


Un análisis exhaustivo de la importancia de la programación en el periodismo de datos moderno

A Comprehensive Analysis of the Importance of Programming in Modern Data Journalism



Andreas Veglis es profesor de tecnología de los medios de comunicación y director del Laboratorio de Informática de los Medios de Comunicación de la Facultad de Periodismo y Comunicación de Masas de la Universidad Aristóteles de Tesalónica. Ha sido y sigue siendo editor, miembro de consejos científicos y revisor en diversas revistas académicas. El profesor Veglis cuenta con más de 200 artículos revisados por pares sobre tecnología de los medios de comunicación y periodismo. En concreto, es autor o coautor de 12 libros y 44 capítulos de libros, ha publicado más de 115 artículos en revistas científicas y ha presentado 154 ponencias en conferencias internacionales y nacionales y ha participado en 50 proyectos de investigación nacionales e internacionales. Sus intereses de investigación incluyen la tecnología de la información en el periodismo, los nuevos medios, el periodismo algorítmico, la IA en el periodismo, la seguridad digital, el periodismo de datos, el *big data* y la verificación de contenidos.

Universidad Aristóteles de Tesalónica, Grecia 

vegkis@jour.auth.gr

ORCID: 0000-0002-0286-2304

Recibido: 31/12/2024 - Aceptado: 17/02/2025 - En edición: 26/02/2026 - Publicado: 01/07/2026

Received: 31/12/2024 - Accepted: 17/02/2025 - Early access: 26/02/2026 - Published: 01/07/2026

Resumen:

En los últimos 20 años, se ha experimentado un desarrollo significativo en el campo del periodismo de datos, que es una especialización basada en la búsqueda de noticias en los datos. Hoy en día, existe una tendencia en el periodismo de datos a emplear técnicas de programación (normalmente Python o R). Esta tendencia proviene principalmente de importantes organizaciones periodísticas que participan en proyectos de periodismo de los macrodatos o *big data* basados en el análisis de datos. El objetivo principal de este estudio es examinar la importancia de la programación en el periodismo de datos actual. Se emplea una metodología mixta, que incluye una revisión bibliográfica y un análisis de estudios de caso. Los hallazgos indican que, si bien la codificación no es necesaria en el periodismo de datos, puede considerarse una necesidad

Abstract:

In the last 20 years, the field of data journalism has seen significant development. Data journalism is a specialization based on finding news stories in data. Today, there is a trend in data journalism to use programming techniques (usually Python or R). This trend comes mainly from prominent journalism organizations engaged in big data journalism projects based on data analysis. The main thrust of this study is to examine the importance of programming in data journalism today. A mixed methodology is employed, including a literature review and case study analysis. The findings indicate that although coding is not necessary in data journalism, it may be viewed as a necessity when used for complex projects that include large datasets and specific types of visualization. The study also discusses hybrid approaches and specific

Cómo citar este artículo:

Veglis, A. (2026). Un análisis exhaustivo de la importancia de la programación en el periodismo de datos moderno. *Doxa Comunicación*, 43, pp. 21-38.

<https://doi.org/10.31921/doxacom.2878>



Este contenido se publica bajo licencia Creative Commons Reconocimiento - Licencia no comercial. Licencia internacional CC BY-NC 4.0

cuando se utiliza para proyectos complejos que incluyen grandes conjuntos de datos y tipos específicos de visualización. El estudio también analiza enfoques híbridos y herramientas específicas que tienen como objetivo cerrar la brecha entre los enfoques basados en herramientas y los basados en código en la práctica del periodismo de datos. Los resultados indican que la programación debería considerarse una habilidad complementaria que los periodistas de datos deberían considerar.

Palabras clave:

Periodismo de datos, programación, codificación, Python, R, metodología de estudio de caso.

tools that aim to bridge the gap between tool-based and code-based approaches in practicing data journalism. The results indicate that programming should be viewed as a complementary skill that should be considered by data journalists.

Keywords:

Data journalism, programming, coding, Python, R, case study methodology.

