



Universidad de Valladolid

CURSO 2021-2022

Facultad de Filosofía y Letras

Grado en Periodismo

**Evolución y desarrollo del Periodismo de
Datos en España. Estudio de caso de El País**

Patricia Hernández Pantoja

Tutora: Pilar Sánchez-García

**Departamento: Historia moderna, Contemporánea y de América,
Periodismo**

Convocatoria: Ordinaria 2022

Evolución y desarrollo del Periodismo de Datos en España. Estudio de caso de El País

AUTORA:

Patricia Hernández Pantoja

TUTORA:

Pilar Sánchez-García

RESUMEN:

Desde el Periodismo de Precisión hasta llegar al actual Periodismo de Datos se necesita recurrir a nuevas tecnologías como el *Big Data* y la minería de datos. En esta nueva rama se acude a la visualización de la información obtenida que evoluciona y mejora a través de la interacción y la multimedialidad enriquecida. En este contexto, el objetivo de esta investigación es observar el progreso y desarrollo de las noticias de Periodismo de Datos en España, en concreto mediante el estudio de caso de la edición digital de *El País*. Para su desarrollo se recurre a una doble metodología tanto para el marco teórico como empírico; primero se realiza una entrevista en profundidad, a Mariano Zafra del equipo narrativo del medio y se completa con la técnica el análisis de contenido. A partir de una tabla de análisis, de elaboración propia, se analizan 3 categorías que incluyen 9 variables, comprendidas de 45 subvariables con una muestra (N=140) seleccionada de informaciones de Periodismo de Datos en la web de *El País* durante siete años (2016-2022). Los resultados muestran que la aplicación de esta nueva especialización periodística se está aplicando en más temáticas y nuevas narrativas.

PALABRAS CLAVES: Periodismo de datos, *Big data*, evolución, El País, narrativa multimedia

Evolution and development of Data journalism in Spain. Case study of EL PAÍS

AUTHOR:

Patricia Hernández Pantoja

TUTOR:

Pilar Sánchez-García

ABSTRACT:

From Precision Journalism to the current Data Journalism you need to resort to new technologies such as Big Data and data mining. This new branch uses the visualization of the information obtained that evolves and improves through interaction and enriched multimedia. In this context, the objective of this research is to observe the progress and development of Data Journalism news in Spain, specifically through the case study of the digital edition of El País. For its development, a double methodology is used for both the theoretical and empirical framework; first an in-depth interview is conducted with Mariano Zafra of the narrative team of the media and the content analysis technique is completed. From an analysis table, of own elaboration, we analyse 3 categories that include 9 variables, this include 45 subvariables, which are applied to a sample (N=140) selected from data journalism information on the website of El País for seven years (2016-2022). The results show that the application of this new journalistic specialization applying in more thematic and new narratives, but remains focused on the national level with 75% of publications.

KEY WORDS: *data journalism, big data, evolution, El País, multimedia narrative*

Índice

1. Introducción	5
1.1. Justificación personal	5
1.2. Justificación académica	6
1.3. Objetivos y preguntas de investigación	8
2. Marco teórico	9
2.1. Evolución del Periodismo de Datos: del Periodismo de Precisión al <i>Big Data</i>	9
2.1.1. Minería de datos	12
2.1.2. Visualización y narrativa del Periodismo de Datos	12
2.2. Evolución del Periodismo de Datos en España y Ley de Transparencia.....	14
2.2.1. Estudio de Caso de <i>El País</i>	16
3. Metodología	17
3.1. Muestra	17
3.2. Tabla de análisis	18
4. Resultados	22
4.1. Resultados de Temática	22
4.2. Resultados de elementos multimedia	24
4.3. Visualización de la infografía	28
5. Discusión y conclusiones	34
6. Bibliografía	37
7. Anexos	40

1. Introducción

1.1. Justificación personal

Desde que aprendes las letras, los números, a leer todo lo que ves es información. Hay veces que esa indagación es necesaria y en otros casos solo es complementaria. Con el periodismo pasa lo mismo. De manera indirecta seleccionamos la información en importante o no.

Empecé a acompañar a mi padre a la redacción de su trabajo el periódico *El Norte de Castilla*. Descubrí que hay gente que escribe y de esta manera ayuda a la gente. Observando a grandes redactores en uno de los medios antiguos de España, empezó el gusanillo por los medios de comunicación. Poder ver desde pequeña la función de informar, de dar visibilidad a cosas que pasaban desapercibidas para las personas normales, incentivó el interés por buscar cosas inéditas.

Para este oficio no hay que ser como los demás, hay que ver más allá. Con un afán al principio por la investigación, por el descubrir sucesos que antes nadie se había preguntado. Siempre con el porqué de lo que ocurría en la mente, hasta que llegó la Universidad y me topé con datos, muchos datos. Analizando redes, publicaciones, interacciones y estudiando campañas políticas una tras otra me di cuenta que de lo que más se fiaba la gente era de los resultados. De ese resumen de datos de los votos, de la repercusión... Estaba bien la explicación que lo acompaña, pero tener de manera visual e interactiva la cifra de participación en unas elecciones, llama más la atención que el proceso de un voto.

Así, he entendido que, con el avance de la sociedad, tecnología y, sobretodo, el aumento de los bulos, los datos en esta profesión de informar y de ayudar, cada vez son más necesarios.

Al buscar información queremos encontrarlo a la primera. Que nuestra principal búsqueda sea clara, concisa y breve, aquí es cuando aparece el Periodismo de Datos de los medios. Esa capacidad de ayudar y mostrar la realidad de una manera diferente, que

Hernández Pantoja, Patricia (2022). *Evolución y desarrollo del Periodismo de Datos en España. Estudio del caso EL País*. TFG. Universidad de Valladolid

permite al lector interactuar y sentirse parte de la noticia, es algo de la actualidad, el lector ha pasado de ser pasivo a convertirse en activo en los medios.

Por lo tanto, cuando leemos los periódicos, vemos la televisión o escuchamos la radio, prestamos atención a aquello que nos va aportar algo o que nos llama la atención. Esta especialización periodística es causante de esta selección inocente que realizamos cada día, nos hace la búsqueda y captación de la información más fácil.

Estos últimos años, con todas las catástrofes que han ocurrido, los datos han sido de vital importancia, y cada búsqueda debía ser clara y agradable a la vista. Porque en la actualidad el mundo de la gestión, el rol que tienen los datos y los procesos de las grandes compañías son necesarias para mostrar la realidad.

Al realizar las búsquedas te das cuenta de lo importante que son los datos, de cómo se tratan y para qué se utilizan. En este estudio se analizará el medio *El País*, uno de los periódicos más consolidados de España, donde podremos analizar su evolución a la hora de tratar los datos y como los plasman de cara al público.

1.1 Justificación académica

Son muchos los estudios que tratan el Periodismo de Datos, pero pocos los que se centran en la evolución y su desarrollo en España. Así esta investigación de esos pocos casos centrados en un solo país y en un solo medio.

Desde del libro de Philip Meyer de 1973, *Precision Journalism. A reporter's Introduction to Social Science Methods* donde explica el comienzo del Periodismo de Precisión, considerado el precedente más cercano al Periodismo de Datos, surgen multitud de estudios al respecto. Artículos centrados en la narrativa multimedia que supone esta especialización, indagaciones sobre los perfiles de los periodistas para poder tratar esta nueva información, los datos, análisis sobre el desarrollo del Periodismo de Datos con el uso de ordenadores y la innovación del uso de infografías para transmitir la investigación.

En 1994 los medios emergieron en Internet, lo que ha provocado cambios en las características principales del periodismo. Aparece el ciberperiodismo, toma prestado de la prensa y es posible el diseño de un lenguaje propio, como es el de los datos (Salaverria, 2004). La periodista especializada en periodismo de precisión, Crucianelli (2014) explica la formación de esta disciplina a partir de otras. Muestra el uso de los ordenadores junto con la implantación de los gráficos para representar los datos, apareciendo en los medios la necesidad de usar infografías.

El mostrar las infografías en los periódicos es otro campo de estudio en el periodismo de datos. *Data Visualization* tiene registrado que en los tres últimos meses del 2008 solo se realizaron 220 visualizaciones, 13 de ella de la propia redacción. Al realizar una comparativa con medios, como *The New York Times*, el uso de infografías en España es menor y más complejo. Por ello surgen observaciones de la infografía, cómo se emplean y su eficacia en el público (Cairo, 2008). Para usar esta herramienta es necesario la especialización de los periodistas creando nuevos perfiles.

Las investigaciones específicas sobre el perfil del periodista de datos incentivan el estudio a profundizar en el desarrollo e integración del Periodismo de Datos en España, ya que es una novedad en la formación del periodista tradicional (Arias, Sánchez-García y Redondo, 2018). Lo que permite que los medios de comunicación muestren interés por esta especialización y así poder invertir, tanto en la formación académica de futuros periodistas como en los actuales informadores de los medios.

La presente investigación ofrece un análisis del desarrollo del Periodismo de Datos en España a través del estudio de caso del diario *El País*. La elección del medio se debe a ser de los pocos periódicos españoles con un equipo de específico para tratar los datos. (Teruel y Blanco, 2016). Se analiza una evolución en un largo periodo del tiempo, del 2016 al 2022, centrado en el tratamiento de los datos y su exposición al público. La investigación se propone observar el inicio en este campo y su mejoría respecto a la tecnología que se aplica, además de observar cómo se utilizan los datos y, sobre todo, para qué.

1.2 Objetivos y preguntas de investigación

El objetivo general de este trabajo es analizar la información especializada de Periodismo de Datos de *El País* en sus noticias y sus características.

A partir de este objetivo general, establecemos dos objetivos específicos:

- O1. Analizar las temáticas de datos en las que se centra *El País* en las informaciones del medio digital.
- O2. Describir la narrativa multimedia y la visualización que utiliza *El País* en las noticias con Periodismo de Datos.
- O3. Descubrir el nivel de interactividad que ofrece *El País* en la narrativa multimedia donde predominan los datos.

Estos objetivos se plantean a partir de unas preguntas de investigación:

- ¿El Periodismo de Datos en España está bien plasmado en los medios?
- ¿*El País* utiliza herramientas adecuadas para sus noticias de Periodismo de Datos?
- ¿Los temas que usan datos son variados?
- ¿La visualización ayuda a la audiencia a entender mejor las noticias que leen?

Con el objetivo de mostrar una aproximación a la evolución en el Periodismo de Datos en España y su libertad de uso se plantean las siguientes hipótesis:

H1: La expansión del Periodismo de Datos ha crecido en la prensa de España en la última década

H2: La narrativa de datos en *El País* se centra en temas políticos y económicos

H3: Las infografías utilizadas por *El País* tienen interacción, *feedback* y prima lo visual

Una vez establecidos los objetivos del estudio es necesario explicar el objeto de estudio, el Periodismo de Datos

2. Marco teórico

La revisión documental realizada para este trabajo, permite ofrecer un marco teórico centrado en tres partes: la primera centrada en la evolución de Periodismo de datos y aplicación del *Big Data*. En segundo lugar, el uso de la minería de datos. Por último, a partir de una entrevista en profundidad la evolución y estudio del Periodismo de Datos en el medio a estudiar, *El País*

2.1 Evolución del Periodismo de Datos: del Periodismo de Precisión al *Big Data*

El *Big Data* hace referencia al conjunto masivo y complejo de información, puede estar estructurada o no. Esta información se recoge durante cierto tiempo y es necesario usar métodos computacionales para sacar conocimiento (Arcila, Barbosa y Cabezuelo, 2016). No es solo cuántos datos sino saber buscar, agregar y cruzar unos datos para obtener otros.

El *Big Data* es un avance tecnológico (Elías, 2015,) y los datos que se utilizan ayudan a explicar lo que sucede y lo que puede ocurrir en el futuro (Videla-Rodríguez, 2018). Boyd y Crawford entienden el *Big Data* desde tres puntos de vista (2012):

1. Tecnológico: incrementa el poder computacional y eficacia de los algoritmos para obtener, analizar y comparar datos.
2. Analista: encontrar patrones que den con tendencias y poder predecir comportamientos futuros
3. Mitológico: conjunto de datos que generen ideas, que, al estar avaladas por números, cuentan con verdad, objetividad y precisión.

