

UNIVERSITAT
JAUME·I

PE0932 - TRABAJO FIN DEL GRADO

Línea A – Investigación

**Análisis electoral del 20-D a partir del Periodismo
de Datos realizado por *El Confidencial*, *El
Español* y *El Mundo***

Grado en Periodismo

Curso: 2015/2016

Marta Gozalbo Berenguer

29217444S

al260420@uji.es

Tutor: Hugo Doménech Fabregat

ÍNDICE

1. INTRODUCCIÓN	4
1.1. Objetivos.....	6
1.2. Hipótesis.....	6
2. MARCO TEÓRICO	7
2.1– Periodismo de Datos como nueva tendencia.....	7
2.2– Orígenes de una nueva disciplina: Movimientos de <i>Open Data</i> y <i>Open Government</i>	8
2.3- Periodismo de Datos no es periodismo con datos.....	12
2.4 - Fases, Metodologías y Herramientas.....	14
2.5 - Escasa formación y Ley de Transparencia en España.....	17
3) DISEÑO METODOLÓGICO	21
3.1- Metodología.....	21
4) ANÁLISIS Y RESULTADOS	24
4.1- Periodismo de datos en <i>El Confidencial</i>	24
4.2- Periodismo de datos en <i>El Español</i>	28
4.3 - Periodismo de datos en <i>El Mundo</i>	30
5) CONCLUSIONES	33
6) BIBLIOGRAFÍA	36
7) BIBLIOGRAFÍA WEB	37
8) ANEXOS	40

RESUMEN: El Periodismo de Datos surge como nueva disciplina que trae consigo nuevas rutinas periodísticas asociadas a las posibilidades que brindan las TIC y caracterizado por la presencia de visualizaciones en las piezas resultantes. Este trabajo se centra en el análisis de este nuevo modelo de periodismo y cómo se ha llevado a cabo en el período electoral en tres medios nacionales: *El Confidencial*, *El Español* y *El Mundo*. Junto al objetivo principal de analizar la cobertura informativa de los comicios a partir de este nuevo modelo de periodismo, se suman el de presentar las diferencias entre nativos e inmigrantes digitales y ofrecer una descripción de la situación de las Unidades de Datos en España. Tres objetivos importantes para conocer el perfil completo de trabajo de este nuevo modelo de prensa que, según diversos autores, ofrece actualmente numerosas oportunidades laborales. De este estudio se desprenden aspectos significativos como que el diario que realiza visualizaciones interactivas y navegables en todas sus piezas, tiene el servicio de programación externalizado, algo que no contemplan los otros dos medios. En lo referente a la cantidad de fuentes empleadas por los tres medios (una de las principales características del Periodismo de Datos es que se recurre a grandes bancos de datos como fuentes), resultan elevadas en los nativos digitales, pero no en el inmigrante digital. A pesar de que esta disciplina cada vez es más conocida y practicada por los profesionales del periodismo, las Unidades de Datos aún son un elemento incipiente en las redacciones españolas y la ausencia de una sección específica en los diarios destinada a su práctica, es una muestra de ello.

Palabras clave: Periodismo de Datos, *Open Data*, *Open Government*, *El Confidencial*, *El Español*, *El Mundo*

ABSTRACT: Data Journalism arises as a new discipline which contributes with new journalistic routines related to the possibilities that ICTs provide and that is characterized by the presence of visualizations in the resultant pieces. This academic paper focus on the analysis of this new model of journalism and how it's been accomplished in the election period in three national media: *El Confidencial*, *El Español* y *El Mundo*. Together with the prime goal of analyzing

the media coverage of the elections through this new model of journalism, it's added the goal of presenting the differences between digital and immigrant natives and offering a description of the Spanish Data Unit's situation. Three important goals to learn the complete profile of work of this new model of press that, in words of some authors, currently offers a great amount of employment opportunities. This study points to significant features such as that the journal that makes interactive and website visualizations in every single one of his pieces has as well an outsider programming service, something not contemplated in the other two media. Concerning to the quantity of sources used by the thrice of the media (one of the major attributes of Data Journalism is going to big data banks as sources) they are elevated on digital natives but not so on digital immigrants. Even when this discipline is increasingly well-known and practiced by journalists, Data Units are still an emerging element in the Spanish national editorial offices and the lack go a specific section in the journals intended for this method is an example.

Keywords: Data Journalism, *Open Data*, *Open Government*, *El Confidencial*, *El Español*, *El Mundo*

1) INTRODUCCIÓN

El periodismo ha desempeñado desde su aparición un papel fundamental en la sociedad, pues ofrece información a la ciudadanía relevante para el correcto funcionamiento de la democracia. En este sentido vigila las acciones de los políticos para asegurarse de que estos desempeñan sus obligaciones para con los ciudadanos de manera correcta y eficaz. De esta forma, ayuda a conformar una opinión pública a través de información que debe ser veraz y rigurosa.

Actualmente nos encontramos en un período de una profunda desafección ciudadana hacia los órganos representativos provocada por los numerosos casos de corrupción y decisiones políticas tomadas a raíz de la crisis económica –desahucios, rescates a bancos, etc.-.

Paralelamente a esta situación, comienza a desarrollarse, gracias a las TIC, un nuevo modelo de periodismo: el Periodismo de Datos. A través de éste, son muchos los periodistas que han realizado trabajos que han destapado casos de corrupción aprovechando datos públicos, una nueva disciplina periodística que se presenta como una gran oportunidad para dotar de calidad a la profesión y que permite un mayor desarrollo en los temas de agenda mediática.

Por otra parte, es el trabajo conjunto de periodistas y programadores lo que ha posibilitado que el periodismo haya dado un salto cualitativo y se estén desarrollando piezas basadas en el Big Data que profundizan en el *cómo* y *por qué* sobre asuntos de máximo interés para el ciudadano, monitorizando así a los centros de poder.

En España, el Periodismo de Datos ha empezado a tener presencia de forma considerable. Y tres son los periódicos que destacan por su trabajo que desarrollan en esta disciplina –*El Confidencial*, *El Español* y *El Mundo*–, aunque ninguno de ellos tiene una sección determinada dedicada a los datos, como sí la tienen desde hace años los medios punteros en Big Data Journalism (*The Guardian* y *The New York Times*), ya que el “Laboratorio” de *El Confidencial*, lo emplean tanto los servicios multimedia como los de datos para trabajos a largo plazo y no diarios.

Sin embargo, pese a que parece que comienza a tener una mayor presencia en los medios españoles, se percibe que siguen abundando fuentes oficiales y noticias de declaraciones frente a las de análisis profundos de datos que dan contexto a la actualidad.

Por ello, en este trabajo se realizará, en primer lugar, una aproximación a esta nueva disciplina: cuándo nace, cuáles son las metodologías y herramientas que utilizan, además de estudiar cómo se está desarrollando en España.

Para investigar cómo se ejerce en las redacciones españolas, se va a partir del análisis de un caso: las elecciones generales del 20-D. De esta forma, se pretende estudiar el tratamiento informativo concedido a un acontecimiento

relevante a nivel democrático como es la cobertura de una campaña electoral y el nivel de monitorización realizado entendido como el escrutinio que se hace sobre asuntos de interés público.

En ese espacio, se examinarán aspectos como el número de piezas, la profundidad de las mismas, el número de visualizaciones y variedad de las fuentes empleadas.

1.1 – Objetivos

El objetivo general de este trabajo es llevar a cabo un análisis exhaustivo de la cobertura informativa de las elecciones generales del 20 de diciembre a partir del Periodismo de Datos desarrollado por los diarios *El Confidencial*, *El Español* y *El Mundo*. Este objetivo general se concreta en los siguientes objetivos específicos.

OBJETIVO 1 = Comparar la cobertura y tratamiento informativo de las elecciones generales del 20-D en tres medios nacionales centrándonos, principalmente, en el uso de sus fuentes.

OBJETIVO 2 = Diseminar las principales diferencias existentes entre el trabajo realizado por nativos e inmigrantes digitales.

OBJETIVO 3 = Describir la situación de las Unidades de Datos en España y sus rutinas productivas en período electoral.

1.2 – Hipótesis

A la hora de determinar las hipótesis relativas a esta investigación, se han tenido en cuenta los objetivos descritos anteriormente, estableciendo una hipótesis como respuesta a cada uno de los objetivos específicos.

HIPÓTESIS 1 = Los análisis profundos de datos que parten o se apoyan en diversidad de fuentes y explican *qué ha ocurrido, por qué ha ocurrido y cómo puede afectar a la ciudadanía* desde diferentes prismas son escasos.

HIPÓTESIS 2 = Los nativos digitales generan mayor cantidad de piezas de datos y visualizaciones que los inmigrantes digitales.

HIPÓTESIS 3 = Los equipos de trabajo de datos son muy reducidos en las redacciones españolas, aspecto que afecta a las piezas resultantes.

2) MARCO TEÓRICO

2.1 – Periodismo de Datos como nueva tendencia

Aún a día de hoy encontramos pocas definiciones completas de qué es el Periodismo de Datos. No obstante, podemos hallar puntos en común que comparten diversos autores respecto a lo que esta nueva disciplina supone. Una nueva disciplina en la que la tecnología adopta un papel principal para el procesamiento de grandes cantidades de información o datos para la posterior narración de historias.

En este sentido, Mar Cabra, periodista miembro del Consorcio Internacional de Periodistas de Investigación determina que este tipo de periodismo se basa en “recopilar y analizar grandes cantidades de información y datos detallados para hacerlos comprensibles a la audiencia a través de artículos, visualizaciones o aplicaciones”¹.

Para Flores y Salinas, el Periodismo de Datos “consiste en usar herramientas estadísticas y de visualización para contar mejor las viejas historias y descubrir nuevas historias que contar” (2012:7).

¹ Cabra, M. (Febrero, 2016) De mayor quiero ser periodista de datos. *I Congreso Internacional sobre Periodismo de Datos. Nuevas narrativas para el Periodismo Especializado*. Congreso dirigido por el grupo In-Data de la Universidad de Málaga, UMA.

Uno de los grandes exponentes de esta nueva disciplina, Aron Pilhofer, considera que el periodismo de datos guarda similitudes con el periodismo tradicional ya que siempre “debe haber una historia y una intención periodística”, pero, como jefe de la sección de *Guardian Visual*, profundiza en la importancia que tiene la visualización en el Periodismo de Datos, pues “en la web tienes alrededor de siete segundos para cazar a un lector y meterlo en tu historia por lo que el diseño lo es todo” (Pilhofer en Gray, Chambers y Bounegru, 2011).

A todo esto, Sandra Crucianelli, periodista especializada en periodismo de investigación y precisión y docente en el Programa de Educación Virtual del Centro Knight para el periodismo en las Américas de la Universidad de Texas, añade que esta es una nueva modalidad en la que se incluye la “incorporación del programador al equipo periodístico” (Crucianelli, 2013:1).

El Periodismo de Datos por tanto, tiene tres características básicas que lo definen: 1) se basa en contar historias valiéndose de grandes volúmenes de datos e información; 2) la visualización interactiva es una parte importante en esta nueva disciplina; 3) en algunos equipos existen sinergias con otras profesiones como la de programador “para desarrollar tareas que van desde la extracción de datos y depuración de los sets hasta el diseño de las aplicaciones de noticias (new apps)” (Crucianelli, 2013:2).