1. Introducción

Desde la concepción del periodismo, su relación con la tecnología ha sido simbiótica, donde uno moldea e influye al otro (Siapera y Veglis, 2012). Recientemente, esta relación ha dado un giro transformador con el avance hacia la datificación, un concepto introducido por Mayer-Schönberger y Cukier (2013), que lo definieron como el proceso de transformar grandes cantidades de información en un recurso escalable para su empleo como producción de conocimiento y así agregar valor económico. La introducción de la datificación en el periodismo puede verse como un cambio significativo porque altera de forma fundamental el proceso de producción, difusión y consumo de noticias. Cabe mencionar que el periodismo ya ha experimentado una transformación causada por la introducción de las tecnologías *big data* (Coddington, 2015), que ha impulsado el desarrollo del periodismo de datos, una especialidad que emplea métodos computacionales para analizar grandes conjuntos de datos y extraer información significativa (Gray, Bounegru y Chambers, 2012; Veglis y Bratsas, 2021).

El periodismo de datos se ha enfrentado al escepticismo de los periodistas reacios a los cambios tecnológicos, sin embargo, con el auge de la digitalización, los datos comenzaron a influir en las redacciones a través de métricas, asignación de recursos, selección de temas y formato de noticias, y, por lo tanto, ganaron atención en las publicaciones científicas (Angelou *et al.*, 2020). Se han empleado diversos términos, como “periodismo basado en datos”, “periodismo interactivo” y “periodismo de *big data*”, para describir esta nueva forma de periodismo (Ausserhofer *et al.*, 2020). Por lo tanto, la datificación se ha convertido paulatinamente en un parámetro importante en la producción de contenido periodístico con el periodismo de datos desempeñando un papel significativo en este proceso (Porlezza, 2023).

La datificación de las herramientas de los periodistas les ha proporcionado la capacidad para informar sobre cuestiones complejas y proporcionar narrativas con base empírica (Lewis y Westlund, 2015). También ha servido para democratizar el campo, permitiendo una mayor participación pública a través de visualizaciones interactivas y plataformas de datos abiertos (Young *et al.*, 2018). Esta transformación ha dado como resultado la redefinición del papel del periodista, que ahora incluye tareas como analista de datos, diseñador visual e intérprete de fenómenos sociopolíticos complejos (Ausserhofer *et al.*, 2020). El espacio periodístico renovado es un nuevo tipo de periodismo en forma de prácticas periodísticas alteradas a través de la creciente incorporación de datos y programación. En este sentido, Ruppert *et al.* (2017) argumentan que los datos se han convertido en una herramienta de poder, visualizando dicha influencia en los sectores público y privado, incluso en el ámbito del periodismo. Los

autores afirman que la llegada del periodismo de datos es parte de la reconfiguración general de las estructuras de poder, con datos, código y algoritmos como los actores más importantes. Por otro lado, el periodismo de datos se practicaba antes de la introducción de la codificación y se centraba en el análisis estadístico a través de la producción periodística asistida por ordenador (CAR) (Parasie y Dagiral, 2013).