Big Data es “la acumulación masiva de datos tal, que supera la capacidad del software habitual para que sean capturados, gestionados y procesados en un tiempo razonable con las herramientas tradicionales” (Mochón, González, 2016, p.25). Se añaden a la definición los términos de volumen, velocidad y variedad, conocido como las3V.

Por estos nuevos términos los datos que se utilizan en el *Big Data* tienen que ser abiertos, es decir, *open data* (Videla-Rodríguez, 2018) y esta información de libre acceso se guarda en la nube virtual llamada *cloud computing* (Elías, 2015). Porque los

periodistas pueden prescindir de ciertas herramientas, pero los datos no es una de ellas (Crucianelli, 2014).

Rodríguez (2018, p. 61) “el *Big Data* se ha convertido en un instrumento imprescindible de gestión de las empresas de comunicación”. Así el uso del *Big Data* en el periodismo es reciente, sobre todo en la comunicación política y estrategias electorales. Un ejemplo de su uso en política es la campaña de 2012 del presidente Barack Obama se desarrolló con tecnología *Big Data* (Elías, 2015).

El Periodismo de Datos hace uso del *Big Data* como herramienta para crear sus informaciones (Videla-Rodríguez, 2018). Esta nueva disciplina se nutre de otras corrientes periodísticas de la investigación, profundidad, precisión y el asistido por computadora entre otras. En cada una se trabaja con grandes volúmenes de datos (Crucianelli, 2014).

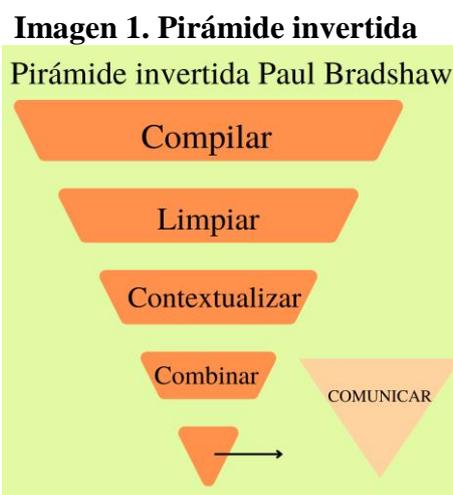
Aún con esto el principal antecedente es el Periodismo de Precisión (Ferrerías, 2012) que nace y desarrolla de la mano de la investigación sociológica (Dader, 1993). El periodismo de precisión es perfeccionado por el periodista Philip Meyer en 1973 (Elías, 2015). Remontando años atrás, 1821, el periódico [*The Guardian*](#) publicó información en una gran tabla las escuelas de Manchester y Salford, plasmaban el número de alumnos que asistían, el gasto anual y quienes recibían educación gratuita (Blanco, Teruel, Martín, 2016).

Ya a finales de los años treinta arrancaba el periodismo de precisión con los lanzamientos de las primeras encuestas científicas y técnicas de experimento (Dader, 1993). Los antecedentes que se relacionan al Periodismo de Datos destacan por la búsqueda de datos por parte del periodista (Blanco, Teruel y Martín, 2016). Meyer en 1973 define el periodismo de precisión como aquel que usa las matemáticas y los métodos de investigación científicas sociales para interpretar datos que explican un hecho noticioso (Elías, 2015).

Algunos antecedentes del Periodismo de Datos se remontan a los años 60 en la consolidación del periodismo de investigación. Se desarrolló la figura del *muckraker* y los periodistas se convierten en valedores de los intereses de la audiencia. Desde el

escándalo del Caso Watergate han aparecido muchos trabajos de investigación en la recuperación y análisis de datos (Flores y Salinas, 2012).

Paul Bradshaw (2012) entiende el Periodismo de Datos como una nueva posibilidad cuando se combina el instinto periodístico y la capacidad de encontrar una historia convincente, con una magnitud y alcance informativo digital disponible. Bradshaw (2012) trata de explicar el proceso a partir de una pirámide invertida: compilación, limpieza de datos, contextualización, combinar y comunicar.



Fuente: elaboración propia a partir de Bradshaw

El Periodismo de Datos se considera una evolución del Periodismo de Precisión ya que comparten rasgos similares (Chaparro, 2013). El Periodismo de Datos usa herramientas de la estadística y de visualización (Flores y Salinas, 2012) y es necesario entrenar y aprender conocimientos de materias alejadas de las preferencias del periodista convencional (Crucianelli, 2014).

Con este avance y evolución de la tecnología, el periodista está sometido a sufrir cambios en su forma de trabajar. Así el perfil del periodista debe tener capacidad de tratar y analizar grandes bases de datos y encontrar información dentro (Flores y Salinas, 2012). El periodista necesita cuatro requisitos: interés por los números, por la programación, tener pasión por la narrativa y un entusiasmo por la expresión artística (Carlos Elías, 2015). Por lo tanto, la labor del periodista está en seleccionar,

contextualizar y jerarquizar la información que es válida e interesante, así es más fácil para el ciudadano (Elías, 2015).

2.1.1 Minería de datos

A medida que pasan los años se generan más datos. El aumento de la información se debe a la evolución de las tecnologías. Aun existiendo el *Big Data* es necesario la aplicación de nuevas tecnologías para procesar la información. Así entra en término la minería de datos.

En término anglosajón *Data Mining*, minería de datos, se originó en 1990 debido a que la tecnología de bases de datos (Daher,2013 Vallejo Ballesteros, 2018) y la disposición de los datos aumentó (Riquelme, Ruiz y Gilbert, 2006). Con la minería de datos se aprovecha la cantidad de datos que se almacenan en distintas áreas (comercio, sanidad...) y la potencia de los ordenadores para hacer las operaciones de análisis de los datos recopilados. (Vallejo Ballesteros et al., 2018). Esta tecnología ayuda a encontrar aquella información que se esconde en los datos, así se descubren patrones y tendencias que más adelante son estructuradas de un modo comprensible para su uso (ESIC, 2018).

La minería de datos ayuda al estudio de algoritmos definidos para mostrar repeticiones de modelos de los datos. Esta nueva práctica puede tener dos tareas: descriptiva o predictiva. Así su objetivo es adelantarse a las salidas y revelar relaciones en los datos (Riquelme, Ruiz y Gilbert, 2006).

Las áreas de investigación que utilizan para la minería de datos: estadística, visualización de datos, bases de datos e inteligencia artificial (Velarde Martínez, 2003). En estos campos se obtienen conocimientos con ciertos patrones de los bancos de datos que se centran en el análisis predictivo (Niño, Illarramendi, 2015).

2.1.2 Visualización y narrativa del Periodismo de Datos

Las visualizaciones de datos se encuentran en un periodo de desarrollo, convirtiéndose en una forma útil de representar la información (Domínguez, 2016). Debido a su gran utilidad, no se usa solo en medios digitales, sino que también tiene un gran papel en prensa gráfica y televisión, el Periodismo de Datos utiliza como herramienta las

visualizaciones para explicar las noticias de una manera más clara y completa (Crucianelli, 2014).

La infografía es popular desde la informatización de las redacciones a finales de los 80 y principios de los 90. Es una representación diagramática de datos. Las infografías se transforman en una herramienta para que la audiencia pueda analizar los datos. Usan la interacción para así poder mejorar la visibilidad, el *feedback*, *affordance*, restricciones y consistencia (Cairo, 2008).

El término de infografía tiene variantes en su definición. Se habla de infografía digital cuando se facilita la comprensión de acontecimientos con dosis estéticas, sustituyendo o acompañando el texto (Salaverría, Díaz, 2003). Por otro lado, está la infografía de prensa que proporciona información elaborada en el periódico escrito, hecha con elementos icónicos y tipográficos, al igual que a digital, facilita la comprensión (Valero, 2000).

En la elaboración de las infografías el periodista debe tener en cuenta una serie de características (Marín, 2015; Salaverría, 2004; Valero; 1999):

1. Utilidad: tiene que servir para algo, debe ser práctica (Marín, 2015) tiene un grado de significación, información y funcionalidad. Su parte informativa ya que responde a las preguntas del periodismo anglosajón. Cuánto más significativa más relevancia tiene la infografía. Es funcional ya que sustituye otras formas de presentación (Valero, 2003).
2. Visualidad: “característica relacionada con la forma que se presenta la infografía al lector y se hace accesible a su conocimiento” (Marín, 2015, p. 6). Tiene que ser comprensible, tener estética y ser icónica (Valero, 2003). “El lector selecciona las imágenes o los elementos que le son familiares y le atraen” (Valero, 1999, p.128).
3. Hipertextualidad: “organizar estructuras discursivas, lineales y no lineales con unidades de contenidos mediáticos” (Salaverría, 2004, p. 41). Proporciona una profundidad que amplía la información (Marín, 2015).

Siguiendo en la infografía digital se mencionan algunos modelos: guerras, eventos deportivos, catástrofes, accidentes, estudios documentales y elecciones (Salaverría, Díaz, 2003). Las visualizaciones se utilizan en una gran variedad de campos, tiene usos y objetivos muy variados. Las representaciones interactivas de grandes cantidades de datos dejan que el usuario sea quien elija lo que quiere ver (Domínguez, 2016). Así, los periodistas dan importancia al concepto de visualización de datos. Aparece como un marco conceptual que incluye la infografía dinámica y multimedia (Herrero, 2015).

Desde el aspecto narrativo, el Periodismo de Datos tenía un problema: cómo construir una narrativa cuando solo existen datos sin explicación. Es en 1967 por Barey Glaser y Anselm Straus hablan de la teoría fundamentada. Explican que es útil para crear hipótesis a partir de los datos de una investigación (Elías, 2015).

En la actualidad se acuña el término *storytelling*. Es una técnica que dispone de lógicas narrativas y herramientas que se pueden vehicular y difundir datos. El reportaje es el género rey para el Periodismo de Datos (Vázquez y Fernández, 2016).

2.2 Evolución del Periodismo de Datos en España y Ley de Transparencia

En los 90 se empezaron a crear gráficos en España. Hasta los 2000 se creó un boom de la infografía en nuestro país, eran 100% papel. En 1998 se utilizaba la herramienta *Flash* para crear elementos interactivos. En distintos medios españoles, entre ellos *El País*, crearon equipos específicos de infografistas. Un equipo era para aquellas publicaciones en papel y otras en digital. En 2008 el programa *Flash* desaparece y desencadena la emigración de algunos de los periodistas de los equipos formados (Zafra, M., 2022, comunicación personal, 9 de mayo 2022). Para crear contenidos de Periodismo datos en este momento el medio recurría a plataformas como el Consorcio Internacional de Periodistas de investigación (ICIJ) (Ferrerías, 2016).

El tardío impulso del Periodismo de Datos en España ha sido consecuencia de una legislación que regulara el acceso a los datos públicos (Martínez, 2016), hasta el 2012 España era el único país de Europa sin una Ley de Acceso a la Información (Flores y Salinas, 2012). En 2006 se aprobó la Ley 27/2006 de acceso a la información

medioambiental (Martínez, 2016). Se fundó en 2011, en España, Civio Fundación Ciudadana por dos ingenieros, Jacobo Elosua y David Cabo (Elías, 2015), esta fundación trabaja por el libre acceso a los datos públicos de la población (Blanco et al. 2016).

Como explican en su página web (civio.es) su objetivo es ayudar a mejorar la transparencia, rendir cuentas a las administraciones y promover el libre acceso a los datos públicos.

También en 2011 nace el Portal de *Datos.gob*, el portal del Gobierno bajo el marco de acciones del Proyecto Aporta. Único punto para acceder a todos los datos públicos para su reutilización (Elías, 2015). En 2013 aparece la Ley de Transparencia, Acceso a la Información Pública y Buen Gobierno que, según el BOE, donde se explica el objetivo:

“Esta Ley tiene por objeto ampliar y reforzar la transparencia de la actividad pública, regular y garantizar el derecho de acceso a la información relativa a aquella actividad y establecer las obligaciones de buen gobierno que deben cumplir los responsables públicos, así como las consecuencias derivadas de su incumplimiento” (BOE, 2013, p. 4)

Sin embargo, hasta un año después de su entrada en vigor, 2014, no se pueden hacer peticiones de información, ni consultar datos. La Ley de Transparencia coincide en el tiempo con el surgir del Periodismo de datos en los medios españoles desde finales de 2013 hasta 2014 con su desarrollo, pero hasta 2015 no se popularizó (La Rosa y Sandoval, 2016).

