

LOS PRIMEROS PASOS DEL PERIODISMO DE BASES DE DATOS EN COLOMBIA.  
ESTUDIOS DE CASO DE EL ESPECTADOR Y LA SILLA VACÍA

MIRIAM FORERO ARIZA

FUNDACIÓN UNIVERSITARIA INPAHU  
FACULTAD DE COMUNICACIÓN, INFORMACIÓN Y LENGUAJE  
PROGRAMA UNIVERSITARIO EN COMUNICACIÓN SOCIAL  
BOGOTÁ  
2013

LOS PRIMEROS PASOS DEL PERIODISMO DE BASES DE DATOS EN COLOMBIA.  
ESTUDIOS DE CASO DE EL ESPECTADOR Y LA SILLA VACÍA

MIRIAM FORERO ARIZA

Trabajo de grado profesional

Directora: Ana Mercedes Suárez Ospina, Comunicadora Social – Periodista, Especialista en  
Periodismo Digital, Cultural e Internacional, Maestría en Educación

Asesora: Consuelo Montaña Arenas, Coordinadora de Investigación, Facultad de  
Comunicación Información y Lenguaje

FUNDACIÓN UNIVERSITARIA INPAHU  
FACULTAD DE COMUNICACIÓN, INFORMACIÓN Y LENGUAJE  
PROGRAMA UNIVERSITARIO EN COMUNICACIÓN SOCIAL  
BOGOTÁ  
2013

A mi amado esposo y  
a Gabrielita que viene en camino

## AGRADECIMIENTOS

Muchos seres han representado aportes importantes para llegar a este paso importante y completar este proyecto, que significa mucho para mí. Agradezco a Dios por llenarme de tantas bendiciones y oportunidades inmerecidas, a mi familia por su apoyo permanente, a mi esposo por ser un soporte para cada paso que doy, a mis profesores por la paciencia con la que compartieron sus conocimientos y guiaron este proceso, a Consuelo Montaña y Ana Mercedes Suárez por creer en mi proyecto desde el principio, a mis compañeros por aportar tanto con sus experiencias e inquietudes, y a los directivos y jefes que han pasado por Consejo de Redacción, donde se gestó y afianzó mi interés por el periodismo de bases de datos.

## TABLA DE CONTENIDO

Resumen .....	8
Abstract .....	8
Introducción .....	9
1. Marco referencial .....	11
1.1 Antecedentes .....	11
1.2 Marco teórico .....	14
1.3 Marco conceptual .....	18
1.3.1 Periodismo de bases de datos (o periodismo de datos) .....	18
1.3.2 Periodismo de investigación .....	24
1.3.3 Derecho de acceso a la información .....	26
1.3.4 Datos abiertos .....	27
1.4 Marco legal .....	29
2. Diseño metodológico .....	31
2.1 Metodología de la investigación .....	31
2.2 Fases de la investigación .....	32
3. ¿Cómo está el periodismo de bases de datos en Colombia? .....	37
3.1 Los medios .....	40
3.2 Los periodistas .....	41
3.3 El Estado .....	44
4. Estudios de caso: El Espectador y La Silla Vacía .....	47
4.1 Doce ejemplos de fiscalización del poder en Colombia .....	47
4.2 Excel, el mejor aliado para un periodista de investigación en El Espectador .....	54
4.3 En la Silla Vacía muchas preguntas se resuelven con bases de datos .....	57
5. Conclusiones .....	61
6. Recomendaciones .....	63
7. Referencias .....	64
Anexos .....	67

## LISTA DE TABLAS

Tabla 1. Medios de comunicación que respondieron la encuesta .....	33
Tabla 2: Condiciones necesarias para el periodismo de bases de datos, según los expertos..	38
Tabla 3: Análisis de contenido - muestra periodística de El Espectador .....	47
Tabla 4: Análisis de contenido - muestra periodística de La Silla Vacía.....	49

## LISTA DE GRÁFICAS

Gráfica 1: Equipos de análisis de bases de datos en los medios encuestados.....	40
Gráfica 2: Periodistas que han recibido capacitación sobre periodismo de bases de datos....	42
Gráfica 3: Uso de la búsqueda avanzada de información y documentos en Internet .....	43
Gráfica 4: Uso de bases de datos obtenidas con derecho de acceso a la información .....	43
Gráfica 5: Uso de bases de datos de consulta en línea.....	43
Gráfica 6: Uso de bases de datos elaboradas por los periodistas .....	43
Gráfica 7: Uso de bases de datos descargadas de Internet.....	44

## RESUMEN

El periodismo de bases de datos es una técnica que aplica los métodos de la investigación social al periodismo, particularmente en la recolección, análisis estadístico, cruces y escrutinio de bases de datos. Dadas las recientes incursiones de diversos medios de comunicación colombianos en dicha tendencia, este trabajo exploró las condiciones en las que periodistas, medios y Estado están aportando para su evolución en el país. A través de una encuesta de diagnóstico, se pudo reconocer el impulso que recibe la técnica, proveniente de periodistas formados en dicha materia, en los principales medios impresos y digitales; mientras que el apoyo de los medios aún es naciente y carece de estructura y definición. Así mismo, los estudios de caso realizados a La Silla Vacía y *El Espectador*, evidenciaron la manera en que sus iniciativas de periodismo de bases de datos fortalecen, tanto sus áreas de investigación como su labor fiscalizadora del poder.

*Palabras clave: periodismo de bases de datos, investigación periodística, medios colombianos, periodismo de precisión*

## ABSTRACT

The database journalism is a technique that applies the methods of social research to journalism, particularly in the collection, statistical analysis, crossing and screening of databases. Given the recent incursions of several Colombian media in this trend, this paper explored the conditions under which journalists, media and Government are contributing to its development in the country. Through a diagnostic survey, it was possible to recognize the impulse given to the technique by journalists who have been trained in that area, in major print and digital media, while the support coming from media is still nascent and lacks structure and definition. Also, the case studies to La Silla Vacía and El Espectador, showed how database journalism initiatives strengthen their research investigative areas and their watchdog work.

*Key words: database journalism, investigative journalism, Colombian media, precision journalism.*

## INTRODUCCIÓN

*“Se está elevando el listón de lo que se requiere para convertirse en periodista. El mundo se ha vuelto tan complicado, el incremento de información disponible tan ingente, que el periodista tiene que ser alguien que criba y no sólo que transmite, un organizador y no sólo un intérprete, así como alguien que reúne y hace accesibles los hechos. En otras palabras, el periodista ahora tiene que ser un administrador de datos acumulados, un procesador de datos y un analista de esos datos”.*

*(Meyer, 1993, p. 25)*

El Knight News Challenge, un programa de becas que apoya iniciativas para aumentar la calidad del periodismo y la innovación mediática, abrió por primera vez en 2012 una convocatoria exclusiva para proyectos que faciliten el acceso y uso de bases de datos; la Universidad Rey Juan Carlos de Madrid y el periódico El Mundo, iniciaron la primera Maestría en Periodismo de Investigación, Datos y Visualización del mundo hispanohablante; y en Colombia, la organización Consejo de Redacción y la Universidad del Rosario trajeron una nómina de expertos internacionales en periodismo de bases de datos a dos de sus principales eventos recientes: el VI Encuentro de Periodismo de Investigación (2013) y la Cátedra Walter Lippman (2012), respectivamente.

Estos sucesos dejan ver la importancia que va adquiriendo a nivel mundial la técnica del periodismo de bases de datos, que consiste en realizar análisis estadísticos, cruces y escrutinio de grandes cantidades de datos públicos, para llegar a hallazgos periodísticos relevantes. Pero más allá del creciente movimiento global, la utilidad que reviste este mecanismo de indagación está en que permite a los periodistas verificar por sí mismos realidades que otros - como instituciones y funcionarios oficiales- presentan, además de encontrar tendencias e irregularidades que no son evidentes a simple vista (Segnini, 2008). Este tipo de periodismo ha permitido destapar escándalos de asignaciones irregulares de subsidios, vigilar el manejo de dineros públicos, identificar tendencias en asuntos sociales como la pobreza y los suicidios y revelar los conflictos de interés de actores políticos, entre otros.

De allí el interés por abordar este asunto en el presente proyecto de grado, pues en Colombia ya se observan algunos trabajos de periodismo de bases de datos en asuntos como la minería, las tendencias electorales y el gasto público en el congreso. Dado que se trata de una realidad naciente en el país y no se conocen investigaciones académicas al respecto, conviene establecer las condiciones actuales en las que se da el periodismo de bases de datos en Colombia, como es el propósito de este trabajo. La indagación hace énfasis en dos medios de comunicación: *El Espectador* y *La Silla Vacía*, sobre los cuales se realizan sendos estudios de caso, ya que ambos publican continuamente historias que involucran el análisis de bases de datos, a partir de un interés claramente definido por promover esa técnica periodística.

Para lograr lo anterior, se plantearon tres objetivos específicos:

- Identificar el grado de reconocimiento y uso del periodismo de bases de datos en los medios de comunicación impresos y digitales en Colombia.
- Tipificar los principales asuntos públicos que *El Espectador* y *La Silla Vacía* abordan desde el periodismo de bases de datos.
- Reconocer las estructuras y métodos de trabajo que aplican esos dos medios de comunicación en sus ejercicios de periodismo de bases de datos.

A partir de una encuesta de diagnóstico y los estudios de caso de los dos medios de comunicación arriba mencionados, la investigación de que da cuenta el presente informe logró analizar las condiciones –en su mayoría favorables, pero apenas iniciales– que se dan en el país para el desarrollo del periodismo de bases de datos, en cuanto a los periodistas, los medios y el Estado. Así mismo, se pudo reconocer la manera en que esa técnica periodística está fortaleciendo la labor investigativa y de fiscalización del poder que realizan *La Silla Vacía* y *El Espectador*; sin dejar de puntualizar en las carencias que aún tiene y el camino que les falta por recorrer.

## 1. MARCO REFERENCIAL

### 1.1 ANTECEDENTES

“Cualquier cosa que se diga del periodismo de datos hoy, no es definitiva. Se trata de una disciplina en evolución” (Crucianelli, 2013). En algunos países se han empezado a producir investigaciones académicas acerca de la naciente tendencia del periodismo de bases de datos:

La tesis de Maestría en Investigación en Documentación ‘*La relación del periodismo de datos con big data y open data*’, desarrollado por Ángeles Mariño (2012) para la Universidad Carlos III de Madrid, busca contextualizar, definir y presentar las herramientas y habilidades que involucra el periodismo de datos, para luego analizar sus vínculos con los conceptos de *big data* y *open data*, que hacen relación a la apertura de datos públicos en aras a facilitar la veeduría ciudadana.

Entre las conclusiones a las que llegó Mariño con su investigación, están el hecho de que el periodismo de datos requiere de diversas habilidades, por lo que se convierte en una práctica ideal para el trabajo en equipo; la aún escasa presencia de esta materia en los programas académicos; y la importancia de que documentalistas, periodistas y demás profesionales vinculados a medios de comunicación reconozcan en el periodismo de datos una oportunidad de reinventar el periodismo tradicional, a través de herramientas que acercan la realidad a una lectura más objetiva y con una visualización más comprensible.

Mariño (2012) encontró, además, que el auge del periodismo de datos se ha dado no sólo por una moda o por los recientes avances tecnológicos, sino que “también ha sido responsable la concientización sobre librar la información para uso y desuso de los usuarios, sean estos periodistas o no” (p. 56). No obstante, descubrió unas preocupantes contradicciones entre la naciente ley de acceso a la información y la ley de protección de datos en España, que pueden bloquear una verdadera accesibilidad a los datos públicos.

Ese hallazgo permite reconocer que es importante investigar qué tanto han podido aprovechar los medios de comunicación esa apertura de datos que facilita la actual movilidad en Internet y el derecho de acceso a la información pública, como lo hace el presente trabajo con respecto al caso colombiano.

Por su parte, la tesis '*Cómo construir informaciones con precisión: guía periodística*', realizada por Patricia J. Blanco Barrios (2000) para optar al título de Licenciada en Comunicación Social de la Universidad Católica Andrés Bello, se enfoca en el periodismo de precisión –que, como se verá más adelante, es una de las raíces del periodismo de datos– y pretende proponer un esquema metodológico para realizar trabajos periodísticos de ese tipo en su país: Venezuela.

Para alcanzar su objetivo, la investigadora, realizó una extensa revisión bibliográfica y consultó las apreciaciones y experiencias de quienes han estudiado o tratado el género del periodismo de precisión. Al igual que la tesis de Mariño y el presente proyecto, el trabajo de Blanco consideró importante examinar el acceso a las fuentes oficiales de información pública, dada sus implicaciones para el ejercicio del periodismo dedicado a analizar datos.

Como resultado de su investigación, Blanco propuso un modelo de periodismo de precisión que considera adecuado para aplicar en su país. Adicionalmente, halló una carencia en los medios de comunicación impresos de Venezuela: no tienen departamentos especializados en esta práctica.

La experiencia de Blanco en Venezuela brinda luces sobre la manera de abordar un diagnóstico sobre el periodismo de bases de datos en un país determinado e identificar las principales prácticas a las que están acudiendo sus periodistas; de este modo, su método de tomar una muestra de trabajos periodísticos que emplearon el periodismo de precisión y analizarlos con base en categorías previamente establecidas, servirá como inspiración para el análisis de contenido que involucra esta investigación.

Para reunir a una comunidad más amplia de personas interesadas por el periodismo de bases de datos e identificar casos llamativos en diversos países, en 2011 se escribió el primer Manual de Periodismo de Datos, como resultado de un taller que realizaron el European Journalism Centre y la Open Knowledge Foundation en el marco del Festival Mozilla de ese año y fue traducido del inglés al español en abril de 2013 por el equipo de Nación Data, del diario *La Nación* de Argentina. El documento, construido de manera colaborativa por cientos de personas, habla de la importancia del periodismo de datos, describe en líneas generales cómo hacerlo y cuáles son sus principales herramientas.

Incluye, además, un relato del ‘detrás de cámaras’ de algunos medios de comunicación que encuentran en el periodismo de bases de datos una de sus principales técnicas de trabajo e investigación: ABC, la BBC, *Chicago Tribune*, *The Guardian*, y Zeit Online. Ese ejercicio servirá de ejemplo para el presente trabajo, pues los aspectos que aborda pueden ser replicados al analizar el caso de los medios colombianos: ¿cómo surgió el uso del periodismo de datos en la redacción?, ¿para qué y por qué lo utilizan?, ¿cómo está conformado el equipo de trabajo?, ¿cómo obtienen los datos?, ¿qué herramientas utilizan?, ¿cuál es el proceso que siguen?, etc.

Actualmente está en curso la elaboración del Manual de Periodismo de Datos Iberoamericano, que busca, además de ofrecer técnicas y herramientas, revisar el estado de la cuestión en países de Iberoamérica y describir los casos más destacados. La amplia conceptualización que realizaron las tesis de España y Venezuela, así como el Manual de Periodismo de Datos, servirán de base para contextualizar el presente trabajo, de modo que haya suficiente claridad sobre la delimitación de los conceptos, herramientas y características a indagar en los medios colombianos.