2.2 – Orígenes de una nueva disciplina. Movimientos de Open Data y Open Government

Respecto a su origen y coincidiendo con Crucianelli, Karma Peiró y Javier Guallar, profesores de la Universitat Ramon Llull, puntualizan en su investigación *Introducció al periodisme de dades: Característiques, Panoràmica i exemples* que este nuevo modelo de periodismo guarda relación con otras especialidades como el periodismo de precisión (Peiró y Guallar, 2013). Un periodismo que para el profesor José Luis Dader, se basa en “desarrollar e

interpretar datos empíricos recolectados, relacionados y verificados mediante métodos científicos de investigación socioestadística o informática” (Dader, 1997:22).

Así pues, se considera a Philip Meyer como el padre de este tipo de periodismo, ya que fue él mismo quien, en 1959, tras cruzar grandes cantidades de datos, descubrió que los seguros escolares resultaban muy caros porque estaban financiando la campaña presidencial estadounidense. Asimismo, Simon Rogers, cita en su libro “Facts are Sacred: The Power of Data” a Meyer, quien define esta disciplina como la “aplicación de métodos científicos de investigación social y comportamental a la práctica del periodismo” (Rogers, 2011: 196).

No obstante, aunque de manera aislada, ya en 1821 encontramos una aproximación al Periodismo de Datos en un artículo publicado en *The Guardian*². Este artículo, mostraba a los lectores una tabla de los colegios de Manchester y Salford en la que comparaba cuántos alumnos había en cada escuela y cuál era su gasto anual. Dicho listado, sirvió para conocer cuántos alumnos recibieron educación gratuita además del índice de pobreza infantil de la ciudad. Permitted escrutar, a través de una filtración, que el número de menores que recibía educación gratuita era de 25.000 y no de 8.000 como aseguraban las cifras oficiales. De esta forma, la sociedad fue consciente de la situación real, en un principio desconocida: que la precariedad infantil triplicaba la cifra que se pensaba inicialmente.

Por otra parte, el Periodismo de Datos también guarda relación con el Periodismo Asistido por Computadora (CAR, en sus siglas anglosajonas). Según Ferreras, este tipo de periodismo es una evolución del Periodismo de Precisión en el que se emplean ordenadores para, a partir de análisis de grandes bases de datos, encontrar “correlaciones estadísticas [...] en todo tipo de documentos informatizados” (Ferreras, 2013:120).

² The Guardian (26 de septiembre de 2011) The first Guardian data journalism: May 5, 1821. Datablog. Recuperado de <http://www.theguardian.com/news/datablog/2011/sep/26/data-journalism-guardian>

Pero Peiró y Guallar también asocian esta disciplina a prácticas periodísticas tradicionales como la profunda documentación y, otras más novedosas, como el *fact checking* o los movimientos *Open Data*. Peiró y Guallar (2013:27) definen el *fact checking* como una especialidad del periodismo “relacionada con el uso de datos para la comprobación de hechos y datos. Poco conocido en nuestro país [...] pero que desde hace años triunfa en Estados Unidos y América Latina. Basado en la comprobación y verificación de informaciones”. Un ejemplo en España se puede observar en el programa El Objetivo de Ana Pastor en La Sexta. Pero, qué es el *Open Data*.

Para hablar de *Open Data*, debemos mencionar en primer lugar qué es el movimiento *Open Government*. Este término y movimiento tiene origen a finales de los años 70 en el panorama político británico. Donde hablar de Gobierno Abierto significaba hablar del “acceso y libertad de información, protección de datos, reforma de las leyes sobre los secretos oficiales y la necesidad, en una democracia sana, de poder obtener información sobre actividades del Gobierno y sobre su disponibilidad para la opinión pública y el ciudadano común” (Chapman y Hunt, 1987:11).

Con este mismo objetivo, unos años más tarde, en el año 2011, se crea la Alianza para el Gobierno Abierto para “proveer una plataforma internacional para reformadores domésticos comprometidos a que sus gobiernos rindan cuentas, sean más abiertos y mejoraren su capacidad de respuesta hacia sus ciudadanos”³. Con todo esto, estipulan que los fundamentos principales de un Gobierno Abierto se centran en: “mayor transparencia, rendición de cuentas y mejora de la capacidad de respuesta hacia sus ciudadanos con el objetivo de mejorar la calidad del gobierno, así como la calidad de los servicios que reciben los ciudadanos”⁴.

Según la profesora de Documentación e Información de la UOC, Montserrat García Alsina, para apoyar esos tres principios, es necesario que se den tres

³ Alianza para el Gobierno Abierto (2011) *Participating Countries*. Recuperado de <http://www.opengovpartnership.org/es/countries>

⁴ Alianza para el Gobierno Abierto (2011) *Misión y Objetivos*. Recuperado de <http://www.opengovpartnership.org/es/acerca-de/misi%C3%B3n-y-objetivos>

circunstancias: acceso y suministro de información, accesibilidad y datos abiertos; siendo el primero de ellos de especial importancia ya que “los gobiernos y las administraciones públicas tienen que garantizar a los ciudadanos el acceso a la información (económica, jurídica u organizativa) generada en el transcurso de sus actividades”⁵. En definitiva, un Gobierno Abierto “comporta abrir los datos gubernamentales a la ciudadanía, lo que implica gestionar una gran cantidad de datos para la toma de decisiones, y facilitar la colaboración y la participación ciudadana”⁶.

Así pues, como uno de los ejes comunes de los principios de Transparencia, Colaboración y Participación de los gobiernos abiertos, el manual de la *Open Knowledge Foundation* (OKF, 2012), determina que la apertura de datos u *Open Data* hace referencia a que los datos pueden ser utilizados libremente, reutilizados y redistribuidos por cualquier persona. Las claves de esta apertura de datos son (OKF, 2012):

- Disponibilidad y acceso. Los datos deben estar disponibles en su totalidad y con un precio razonable de reproducción, preferiblemente mediante la descarga a través de Internet. Además de permitir su modificación.
- Reutilización y redistribución en condiciones óptimas incluyendo el intercambio y combinación con otras bases de datos.
- Participación universal. Todos deben ser capaces de utilizarlos, reutilizarlos y redistribuirlos. No deben existir restricciones contra empresas, grupos o personas.

Pudiendo desarrollarse en diversos ámbitos -cultural, científico, financiero, medioambiental- desde OKF contemplan 3 razones por las cuáles los datos deben ser abiertos: “Transparencia, Liberalización de valor social y comercial y Participación y compromiso”⁷.

⁵ Alsina, Montserrat (2014). Open Government, open data, big data y transparencia: la información como nexo de unión. COMeIN, 39. Recuperado de <http://www.uoc.edu/divulgacio/comein/es/numero39/articulos/Article-Montserrat-Garcia-Alsina.html>

⁶ Ibid.

⁷ Open Knowledge Foundation (2012) *What is Open?* Recuperado de <https://okfn.org/opendata/>

Así pues, las semejanzas que Peiró y Guallar estipulan que tiene Periodismo de Datos y el *Open Data* o “datos en abierto” son referidos “a la capacidad de capturar, tratar, modificar, reutilizar y difundir datos. En definitiva, gestionar datos sin restricciones”. Investigar bases de datos de naturaleza pública que gestionan las administraciones con el objetivo de hacer esos datos comprensibles a los ciudadanos y garantizar la rendición de cuentas. No obstante, la práctica del periodismo de datos, como comentábamos unas líneas más arriba, no es en absoluto una práctica sencilla, pues, lejos de todas las herramientas que se deben emplear la manera de conseguir la materia prima en ocasiones puede resultar dificultosa o llevar un tiempo mucho mayor al habitual, pues no siempre las bases de datos se encuentran abiertas o su cifrado no resulta sencillo de interpretar.

2.3 – Periodismo de Datos no es periodismo con datos

En cualquier caso, lo que es indiscutible, es que la principal fuente de trabajo para esta nueva disciplina son los datos. Y respecto a ellos, Paul Bradshaw, profesor de la Birmingham City University apunta en *The Data Journalism Handbook* que “pueden ser la fuente del periodismo de datos, o pueden ser la herramienta con la que se narra la historia o ambas cosas. Como cualquier fuente, debe tratarse con escepticismo; y como cualquier herramienta debemos ser conscientes de cómo puede modelar y limitar las historias que se crean con las misma” (Bradshaw en Gray et al., 2011).

Por ello, Philip Meyer, señala que en una época de abundancia informativa, resulta de especial importancia el procesamiento de los datos, dividiendo éste en dos niveles: “1) análisis para encontrar sentido y estructura en el flujo sin fin de datos y 2) presentación de esa información para meter lo que es importante y relevante en la cabeza del consumidor” (Meyer en Gray et al., 2011).

Y a esto también se refiere Antonio Delgado, jefe de la Unidad de Datos del periódico *El Español*, cuando dice que por ello es esencial “la función del

periodista para que busque patrones y tendencias –entre todos esos datos–⁸ y así se puedan traducir para contar historias. Pues como considera Eva Belmonte (2015), de la Fundación Ciudadana Civio, “los datos en sí mismos no dicen nada si no son tratados”⁹. Y he aquí la primera particularidad del periodismo de datos: “saber hacer preguntas adecuadas a los datos para hallar historias que contar”¹⁰.

Así pues, se advierte otra singularidad del Periodismo de Datos: que el periodista sepa trabajar con grandes cantidades de información para optimizar el tiempo. Una tarea realizable gracias a la evolución tecnológica, que ha permitido la digitalización de todo lo que sucede en el mundo y desarrollar programas informáticos o *software* que posibilitan al periodista procesar toda la base de datos y extraer los más significativos o relevantes en un espacio de tiempo relativamente breve.

A este respecto se refiere Liliana Bounegru cuando asegura que “el surgimiento de la etiqueta *Periodismo de Datos* al comienzo de siglo indica una nueva fase en la que el mero volumen de los datos libremente disponibles online –combinado con herramientas sofisticadas centradas en el usuario, la autoedición y las herramientas de colaboración abierta (*crowdsourcing*)- permite a más gente trabajar con más datos de modo más fácil que nunca” (Bounegru en Gray et al., 2011).

Para el co-fundador de *Datawrapper*¹¹, Mirko Lorenz (2011), resulta indispensable que el periodista trabaje con datos y modifique sus rutinas productivas habituales para “aportar una visión más profunda de lo que sucede a nuestro alrededor y cómo puede afectarnos”, además, argumenta que utilizar

⁸ Delgado, A. (2014) *Periodismo de datos. Una introducción a esta disciplina del periodismo*. Recuperado en <http://mugi.webs.upv.es/wp-content/uploads/2014/10/periodismo-datos-valencia-adelgado.pdf>

⁹ Belmonte, E. (2015, julio 28) Cómo hacer periodismo de datos con el ejemplo de Medicamentalia [Archivo de vídeo] Recuperado en <http://escuela.civio.es/como-hacer-periodismo-de-datos-con-el-ejemplo-de-medicamentalia/>

¹⁰ Delgado, A. (Febrero, 2016) Periodismo de Datos en España: Más y mejor periodismo. En A. Teruel (Organizador/Presidencia), *I Congreso Internacional sobre Periodismo de Datos. Nuevas narrativas para el Periodismo Especializado*. Congreso dirigido por el grupo In-Data de la Universidad de Málaga, UMA.

¹¹ Herramienta online para crear gráficos y mapas. Disponible en: <https://datawrapper.de/>

los datos permite transformar “algo abstracto en algo que todos pueden entender y con lo que pueden relacionarse”. Es decir, trabajar con datos puede facilitar al periodista “revelar cómo una amenaza abstracta (como el paro) – por ejemplo- afecta a la gente de acuerdo a su edad, género o nivel de educación” (Lorenz en Gray et al., 2011).