Con el acceso libre a los datos se evita que las informaciones periodísticas sean resultados de filtraciones o de réplica institucional (Escudero 2016, como se citó en La Rosa y Sandoval, 2016). Esta herramienta permite al periodismo dotar de instrumentos para consultar los datos que se ofrecen y descubrir la búsqueda y enfoque de los datos (Blanco, Teruel y Martín, 2016).

2.2.1 Estudio de Caso *El País*

Para el desarrollo de este apartado se ha realizado una entrevista en profundidad a Mariano Zafra, redactor jefe de Narrativas Visuales y Gráficos en *El País*¹. Con esta cita se contrasta la documentación encontrada, a su vez se actualiza y completa.

Teruel y Blanco (2016) afirman que *El País* establece su propio portal de transparencia, así los lectores podían generar preguntas y las gestionaba el medio. Además, explican que en este medio de una manera u otra siempre ha habido Periodismo de Datos.

Hasta 2016 el periódico no crea un *Equipo Express* inspirado en el *New York Time* en publicaciones de 2015. Es la denominación que se da al conjunto de trabajadores de varias especializaciones, cuyo objetivo es hilar la información con la visualización de datos. Esta iniciativa de David Alandite supone juntar varios perfiles para crear contenidos digitales con texto, gráfico y un alto grado de interactividad (Valero y Carvajal, 2017). Sin embargo, Noemí Ramírez (citado en Ferreras, 2016) confirmó que el medio no contaba con equipo específico de Periodismo de Datos.

El periódico *El País*, en 2016, forma un equipo *ad hoc* formado entre tres y seis personas, cada uno de una rama: infografistas, desarrolladores y redactores. Sus trabajos de *computational journalism* permiten visualizaciones interactivas en grandes cantidades de información a partir de un *software* de escucha activa, así el lector se permitía orientar su lectura (Teruel y Blanco, 2016).

La diferencia entre el *Equipo Express* y el *ad hoc*, está en su duración. El primero es una unidad que continua hasta la actualidad. El *ad hoc* se creaba en casos de necesidad para las noticias de datos y colaboraban con empresas especializadas.

Así, Mariano Zafra explica que hasta 2016-2017 no se creó el equipo de Periodismo de Datos con los perfiles de periodistas idóneos. Ahora el medio dispone

¹ La entrevista en profundidad se solicitó por correo electrónico. El periodista Mariano Zafra accedió a realizarla por teléfono el 9 de mayo de 2022. Fueron 10 preguntas abiertas relativas a la infografía y el equipo. Durante la conversación surgieron más temas, de los preparados, para completar el apartado del marco teórico. La cita tuvo una duración de 1 hora y 15 minutos.

con Daniele Grasso, Kiko Llaneras y Jorge Andrino, periodistas de datos. El periódico junta el grupo de narrativas visuales y de datos para así tener “una mirada más profunda”, explica Zafra.

La escasa presencia de Periodismo de Datos es por la falta de recursos económicos, medios que no ven la rentabilidad de su uso y la formación de los periodistas (Ferrerías, 2016). Para la implantación de estos perfiles en una redacción es necesario invertir y tener confianza. “No se tarda lo mismo un reportaje de datos que uno generalista”, explica Zafra (2022).

Valero y Carvajal (2017) insisten en la contratación de esta plantilla para poder contextualizar información tratando y visualizando los datos. Según el Informe Anual de la Profesión Periodística de 2020, el 70% de periodistas exige una mayor especialización. Zafra (2022) de acuerdo con esta idea recalca “al equipo de Periodismo de Datos es necesario desde siempre, pero antes no existía”.

3. Metodología

Además de la entrevista en profundidad referida en el marco teórico, para la investigación sobre la evolución y desarrollo del Periodismo de Datos en España, mediante un estudio de caso de *El País* se recurre también a la metodología de análisis de contenido.

Este método se utiliza en investigaciones para descubrir el ADN de los mensajes mediáticos. Es el proceso ideado para analizar el contenido de informaciones archivadas (Igartúa y Humanes, 2014), Además, este método es sistemático, objetivo y cuantitativo (Wimmer y Dominick, 1996, como se citó en Valero 1999).

Para el análisis de contenido se analizan 3 categorías distintas: temática, elementos multimedia y visualización de la infografía. Cada categoría tiene un total de 9 variables distintas y, en esta división, se cuenta con un total de 45 subvariables.

3.1 Muestra

La investigación parte de una muestra de N=140 noticias on-line en el medio a estudiar. La muestra temporal se obtiene del periodo comprendido en el año 2016 y termina en la

actualidad, 2022. Se elige este año de partida por la recomendación del periodista Mariano Zafra, Redactor Jefe de Narrativas Visuales y Gráficos en *El País*. Además, algunos artículos, como el de Eva Domínguez 2016, señalan el 2016 como el comienzo del equipo de Periodismo de Datos en el medio *El País*.

La selección de la muestra incluye 20 noticias de cada año. El año 2022 no se analiza entero, se recogen las noticias publicadas desde enero hasta mayo, pero aun así se han obtenido unidades representativas. La elección de las unidades de análisis se ha hecho a través de la página web de *El País*, mediante la búsqueda de artículos hechos por el equipo de Periodismo de Datos del medio. Se especifica en el año de inicio y se selecciona una primera fase de observación si tiene alguna de las características de la tabla de análisis. Una vez recopiladas las 20 noticias de cada año se estudian y los resultados se codifican en el [Excel](#) separado por distintas hojas.

3.2 Tabla de análisis

La tabla de análisis se ha creado mezclando tablas de varios estudios para hacerla completa y específica para esta investigación. El análisis se compone de 9 variables divididas en categorías: temática (C1), elementos multimedia (C2), visualización de la infografía (C3).

El primer apartado centrado en la temática (C1) cuenta con dos variables:

- **Ámbito:** se refiere a las noticias que están dirigidas a una geografía en concreta. En este caso especifica en lo nacional, España, o internacional, cualquier otro país.
- **Sección Temática:** son los espacios que reserva el medio para hablar de un tema en concreto. Formado por seis temas: política, economía, deportes, sociedad, sanidad y otros (atentados, catástrofes naturales, guerra...)

En el segundo apartado se analiza los elementos multimedia (C2) que puede tener o no la noticia analizada. Se han seleccionado tres variables distintas:

- **Hipertextualidad:** “relacionar de manera no sucesiva palabras, frases, documentos mediante hipervínculos” (Larrondo, 2009, p. 60). Y en los hipervínculos podemos diferenciar entre:
 - **Micronavegación:** aquellos enlaces que llevan al lector a una navegación interna de la web.

- Macronavegación: los enlaces que permiten al usuario visitar otros sitios web.

Existe la posibilidad de que un artículo cuente con estos dos tipos de navegación o con ninguno. Por ello se añaden las variables de ambas o de nada.

- Base de datos: es la referencia donde se han encontrado los datos que muestra la noticia. Estas bases pueden ser varias:
 - Excel: un documento de Office Excel donde están todos los datos en bruto
 - CSV: archivo en el cual los datos se guardan en un formato específico, suele ser en forma de tabla
 - Web: los datos se han encontrado en una URL específica como puede ser las estadísticas del INE
 - PDF: documento de adobe en el que se encuentran los datos de la noticia
 - Nada: es posible que algunos artículos no señalen de donde han sacado los datos
- Multimedia: la publicación reúne distintos medios comunicación mezclados. Se han seleccionado cuatro:
 - Imágenes: fotografías que se encuentren en la noticia, son estáticas
 - Vídeo: grabación de imágenes o de explicaciones, en movimiento
 - Sonidos: elementos de audio, suele ser en formato de podcast
 - Infografías: elementos visuales que ayudan al lector a comprender los datos que se explican en el texto

Por último, en la última categoría de análisis la visualización de la infografía (C3). Esta categoría está formada con cuatro variables:

- Presentación: hace referencia a que herramienta se utiliza para explicar los datos. Se han seleccionado ocho:
 1. Gráfico: representación de datos a partir de una estructura
 2. Diagrama: visualización de una variación entre varios elementos

3. Mapa: presentar un territorio de manera gráfica con las características de un territorio con aspectos importantes
 4. Tabla: conjunto de datos colocados en filas y columnas
 5. Línea del tiempo: recta dividida para explicar la evolución de algo en concreto en una sucesión de años
 6. *Fotoscrolling*: cambio sucesivo de imágenes estáticas mientras se avanza en la lectura
 7. *Videoscrolling*: cambio de grabaciones al mover el cursor por la lectura
 8. Otros: comprenden aquellas presentaciones que no están mencionadas o no se sabría clasificar
- Exposición de datos: una vez elegida la presentación hay que especificar en cómo aparecen los datos explicados. Se representan en barras, sectores, línea, pirámide, puntos u otros
 - Según su contenido (Mier, 2016, citado en Juan Díez, 2019): según el contenido que narra se diferencia entre:
 1. Arquitectónica: mostrar lo que tiene una instalación, inmueble...
 2. Informativa: ofrece datos que aclaran la información con gráficas
 3. Didáctica: enseñar el funcionamiento de algo
 4. Acontecimiento: explica un hecho con visualizaciones
 5. Personaje: parten de una entrevista y se muestra al personaje
 6. Geográfica: información en mapas sobre población o en relación a la geografía
 7. Híbridas: parten de un hilo conductor y cuando con varias de las mencionadas. La más completa.
 - Para la variable de interactividad se ha diferenciado entre tres opciones. Las denominaciones se deben a la entrevista realizada a Mariano Zafra. Explicó que a la hora de empezar una infografía pensaban primero en su interactividad. Si iba a ser *clickativo*, es decir, el usuario no puede variar la visualización, pero puede concretar los datos que le interesan. Y algo es interactivo cuando el lector puede cambiar, buscar y jugar con los datos. También es posible que no exista interactividad.

Tabla 1. Ficha de análisis de informaciones periodísticas de Periodismo de Datos

C.1 TEMÁTICA	Ámbito	<ol style="list-style-type: none"> 1. Nacional 2. Internacional
	Sección temática	<ol style="list-style-type: none"> 1. Política 2. Economía 3. Deportes 4. Sociedad 5. Sanidad 6. Otros (catástrofes naturales, atentados...)
C.2 ELEMENTOS MULTIMEDIA	Hipertextualidad	<ol style="list-style-type: none"> 1. Micronavegación 2. Macronavegación 3. Ambas 4. Nada
	Base de datos	<ol style="list-style-type: none"> 1. Excel 2. CSV 3. Web 4. PDF 5. Nada
	Multimedia	<ol style="list-style-type: none"> 1. Imágenes 2. Vídeos 3. Sonidos 4. Infografías
C.3 VISUALIZACIÓN DE LA INFOGRAFÍA	Presentación	<ol style="list-style-type: none"> 1. Gráfico 2. Diagrama 3. Mapa 4. Tabla 5. Línea del tiempo 6. <i>Fotoscroling</i> 7. <i>Videoscroling</i> 8. Otros
	Exposición de datos	<ol style="list-style-type: none"> 1. Barras 2. Sectores 3. Línea 4. Pirámide 5. Puntos 6. Otros
	Contenido	<ol style="list-style-type: none"> 1. Arquitectónico 2. Informativa 3. Didáctica 4. Acontecimiento 5. Personaje 6. Geográfica 7. Híbrida
	Interactividad	<ol style="list-style-type: none"> 1. Nada 2. <i>Clickativo</i> 3. Interactivo

Fuente: elaboración propia a partir de Palau-Sampio y Sánchez-Pilar (2019, citado en Juan Díez, 2019)

4. Resultados

Tras analizar las 140 unidades de análisis extraídas de la web digital *El País*, se presentan a continuación los resultados ordenados por categorías y teniendo en cuenta sus variables. Los resultados de la C1, Temática, con dos gráficos, uno para mostrar el ámbito y otro la sección temática. Las deducciones de la C2, los Elementos Multimedia, uno para la hipertextualidad, otro para la base de datos, la multimedialidad. Y un cuarto para explicar un hallazgo en las variables. Por último, se explican los resultados de la C3, Visualización de la Infografía, con sus respectivos gráficos a la presentación, exposición de datos, contenido e interactividad, en esta categoría se amplía a dos gráficos fuera de las variables, así se podrá explicar una novedad observada

Para la exposición de los resultados se usan gráficos siguiendo las tablas obtenidas en el análisis de los datos guardadas en el [Excel](#). Se representa una comparación final de las variables y ayudan a mostrar de una manera más visual los resultados.

