

Estándares de calidad en el periodismo de datos: fuentes, narrativas y visualizaciones en los *Data Journalism Awards 2019*

Quality standards in data journalism: sources, narratives and visualizations in the *Data Journalism Awards 2019*

Alba Córdoba-Cabús

Cómo citar este artículo:

Córdoba-Cabús, Alba (2020). "Estándares de calidad en el periodismo de datos: fuentes, narrativas y visualizaciones en los *Data Journalism Awards 2019*". *Profesional de la información*, v. 29, n. 3, e290328.

<https://doi.org/10.3145/epi.2020.may.28>

Artículo recibido el 04-11-2019
Aceptación definitiva: 02-03-2020



Alba Córdoba-Cabús

<https://orcid.org/0000-0002-3519-0583>

Universidad de Málaga
Departamento de Periodismo
León Tolstoi, s/n. 29071 Málaga, España
albacordoba@uma.es

Resumen

Este estudio examina las características de los trabajos de periodismo de datos con mayor reconocimiento internacional, con la intención de ofrecer una radiografía actual y extraer elementos comunes en el ejercicio de esta práctica periodística. Mediante un análisis de contenido se diseccionan 42 proyectos nominados a los *Data Journalism Awards 2019*, seleccionados por la *Global Editors Network* por cumplir estándares de calidad en el marco de esta especialidad. Las piezas se examinan desde tres perspectivas: las propiedades de la historia, los datos y las visualizaciones empleadas. Entre otras constataciones, los resultados evidencian la transversalidad de esta técnica de tratamiento de la actualidad, moldeable para ser empleada en cualquier área temática, y el predominio de temas como la salud, la ciencia y el medioambiente, debido en parte a circunstancias relacionadas con la urgencia climática y los desafíos medioambientales establecidos. Existe variedad de fuentes pero se detecta una fuerte dependencia de las documentales procedentes de gobiernos u oficinas públicas (59,52%) y un uso mayoritario de las infografías como forma de visualización (45,24%), presentadas generalmente en forma de *scrollytelling*. Asimismo, se deduce que la calidad está condicionada por la naturaleza de las fuentes, el carácter innovador del análisis estadístico realizado, la forma de recopilar la información y la complejidad de las visualizaciones incorporadas. El periodismo de datos es una práctica periodística en constante cambio, por eso resulta necesario repensar los parámetros de calidad a medida que evoluciona como especialización para adaptarlos a las nuevas formas de materialización de la información.

Palabras clave

Periodismo de datos; Periodismo digital; Nuevas narrativas; Premios; *Data Journalism Awards*; Calidad; Visualización de información; Visualizaciones; Datos; Fuentes; Infografías; Análisis de contenido.

Abstract

The purpose of this study is to examine the characteristics of the most internationally recognized data journalism projects, with the aim of providing an in-depth snapshot of the current situation as well as identifying the common elements in the exercise of this journalistic practice. A content analysis is carried out to dissect 42 projects nominated for the *Data Journalism Awards 2019*, which were selected by the *Global Editors Network* for complying with the quality standards of this specialty. The pieces are examined from three different perspectives: the features of the story, the

data, and the visualizations used. Among other findings, the results show the transversality of this treatment technique, which can be shaped to be applied in any subject area, and the prevalence of topics such as health, science, and the environment, partly due to circumstances related to the climate emergency together with the current environmental challenges. There is a wide variety of sources, but a strong dependence on documentaries from governments or public offices (59.52%) can be observed as well as widespread use of infographics as a form of visualization (45.24%), generally presented as scrollytelling. It is also deduced that quality is conditioned by the nature of the sources, the innovative character of the analysis carried out, the way in which the information is compiled, and the complexity of the visualizations involved. Data journalism is an ever-changing journalistic practice, hence the necessity to rethink its quality parameters as it evolves as a specialization to adapt them to the new forms of materializing information.

Keywords

Data journalism; Digital journalism; New narratives; Awards; *Data Journalism Awards*; Quality; Information visualization; Visualizations; Data; Sources; Infographics; Content analysis.

1. Introducción

La datificación y la digitalización de la sociedad han transformado la manera en la que trabajamos y pensamos (Loosen, 2018; Coddington, 2015), propiciando la aparición de perfiles profesionales dedicados a extraer información y representar la realidad escondida tras grandes volúmenes de cifras, como es el caso del periodista de datos.

Los valores, rutinas y normas del periodismo tradicional se mantienen, pero la adaptación constante a tecnologías y programas con los que trabajar provoca que el periodismo de datos se encuentre en constante evolución (Tandoc; Soo-Kang, 2017). En este trabajo se entiende que el periodismo de datos es una instrumento periodístico –aplicable a cualquier área temática– que posibilita el manejo de grandes cantidades de datos, sirviéndose de la estadística y publicando resultados aprovechando las visualizaciones. A simple vista parece que no existen diferencias entre el periodismo de precisión o asistido por computadora y el actual periodismo de datos. Sin embargo, Philip Meyer (Gray; Bounegru; Chambers, 2012) atestigua que la principal distinción radica en la enorme cantidad de información numérica a la que se enfrenta el profesional, impulsado por el auge de la cultura de datos abiertos y el surgimiento de nuevas extensiones tecnológicas. Sandoval-Martín y La-Rosa (2018) especifican que la diferencia fundamental radica en el contexto, debido a la aparición de necesidades surgidas a raíz de los datos masivos como el procesamiento de cantidades ingentes de información, el cumplimiento de los ideales de transparencia y datos abiertos o la confección de visualizaciones para facilitar la interpretación de los datos.

Los números por sí solos no poseen valor, es el periodista el encargado de buscar historias tras ellos y mostrarlas a la audiencia de una manera atrayente (Sánchez-Bonvehí; Ribera, 2014; Hammond, 2015). Bradshaw (2011) planteó la “pirámide invertida del periodismo de datos”, estableciendo las fases a seguir para examinar grandes cantidades de información: compilar, cribar, contextualizar, combinar y comunicar. Estos cinco estadios implican la familiarización de los reporteros con aplicaciones analíticas y el aprendizaje constante de técnicas para acceder a las bases de datos en bruto y aplicar la lógica periodística para ayudar a la audiencia a comprender las historias (Tabary; Provost; Trottier, 2016; Paraise; Darigal, 2013).

El nivel de desarrollo del periodismo de datos depende del contexto en el que emerge, a lo que se une una serie de limitaciones para su evolución como práctica mediática (Uskali; Kuutti, 2015; Howard, 2014). De-Maeyer et al. (2015) aglutinan estos obstáculos en tres grandes grupos:

- dentro de las organizaciones: en términos de tiempo, recursos y flujos de trabajo;
- fuera de las organizaciones: dificultad para acceder a los datos y el formato de estos;
- individuales: formación.