Para la elaboración de este trabajo se realizó una búsqueda en las principales facultades de comunicación social de Bogotá: Fundación Universitaria Inpahu, Pontificia Universidad Javeriana, Universidad de la Sabana, Universidad Sergio Arboleda, Universidad Externado

de Colombia, Universidad Jorge Tadeo Lozano y Universidad Central. En ninguna de ellas se encontró una tesis o investigación académica relacionada con el periodismo de bases de datos –o con alguno de los otros términos por los que es conocida esta técnica (periodismo asistido por computador, periodismo de precisión, etc.)-, lo cual es comprensible, pues esta es una práctica reciente en el continente y aún poco explorada en Colombia; pero además impulsa aún más el desarrollo del presente trabajo, dada la necesidad de empezar a indagar el campo del periodismo de bases de datos.

## **1.2 MARCO TEÓRICO**

Una de las grandes controversias, aún sin resolver, en el estudio de la comunicación social, es si ésta puede tener el carácter de ciencia o no. Como denota Berganza (2005), el debate se torna más polarizado cuando se analiza el área periodística y su posibilidad de realizar investigación de tipo científico.

Por una parte, los contradictores de esta idea –Spengler (1989), B. Philips (1977) y G. Sartori (1998), ejemplos citados por Berganza (2005)- no tardan en resaltar las principales falencias del periodismo como prueba de su incapacidad de generar conocimiento científico. Para ellos, la influencia de intereses económicos y políticos, el magnetismo de la opinión pública, la poca profundidad y trascendencia de los contenidos, y el conformismo de los periodistas, para quienes la simple familiaridad con las cosas o el mero conocimiento intuitivo son suficientes, no permiten que su actividad tenga un estatus científico, capaz de reflejar una realidad de manera completa y objetiva.

Sin embargo, Meyer (1993), padre del llamado periodismo de precisión, considera que esos errores no se dan por una ausencia de esfuerzo o capacidad, sino porque hace falta aplicar las prácticas de la ciencia, emplear un cuerpo de conocimientos en el cubrimiento periodístico. De este modo, Meyer (1993) aconseja “tratar al periodismo como si fuera una

ciencia, adoptando el métodos científico, la objetividad científica y los ideales científicos en el proceso complejo de la comunicación de masas” (p. 31).

El físico Lawrence Cranberg (1989), citado por Meyer, coincide en que el periodismo comparte con la ciencia la búsqueda de conocimientos y la persecución de hechos ocultos, por lo que puede aplicar sus mismas reglas en la labor investigativa, no sin antes resaltar la necesidad de una cualificación apropiada para que el periodista pueda convertirse en un científico en ejercicio.

Esa adopción de elementos de recolección de datos, verificación exhaustiva y sistematización de procesos investigativos se presenta como una solución frente a las debilidades del periodismo. Resuelve, incluso, problemas que eran defendidos en la concepción tradicional del periodismo como la pasividad –no crear noticias sino esperar a que otros las generen- y la inocencia –emplear una mirada espontánea, sin conocimientos especializados- (Meyer 1993); pues el periodista que acude al método científico es activo, propone, adquiere experticia, profundiza y descubre, así, verdades ocultas relevantes para la sociedad.

Esta tendencia no pretende alcanzar la infalibilidad del periodismo, pues está claro que la misma ciencia reconoce sus limitaciones, admite que siempre quedan hechos ocultos o verdades que no se alcanzan a conocer completamente; lo cual puede ocurrir también con la labor de informar, incluso si esta aplica toda la rigurosidad del método científico (Cranbert, 1989, citado por Meyer, 1993).

El presente trabajo se soporta en la teoría de estos últimos, quienes creen que la actividad periodística – específicamente su rama investigativa, siempre y cuando cuente con las herramientas y la preparación adecuada- puede adquirir un carácter científico. De hecho, basta leer a Wimmer y Dominick (2002) para descubrir que el periodismo de alta calidad, realizado con los más exigentes estándares éticos, comparte con el método científico las principales características que describen los dos escritores, a saber:

- La investigación científica es pública. “El avance científico depende de la información disponible sin reserva” (Wimmer y Dominick, 2002, p. 11). Del mismo modo que estos dos autores recomiendan que los investigadores científicos den a conocer con claridad los métodos y pasos empleados para determinado hallazgo, el periodista de investigación debe indicar con claridad en sus publicaciones las fuentes documentales o personales y la reportería que le permitieron llegar a determinada conclusión. El periodismo de bases de datos va incluso más allá; la tendencia es hacia una transparencia total, por lo que muchos medios de comunicación especializados en dicha técnica incluyen en sus publicaciones artículos enteros para describir cómo se realizó la investigación y secciones en las que dejan a disposición la totalidad de los datos utilizados en su trabajo, como el caso del *Datablog*, del periódico *The Guardian* y la sección Nación Data, del diario argentino *La Nación*.
- La ciencia es objetiva. Debe cumplir unas reglas metodológicas de modo que dos observadores, siguiendo cada uno por su cuenta idéntico procedimiento, lleguen a la misma conclusión. En cuanto al periodismo, aunque existe una discusión casi estéril sobre su capacidad de ser realmente objetivo, ya que el simple hecho de seleccionar los temas y las fuentes a cubrir implica una intromisión del punto de vista del periodista (Herrán y Restrepo, 1991), este rasgo de la objetividad alcanza su mayor expresión en el periodismo investigativo, especialmente con la técnica de análisis de bases de datos. Crucianelli (2012a) resalta el aspecto científico de dicha técnica periodística, pues las publicaciones que la emplean deben explicar claramente el camino recorrido para llegar a determinada conclusión y mostrar los datos originales de donde surgió, de modo que otro periodista pueda llegar a la misma conclusión usando esas mismas herramientas.. Adicionalmente, Wimmer y Dominick recalcan la importancia de que en el método científico se le dé mayor relevancia a los hechos que a la interpretación de los mismos, lo cual coincide con numerosas recomendaciones éticas hacia los periodistas, según las cuales éstos

deben limitarse a los hechos y distinguirlos claramente de lo que constituya comentarios u opinión (Herrán y Restrepo, 1991).

- La ciencia es empírica. Aborda aspectos del mundo susceptibles de ser conocidos y medidos, por lo cual rechaza las explicaciones metafísicas de los fenómenos. Del mismo modo, el periodismo de investigación se ocupa de realidades palpables y experimentables, y basa su actividad en la observación y la confrontación, con lo cual busca que sus conclusiones estén altamente verificadas a través de fuentes confiables, preferiblemente documentales. Sus principales exponentes incluyen la formulación de hipótesis como parte fundamental del proceso investigativo, entendiéndola como una posible explicación al asunto que se indaga, la cual es verosímil y susceptible de verificación; dicha hipótesis sirve como guía que puede ser probada, reformulada o rebatida durante una pesquisa periodística (Santoro, 2004); de hecho, este experimentado periodista argentino indica que una de los indicios para diferenciar una investigación periodística seria y profunda de las meras primicias es que en la primera “el autor fundamenta lo que sostiene en entrevistas, documentos, fuentes o citas bibliográficas” (p. 22). Más aún, como herramienta del periodismo investigativo, el periodismo de bases de datos encierra una alta dosis de empirismo, ya que su material de trabajo son los datos recolectados de manera tangible, precisa y rigurosa.
- La ciencia es sistemática y acumulativa. Los investigadores recurren a estudios previos para construir su propio trabajo con base en ellos y aprovechar así los conocimientos previamente obtenidos e identificar los factores principales. Los periodistas de investigación recomiendan, también, iniciar la pesquisa con un ejercicio de búsqueda de antecedentes, tanto en los archivos de los medios de comunicación, como en los documentos públicos que dan cuenta de las acciones previas al caso que se quiere investigar (Santoro, 2004).

- La ciencia es predictiva. Los científicos buscan desarrollar teorías porque estas, una vez probadas, sirven para predecir sucesos futuros con base en el análisis de los presentes. Aunque el periodismo no busca establecer teorías, sí logra identificar el *modus operandi* de los diversos tipos de corrupción, con lo cual muchos periodistas se apoyan en los hallazgos de sus colegas para replicar investigaciones similares en sus lugares de trabajo. De ahí que hayan surgido muchas redes de periodistas de investigación, quienes comparten sus métodos y se reúnen en congresos de diversa periodicidad para exponer sus trabajos, que se convierten en ‘inspiración’ para los asistentes. Adicionalmente, muchas investigaciones periodísticas surgen cuando los actos de corrupción aún están en proceso, con lo cual se logra alertar a las autoridades y a la sociedad sobre lo que podría pasar a futuro si no se toman medidas frente a los hechos que se denuncian.

### **1.3 MARCO CONCEPTUAL**

Como es lógico, el principal concepto sobre el que se construye el presente proyecto es el de ‘periodismo de bases de datos’. Adicionalmente, se documentarán tres conceptos que están estrechamente relacionados con el primero: periodismo de investigación, derecho de acceso a la información y datos abiertos.

#### **1.3.1 Periodismo de bases de datos (o periodismo de datos)**

Esta práctica se deriva de dos tendencias que aparecieron con anterioridad: el periodismo de precisión y el periodismo asistido por computador.

El periodismo de precisión surgió en los años 60 como respuesta a las tradicionales críticas que se le hacen al periodismo en cuanto a su falta de verificación y de metodología rigurosa. El reportero y académico estadounidense Philip Meyer es considerado precursor de dicha corriente, ya que la bautizó como tal en su libro ‘Precision Journalism,. A Reporter’s Introduction to Social Science Methods’, en 1973; y la define como “la

aplicación de los métodos científicos de investigación social y psicosocial en la actividad periodística” (Meyer, 1993, p. 22). Por su parte, José Luis Dader (1997), discípulo español de Meyer, lo entiende como “la información periodística que, sobre cualquier asunto de trascendencia social, aplica o analiza sistemáticamente métodos empíricos de investigación científica, de carácter numérico o no numérico, con especial inclinación al campo de las ciencias sociales” (p. 22).

Las técnicas más empleadas por el periodismo de precisión son la encuesta y el sondeo de opinión, aunque también recurre al experimento psicosocial, el análisis de contenido y el rastreo de listados alfanuméricos o de bases de datos. Esta última herramienta es la que dio origen en Estados Unidos al término *database journalism*, es decir, periodismo de bases de datos, que inicialmente analizaba listados enormes de manera ‘manual’, pero que luego evolucionó con la digitalización de los archivos impresos en diversas entidades.

Si el periodismo de precisión se vio impulsado por el surgimiento y propagación de los computadores en las salas de redacción, esta importante novedad se convirtió en la cuna de lo que se empezó a conocer como el periodismo asistido por computador (PAC), que se enriquece, además, con la llegada de Internet. Uno de sus padres, Brant Houston (2004), lo entiende como el aprovechamiento de cuatro tipos de herramientas tecnológicas: recursos en línea, *software* de hojas de cálculo o administradores de bases de datos, *software* estadístico y sistemas de información geográfica (GIS por su sigla en inglés).

Un poco más general, Garrison, B. (1998), citado por Segnini (2008), define el PAC como "cualquier proceso que utiliza la ayuda de las computadoras durante la recolección de insumos para noticias" (p. 45). Entre las ventajas que resalta Giannina Segnini del PAC, están “el fácil y rápido acceso a la información a través de Internet, la agilización de procesos matemáticos o estadísticos que solían ser tediosos y ajenos al periodista y la posibilidad de obtener y almacenar voluminosa información en poco espacio” (p. 45).

Al conocer las herramientas y el tipo de trabajos que realiza el periodismo de bases de datos, se entiende por qué las dos corrientes descritas anteriormente son consideradas sus raíces, pues la técnica que estudia el presente proyecto retoma y potencia el análisis –tanto estadístico como mediante cruces- de bases de datos que surgió con el periodismo de precisión, pero además aprovecha las ventajas que ofrecen programas de informática más modernos y sencillos como Excel, los sistemas de mapeo, las aplicaciones gratuitas en Internet para la limpieza y cruce de datos, así como la gran cantidad de datos abiertos que, cada vez más, se encuentran en la web.

Mariño (2012) define el periodismo de datos como “el conjunto de herramientas y prácticas que utilizan datos para la construcción de una historia (o noticia). Y estos datos pueden tanto acompañar una historia como ser fuentes de una investigación” (p. 14). Sin embargo, una parte de esta definición puede llevar a confusiones, ya que alude simplemente al uso de ‘datos’, término que algunos pueden entender de manera general como cualquier pieza de información y no como un conjunto -grande o pequeño- de datos (bases de datos), que es a lo que acude el periodismo de datos y a lo que seguramente se refieren Mariño (posiblemente se trata de una confusión idiomática).

La investigadora del presente proyecto considera que así como “todo periodismo supone investigación [y] periodista que no investiga no es periodista” (Gorriti, 1999), toda actividad periodística debe involucrar el análisis de diversos tipos de datos, los cuales a veces consisten en la información recolectada a través de entrevistas, de revisión documental, de fenómenos observados, entre otros métodos, y no necesariamente provienen de bases de datos. Por lo tanto, para mayor claridad y para aproximarse a un término más preciso, este trabajo emplea la expresión ‘periodismo de bases de datos’ (traducción literal del término inglés *database journalism*) en lugar de ‘periodismo de datos’ (*data journalism*), para referirse a la práctica específica que extrae conclusiones relevantes, con criterio periodístico, a partir del análisis de bases de datos que contienen información veraz, confiable y de interés público.

El Manual de Periodismo de Datos (Aisch y otros, 2012) intenta ampliar esa definición simple de periodismo que se hace con datos, enumerando algunas de sus principales utilidades: encontrar relaciones entre miles de documentos, automatizar el proceso de recolección de información, ayudar a entender historias complejas mediante visualizaciones e infografías interactivas y ampliar las posibilidades de búsqueda de información que tienen las personas al publicar las bases de datos completas en Internet.

Sin embargo, existe una controversia con respecto a los últimos dos puntos pues ambos figuran más como elementos anexos al periodismo de bases de datos que como características constitutivas de éste, en las conceptualizaciones de reconocidos periodistas expertos en periodismo de investigación y periodismo de bases de datos.

Tanto Sandra Crucianelli como Giannina Segnini, principales exponentes latinoamericanas de esta especialidad desde su labor en Argentina y Costa Rica, respectivamente, consideran que la sola visualización de grandes cantidades de datos –por elaborada y llamativa que sea- o la mera publicación de las tablas de datos originales para consulta de los usuarios no representa periodismo de bases de datos.

"[El Análisis] es el paso más importante en periodismo de base de datos, no se trata de publicar datos, es hacer lo que hacemos los periodistas desde tiempos inmemoriales, sacar una conclusión", aseguró Segnini durante la Conferencia Latinoamericana de Periodismo de Investigación 2013 (Colpin), según reseña Ramírez (2012). Crucianelli (2012a), por su parte, presenta el procesamiento y análisis estadístico o matemático de los datos como requisito indispensable para diferenciar el periodismo de bases de datos de otras iniciativas digitales.

Keenan (2010) explica con claridad esta postura, mediante el caso de los cables diplomáticos revelados por *Wikileaks* que, según el autor, habrían tenido un impacto mínimo si sólo se hubieran publicado los documentos sin pasar por el procesamiento y análisis de experimentados periodistas de los medios con los que Julian Assange hizo alianzas en diversos países.

El presente trabajo se apoya en las posturas de Crucianelli, Keenan y Segnini, por lo cual sólo se tendrán en cuenta para el análisis los casos periodísticos en los que se utilice el procesamiento de bases de datos, más allá de las visualizaciones que los acompañen.