Los gráficos de la investigación son de varios tipos: barra, línea o circulares, depende de la variable que se explique, las herramientas que se han utilizado para su elaboración son *Datawrapper* y *Flourish*

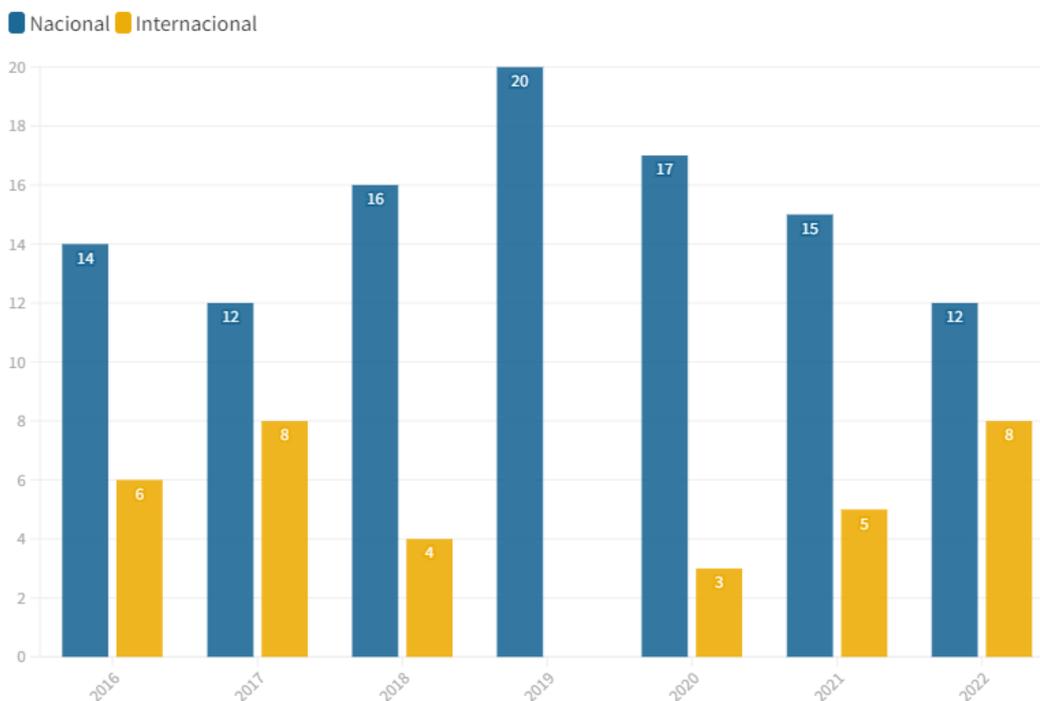
4.1 Resultados de Temática (Ámbito / Sección temática)

El análisis de la categoría Temática (C1) refleja cómo la mayoría de las informaciones de Periodismo de Datos analizadas en *El País* se centran en un 75% en el ámbito nacional y un 25% en noticias internacionales.

Durante los siete años analizados también se observa una mayor publicación de ámbito nacional, la media es de más del 50% cada año. Resulta destacable que, en el año 2019 de las 20 noticias analizadas, el 100% son de ámbito nacional. A medida que van pasando los años se observa que las noticias internacionales no superan el 40% en ningún período analizado

Gráfico 1: Análisis del ámbito territorial de las informaciones de Periodismo de Datos en *El País* (2016-2021)

ÁMBITO



Fuente: elaboración propia con Flourish

En cuanto a la sección temática se observa, con el paso de los años, la variedad de asuntos centrados en base de datos. La gran diferencia entre 2016 y 2022 es el uso de más contenidos. En 2016, de las 20 muestras analizadas 13 son de Política, el 65%. Sin embargo, en 2022 existe un equilibrio de las publicaciones, cada sección ronda en el 25%.

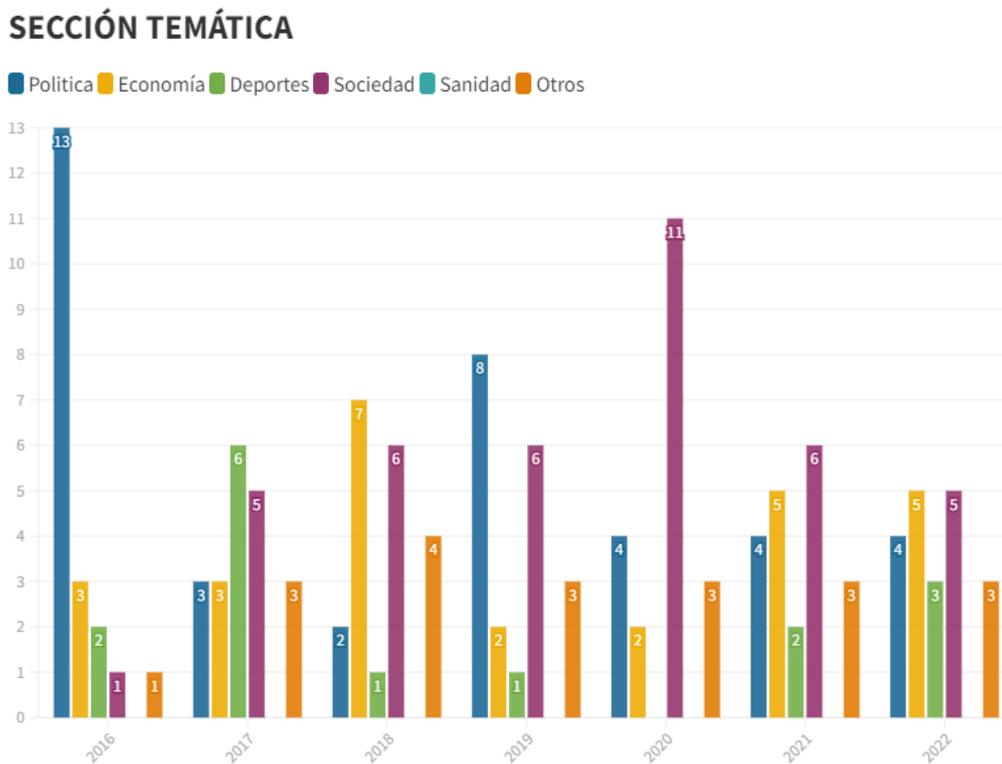
Para aclarar el dato de Sociedad en el año 2020 que es un 55% de las noticias debido a la COVID-19. Este resultado lo confirma el propio responsable de *El País* entrevistado para este trabajo Mariano Zafra quien asegura que las publicaciones de pandemia se catalogan en la sección Sociedad.

La sección centrada en Economía va ascendiendo cada año, de un 15% en 2016 pasa a un 35% de las publicaciones en 2018. En los siguientes años se regula al 25% con 5 noticias por año. La muestra analizada de ‘otros’ evoluciona de existir una noticia en 2016, el 5%, a publicarse 3, el 15%, en el último año. Respecto a otras temáticas, cabe destacar que en los siete años analizados han aumentado un 10% las noticias sobre

catástrofes naturales, terrorismo... La sección de Deportes pasa de un 10% con dos publicaciones anuales a un 30% que son 6 noticias por años.

De manera jerárquica los resultados, respecto a la muestra total N=140, en los últimos siete años, encontramos en primer lugar Sociedad con 40 publicaciones, un 28'57%, seguido de Política con 38 noticias, un 27'24% y tercero la Economía con 27 apariciones, 19'29%. Fuera del rango de los tres primeros continua 'otros' con 20, el 14'39% y Deportes con 15, 10'71%.

Gráfico 2: Análisis de la sección temática de las informaciones de Periodismo de Datos en *El País* (2016-2021)



Fuente: elaboración propia con Flourish

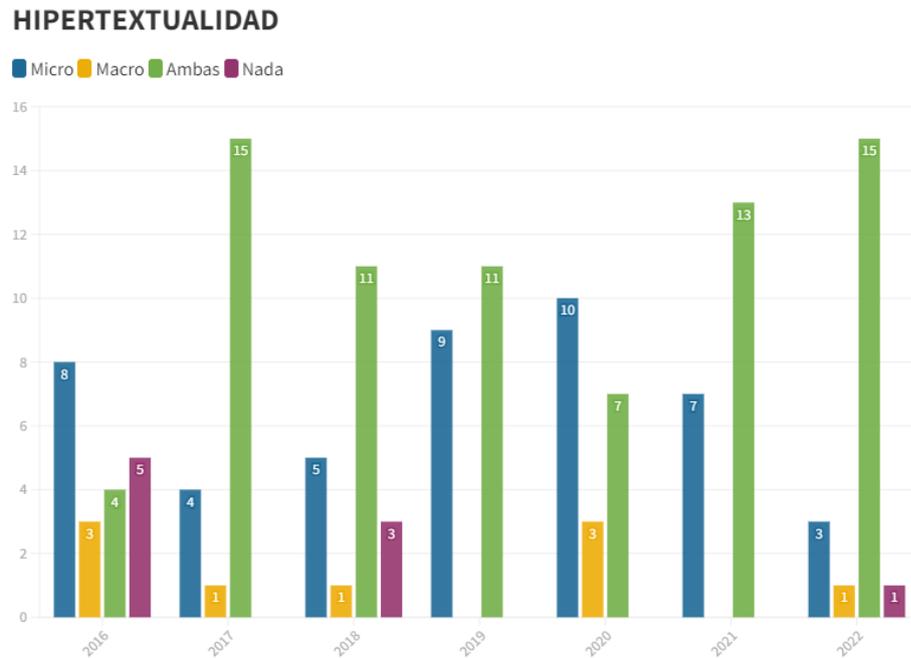
4.2 Resultados de elementos multimedia (Hipertextualidad/Base de datos/Multimedia)

Tras realizar el análisis de los enlaces de las noticias se discurre que el medio realiza gran uso tanto de micro como de macro navegaciones. Se han encontrado 76 noticias con estos dos tipos de navegaciones, es decir, el 54'28% permite al usuario aumentar la información. No obstante, el 35'85%, es decir, 46 informaciones en las que solo se utiliza la micronavegación y provoca que el usuario no salga de la web.

De una manera gradual las informaciones de datos en *El País* se van haciendo más completas, creando más noticias con ambas navegaciones. Respecto a la macronavegación, la capacidad de enlazar a otras webs, en el análisis de noticias de Periodismo de Datos existe un desuso. Esta variable en vez de ir en aumento, se ha reducido empezando a tres noticias, que es el 15%, en 2016 y disminuye hasta el 5% en el último año, 2022, que es una noticia solo con macronavegación.

En cambio, la micronavegación se ha regulado con el tiempo. Empieza con un 40%, son 8 publicaciones en 2016, subió hasta aparecer en el 50% de la muestra anual, que es 20, en 2020. En la actualidad forma el 15% de las noticias.

Gráfico 3: Análisis de la hipertextualidad en las informaciones de Periodismo de Datos en *El País* (2016-2021)



Fuente: elaboración propia con Flourish

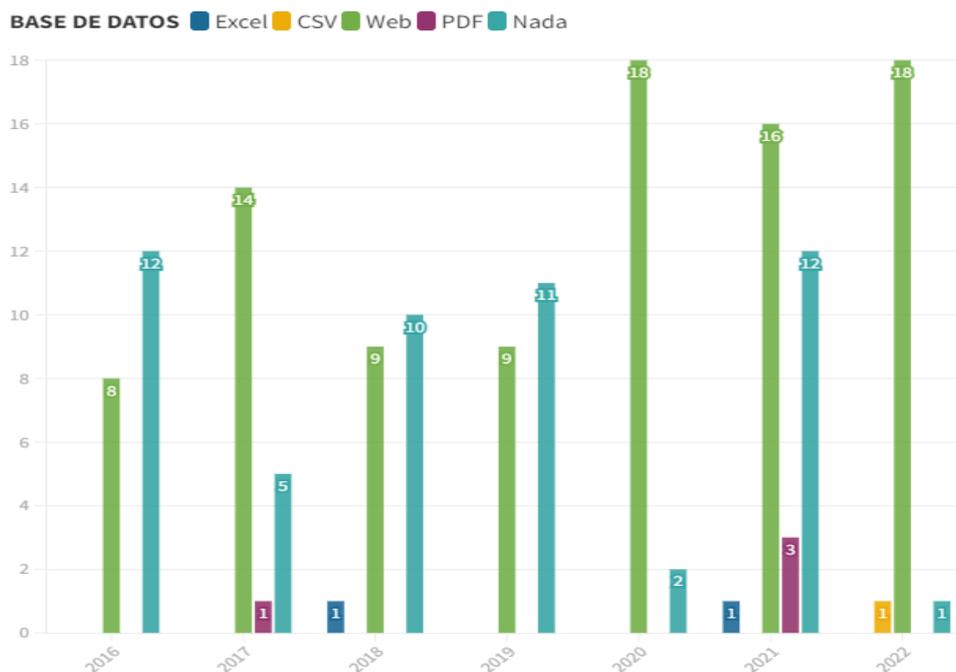
Una vez estudiado los enlaces que utiliza el periódico, observamos la procedencia de los datos de las noticias, una cuestión relevante para este estudio si tenemos en cuenta que el Periodismo de Datos se caracteriza también la transparencia y ofrecer la base de datos. Permitir al usuario ver los datos en bruto. Con el análisis realizado se observa que esta característica incrementa con los años.