Paraise y Darigal (2013) y Tabary, Provost y Trottier (2016) hacen hincapié en la dependencia gubernamental, mientras que otros como Fink y Anderson (2015) y Rogers, Schwabish y Bowers (2017) concretan y sugieren que no se trata tanto de falta de formación como de escasez de herramientas.

1.1. Estándares de oro del periodismo de datos

Al tratarse de una disciplina que evoluciona en función del contexto y requerir adaptación constante, se generan discrepancias conceptuales en torno a ella. A su vez, no existe un consenso claro sobre cuáles son los “estándares de oro” que dotan de calidad y éxito a las historias con datos, lo que dificulta su evolución como especialización. Alexander y Vetere (2011) plantearon en su momento un repertorio heterogéneo de parámetros:

- ofrecer confianza en los datos que se presentan;
- disponer de factor sorpresa;
- ser transparente en los procesos;
- contar con un líder;
- emplear un lenguaje accesible;
- permitir el acceso a las cifras.

Sin embargo, estos criterios han quedado obsoletos y se obvian elementos presentes en la materialización de este tipo de proyectos.

La llegada en 2012 de los *Data Journalism Awards* —organizados por la *Global Editors Network* y dirigidos por Simon Rogers— supuso la creación de los primeros galardones internacionales que reconocen los trabajos sobresalientes de periodismo de datos a nivel mundial. Estos premios homenajean cada año a las mejores investigaciones sustentadas en cifras, valorando la calidad e innovación en el uso y el análisis de datos, el contenido y la presentación de la información (*Global Editors Network*, 2019). El jurado está conformado por profesionales reconocidos del periodismo de datos, entre los que se encuentran Alberto Cairo, de la *University of Miami School of Communication*; Aron Pilhofer, de la *Temple University*; o Liliana Bounegru, del *Oxford Internet Institute*.

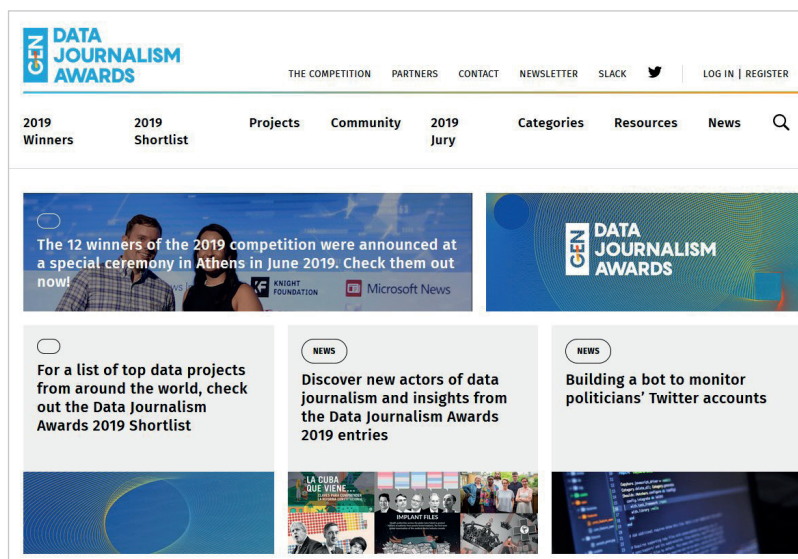


Figura 1. *Data Journalism Awards*
<https://www.datajournalismawards.org>

El periodismo de datos está presente en toda la bibliografía académica de producción digital de información y los perfiles emergentes de digitalización de la sociedad. Los medios lo emplean como un arma ante el periodismo de declaraciones y se presenta como clave para la recuperación de la credibilidad y la democratización, ofreciendo objetividad, técnicas para controlar la actividad de los gobiernos e incrementar la implicación de la ciudadanía en los asuntos políticos (**Cushion**; Lewis; Callaghan, 2016; Cortés-del-Álamo; Luengo-Cruz; Elías, 2018; Paraise; Darigal, 2013). De ahí que sea precisamente la política una de las áreas en las que esta disciplina tiene mayor presencia (**Wright**; Doyle, 2019).

Durante el último lustro diversos autores han abordado los *Data Journalism Awards*. Loosen, Reimer y De-Silva-Schmidt (2017) examinaron los nominados de 2013 a 2016 (n=225) y a través de un análisis de contenido detectaron la frecuente aparición de la política como tema principal, dependencia de las fuentes públicas, empleo de visualizaciones simples como gráficos de barras o mapas y uso de la narración y la visualización como forma de comunicación más común. Ojo y Heravi (2018) plantearon un análisis de los ganadores de 2013 a 2016 (n=44) a partir del cual concluyeron que la mayoría de las historias tenían como propósito informar, se basaban en bases de datos públicas y la forma más frecuente de comunicación era la narración y la visualización, con mayor presencia de gráficos y mapas con anotaciones. Young, Hermida y Fulda (2018) estudiaron los proyectos canadienses nominados y premiados en los *Data Journalism Awards* de 2012 a 2015 (n=26), coincidiendo en que la narración y la visualización eran las formas de publicación más utilizadas, destacando el estilo diegético conducido por el autor y los mapas dinámicos como elemento visual más común. Otros trabajos e investigaciones de la última década, aunque emplean muestras distintas, alcanzan conclusiones similares (Paraise; Darigal, 2013; Knight, 2015; De-Maeyer et al., 2015; Tabary; Provost; Trottier, 2016; Stalph, 2018; Stalph; Borges-Rey, 2018; Appelgren, 2018; Córdoba-Cabús, 2018; Lowrey; Hou, 2018; Váñez; Codina, 2018; Usher, 2019). Diversos expertos señalan que el principal punto de mejora de las piezas de periodismo de datos debe ser la incorporación de testimonios a las historias (Paraise; Darigal, 2013; Gray; Bounegru; Chambers, 2012).

“ Para la aplicación práctica del periodismo de datos se precisa, principalmente, apoyo institucional ”

Tomando como punto de partida los *Data Journalism Awards 2019*, esta investigación tiene como objetivo examinar las características de los trabajos de periodismo de datos con mayor reconocimiento internacional, con la intención de ofrecer una radiografía actual y extraer elementos comunes en el ejercicio de esta práctica periodística entre los nominados a este galardón. De esta forma, y teniendo en cuenta para ello las propiedades de las historias, las cifras y las visualizaciones, se pretenden fijar los estándares de calidad de la disciplina presentes en estos trabajos.