A continuación se mencionan algunas de las principales características anexas –pero no obligantes- al periodismo de bases de datos:

- Se fortalece con el derecho de acceso a la información pública y las acciones de datos abiertos (ambos conceptos se exponen más adelante en este capítulo) (Mariño, 2012).
- Presenta a su audiencia los documentos y datos originales con los que trabajó (Crucianelli, 2012b).
- Explica el método que siguió la investigación para que pueda ser sometida a crítica (Crucianelli, 2012b).
- Lo ideal es que no se pierda información, sino que todos los datos recolectados, aunque no todos se utilicen en la historia periodística, sean reunidos en una base de datos general (Segnini, 2011).
- Entre las ventajas que ofrece están: ayuda a los periodistas a encontrar en grandes cantidades de datos las noticias que estos esconden y que otros no ven; se convierte en una herramienta útil para la labor de ‘perro guardián’ que tiene el periodismo (Crucianelli, 2013); elimina la dependencia de las filtraciones o de las fuentes oficiales, que suelen entregar sólo piezas de información según su conveniencia (Segnini, 2011).
- Propicia el trabajo en equipo con personas de otras disciplinas como desarrolladores, estadísticos, analistas, infografistas, etc. (Mariño, 2012).
- Las técnicas de procesamiento de datos más comunes son: análisis estadístico, análisis georreferencial, cruce de bases de datos y algoritmos para análisis automatizado de textos (en etapa experimental).
- Las principales herramientas para obtener datos son: extracción de datos de Internet (conocido como *scarping*), solicitudes mediante derecho de acceso a la

información, construcción de bases de datos propias a partir de diversas fuentes y filtraciones (en muy pocos casos).

- Estas son algunas de las herramientas –entre *software* libre y pago- que emplea: Excel, *Google Refine*, *Google Fusion Tables*, *Talend*, *Many Eyes*.
- Los principales pasos que se pueden seguir, según Segnini en la reseña de Ramírez (2012), son: obtener los datos, limpiarlos, analizarlos, verificarlos (mediante reportería) y crear visualizaciones.

Los siguientes son algunos ejemplos de los mejores trabajos que se hacen actualmente en el mundo, a través del periodismo de bases de datos:

- 156 candidatos a alcaldías tienen morosidad, sanciones o condenas, *La Nación* (Costa Rica): en diciembre de 2010, antes de las elecciones regionales, los periodistas cruzaron la base de datos de todos los candidatos con registros criminales, personas con prohibición para ocupar cargos públicos y deudores morosos de seguridad social. Encontraron cinco candidatos con condenas, 27 inhabilitados y muchos deudores morosos del Estado. Luego de verificar manualmente los resultados y solicitar la versión de los candidatos, publicaron un mapa interactivo con los hallazgos de la investigación.
- Gastos del Senado, *La Nación* (Argentina): a comienzos de 2013 el equipo de investigación construyó una gran base de datos a partir de 30.000 PDFs relacionados con los gastos del Senado. Para poderla procesar tuvieron que descargar todos los documentos, eliminar la protección de los PDFs, convertirlos en archivos reutilizables, analizar los datos por categorías, reorganizar la información y generara diagramas de Gantt para ver líneas del tiempo. Descubrieron, entre otras cosas, que el vicepresidente usó fondos de emergencia para comprar muebles de lujo y cobró viáticos por viajes que no realizó.
- Cracking the codes, Public Integrity (Estados Unidos): esta organización sin ánimo de lucro analizó en 2012 los miles de registros de pagos del sistema de salud estadounidense (Medicare) entre 1999 y 2008 y encontró que miles de profesionales de la salud estaban cobrando servicios superiores a los realmente prestados.

Además, hay medios que tienen una apuesta específica de dedicar tiempo, recursos y un equipo completo dedicado al periodismo de bases de datos: Nación Data, proyecto de *La Nación*, de Argentina; Datablog, de *The Guardian*, en el Reino Unido; y Nerd Blog, de ProPublica, en Estados Unidos, son proyectos en los que cada medio publica las bases de datos a partir de las cuales realiza sus investigaciones, comparte herramientas y artículos de expertos en el tema, describe cómo realizó cada investigación y reúne las visualizaciones interactivas de sus trabajos. Al igual que los anteriores, *La Nación*, de Costa Rica, tiene un equipo conformado por periodistas y desarrolladores para adelantar este tipo de investigaciones, pero aún no tiene un proyecto en la web para narrar lo que realizan.

### **1.3.2 Periodismo de investigación**

Antes de explorar el concepto de periodismo de investigación, es importante ver su relación con el periodismo de bases de datos. En síntesis, este último es una herramienta que enriquece y potencia al primero, ya que sus técnicas “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” (Segnini, 2008, p. 46).

Para definir el periodismo de investigación, Gorriti (1999) empieza por exponer la controversia que suele despertar el término, pues la ética periodística indica que toda labor en ese oficio supone la investigación y por tanto no debería haber una especialidad para investigar en el periodismo. Aclara, entonces que:

El periodismo de investigación es simplemente periodismo que ha tenido (a veces) más tiempo para aplicar técnicas específicas de averiguación respecto a temas o realidades que se resisten a ser revelados. Sus principios son los de toda disciplina de investigación, desde la epidemiología a la paleontología. Pero sus reglas son las del periodismo en general. Sólo se distingue en la práctica de otras

formas de periodismo por la aplicación más frecuente y relativamente especializada de las mencionadas técnicas de averiguación.

Esa posibilidad adicional que tiene el periodismo de investigación para aplicar técnicas avanzadas de indagación, hace que sus productos se diferencien de los trabajos cotidianos: los hallazgos están claramente sustentados en numerosas fuentes personales, documentales o bibliográficas; no se rigen por teorías conspirativas para explicar la realidad, sino que indagan datos verificables; y dan siempre crédito a quien corresponda (otros colegas, la justicia, etc) al utilizar sus descubrimientos (Santoro, 2004).

Hay tres características imprescindibles a las que los tradicionales estudiosos del periodismo de investigación acuden para definirlo:

- Se trata de investigaciones propias del periodista, no de pesquisas adelantadas por un grupo académico, la justicia, los científicos, etc.
- Revela lo que algún poder intenta mantener oculto y, para esto, supera los obstáculos de acceso a la información e, incluso, de seguridad.
- Trata asuntos de interés público sin entrometerse en la vida privada de las personas, salvo en las ocasiones en las que ésta afecte la actuación legal de servidores públicos. (Santoro, 2004).

En cuanto al método, Santoro (2004) enumera los diez puntos que ha definido como su plan de trabajo. Además de mencionarlos, a continuación se presentan los puntos de encuentro con el periodismo de bases de datos:

- 1°. Búsqueda de historias para investigar. A veces las bases de datos constituyen una rica fuente de ideas sobre nuevas historias a abordar.
- 2°. Delimitación de la historia.
- 3°. Formulación de hipótesis.
- 4°. Organización del archivo.
- 5°. Recolección de información. En una investigación que se apoya en el periodismo de bases de datos, este punto corresponde a la obtención de los datos, los cuales

pueden llegar a representar una porción significativa de la información a recolectar.

- 6°. Cruzamiento de datos. Equivale al análisis en el periodismo de bases de datos.
- 7°. Diseño de la estructura de la nota. Puede implicar, también el diseño de una visualización para reflejar claramente los datos procesados.
- 8°. Verificación de la información. También necesaria en el periodismo de bases de datos.
- 9°. Redacción.
- 10°. Edición.

### **1.3.3 Derecho de acceso a la información**

La información es una herramienta fundamental de la democracia y el acceso a aquella ha facilitado la consolidación de los sistemas democráticos en países latinoamericanos, a través de la participación ciudadana en asuntos públicos (Relatoría Especial para la Libertad de Expresión, 2012).

“Los gobiernos democráticos, que tienen sensibilidad democrática, que lo que quieren es cumplir los deberes que el derecho les asigna, ven en el derecho de acceso una herramienta para saber dónde están fallando, cómo hacerlo mejor” (Botero, 2009).

En este sentido, se entiende el acceso a la información como el derecho que tienen los ciudadanos a buscar y obtener los datos que necesiten para tomar decisiones, y conocer sus derechos y sus deberes. La importancia de este derecho y de su conocimiento y uso por parte de los periodistas, radica en el hecho de que se convierte en un mecanismo para vigilar y promover la transparencia de las actuaciones del Estado, exigir la rendición de cuentas, saber cómo se está administrando el dinero y demás bienes públicos, así como defender los derechos fundamentales de las personas y de la sociedad (Fundación para la Libertad de Prensa –FLIP-, 2011).

La Relatoría Especial para la Libertad de Expresión (2012) señala dos principios rectores para este derecho:

- Principio de máxima divulgación: indica que toda la información en manos del Estado debe ser pública y accesible, mientras que los casos en los que determinada información se mantenga en secreto serán la excepción de dicha regla, mediante un régimen legal establecido por cada Estado.
- Principio de buena fe: es necesario que quienes están obligados a ejecutar el derecho de acceso a la información actúen de buena fe; es decir, provean ayuda a los ciudadanos, promuevan la transparencia, aseguren la aplicación del derecho, etc.

#### **1.3.4 Datos abiertos**

Con la evolución de Internet y la difusión del concepto de acceso a la información, ha surgido una tendencia internacional de apertura de datos, que ofrece la posibilidad de que enormes cantidades de información –generalmente estadística- sobre la ejecución de los gobiernos y sobre las principales cifras de cada país, estén disponibles para consulta fácil e inmediata por parte de los ciudadanos.

Incluso el caso *Wikileaks*, aunque se trata de una filtración y no de una obtención legal de la información, puso en evidencia la importancia de que en el mundo se dé una apertura y transparencia de los datos (Mariño, 2012).

Estas son algunas definiciones de la tendencia conocida como *open data* o datos abiertos:

Los datos abiertos son datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, y que se encuentran sujetos, cuando más, al requerimiento de atribución y de compartirse de la misma manera en que aparecen. (Open Knowledge Foundation, 2010).

Información de carácter público que se pone a disposición del ciudadano o usuario, en orden a una mayor transparencia de acción. Se trata simplemente de

abrir al público información que se aloja en los servidores propios o en la nube y que son de autoría de la institución local, nacional o internacional. Puede tratarse de informes, de bases de datos, de encuestas, de estadísticas y demás, en formatos compatibles con los potenciales usuarios y terceros. (Mariño, 2012, p. 52).

Son todos aquellos datos primarios, sin procesar, en formatos estándar e interoperables que facilitan su acceso y permiten su reutilización, los cuales están bajo la custodia de las entidades públicas y que pueden ser obtenidos y ofrecidos sin reserva alguna, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos. (Gobierno de Colombia, Catálogo de Datos Abiertos).

Esta última definición es la más acorde al interés de este proyecto, ya que incluye las características por las que la tendencia de datos abiertos es importante para el periodismo de bases de datos: amplía enormemente la gama de fuentes disponibles para los periodistas y ofrece una verdadera accesibilidad de los datos al presentarlos en formatos que permiten su reutilización, reorganización y análisis.

Sólo un aspecto habría que agregar al concepto de datos abiertos y es la posibilidad de que la apertura de información se dé no sólo en entidades públicas, sino también en todas aquellas -como medios de comunicación, organizaciones de la sociedad civil, etc.- que durante su actividad reúnan datos de interés público y deseen compartirlos, siempre que no vulneren derechos individuales como el de la intimidad y el buen nombre.

Precisamente, de la sociedad civil, así como de los gobiernos, han surgido diversas iniciativas de datos abiertos. Estos son algunos ejemplos:

- *They work for you* (Ellos trabajan para usted): obtiene y publica debates, respuestas escritas, discursos y votaciones del parlamento británico.
- Opendata Latinoamérica: reúne bases de datos de gobiernos e instituciones, a través de la colaboración de organizaciones civiles de diversos países latinoamericanos, y

de información recolectada en talleres y *hackatones* (reuniones de periodistas y desarrolladores).

#### **1.4 MARCO LEGAL**

La legislación colombiana protege la labor periodística, desde su Constitución, así como el derecho de acceso a la información pública:

Artículo 73: La actividad periodística gozará de protección para garantizar su libertad e independencia profesional.

Artículo 23: Toda persona tiene derecho a presentar peticiones respetuosas a las autoridades por motivos de interés general o particular y a obtener pronta resolución. El legislador podrá reglamentar su ejercicio ante organizaciones privadas para garantizar los derechos fundamentales.

Artículo 74: Todas las personas tienen derecho a acceder a los documentos públicos salvo los casos que establezca la ley. El secreto profesional es inviolable.

En cuanto a la regulación de estos principios, el 20 de junio de 2012 fue aprobada en el Congreso la ‘Ley de Transparencia y de Acceso a la Información Pública’, que luego fue revisada por la Corte Constitucional y declarada exequible, con algunas modificaciones y condiciones, el 10 de mayo de 2013. El proyecto de ley fue una iniciativa de la alianza *Más Información, Más Derechos*, conformada por Transparencia por Colombia, la Fundación para la Libertad de Prensa (Flip), el Centro de Estudios de Derecho Justicia y Sociedad (Dejusticia), el Proyecto Antonio Nariño (PAN) y Corporación Ocasá.

Por lo reciente que está la terminación de este proceso, al momento de realizar este trabajo aún no se puede observar la implementación de la ley; sin embargo, hasta este año el país ha

contado con otras herramientas legales para proteger y regular el derecho de acceso a la información, a través de la Ley 57 de 1985, que establece la publicidad de los actos y documentos oficiales, y del Código Contencioso Administrativo, que en su capítulo III determina las actuaciones administrativas en cuanto al derecho de petición en interés particular.

En términos generales, estas son las características actuales del acceso a la información:

- Cualquier ciudadano puede obtener información sobre las actuaciones y el funcionamiento del Estado, así como la que administren todas las oficinas y dependencias públicas. Para obtenerla puede solicitarla de manera escrita (lo que se conoce como el ‘derecho de petición’) u oral.
- La entidad a la que se dirige la solicitud tiene 15 días hábiles para responder y 10 en el caso de solicitudes periódicas. En caso de no ser la administradora de la información solicitada, debe dirigir la solicitud a la entidad correspondiente y hacerlo saber así al solicitante.
- Las respuestas deben ser de fondo, consecuentes con la solicitud hecha, no evasivas.
- Si la entidad no responde en el tiempo establecido, se aplica la acción de tutela.
- Si la entidad considera que no debe entregar la información solicitada, debe argumentar su negativa y sólo puede basarse en una ley –no un decreto o acto administrativo- que específicamente establezca la reserva sobre dicha información.
- Cuando haya una decisión negativa, se debe notificar al solicitante y al Ministerio Público. Quien solicita puede acudir al recurso de insistencia.
- No se puede cobrar el acceso a datos públicos. Sólo se puede cobrar el valor de fotocopias, cuando sea necesario, o las tarifas que establezca la ley para la expedición de documentos o certificados.

Por otra parte, el gobierno colombiano se unió, el 20 de septiembre de 2011, a la *Open Government Partnership* (Alianza de Gobierno Abierto), un proyecto en el que varios países se comprometen a tomar acciones para promover la transparencia, empoderar a los ciudadanos y aprovechar las nuevas tecnologías para fortalecer la democracia.

## 2. DISEÑO METODOLÓGICO

### 2.1 METODOLOGÍA DE LA INVESTIGACIÓN

Este proyecto, que busca identificar las condiciones actuales en las que se da el periodismo de bases de datos en Colombia, con especial énfasis en los medios de comunicación *El Espectador* y *La Silla Vacía*; tiene un enfoque cualitativo, ya que reúne los datos necesarios en una variedad de formas (anotaciones durante las entrevistas, transcripciones, observación, documentación textual), más allá de la numérica, y los va evaluando durante el proceso de investigación, hasta llegar a una conclusión con un método inductivo.