En los primeros años, prima la carencia de los datos en bruto. En 2016 son 12 noticias las que no redirigen al lector a la base de datos, es decir, más de la mitad de las noticias de ese año analizadas, un 60%. En 2017 se enlaza en 1 sola noticia a un PDF, el 5%, y en 2018 también fue un 5% enlazado a un documento Excel.

El desarrollo de esta aportación se observa en los dos últimos años, en 2021 ya encontramos varias publicaciones con PDF, el 15% y un 5% con Excel. Sin embargo, en la actualidad solo una deriva a CSV, 5%, y prima las referencias de web con el 90% de las noticias de 2022. Con el paso del tiempo se puede ver que son menos las noticias que tienen ausencia de su fuente de datos. Esta evolución pasa de 12 noticias sin base de datos a 1 sola, es decir, se ha reducido el 55% de las publicaciones que no muestran su procedencia de datos.

Por lo tanto, al establecer un orden de mayor a menor uso, encontramos en primer lugar con 92 noticias, 65'71%, donde su base de datos es una web. Aun así, son 53 noticias, corresponde al 37'86%, que no enlazan o señalan la fuente de sus datos. Seguido de las noticias con PDF que son cuatro, el 2'86%, con Excel encontramos dos, 1'43%, y con CSV, la más actual, solo una, 0'71%

Gráfico 4: Análisis de las bases de datos en las informaciones de Periodismo de Datos en El País (2016-2021)

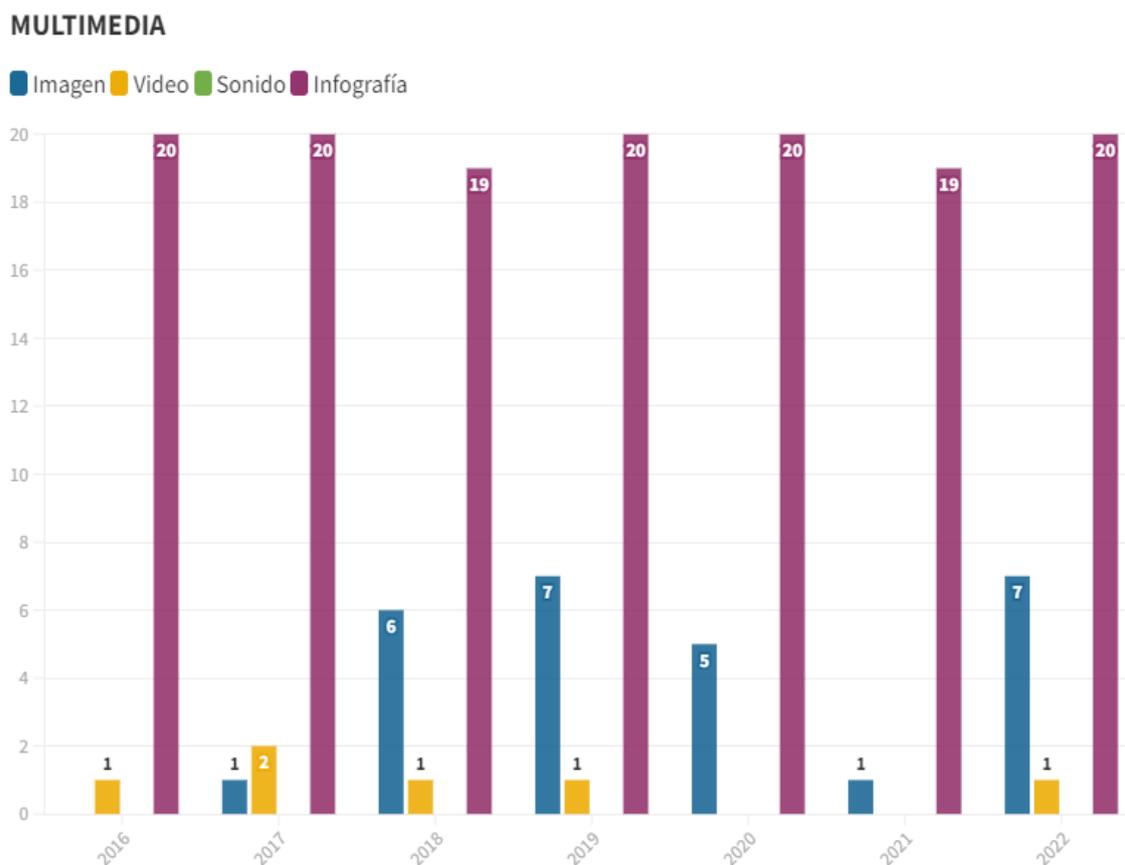


Fuente: elaboración propia con Flourish

Respecto a la narrativa multimedia, los datos encontrados se explican desde dos resultados distintos, se ha dado la situación de noticias con más de un elemento multimedia: imágenes, vídeo, sonido e infografía, por lo tanto, para poder explicar estos datos se han realizado dos gráficos.

El Gráfico 5 muestra la totalidad de los elementos multimedia en las noticias. Al estar centrado el análisis en noticias de datos, el elemento principal es la infografía, por lo tanto, el 98'57% de las publicaciones cuentan con el elemento principal visual. De manera individual el uso de las imágenes en las noticias de datos, asciende de un 5%, con una sola noticia, a un 35% que son 7 noticias con este elemento multimedia. Sin embargo, el uso de vídeos no asciende y se mantiene en un 5% durante todos los años.

Gráfico 5: Análisis de los elementos multimedia en las informaciones de Periodismo de Datos en *El País* (2016-2021)



Fuente: elaboración propia con Flourish

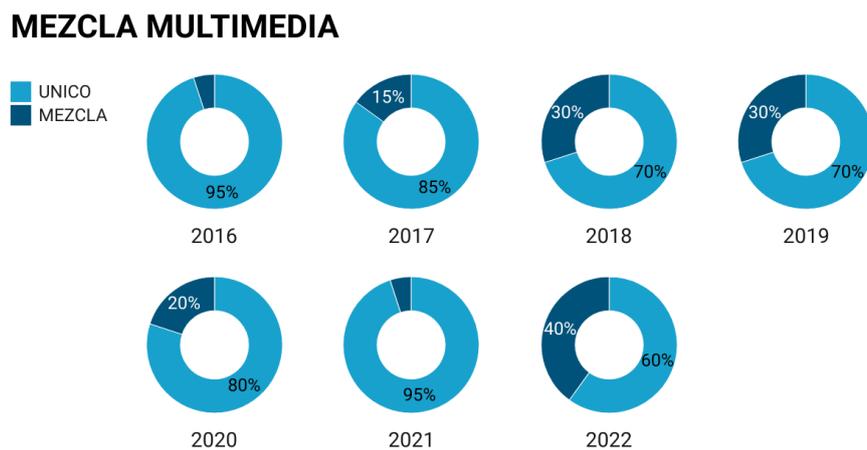
El Gráfico 6 muestra cuántas noticias, de las 140, solo tienen infografía y cuántas se complementan con otros elementos audiovisuales cada año. El contraste de la

‘mezcla’ que hace referencia a las noticias que tienen más de un elemento multimedia y ‘único’ a las noticias que solo tienen infografía.

Con esta diferenciación se observa mejor el aumento de elementos multimedia en las noticias. El Periodismo de Datos en el medio empezó en 2016 con un 95% de noticias eran solo con infografía. En los años consecutivos se empieza a mezclar la infografía con otros elementos. En 2018 y 2019, se alcanza el 30%, con 6 publicaciones con algún vídeo o imagen además de la visualización de datos.

Así, el último año, 2022 es el que más variedad de elementos multimedia obtiene con 8 noticias, que representan un 40% con algún elemento más aparte de la infografía.

Gráfico 6: Análisis de la mezcla multimedia en las informaciones de Periodismo de Datos en *El País* (2016-2021)



Source: FUENTE PROPIA • Created with Datawrapper

Fuente: elaboración propia con *Datawrapper*

4.3 Visualización de la infografía (Presentación/Exposición de datos/Contenido/Interactividad)

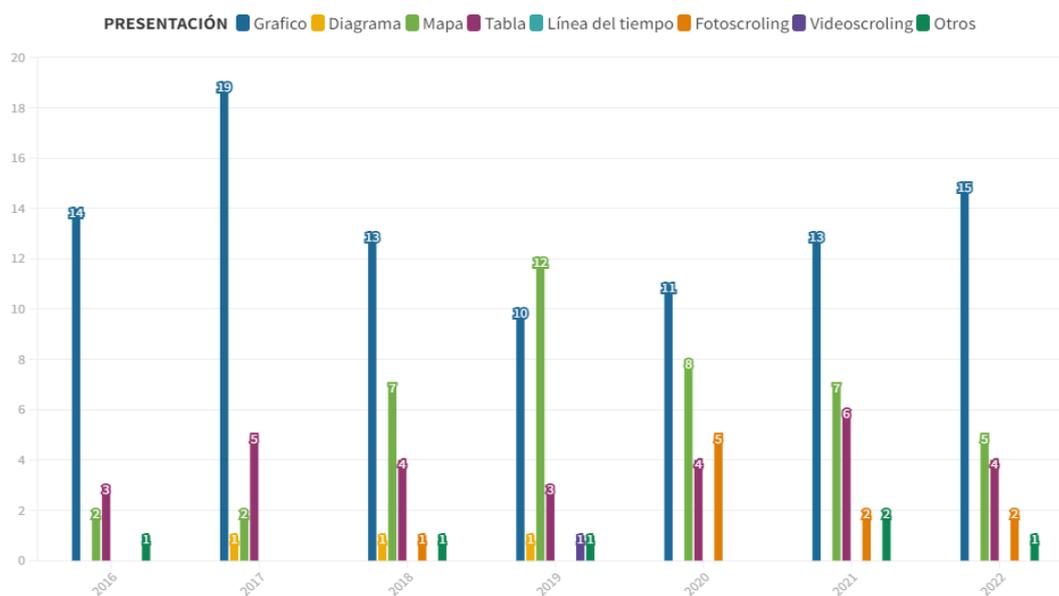
Una vez expuesta la variable de noticias con infografía, se explica su presentación por tratarse de una característica de relevancia en el Periodismo de Datos. En esta observación ocurre lo mismo que con el análisis multimedia, se ha dado con noticias que presentan los datos con distintos tipos. Por ello se explica primero de manera general el uso de cada presentación (Gráfico 7) y después se explica y analiza las noticias en las que se han encontrado más de un elemento de presentación (Gráfico 8).

De manera evolutiva en el gráfico se observa que con los años el medio ha ido mezclando e incluyendo variedad de métodos de visualización de datos en las noticias. En un primer momento, en 2016 y 2017, prima el uso del gráfico, 70% y 95% de manera respectiva, con alguna variación del uso de las tablas entre un 15% y un 25%. El año más equilibrado es 2019 donde hay una gran variedad, el uso de mapas con un 60%, se encuentra en 12 publicaciones. Usa el gráfico 10 veces, un 50%. Y las tablas en 3 publicaciones con el 15%. Diagrama, *Videoscroling* y otros tiene cada uno un 5% de uso.

Precede el año 2021 como el más completo. Un 65% de sus noticias emplean gráficos. En 7 publicaciones usa mapas, un 35%. Las tablas son visualizadas en el 30% de las noticias de la muestra anual. *Fotoscroling* y otros rondan en el 10%.

