reputación.

¿Qué saberes o formación deben tener quiénes quieren hacer periodismo de datos?

Lo clave de hacer periodismo de datos es entender que es una labor en equipo. Los lobos solitarios no existen en este nuevo escenario ecosistémico en donde conviven profesionales de diferentes saberes, que en carta blanca son:

- Periodistas.
- Ingenieros de sistemas
- Diseñadores.
- Analistas de datos.

De hecho la riqueza de esta disciplina que se abrió un campo y se ganó un espacio respetable en las redacciones de The Guardian, The New York Times, The Washington Post, La Nación de Costa Rica o La Nación de Buenos Aires, entre otros, es que sus coordinadores creen que la riqueza de un buen equipo de periodismo de datos está en la suma de las disciplinas, que al final se complementan.

“La clave reside en que cada uno de los miembros entienda cuáles son las habilidades del resto de las personas que integran el equipo como para poder trabajar los datos de la mejor manera posible y saber pedir las cosas. En este sentido, es clave poder pensar desde los inicios de un proyecto las posibles historias que se quieren acercar a los lectores y cómo contarlas”, precisa Romina Colma, responsable de la producción de datos de la Unidad de La Nación, Argentina.

Y por qué es necesario es interdisciplinariedad, justamente porque casi siempre los datos no son comprensibles desde el primer momento en que se abordan. Sus lenguajes técnicos implican análisis diversos. De hecho por eso existiendo un universo tan variado de información en la red, mucho la desconocen, incluso la ciudadanía, lo cual la desalienta para acercarse. Es entonces cuando el periodismo de datos cumple su papel: analizar, traducir, transformar, crear una mejor forma y suministrar la información de manera pública y libre. En ese sentido lo que hace el periodista de datos es derribar esa barrera que existe entre los gobiernos y las empresas con la comunidad. Zanjear la incomprensión y canalizarla hacia la claridad, la transparencia y la verdad.

¿Con qué recursos trabajan los periodistas de datos?

El listado es grande y con tendencia a crecer, dado que la tecnología avanza en medio del ensayo error que conlleva a la prueba y el éxito. Por tanto, de seguro que lo mencionados ahora en cuestión de semana ya sean revaluados por herramientas mucho más óptimas y en criterios de usabilidad, más comprensibles. Algunas:

- **Excel:** La hoja de cálculo que permite trabajar con tablas de datos, gráficos, bases de datos, macros, y otras aplicaciones avanzadas. Básicamente sirve para el almacenamiento de datos; los cálculos completos, como presupuestos; precisar matemáticamente comparaciones, llevar estadísticas y organizar encuestas...
- **Google spreadsheets:** Es un servicio vía web de hojas de cálculo, parecido al Excel, solo que no ofrece la funcionalidad de realizar gráficos a partir de los datos de las tablas.
- *Google Refine:* Es una herramienta digital de algoritmos que sirve para limpiar datos.
- **Screen scraping:** Para obtener datos de sitios de la red. Se trata de extraer contenido estructurado de una página normal de la red con la ayuda de un programa de recuperación de información.
- **Tableau:** Esta herramienta se basa en analizar datos, filtrarlos, ordenarlos, realizar cálculos en ellos, reorganizarlos, resumirlos.
 - Tableau desktop es una aplicación de visualización de datos que le permite analizar prácticamente cualquier tipo de dato estructurado y producir

atractivos gráficos, paneles de control e informes altamente interactivos en solo minutos.

- Tableau Server es una solución de inteligencia comercial que ofrece análisis visual desde el navegador
- Tableau Reader es una aplicación de visualización gratis que le permite a cualquier persona leer e interactuar con libros de trabajo empaquetados creados con Tableau Desktop.
- **Open Refine:** Es una herramienta de poder para trabajar con datos desordenados, limpiarlos, transformándolos de un formato a otro. Tiene vinculación con bases de datos como Freebase.
- **Document Cloud:** Una herramienta de código abierto que alberga miles de documentos de fuentes primarias de las salas de redacción. Se usa no para guardar solamente, sino para compartir con los cientos de lectores, siendo fiel a la premisa de datos abiertos:
- **CartoDB:** Una plataforma que funciona como una base de datos y herramienta de visualización basada en la nube que permite a los usuarios analizar, visualizar y construir aplicaciones con datos.
- *Talend:* Para extraer, transferir y cargar datos.
- **API de visualización de Google:** Para visualizar datos.
- *Fusion Tables:* Herramienta de mapeo de Google.
- **I2:** Software basado en inteligencia (*Analyst's Notebook* y *iBase*) para analizar datos.
- *'R':* Herramienta para analizar conjuntos grandes de datos.
- **Mindmister:** para hacer mapas
- **Wolfram Alpha:** Herramienta que permite hacer comparaciones y llegar a valiosos hallazgos.
- **Waybach machine:** Herramienta que sirve para hacer arqueología en la WEB y hallar la pista de cualquier documento, desde su origen.
- **Spokeo.com:** Herramienta para buscar personas en la WEB.
- **Tin eye:** Herramienta para buscar fotografías.
- **Ping it:** Para buscar lo que se registra en la WEB con énfasis en redes sociales y popularidad.