2. Método

Con la intención de examinar la materialización del periodismo de datos de calidad, se optó por aplicar el análisis de contenido (Krippendorff, 2002; Igartua, 2006) a los trabajos nominados a los *Data Journalism Awards* en su edición de 2019 (tabla 1), con propósito descriptivo univariante y bidimensional. La elección del universo vino determinada por el hecho de que estos premios están considerados el mayor reconocimiento internacional de la especialidad.

La investigación se centró en analizar piezas únicas basadas en datos; por ello, si la nominación no hacía referencia a un proyecto concreto, sino que recaía sobre el conjunto de la redacción, un portfolio individual o aplicaciones y webs

completas, el trabajo se excluía del análisis. La muestra final quedó conformada por 42 piezas, procedentes en su mayoría de Estados Unidos (50%) –el siguiente país en cuanto a número de proyectos es Reino Unido (7,14%)–. Los trabajos fueron confeccionados, sobre todo, por las versiones web de periódicos tradicionales (48,84%), a gran distancia de las organizaciones periodísticas como *ProPublica* (23,26%) o las agencias de noticias (18,60%).

Tabla 1. Resumen de los proyectos de los *Data Journalism Awards 2019*

		2019
Proyectos presentados	Frecuencia	607
Proyectos nominados	Frecuencia	103
	% de los presentados	16,9
Proyectos adecuados para el análisis	Frecuencia	42
	% de los nominados	40,77
Proyectos ganadores	Frecuencia	13
	% de proyectos ganadores respecto a los nominados	12,62
	Proyectos ganadores analizados	6

Con este corpus de análisis podríamos examinar la disciplina sin restricciones temáticas ni geográficas y alcanzar resultados significativos. No en vano, **Appelgren** (2018) asegura que el estudio de los nominados a estos premios influye en el desarrollo del periodismo de datos como práctica mediática. Conocer a fondo las propiedades de las historias aportaría una idea certera sobre cómo plasman la información los grandes trabajos de periodismo de datos.

Para la operacionalización de las variables se utilizaron elementos empleados frecuentemente en la bibliografía científica, incorporando consideraciones propias. **Ojo** y **Heravi** (2018) plantean que lo adecuado es combinar distintos análisis para conformar un constructo robusto que permita analizar esta práctica periodística en su totalidad. El libro de códigos utilizado (tabla 2) agrupaba las variables en tres conjuntos: propiedades de la historia –relacionado con la narrativa–, datos y visualizaciones.

Tabla 2. Variables para el análisis

Variables	Opciones	Referencias
Temática	Política, Sociedad, Economía, Salud/Ciencia/Medioambiente, Educación, Deportes, Cultura/Arte, Otro	Loosen; Reimer; De-Silva-Schmidt (2017), Stalph (2018), Tandoc; Soo-Kang (2017)
Estilo narrativo	Explicativo, Exploratorio, Híbrido	Segel; Heer (2010), Ojo; Heravi (2018), Young; Hermida; Fulda (2018)
Tipo de historias	Solo por los hechos, Noticias basadas en datos, Datos locales contando historias, Análisis y antecedentes, Investigaciones de buceo profundo	Rogers (2014)
Dimensión	Pasado, Presente, Futuro	Davenport (2014)
Elementos de foco	Qué ocurre, Por qué, Cómo abordar el problema	Davenport (2014)
Propósito	Informar, Persuadir, Entretener, Explicar, Desconocido	Slaney; Kelliher (2012), Lee; Herny-Riche; Isenberg; Carpendale (2015), Gray; Chambers; Bounegru (2012)
Manera de comunicar	Visualización, Narración, Comunicación social, Humanización, Personalización, Utilización	Bradshaw (2011)
Fuente	Gobierno u oficina pública, Corporaciones privadas, Otras organizaciones, Fuentes propias, otras	Loosen; Reimer; De-Silva-Schmidt (2017), Knight (2015)
Qué muestran los datos	Comparación, Conexiones y flujos, Cambios en el tiempo, Jerarquía, Otros	Kang (2015)
Visualización	Tabla, Gráfico estático, Mapa, Gráfico interactivo, Elementos gráficos, Visualización animada, Infografía, Otros	Appelgren (2018), Young; Hermida; Fulda (2018), Knight (2015), Coddington (2015), Segel; Heer (2010)
Ratio de visualizaciones	Predomina el texto, Predominan las visualizaciones, Equilibrado, Solo texto, Solo visualizaciones	Stalph (2018)

La incorporación en el análisis de la parte narrativa y de la visual resultaba especialmente importante por ser las dos formas de comunicación más frecuentes en el periodismo de datos (**Loosen; Reimer; De-Silva-Schmidt**, 2017; **Córdoba-Cabús**, 2018; **Young; Hermida; Fulda**, 2018; **Ojo; Heravi**, 2018).

Las propiedades de las historias se estudiaron a través de conjuntos de variables categóricas nominales y de escala centradas en la parte narrativa de los trabajos. La temática se estableció en función de las categorizaciones propuestas en

las investigaciones de **Loosen, Reimer y De-Silva-Schmidt (2017)**, **Stalph (2018)** y **Tandoc y Soo-Kang (2017)**, quedando estructurada así:

- política
- sociedad
- economía y negocios
- salud, ciencia y medioambiente
- educación
- deportes
- cultura y arte
- otros.

El estilo narrativo se evaluó en base a la propuesta de **Segel y Heer (2010)**, reproducida más tarde por otros autores como **Ojo y Heravi (2018)** y **Young, Hermida y Fulda (2018)**, distinguiendo entre el estilo:

- explicativo o conducido por el autor: caracterizado por poseer un orden lineal, fuerte carga de mensajes y reducida interactividad;
- exploratorio o impulsado por el lector: no presenta un orden concreto, los mensajes son escasos y la interacción del usuario es libre;
- híbrido: en el que se combinan ambos.

El tipo de historia se clasificó teniendo en cuenta el tipo de dato, la manera de recopilar la información y la audiencia a la que se dirigía. En este apartado se podía seleccionar entre: piezas con datos públicos con los que se confecciona una visualización, noticias basadas en datos, datos locales contando historias, análisis y antecedentes o investigaciones de buceo profundo (**Rogers, 2014**). Tanto la dimensión temporal –si aborda el pasado, el presente o el futuro– como los elementos de la historia –si responde a qué ocurre, por qué y cómo– fueron delimitados según el criterio de **Davenport (2014)**. Por su parte, cuando se menciona el propósito del trabajo publicado se habla de diferenciar entre una finalidad desconocida, informar, persuadir, entretener y explicar (**Kelliher; Slaney, 2012; Lee et al., 2015; Gray; Bounegru; Chambers, 2012**), mientras que la manera de comunicar vino determinada por la propuesta de **Bradshaw (2011)**: visualización, narración, comunicación social, humanización, personalización y utilización.