En este sentido, se debate a menudo en la industria de los medios la necesidad de emplear la codificación en el periodismo de datos. Por un lado, quienes lo defienden afirman que las técnicas de codificación son importantes para los periodistas (y especialmente para los periodistas de datos) en el panorama mediático actual. Argumentan que la codificación facilita a los periodistas recopilar, analizar y presentar datos de formas nuevas y creativas, lo que les permite detectar patrones ocultos y proporcionar una explicación exhaustiva de cuestiones complejas. Además, afirman que la utilidad de la codificación para la narración y el periodismo de investigación ha aumentado (Simon, 2021). A este respecto, Porlezza (2023) argumenta que la codificación es un prerequisite para la práctica del periodismo de datos, que implica la recopilación, el análisis y la visualización de datos para contar historias y revelar lo percibido. Para gestionar la información de forma eficaz, los periodistas suelen utilizar lenguajes de programación (como Python y R) para depurar datos, realizar análisis estadísticos y crear visualizaciones de datos. Las técnicas de codificación permiten a los periodistas utilizar *big data*, la automatización del procesamiento de datos y su visualización interactiva. Sin embargo, cabe señalar que la integración de la codificación con la práctica periodística también plantea algunas dudas. En particular, los críticos insisten en que los periodistas deben centrarse en la narración y la recopilación de información, y dejar la parte técnica a los expertos en datos y a los desarrolladores. Se teme que dominar la codificación pueda distraer a los periodistas de sus tareas principales y, por lo tanto, comprometer la integridad periodística (Cruz, *sf*; Hannaford, 2015). Heravi y Lorenz (2020) afirman que el uso de la codificación en el periodismo de datos es un tema a debatir, si bien la habilidad para la codificación puede ser útil en el marco de las historias basadas en datos, cabe señalar que no todos los periodistas de datos están obligados a ser especialistas en codificación. Porlezza (2023) no desmiente esta conclusión, pero afirma que el conocimiento de la codificación puede fortalecer la capacidad del periodista en una gestión adecuada de los datos. La codificación no se considera un componente central del periodismo de datos, sino una herramienta que puede ayudar significativamente en el proceso (Heravi y Lorenz, 2020).

Es crucial encontrar el equilibrio entre las ventajas y desventajas de la programación en el periodismo, ya que los periodistas que controlan las técnicas de programación pueden extraer datos llamativos y de este modo atraer audiencias mediante el uso de dichas técnicas; sin embargo, es importante que no se comprometan los estándares ni la ética periodística. Colman *et al.* (2018) explican los problemas éticos que surgen cuando los periodistas aplican la programación en su trabajo.

Por lo tanto, el uso de la programación en el periodismo presenta tantas oportunidades como desafíos, y es fundamental que los periodistas evalúen críticamente los beneficios y riesgos asociados a la programación, considerando el contexto específico y los objetivos de sus reportajes. Al aprovechar las habilidades de programación de forma responsable y ética, los periodistas pueden maximizar el poder de la tecnología para mejorar su narrativa y ofrecer un periodismo impactante.

El objetivo principal de este estudio es examinar la importancia de la programación en la práctica del periodismo de datos en la actualidad y con este fin se formularon dos preguntas de investigación. En concreto, el debate sobre si la programación es esencial o complementaria en el periodismo de datos moderno es la base de la primera pregunta de investigación:

PI1: ¿Debe considerarse la codificación como algo imprescindible para practicar el periodismo de datos?

Además, la discusión en los casos en que se utiliza programación apunta a la idea de que las características de un proyecto específico probablemente determinan la necesidad de enfoques basados en código, lo que nos ayuda a articular la segunda pregunta de investigación:

PI2: ¿Qué características de un proyecto de periodismo de datos influyen en la decisión de utilizar la codificación?

1.1. *Practicar el periodismo de datos*

Como señaló Knight (2015), se atribuye a Simon Rogers la primera utilización del término *data journalism* en una publicación del blog Guardian Insider. Este concepto abarca un flujo de trabajo que comienza con el análisis de datos, seguido del filtrado y la visualización de estos para complementar la narrativa, como lo describe Lorenz (2010). Integra aspectos como hojas de cálculo, gráficos, análisis de datos y noticias importantes (Rogers, 2008) y se centra esencialmente en la creación de gráficos de noticias, incorporando elementos de diseño e interactividad (Bradshaw, 2018; Lorenz, 2010; Rogers, 2008). Megan Knight (2015) define el periodismo de datos como una narrativa basada principalmente en datos numéricos o que involucra datos o visualización de forma significativa.

Veglis y Bratsas (2017a) ofrecieron una definición que resume con mayor eficacia la importancia de la visualización y la interactividad en el periodismo de datos. La describen como un proceso consistente en extraer información valiosa de los datos, elaborar artículos basados en esta información e integrar visualizaciones (a veces interactivas) para facilitar la comprensión del lector o permitirle interactuar con datos relevantes. También propusieron la categorización del proceso de periodismo de datos en seis etapas distintas: Recopilación de datos, Limpieza de datos, Comprensión de datos, Validación de datos, Visualización de datos y Redacción de artículos (2017b).