El trabajo empleó técnicas etnográficas de investigación: encuestas, análisis de contenido y entrevistas, con el fin de realizar dos estudios de caso –el de *El Espectador* y el de *La Silla Vacía*–, ya que esta metodología permite recolectar suficiente información para comprender un fenómeno y describirlo, como en este caso se pretende hacer con respecto al periodismo de datos y su aplicación en Colombia. Las características del estudio de caso que enuncia Merriam (1988), citada por Wimmer y Dominick (2002), se ajustan perfectamente al trabajo que se busca desarrollar con este proyecto:

- Particularista: los estudios de caso se concentran en una situación particular. Del mismo modo, esta investigación se ciñe únicamente a la práctica del periodismo de bases de datos en los medios de comunicación escritos (prensa e Internet) en Colombia, con especial énfasis en dos de ellos: *El Espectador* y *La Silla Vacía*.
- Descriptivo: el resultado de un estudio de caso es una descripción detallada del asunto estudiado y así se espera que sea el producto final del presente proyecto, con respecto a los dos medios mencionados.
- Heurístico: ayuda a entender el fenómeno que se investiga. Del mismo modo, este trabajo pretende comprender cómo se está dando el periodismo de bases de datos en los medios estudiados y alcanzar así una nueva perspectiva que permita seguir avanzando sobre esta técnica periodística en el país.

- Inductivo: precisamente, por acudir a un caso particular para comprender a través de él la situación actual del periodismo de bases de datos en Colombia, de una manera más general, esta investigación es inductiva.

## **2.2 FASES DE LA INVESTIGACIÓN**

La investigación se desarrolló en cuatro fases que permitieron estudiar el fenómeno del periodismo de bases de datos desde igual número de fuentes que le atañen: los expertos que lo han estudiado y experimentado, los editores y directivos que determinan las líneas de acción en medios impresos y digitales, los contenidos producidos en dos medios de comunicación (*El Espectador* y *La Silla Vacía*) que emplean esa técnica periodística y los periodistas que utilizan ese método de trabajo en dichos medios.

### **2.2.1 Fase 1: Documentación**

Durante un periodo de aproximadamente seis meses se mantuvo un proceso de documentación con respecto a los principales aspectos del periodismo de bases de datos. Las principales acciones realizadas y fuentes consultadas se pueden clasificar así:

- Bibliografía y cibergrafía: libros como el *Manual de Periodismo de Datos* y la guía escrita por Giannina Segnini sobre periodismo asistido por computador, así como el material del *blog* Nación Data del diario *La Nación* de Argentina, fueron fundamentales para comprender el concepto del periodismo de bases de datos, sus características, utilidades y principales herramientas. Además, se recurrió al material bibliográfico de Philip Meyer, José Luis Dader y Brant Houston, precursores de las raíces del periodismo de bases de datos, para vislumbrar sus orígenes.
- Seguimiento a leyes y programas de gobierno: la participación en los primeros foros para proponer la Ley de Transparencia y Acceso a la Información y el seguimiento de su proceso hasta que fue aprobada por la Corte Constitucional en medio del

desarrollo de este proyecto, permitieron observar los primeros pasos de transformación de una herramienta que es fundamental para el periodismo de bases de datos. Del mismo modo, se hizo seguimiento a los programas Gobierno Abierto y Gobierno en Línea del Estado colombiano.

- Diálogo con expertos: se consultó a Giannina Segnini, fundadora y editora de la Unidad de Investigación del periódico *La Nación*, de Costa Rica; Sandra Crucianelli, becaria Knight, fundadora y editora de Sololocal.info; y Brant Houston, presidente de la junta directiva de Investigative News Network (Red de Noticias de Investigación), entorno a las condiciones que ellos consideran necesarias con respecto a los periodistas, los medios de comunicación y el Estado, para que el periodismo de bases de datos surja y avance en un país.

### 2.2.2 Fase 2: Diagnóstico

Se diseñó un cuestionario para ser aplicado a directores, editores y jefes de redacción de medios de comunicación escritos (incluye impresos y digitales) de diferentes ciudades colombianas. Con el fin de alcanzar un diagnóstico básico del grado de uso que se le da al periodismo de bases de datos en dichas organizaciones periodísticas, el cuestionario tuvo un enfoque de encuesta descriptiva, ya que esta intenta, en palabras de Wimmer y Dominick (2002), “retratar o documentar condiciones o actitudes actuales, es decir, se describe lo que existe en el momento” (p. 54). Once medios, de quince contactados, respondieron la encuesta, a través de Internet:

**Tabla 1. Medios de comunicación que respondieron la encuesta**

Nombre del medio	Tipo	Ciudad	Persona que respondió
<b>Dinero</b>	Impreso	Bogotá	Aldemar Moreno, editor de negocios
<b>El Colombiano</b>	Impreso	Medellín	Nelson Matta, editor del área de Investigaciones
<b>El Espectador</b>	Impreso	Bogotá	Elber Gutiérrez, jefe de redacción

<b>El Meridiano de Córdoba</b>	Impreso	Montería	Ginna Morelo, editora general
<b>El País</b>	Impreso	Cali	Fabio Posada, editor de investigación
<b>El Tiempo</b>	Impreso	Bogotá	Martha Soto, editora de la Unidad Investigativa
<b>La Patria</b>	Impreso	Manizales	Fernando Ramírez, editor de noticias
<b>La Silla Vacía</b>	Web	Bogotá	Juan Esteban Lewin, periodista
<b>Semana</b>	Impreso	Bogotá	Carlos Eduardo Huertas, editor de investigaciones
<b>Vanguardia Liberal</b>	Impreso	Bucaramanga	Diana Saray Giraldo, editora general
<b>Verdad Abierta</b>	Web	Bogotá	César Molinares Dueñas, editor

### 2.2.3 Fase 3: Estudios de caso

Los medios seleccionados para hacer un estudio más detallado de ellos fueron *El Espectador* y *La Silla Vacía*, ya que ambos tienen una apuesta consciente y definida por el periodismo de bases de datos. Además, el análisis de estos casos permitirá ver el manejo que se le da a esta práctica desde dos orillas diferentes: un medio de comunicación tradicional impreso y un medio exclusivo para la web. Los estudios se dieron en dos pasos:

#### Análisis de contenido

Se tomó una muestra de seis trabajos periodísticos que involucran el análisis de bases de datos en cada uno de los medios a estudiar, publicados entre 2012 y mayo de 2013. Este corpus fue establecido junto con los periodistas encargados del periodismo de bases de datos en cada medio, teniendo en cuenta los ejemplos identificados por la investigadora y los que ellos consideraron más representativos del tipo de trabajo que realiza su medio. El

fin de este instrumento de investigación fue identificar las principales temáticas y herramientas empleadas en los trabajos seleccionados, para hacer un análisis comparativo.

Para caracterizar los artículos periodísticos, se establecieron los siguientes grupos de categorías. En algunos casos se consultó a los autores para conocer las herramientas empleadas:

Datos básicos:

- Título del artículo
- Enlace web
- Fecha de publicación
- Autor (es)

Análisis temático:

- Sección o proyecto: se indica en qué área fue incluido cada artículo, según el criterio y la estructura que emplea cada medio.
- Área temática: aspecto de la realidad que aborda el artículo (política, social, ambiental, económica, internacional, ect.).
- Tipo de fiscalización: las opciones aquí fueron establecidas según el tipo de uso, generalmente veedor, que se le da al periodismo de bases de datos, buscando responder qué se intentó fiscalizar en cada ocasión: contratación pública, figuras públicas, gasto público, sector industrial.
- Tipo de resultado: se indica si la investigación hecha para el artículo periodístico arrojó una denuncia, sea de un hecho probadamente ilegal o de una situación moralmente indignante; un llamado de atención ante determinada situación; o un perfil descriptivo, sea de personas, instituciones o sectores.

Análisis del tipo de publicación:

- Género periodístico: por el tipo de temas que trabaja el periodismo de bases de datos, aquí las opciones estarán limitadas a noticia o reportaje.
- Extensión del texto: medido en número de palabras, incluyendo títulos y sumarios.

- Infografía: estática, interactiva o ninguna.
- Explicación: se indica si el medio describió el análisis de datos realizado, sea de manera somera (solo una mención), detallada o ninguna.
- Datos en bruto: se indica si los datos originales a partir de los cuales el medio hizo el análisis fueron publicados o no.

Análisis de herramientas empleadas para la obtención y análisis de los datos:

- Solicitud de información pública mediante derecho de petición (DP)
- Bases de datos que permiten consultas individuales en la web (BW)
- Bases de datos descargables de Internet (BD)
- Bases de datos construidas por los periodistas o por el medio (BC)
- Análisis geográfico (AG)
- Análisis estadístico (AE)
- Cruce de datos (CD)

#### Entrevistas a profundidad

Se realizaron entrevistas intensivas con los protagonistas de la producción de piezas periodísticas de análisis de bases de datos: Alexander Marín en *El Espectador* y Juanita León y Juan Esteban Lewin en *La Silla Vacía*. Se buscó conocer sus principales proyectos, los recursos humanos y tecnológicos con los que trabajan, sus dinámicas periodísticas y sus consideraciones sobre el acceso a la información pública en Colombia. El informe de estas entrevistas será presentado en formato de perfil periodístico de cada medio.

#### **2.2.4 Fase 4: Análisis**

Finalmente, los elementos recolectados en las fases 2 y 3 fueron analizadas con base en las características y condiciones del periodismo de bases de datos que se identificaron en la fase de documentación, así como en el marco referencial. Los capítulos 3 y 4 del presente informe dan cuenta de los resultados.

### 3. ¿CÓMO ESTÁ EL PERIODISMO DE BASES DE DATOS EN COLOMBIA?

Con el fin de conocer las condiciones necesarias para que en un país pueda surgir y evolucionar el periodismo de bases de datos, se entabló un diálogo con tres expertos:

**Brant Houston.** Es considerado precursor del periodismo asistido por computador, autor de tres ediciones de *Computer-Assisted Reporting: A Practical Guide*. Es presidente de Investigative News Network y del Fund for Investigative Journalism, además de cofundador de la Global Investigative Journalism Network - GIJN. Fue director ejecutivo de Investigative Reporters and Editors (IRE) durante 12 años. En abril de 2013 participó como conferencista en el VI Encuentro de Periodismo de Investigación en Bogotá, lo cual permitió conversar con él en persona acerca del asunto que se trata en este apartado.

**Giannina Segnini.** Editora de la Unidad de Investigación que fundó en el diario La Nación, de Costa Rica. Ha recibido el Gran Premio SIP a la Libertad de Prensa de la Sociedad Interamericana de Prensa (2005), el Premio Ortega y Gasset de España (2005); en dos oportunidades el Premio a la Mejor Investigación Periodística de un Caso de Corrupción en América Latina y el Caribe, del Instituto Prensa y Sociedad; y mención especial en el Premio Maria Moors Cabot, de la Universidad de Columbia. Para el ejercicio de conceptualización de este trabajo fue contactada mediante correo electrónico.

**Sandra Crucianelli.** Periodista argentina, fundadora y editora del medio digital Solo Local.info. Se especializó en periodismo de investigación, periodismo de precisión, fuentes digitales y periodismo de bases de datos. Es miembro del consejo asesor del Centro de Periodismo Digital de la Universidad de Guadalajara, docente del Media Center de la Universidad Internacional de La Florida y del Centro Knight para el Periodismo en las Américas. Autora del libro *Herramientas Digitales para Periodistas*. En diálogo con ella pidió tomar sus ideas sobre las condiciones para el periodismo de bases de datos, del artículo *Periodismo de Datos: Lo bueno, lo malo y lo feo*, publicado por ella en 2012.

La pregunta que dirigió la conversación fue: ¿cuáles son las condiciones necesarias o favorables -en cuanto a los periodistas, los medios de comunicación y el Estado- para que en un país surjan y se consoliden proyectos de periodismo de datos? Las respuestas de los tres expertos encuentran varios puntos de consenso y algunas ideas independientes que enriquecen el diálogo, como se puede ver en la siguiente tabla de categorización:

**Tabla 2: Condiciones necesarias para el periodismo de bases de datos, según los expertos**

	<b>Brant Houston</b>	<b>Giannina Segnini</b>	<b>Sandra Crucianelli</b>
<b>Periodistas</b>	Habilidades para manejar programas como Excel, Access, Google Fusion.	Dominio de herramientas y programas para procesar datos.	Conocer los principios básicos del programa Excel o similar.
	Habilidades de búsqueda en línea para extraer bases de datos.		Habilidades desarrolladas en materia de búsquedas en la Web.
		Capacidad de análisis de datos.	Conocimientos de matemática y estadística.
	Criterio para convertir datos en historias periodísticas.	Actitud crítica.	
	Estar dispuestos a elaborar sus propias bases de datos.		
	Saber consultar bases de datos nacionales y de otros países.		
			Entrenamiento permanente, muchas veces autodidacta.
			Entender el funcionamiento del Estado y saber interpretar leyes.
			Dominio de inglés.

<b>Medios</b>		Proporcionar tiempo a los periodistas.	
	Crear equipos con varias personas que sepan procesar datos, y dos con criterio periodístico que los entiendan.	Integrar equipos multidisciplinarios (ingenieros, estadísticos, geógrafos, sociólogos, etc.).	
	No depender sólo del 'nerd' que sabe de todo y al que todos acuden.		
		Dotar de recursos tecnológicos.	
<b>Estado</b>	Adquirir la rutina de subir a la web todo tipo de datos sobre el Estado.	Poner a disposición de los ciudadanos todos los datos sobre gestión del Estado.	Datos abiertos: reutilizables, descargables, actualizados, ordenados.
		Garantizar acceso gratuito o a bajo costo de los datos de registro públicos.	
		Promulgar legislación y normativa que facilite el acceso a información relevante y actualizada.	

Como se puede observar, el gran consenso en cuanto a los periodistas es que necesitan saber manejar programas de análisis de datos, para lo cual es indispensable la capacitación permanente. Por lo demás, la mayoría de requerimientos mencionados para los periodistas no distan de lo que debería tener cualquier reportero que desee hacer su trabajo con cierto nivel de profundidad e investigación (conocimiento del Estado, actitud crítica, habilidades de búsqueda en Internet).

En cuanto al papel de los medios de comunicación, es evidente que se requiere compromiso para proveer los recursos, tiempos y equipos periodísticos necesarios. El trabajo de un

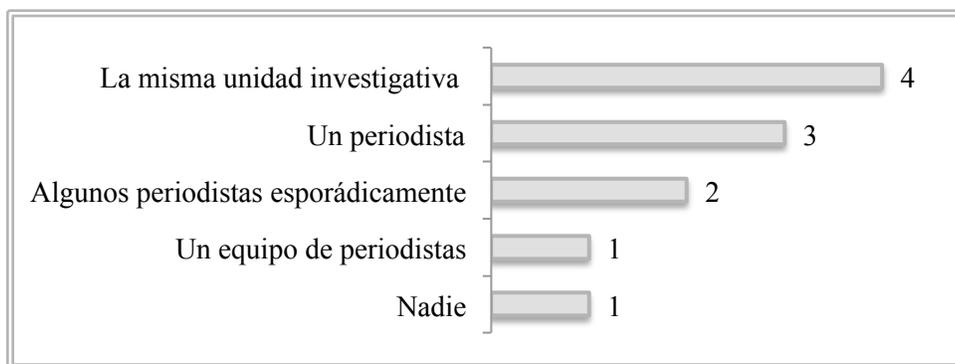
periodista en solitario no da frutos en el periodismo de bases de datos, por el contrario, la interdisciplinariedad se convierte en su gran aliada.