“ El periodismo de datos es una herramienta transversal que puede ser incorporada en cualquier área temática ”

Al tratarse de informaciones sustentadas en cifras, se estimó adecuado dedicarles un apartado específico, ofreciendo así la posibilidad de conocer su procedencia, naturaleza, forma en la que se presentan e información que aportan. En función de las investigaciones de **Knight (2015)** y **Loosen, Reimer y De-Silva-Schmidt (2017)** se confeccionó la siguiente clasificación de fuentes:

- gobiernos u oficinas públicas;
- corporaciones privadas;
- otras organizaciones como ONGs, institutos de investigación o universidades;
- fuentes propias;
- otras.

El tipo de información que proporcionan los datos se determinó en base a la distribución establecida por **Kang (2014)**, discerniendo entre comparar, plasmar cambios en el tiempo, mostrar relaciones entre variables, evidenciar jerarquía y otros.

Las visualizaciones se entienden como un elemento clave para el ejercicio del periodismo (**Engebretsen; Kennedy; Weber, 2018**) y más concretamente del periodismo de datos, constituyendo una parte fundamental de la última fase de la pirámide invertida de la especialidad: la comunicación. Por sí solas son capaces de contar una historia y se pueden adaptar a la audiencia específica a la que se dirigen (**Barlow, 2014**), convirtiéndose en un elemento para ser interpretado (**Cairo, 2017**). Por eso se dedica un espacio exclusivo para su examen, contabilizando el número y especificando sus tipos. Existen multitud de clasificaciones según los tipos de visualizaciones, por lo que en este trabajo se planteó una combinación tratando de abarcar todas las posibilidades (**Segel; Heer, 2010; Coddington, 2015; Knight, 2015; Appलगren, 2018; Young; Hermida; Fulda, 2018**):

- tabla o lista: representación de datos o información en formato cuadrícula o lista;
- línea de tiempo: plasmar en una visualización acontecimientos o hechos ordenados cronológicamente;
- gráfico estático: representación visual de información bidimensional. No se identificará el tipo de gráfico que incorpora (de barras, sectores, dispersión, etc.), sino la incorporación o no de interactividad;
- gráfico interactivo: representación visual de información bidimensional. Se entiende por interactividad aquellas acciones que puede realizar el usuario con el gráfico (buscar, filtrar, seleccionar...);
- mapa con o sin interactividad: un identificador de ubicación, una fecha gráfica con una o más ubicaciones identificadas;
- elementos gráficos: vídeos, fotos o ilustraciones;
- visualización animada: la que plasma información representada de manera animada. El movimiento es de forma automática, el usuario no ejecuta ninguna acción. Ejemplos: gif, iconos, secuencias de gráficos en movimiento, etc.;

- infografía: combinación de imágenes o ilustraciones, texto y gráficos en una misma visualización. Aunque la infografía incorpore elementos en movimiento e interactivos no se categorizará como visualización animada;
- otros: aquí se incluirá cualquier otra forma de visualización que no pueda contabilizarse en las opciones anteriores.

Siguiendo las indicaciones de **Stalph** (2018) se valoró qué predominaba más en los artículos: si el texto o las visualizaciones. Al igual que en su investigación, en este trabajo se limitó el análisis de visualizaciones por pieza a un máximo de ocho. En los trabajos con varias pantallas se analizó exclusivamente la principal –contabilizándose como una única visualización, teniendo en cuenta los datos y las representaciones empleadas–. Cuando la historia se contaba a través de una infografía, se examinó y categorizó como una única visualización.

3. Resultados

A continuación se describen de manera pormenorizada las características de los trabajos nominados a los *Data Journalism Awards 2019*, de acuerdo con las variables del análisis.

3.1. Propiedades de las historias

El tema más frecuente en los artículos examinados es el relacionado con la salud, la ciencia y el medioambiente (gráfico 1), centrándose, especialmente, en este último. Los trabajos localizados se esfuerzan por otorgar visibilidad al cambio climático, los desastres naturales y las energías renovables, como por ejemplo los proyectos *To see how levees increase flooding, we built our own* –ProPublica– o *Concrete and coral* –Reuters–. La investigación de ProPublica refleja cómo la construcción de diques para proteger zonas concretas de inundaciones perjudica a otras áreas, mientras que la de Reuters se centra en mostrar el ritmo y escala de construcción de los arrecifes y terrenos recuperados en diferentes islas del Mar de la China Meridional. Los artículos que tratan problemas o asuntos sociales –concernientes a colectivos sociales, sus relaciones e instituciones, así como informes de delitos, medios de comunicación, religión e historia– ocupan el segundo lugar. Cerca del 17% de las piezas analizadas abordan asuntos políticos, en proporción similar a las recogidas bajo la categoría “Otra”, en la que se agrupan las coberturas de accidentes aéreos, temas candentes en redes sociales, así como cuestiones sin cabida en el resto de los items. Por su parte, “Economía” y “Educación” son los temas que menor representación obtienen.

El estilo narrativo predominante en los nominados a los *Data Journalism Awards 2019* es el explicativo (76,19%). En un segundo plano quedan las piezas híbridas (14,29%), que mezclan el estilo conducido por el autor y el impulsado por el lector, y las exploratorias (9,52%).

Si se clasifica la historia en función del tipo de dato, la manera de recabarlos y la audiencia a la que se dirige, más de la mitad de los trabajos analizados son noticias basadas en cifras (40,48%) y piezas en las que la información es pública y se respalda con análisis y visualizaciones (21,43%). Menor representación alcanzan las investigaciones exhaustivas de grandes conjuntos de datos (14,29%), los trabajos en los que se analizan y explican los antecedentes (14,29%), así como los proyectos que emplean datos locales (9,52%). Pese a ser “Salud, ciencia y medioambiente” el tema con mayor peso en la muestra, las investigaciones de buceo profundo están principalmente relacionadas con cuestiones sociales. Claro ejemplo de ello es el proyecto *The force report* de *NJ Advance Media*, una investigación de 16 meses durante los cuales el medio elaboró una base de datos sobre el uso de la fuerza por parte de la policía en Estados Unidos.

En general, las noticias no cubren un único período de tiempo. Del análisis de la muestra se desprende que más del 90% de los casos dedican espacio a tratar asuntos del presente, mientras que casi el 60% incorpora elementos diacrónicos. Las referencias al futuro suelen ser menores, identificándose sólo un 10% de piezas que las incorporan. Las predicciones figuran, principalmente, en artículos medioambientales en los que se vaticina cuándo volverá a ocurrir otro desastre o qué habría pasado si se hubiesen tomado medidas, como ocurre en *The planet’s hidden climate change* de Reuters.