2.4 – Fases, Metodologías y Herramientas

Conocer las distintas fases que se llevan a cabo en el Periodismo de Datos, posibilita que comprendamos cómo se transforman las rutinas en esta disciplina dentro de la misma profesión. Cambian aspectos como las fuentes, ya que ahora son los datos abiertos y las iniciativas ciudadanas los que constituyen la materia prima para la elaboración de las piezas de este nuevo modelo de periodismo. También cambian la forma de confeccionar las piezas y difundirlas, ya que las visualizaciones, suponen una irrupción en la tipología de géneros informativos.

Una de las características principales de esta nueva disciplina es que tiene unas fases muy diferenciadas. El profesor Paul Bradshaw (2011) define en su pirámide invertida, las cinco etapas del Periodismo de Datos: Compilar, limpiar, contextualizar, combinar y finalmente, comunicar o narrar una historia a partir de los datos relevantes.

1. Sobre la primera fase –**compilación**–, Antonio Delgado, asegura que se pueden obtener por dos vías: datos abiertos o filtraciones¹². Y vincula el primero al movimiento *Open Government*, que comentábamos anteriormente, y que muchos ayuntamientos han empezado a practicar publicando datos que generan, sobre todo desde el punto de vista económico, como es el caso *Datashare* del Ayuntamiento de Madrid. En esta primera etapa, es habitual que los periodistas utilicen la técnica de *scraping* para obtener datos de las páginas web de manera estructurada cuando no están disponibles a través de descarga directa o no es posible acceder a ellos. La ejecución de esta tarea se realiza a través de

¹² Ibid.

programas como *Daper* o *import.io* (si se desea evitar la programación) y *Python* (con programación).

Es común que el periodista se encuentre con archivos en diversidad de formatos: CSV, Excel, BBD, SQL, HTML o PDF, siendo el primero de ellos el preferido por los profesionales y el último el menos recomendable por comportar una mayor cantidad de trabajo para organizar la base de datos.

La razón por la que el PDF no es un buen documento con el que trabajar se justifica en que, al tratarse de una imagen, los periodistas deben emplear distintos programas para '*hackear*' el archivo y poder organizar los datos en tablas que les permitan tanto a ellos como a los programadores, hacer preguntas al documento. Así, entre dichos programas encontramos algunos de pago como *Abbyy FineReader*, *Adobe Acrobat Pro* o *Nitro PDF* y otros gratuitos como *PDF to Excel online*, *CometDocs*, *Zamzar* o, el más empleado, *Tabula*.

2. En cuanto la fase de compilación finaliza, se debe poner en marcha un proceso de **limpieza**, ya que aunque hayamos rescatado estos datos, no significa que estén listos para utilizarse. A menudo las hojas de cálculo de Excel con las que trabajan los periodistas "deben ser depuradas, porque hay columnas extrañas, celdas fusionadas que lo único que hacen es no ayudar a comprender la información" (Rogers en Gray et al., 2011). Por ello, se debe limpiar los documentos de manera que la información se presente de manera ordenada. Tarea para la cual resulta muy efectivo el uso de la herramienta *Open Refine*¹³, con la que se realiza una estandarización de los datos de forma rápida y efectiva.
3. Para **contextualizar** esos datos, "hay que insertar el dato material en un marco de referencia mental que ayude a su comprensibilidad" (Meyer, 1993:33). Como muchos de estos datos en ocasiones vienen cifrados con un código oficial, -cada municipio, hospital, país o escuela, tiene su propio código identificativo- la fase anterior (limpieza) no sólo permite

¹³ Herramienta de descarga gratuita ofrecida por Google. Disponible en: <http://openrefine.org/download.html>

eliminar celdas o datos no relacionados con la historia, sino que invita a que los profesionales de la información “puedan comparar cosas comparables” (Rogers en Gray et al., 2011), por tanto, da también paso a la cuarta fase: la combinación o comparación de datos.

4. En esta etapa de **combinación**, es en la que Antonio Delgado, dice, debemos encontrar patrones y tendencias que nos ayuden a contar una historia. Datos que cobren sentido al juntarse con otros ¹⁴. Para relacionar datos y extraer conclusiones, podemos realizar cálculos en las variantes con programas como *Tableau Public*, *Numbers* o *R*.
5. Finalmente, en el período de **comunicación** se debe seleccionar cuál será el producto más apropiado para publicar: texto, gráficos, visualizaciones interactivas, vídeos... Ya que el Periodismo de Datos “no son gráficos y visualizaciones. Sino que es contar una historia de la mejor manera posible” (Pilhofer en Gray et al., 2011). Y también se decidirá el enfoque que tendrá el producto periodístico. En función del producto que elijamos que tendrá nuestra pieza de datos, podremos utilizar herramientas específicas para una visualización más efectiva: *DataWrapper*, *Infogr.am*, *Silk*, *Google Fusion Tables*, *StoryMap.js*, *Timeline.js*, y más.

Para Bradshaw (2011), existen seis maneras diferentes para llevar a cabo esta última fase:

- La *visualización* como la manera más rápida para comunicar los resultados. No es sencilla, pero puede hacer que la comunicación sea muy efectiva.
- La *narración* basada en el artículo tradicional.
- La *comunicación social* valiéndose de iniciativas como el *crowdsourcing*.

¹⁴ Ibid, p.13.

- *Humanizar* los datos con el fin de enganchar al espectador. Por ejemplo, grabar una vida afectada por dichos datos.
- *Personalizar* los datos para que cada usuario pueda aportar información de manera que se conozca una situación general a partir de diversas situaciones particulares.
- Y por último, *utilizar* los datos para generar aplicaciones o herramientas basadas en la información. Para el periodista, la manera más compleja de comunicar datos.

2.5- Escasa formación y Ley de Transparencia en España

Para comenzar a definir cómo se introduce el Periodismo de Datos en España, se debe remontar al lector de esta investigación a diciembre de 2011, año en que David Cabo y Jacobo Elousa deciden poner en marcha la *Fundación Ciudadana Civio*. Motivados por el trabajo que hacen organismos como *MySociety* en Reino Unido, *Ciudadano Inteligente* en Chile o *Sunlight Foundation* en Estados Unidos, se disponen a poner en marcha una organización que garantice ciertos derechos civiles. Así, *Civio* surge con el fin de “ayudar a mejorar la transparencia -en España-, exigir rendición de cuentas a las administraciones y promover el libre acceso a los datos públicos”¹⁵.

Desde entonces y hasta ahora, con la colaboración de programadores y periodistas, la fundación ha publicado proyectos como *¿Dónde van mis impuestos?*, *Tuderechoasaber.es* (junto a *Access Info Europe*), *El Indultómetro* o *El BOE nuestro de cada día* haciéndose valer de los datos obtenidos a través de peticiones de información a las administraciones.

Por otra parte, se desarrollan proyectos paralelos que van en la misma dirección que persigue *Civio*. En octubre de 2012, se constituye oficialmente la delegación española de la *Open Knowledge Foundation*, (*OKFN-Spain* en España). A partir de su creación, comienzan a promover actividades de las que se desprenden la importancia de la existencia de un organismo así (apertura de

¹⁵ Fundación Ciudadana Civio (2012) Breve historia de Civio. Recuperado de <http://www.civio.es/breve-historia-de-civio/>

datos, reutilización, libre decisión, etc.) y convocan siete meses después de su nacimiento las *I Jornadas de Periodismo de Datos y Open Data* en Barcelona, Madrid, Donostia y Sevilla simultáneamente.

Sin embargo, aún a día de hoy la práctica del Periodismo de Datos en España es muy escasa. Quizá una de las causas podamos hallarla, como desarrolla la profesora de la universidad Complutense de Madrid, Pilar José López, en su doctorado *Antecedentes y fundamentos del periodismo de datos* en el retraso de los planes de estudio en las universidades españolas respecto a los nuevos perfiles de redactores que se demandan actualmente. Pues “la no inclusión de asignaturas como programación básica, deja a los futuros redactores españoles en una clara desventaja”¹⁶. Para formarse en Periodismo de Datos en España se puede optar por cursar el “Máster en Periodismo de Investigación, Datos y Visualización” que ofrece *El Mundo* de manera conjunta con la Universidad Rey Juan Carlos de Madrid y en el que colabora *Google*, o realizar el Máster en la Universidad de Deusto de “Experto en Análisis, Investigación y Comunicación de Datos”.

Mientras, facultades prestigiosas del mundo como la Universidad de Columbia de Nueva York ya incluyen el Periodismo de Datos y el tratamiento del Big Data en sus planes de estudio. Tanto es así, que incluso 5 alumnos de esta propia universidad han participado en el reciente proyecto *#PanamaPapers*¹⁷ procesando parte de esos 11 millones y medio de documentos internos de Mossack Fonseca y colaborando con el *Consortio Internacional de Periodismo de Investigación* a destapar todo el entramado existente alrededor de las entidades ‘offshore’.

Aparte de esta formación oficial en España, son numerosos los congresos que se realizan alrededor de la península sobre esta nueva disciplina. En este sentido, *Medialab Prado* –un programa del área de las artes del Ayuntamiento

¹⁶ López, P. J. (Febrero, 2016) *Antecedentes y fundamentos del periodismo de datos. I Congreso Internacional sobre Periodismo de Datos. Nuevas narrativas para el Periodismo Especializado*. Congreso dirigido por el grupo In-Data de la Universidad de Málaga, UMA.

¹⁷ Columbia Journalism School (2016) *Columbia Journalism Partners on Panama Papers*. Recuperado de <http://www.journalism.columbia.edu/page/1253-columbia-journalism-partners-on-panama-papers/748>

de Madrid– ofrece un “programa de actividades compuesto por talleres de producción y de formación, seminarios y debates, reuniones de diferentes grupos de trabajo, muestras de proyectos, conferencias”¹⁸. En este marco, han realizado jornadas que giran en torno al Periodismo de Datos como el “III Taller de producción de periodismo de datos” o “Visualizar’15”.

Para Montse Quesada y Cristina Ribes, docentes en la Universidad Pompeu Fabra, la situación actual para el desarrollo de esta disciplina es complicada en España no solo por la formación, sino porque los medios de comunicación españoles “no han creado secciones específicas de *Open Data* que trabajen a pleno rendimiento ni han apostado por formar a sus redactores en las nuevas herramientas digitales. En lugar de ello, son pequeños equipos de profesionales independientes que nutren con sus proyectos a la mayoría de los medios clásicos”¹⁹.

Pero más allá de la escasa oferta formativa existente, acertamos otras posibles razones por las que el Periodismo de Datos tiene un menor desarrollo en España. Hasta diciembre de 2013, no se aprueba la *Ley de la Transparencia, Acceso a la Información Pública y Buen Gobierno*. Esta ley tiene como objeto incrementar y reforzar...

...la transparencia en la actividad pública –que se articula a través de obligaciones de publicidad activa para todas las Administraciones y entidades públicas–, reconoce y garantiza el acceso a la información – regulado como un derecho de amplio ámbito subjetivo y objetivo– y establece las obligaciones de buen gobierno que deben cumplir los responsables públicos así como las consecuencias jurídicas derivadas de su incumplimiento –lo que se convierte en una exigencia de responsabilidad para todos los que desarrollan actividades de relevancia pública-²⁰.

¹⁸ Medialab Prado (2007) *Qué es Medialab Prado*. Recuperado de http://medialab-prado.es/article/que_es

¹⁹ Quesada, M. & Ribes, C. (Febrero, 2016) *El Periodismo de datos en Catalunya. I Congreso Internacional sobre Periodismo de Datos. Nuevas narrativas para el Periodismo Especializado*. Congreso dirigido por el grupo In-Data de la Universidad de Málaga, UMA.