De manera general se observa cómo el gráfico y el mapa se utiliza en todos los años, ya sea en mayor o menor medida, aparecen. En datos totales el gráfico se emplea en 95 noticias, un 67'86%, y el mapa en 43 publicaciones, 30'71%. En tercer lugar, se encuentra la tabla con 29 veces, corresponde al 20'71%. El *Fotoscroling*, es la presentación más moderna, ocupa solo el cuarto lugar con 9 usos, 7'14%.

Gráfico 7: Análisis de presentación en las informaciones de Periodismo de Datos en *El País* (2016-2021)



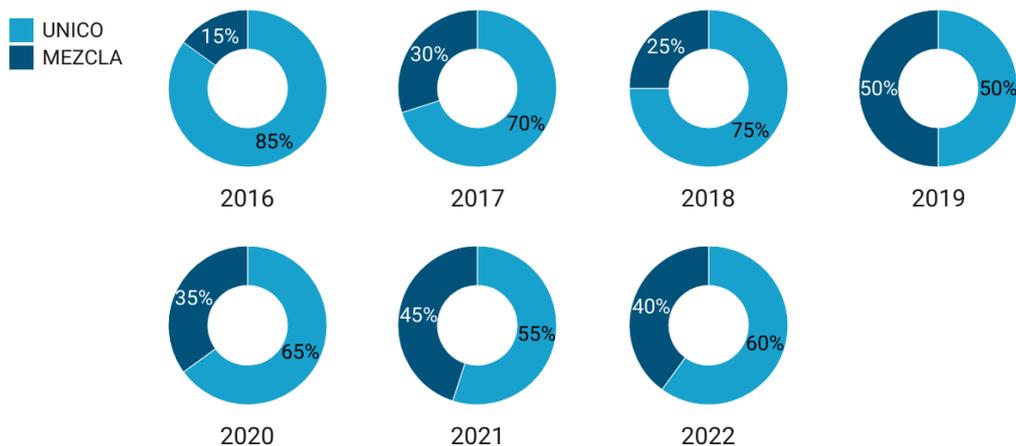
Fuente: elaboración propia con Flourish

El resultado final en los datos se observa que la suma de todos supera los 140 de la muestra, esto se debe porque en varias publicaciones se han encontrado variaciones: uso de gráfico y mapa, tabla con mapa y otros.

A continuación, se analiza cómo con los años el porcentaje de noticias que utilizan más recursos de presentación va en aumento. El año 2019, como se explica antes, es el más completo con la mitad de las noticias, 50%, aparece más de un elemento. En los primeros años, con un 15% en 2016 hasta un 25% en 2018, el medio opta siempre de la misma manera para explicar los datos. En los dos últimos años, 2021 y 2022, es muy parecido rondan el 40% y 45% de mezcla de presentación.

Gráfico 8: Análisis de mezcla de presentación de los datos en las informaciones de Periodismo de Datos en *El País* (2016-2021)

MEZCLA PRESENTACIÓN



Source: FUENTE PROPIA • Created with Datawrapper

Fuente: elaboración propia con *Datawrapper*

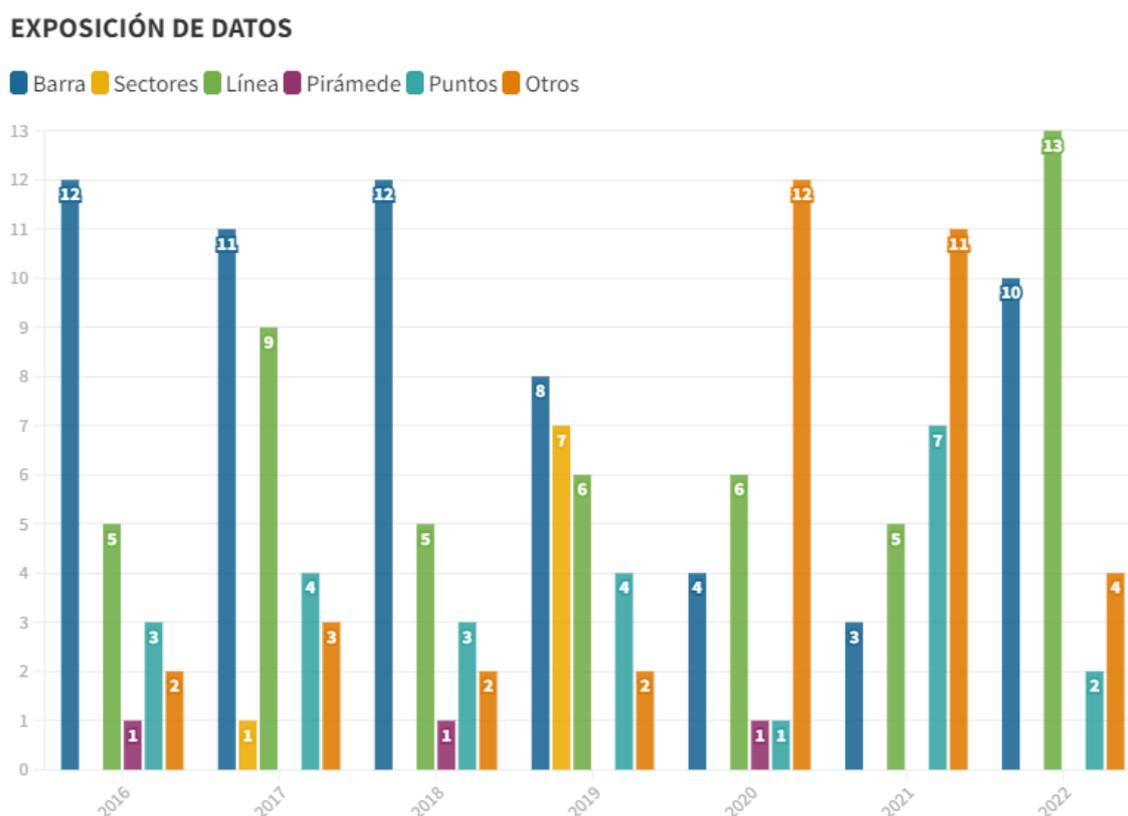
Los datos están presentados de diversas formas, como se ha observado. Pero además se analiza la manera en la que están expuestos, es decir, qué elemento gráfico utilizan para mostrar los datos. La exposición de la información es muy voluble en todos los años (Gráfico 9).

En los tres primeros años, 2016, 2017 y 2018, donde las opciones de creación eran básicas, prima el uso de las barras. En esos tres años las usan en 56'66%, es decir, 35 de las 60 muestras de esos años, se expone en barras. La presentación de datos en 2019 es el momento de más variación de exposiciones. Utiliza cada variable, la barra en

un 40%, el uso de sectores en un 35%, las líneas con un 30%, los puntos en un 20% y otros tipos de exposiciones un 10%.

En los últimos años se nota el desarrollo de las herramientas digitales ya que los datos de la opción ‘otros’ (terrorismo, catástrofes naturales) aumenta de manera considerada. Evolucionan de un 10%, en 2016, de las noticias a un 20%, en 2022. Esta sección cubre aquellos gráficos novedosos que no tienen un término claro, o algunos que mezclan. Se han encontrado representaciones de temperatura con puntos. También visualizaciones de puntos con barras. La exposición en línea del mismo modo se desarrolla de manera exponencial, al principio con 5 noticias, un 25%, aumenta hasta en 13 publicaciones, un 65%, con lo que crece un 37%.

Gráfico 9: Análisis de la exposición de datos en las informaciones de Periodismo de Datos en *El País* (2016-2021)



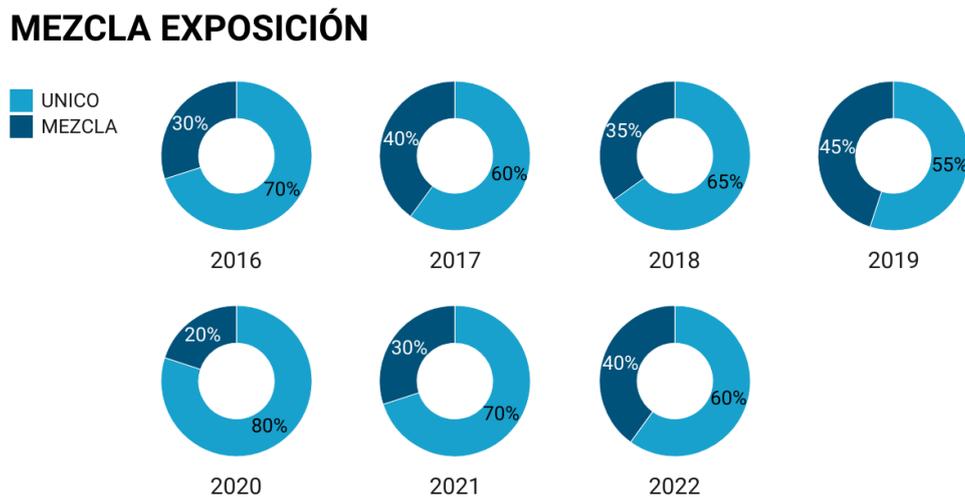
Fuente: elaboración propia con Flourish

En este análisis también se da la situación de mezcla que se observa en las dos variables anteriores. Hay noticias que usan un gráfico de barras y además añaden un mapa. Por ejemplo, para explicar unas elecciones muestran los votos por provincia y además en un

gráfico de barras, explican cómo ha quedado cada partido político. Para mostrar estas noticias con variedad de exposiciones se compone otro análisis (Gráfico 10).

La mezcla de modelos de visualizaciones en una noticia se observa en todos los años. La media de la combinación está en el 40% de las publicaciones. El año en el que menos se varia es en el 2020, solo un 20% son noticias con variedad de exposiciones. El 2019 sigue destacando sobre el resto con un 45% de publicaciones, con mezcla de tipos de muestras. Los dos últimos años, que cuentan con más variedad digital, obtienen un 30%, en 2021, y un 40%, en 2022.

Gráfico 10: Análisis de la mezcla de exposición de las informaciones de Periodismo de Datos en *El País* (2016-2021)



Source: FUENTE PROPIA - Created with Datawrapper

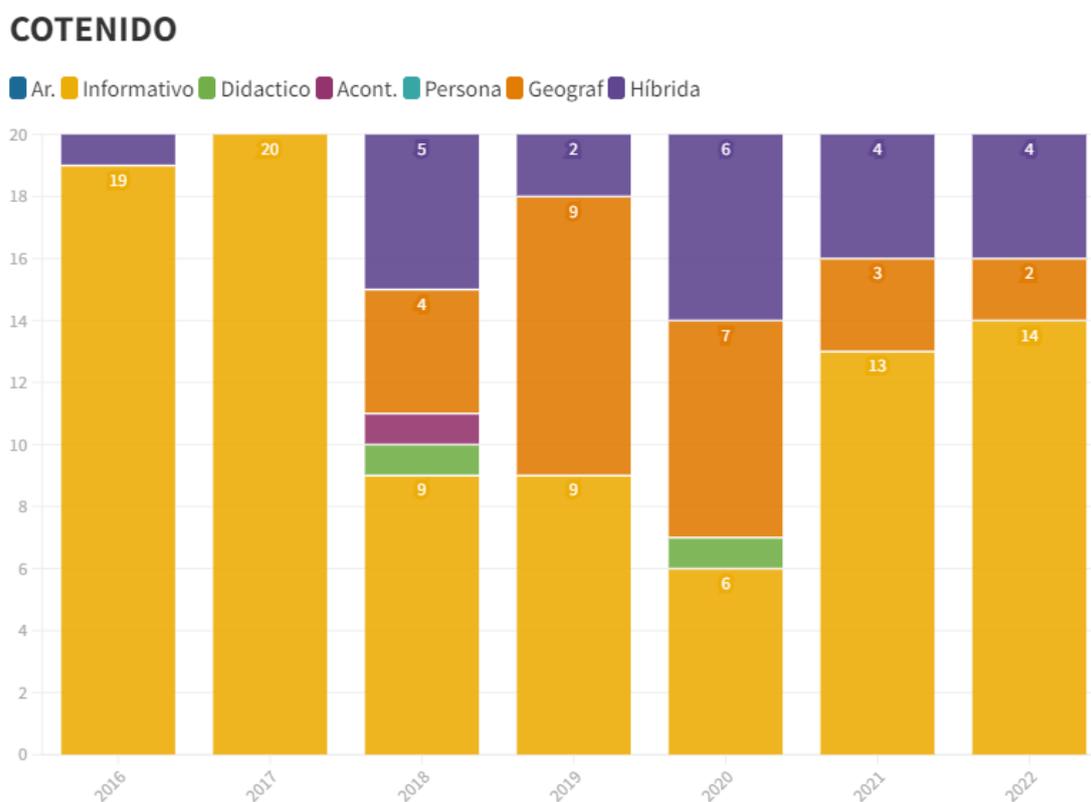
Fuente: elaboración propia con Flourish

A continuación, se muestran los resultados de la tipología de las infografías. La variable de contenidos abarca: arquitectónico, informativa, didáctica, acontecimiento, personaje, geografía e híbrida. En el análisis (Gráfico 11) se observa una mayoría por el contenido informativo con un 64'29% de noticias (90). El segundo más usado es la geografía con un 17'86%, 25 noticias de la muestra total. Y el tercero es la variable híbrida con un 15'71%, 21 noticias en los siete años.