El principal propósito de las piezas de periodismo de datos es informar a la audiencia (n=42). Sin embargo, no todas las historias poseen un único objetivo. Cerca de la mitad de los trabajos están interesados también en explicar (n=20), localizando solo un caso en el que la finalidad es diferente: persuadir. El foco central de las historias es contar qué ocurre u ocurrió (97,62%), incorporando en menor medida los factores que causan el hecho (42,86%) y más raramente las posibles soluciones para mejorar la situación (7,14%). Los proyectos que incluyen los tres elementos son los que cubren temas asociados a la salud, la ciencia y el medioambiente.

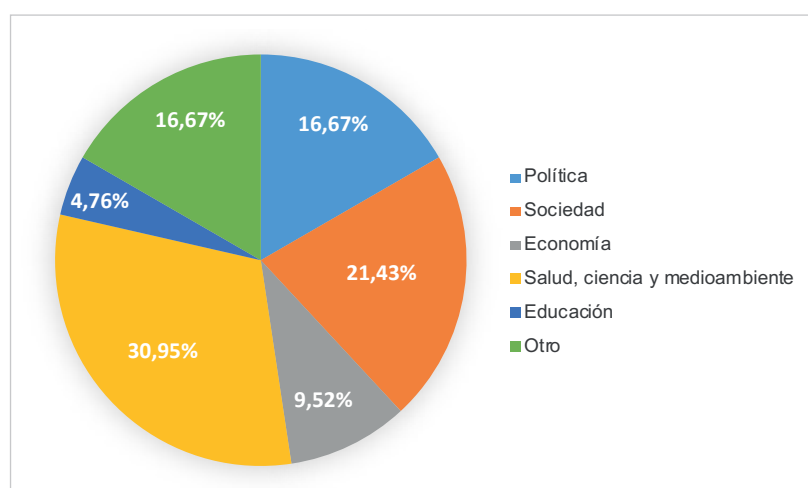


Gráfico 1. Temática de las historias nominadas

Como se observa en la tabla 3, todas las historias analizadas se componen de elementos narrativos y visualización, combinándolos en ocasiones con humanización –que consiste en transformar los datos en testimonios–, personalización –explicar cómo afectan los datos a la ciudadanía– y, en ocasiones, con utilización –crear una aplicación complementaria de las cifras empleadas–. Resulta especialmente llamativa la inexistencia de piezas que incorporen elementos de comunicación social –*crowdsourcing*–.

3.2. Datos y fuentes

En la contabilización de las fuentes de datos empleadas por cada artículo se encontraron valores extremos –cinco piezas contenían más de 20 fuentes– que fueron eliminados para no desvirtuar los estadísticos. Las cifras denotan la inclusión de una media cercana a las tres fuentes por trabajo, siendo lo más frecuente el uso de una sola (42,11%, n=16).

La mayor parte de las piezas de la muestra confía en fuentes públicas procedentes del Gobierno u organismos oficiales del tipo ministerios, el *National Center for Education Statistics*, el *National Institute of Securities Markets* o la *Central Pollution Control Board* (tabla 4). El segundo grupo (35,71%) consiste en piezas que utilizan datos de instituciones como ONGs, universidades o institutos de investigación. Algunos ejemplos de los aquí incluidos son *Save the Children*, la *University of Southampton* o la *Boston University's School of Public Health*. Aproximadamente el 30% de los trabajos analiza cifras que la organización correspondiente recopiló (mediante encuestas, *scraping* o buscando en sus propios archivos). Esta cuota es relativamente superior a las procedentes de filtraciones, análisis de emisiones televisivas –incorporadas ambas en la categoría “Otras”– y datos obtenidos de corporaciones privadas, como pueden ser *IBM* o *S&P Global Market Intelligence*.

Si se analizan en función de la temática, se observa que los asuntos políticos y económicos son los únicos en los que las fuentes principales no son las gubernamentales. Los datos proceden principalmente de filtraciones o de programas de televisión (42,83% en el caso político), así como de otras organizaciones (75% en economía). Las cifras extraídas por el propio equipo se dan con mayor frecuencia en temas sociales y en los relacionados con ciencia, salud y medioambiente. Los trabajos en los que la organización recaba los datos suelen obtener mayor reconocimiento, como es el caso del proyecto galardonado *Radmesser –Der Tagesspiegel–*. El medio recopiló, mediante sensores construidos por ellos mismos y con el apoyo de 100 voluntarios, la distancia entre coche y bicicleta en los adelantamientos para comprobar si se cumplía la legislación en materia de seguridad vial.

Estos valores numéricos se incorporan para proporcionar información y mostrar asociaciones y conexiones (57,14%), así como cambios temporales (52,38%). En menor medida, los datos se centran en comparar valores (28,57%).

3.3. Visualizaciones

Uno de los elementos distintivos del periodismo de datos son las visualizaciones. Esto incluye desde tablas y diagramas, que sirven para representar la información en que se basa la historia, hasta fotografías o vídeos que no están relacionados con las cifras, sino que tienen una función meramente ilustrativa.

De acuerdo con el análisis, las piezas de la muestra contienen de media más de dos visualizaciones ($\bar{x} = 2,64$). Sin embargo, lo más frecuente es el uso de un único elemento visual (50%).

La tabla 5 refleja el predominio de las infografías (45,24%). Estas se presentan estructuradas como una *scrollytelling* –manera de contar una historia en la que el lector usa la tecla de desplazamiento para descubrir el contenido– y su uso es más frecuente en los temas de salud, ciencia y medioambiente, presentando menor incidencia en economía y educación. Además, las infografías son las únicas visualizaciones presentes en aquellas piezas consideradas completas por atender a los tres focos centrales de la historia (qué, por qué y cómo abordar el problema). Los elementos gráficos (ilustración, foto y vídeo) son el segundo tipo más empleado, seguido de los gráficos estáticos, los gráficos interactivos, los mapas, las tablas y las visualizaciones animadas.