Recopilación de datos: Esta fase inicial de un proyecto de periodismo de datos comienza con una pregunta que requiere datos o un conjunto de ellos que necesitan ser explorados. La recopilación de datos puede darse por varias vías, como por ejemplo la recepción directa de datos de una organización (a menudo como datos abiertos), el uso de técnicas avanzadas de búsqueda (principalmente con la ayuda de la búsqueda avanzada de Google), el empleo de motores de búsqueda que se centren en conjuntos de datos (por ejemplo, la herramienta Data Search), *web scraping* o raspado web (con herramientas de raspado o mediante aplicaciones ampliamente utilizadas que admiten raspado web), la conversión de formatos de documentos para permitir el análisis (por ejemplo, extracción de tablas de archivos PDF y exportación a formato CSV), o la recopilación de datos a través de la observación, encuestas, formularios en línea o *crowdsourcing* (Veglis y Bratsas, 2017a).

Limpieza de datos: También conocida como depuración de datos, implica la detección y corrección de registros erróneos o corruptos dentro de un conjunto de datos (Wu, 2013). Conlleva eliminar errores humanos y estandarizar los formatos de datos para que sean coherentes con otros datos utilizados por periodistas (Veglis y Bratsas, 2017b). Para esta tarea, se pueden utilizar aplicaciones de uso general como Microsoft Excel y Hojas de Cálculo de Google, que incluyen numerosas funciones, algunas de las cuales pueden utilizarse para la limpieza de datos de forma directa o indirecta (Bauzon *et al.*, 2021; Guerrero *et al.*, 2019; Setiyanto y Setiawan, 2022).

Comprensión de datos: En esta etapa, los periodistas deben descifrar varios códigos en conjuntos de datos que representan categorías, clasificaciones o ubicaciones, junto con terminología especializada, y a menudo implica añadir datos adicionales para que los existentes sean significativos. Para poder realizar estas tareas, los periodistas deben tener conocimientos de datos y ser capaces de comprender, articular y analizarlos de forma crítica (Veglis y Bratsas, 2017a). Para este propósito, puede ser necesario encontrar datos adicionales, para lo cual se pueden emplear las herramientas utilizadas en la etapa anterior, y también pueden ser útiles las herramientas para combinar conjuntos de datos. Dicho proceso se puede lograr de manera más simple con aplicaciones que admiten una fácil combinación de datos, como Tableau (Balaji, *et al.*, 2021; Batt, *et al.*, 2020; Loth, 2019) y en este sentido, es interesante señalar que la combinación de conjuntos de datos se puede lograr con la ayuda de aplicaciones de hojas de cálculo de uso general como Microsoft Excel o Google Sheets, aunque no es un proceso fácil porque estas aplicaciones no soportan dichos procesos de forma directa (Guerrero, Guerrero y Rauscher, 2019).

Validación de datos: Esto implica la verificación cruzada de los datos originales y la adquisición de información adicional de las fuentes para enriquecerlos (Silverman, 2014; Veglis, 2013). Es importante reconocer que los conjuntos de datos, como cualquier fuente, tienen sesgos y objetivos inherentes y que los periodistas deben investigar el origen, el propósito y la metodología de recopilación de un conjunto de datos (Bradshaw, 2018), algo que puede hacerse explorando el historial de creación del conjunto de datos, encontrando referencias o utilizando otras fuentes de información relacionadas con el tema de investigación (Silverman, 2014; Veglis y Bratsas, 2017a). Por lo tanto, se pueden utilizar técnicas de búsqueda de uso general como las empleadas en la etapa de compilación de datos.