Lo que puede hacer el Periodismo de Datos

- **Declaraciones Juradas Abiertas. La Nación / Buenos Aires**
<http://interactivos.lanacion.com.ar/declaraciones-juradas/#pd=0>
El periódico puso a disposición de las personas las declaraciones juradas de los principales funcionarios de los 3 poderes. Este trabajo implicó un esfuerzo conjunto con 3 ONGs (Poder Ciudadano, Fundación Directorio Legislativo y la Asociación Civil por la Igualdad y la Justicia) y más de 30 voluntarios. Terminaron publicando más de 1500 declaraciones juradas. Esos son 1500 documentos públicos que las personas ya no tendrán que pedir a través de un pedido formal, debido a que hubo un gran equipo que hizo disponible la información junto con sus documentos originales.
- **No causar daño**
<http://www.lasvegassun.com/hospital-care/>
El Sun analizó más de 2.900.000 de registros de aranceles hospitalarios, que revelaron más de 3600 lesiones, infecciones y errores quirúrgicos evitables. Obtuvieron datos a través de un pedido de acceso a archivos públicos e identificaron más de 300 casos en que los pacientes murieron por errores que pudieron haberse prevenido.
- **SALARIOS DE EMPLEADOS DEL ESTADO**
<http://www.texastribune.org/library/data/government-employee-salaries/>
Este proyecto reúne en una base de datos la información de los salarios de 660.000 empleados estatales para que los usuarios busquen y ayuden a generar historias.
- **MISTERIO DE ASESINATOS POR RESOLVER**
<http://projects.scrippsnews.com/magazine/murder-mysteries/>
A partir de datos oficiales y pedidos de acceso a registros públicos, se creó una base de datos, que incluye el detalle demográfico de más de 185.000 asesinatos no resueltos, y luego diseñó un algoritmo para buscar patrones que sugieran la posible presencia de asesinos seriales.
- **Cobertura de los disturbios en el Reino Unido por el Datablog de The Guardian**
<http://www.theguardian.com/uk/series/reading-the-riots>
El diario usó periodismo de datos a gran escala para permitir al público comprender mejor quién saqueaba y por qué, cuando en el verano de 2011, hubo una oleada de disturbios en el Reino Unido y se culpó a los medios sociales por causar los disturbios, sugiriendo que había habido incitación desde estas plataformas y que los disturbios fueron organizados a través de Facebook, Twitter y Blackberry Messenger (BBM)

¿A dónde nos conduce esta especialización del Periodismo de Datos?

Quizá la más importante conclusión tenga que ver con la siguiente reflexión: Los periodistas son subjetivos por naturaleza, aun cuando en la escuela de periodismo enseñen que hay que ser objetivos por encima de cualquier cosa. Desde el mismo momento en que nuestro instinto nos indica por dónde meternos y a quien entrevistar, estamos dejando salir nuestras ideas, gustos, preferencias y muchas veces prejuicios. De allí que resulte más válido y responsable apuntarle a ser subjetivamente honestos. Pues bien, el Periodismo de Datos pareciera acercarnos a esa esquiva mirada neutra, porque son los datos los que al cruzarlos y analizarlos nos conducen a verdades incontrovertibles. Esta especialidad no va con las filtraciones en el sentido literal de que la pieza tiene su origen y fin en lo que otra fuente quiera decirnos y/o revelarnos. Ahora serán los datos los que a partir de su despiece y filtraciones, hablen.

El Periodismo de Datos es una herramienta poderosa que nos conduce a un universo de verdades que se unifican a partir de las piezas fragmentadas que son los datos, los cuales van más allá de una simple declaración y por supuesto mucho más lejos que la denuncia, cuando se surten los pasos del refinamiento de los mismos.

“El periodismo basado en datos aumenta la credibilidad y la independencia. Por primera vez en la historia, el periodismo de investigación cuenta con herramientas poderosas para descubrir historias completas, no sólo las piezas que alguna fuente quiere revelar”.
Gianina Segnini, experta en Periodismo de Datos.

Bibliografía:

Entrevista a Gianina Segnini, de la Red Global de Editores.

<http://consejoderedaccion.org/component/content/article/48-rotador-3/98-giannina-segnini-periodismo-de-datos-aumenta-credibilidad>

Entrevista a Romina Colman, Responsable de Datos de La Nación, Argentina. CdR / 2014