²⁰ Boletín Oficial del Estado (BOE) del 9 de diciembre de 2013, número 295, página 1. Disponible en: <https://www.boe.es/buscar/pdf/2013/BOE-A-2013-12887-consolidado.pdf>

Sin esta ley, los periodistas no podían acceder anteriormente de forma libre a los datos que manejan las administraciones públicas, por lo que el desarrollo de esta disciplina era muy escaso, pues todos los proyectos debían elaborarse partiendo de peticiones oficiales a la administración.

Sin embargo, a pesar de su aprobación, organismos como *Access Info Europe*, que se encargan de promover la protección del derecho al acceso a la información en Europa, indica que esta *Ley de Transparencia* no cumple con los requisitos mínimos internacionales. Victoria Anderica, coordinadora del organismo, describe que, si se hubiera considerado un derecho fundamental amparado por los artículos 20 de libertad de información y 23 de libre participación política, y no enmarcarse bajo el artículo 105 b) que garantiza el acceso de los ciudadanos a los archivos y registros administrativos, “se tendría un ámbito de aplicación más amplio y unas garantías de protección ante los tribunales más reforzadas”²¹.

A pesar de todos estos hándicaps para ejercer el Periodismo de Datos en España y que las leyes nacionales no cumplan estándares o requisitos que el Tribunal Europeo de Derechos Humanos, las Naciones Unidas o la Corte Interamericana de Derechos Humanos, reconocen como derechos fundamentales, la práctica del Periodismo de Datos en nuestro país es cada vez mayor. El Informe Anual de la Profesión Periodística que publica todos los años la Asociación de Prensa de Madrid²², indicaba en 2014 que este nuevo modelo de periodismo está en auge y cada vez es mayor el número de profesionales españoles que la reconocen o practican. El 69,3% de los periodistas asegura conocerlo y el 7% trabaja en ello, quince puntos y un punto por encima, respectivamente, que el año anterior.

²¹ Zoom News (2013) Los seis agujeros de la Ley de Transparencia. Recuperado de http://www.zoomnews.es/150714/actualidad/espana/seis-agujeros-ley-transparencia?cx_source=carrusel

²² Asociación de Prensa de Madrid (2014) Informe Anual de la Profesión Periodística. Recuperado de http://www.apmadrid.es/images/stories/Informe%20profesion_2014_def_baja.pdf

3) DISEÑO METODOLÓGICO

3.1 – Metodología

La metodología a aplicar en el presente trabajo de investigación partirá del estudio de tres periódicos: *El Confidencial*, *El Español* y *El Mundo*. Tras la comprobación de que los tres diarios publican con frecuencia piezas de datos basadas en el análisis de grandes cantidades de información, se considera oportuno sustentar el trabajo en el estudio de dos nativos digitales y un inmigrante digital.

Tras la aproximación académica inicial del presente ensayo se ha podido definir qué es el Periodismo de Datos, cuáles son sus antecedentes, qué le diferencia del periodismo con datos y cuáles son las fases de trabajo de esta disciplina. Posteriormente, se ha descrito y detallado su situación en España.

A continuación se han planteado tres hipótesis principales. Para poder resolverlas empíricamente, se ha seleccionado un período determinado para analizar las piezas de Periodismo de Datos que han publicado los tres diarios previamente mencionados. De esta forma, la muestra de este estudio queda conformada por los trabajos de esta disciplina publicados entre el 15 y 25 de diciembre en estos tres medios españoles, ya que durante los días previos a los comicios y los días posteriores, se genera gran cantidad de datos y material de carácter público (CIS) y privado (diversidad de empresas sociológicas) que permite tanto su análisis como su representación en visualizaciones.

Para resolver las dos primeras hipótesis se realizará un análisis cuantitativo en el que, a través de la siguiente rejilla se estudien las sucesivas variables:

		El Confidencial	El Español	El Mundo
Nº de piezas				
Nº de visualizaciones	Estáticas			
	Interactivas			
Posee	piezas formadas			

únicamente por visualizaciones			
Nº de piezas basadas en simple plasmación de resultados			
Nº máximo de fuentes utilizadas			
Nº mínimo de fuentes utilizadas			
Utiliza únicamente sondeos y encuestas como fuentes			
Utiliza fuentes de distinta naturaleza			
Variedad temática			

Elaboración propia.

A través del estudio de estas variables, se podrá conocer si diversas cuestiones: volumen de trabajo en 10 días, número de visualizaciones confeccionadas y el carácter estático o interactivo de las mismas, pues las interactivas requieren de un mayor trabajo de programación que las estáticas y esto permite conocer lo que cada medio apuesta por esta tarea y, en su defecto, por la incorporación del programador al equipo periodístico. Además, podremos conocer qué cantidad de piezas se presentan sin texto que las acompañe (algo a lo que periodistas como Pilhofer son contrarios por considerar que todo trabajo visual debe ir acompañado de texto que lo contextualice).

Por otra parte, se expresará cuántas de las piezas desarrolladas por los medios se basan en una simple plasmación de resultados sin analizar o interpretarlos. Conocer este dato resulta significativo ya que el Periodismo de Datos tiene una clara finalidad analítica.

Por último, y respecto al uso de fuentes, esta rejilla nos admite averiguar cuales son las piezas en las que se utilizan mayor y menor número de fuentes respectivamente, de modo que comprendamos si existe un equilibrio entre todas las piezas, si hay algún desequilibrio entre los tres medios, así como si los medios apoyan sus análisis electorales en fuentes documentales (bases de

datos) o también se ayudan de voces expertas que les ayuden a interpretar la información obtenida o voces implicadas en los procesos participativos (votantes o candidatos políticos), de manera que profundicen más en sus trabajos.

Por otra parte, se estudiará de manera paralela la forma de comunicación de Bradshaw que predomina en cada medio por hallarse presente en la totalidad de sus piezas.

Se pretende así conseguir alcanzar el primer objetivo, conocer cómo se ha realizado la cobertura informativa de las elecciones a partir del Periodismo de Datos en los tres medios y poder realizar una comparativa entre los mismos.

Para resolver la tercera hipótesis –que las Unidades de Datos españolas son muy reducidas- se opta por un análisis cualitativo. Se utilizará la entrevista en profundidad a los jefes de las unidades correspondientes vía e-mail como herramienta. El canal que se elige para la ejecución de dichas entrevistas queda justificado por la imposibilidad de hacerlo de manera presencial y el escaso tiempo del que se dispone.

La batería de preguntas que formarán dichas entrevistas en profundidad serán las siguientes:

- ¿Quiénes forman la Unidad de Datos de El Confidencial?
- ¿Qué formación tienen?
- ¿Cuenta el medio con programadores que realizan las visualizaciones?
¿O es el equipo Multimedia quien las confecciona?
- ¿La Unidad de Datos se encuentra integrada en alguna sección del resto de la redacción o es un departamento propio?
- En período de elecciones, ¿se trabaja de manera conjunta con el resto de la redacción que no trabaja con datos?
- ¿Varían las rutinas productivas en este período? ¿Cómo?
- ¿Qué fuentes se suelen consultar? ¿Se realizan peticiones de información?

- ¿Qué criterios determinan que la Unidad de Datos trabaje unas noticias y el resto de secciones otras?
- ¿Qué grado de protagonismo tienen las visualizaciones en las piezas de datos?

Con las respuestas obtenidas se tratará de conocer la situación de las unidades de datos dentro de las redacciones españolas y sus rutinas en el período electoral.

4) ANÁLISIS Y RESULTADOS

4.1 – Periodismo de Datos en *El Confidencial*

La Unidad de Datos de *El Confidencial* es de las más consolidadas en la prensa nacional. Para hallar las piezas de Periodismo de Datos que el equipo ha realizado entre los días señalados se ha optado por diversas vías, ya que en la sección de “Laboratorio” (*ElConfidencial.LAB*) sólo se pueden encontrar los trabajos más amplios realizados a largo plazo. Así pues, se han buscado las piezas por *tags*, pero también por el nombre de los periodistas y equipo técnico que trabaja esta disciplina (Daniele Grasso, Jesús Escudero, Adrián Blanco y Antonio Esquembre) y a través de la búsqueda avanzada de Twitter que nos ha permitido restringir el período establecido.

Con todo esto, se han hallado un total de 8 piezas de Periodismo de Datos abarcando distintos ítems de interés público: atención de los políticos a las preocupaciones ciudadanas a través de Twitter –estudio del nivel de compromiso y análisis de cómo varía la comunicación política-, test de afinidad ideológica a partir de análisis de programas electorales, análisis de los discursos de los políticos en los diferentes debates a través de la técnica *fact checking* señalando en qué han mentido o expresado imprecisiones y verdades, estudio de resultados electorales, un juego interactivo de pactos, análisis de una posible Ley de Circunscripción Única y crítica a la actual Ley Electoral, búsqueda de perfiles de votantes y, por último, un reportaje que

aborda distintas cuestiones que se desprenden de los resultados de los comicios.

Entre todas las piezas, no encontramos ninguna que posea gráficos sin texto o textos sin gráficos; así pues, hay un equilibrio entre los aspectos formales y de contenido. Entre todas los artículos, sólo dos no cuentan con visualizaciones interactivas, ya que el resto todas incluyen al menos una, sumando así: 11 visualizaciones interactivas a partir de gráficos de barras, mapas de calor e interactivos más desarrollados; y 21 visualizaciones estáticas formadas por gráficos de barras, quesos/hemiciclo, esquemas, mapas de calor o gráficos lineales.

Por otra parte, *El Confidencial* ha empleado hasta 16 fuentes distintas en uno de los trabajos. Así pues, para llevar a cabo una pieza de *fact check*, analizan un total de 51 afirmaciones y los periodistas deben recurrir a distintas fuentes: CIS, métricas de Twitter, Eurostat, EPA, Banco Mundial, PIB, INE, datos abiertos de los distintos Ministerios... Además de fuentes expertas como politólogos o periodistas especializados. No obstante, hay piezas que realizan un análisis a partir de una sola fuente.

Respecto a la forma de comunicación, observamos una predominancia de la narración y la visualización, pues ambas están presentes en los 8 artículos analizados. Sin embargo, destaca que hayan empleado en dos de ellas también la utilización, ya que según Bradshaw es la más compleja para llevar a cabo por los periodistas. Así, han creado aplicaciones a partir de la información y los datos extraídos de sus análisis.

El diario no solo realiza análisis de resultados electorales, sino que, tanto previa como posteriormente, también elabora artículos de datos sobre diversos temas. En total, 6 de los 8 artículos no versan sobre los resultados electorales.

Destaca el trabajo *La resaca electoral del 20-D explicada en diez mapas y dos gráficos*, la última de entre todas las piezas, por el contexto que aporta a la actualidad. A través del uso de diversas fuentes explican cuál ha sido “El

partido más votado provincia a provincia”, dónde se encuentran “Los caladeros de voto de cada partido”, quién ha vencido en el duelo “Podemos vs. Ciudadanos”, por qué nos hallamos ante “El fin del bipartidismo” o que ha habido “Menos abstención que hace cuatro años”.

Finalmente, para estudiar el trabajo de datos llevado a cabo por el diario, se debe señalar que *El Confidencial*, indica en cada pieza cuál ha sido la metodología para llevarla a cabo, qué *software* emplea, e incluso en algunas permite la descarga directa de archivos de datos para que el lector los consulte en formato CSV.