Visualización de datos: Es la representación visual de información abstracta para su análisis y comunicación (Cairo, 2012). Dado que la información estadística es abstracta, su transformación en una representación física requiere comprender la percepción visual y la cognición. Una visualización de datos eficaz sigue principios de diseño basados en la percepción humana (Card *et al.*, 1999; Few, 2013). Con este fin, diversas aplicaciones (de escritorio o en línea) permiten crear visualizaciones (Protopsaltis *et al.*, 2020), y la mayoría de ellas ofrecen múltiples plantillas para crear diferentes tipos de visualización, algunos ejemplos típicos son Infogram, Piktochart, DataWrapper y Tableau.

1.2. Soluciones disponibles para la adopción de la codificación

En el sector de los medios de comunicación, existen dos soluciones alternativas para el uso de la codificación: los lenguajes de programación Python y R (en algunos casos Ruby) (Parasie y Dagiral, 2013), que ofrecen capacidades robustas para el procesamiento y la visualización de datos. Python, que se caracteriza por su simplicidad y legibilidad, puede utilizarse para tareas como el *web scraping*, el análisis de datos y la visualización interactiva, además de ofrecer amplias bibliotecas, como Pandas y Matplotlib, que permiten a los periodistas gestionar eficientemente grandes conjuntos de datos y presentar información compleja en un formato accesible. Python admite la integración con *frameworks web*, una característica que facilita la creación de presentaciones online interactivas (Lutz, 2001).

El lenguaje de programación R goza de gran prestigio por sus capacidades de análisis estadístico y visualización de datos, lo que le convierte en la opción ideal para el estudio exhaustivo de conjuntos de datos complejos, y además ofrece numerosos paquetes

que facilitan el descubrimiento de patrones y tendencias ocultos en los datos. R Markdown por su parte ofrece una plataforma que permite combinar código, visualizaciones y narrativas que mejoran la redacción en el periodismo de datos (Crawley, 2012). Ambos, Python y R, son de código abierto, lo que fomenta un entorno colaborativo con una amplia comunidad de desarrolladores que contribuyen a su constante mejora y ofrecen soporte. Este aspecto es esencial para los periodistas, que a menudo dependen de recursos compartidos y del apoyo de la comunidad para aprender y resolver problemas. Por otro lado, existe una gran cantidad de recursos educativos y tutoriales en línea para ambos lenguajes de programación, que facilita su adopción por parte de los periodistas (Appelgren y Nygren, 2014; Bradshaw, 2018; Bounegru *et al.*, 2018; Hamilton, 2016; Knight, 2015).

Aunque configurar Python o R en un ordenador no es tarea fácil, existen soluciones alternativas que utilizan la computación en la nube, como Google Colab, que permite escribir y ejecutar código Python a través de un navegador (Tock, 2019). En concreto, Google Colab es un entorno Jupyter Notebook gratuito basado en la nube que permite escribir y ejecutar Python en un navegador web con acceso gratuito a las GPU y TPU, lo que es idóneo para el aprendizaje automático y la ciencia de datos (Burke, 2023). Por lo tanto, los periodistas no necesitan configurar un ordenador personal para usar Python.

Otro desarrollo reciente es la integración de Excel con Python donde Microsoft combina las capacidades computacionales avanzadas de Python con las funciones de hoja de cálculo de Excel (Kinnestand, 2023). Los usuarios pueden usar Python directamente en las celdas de Excel aprovechando sus bibliotecas para la manipulación y visualización de datos¹. Esta integración en Excel permite a los periodistas migrar poco a poco a la programación en Python mientras utilizan el entorno habitual de Excel.