Finalmente, la función del Estado está en la cultura de datos abiertos, la cual implica no sólo permitir que quienes quieran conocer determinada información puedan acceder a ella gratuitamente, sino tener una actitud proactiva: subir permanentemente datos a la web, en formatos legibles y reutilizables, así como proteger ese derecho mediante legislación.

Con base en estas disertaciones, se empleó la encuesta aplicada a nueve medios impresos y dos digitales, las entrevistas en *La Silla Vacía* y *El Espectador* y la revisión documental, para analizar las condiciones del periodismo de bases de datos en Colombia:

### 3.1 LOS MEDIOS

Todos los editores o jefes de redacción de los medios encuestados aseguraron conocer el periodismo de bases de datos. En general, se observa un deseo por apoyar el uso de esa técnica en sus medios. En la mayoría hay alguien que aprovecha el análisis de bases de datos en su labor periodística, sólo uno dijo que nadie lo hace. Los que sí involucran la herramienta, lo hacen a través de equipos periodísticos (5 medios), un periodista solo (3 medios) o varios de la redacción de manera esporádica (2 medios), como se observa en la siguiente gráfica:



**Gráfica 1: Equipos de análisis de bases de datos en los medios encuestados**

Sin embargo, al entrar en los detalles sobre cómo está organizado el medio para realizar ese tipo de trabajos, se observa una falta de estructura. Cuando es un único periodista que emplea por su cuenta el análisis de bases de datos, se trata de la misma unidad investigativa o de otros miembros de la redacción que, de manera ocasional y ‘participativa’ acuden a esa herramienta de investigación. Adicionalmente, llama la atención que no se involucran expertos de otras áreas en estos equipos de trabajo.

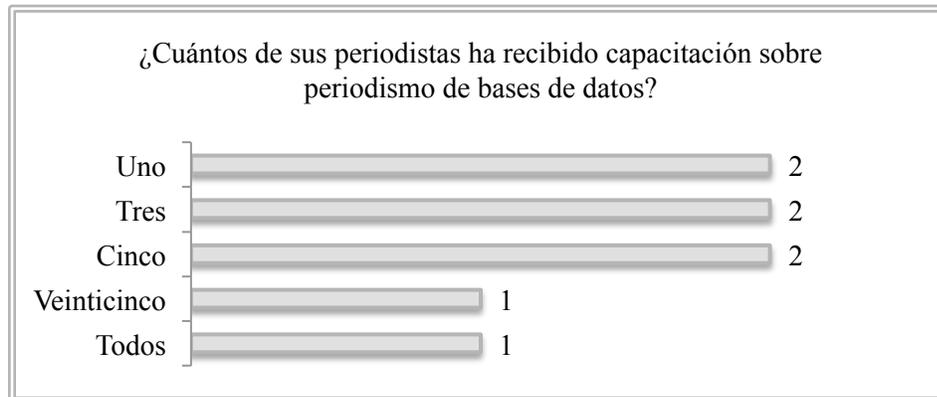
En este sentido, el periodismo de bases de datos parece ser más una iniciativa de algunos periodistas en medio de su trabajo cotidiano, que un área impulsada de manera consciente y estructurada por parte de los medios (con algunas excepciones como la de los dos medios seleccionados para los estudios de caso que se describen en el siguiente capítulo). Otro factor que evidencia este hecho es que sólo dos de los medios encuestados dijeron haber pagado la capacitación de sus periodistas en áreas del periodismo de bases de datos.

Los ejemplos que enviaron los encuestados sobre reportajes que involucran el periodismo de bases de datos en cada medio, dejan ver que es muy frecuente la consulta de bases de datos en línea; en otras ocasiones se observa el uso de herramientas más avanzadas como el cruce de grandes cantidades de datos. En todo caso, sería necesario hacer un estudio muy detallado de los casos vistos para saber qué tan avanzada está esa técnica en cada medio – incluso, en algunas de las respuestas se observan confusiones sobre lo que es y lo que no es periodismo de bases de datos (estudio de informes, documentos, etc.); pero se alcanza a observar que hay una tendencia creciente a la utilización de la mencionada técnica.

### **3.2 LOS PERIODISTAS**

En cuanto a las condiciones de los periodistas, se preguntó acerca de su formación específica y sobre las herramientas de obtención de datos que más utilizan. En ocho de los once medios encuestados al menos algún periodista ha tomado capacitaciones sobre

periodismo de bases de datos. Incluso, llama la atención que, en muchos casos, la formación la ha recibido más de un miembro de la redacción:



**Gráfica 2: Periodistas que han recibido capacitación sobre periodismo de bases de datos**

En esa capacitación hay un evidente liderazgo de Consejo de Redacción (CdR), la organización de periodistas que promueve el periodismo de investigación en el país, que tiene periodistas asociados en diversas ciudades y que ha dictado varios talleres sobre herramientas digitales de investigación. De hecho, cinco de los ocho medios cuyos periodistas han tomado formación en este tema dicen que lo han hecho a través de CdR.

Con respecto a las herramientas para obtener datos, el siguiente es el ranking de las ‘favoritas’, según el número de medios que dijeron usarlas con mucha frecuencia:

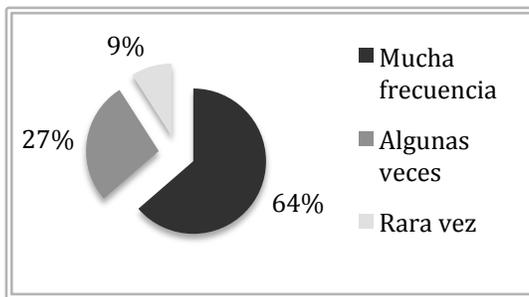
- 1°. Búsqueda avanzada de información y documentos en Internet
- 2°. Bases de datos obtenidas mediante derecho de acceso a la información pública
- 3°. Bases de datos de consulta en línea
- 4°. Bases de datos elaboradas por los periodistas
- 5°. Bases de datos descargadas de internet

Este resultado coincide con el estado inicial en el que se encuentra el periodismo de bases de datos en el país. Las herramientas más usadas son las más difundidas por los avances tecnológicos (buscadores en Internet, bases de datos en línea) y las tradicionales (derecho

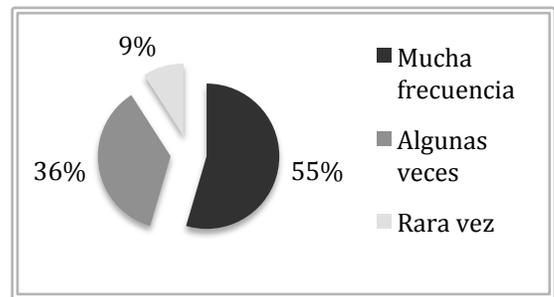
de petición); con lo cual se evidencia la transición que se está dando hacia la exploración de métodos más avanzados –en cuanto al periodismo de bases de datos, cabe resaltar-. Es posible que, en algunos años, los últimos dos puntos de este ranking tengan niveles superiores de uso, si la tendencia del análisis periodístico de datos sigue creciendo.

Cabe, además, resaltar que es muy positivo el empleo que se hace del derecho de acceso a la información –mediante solicitudes escritas o consultas en la web-, pues deja ver la inquietud de los periodistas por obtener información concreta sobre el funcionamiento y ejecución de la función pública, para realizar la veeduría que les corresponde.

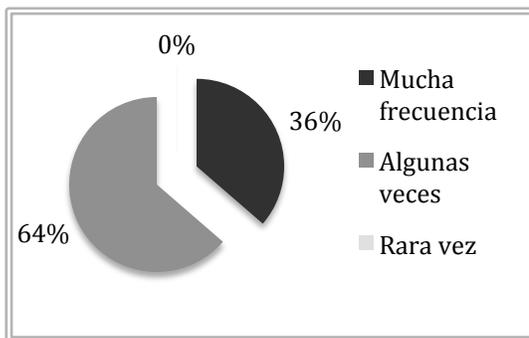
Las siguientes fueron las respuestas con respecto al uso de cada herramienta:



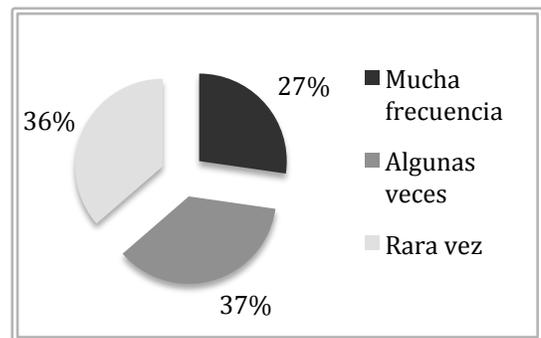
**Gráfica 3: Uso de la búsqueda avanzada de información y documentos en Internet**



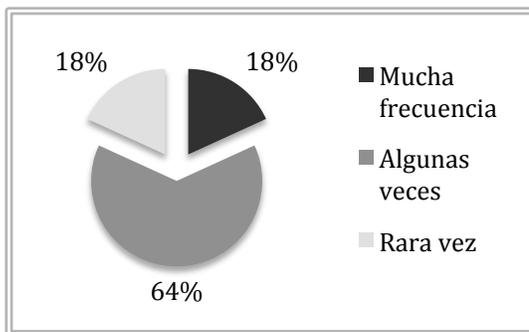
**Gráfica 4: Uso de bases de datos obtenidas con derecho de acceso a la información**



**Gráfica 5: Uso de bases de datos de consulta en línea**



**Gráfica 6: Uso de bases de datos elaboradas por los periodistas**



**Gráfica 7: Uso de bases de datos descargadas de Internet**

### 3.3 EL ESTADO

Con respecto a la nueva Ley de Transparencia y Acceso a la Información –aprobada en mayo de 2013-, ésta ha generado una fuerte polarización, pues algunos periodistas consideran que pueden llevar a una mayor ocultación de información, por la manera en que quedaron redactados algunos puntos. Entre los reparos de Donadío (2012) está el hecho de que quedaron exceptuados de la ley los documentos, bases de datos y contratos relacionados con la defensa y la seguridad nacional, el orden público y las relaciones internacionales. Ese artículo fue declarado inconstitucional por la Corte Constitucional, según *El Espectador* (2013).

Otras de las preocupaciones que expresa Donadío tienen que ver con el hecho de que se crea una nueva categoría de reserva para los documentos que contengan opiniones de servidores públicos, dentro de los procesos deliberativos que éstos tengan; y el hecho de que deja en manos de funcionarios públicos determinar cuándo el daño causado a un interés protegido es mayor al interés general de acceso a la información. Los promotores de la ley han hecho énfasis en que aún así se mantiene la obligación de contar con una reserva legal para negar el acceso a cualquier información pública.

*El Espectador*, en su editorial publicado el 12 de mayo de 2013, considera que, a pesar de las confusiones, la ley tiene cosas buenas que el gobierno puede aprovechar si logra una buena aplicación de la misma. Para el medio, el problema principal es la ‘cultura del secreto’, por la que muchos funcionarios públicos creen que lo mejor es mantener la mayor cantidad de información en reserva, esquivando, en muchas ocasiones, la ley.

Mientras se espera la implementación de la nueva ley para poder evaluarla, hay varios datos que hablan de cómo perciben actualmente la situación del acceso a la información los periodistas colombianos. Para empezar, en la Segunda encuesta nacional a periodistas sobre libertad de expresión y acceso a la información, que realiza el Proyecto Antonio Nariño – PAN- (2013), el 53% de los entrevistados dijo observar una mejoría en este aspecto en su departamento durante el último año. Adicionalmente, los periodistas de La Silla Vacía y de *El Espectador* –Juan Esteban Lewin y Alexander Marín, respectivamente- que fueron entrevistados para los estudios de caso del presente trabajo afirmaron que también ven avances, principalmente en lo que tiene que ver con el acceso a información en Internet.

Sin embargo, persisten los obstáculos. Ambos dicen haber recibido, en algunas ocasiones, respuestas evasivas, códigos en lugar de nombres concretos, estadísticas en lugar de datos específicos, o negativas sin fundamento por parte de entidades del Estado, a lo que reaccionan acudiendo a las herramientas legales que existen: derecho de insistencia y tutela (sólo Alexander dijo haberla usado en una ocasión). Los encuestados por el PAN (2013) también han enfrentado este tipo de obstáculos: el 57% trabas administrativas, al 51% le han exigido revelar el motivo de su solicitud, al 47% le han pedido no revelar la fuente.

Con respecto a la apertura de datos, a través de la *Open Government Partnership* (Alianza de Gobierno Abierto) el Estado colombiano ha iniciado una serie de proyectos para publicar información que sirva para el control ciudadano como el Portal de Transparencia Económica y el Catálogo de Datos Abiertos. Esta estrategia todavía está en etapa de implementación y, como señala la mencionada Alianza, aún no se tienen indicadores para medir el éxito de la misma. Precisamente, el portal de datos abiertos, que es el más

importante en esta línea, sólo tiene 39 bases de datos publicadas al momento de revisión para este proyecto (mayo de 2013), y algunas son poco relevantes para el análisis periodístico y para un verdadero control de la administración pública: ubicaciones de las instituciones educativas oficiales en Barranquilla, sitios de interés en Acacías, etc. De todos modos, existen muchos otros sitios web de entidades oficiales que cada vez más publican bases de datos en línea o las dejan consultar allí; la plataforma Zoom Online, de Consejo de Redacción, reseña 140 de ellas.

En cuanto a su uso, los dos periodistas entrevistados dicen que cada vez encuentran más páginas web de este tipo y las usan con frecuencia. Las diferencias se ven en el tipo de utilidad que encuentran: Lewin, por ejemplo, pudo descargar una base de datos del Portal de Transparencia Económica, lo que le permitió, fácil y rápidamente, hacer un análisis para publicar un artículo periodístico de utilidad para los lectores. Marín, por su parte, emplea bastante el Portal Único de Contratación, pero en cada investigación debe descargar uno por uno cientos de documentos, clasificarlos y reorganizarlos para poderlos procesar; lo cual demuestra que para el periodismo de bases de datos resulta mucho más útil la apertura de grandes cantidades de datos en formatos reutilizables, que la opción de hacer consultas individuales. De hecho, la dificultad que manifestaron ambos periodistas está en que, muchas veces, se encuentran con bases de datos –obtenidas en Internet o mediante derecho de petición- que están en formato PDF, escaneadas, protegidas con clave u organizadas de manera que su análisis se hace demasiado complicado.

Finalmente, se observan algunas iniciativas de organizaciones civiles que pueden ser útiles para el acceso a datos de interés por parte de los periodistas. Entre ellas están Cuentas Claras, desarrollado por Transparencia por Colombia y donado al Consejo Nacional Electoral, en el que se puede ver la rendición de cuentas de las campañas electorales; y Congreso Visible, con el que la Universidad de los Andes hace seguimiento al Congreso registrando en una base de datos toda la actividad legislativa de esa corporación, así como los datos principales de los congresistas.