Tabla 1

		El Confidencial
Nº de piezas		8
Nº de visualizaciones	Estáticas	21
	Interactivas	11
Posee piezas formadas únicamente por visualizaciones		No
Nº máximo de fuentes utilizadas		16
Utiliza únicamente sondeos y encuestas como fuentes		Nunca
Utiliza fuentes de distinta naturaleza		Siempre

Fuente: Elaboración propia.

Antes de hablar de la situación de la redacción y de sus rutinas (análisis cualitativo), se realizará a continuación un breve recorrido del diario con el objetivo de contextualizar y situar al lector de este trabajo de investigación.

El tratamiento de datos en el periódico les ha hecho consolidarse como referente en Periodismo de Datos en España. Tanto es así que en febrero de 2016, la *Digital News Initiative* de Google elige a *El Confidencial* para fomentar

la innovación periodística²³, por la que la cual el “Laboratorio” de este diario, deberá crear nuevos productos y experimentos en el marco de innovación digital en los medios de comunicación.

Algunas de las piezas más sonadas que el diario digital ha elaborado hasta el momento han sido *The Migrants Files*, trabajo galardonado con el premio Data Journalism Awards 2014 y el European Press Pritze 2015, o Generación E.

Una vez se ha situado al lector con una breve introducción sobre el recorrido del diario con sus trabajos de datos, podemos comenzar con el desarrollo de las rutinas que realizan y su situación dentro del medio.

Respecto a las visualizaciones que acompañan los análisis, Daniele Grasso afirma que “el equipo cuenta con un programador que se encarga de hacer programaciones sencillas. Por lo general, todos tenemos conocimientos básicos de HTML que nos sirven para hacer trabajos simples. Además, muchos de los programas dedicados a las visualizaciones, no requieren conocimientos de C++. Sin embargo, en *El Confidencial*, en algunos trabajos externalizamos el servicio de programación para que otros se encarguen de realizar los interactivos más complejos y navegables”. Además, éstas tienen un carácter protagonista en el periódico, pero, el equipo tiene como norma que éstas siempre deben ir acompañadas de “texto que las apoyen, complementen, contextualicen o analicen”.

Respecto a la cuestión sobre la integración con otras secciones y si comparten rutinas, *El Confidencial* expone que la colaboración con otras secciones es una constante. A la hora de proponer temas, el trabajo es bidireccional, es decir, la propuesta temática puede partir de cualquier sección y buscar colaboración con los periodistas de datos para crear una pieza conjunta. “La Unidad de Datos provee visualizaciones al resto de secciones del periódico con el fin de incluirlas como apoyo al resto de artículos. [...] Puede que el equipo de datos

²³El Confidencial (25 de febrero de 2016) *Google elige a El Confidencial como medio para fomentar la innovación periodística.* Blogs. Recuperado de http://blogs.elconfidencial.com/comunicacion/en-contacto/2016-02-25/google-selecciona-confidencial-laboratorio-innovacion-medios-comunicacion-periodismo_1158461/

proponga un tema a una sección y se acuerda hacerla conjunta con ellos o firmada solo por el equipo de datos. O que cualquier sección busque en la Unidad de Datos un apoyo en forma de visualizaciones”.

Por último, en relación al emplazamiento, los periodistas de datos, aunque sean un equipo propio, se ubican en un conjunto de mesas que amplían las secciones de “Investigación” e “Internacional”.

4.2 – Periodismo de Datos en El Español

El periódico de Pedro J. Ramírez ha contado desde su nacimiento con piezas de datos. Para llevarlas a cabo, el diario cuenta con un equipo de cuatro profesionales: Kiko Llaneras, Antonio Delgado, Patricia López y Nicolás Pérez.

Para cubrir los comicios generales, el medio ha realizado un total de 6 piezas de datos que suman 33 visualizaciones de las que tan solo dos son interactivas, ya que el resto presentan formato JPEG. La pieza en la que encontramos las visualizaciones interactivas no incluyen texto, y las acompaña un breve que indica qué plasman los mapas –resultados provinciales y municipales del Congreso y Senado-, sin que podamos consultar un análisis narrativo de los mismos. Para ello, debemos remitirnos a una pieza publicada 13 horas antes, de lo que se desprende que anteponen el análisis a las visualizaciones.

Por lo que respecta a las fuentes, *El Español* recurre hasta a 15 fuentes diferentes para elaborar una de las piezas analizadas. Todas son resultados de encuestas realizadas por distintas empresas sociológicas como: GAD3, Metroscopia, My Word, Celestetel, Sigma Dos, Invymarck, Simple Lógica, NC Report, Encuestamos, Demoscopia Servicios, DYM, GESOP, GETS, TNS Demoscopia y CIS. Además, en una de las noticias en la que tratan de definir a los nuevos votantes mayores de edad, también se apoyan en fuentes personales que aportan su visión particular sobre estas elecciones. De esta forma, optan por humanizar los números. Además, *El Español* es el único de los tres diarios que ha empleado como mínimo en todas sus piezas 2 fuentes.

Asimismo, en las 6 piezas hallamos una predominancia por la narración y visualización como forma de comunicar por estar presentes en todas las piezas de datos. No obstante, 2 de estos trabajos están formados por visualizaciones interactivas que carecen de análisis, algo contrario a lo que recomienda Eva Belmonte (que los datos no dicen nada si no son interpretados), o Pilhofer como citábamos en la metodología del presente trabajo.

El diario no solo realiza análisis de resultados electorales, sino que días previos a los comicios, también elabora artículos de datos sobre diversos temas como la brecha generacional, el peso del voto joven o el grado de implicación política del electorado juvenil vinculado a la abstención de votos... En total, 2 de los 6 artículos no versan sobre los resultados electorales.

Para finalizar, cabe señalar que en las piezas de mayor dificultad –por los cálculos estadísticos que se deben realizar para su elaboración- el medio indica la metodología practicada explicada paso por paso.

Tabla 2

		El Español
Nº de piezas		6
Nº de visualizaciones	Estáticas	31
	Interactivas	2
Posee piezas formadas únicamente por visualizaciones		Sí (2)
Nº máximo de fuentes utilizadas		15
Nº mínimo de fuentes utilizadas		2
Variedad temática		(2/6)

Fuente: Elaboración propia.

El análisis cualitativo no ha podido llevarse a cabo con este medio al no obtener respuesta de ningún responsable por distintas vías (teléfono, e-mail y Twitter).

4.3 – Periodismo de Datos en El Mundo

A diferencia de los dos anteriores, el diario *El Mundo* es un inmigrante digital. El equipo de datos está formado por cinco personas: Hugo Garrido (coordinador), Javier J. Barriocanal (diseñador e infografista), Pablo Medina (programador), Paula Guisado y Marta Ley (redactoras).

De la misma manera que con *El Confidencial* y *El Español*, se ha debido recurrir al rastreo de los trabajos de datos a partir de los nombres de sus autores, búsquedas avanzadas de Twitter y *tags*. No obstante, este diario no tiene bien definidas las piezas, puesto que no todas están identificadas con la etiqueta “Periodismo de Datos” como sí ocurre con los nativos digitales. Así, en este caso, se ha debido realizar un seguimiento pormenorizado de las publicaciones de este periódico referidas a las elecciones consultando todos los trabajos publicados por el diario en las fechas señaladas.

En el período establecido, *El Mundo* ha trabajado un total de 5 piezas de datos. Entre ellas se pueden contabilizar 6 visualizaciones interactivas y 11 gráficos estáticos. Sin embargo, de entre todas las visualizaciones, cabe resaltar que tanto uno de los mapas interactivos como una tabla interactiva, muestran la misma información. Y, por otra parte, una de las visualizaciones está presente en dos noticias distintas. Además, encontramos una pieza formada únicamente por una visualización interactiva sin texto que la contextualice ni analice.

Por lo que respecta a las fuentes, la mayoría proceden de datos oficiales como barómetros del CIS, Censo Electoral de los Residentes Ausentes (CERA) o INE. Tan sólo en una pieza emplean fuentes no documentales, aportando así opiniones expertas de doctores en ciencia política. La pieza en la que mayor cantidad de fuentes han empleado, no supera las 4.

Respecto a la forma de comunicación, observamos un predominio de la narración y la visualización, al estar presentes ambas en la totalidad de las piezas. Asimismo, como elemento diferenciador de los trabajos realizados por los nativos digitales, *El Mundo* incluye un vídeo en uno de sus artículos.

Finalmente, conviene mencionar que la variedad temática es inferior a los otros diarios, pues de las 5 piezas, sólo 2 no versan sobre resultados electorales. Además, sólo hallamos trabajos de datos a partir del día de las elecciones.

Tabla 3

		El Mundo
Nº de piezas		5
Nº de visualizaciones	Estáticas	11
	Interactivas	6
Nº de piezas basadas en simple plasmación de datos		4
Nº máximo de fuentes utilizadas		4
Utiliza fuentes de distinta naturaleza		No siempre
Variedad temática		(2/6)

Fuente: Elaboración propia.

Respecto al equipo de datos, el diseñador Javier J. Barriocanal asegura en la entrevista realizada que el periódico “cuenta con un equipo propio de desarrollo –conformado por él y Medina- que realiza todo lo que incumbe a la programación”, por tanto, no es un servicio externalizado. El equipo de datos de *El Mundo* “es un departamento propio, pero se encuentra inserto en la sección de Nacional”. En dicha entrevista, el diseñador afirma que, de manera general, los temas son propios por lo que la Unidad de Datos trabaja de forma individual, “aunque en ocasiones colabora con el resto de secciones de la redacción”.

Por último, para trabajar en período de elecciones, “la dirección convoca una reunión para conocer qué se va a hacer de cara a los comicios. Entre lo que

ésta propone y lo que sugiere la unidad de datos, se llega a una serie de acuerdos que determinan qué temas realizará cada departamento”, afirma.

Tabla 4

		El Confidencial	El Español	El Mundo
Nº de piezas		8	6	5
Nº de visualizaciones	Estáticas	21	31	11
	Interactivas	11	2	6
Posee piezas formadas únicamente por visualizaciones		No	Sí (2)	Sí (1)
Nº de piezas basadas en simple plasmación de datos		2	1	4
Nº máximo de fuentes utilizadas		16	15	4
Nº mínimo de fuentes utilizadas		1	2	1
Utiliza únicamente sondeos y encuestas como fuentes		Nunca	2	1
Utiliza fuentes de distinta naturaleza		Siempre	No siempre	No siempre
Variedad temática		(6/8)	(2/6)	(2/6)

Fuente: Elaboración propia.

Pese a que los tres diarios parten de la misma disciplina observamos diferencias considerables en los estilos de los artículos resultantes. Así, en *El Español* se aprecia un Periodismo de Datos basado en el análisis y elaboración de visualizaciones –en su mayoría estáticas–, vinculado al estilo sociológico. *El Confidencial*, por su parte, realiza un Periodismo de Datos que ofrece contexto, visualizaciones navegables y análisis. Y por último, *El Mundo* es el único que recurre a narrativas tradicionales para nutrir sus piezas de datos como vemos en uno de los trabajos en los que se hacen valer de un vídeo multimedia con opiniones de dos periodistas de la propia casa.

De entre todas las piezas, se hallan 3 formadas exclusivamente por visualizaciones, sin texto que las contextualice y analice. Algo significativo y contrario a la estimación que hacía en el marco teórico Eva Belmonte y Pilhofer.