Tabla 3. Manera de comunicar las historias con datos

Manera de comunicar	Frecuencia	Porcentaje
Visualización	42	100
Narración	42	100
Comunicación social	0	0
Humanización	11	26,11
Personalización	4	9,52
Utilización	3	7,14

Tabla 4. Clasificación de fuentes según el tipo

Tipo de fuentes	Frecuencia	Porcentaje
Gobierno u oficina pública	25	59,52
Corporaciones privadas	4	9,52
Otras organizaciones	15	35,71
Fuente propia	13	30,95
Otras	7	16,67

Tabla 5. Tipo de visualización empleada en los nominados

Tipo de visualización	Frecuencia	Porcentaje
No hay	1	2,38
Tabla	3	7,14
Gráfico estático	10	23,81
Mapa	6	14,29
Gráfico interactivo	7	16,67
Elementos gráficos	13	30,95
Visualización animada	3	7,14
Infografía	19	45,24
Otros	5	11,90

Si se analiza la interactividad de los gráficos, llama la atención su presencia sobre todo en asuntos políticos, sociales y económicos. Dentro de los elementos interactivos localizados, destacan aquellos con funciones autorrellenables por el usuario, enmarcándose estas en el apartado de “Otros”. Esta característica se puede apreciar en el proyecto *How to forecast an American’s vote*, de *The Economist*, en el apartado “Build an American voter”.

4. Conclusiones y discusión

Este trabajo se proponía como objetivo principal identificar las propiedades y elementos comunes de los proyectos nominados a los *Data Journalism Awards 2019*, ofreciendo así una radiografía actual de los trabajos que cuentan con los “estándares de oro” del periodismo de datos.

Pese a la consideración tan extendida de que la política es una de las áreas en las que esta disciplina tiene mayor presencia, del análisis de la muestra se desprende la preponderancia de proyectos relacionados con salud, ciencia y medioambiente (30,95%), debido, en parte, a circunstancias relacionadas con la emergencia climática—evidenciada en varios informes recientes como el del *Intergovernmental Panel on Climate Change* o el *Informe anual de 2018 de ONU Medio Ambiente*— y con los desafíos medioambientales. Esto confirmaría la transversalidad de la herramienta, que puede ser incorporada a cualquier área temática. Los asuntos políticos suelen estar más presentes en el periodismo de datos diario, cometiendo en algunas ocasiones el error de ajustar esta herramienta a la simple visualización. En estos casos, la necesidad de inmediatez eclipsa al periodismo de datos más analítico.

Queda patente, con el análisis del estilo narrativo, cierta escasez de interactividad. **Stabe** (2016) ya aventuró que trasladar la responsabilidad del diseñador al lector era un error y así lo han entendido los creadores de los proyectos nominados (el 76,19% opta por un estilo explicativo, conducido por el autor). Por eso se prevé que el uso de la interactividad siga en esta dirección, incorporando secuencias de gráficos estáticos con anotaciones simples y mostrando variaciones de un mismo conjunto de datos, limitando así las funciones de los usuarios.

Por lo general, los proyectos nominados en 2019 emplean cifras actuales y se centran en explicar qué ocurre u ocurrió, dejando a un lado las explicaciones sobre por qué suceden o cómo abordar las situaciones, lo cual implicaría mayor interpretación y análisis por parte del periodista. Más de la mitad de los artículos son noticias basadas en datos y piezas en las que estos se complementan con análisis y visualizaciones. La escasa presencia de investigaciones de buceo profundo está directamente relacionada con las limitaciones de tiempo y recursos a las que se enfrenta el periodismo de datos. Es por eso por lo que, excepto en la categoría “Mejor uso de datos en noticias de última hora”, la mayoría de los proyectos presentan temas atemporales. Lo mismo ocurre con las historias que incorporan datos locales: al ser pequeñas organizaciones, sus recursos son escasos y les resulta mucho más complicado alcanzar visibilidad y reconocimiento.

La manera más habitual de comunicar este tipo de trabajos es mediante narración y visualización, combinándolo en ocasiones con humanización—testimonios—. Es este uno de los puntos a explotar, ya que los propios periodistas consideran imprescindible transformar los datos en historias o encontrar a personas afectadas por las cifras para atraer a la audiencia (**Paraise; Darigal, 2013; Gray; Bounegru; Chambers, 2012**).

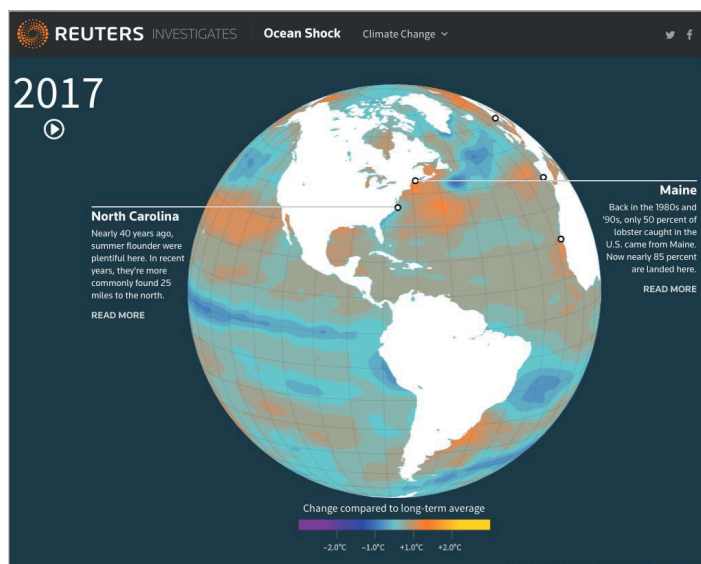


Figura 2. Proyecto *The planet’s hidden climate change* de Reuters. <https://www.reuters.com/investigates/special-report/ocean-shock-warming>



Figura 3. Proyecto *Every time Ford and Kavanaugh dodged a question, in one chart* de Vox. <https://www.vox.com/policy-and-politics/2018/9/28/17914308/kavanaugh-ford-question-dodge-hearing-chart>

Existe variedad de fuentes, pero se imponen las documentales procedentes del Gobierno u oficinas públicas. Esto explicaría el tipo de historias, la dimensión de tiempo y los elementos de foco más comunes. Son los asuntos políticos y económicos los únicos en los que las fuentes principales no son las gubernamentales. Esto puede deberse a la búsqueda de elementos diferenciadores que permitan un tratamiento de datos innovador, para conseguir diferenciarse en temáticas ya muy explotadas. El rasgo de calidad radicaría tanto en la combinación de las fuentes como en los análisis aplicados. La fuerte dependencia de este tipo de fuentes puede ser una limitación, ya que los periodistas confían en el procesamiento que las instituciones hacen de los datos. Aquellos proyectos en los que las cifras son recopiladas por la organización (mediante encuestas, *scraping* o indagando en sus propios archivos), siendo fuentes únicas o exclusivas con las que sólo cuenta el equipo, suelen alcanzar mayor reconocimiento.