#### 4. ESTUDIOS DE CASO: EL ESPECTADOR Y LA SILLA VACÍA

*El Espectador* es un periódico de circulación diaria, nacional, que completó 126 años de existencia en marzo de 2013. Ese mismo mes La Silla Vacía, un medio digital que se interesa por cubrir el poder en Colombia, cumplió sus primeros 4 años. Ambos publican continuamente historias que involucran el análisis de bases de datos, a partir de un interés claramente definido por promover ese tipo de trabajos; por eso, y por el contraste que representa explorar dos medios de comunicación importantes para el país pero con plataformas, estructuras y orígenes diferentes, fueron seleccionados para realizar los estudios de caso cuyos resultados se exponen a continuación.

##### 4.1 DOCE EJEMPLOS DE FISCALIZACIÓN DEL PODER EN COLOMBIA

En cada medio fueron seleccionadas seis muestras periodísticas que fueron analizadas para reconocer en cada una la clase de temática trabajada, cómo fue abordada, qué herramientas se emplearon en la investigación y cómo resultó la publicación final, teniendo en cuenta lo que es característico del periodismo de bases de datos.

Para el estudio se elaboraron las siguientes fichas<sup>1</sup> que incluyen 20 características.

**Tabla 3: Análisis de contenido - muestra periodística de El Espectador**

Título: Los contraticos de la Cámara			
Fecha de publicación: <sup>2</sup> 29 de julio de 2012		Autor (es): Alexander Marín	
<b>Análisis temático</b>			
Sección	Área	Fiscalización	Resultado
Tema del día	Política	Contratación pública	Denuncia (caso indignante)

<sup>1</sup> Para conocer la definición de cada categoría y el significado de las siglas utilizadas, es necesario remitirse al capítulo 2 del presente informe: Diseño Metodológico.

<sup>2</sup> En el caso de El Espectador, la fecha indicada en la tabla es la que aparece en la versión digital de cada artículo, que a veces difiere un día de la fecha de publicación impresa.

Análisis del tipo de publicación						
Género	Extensión	Infografía	Explicación	Datos en bruto		
Reportaje	1.475	Estática	Somera	NO		
Análisis de herramientas						
DP	BW	BD	BC	AG	AE	CD
SÍ	SÍ	NO	NO	NO	SÍ	NO

Título: Siniestros en la Cámara						
Fecha de publicación: 16 de septiembre de 2012			Autor (es): Alexander Marín			
Análisis temático						
Sección	Área	Fiscalización	Resultado			
Investigación	Política	Gasto público	Llamado de atención			
Análisis del tipo de publicación						
Género	Extensión	Infografía	Explicación	Datos en bruto		
Reportaje	1.339	Estática	Somera	NO		
Análisis de herramientas						
DP	BW	BD	BC	AG	AE	CD
NO	SÍ	SÍ	NO	NO	SÍ	NO

Título: Gasolina a chorros						
Fecha de publicación: 16 de octubre de 2012			Autor (es): Felipe Morales y Alexander Marín			
Análisis temático						
Sección	Área	Fiscalización	Resultado			
Investigación	Política	Gasto público	Denuncia (caso indignante)			
Análisis del tipo de publicación						
Género	Extensión	Infografía	Explicación	Datos en bruto		
Reportaje	471	Estática	NO	NO		
Análisis de herramientas						
DP	BW	BD	BC	AG	AE	CD
SÍ	NO	NO	NO	NO	SÍ	NO

Título: Así está el mapa minero del país						
Fecha de publicación: 8 de abril de 2013			Autor (es): Alexander Marín, Viviana Londoño y Jonathan Rwoilsonh			
Análisis temático						
Sección	Área	Fiscalización	Resultado			
Infografía	Economía	Sector industrial	Perfil (sectorial)			
Análisis del tipo de publicación						

Género	Extensión	Infografía	Explicación	Datos en bruto		
Gráfico	129	Estática	Somera	NO		
<b>Análisis de herramientas</b>						
DP	BW	BD	BC	AG	AE	CD
SÍ	NO	NO	NO	SÍ	SÍ	SÍ

<b>Título: Los empresarios Uribe Moreno</b>						
Fecha de publicación: 20 de abril de 2013			Autor (es): Alexander Marín			
<b>Análisis temático</b>						
Sección		Área		Fiscalización		Resultado
Economía		Economía		Figuras públicas		Perfil (personal)
<b>Análisis del tipo de publicación</b>						
Género	Extensión	Infografía	Explicación	Datos en bruto		
Reportaje	1.660	Ninguna	Somera	NO		
<b>Análisis de herramientas</b>						
DP	BW	BD	BC	AG	AE	CD
NO	SÍ	SÍ	SÍ	NO	SÍ	SÍ

<b>Título: AH1N1, una gripe costosa para el erario</b>						
Fecha de publicación: 4 de mayo de 2013			Autor (es): Alexander Marín			
<b>Análisis temático</b>						
Sección		Área		Fiscalización		Resultado
Nacional		Salud		Gasto público		Denuncia (ilegalidad)
<b>Análisis del tipo de publicación</b>						
Género	Extensión	Infografía	Explicación	Datos en bruto		
Reportaje	1.891	Estática	Detallada	NO		
<b>Análisis de herramientas</b>						
DP	BW	BD	BC	AG	AE	CD
SÍ	NO	NO	SÍ	NO	SÍ	SÍ

**Tabla 4: Análisis de contenido - muestra periodística de La Silla Vacía**

<b>Título: El top de los magistrados viajeros</b>						
Fecha de publicación: 6 de enero de 2012			Autor (es): Juan Esteban Lewin			
<b>Análisis temático</b>						
Sección		Área		Fiscalización		Resultado
Historias		Política		Gasto público		Llamado de atención

<b>Análisis del tipo de publicación</b>						
Género	Extensión	Infografía	Explicación	Datos en bruto		
Reportaje	1.411	NO	Somera	SÍ		
<b>Análisis de herramientas</b>						
DP	BW	BD	BC	AG	AE	CD
NO	NO	SÍ	NO	NO	SÍ	NO

<b>Título: Las fichas para la reelección del Procurador</b>						
Fecha de publicación: 1 de Noviembre de 2012			Autor (es): Juan Esteban Lewin			
<b>Análisis temático</b>						
Sección	Área	Fiscalización	Resultado			
Historias	Política	Figura pública	Denuncia (caso indignante)			
<b>Análisis del tipo de publicación</b>						
Género	Extensión	Infografía	Explicación	Datos en bruto		
Reportaje	321	Interactiva	Detallada	SÍ		
<b>Análisis de herramientas</b>						
DP	BW	BD	BC	AG	AE	CD
SÍ	SÍ	NO	SÍ	SÍ	SÍ	SÍ

<b>Título: Así son los que mandan en las empresas más grandes del país</b>						
Fecha de publicación: 4 de Febrero de 2013			Autor (es): David Graaff, Juan Pablo Pino, Juan Esteban Lewin			
<b>Análisis temático</b>						
Sección	Área	Fiscalización	Resultado			
Historias	Política	Sector industrial	Perfil (sectorial)			
<b>Análisis del tipo de publicación</b>						
Género	Extensión	Infografía	Explicación	Datos en bruto		
Reportaje	382	Interactiva	Somera	SÍ		
<b>Análisis de herramientas</b>						
DP	BW	BD	BC	AG	AE	CD
NO	SÍ	NO	SÍ	SÍ	SÍ	SÍ

<b>Título: El nuevo quién es quién del Banco de la República</b>						
Fecha de publicación: 21 de Febrero de 2013			Autor (es): Juan Esteban Lewin			
<b>Análisis temático</b>						
Sección	Área	Fiscalización		Resultado		
Historias	Económica	Figuras públicas		Perfil (institucional)		
<b>Análisis del tipo de publicación</b>						
Género	Extensión	Infografía		Explicación	Datos en bruto	
Noticia	216	Interactiva		Somera	NO	
<b>Análisis de herramientas</b>						
DP	BW	BD	BC	AG	AE	CD
NO	NO	NO	SÍ	SÍ	SÍ	NO

<b>Título: El agua que pasa por debajo del descontento en la Unidad Nacional</b>						
Fecha de publicación: 3 de Abril de 2013			Autor (es): Juan Esteban Lewin			
<b>Análisis temático</b>						
Sección	Área	Fiscalización		Resultado		
Historias	Política – social	Gasto público		Llamado de atención		
<b>Análisis del tipo de publicación</b>						
Género	Extensión	Infografía		Explicación	Datos en bruto	
Reportaje	367	Interactiva		Detallada	SÍ	
<b>Análisis de herramientas</b>						
DP	BW	BD	BC	AG	AE	CD
NO	NO	SÍ	SÍ	NO	SÍ	SÍ

<b>Título: Así va la ejecución del gobierno</b>						
Fecha de publicación: 9 de mayo de 2013			Autor (es): Juan Esteban Lewin			
<b>Análisis temático</b>						
Sección	Área	Fiscalización		Resultado		
Historias	Economía	Gasto público		Perfil (sectorial)		
<b>Análisis del tipo de publicación</b>						
Género	Extensión	Infografía		Explicación	Datos en bruto	
Reportaje	942	NO		Detallada	SÍ	

Análisis de herramientas						
DP	BW	BD	BC	AG	AE	CD
NO	NO	SÍ	NO	NO	SÍ	NO

La primera gran conclusión al analizar estos dos grupos de muestras periodísticas es que los dos medios están haciendo una labor muy importante de fiscalización del poder, principalmente en cuanto a la vigilancia del gasto público (50% de los artículos), a través del análisis de bases de datos. Sus investigaciones en este sentido han permitido revelar asuntos relevantes como el desperdicio de fondos en la compra y distribución de tratamientos para la gripa AH1N1, y la concentración de la inversión en agua para los municipios en manos de determinados partidos políticos. Así mismo, se observaron ejercicios interesantes al investigar figuras del poder, indagando sus propiedades, relaciones políticas, intereses particulares, hojas de vida, etc.

El área prioritaria en ambos medios es la política y, en menor medida, la económica. Esto sería de esperarse en La Silla Vacía, pues dicho medio trata, casi de manera exclusiva, asuntos políticos; pero al encontrar una tendencia tan similar en ambos medios, se hace evidente que uno de sus principales motores para el uso del periodismo de bases de datos es aprovecharlo como herramienta para vigilar el poder político y, de este modo, cubrir este asunto desde ángulos y fuentes diferentes a las tradicionales.

Los resultados de los trabajos analizados fueron, mayoritariamente (4 de 6), denuncias de asuntos ilegales o que generan cierta indignación por el manejo de dineros públicos, en el caso de *El Espectador*. Aunque trabajan temas muy parecidos, de los trabajos de La Silla Vacía, sólo uno llegó a ese nivel de denunciar una irregularidad (el de los nombramientos del procurador); los demás tienen un tono de llamado de atención sobre asuntos de interés público (los viajes de los magistrados, la inversión en agua) o generan perfiles que permiten conocer a mayor profundidad a quienes están fuertemente relacionados con el poder político y económico.

Para llegar a estos hallazgos, ambos medios acudieron, principalmente a análisis de tipo estadístico con base en la información que recolectaron. La técnica de cruce de datos se encontró en la mitad de los artículos estudiados tanto del medio impreso como del digital; y la de análisis geográfico apareció más en los trabajos de La Silla Vacía. En este último, además, llama la atención comprobar que los casos en los que se combinaron más herramientas de análisis y más fuentes para la obtención de datos, los resultados fueron más llamativos y relevantes (el artículo sobre el procurador, que tuvo gran impacto, y el trabajo sobre las empresas más grandes del país).

Mientras las similitudes entre los dos medios tienen que ver con asuntos de fondo –las temáticas que trabajan, la utilidad de éstas para fiscalizar la administración pública y las técnicas de análisis–, las diferencias se dan en asuntos de forma: las herramientas para obtener datos y la manera en que presentan los resultados de sus investigaciones. A continuación se exponen estos aspectos:

Ambos medios acuden al estilo del reportaje para presentar sus hallazgos, lo cual concuerda con la tendencia mundial, pues, por la naturaleza del trabajo con bases de datos, es difícil desarrollar crónicas u otro tipo de formato. Sin embargo, la extensión sí varía notablemente; mientras los textos de *El Espectador* tienen una extensión promedio de 1.161 palabras, los de La Silla Vacía tienen un promedio de 539. Además de la plataforma que emplea cada medio (impresa y digital, respectivamente), lo anterior se explica porque La Silla recurre mucho más al uso de infografías interactivas con gran cantidad de contenido.

Precisamente, esa es una de las características del periodismo de bases de datos en las que el mencionado medio digital le lleva una gran ventaja a *El Espectador*. Para empezar, sus infografías no sólo son dinámicas, sino que se vuelven indispensables para complementar el contenido del artículo, mostrar los datos originales y profundizar en el asunto abordado; por el contrario, las del periódico son todas estáticas, incluso en su versión digital.

En lo que tiene que ver con acciones de transparencia por parte de los medios, La Silla Vacía ofrece explicaciones más detalladas sobre el procedimiento de recolección y análisis de datos empleado en cada investigación, aunque todavía no se nota en la totalidad de sus trabajos (la mitad de los estudiados). Así mismo, en cinco de los seis artículos observados, ese medio permite ver la base de datos original con la que trabajó para el análisis (en el caso del texto sobre las empresas más grandes, lo hace a través de listas y *rankings* dentro de la visualización); mientras que *El Espectador* no ofrece esa posibilidad en ninguno de los casos. Más allá de seguir una tendencia internacional que se está dando en los grandes medios líderes en periodismo de bases de datos, esta práctica de publicar los datos en bruto le ha permitido a La Silla Vacía aumentar y fortalecer su interacción con los usuarios; el caso más llamativo es el del artículo ‘Las fichas para la reelección del Procurador’, en el que la colaboración de éstos fue fundamental.

Con respecto a las herramientas para obtener los datos, *El Espectador* parece acudir con mayor frecuencia al derecho de petición para realizar este tipo de trabajos, mientras que La Silla Vacía tiende más a la creación de sus propias bases de datos, alimentadas con información de diversas fuentes personales, documentales y digitales.

Finalmente, durante el análisis fue llamativo ver casos en los que hizo falta equilibrar la información solicitando la versión de las personas cuyas actuaciones están siendo vigiladas y cuestionadas: los magistrados de las Altas Cortes y los representantes a la Cámara, en trabajos de La Silla Vacía y *El Espectador*, respectivamente.

#### **4.2 EXCEL, EL MEJOR ALIADO PARA UN PERIODISTA DE INVESTIGACIÓN EN EL ESPECTADOR**

En enero de 2011 Jhon Alexander Marín Correa salió del diario *La Patria*, en Manizales y, como diría su antiguo jefe, ‘pasó a mejor vida’, porque se fue a un medio de circulación nacional en Bogotá: *El Espectador*.

Llegó gracias a que Jorge Cardona, el editor del periódico, tuvo la idea –que le aprobó la Junta Directiva- de reforzar las cuatro secciones principales del medio llevando un nuevo periodista a cada una para que, cuando alguno del equipo quisiera hacer una investigación pudiera dedicarse a ella por un tiempo, mientras los otros cubrían el día a día. De este modo Alexander estuvo un año en el área judicial, pero cuando el proyecto inicial de Cardona no se cumplió, fue sacado de allí para convertirse en periodista de investigación.

En esa área trabaja también Norbey Quevedo, pero sus labores nunca se cruzan, ya que éste emplea su método personal de trabajo, mientras que Alexander tiene la misión de aplicar técnicas de periodismo de bases de datos y apoyar a sus colegas de la redacción cuando quieran hacer algún trabajo de ese tipo. Incluso, gracias a la experiencia que ha adquirido en el tema, también lo consultan para presentar derechos de petición.