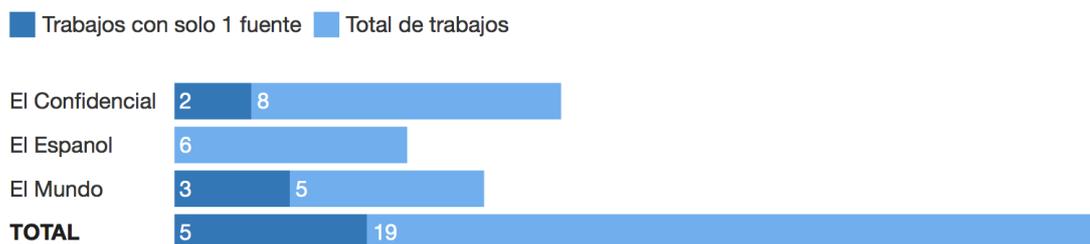
Con respecto a la temática, se observa un nexo común en todos los diarios: de entre todas las piezas de cada medio, al menos 2 son dedicadas al análisis de sondeos electorales y representación visual de los resultados. No obstante, *El Español* y *El Confidencial*, han elaborado artículos de datos durante los días previos a la cita electoral abarcando así mayor diversidad de temas que *El Mundo* no ha tratado. Por tanto, se acercan más a lo que recomienda Lorenz, que se profundice en distintos aspectos para que se entienda cómo los datos pueden afectar al ciudadano.

5) CONCLUSIONES

Como conclusión se trata de responder a las hipótesis que, ligadas a los objetivos, se plantean al inicio de esta investigación. Alrededor de la temática electoral, se busca responder a las cuestiones relativas a la cobertura informativa de Periodismo de Datos que los periódicos comentados realizan.

1. Aunque el uso de fuentes es elevado en algunos trabajos (16 y 15) hay 5 piezas que se apoyan en tan solo 1 fuente. Esto supone que un 26% de los trabajos de datos no se apoyan en diversidad de fuentes para profundizar en variables o aspectos como *qué ha pasado y cómo afecta a los ciudadanos* los resultados electorales. Al significar apenas $\frac{1}{4}$ de la totalidad de piezas, no se confirma la primera hipótesis que planteaba que los trabajos apoyados en pocas fuentes eran los más abundantes.

HIPÓTESIS 1 - USO DE UNA ÚNICA FUENTE



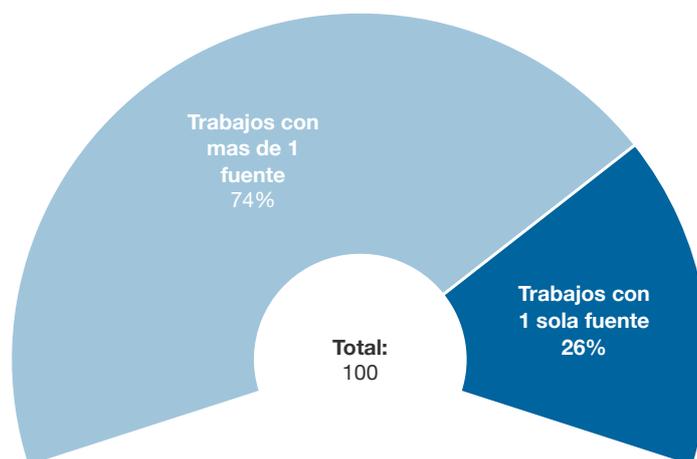
Elaboración propia

Source: [Bundestag](#)

Created with [Datawrapper](#)

Elaboración propia.

HIPÓTESIS 1 - USO DE UNA ÚNICA FUENTE



Elaboración propia

Source: [Bundestag](#)

Created with [Datawrapper](#)

Elaboración propia.

2. Lo segundo que se cree conveniente evidenciar es que se han localizado más artículos en los diarios nativos digitales que en el inmigrante digital. Sucede lo mismo con el número de visualizaciones, que es un 50% menor en el caso del nativo digital. Por lo que se confirma la segunda hipótesis: los inmigrantes digitales trabajan más cantidad de piezas de datos y más visualizaciones. Además, los análisis de los nativos digitales son más profundos porque se apoyan en mayor número de fuentes (triplican la cantidad de fuentes que emplea el inmigrante digital) y estudian las

consecuencias y causas de los acontecimientos. Con esto quedan definidas las principales diferencias entre los tres diarios y se alcanza el segundo objetivo.

HIPÓTESIS 2 - Nativos digitales vs. Inmigrante digital



Elaboración propia

Source: [Bundestag](#)

Created with [Datawrapper](#)

Elaboración propia.

3. Es conveniente señalar la dificultad para hallar todas las piezas de datos. Se ha debido recurrir a la búsqueda por Twitter, nombre de autores y *tags* en los tres diarios, e incluso en el caso del nativo digital se ha debido realizar búsqueda por fechas para la comprobación de artículos uno por uno al no identificarse todos los trabajos de datos con etiquetas o elementos identificativos.

Ninguno de los diarios cuentan con una sección específica dedicada a los datos como sí sucede en medios anglosajones como el *The Guardian* y *The New York Times*. Esto certifica el retraso de las unidades de datos españolas respecto a sus homólogas extranjeras y la situación todavía incipiente del Periodismo de Datos en los medios españoles.

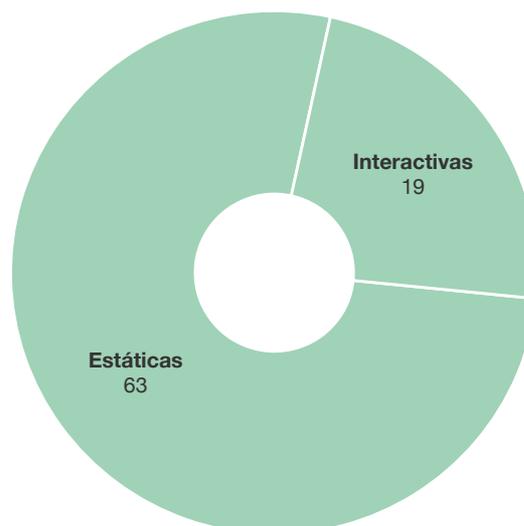
Por otra parte, no es casual que el diario que realiza mayor cantidad de visualizaciones interactivas tenga parcialmente externalizado dicho servicio de programación (los trabajos más complejos los realiza una empresa ajena al periódico). Demostrando así que la inserción de los programadores profesionales es cada vez más necesaria en estos equipos periodísticos para dotar de calidad a los resultados infográficos.

Además, las redacciones (al menos las de *El Mundo* y *El Confidencial* que son las que han respondido a las preguntas de este estudio) aún se

encuentran insertas dentro de otras secciones, lo que confirma lo que considera Montse Quesada, que en España los medios aún no han apostado por crear secciones dedicadas únicamente a los datos y al *Open Data*.

Por tanto, se confirma la tercera hipótesis: las redacciones en España todavía son muy reducidas y la poca formación de los periodistas en materia de programación, así como la no inclusión del programador en todos los equipos de datos, afectan, al menos, al resultado de las visualizaciones que, en su mayoría (en el presente trabajo, de las 82 visualizaciones, 63 eran fijas), terminan siendo estáticas por suponer un trabajo más elemental.

HIPÓTESIS 3 - Redacciones españolas y programadores



Elaboración propia

Source: [Bundestag](#)

Created with [Datawrapper](#)

Elaboración propia.

6) BIBLIOGRAFÍA

- Chapman, R. A. & Hunt, M. (1987). *Open Government. A study of the Prospects of Open Government within the Limitations of the British*

Political System, Routledge Revivals: Londres y Nueva York.

- Dader, J. L. (1997). *Periodismo de precisión: la vía socioinformática de descubrir noticias*. Madrid: Síntesis.
- Ferreras Rodríguez, E. M. (2013). Aproximación teórica al perfil profesional del 'Periodista de datos', *Icono 14*, volumen 11 (2), pp.115-140. doi: 10.7195/ri14.v11i2.573
- Gray, J. Chambers, L. & Bounegru, L. (2011). *The Data Journalism Handbook*. Reino Unido: Open Knowledge Foundation. Recuperado de <http://datajournalismhandbook.org/1.0/en/>
- Meyer, P. (1993). *Periodismo de precisión. Nuevas fronteras para la investigación periodística*. Barcelona: Bosch
- Rogers, Simon (ed.) (2011): *Facts are Sacred. The Power of Data*. Londres: Guardian Books.

7) BIBLIOGRAFÍA WEB

- Alianza para el Gobierno Abierto (2011) *Participating Countries*. Recuperado de <http://www.opengovpartnership.org/es/countries>
- Alsina, Montserrat (2014). Open Government, Open Data, Big Data y Transparencia: la información como nexo de unión. COMeIN, 39. Recuperado de <http://www.uoc.edu/divulgacio/comein/es/numero39/articles/Article-Montserrat-Garcia-Alsina.html>
- Belmonte, E. (2015, julio 28) Cómo hacer periodismo de datos con el ejemplo de Medicamentalia [Archivo de vídeo] Recuperado de <http://escuela.civio.es/como-hacer-periodismo-de-datos-con-el-ejemplo-de-medicamentalia/>
- Boletín Oficial del Estado (BOE) del 9 de diciembre de 2013, número 295, página 1. Recuperado de <https://www.boe.es/buscar/pdf/2013/BOE-A-2013-12887-consolidado.pdf>
- Bradshaw, P. (2011) The inverted pyramid of data journalism. *Online Journalism Blog*. Recuperado de

<https://onlinejournalismblog.com/2011/07/07/the-inverted-pyramid-of-data-journalism/>

- Bradshaw, P. (2011) 6 ways of communicating data journalism (The inverted pyramid of data journalism part 2). *Online Journalism Blog*. Recuperado de <https://onlinejournalismblog.com/2011/07/13/the-inverted-pyramid-of-data-journalism-part-2-6-ways-of-communicating-data-journalism/>
- Columbia Journalism School (2016) *Columbia Journalism Partners on Panama Papers*. Recuperado de <http://www.journalism.columbia.edu/page/1253-columbia-journalism-partners-on-panama-papers/748>
- Crucianelli, S. (2013). *¿Qué es el periodismo de datos?* Cuaderno de periodistas, 26, p. Recuperado de <http://www.cuadernosdeperiodistas.com/que-es-el-periodismo-de-datos/>
- Delgado, A. (2014) *Periodismo de datos. Una introducción a esta disciplina del periodismo*. Recuperado en <http://mugi.webs.upv.es/wp-content/uploads/2014/10/periodismo-datos-valencia-adelgado.pdf>
- El Confidencial (25 de febrero de 2016) *Google elige a El Confidencial como medio para fomentar la innovación periodística*. Blogs. Recuperado de http://blogs.elconfidencial.com/comunicacion/en-contacto/2016-02-25/google-selecciona-confidencial-laboratorio-innovacion-medios-comunicacion-periodismo_1158461/
- Flores, J. & Salinas, C. (2012) Sinergias en la construcción del Nuevo Periodismo derivadas del Data Journalism y el Transmedia Journalism. En Ortega, F. & Cardeñosa, L. (Eds), *III Congreso Internacional Comunicación 3.0* (pp. 7) Recuperado de http://www.2ip.es/wp-content/uploads/2013/01/sinergias_construccion_nuevoperiodismo.pdf
- Fundación Ciudadana Civio (2012) Breve historia de Civio. Recuperado de <http://www.civio.es/breve-historia-de-civio/>
- Guallar, J. y Peiró, K. (2013) *Introducció al periodisme de dades. Característiques, panoràmica i exemples*. COBDC, 57, p.22-36 Recuperado de <http://www.raco.cat/index.php/Item/article/viewFile/269701/357321>