Otro enfoque disponible para trabajar con Python posibilita el uso de interfaces de IA como ChatGPT, con la ayuda de complementos específicos². Sin embargo, estas soluciones aún se encuentran en una etapa temprana de desarrollo y sus complementos aparecen y desaparecen con rapidez. Con la ayuda de complementos específicos, ChatGPT podrá servir como interfaz intuitiva con sus capacidades de procesamiento de lenguaje natural para periodistas que no posean grandes conocimientos de programación o carezcan de interés o tiempo para aprender a programar. Esto podría implicar la consulta de conjuntos de datos, la generación de resúmenes estadísticos o la creación de visualizaciones mediante comandos de texto. De este modo, el uso de Python para el periodismo de datos se simplifica, permitiendo a los periodistas centrarse más en la narrativa y menos en los aspectos técnicos del manejo de datos (Knight, 2015). El uso de ChatGPT como interfaz con Python también puede considerarse una democratización del periodismo de datos, porque permite que un mayor número de periodistas participen en la narración basada en datos. Este enfoque se alinea con la tendencia hacia formas de periodismo más accesibles y colaborativas (Hamilton, 2016). La IA, al reducir la barrera de entrada al periodismo de datos, juega un papel importante en la evolución del campo haciéndolo más inclusivo y diverso.

1 <https://www.microsoft.com/es-es/microsoft-365/python-en-excel>

2 <https://openai.com/index/chatgpt-plugins/>

1.3. ¿Periodismo de datos con herramientas o basado en código?

Partiendo de la discusión previa, se realiza una comparación directa entre los enfoques basados en herramientas y en código para la práctica del periodismo de datos. La comparación se basa en las posibilidades que ofrece cada enfoque para facilitar las distintas etapas del desarrollo del periodismo de datos, como lo describen Veglis y Bratsas (2017b) y se presenta en la Sección 1.1. Para facilitar dicha comparación, se ha elaborado la Tabla 1, que incluye herramientas indicativas (servicios y aplicaciones) que pueden emplearse en las cinco etapas del periodismo bajo el enfoque basado en herramientas, junto con las capacidades disponibles que ofrecen los dos lenguajes de programación mencionados (Python y R), ya sea de manera directa o con la ayuda de bibliotecas o paquetes específicos.

Tabla 1. Comparación directa de las posibilidades que ofrecen los enfoques basados en herramientas y basados en código

Etapa del periodismo de datos	Herramientas utilizadas	Posibilidades de codificación (Python/R)
Recopilación de datos	<i>Motores de búsqueda:</i> Google, Bing, etc. <i>Herramienta de búsqueda de datos:</i> Búsqueda de datos de Google <i>Web scraping:</i> Microsoft Excel, Google Sheet. <i>Raspado de PDF:</i> Tabula	Realice búsquedas web en Python a través de la API de búsqueda personalizada de Google y en R con la ayuda del paquete GoogleSearchR. El raspado web es compatible tanto con Python (es decir, BeautifulSoup 4) como con R (es decir, inversión).
Limpieza de datos	<i>Aplicaciones de propósito general:</i> Microsoft Excel, Google Sheet. <i>Herramienta de limpieza dedicada:</i> OpenRefine	La limpieza de datos se puede implementar tanto con Python como con R, ya que ambos lenguajes de programación admiten operaciones de limpieza de datos.
Comprensión de datos	<i>Aplicaciones/herramientas que admiten la combinación de conjuntos de datos:</i> Tableau, Microsoft Excel, Google Sheet.	Python admite la combinación de conjuntos de datos, principalmente mediante la biblioteca Pandas. Lo mismo aplica para R.
Validación de datos	No se utilizan aplicaciones ni herramientas adicionales	No aplicable
Visualización de datos	<i>Herramientas/aplicaciones de visualización:</i> Infogram, Datawrapper, Tableau, Piktochart, etc.	Ambos lenguajes de programación poseen amplias capacidades para crear visualizaciones.

Aunque de la Tabla 1 se desprende fácilmente que ambos enfoques son compatibles con todas las etapas del periodismo de datos, cabe destacar que, en el enfoque basado en código, todas las etapas del periodismo de datos se gestionan mediante la misma plataforma (es decir, Google Colab), mientras que en el enfoque basado en herramientas no existe la interoperabilidad, por lo que los periodistas deben exportar, guardar, importar y cargar manualmente los datos de una herramienta a otra. Esto añade complejidad al proceso y genera la posibilidad de errores que pueden ralentizar la producción de artículos de periodismo de datos.