Como método de trabajo, Alexander abre una tabla de Excel cada vez que tiene un nuevo tema y empieza a copiar allí datos como nombres, números de NIT, empresa, cargo, etc., de las personas relacionadas con el asunto. Por eso ese programa se ha convertido en su gran aliado, pues le ayuda, tanto a organizar el material de sus investigaciones, como a realizar análisis de datos, mediante filtros, tablas dinámicas y una que otra fórmula. También crea carpetas en su computador y allí organiza las bases de datos, los documentos que descarga de la web, los soportes de su investigación, etc.

El proceso podría resumirse así: se define un tema, sea por coyuntura, porque está dentro de las áreas de interés o porque apareció mientras se exploraban sitios web o mientras se desarrollaba otra investigación; luego Alexander hace un exploración de las bases de datos para ver si ‘hay tema’, es decir, si hay probabilidades de sacar una historia de ellas; con eso en mano, se le presenta la propuesta al editor. Una vez aprobado el tema, se obtiene toda la información necesaria y se hace una depuración de la base de datos para trabajar sólo con la información que será útil, posteriormente se analizan los datos y, finalmente, se verifican los resultados. De hecho, durante el proceso, este riguroso periodista tiene la costumbre de

confrontar la información que le entregan con la que está en bases de datos de consulta en línea, por ejemplo, para confirmar que determinada cédula sí corresponde a la persona que dice la tabla.

Para estos trabajos cuenta con la ayuda de un practicante y, en ocasiones, los realiza de la mano de otros periodistas de la redacción, según el área que se esté tratando. La función principal del practicante es apoyar la búsqueda de datos; para esto, Alexander les enseña desde el comienzo a conocer bases de datos públicas que se pueden consultar en línea y luego algunos indicios de cómo crear y administrar sus propias bases de datos.

“El hecho de llegar como practicante a un sección de investigación genera un problema y es que se saltan un paso: el del día a día. No tienen el ritmo de qué es la noticia, la premura de un cierre, cómo saber qué es lo más importante en una historia, entonces se pierden ese paso que le forma a uno el criterio periodístico”, afirma Alexander con tono de preocupación. Ante este reto, también entrena a los practicantes en acceso a la información, en cómo dar la pelea cuando las fuentes no responden rápidamente lo que se está buscando. Además les explica la ventaja del derecho de petición frente a las filtraciones: éstas siempre tienen intereses detrás y hay que pasar a confirmar toda la información que se recibe, mientras que la que viene de derechos de petición ya es oficial.

Como editor, Jorge Cardona, a quien Alexander le pide buscar practicantes “que al menos sepan sumar”, pues considera necesario dejar la creencia de que el periodismo y los números no tienen nada que ver, es el encargado de aprobar, blindar (garantizar la veracidad y sostenibilidad de todos los datos) y hacer corrección de estilo en los trabajos del área.

Alexander ha tomado cursos de matemáticas para periodistas, periodismo de investigación, periodismo asistido por computador y periodismo de datos, lo cual, según él, ha sido una de las ventajas de hacer parte de la asociación de periodistas llamada Consejo de Redacción. Esto le ha servido para desarrollar el tipo de trabajo que hoy hace en *El Espectador*, pero reconoce que aún es un área muy reciente en el medio –de hecho, nadie más aplica el

análisis de bases de datos en la redacción-, por lo que falta hacer crecer más la cultura de los datos para que algún día este periódico de 126 años se atreva a publicar en Internet la información en bruto: las bases de datos a partir de las cuales surgen muchas de las historias que desarrolla Álex, y también para que se den cosas como el pago de suscripciones a servicios de datos en línea como información de empresas, de propiedades, de cortes penales, etc.

Entre las cosas que están por venir, o al menos en las que Alexánder espera ver avances, está el trabajo con visualizaciones interactivas. El reportero reconoce que esta es una carencia, pues, aunque trabaja de la mano con el único infografista de la publicación impresa, quien le da muy buenas ideas para generar gráficas que acompañen sus reportes, hace falta producir visualizaciones más animadas para la web, que permitan explicar las investigaciones de una manera más clara en esa plataforma. La otra necesidad tiene que ver con la conglomeración de toda la información que tiene; aunque el periódico no tiene pensado crear una gran base de datos con el material que obtienen todos los periodistas, Alexánder sí espera, en algún momento, tener el tiempo para juntar todos los datos con los que ha trabajado y lograr así una herramienta de cruce de información diversa que, seguramente, enriquecerá aún más sus investigaciones.

#### **4.3 EN LA SILLA VACÍA MUCHAS PREGUNTAS SE RESUELVEN CON BASES DE DATOS**

“Desde que concebimos La Silla Vacía la vimos como un laboratorio de experimento sobre el futuro del periodismo”, recuerda Juanita León, directora de uno de los medios digitales más influyentes en el país, cuando reflexiona sobre porqué usan el periodismo de bases de datos. Dice que le gusta porque es la forma natural de contar historias en la web, ya que permite varios niveles de información y ayuda a profundizar más.

El periodista que más realiza ese tipo de trabajos, aunque no sea muy consciente de ello, es Juan Esteban Lewin. Este abogado e historiador, que llegó al periodismo por ser un muy buen usuario de La Silla Vacía, tiene como método natural de trabajo empezar a elaborar bases de datos, también en Excel, para llenarlas de información organizada a medida que avanzan sus investigaciones. Como no ha tomado ningún curso al respecto, le adjudica esa habilidad al gusto por el análisis cuantitativo de ciencias sociales que aprendió en la universidad.

Tal vez por esa misma formación de carreras estrictas y metodológicas, antes de hablar sobre periodismo de bases de datos Juan Esteban se cuestiona acerca de qué es eso, dónde empieza y dónde termina. La inquietud surge porque quiere saber si la historia puede ser sólo la base de datos: explicar cómo se armó, qué se puede encontrar en ella y dejarla a disposición de los usuarios para que ellos “jueguen” con la información; o si definitivamente se tiene que hacer una historia periodística, sacar una o cinco grandes conclusiones de los datos para, ahí sí, invitar al lector a seguir buscando. Se pregunta si unos pocos análisis estadísticos, no muy sofisticados, ya son periodismo de bases de datos.

Después de conversar sobre lo que se ha hablado de esta discusión en escenarios académicos en todo el mundo y encontrar que los expertos consideran que siempre debe haber un análisis periodístico, por sencillo que sea, Juan Esteban se convence. “Si es así, entonces yo hago mucho periodismo de bases de datos. Solo porque así funciona”.

Entre los proyectos que, en este sentido, ha desarrollado La Silla Vacía, están la sección ‘Quién es quién’. A pesar de que no fue concebida como una base de datos, error que reconocen tanto los periodistas como la editora creativa, Olga Lucía Lozano, en eso se va a convertir dentro de poco. De todos modos, ya ha permitido cruces interesantes al exponer las relaciones que se dan entre figuras del poder en Colombia, pues se trata de una especie de Facebook que deja ver quién es cada una de las personas allí registradas, su trayectoria y las cercanías de todo tipo que hay entre ellas.

El impacto de esta herramienta es un motor que le indica a Juanita que vale la pena seguir trabajando en el área de las bases de datos, a pesar de los costos que requiere. Esa dificultad en términos de recursos, en el caso de La Silla Vacía, se expresa en que consume mucho tiempo de un solo periodista, quien además debe hacer otras cosas para mantener la producción diaria del medio; también en el hecho de que la diseñadora, con quien es absolutamente necesario el diálogo en estos trabajos, no tiene dedicación de tiempo completo sino que trabaja como *freelance*.

Sin embargo, el impacto se ha notado en esa y otras experiencias como la del artículo que destapó las fichas del procurador para ser reelegido: fue una investigación que muchos medios replicaron, que despertó movimientos de indignación entre los ciudadanos y que se escuchó durante un buen tiempo. Por su parte, la publicación sobre las juntas directivas de las empresas más grandes del país ha alcanzado más de 21 mil visitas en un mes y medio.

Detrás de esos trabajos están Olga Lucía, que siempre está pensando cosas nuevas para el medio; Juan Esteban, que trabaja con las bases de datos, las elabora, las limpia, las organiza y las analiza; algún practicante ocasional que ayuda a digitalizar información y a alimentar las tablas de datos; Juanita León, quien se encarga de darle la mirada editora a todos los hallazgos; y, muchas veces, la diseñadora, Lorena Parra. Sobre la interacción con esta última, Lewin indica que favorece el trabajo, pues ella tiene el lenguaje gráfico en la cabeza, le ayuda a aterrizar ideas y a encontrar soluciones.

El origen de los temas es, casi siempre, una pregunta periodística. Juan Esteban, lo describe con un ejemplo: si un día se pregunta ¿de dónde vienen los miembros de las Farc que están participando en las negociaciones de La Habana o qué tendencia está mayormente representada?, entonces empieza a construir una base de datos que incluya información sobre los estudios, la edad, participación política, región de origen, etc., de los negociadores; de este modo reúne suficiente información para analizarla y resolver su pregunta periodística. Las respuestas también llegan a partir de bases de datos obtenidas

mediante derechos de petición, cuya información, generalmente necesita ser complementada, o desde tablas de datos descargadas de Internet.

Un aspecto característico de este tipo de publicaciones en La Silla Vacía es que, en la mayoría de los casos, incluyen la base de desde la cual surgió la investigación. De hecho, a mediados de mayo de 2013 el medio lanzó su tercera remodelación que, además de un cambio de colores, un avance tecnológico y una reorganización de la página principal, incluye dos secciones nuevas: La Silla al Instante y Hágame el Cruce. En esta última se podrán encontrar todas las investigaciones que trabajan sobre bases de datos y acceder a esa información en bruto para hacer cruces, encontrar más hallazgos interesantes, etc.

El hecho de compartir grandes cantidades de datos recopilados y elaborados por un medio o cualquier entidad, a pesar de ser una tendencia mundial que cada vez crece más, suele generar dudas, pues detrás hay un enorme esfuerzo que podría ser valorado de mejor manera. Juanita confiesa que se ha preguntado lo mismo, ha pensado opciones para cobrar servicios informativos como el de Quién es Quién, cuando llegue a robustecerse tanto que valga la pena. Sin embargo, además del asunto administrativo que implica costos y dedicación, hay una consideración importante para la directora de este medio: en términos filosóficos La Silla Vacía está construida como una red abierta y comenzar a poner puertas también puede significar un precio alto.

Aunque Juan Esteban piensa que el periodismo de bases de datos que hacen es muy sencillo y que hay un nivel mucho más complejo, una especie de ‘Olimpo’ al que quisiera llegar algún día, lo cierto es que La Silla Vacía ya tiene un camino avanzado en cuanto a sus principios de transparencia y a la experiencia adquirida. Lo que viene es seguir avanzando en ese camino, por ejemplo, mediante la renovación de Quién es Quién y de otras bases de datos para lograr que todas tengan el mismo esquema interno, de modo que puedan ‘dialogar’, cruzar datos cuando sea necesario.

## 5. CONCLUSIONES

Hay una creciente tendencia a usar el análisis de bases de datos como técnica investigativa en los principales medios impresos y digitales en Colombia. Sin embargo, en la mayoría de los casos, la iniciativa proviene de periodistas que han aprendido el manejo de bases de datos, y no de los medios, por lo que el apoyo de estos aún carece de estructura y definición: no se observan equipos de trabajo claramente organizados para producir investigaciones de ese tipo, pocos medios pagan capacitaciones al respecto para sus periodistas y faltan ayudas como la adquisición de servicios informativos en línea.

La formación de reporteros en técnicas específicas de análisis de datos ha sido fundamental para ese surgimiento del periodismo de bases de datos en el país. Al respecto, la organización Consejo de Redacción ha aportado gran parte del entrenamiento recibido por quienes actualmente ejercen ese rol en los principales medios. De este modo, los periodistas en Colombia hacen uso de instrumentos básicos como la obtención de datos a través de Internet y de derechos de petición, y empiezan incursionar en herramientas más avanzadas como la construcción de bases de datos propias y el cruce de grandes cantidades de información.

El aporte del Estado en este proceso se ve un poco más relegado. Aunque cada vez son más las opciones de consulta y descarga de bases de datos oficiales en línea, los periodistas se siguen enfrentando con bloqueos como información dispersa, en formatos y estructuras que no permiten su reutilización, procesamiento y análisis.

En cuanto al caso específico de La Silla Vacía y *El Espectador*, la labor que realizan estos dos medios a partir del análisis de bases de datos, representa un aporte importante para el país en lo que tiene que ver con la fiscalización de los poderes políticos y económicos; particularmente la vigilancia del gasto público. La inclusión de esa técnica periodística en sus dinámicas de trabajo ha enriquecido el cubrimiento que hacen, principalmente de

asuntos políticos, permitiéndoles llegar a hallazgos relevantes sin necesidad de depender de fuentes oficiales o filtraciones; además, se ha visto fortalecida el área investigativa, especialmente en el medio impreso.

En ambos hay un solo periodista que convirtió la utilización de las bases de datos en su método natural de trabajo. En sus procedimientos se observa la aplicación de las técnicas de investigación de las ciencias sociales al periodismo y el recurso de una metodología completa que incluya el planteamiento de hipótesis –a veces con forma de pregunta periodística-, la recolección de datos, procesos de análisis, verificación y confrontación; lo cual confirma que tanto *El Espectador* como La Silla Vacía están dentro de los primeros medios que incursionan en el periodismo de bases de datos en el país.

Este último tiene un amplio camino avanzado en lo que se refiere al empleo de visualizaciones interactivas que facilitan la transmisión y comprensión de los hallazgos obtenidos durante indagaciones de este tipo; así mismo, tiene la delantera en la filosofía de transparencia del medio, gracias a la cual creó una nueva sección en su portal, llamada ‘Hágame el cruce’, en la que publica los insumos originales de los artículos que realiza a partir de bases de datos.

Entre las carencias que presentan los dos medios estudiados, están la ausencia de un sistema que reúna los grupos de datos recolectados, no sólo por el reportero que hace periodismo de bases de datos, sino por toda la redacción, a lo largo de su ejercicio periodístico; esta herramienta –que La Silla Vacía ya tiene en mente- les facilitaría investigaciones a mayor escala y enriquecería cada vez más la reportería digital de todos los periodistas. Hace falta, además, fortalecer la capacitación de varios miembros de los equipos de trabajo en técnicas de periodismo de bases de datos –como lo ha empezado a hacer Alexander Marín, de *El Espectador*-, lo cual sería un fuerte impulso para que ese mecanismo de indagación periodística avance y evolucione en ambos medios, elevando la calidad de su labor investigativa y el nivel de la vigilancia del poder en Colombia.