- Open Knowledge Foundation (2012) *What is Open?* Recuperado de <https://okfn.org/opendata/>
- Pilhofer, A. (2013, diciembre 23) Los datos pueden ser una fuente y contar historias [Archivo de video]. Recuperado de http://blogs.cccb.org/lab/es/video_aron-pilhofer-les-dades-poden-ser-una-font-i-explicar-histories/
- The Guardian (26 de septiembre de 2011) The first Guardian data journalism: May 5, 1821. Datablog. Recuperado de <http://www.theguardian.com/news/datablog/2011/sep/26/data-journalism-guardian>

8) ANEXOS

The screenshot shows a Gmail interface with the subject "Preguntas sobre Periodismo de Datos en EC - TFG Marta Gozalbo". The email is from Marta Gozalbo Berenguer to ablanco. The content of the email is as follows:

Hola, buenos días Adrián. Acabo de hablar contigo por teléfono. Soy Marta Gozalbo. Estas serían las preguntas que finalmente me gustaría que respondieras:

- ¿Quiénes forman la Unidad de Datos de El Confidencial?
- ¿Qué formación tienen? Es decir, ¿tienen alguna especialización?
- ¿Cuenta el medio con programadores que realizan las visualizaciones? ¿O es el equipo Multimedia quien las confecciona?
- ¿La Unidad de Datos se encuentra integrada en alguna sección del resto de la redacción o es un departamento propio?
- En período de elecciones, ¿se trabaja de manera conjunta con el resto de la redacción que no trabaja con datos?
- ¿Varían las rutinas productivas en este periodo? ¿Cómo?
- ¿Qué fuentes se suelen consultar en periodo electoral? ¿Se realizan peticiones de información?
- ¿Qué criterios determinan que la Unidad de Datos trabaje unas noticias y el resto de secciones otras?
- ¿Qué grado de protagonismo tienen las visualizaciones en las piezas de datos?

Muchísimas gracias de antemano.
Saludos, Marta.

Entrevista a Adrián Blanco, periodista de datos de *El Confidencial*.

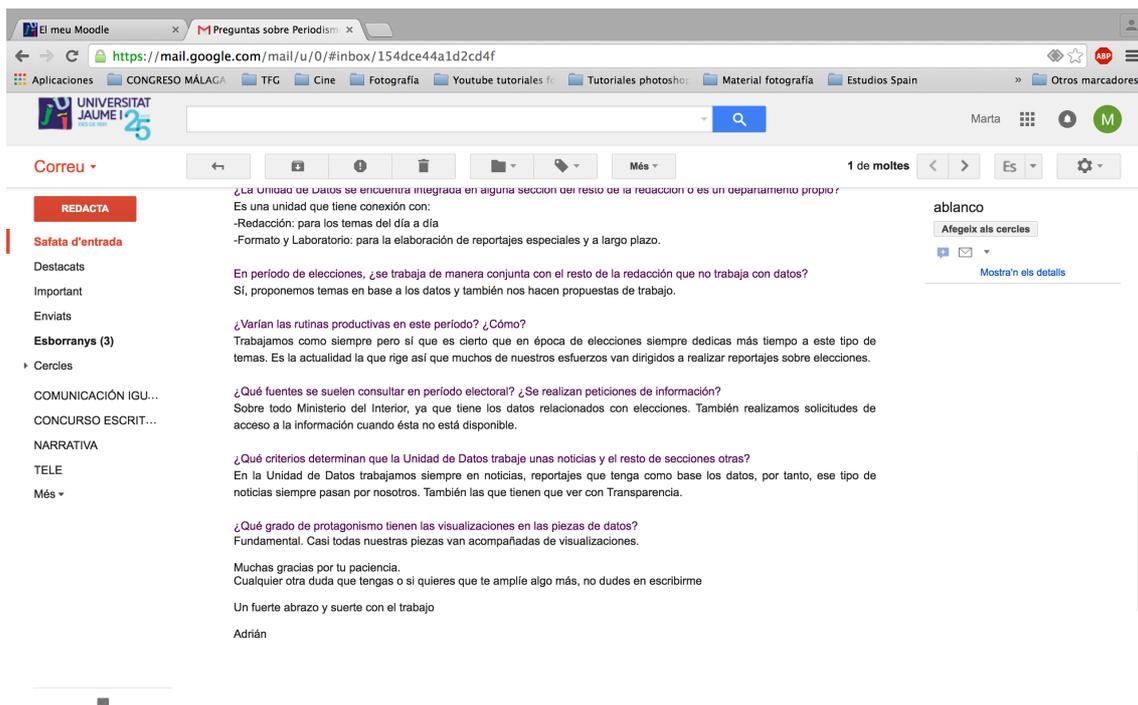
The screenshot shows a Gmail interface with the subject "Preguntas sobre Periodismo de Datos en EC - TFG Marta Gozalbo". The email is a reply from Adrián Blanco to Marta. The content of the email is as follows:

Hola Marta,

Disculpas de nuevo de parte de los tres. Te respondo a las preguntas que nos planteabas. Espero que sean de ayuda:

- ¿Quiénes forman la Unidad de Datos de El Confidencial?
Somos tres periodistas de datos: Daniele Grasso, Jesús Escudero y Adrián Blanco. Contamos con el apoyo de un programador, Antonio Esquembre.
- ¿Qué formación tienen? Es decir, ¿tienen alguna especialización?
Sí, un master en periodismo de datos y especialización por cuenta propia en estadística, análisis de datos, visualización, unas pinceladas de lenguajes de programación
- ¿Cuenta el medio con programadores que realizan las visualizaciones? ¿O es el equipo Multimedia quien las confecciona?
Sí. Tanto programadores como nosotros mismos nos encargamos de realizar las visualizaciones.
- ¿La Unidad de Datos se encuentra integrada en alguna sección del resto de la redacción o es un departamento propio?
Es una unidad que tiene conexión con:
-Redacción: para los temas del día a día
-Formato y Laboratorio: para la elaboración de reportajes especiales y a largo plazo.
- En período de elecciones, ¿se trabaja de manera conjunta con el resto de la redacción que no trabaja con datos?
Sí, proponemos temas en base a los datos y también nos hacen propuestas de trabajo.
- ¿Varían las rutinas productivas en este periodo? ¿Cómo?
Trabajamos como siempre pero sí que es cierto que en época de elecciones siempre dedicamos más tiempo a este tipo de temas. Es la actualidad la que rige así que muchos de nuestros esfuerzos van dirigidos a realizar reportajes sobre elecciones.

Entrevista a Adrián Blanco, periodista de datos de *El Confidencial*.



Entrevista a Adrián Blanco, periodista de datos de *El Confidencial*.

Aparte de esta entrevista por e-mail a Adrián Blanco, se decidió tratar de hacer otra a un miembro más del equipo de datos por la brevedad de las respuestas. De esta forma, se contactó vía telefónica con Daniele Grasso, que amplió las respuestas de su compañero. La entrevista se encuentra transcrita a continuación. La elección del canal, en este caso, quedó sujeta al condicionante del tiempo que quedaba para la entrega del proyecto.

TRANSCRIPCIÓN ENTREVISTA TELEFÓNICA A DANIELE GRASSO (*El Confidencial*)

¿Quiénes forman la Unidad de Datos de *El Confidencial*?

Jesus Escudero y yo, Daniele Grasso fuimos las dos primeras personas que comenzamos a realizar periodismo de datos en *El Confidencial* en el año 2013. Jesús en la sección de “Kotizalia” y yo en la de “Nacional”. Actualmente, también trabaja con datos Adrián Blanco y Antonio Esquembre, que es nuestro programador.

¿Qué formación tienen?

Los tres somos licenciados en Periodismo e hicimos el “Máster en Periodismo de Investigación, Datos y Visualización” de *El Mundo* y la Universidad Rey Juan Carlos. Más tarde hemos ido adquiriendo conocimientos sobre estadística, análisis o programación básica a través de cursos.

¿Cuenta el medio con programadores que realizan las visualizaciones?

¿O es el equipo Multimedia quien las confecciona?

El equipo cuenta con un programador que se encarga de hacer programaciones sencillas. Por lo general, todos tenemos conocimientos básicos de HTML que nos sirven para hacer trabajos simples. Además, muchos de los programas dedicados a las visualizaciones, no requieren conocimientos de C++. Sin embargo, en *El Confidencial*, en algunos trabajos externalizamos el servicio de programación para que otros se encarguen de realizar los interactivos más complejos y navegables.

Con el equipo Multimedia se colabora constantemente. Está formado por 6 personas, pero tienen conocimientos que nada tienen que ver con la programación. Unos se encargan de vídeo e imagen, otros son diseñadores...

¿La Unidad de Datos se encuentra integrada en alguna sección del resto de la redacción o es un departamento propio?

La Unidad de Datos es un equipo propio, pero nos situamos en la sección de “Internacional” e “Investigación”. Dependiendo del tema, trabajamos con la redacción, temas diarios, o con Laboratorio, para temas más amplios.

En período de elecciones, ¿se trabaja de manera conjunta con el resto de la redacción que no trabaja con datos?

La Unidad de Datos provee visualizaciones al resto de secciones del periódico con el fin de incluirlas como apoyo al resto de artículos. Las propuestas de temas y el desarrollo de estos parten de ambas partes. Puede que el equipo de datos proponga un tema a una sección y se acuerda hacerla conjunta con ellos o firmada solo por el equipo de datos. O que cualquier sección busque en el equipo de datos un apoyo en forma de visualizaciones.

¿Varían las rutinas productivas en este período? ¿Cómo?

Sí porque en un período concreto, se intenta trabajar sólo piezas de elecciones. En una fase tan determinada, también rige la actualidad.

¿Qué fuentes se suelen consultar? ¿Se realizan peticiones de información?

Generalmente acudimos a bases de datos abiertas porque lo otro es un proceso lento, que además puede no servir para nada porque puede que nunca te llegue la información que necesitas. En período de elecciones se utilizan muchas fuentes como el INE, Eurostat, EPA... Y los datos que facilita el Ministerio de Interior. Y luego también el equipo de datos realiza con ellos análisis estadísticos o busca la interpretación de algunas variables.

¿Qué criterios determinan que la Unidad de Datos trabaje unas noticias y el resto de secciones otras?

Depende de quién parta la propuesta y de la reunión con la dirección. Generalmente los trabajos de datos requieren más análisis. Y aunque se busca sacar un tema diario, no siempre se consigue. El resto de secciones sí lo hacen.

Pero todos los temas que impliquen datos o Transparencia, pasa por la Unidad de Datos.

¿Qué grado de protagonismo tienen las visualizaciones en las piezas de datos?

El protagonismo de todos los trabajos recaen en la parte visual, es lo que más llama la atención del lector. Por eso es importante que sea atractiva. Pero en la Unidad de Datos tenemos claro que siempre se tienen que completar con una explicación, con texto que las apoye, complemente, contextualice o analice.

Para la elaboración de la entrevista con *El Mundo*, tuve que ponerme en contacto con ellos vía telefónica ya que, en su web tan sólo presentan una dirección de correo destinada a aquellos que quieran hacer una inversión publicitaria y otra dirección para los accionistas de Unidad Editorial. La otra opción que presentaba la web periódico para ponerte en contacto con ellos consistía en rellenar un formulario, pero sólo para cuestiones técnicas. Aunque lo rellené en distintas ocasiones, nunca obtuve respuesta. De este modo, decidí llamar a la centralita de Unidad Editorial y solicité que me pusieran en contacto con alguna de las 5 personas que trabajan en la Unidad de Datos de *El Mundo*.

TRANSCRIPCIÓN ENTREVISTA TELEFÓNICA A JAVIER J. BARRIOCANAL (*El Mundo*)

¿Quiénes forman la Unidad de Datos de *El Mundo*?