Los asuntos políticos suelen estar más presentes en el periodismo de datos diario, cometiendo en algunas ocasiones el error de ajustar esta herramienta a la simple visualización

Las visualizaciones más frecuentes son las infografías, presentadas en su mayoría como una *scrollytelling*. El uso de los distintos tipos de visualizaciones va en consonancia con la adaptación constante del periodismo de datos a las nuevas herramientas y tecnologías, lo que implica una actualización continua del periodista. La infografía aporta distinción y calidad, ya que se considera uno de los elementos con datos más complejos por el nivel de interpretación y análisis requerido (Knight, 2015).

De lo anterior se deduce que la calidad en el periodismo de datos practicado por los nominados a los *Data Journalism Awards 2019* viene marcada por cuatro factores:

- la naturaleza de las fuentes, obteniendo mayor reconocimiento las piezas en las que se combinan varias y en las que se cuenta con información exclusiva;
- el carácter innovador del análisis realizado, destacando aquellas que presentan una perspectiva poco frecuente en temas más explotados como la política;
- la forma de recopilar la información, valorando positivamente los proyectos en los que la propia organización recaba los datos;
- las visualizaciones incorporadas, siendo las infografías presentadas como una *scrollytelling* las más complejas en cuanto a diseño e interpretación previa de la información.

El periodismo de datos es una práctica en constante cambio, por eso resulta necesario repensar los parámetros de calidad a medida que evoluciona como especialización para adaptarlos a las nuevas formas de materialización de la información.

Las limitaciones de esta investigación se encuentran principalmente en la muestra seleccionada. Los 42 trabajos examinados no fueron elegidos al azar y son proyectos confeccionados en un período de tiempo concreto, el bienio 2018-2019, por lo que no hay posibilidad de establecer tendencias o comparativas. La variabilidad de las publicaciones complicó el análisis de contenido y eso implica un sesgo en las variables estudiadas. Además, no se ha podido explicar cuántas personas intervinieron en las publicaciones, durante cuántos meses o de cuánto presupuesto disponían, para lo que hubiese sido necesario realizar entrevistas con los periodistas de datos. Aun así, se considera que las pautas identificadas en el análisis pueden servir para delimitar de forma más precisa los estándares de calidad del periodismo de datos tal y como se practica en la actualidad.

El periodismo de datos es una práctica periodística en constante cambio, por eso resulta necesario repensar los parámetros de calidad a medida que evoluciona como especialización para adaptarlos a las nuevas formas de materialización de la información

El periodismo de datos es una práctica en constante cambio, por eso resulta necesario repensar los parámetros de calidad a medida que evoluciona como especialización para adaptarlos a las nuevas formas de materialización de la información.

5. Referencias

Alexander, Stephanie; Vetere, Colleen (2011). "Telling the data story the right way". *Healthcare financial management*, v. 65, n. 10, pp. 104-110.

Appelgren, Ester (2018). "An illusion of interactivity. The paternalistic side of data journalism". *Journalism practice*, v. 12, n. 3, pp. 308-325.
<https://doi.org/10.1080/17512786.2017.1299032>

Barlow, Mike (2014). *Data visualization: A new language for storytelling*. Sebastopol, CA: O'Reilly Media. ISBN: 978 1 491 94503 2

Bradshaw, Paul (2011). "6 ways of communicating data journalism". *Online journalism blog*, 15 July.
<https://bit.ly/3dbcjDI>

Cairo, Alberto (2017). "Visualización de datos: una imagen puede valer más que mil números, pero no siempre más que mil palabras". *El profesional de la información*, v. 26, n. 6, pp. 1025-1028.
<https://doi.org/10.3145/epi.2017.nov.02>

- Coddington, Mark** (2015). "Clarifying journalism's quantitative turn. A typology for evaluating data journalism, computational journalism, and computer-assisted reporting". *Digital journalism*, v. 3, n. 3, pp. 331-348.
<https://doi.org/10.1080/21670811.2014.976400>
- Córdoba-Cabús, Alba** (2018). "Análisis del periodismo de datos en la campaña electoral del 20D a través de las ediciones digitales de diarios generalistas". *Estudios sobre el mensaje periodístico*, v. 24, n. 1, pp. 137-154.
<https://doi.org/10.5209/ESMP.59942>
- Cortés-Del-Álamo, Helena-María; Luengo-Cruz, María; Elías, Carlos** (2018). "Periodismo de datos y transparencia al margen de los grandes medios, un estudio comparativo de Civio y Propública". *Icono 14*, v. 16, n. 2, pp. 66-87.
<https://doi.org/10.7195/ri14.v16i2.1177>
- Cushion, Stephen; Lewis, Justin; Callaghan, Robert** (2016). "Data journalism, impartiality and statistical claims. Towards more independent scrutiny in news reporting". *Journalism practice*, v. 11, n. 10, pp. 1751-2794.
<https://doi.org/10.1080/17512786.2016.1256789>
- Davenport, Thomas H.** (2014). "10 kinds of stories to tell with data". *Harvard business review*, 5 May.
<https://hbr.org/2014/05/10-kinds-of-stories-to-tell-with-data>
- De-Maeyer, Juliette; Libert, Manon; Domingo, David; Heinderyckx, François; Le-Cam, Florence** (2015). "Waiting for data journalism. A qualitative assessment of the anecdotal take-up of data journalism in French-speaking Belgium". *Digital journalism*, v. 3, n. 3, pp. 432-446.
<https://doi.org/10.1080/21670811.2014.976415>
- Engebretsen, Martin; Kennedy, Helen; Weber, Wibke** (2018). "Data visualization in Scandinavian newsrooms. Emerging trends in journalistic visualization practices". *Nordicom review*, v. 39, n. 2, pp. 3-18.
<https://doi.org/10.2478/nor-2018-0007>
- Fink, Katherine; Anderson, Christopher W.** (2015). "Data journalism in the United States. Beyond the 'usual suspects'". *Journalism studies*, v. 16, n. 4, pp. 467-481.
<https://doi.org/10.1080/1461670X.2014.939852>
- Global Editors Network* (2019). *Eligibility & rules*.
<https://datajournalismawards.org/eligibility-rules>
- Gray, Jonathan; Bounegru, Liliana; Chambers, Lucy** (2012). *The data journalism handbook*. Sebastopol: O'Reilly Media. ISBN: 978 1 449330064
- Hammond, Phil** (2015). "From computer-assisted to data-driven: Journalism and big data". *Journalism*, v. 18, n. 4, pp. 408-424.
<https://doi.org/10.1177/1464884915620205>
- Howard, Alexander-Benjamin** (2014). *The art and science of data-driven journalism*. New York: Tow Center for Digital Journalism.
<https://doi.org/10.7916/D8Q531V1>
- Igartua, Juan-José** (2006). *Métodos cuantitativos de investigación en comunicación*. Barcelona: Bosch. ISBN: 978 84 97902717
- Kang, Martha** (2015). "Exploring the 7 different types of data stories". *Mediashift*, 15 June.
<https://mediashift.org/2015/06/exploring-the-7-different-types-of-data-stories>
- Kelliher, Aisling; Slaney, Malcolm** (2012). "Tell me a story". *IEEE multimedia*, v. 19, n. 1, pp. 1139-1148.
<https://doi.org/10.1109/MMUL.2012.13>
- Knight, Megan** (2015). "Data journalism in the UK: a preliminary analysis of form and content". *Journal of media practice*, v. 16, n. 1, pp. 55-72.
<https://doi.org/10.1080/14682753.2015.1015801>
- Krippendorff, Klaus** (2002). *Metodología de análisis de contenido: Teoría y práctica*. Barcelona: Paidós. ISBN: 84 7509 627 1
- Lee, Bongshin; Riche, Nathalie H.; Isenberg, Petra; Carpendale, Sheelagh** (2015). "More than telling a story: Transforming data into visually shared stories". *IEEE computer graphics and applications*, v. 35, n. 5, pp. 84-90.
<https://doi.org/10.1109/MCG.2015.99>
- Loosen, Wiebke** (2018). "Four forms of datafied journalism. Journalism's response to the datafication of society". *Communicative figurations*, n. 18.
https://www.kofi.uni-bremen.de/fileadmin/user_upload/Arbeitspapiere/CoFi_EWP_No-18_Loosen.pdf
- Loosen, Wiebke; Reimer, Julius; De-Silva-Schmidt, Fenja** (2017). "Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013-2016". *Journalism*, online first.
<https://doi.org/10.1177/1464884917735691>