## 6. RECOMENDACIONES

- A los medios de comunicación colombianos que no lo han hecho, se les aconseja incursionar en el periodismo de bases de datos, pues, a pesar de los pequeños costos adicionales que pueda representar al comienzo, genera frutos que vale la pena aprovechar, como el fortalecimiento del área investigativa, la producción de mayores trabajos con independencia de las fuentes oficiales y el robustecimiento de la labor fiscalizadora del poder que deben realizar los medios.
- Para esto, los periodistas deben capacitarse constantemente en las técnicas de procesamiento de bases de datos y en labores afines a esa área, como el trabajo con equipos interdisciplinarios, la visualización de datos, el funcionamiento del Estado, la vigilancia del poder público, los mecanismos de acceso a la información, entre otros.
- Los medios deben ofrecer un apoyo más decidido a estas iniciativas, mediante la conformación de equipos de trabajo robustos que favorezcan la interdisciplinariedad –periodistas, ingenieros, diseñadores, analistas-, el acceso a las principales bases de datos pagas de consulta en línea y la disposición de recursos tecnológicos, además de la mencionada capacitación.
- Quienes ya aplican el periodismo de bases de datos, por ejemplo La Silla Vacía y *El Espectador*, pueden empezar a explorar áreas diversas de trabajo, más allá de la política y económica que, por supuesto, son importantes. Los temas deportivos, ambientales, de salud, sociales, incluso de entretenimiento, también son susceptibles de ser indagados desde la mencionada técnica, y generan conexiones significativas con las audiencias.
- Es importante que el programa de Datos Abiertos del Estado propicie una verdadera apertura de información relevante para el control social que, en parte, realiza la prensa. Para esto, se debe exigir a los entes oficiales que ofrezcan acceso en línea a bases de datos reutilizables, suficientes, actualizadas y de interés público.

## 7. REFERENCIAS

Aisch, Gregor; Alfter, Brigitte; Anderton, David; Ball, James; Barr, Caelainn; Berruezo, Mariana;... Wu, Chrys (2012). *Manual de Periodismo de Datos* (trad. ). Recuperado de <http://interactivos.lanacion.com.ar/manual-data/index.html>

Berganza Conde, María Rosa (2005). *Investigar en comunicación: Guía práctica de métodos y técnicas de investigación social en Comunicación*. Madrid: McGraw-Hill.

Botero, Catalina (30 de diciembre de 2009). *Importancia del derecho de acceso a la información* [respuesta en una entrevista del programa Sociedad Transparente]. Recuperado de [http://www.youtube.com/watch?feature=player\\_embedded&v=TgWIjs5TioQ#](http://www.youtube.com/watch?feature=player_embedded&v=TgWIjs5TioQ#)

Crucianelli, Sandra (23 de octubre de 2012a). ¿Qué es el periodismo de datos? [respuesta en una entrevista del portal Periodismo y Otras Yerbas]. Recuperado de <http://periodismoyotrasyerbas.blogspot.com/2012/10/sandra-crucianelli-introduccion-al.html>

Crucianelli, Sandra (2012b). *Periodismo de Datos: Lo bueno, lo malo y lo feo*. Recuperado de <http://blogs.lanacion.com.ar/data/periodismo-con-bases-de-datos/periodismo-de-datos-lo-malo-lo-bueno-y-lo-feo/>

Crucianelli, Sandra (2013). *12 puntos para entender el periodismo de datos*. Recuperado de <https://knightcenter.utexas.edu/es/blog/00-13717-12-puntos-para-entender-el-periodismo-de-datos>

Dader, José Luis (1997). *Periodismo de precisión la vía socioinformática de descubrir noticias*. Madrid, España: Síntesis.

Donadío, Alberto (2012). *La llave de la transparencia*. Bogotá, Colombia: Sílabas Editores.

Fundación para la Libertad de Prensa (2011). *Manual para el acceso a la información*. Bogotá: Autor.

Gobierno de Colombia. Catálogo de Datos Abiertos [portal web]. Recuperado de [www.datos.gov.co](http://www.datos.gov.co)

Gorriti, Gustavo (1999). Verdades tácticas y estratégicas. *Pulso del Periodismo*. Recuperado de <http://www.saladeprensa.org/art37.htm>

Herrán, María T. y Restrepo, Javier D. (1991). *Ética para periodistas*. Bogotá: Tercer Mundo Editores

Houston, Brant (2004). *Computer Assisted Reporting. A practical guide* (3ra edición). Boston: Bedford/St. Martin's.

Keenan, John (2010). *Why raw data sites need journalism*. Recuperado de <http://www.guardian.co.uk/commentisfree/libertycentral/2010/aug/11/raw-data-journalism-wikileaks>

Mariño, Ángeles (2012). *La relación del periodismo de datos con big data y open data*. (Tesis inédita de maestría). Universidad Carlos III de Madrid, España.

Meyer, Philip (1993). *Periodismo de precisión nuevas fronteras para la investigación periodística*. Barcelona: Bosch.

Open Government Partnership. Colombia [portal web]. Recuperado de <http://www.opengovpartnership.org/node/%2055>

Open Knowledge Foundation. *El manual de Open Data, 2010*. Recuperado de <http://opendatahandbook.org/es/index.html>

Proyecto Antonio Nariño. *Segunda encuesta nacional a periodistas sobre libertad de expresión y acceso a la información, 2013*. Recuperado de <http://pan.org.co/noticias.php?noticia=471>

Ramírez, Fernando (2012). *Periodismo de datos, periodismo de siempre*. Recuperado de <http://www.lapatria.com/blogs/periodismo-de-datos-periodismo-de-siempre>

Relatoría Especial para la Libertad de Expresión, Comisión Interamericana de Derechos Humanos, Organización de Estados Americanos. *El derecho de acceso a la información pública en las Américas. Estándares Interamericanos y comparación de marcos legales, 2012.* Washington, D.C. Recuperado de [http://www.oas.org/es/cidh/expresion/temas/acceso\\_informacion.asp](http://www.oas.org/es/cidh/expresion/temas/acceso_informacion.asp)

Santoro, Daniel. (2004). *Técnicas de investigación. Métodos desarrollados en diarios y revistas de América Latina.* México: Fondo de Cultura Económica : Fundación para un Periodismo Iberoamericano

Segnini, Giannina (2008). *Introducción al Periodismo Investigativo y al Periodismo Asistido por Computadora.* Agencia de los Estados Unidos para el Desarrollo Internacional – USAID. Ciudad de Guatemala: Litogres, S.A.

Segnini, Giannina (2011). *Giannina Segnini: "El periodismo de datos aumenta la credibilidad y la independencia"* [entrevista con la Red Global de Editores] (trad. ). Recuperado de <http://www.consejoderedaccion.org/component/content/article/48-rotador-3/98-giannina-segnini-periodismo-de-datos-aumenta-credibilidad>

Un paso importante (12 de mayo de 2013). *El Espectador.* Recuperado de <http://www.elespectador.com/opinion/editorial/articulo-421688-un-paso-importante>

Wimmer, Roger D. y Dominick, Joseph R. (2002). *Introducción a la investigación de medios masivos de comunicación.* México: Thomson Editores

## ANEXOS

### Encuesta aplicada en la fase 2

Por favor, indique el nombre de su medio

Su nombre:

Su cargo:

1. ¿Conoce el periodismo de bases de datos (o periodismo de datos)?

- Sí
- No

2. ¿Su medio tiene unidad investigativa?

- Sí
- No

3. ¿En la redacción del medio hay alguien que se encargue del análisis de bases de datos periodísticas?

- Sí, un periodista
- Sí, un equipo de periodistas
- Sí, la misma unidad investigativa
- Sí, un equipo de periodistas y otros expertos
- Sí, un experto en sistemas, no periodista
- No, nadie
- Otro:

4. Si el encargado de análisis de bases de datos es un equipo de trabajo, ¿cuántos son y cómo está conformado el grupo?

5. Indique la frecuencia con la que sus periodistas utilizan los siguientes recursos en sus investigaciones:

	Mucha frecuencia	Algunas veces	Rara vez	Nunca
Bases de datos de consulta en línea				
Bases de datos descargadas de internet				
Bases de datos obtenidas mediante derecho de acceso a la información pública				
Bases de datos elaboradas por los periodistas				
Búsqueda avanzada de información y documentos en Internet				

6. ¿Alguno de sus periodistas ha recibido capacitación sobre periodismo de datos?
  - Sí
  - No
7. ¿Cuántos periodistas han recibido formación en periodismo de datos?
8. Mencione algunas de las capacitaciones (tema e instructor) que han recibido los periodistas
9. ¿El medio ha pagado alguna de esas capacitaciones?
  - Sí
  - No
10. Si su medio ha realizado alguna investigación o reportaje periodístico que involucre el periodismo de datos, por favor indique acá algunos ejemplos.

### **Entrevistas realizadas en la fase 3**

#### Recursos humanos y tecnológicos

- ¿Desde cuándo existe esa iniciativa de realizar investigaciones a partir del análisis de bases de datos en el medio? ¿Cómo surgió la idea?
- ¿Quiénes trabajan en el área de análisis de bases de datos? ¿Cómo está conformado el equipo de trabajo?
- ¿Qué capacitación tienen? ¿Algún tipo de formación específica o conocimientos en estadística, matemática aplicada al periodismo, etc?
- ¿Interactúan con otros miembros de la redacción y con especialistas en otras áreas (ingenieros, diseñadores, estadísticos, abogados, etc.)?
- ¿Con qué equipos tecnológicos trabajan?
- ¿Qué programas utilizan para la depuración, análisis y visualización de datos? ¿Son gratuitos o pagos? ¿El medio paga esos programas cuando es necesario?
- ¿Están suscritos a algún tipo de base de datos en línea? ¿El medio paga esas suscripciones?

#### Agenda y dinámicas periodísticas

- ¿Cómo se define la agenda? ¿Quién propone los temas? ¿De dónde salen?
- Describa el proceso de elaboración de los trabajos que involucran análisis de bases de datos. ¿Quién depura los datos, quién los analiza?, ¿cómo?
- ¿Qué papel cumple su editor en este proceso de análisis de bases de datos?
- ¿Qué tanto tiempo les da el medio para realizar este tipo de trabajos?
- ¿Deben realizar, además, una producción diaria de artículos y noticias?

- ¿El medio le ha impedido trabajar o publicar determinado tema que requiera el análisis de bases de datos?
- ¿Dónde quedan almacenadas las bases de datos que se utilizan en cada uno de estos procesos? ¿El medio tiene alguna política o plan de construir sus propias bases de datos?

#### Acceso a la información

- ¿Cómo obtienen las datos que analizan?
- ¿Cómo ve el acceso a la información pública en Colombia?
- ¿Le han puesto obstáculos para entregarle información pública?, ¿de qué tipo? ¿qué hace en esos casos?
- ¿Ha accedido a otro tipo de conjuntos de datos que no necesariamente sean del sector público?, ¿cómo?

#### **Proyectos e iniciativas de periodismo de bases de datos mencionados en el documento**

- 156 candidatos a alcaldías tienen morosidad, sanciones o condenas, La Nación (Costa Rica)  
[www.nacion.com/2010-11-24/EIPais/NotaPrincipal/EIPais2585950.aspx](http://www.nacion.com/2010-11-24/EIPais/NotaPrincipal/EIPais2585950.aspx)  
 Descripción: [www.consejoderedaccion.org/component/content/article/48-rotador-3/98-giannina-segnini-periodismo-de-datos-aumenta-credibilidad](http://www.consejoderedaccion.org/component/content/article/48-rotador-3/98-giannina-segnini-periodismo-de-datos-aumenta-credibilidad)
- Cracking the codes, Public Integrity (Estados Unidos)  
[www.publicintegrity.org/health/medicare/cracking-codes](http://www.publicintegrity.org/health/medicare/cracking-codes)  
 Descripción: [www.publicintegrity.org/2012/09/06/10819/methodology-cracking-codes](http://www.publicintegrity.org/2012/09/06/10819/methodology-cracking-codes)
- Catálogo de Datos Abiertos (Colombia)  
[www.datos.gov.co](http://www.datos.gov.co)
- Cuentas Claras (Colombia)  
[www.cnecontasclaras.com](http://www.cnecontasclaras.com)
- Congreso Visible (Colombia)  
[www.congresovisible.org](http://www.congresovisible.org)
- Datablog, proyecto de The Guardian (Reino Unido)  
[www.guardian.co.uk/news/datablog](http://www.guardian.co.uk/news/datablog)
- Fundación Ciudadana Civio (España)  
[www.civio.es](http://www.civio.es)
- Gasto Público Bahiense (Argentina)  
[www.gastopublicobahiense.org](http://www.gastopublicobahiense.org)

- Gastos del Senado, La Nación (Argentina)  
[www.lanacion.com.ar/gastos-en-el-senado-t49163](http://www.lanacion.com.ar/gastos-en-el-senado-t49163)  
 Descripción: <http://blogs.lanacion.com.ar/ddj/data-driven-investigative-journalism/argentina-senate-expenses/>
- Nación Data, proyecto de La Nación (Argentina)  
<http://blogs.lanacion.com.ar/data/about/>
- Open Government Partnership (mundial)  
<http://www.opengovpartnership.org/>
- Opendata Latinoamerica (diversos países del continente)  
[www.opendatalatinoamerica.org](http://www.opendatalatinoamerica.org)
- Portal de Transparencia Económica (Colombia)  
[www.pte.gov.co](http://www.pte.gov.co)
- They Work for You (Reino Unido)  
[www.theyworkforyou.com](http://www.theyworkforyou.com)
- Unidad Investigativa de La Nación (Argentina)  
[www.nacion.com/2013-05-20/Investigacion.aspx](http://www.nacion.com/2013-05-20/Investigacion.aspx)  
 Descripción: [www.consejoderedaccion.org/component/content/article/48-rotador-3/98-giannina-segnini-periodismo-de-datos-aumenta-credibilidad](http://www.consejoderedaccion.org/component/content/article/48-rotador-3/98-giannina-segnini-periodismo-de-datos-aumenta-credibilidad)
- Zoom Online (Colombia)  
[www.zoomonline.info](http://www.zoomonline.info)

### **Enlaces a los artículos incluidos en el análisis de contenido**

- Los contraticos de la Cámara (El Espectador)  
<http://www.elespectador.com/impreso/temadeldia/articulo-363617-los-contraticos-de-camara>
- Siniestros en la Cámara (El Espectador)  
<http://www.elespectador.com/noticias/investigacion/articulo-375243-siniestros-camara>
- Gasolina a chorros (El Espectador)  
<http://www.elespectador.com/noticias/investigacion/articulo-381582-gasolina-chorros>
- Así está el mapa minero del país (El Espectador)  
<http://www.elespectador.com/noticias/infografia/articulo-414719-asi-esta-el-mapa-minero-del-pais>

- Los empresarios Uribe Moreno (El Espectador)  
<http://www.elespectador.com/noticias/economia/articulo-417371-los-empresarios-uribe-moreno>
- AH1N1, una gripe costosa para el erario (El Espectador)  
<http://www.elespectador.com/noticias/nacional/articulo-420159-ah1n1-una-gripe-costosa-el-erario>
- El top de los magistrados viajeros (La Silla Vacía)  
<http://www.lasillavacia.com/historia/el-top-de-los-magistrados-viajeros-30574>
- Las fichas para la reelección del Procurador (La Silla Vacía)  
<http://www.lasillavacia.com/historia/las-fichas-para-la-reeleccion-del-procurador-36860>
- Así son los que mandan en las empresas más grandes del país (La Silla Vacía)  
<http://www.lasillavacia.com/historia/asi-son-los-que-mandan-en-las-empresas-mas-grandes-del-pais-40582>
- El nuevo quién es quién del Banco de la República (La Silla Vacía)  
<http://www.lasillavacia.com/historia/el-nuevo-quien-es-quien-del-banco-de-la-republica-41582>
- El agua que pasa por debajo del descontento en la Unidad Nacional (La Silla Vacía)  
<http://www.lasillavacia.com/historia/el-agua-que-pasa-por-debajo-del-descontento-en-la-unidad-nacional-43343>
- Así va la ejecución del gobierno (La Silla Vacía)  
<http://www.lasillavacia.com/historia/as%C3%AD-va-la-ejecuci%C3%B3n-del-gobierno-44044>