El Mundo Data lo formamos 5 personas. Hugo Garrido, que tiene perfil de informático y es el Coordinador de la sección, Pablo Medina es el programador, Javier J. Barriocanal (infografista y visualizador), y Paula Guisado y Marta Ley son las redactoras.

¿Qué formación tienen?

Paula Guisado y Marta Ley estudiaron el Máster de Unidad Editorial. El resto tenemos formación específica de informático, programador y diseñador gráfico.

¿Cuenta el medio con programadores que realizan las visualizaciones?

¿O es el equipo Multimedia quien las confecciona?

El equipo lo formamos el departamento de Desarrollo. En él trabajamos Pablo Medina, que es el programador, y yo, que soy diseñador y también hago la programación ligera y el CSL.

¿La Unidad de Datos se encuentra integrada en alguna sección del resto de la redacción o es un departamento propio?

La Unidad de Datos es un departamento propio, pero se encuentra dentro de “Infografía” o “Nacional”.

En período de elecciones, ¿se trabaja de manera conjunta con el resto de la redacción que no trabaja con datos?

Depende del tema. Cuando hacemos algún tema que se ha desarrollado en la sección de Nacional, pues sí. Pero, generalmente los temas son propios y los hacemos de forma individual.

¿Varían las rutinas productivas en este período? ¿Cómo?

Sí porque, además de hacer muchas reuniones, se le suele dar prioridad a todos los temas de actualidad política.

¿Qué fuentes se suelen consultar? ¿Se realizan peticiones de información?

Sobre todo los datos que ofrece el Ministerio del Interior con las empresas a las que contrata.

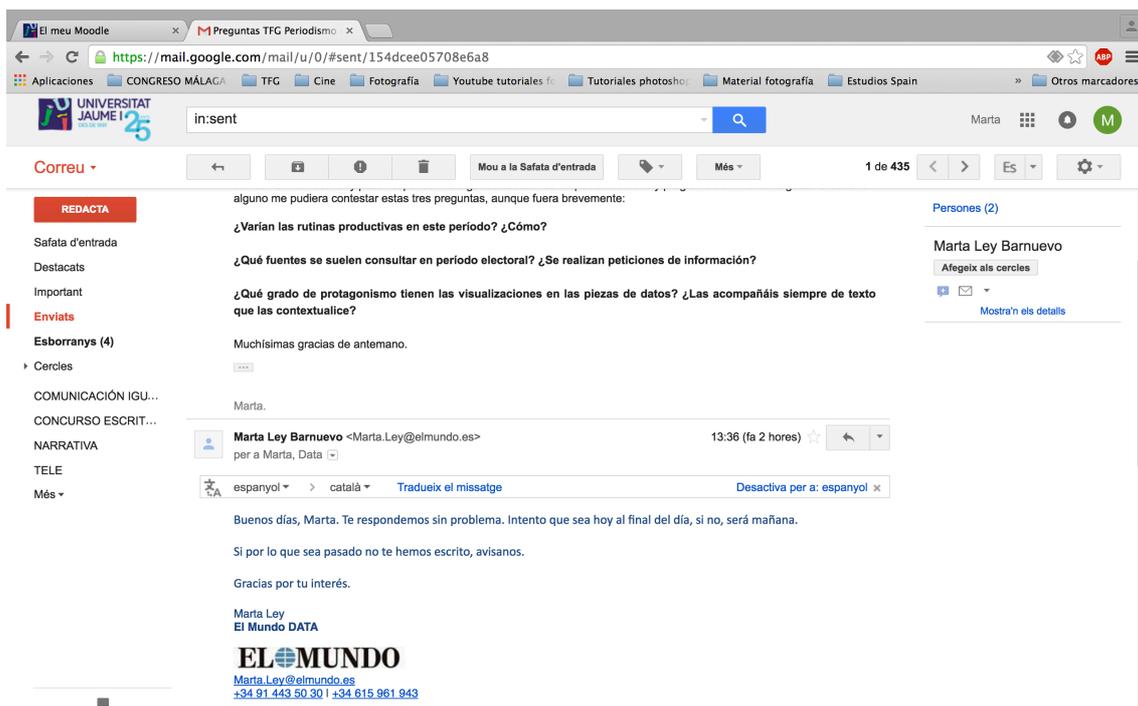
¿Qué criterios determinan que la Unidad de Datos trabaje unas noticias y el resto de secciones otras?

Ante temas tan concretos, hay una jerarquía. Generalmente hay una serie de reuniones que convoca la dirección para ver qué vamos a hacer de cara a las elecciones. Nosotros hacemos nuestras propuestas, pero la dirección también quiere que determinados temas queden cubiertos. Así que entre lo que propone la dirección y lo que proponemos nosotros, se llegan a acuerdos.

¿Qué grado de protagonismo tienen las visualizaciones en las piezas de datos?

Son fundamentales. Facilitan mucho la comprensión de la información, que en ocasiones parece que sea un tanto abstracta.

Del mismo modo que pasó con *El Confidencial*, esta vez traté de ponerme en contacto con alguien más de la sección de datos. Esta vez fue Barriocanal quien me facilitó un correo electrónico. Aunque la periodista Marta Ley mostró interés en ampliar tres cuestiones que el diseñador de *El Mundo* resolvió de manera escueta, no volvió a contestar al correo electrónico.



Entrevista a Marta Ley, periodista de datos de *El Mundo*.

EXECUTIVE ABSTRACT

Electoral analysis of 20-D through Data Journalism by *El Confidencial*, *El Español* and *El Mundo*

Data Journalism appears as a new discipline driven by the development of the Information & Communication Technologies (ICTs). A new model of journalism that shows up as a massive opportunity in order to endow with quality to the profession and that allows a bigger development on media calendar issues.

A big amount of theorists link this new speciality to precision journalism associated with Philip Meyer and Computed Assisted Reporting (CAR) because of its high technological content. Furthermore, some of them think its appearance couldn't have been possible, or it would have suffered a gradual introduction, if not for the *Open Government* and *Open Data* movements, as both of them have been two important factors for the sake of transparency and open access to information. These allow journalists to handle databases without the need to seek formal requests of information to the governments. A task which meant a complication to those journalism professionals dedicated to investigation, as these requests went always hand in hand with long waits.

Even if the circumstances in Spain are not the most suitable -the Spanish Ley de Transparencia has been labelled by *Access Info Europe* as a law that doesn't comply to the minimum standards of international transparency-, journalists are able to work with accessible data through different programs of their own to every one of the phases Data Journalism works with. Programs that come with programming functions - excluding those that are specifically for the interactive visualizations design, that functions with a bigger complexity and requires specific knowledge-.

This new speciality has features that define it specially like the use it makes of big databases, the importance of visualizations in the resulting journalistic pieces and the valuable programming task to optimize working time and efficient search between databases that allow to detect patterns and trends valid to the later storytelling.

The goal of this investigation is to analyze the election coverage through the Data Journalism conducted on the Spanish newspapers *El Confidencial*, *El Español* and *El Mundo*. This analysis entails an study of the diversity of information sources each media uses in order to support their election analysis and to contextualize the present through different points of view. This way, we'll try to know if data pieces are supported merely on documentary sources or, on the contrary, are also backed on expert sources, such as political scientists that help them to read the processed data. Also, to learn if it gives voice to those individuals involved in participation processes (voters, candidates to the Congreso and candidates to the Senado).

On the other hand, we'll study the difference between the work by digital natives and digital immigrants. Further, we pretend to create a brief approximation to the Spanish Data Units situation and the productive routines that unfold on a election period.

This study focuses on three principal scenarios that try to give an answer to the goals described above. The first of them answers to the sources used by each media, their diversity or categorization that allow to contextualize the current political situation through different points of view. The initial overview is that, even if the sources are a distinguishing element of Data Journalism, the amount used and its diversity are low.

On the other hand, the second scenario considers digital natives as better adapted to this new kind of journalism, composing a bigger quantity of articles and a larger number of website visualizations.

Finally, the third scenario raises that, being an emerging discipline in Spain, the working teams are small in the national editorial offices and this affects the pieces they fulfill, consequence of the small proportion of thoroughness achieved by quantity and diversity used and because of the development of visualizations in every article.

Therefore, this poses as an observational and comparative investigation that responds both to the analysis of the formal aspects and the content of the three chosen media. To set up the sample three journals have been chosen that frequently create pieces or works of Data Journalism. Regarding the temporary sample, it was better thought to select 5 days prior and 5 days after the date with the ballot boxes, being that a period when the sociological business produce a big amount of data that journalists can use in order to conduct analysis and interpret them by visualizations that make their data comprehension easier. In total, the sample is made up by 19 pieces.

The classification parameters contemplated to this quantitative analysis through a chart is divided in 9 groups. The first of those groups considers the number of pieces that allows to know the periodicity or the volume of work generated by the three journals in the ten days selected. The second explores the number of confectioned visualizations and the static or interactive nature of these, as the interactive ones require a huger programming work than the static ones, which allows to know what every media banks on that task and the incorporation of a programmer to the journalistic team.

Concerning the third parameter, the quantity of pieces that show visualizations is analyzed without these being followed by a text to contextualize, something not akin to what Pilhofer considers, that Data Journalism doesn't merely consist in write down data without visualizations, that it has to tell a story the best way affordable so they should be followed by a text that includes the context they represent.

Fourthly, hoy many of the pieces developed by media has been studied in order to learn how many of them are based on a basic expression of results with no analysis or interpretation. Learning this it's meaningful because Data Journalism has a clear analytical intent. Further, the fifth and sixth parameter contemplates which are the pieces in which a bigger and lesser amount of sources are used in favor of knowing if it exists a balance between all the pieces, and also to detect if there is a imbalance or distortion between the three

journals as, indeed, it's been verified. The digital immigrant doesn't even get to use 1/3 of the quantity used on the other two journals.

Talking about the sources, the fact of journals getting support only by documentary sources or not has been noted, as well as if they give voice to different individuals involved or experts in the field.

Finally, it's important to learn the variety of topics in a sense in which, achieving so, we have come to realize which one is the media that has worked on more topics of their own beyond analyzing polls or surveys about voter intention or election results opposite to the media that has complied in a more stricter way to that topic.

All of this quantitative research has allowed to obtain significant conclusions regarding the first two presented scenarios: that a 26% of the works use only one source to compile a piece of Data Journalism, that the variety of topics in *El Español* and *El Mundo* is scarce since it's only 1/3 of the pieces compiled (the vast majority of their articles analyze only surveys about voter intention and election results), that only one of the journals uses different types of sources in every single one of their articles, giving voice not only in a number of articles to the digital native but duplicating that visualization's number.

On the other hand, a qualitative research has been conducted as well. The journalistic routines of the three media has been analyzed through in-depth interviews that have allowed to learn what level of establishment this new discipline has in the Spanish national editorial offices: not everyone of them have a programmer of their own, some of them have an outside service for the production of more elaborated visualizations instead of inserting the figure of a programmer to their journalistic team. In that sense, more simplistic visualizations are done by programs that don't require programming. In the other media, the visualization works are done between the programmer and the designer.

Plus, even if they consider the Data Unit as a department of their own, the team is inset on other sections already created (in one of the cases in Nacional - cutting back the variety of topics the newspaper can deal with only to the national level- and the other one in International and Investigation). They don't have either a specific section inside the media and the challenge to locate data work not being classified -neither in a section nor through labels- is elevated. However, Data United works with topics of their own, even if the collaboration with the rest of the editorial office is constant.