A medida que pasan los años las noticias de Periodismo de Datos desarrollan variados contenidos. En el primer año, 2016, solo se usa un 28'57%, dos tipos de

variables, en 2018 aumenta hasta el 71'42% variando hasta en 5 contenidos. Se encuentran contenidos informativos, didácticos, acontecimiento, geografía e híbridos. En 2020 también existe pluralidad con un 57'14%, usando 4 tipos distintos.

Gráfico 11: Análisis del contenido de la infografía en las informaciones de Periodismo de Datos en *El País* (2016-2021)



Fuente: elaboración propia con Flourish

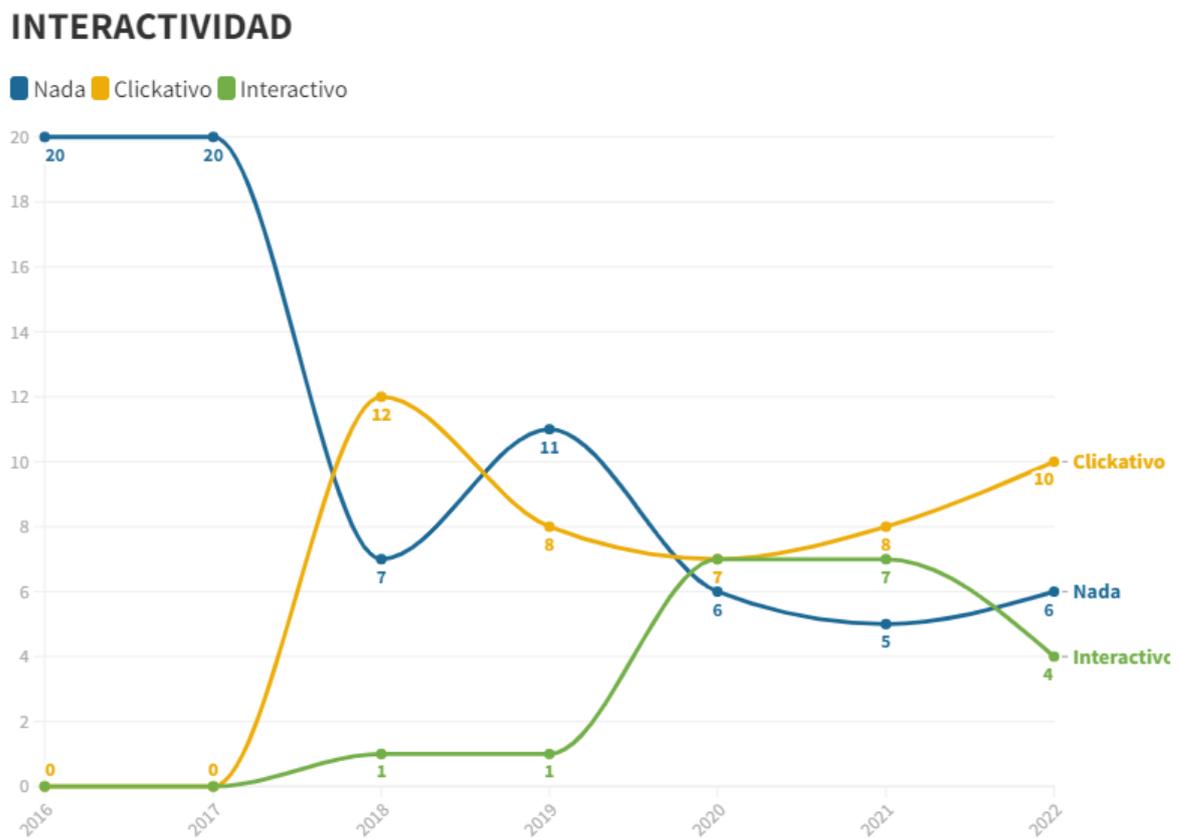
Por último, se explica la interactividad. Esta opción de las infografías muestra de manera la evolución que tiene el medio al respecto. Formado por tres variables: *clickativo*, interactivo y nada. El Gráfico 12 de líneas muestra el crecimiento y el descenso de cada una a lo largo de los siete años analizados.

En los dos primeros años, 2016 y 2017, la interactividad es nula. El 100%, de las 40 informaciones, en esos dos años no tenía interactividad. A partir de 2018, la interactividad *clickativa* aumenta al 60%, son 12 noticias. Las que no tienen interactividad desciende un 65%, porque en 2018 hay solo un 35% de publicaciones sin interactividad. El año 2020 es el más completo respecto a esta variable. Hay el mismo número de publicaciones con participación *clickativo*, 35% y con opción interactivo,

35%, 7 noticias con cada interactividad. Y las noticias sin interactividad desciende en 2020 hasta el 30%.

A pesar del crecimiento de la interactividad, en los siete años de análisis prima la carencia de este *feedback* con un 53'57% de la muestra total, n=140. La opción de noticias con infografías de *Clickativo* es de un 32'14% y la capacidad de interactivo, de cambiar el gráfico es de un 14'29%.

Gráfico 12: Análisis de la interactividad de las informaciones de Periodismo de Datos en El País (2016-2021)



Fuente: elaboración propia con Flourish

5. Discusión y conclusiones

En el presente trabajo se obtienen las siguientes conclusiones que muestran la evolución y desarrollo del Periodismo de Datos en el medio *El País*.

Cumple con el objetivo principal de verificar el progreso y el avance de esta especialización periodística en el estudio de caso *El País*. Por lo tanto, se da también respuesta a los objetivos secundarios. El primero (O1) determina las temáticas

principales en el campo de la información de datos, Social y Política. Al observar en detalle la visualización y la narrativa de datos el segundo objetivo (O2) permite demostrar la evolución de esta novedosa manera de plasmar la información. El tercer objetivo (O3) responde al tipo de interactividad que predomina en el medio, que según el análisis continua en desarrollo.

En relación a la confirmación o refutación de las hipótesis, la primera (H1), que concierne a si la expansión de Periodismo de Datos ha crecido en la prensa de España en la última década, queda confirmada. En los siete años seleccionados se observa un crecimiento en esta especialización, durante los primeros años de búsqueda las publicaciones eran escasas y en los dos últimos la elección ha sido más variada.

En la segunda hipótesis (H2), donde se expone que la narrativa de datos en *El País* se centra en temas políticos y económicos, queda refutada. Como se ha podido observar las noticias que priman los datos son los de Sociedad (40) y Política (38). En la entrevista realizada a Mariano Zafra se le planteaba esta hipótesis, para él las temáticas más repetidas eran las de Ciencia y la de ámbito Internacional. Por lo tanto, la idea planteada por el entrevistado queda también refutada ya que no se han encontrado informaciones del campo de la ciencia y el análisis demuestra que hay más publicaciones nacionales, un 75% de la muestra.

La tercera hipótesis (H3), se refiere a si la infografía utilizada en *El País* tiene interacción, *feedback* y prima lo visual, se refuta en parte. En los siete años analizados abunda lo visual, por esta parte queda confirmada. Por otro lado, está la interacción y el *feedback* no que está presente en todos los años, hasta 2018 no aparecen estas características, esta carencia representa el 53'57% de la muestra, de esta manera se rebate la infografía consta de interactividad y reciprocidad.

Una vez respondidas las hipótesis y mostrar los objetivos, se añaden algunas conclusiones concretas a partir de los resultados y la observación de la muestra. Se ha logrado responder a las preguntas de investigación, en primer lugar, el Periodismo de Datos se encuentra en pleno desarrollo y auge en los medios de España, lo que ha provocado una mayor variedad temática, así se responde a las dos primeras cuestiones. Respecto a la incógnita de las herramientas utilizadas, en el estudio existe una mayor

aplicación y dedicación a las visualizaciones, provocado por la mejora de la interactividad y de la mezcla de distintas formas de presentar los datos, se puede observar un avance y variedad de aplicaciones. Además, con la evolución de los años y la mejora de la muestra de datos, permite a los usuarios un mayor entretenimiento y participación en la información.

La presente investigación cuenta con ciertas limitaciones como el análisis de un solo medio y una muestra limitada. Se considera un punto de partida para futuros estudios que podrían centrarse en una comparación entre varios medios audiovisuales, como el de Antena3 y RTVE, ya que cada vez son más informaciones las que recurren a distintas formas de plasmar los gráficos, con hologramas o realidades en 3D.

6. Bibliografía

- Arcila, C., Barbosa, E. y Cabezuelo, F., (2016). Técnicas big data: análisis de textos a gran escala para la investigación científica y periodística. *El profesional de la información*, 25 (4). 623-631.
- Arias, D., Sánchez-García, P. y Redondo, M. (2018). El perfil del periodista de datos en España: Formación, fuentes y herramientas, *Doxa Comunicación*. 27. 273-293.
- Blanco, E., Teruel, L., Martín, V., (2016). *El Periodismo de datos en España. Más calidad y valor añadido a la información*. En Periodismo de datos, Blanco Castilla & M. Quesada (Coords.), CAC112 (pp. 13–37). La Laguna (Tenerife). Latina.
- Bradshaw, P. (2019, 22 mayo). *¿Qué es el periodismo de datos?* Fundación Gabo. <https://fundaciongabo.org/es/que-es-el-periodismo-de-datos>
- Cairo, A. (2008). *Infografía 2.0: visualización interactiva de información en prensa*. Almut ediciones.
- Chaparro, M. A. (Ed.). (2013). *La evolución del periodismo de precisión: el blog de The Guardian sobre periodismo de datos*. Actas del I Congreso Internacional Comunicación y Sociedad. Logroño.
- Crucianelli, S. (2014, 4 febrero). *¿Qué es el periodismo de datos?* Cuadernos de Periodistas. <https://www.cuadernosdeperiodistas.com/que-es-el-periodismo-de-datos/>
- Dader, J.L, y Gómez, P. (1993). Periodismo de Precisión: Una nueva metodología para transformar el periodismo, *Análisis* (15), 99-116.
- Díaz-Noci, J y Salaverría. (2003). *Manual de redacción ciberperiodística*. Ariel Comunicación.
- Díez Regidor, Juan (2019). *Análisis de la narrativa multimedia en el Periodismo de Datos. Estudio del caso de El Confidencial y El Mundo*. Trabajo de Fin de Grado en Periodismo. Universidad de Valladolid
- Domínguez, E., (2016). *El reto narrativo de integrar datos en los formatos audiovisuales e interactivos*. En Periodismo de datos, Blanco Castilla & M. Quesada (Coords.), CAC112 (pp. 87–105). La Laguna (Tenerife). Latina
- Elías Pérez, C. (2015). *Big Data y periodismo en la sociedad red*. Síntesis.
- Ferreras, E. M. (2016). El periodismo de datos en España. *Estudios sobre el mensaje periodístico*, 22(1), 255–272. <https://revistas.ucm.es/index.php/ESMP/article/view/52594>

Hernández Pantoja, Patricia (2022). *Evolución y desarrollo del Periodismo de Datos en España. Estudio del caso EL País*. TFG. Universidad de Valladolid