- Lowrey, Wilson; Hou, Jue** (2018). "All forest, no trees? Data journalism and the construction of abstract categories". *Journalism*, first online.
<https://doi.org/10.1177/1464884918767577>
- Ojo, Adegboyega; Heravi, Bahareh** (2018). "Patterns in award winning data storytelling. Story types, enabling tools and competences". *Digital journalism*, v. 6, n. 6, pp. 693-718.
<https://doi.org/10.1080/21670811.2017.1403291>
- Paraise, Silvian; Darigal, Eric** (2013). "Data-driven journalism and the public good: Computer assisted-reporters and programmer-journalists in Chicago". *New media & society*, v. 15, n. 6, pp. 853-871.
<https://doi.org/10.1177/1461444812463345>
- Rogers, Simon** (2014). "Data journalism is the new punk". *British journalism review*, v. 25, n. 2, pp. 31-34.
<https://doi.org/10.1177/0956474814538181>
- Rogers, Simon; Schwabish, Jonathan; Bowers, Danielle** (2017). *Data journalism in 2017. The current state and challenges facing the field today*. Google News Lab.
<https://newslab.withgoogle.com/assets/docs/data-journalism-in-2017.pdf>
- Sánchez-Bonvehí, Clàudia; Ribera, Mireia** (2014). "Visualización de la información en la democratización de los datos: propuestas desde el periodismo y la narratividad". *El profesional de la información*, v. 23, n. 3, pp. 311-318.
<https://doi.org/10.3145/epi.2014.may.11>
- Sandoval-Martín, María-Teresa; La-Rosa, Leonardo** (2018). "Big data as a differentiating sociocultural element of data journalism: the perception of data journalists and experts". *Communication & society*, v. 31, n. 4, pp. 193-209.
<https://revistas.unav.edu/index.php/communication-and-society/article/view/35681>
- Segel, Edward; Heer, Jeffrey** (2010). "Narrative visualization: Telling stories with data". *IEEE transactions on visualization and computer graphics*, v. 16, n. 6, pp. 1139-1148.
<https://doi.org/10.1109/TVCG.2010.179>
- Stabe, Martin** (2016). "Why the FT creates so few clickable graphics". *Financial Times*, 3 October.
<https://www.ft.com/content/c62b21c6-7feb-11e6-8e50-8ec15fb462f4>
- Stalph, Florian** (2018). "Classifying data journalism. A content analysis of daily data-driven stories". *Journalism practice*, v. 12, n. 10, pp. 1332-1350.
<https://doi.org/10.1080/17512786.2017.1386583>
- Stalph, Florian; Borges-Rey, Eddy** (2018). "Data journalism sustainability. An outlook on the future of data-driven reporting". *Digital journalism*, v. 6, n. 8, pp. 1078-1089.
<https://doi.org/10.1080/21670811.2018.1503060>
- Tabary, Constance; Provost, Anne-Marie; Trottier, Alexandre** (2016). "Data journalism's actors, practices and skills: A case study from Quebec". *Journalism*, v. 17, n.1, pp. 66-84.
<https://doi.org/10.1177/1464884915593245>
- Tandoc, Edson; Soo-Kwang, Oh** (2017). "Small departures, big continuities? Norms, values and routines in The Guardian's big data journalism". *Journalism studies*, v. 18, n. 8, pp. 997-1015.
<https://doi.org/10.1080/1461670X.2015.1104260>
- Usher, Nikki** (2019). "What is data journalism for? Cash, clicks, and cut and trys". In: Gray, Jonathan A.; Bounegru, Liliana. *The data journalism handbook 2: Towards a critical data practice*. Amsterdam: Amsterdam University Press.
<https://datajournalism.com/read/handbook/two/reflections/what-is-data-journalism-for-cash-clicks-and-cut-and-trys>
- Uskali, Turo I.; Kuutti, Heikki** (2015). "Models and stream of data journalism". *The journal of media innovations*, v. 2, n. 1, pp. 77-88.
<https://doi.org/10.5617/jmi.v2i1.882>
- Vállez, Mari; Codina, Lluís** (2018). "Periodismo computacional: evolución, casos y herramientas". *El profesional de la información*, v. 27, n. 4, pp. 759-768.
<https://doi.org/10.3145/epi.2018.jul.05>
- Wright, Scott; Doyle, Kim** (2019). "The evolution of data journalism: A case study of Australia". *Journalism studies*, v. 20, n. 13, pp. 1811-1827.
<https://doi.org/10.1080/1461670X.2018.1539343>
- Young, Mary-Lynn; Hermida, Alfred; Fulda, Johanna** (2018). "What makes for great journalism?". *Journalism practice*, v. 12, n. 1, pp. 115-135.
<https://doi.org/10.1080/17512786.2016.1270171>