- Flores, J. y Salinas, C. (2013). El periodismo de datos como especialización de las organizaciones de noticias en internet. *Correspondencia y análisis* 3. 15-34
- Herrero Solana, V., & Rodríguez-Domínguez, A. M. (2015). Periodismo de datos, infografía y visualización de la información: un estudio de El País, El Mundo, Marca y El Correo. *Universitat de Caralunya*. 34. 1-10
- Igartua, J. J., & Humanes, M. L. (2004). El método científico aplicado a la investigación en comunicación social. *Journal of health communication*, 8(6), 513-528.
- Larrondo, A., (2009). La metamorfosis del reportaje en el ciberperiodismo: concepto y caracterización de un nuevo modelo narrativo. *Comunicación y Sociedad*. 22 (2). 59-88
- La Rosa, L. y Sandoval, T., (2016). La insuficiencia de la Ley de Transparencia para el ejercicio de Periodismo de datos en España. *Revista Latina de Comunicación Social*. 71. Pp 1208-1229
- Ley de Transparencia, Acceso a la Información pública y buen gobierno 19 de 2013. Se transforma la ley de transparencia, acceso a la información pública y buen gobierno. 9 diciembre 2013. D.O.No. 295
- Marín Ochoa, B.E. (2010). *La infografía digital, una nueva forma de comunicación*. En Congreso Euro-Iberoamericano de Alfabetización Mediática y Culturas Digitales Sevilla: Universidad de Sevilla.
- Martínez, L. (2016). *Plataformas y proyectos independientes impulsores del Periodismo de datos en España*. En Periodismo de datos, Blanco Castilla & M. Quesada (Coords.), CAC112 (pp. 39–64). La Laguna (Tenerife). Latina
- Niño, M., & Illarramendi, A. (2015). Entendiendo Big Data: antecedentes, origen y desarrollo posterior. *Dyna: New technologies*, 2(1). <https://doi.org/10.6036/NT7835>
- Salaverría, R., (2004). Diseñando el lenguaje para el ciberperiodismo. *Chaqui*. 86. 38-45
- Salaverría, R., (2005). Hipertexto periodístico: mito y realidad. Universidad Ramon Llull, Barcelona. III Congrés internacional Comunicació I Realitat. *Trípodos*. 517-524
- Riquelme, J., Ruiz, R. Y Gilbert, K., (2006). Minería de datos. *Conceptos y tendencias Inteligencia Artificial*. 29. PP 11-18
- Valero Sancho, J. L. (1999-2000). La infografía de prensa. *Ámbitos*. 3. 123-131 <https://www.recimundo.com/index.php/es/article/view/182>
- Valero Sancho, J.L. (2003). El relato en la infografía digital, en *Manual de redacción ciberperiodística*, Díaz-Noci y Salaverría (Coords). Ariel Comunicación.
- Valero Sancho, J. L. (2014). La visualización de datos. *Revista Internacional de Comunicación*, 25, 1–14. <http://institucional.us.es/ambitos/?p=1136>

- Valero, J.M Y Carvajal, M. (2017). Organización de equipos multidisciplinares: el periodismo de datos en *El País Textual & Visual Media*. 10. 247-262
- Vallejo Ballesteros, H. F., Guevara Iñiguez, E., & Medina Velasco, S. R. (2018). Minería de Datos. *Revista Científica Mundo de la Investigación y el Conocimiento*, 2, 339–349.
- Vázquez, P. y Fernández, T. (2016). *Storytelling como recurso discursivo en el Periodismo de datos*. En *Periodismo de datos*, Blanco Castilla & M. Quesada (Coords.), CAC112 (pp. 137-158). La Laguna (Tenerife). Latina
- Videla-Rodríguez, J., (2018). *Big data y comunicación: una relación imprescindible*. https://www.researchgate.net/publication/328957109_Big_data_y_comunicacion_una_relacion_imprescindible

7. Anexos

Anexo 1: Hoja de cálculo del análisis del año 2016

	AMBITO	SECCION	HIPERTEXT	BASE DE DATOS	MULTIMEDIA	PRESENTACION	Exposicion	Contenido	INTERACTIVIDAD	
26-jul	1	1	1	4	5	4	4	6	2	1
29-jul	1	1	1	3	3	4	4	1	2	1
02-ago	1	2	2	4	5	4	1	5	2	1
05-ago	2	3	2	2	3	4	1	4	2	1
12-ago	2	3	1	1	5	4	1	5	2	1
12-ago	1	1	1	4	5	4	1	1	2	1
16-ago	1	1	1	1	5	4	1	1	2	1
22-sep	1	1	1	1	5	4	1	6	2	1
22-sep	1	1	1	4	5	4	1	1	2	1
31-dic	2	2	2	3	3	4	1	1,3	2	1
02-oct	1	1	1	1	5	4	1,4	1	2	1
09-oct	1	1	1	1	5	2,4	1	1,3	2	1
16-oct	1	1	1	3	3	4	1	1,3	2	1
16-oct	1	4	1	1	5	4	1	1	2	1
16-oct	1	1	1	1	5	6	8	6	2	1
16-oct	1	1	3	3	3	4	1	1,3	2	1
17-oct	1	6	4	4	3	4	1,3	1,2	7	1
06-nov	2	1	2	2	3	4	1	1	2	1
09-nov	2	1	2	2	3	4	1,3	3,5	2	1
25-nov	2	2	2	1	5	4	1	3	2	1
TOTALES										
	Nacional 14	Política 13	Micro 8	Excel	Imagen	Grafico 14	Barra 12	Ar.	Nada 20	
	Internacional 6	Economía 3	Macro 3	CSV	Video 1	Diagrama	Sectores	Informativo 19	Clickativo	
		Deportes 2	Ambas 4	Web 8	Sonido	Mapa 2	Línea 5	Didactico	Interactivo	
		Sociedad 1	Nada 5	PDF	Infografía 20	Tabla 3	Pirámede 1	Acont.		
		Sanidad 0		Nada 12		Línea del tiempo	Puntos 3	Persona		
		Otros 1				Fotoscroling	Otros 2	Geograf		
						Videoscroling		Híbrida 1		
						Otros 1				

Anexo 2: Hoja de cálculo del análisis del año 2019

	AMBITO	SECCION	HIPERTEXT	BASE DE DATOS	MULTIMEDIA	PRESENTACION	Exposicion	Contenido	INTERACTIVIDAD	
24-ene	1	1	4	3	5	4	1,3	1	6	1
28-mar	1	1	6	3	3	4,1	3	2	6	1
15-jun	1	1	1	3	3	4	3	5	6	1
28-mar	1	1	4	3	3	2,4,1	1,3	3,2	7	1
24-ene	1	1	4	3	3	1,4	2	2	2	1
07-feb	1	1	6	3	3	1	4,1	3,1	2	2
21-oct	1	1	4	3	3	4	3,4	5	6	2
16-feb	1	1	1	1	5	4	1	1	2	1
26-may	1	1	1	1	5	4	3	2	6	2
15-jul	1	1	1	1	5	4,1	3,1	2,1	6	1
28-mar	1	1	6	1	5	1,4	2,4	2	2	1
21-jun	1	1	3	3	3	4	3,1	2,3	2	2
02-may	1	1	1	1	5	4	3,1	2,5	6	2
21-feb	1	1	2	1	5	4	3	2	6	3
24-ene	1	1	2	1	5	1,4	1	1,3	2	2
23-abr	1	1	1	1	5	4	8,1	6,1	2	1
14-feb	1	1	4	3	5	4	3	2	6	2
13-feb	1	1	4	3	3	4	1,3	3,1	7	2
03-nov	1	1	1	1	5	4	6	6	2	1
02-oct	1	1	1	3	3	4	2	3,5,1	2	1
	Nacional 20	Política 8	Micro 9	Excel	Imagen 7	Grafico 10	Barra 8	Ar.	Nada 11	
	Internacional	Economía 2	Macro	CSV	Video 1	Diagrama 1	Sectores 7	Informativo 9	Clickativo 8	
		Deportes 1	Ambas 11	Web 9	Sonido	Mapa 12	Línea 6	Didactico	Interactivo 1	
		Sociedad 6	Nada	PDF	Infografía 20	Tabla 3	Pirámede 0	Acont.		
		Sanidad 0		Nada 11		Línea del tiempo	Puntos 4	Persona		
		Otros 3				Fotoscroling	Otros 2	Geograf	9	
						Videoscroling 1		Híbrida 2		

Anexo 3: Hoja de cálculos sobre los totales de la muestra analizada

	2016	2017	2018	2019	2020	2021	2022	TOTALES	PORCENTAJES	
ÁMBITO										ÁMBITO
Nacional	14	12	16	20	17	15	12	106	75%	Nacional
Internacional	6	8	4	0	3	5	8	34	25%	Internacional
	2016	2017	2018	2019	2020	2021	2022			
Política	13	3	2	8	4	4	4	38	27,14	Política
Economía	3	3	7	2	2	5	5	27	19,29	Economía
Deportes	2	6	1	1	0	2	3	15	10,71	Deportes
Sociedad	1	5	6	6	11	6	5	40	28,57	Sociedad
Sanidad	0	0	0	0	0	0	0	0	0,00	Sanidad
Otros	1	3	4	3	3	3	3	20	14,29	Otros
	2016	2017	2018	2019	2020	2021	2022	140		
Micro	8	4	5	9	10	7	3	46	32,86	Micro
Macro	3	1	1	0	3	0	1	9	6,43	Macro
Ambas	4	15	11	11	7	13	15	76	54,29	Ambas
Nada	5	0	3	0	0	0	1	9	6,43	Nada
BASE DE DATOS	2016	2017	2018	2019	2020	2021	2022			BASE DE DATOS
Excel	0	0	1	0	0	1	0	2	1,43	Excel
CSV	0	0	0	0	0	0	1	1	0,71	CSV
Web	8	14	9	9	18	16	18	92	65,71	Web
PDF	0	1	0	0	0	3	0	4	2,86	PDF
Nada	12	5	10	11	2	12	1	53	37,86	Nada
MULTIMEDIA	2016	2017	2018	2019	2020	2021	2022			MULTIMEDIA
Imagen	0	1	6	7	5	1	7	27	19,29	Imagen
Video	1	2	1	1	0	0	1	6	4,29	Video
Sonido	0	0	0	0	0	0	0	0	0,00	Sonido
Infografía	20	20	19	20	20	19	20	138	98,57	Infografía
MEZCLA	5%	15%	30%	30%	20%	5%	40%			MEZCLA
PRESENT	2016	2017	2018	2019	2020	2021	2022			PRESENT
Grafico	14	19	13	10	11	13	15	95	67,86	Grafico
Diagrama	0	1	1	1	0	0	0	3	2,14	Diagrama
Mapa	2	2	7	12	8	7	5	43	30,71	Mapa
Tabla	3	5	4	3	4	6	4	29	20,71	Tabla
Línea del tiempo	0	0	0	0	0	0	0	0	0,00	Línea del tiempo
Fotoscroling	0	0	1	0	5	2	2	10	7,14	Fotoscroling
Videoscroling	0	0	0	1	0	0	0	1	0,71	Videoscroling
Otros	1	0	1	1	0	2	1	6	4,29	Otros
Mezcla	15%	30%	25%	50%	35%	45%	40%			Mezcla
EXPO	2016	2017	2018	2019	2020	2021	2022			EXPO
Barra	12	11	12	8	4	3	10	60	42,86	Barra
Sectores	0	1	0	7	0	0	0	8	5,71	Sectores
Línea	5	9	5	6	6	5	13	49	35,00	Línea
Pirámede	1	0	1	0	1	0	0	3	2,14	Pirámede
Puntos	3	4	3	4	1	7	2	24	17,14	Puntos
Otros	2	3	2	2	12	11	4	36	25,71	Otros
Mezcla	30%	40%	35%	45%	20%	30%	40%			Mezcla
	2016	2017	2018	2019	2020	2021	2022			
Ar.	0	0	0	0	0	0	0	0	0,00	Ar.
Informativo	19	20	9	9	6	13	14	90	64,29	Informativo
Didactico	0	0	1	0	1	0	0	2	1,43	Didactico
Acont.	0	0	1	0	0	0	0	1	0,71	Acont.
Persona	0	0	0	0	0	0	0	0	0,00	Persona
Geograf	0	0	4	9	7	3	2	25	17,86	Geograf
Híbrida	1	0	5	2	6	4	4	22	15,71	Híbrida
Mezcla	2	1	5	3	4	3	3	21		
Porcentaje	28,57	14,29	71,43	42,86	57,14	42,86	42,86			
	2016	2017	2018	2019	2020	2021	2022			
Nada	20	20	7	11	6	5	6	75	53,57	Nada
Clickativo	0	0	12	8	7	8	10	45	32,14	Clickativo
Interactivo	0	0	1	1	7	7	4	20	14,29	Interactivo