2. Metodología

Para este estudio, se ha seleccionado un enfoque mixto que incluye una revisión bibliográfica y estudios de caso. La metodología de estudio de caso es un enfoque de investigación de amplio uso y que permite respaldar el estudio de fenómenos o procesos complejos. Se trata de una técnica de investigación cualitativa empleada comúnmente en diversas disciplinas, como las ciencias sociales, la educación y los negocios. Esta metodología implica una exploración exhaustiva de uno o más casos, utilizando diversas técnicas de recopilación de datos, como entrevistas, documentos y observaciones, y ofrece una perspectiva holística que genera información rica y detallada sobre el tema de investigación (Kratochwill *et al.*, 2013; Tellis, 1997).

A pesar de sus ventajas porque proporciona perspectivas integrales y profundas, y contribuye tanto al desarrollo teórico como a las aplicaciones prácticas, como sugiere Yin (2014), la metodología de estudio de caso presenta limitaciones. Baxter y Jack (2008) advierten sobre la limitada generalización de los hallazgos de los estudios de caso debido al hecho de que se centran en solo uno o pocos casos.

Para el estudio, se seleccionaron dos artículos de periodismo de datos, el primero se nutre de datos que pueden caracterizarse como *big data* y el segundo, en datos de tamaño moderado. Ambos fueron publicados por iMedLAB³, una organización sin ánimo de lucro que busca mejorar la transparencia, la credibilidad y la independencia en el periodismo, fundada en 2018 con financiación de la Fundación Stavros Niarchos (SNF)⁴. Se eligió esta organización porque publica artículos en inglés (y, en algunos casos, también en griego) además de conjuntos de datos que se han empleado en sus artículos de periodismo de datos. También proporciona información sobre el proceso y las herramientas empleadas durante la recopilación de conjuntos de datos y la creación de visualizaciones, aspecto crucial para el propósito de este estudio. Los dos artículos seleccionados son los siguientes:

Artículo 1: Aplicación MyCoast: Análisis de datos sobre contratos de arrendamiento de playas en Grecia (publicado el 2 de agosto, 2024), escrito por Thanasis Trouboukis, disponible en:

<https://lab.imedd.org/en/analysame-ta-dedomena-gia-tis-ellinikes-paralies-tou-my-coast/>

Artículo 2: En Grecia, en las elecciones europeas más personas se abstuvieron de votar que las que participaron (publicado el 18 de junio, 2024), escrito por Kelly Kiki y Chrysoula Marinou, disponible en:

<https://lab.imedd.org/en/more-people-abstained-from-voting-than-participated-in-the-european-elections/>

3 <https://lab.imedd.org/>

4 <https://www.snf.org/es/>

3. Análisis de los estudios de caso

A continuación, se presenta brevemente el contenido de cada artículo. El artículo 1 informa sobre un problema bastante común en Grecia durante el verano, utilizando datos de la aplicación MyCoast, una aplicación que permite a los griegos verificar si se cumplen los términos de la concesión en las playas que visitan y, en caso necesario, enviar una denuncia. El artículo incluye varias visualizaciones estáticas y dos más complejas (ver figura 1) que ofrecen cierto grado de interactividad (el usuario puede desplazarse y modificar los datos presentados en la visualización) y un mapa detallado de los datos utilizados, el conjunto de datos también se puede descargar en formato CSV. La inspección del conjunto de datos reveló que comprendía 9170 filas con un tamaño total de 4,26 MB, lo que da como resultado un conjunto de datos bastante amplio que puede caracterizarse como *big data*. El conjunto de datos estaba disponible públicamente y, por lo tanto, podemos asumir que se obtuvo directamente del gobierno griego sin ningún proceso especial (es decir, raspado web, búsqueda). El equipo de iMed estudió un total de 9170 concesiones costeras y analizó 3957 contratos para obtener las consideraciones financieras de cada arrendamiento, habiendo gestionado y vinculado los tipos de datos relacionados. Cabe destacar que el 43,0 % de los registros de la aplicación MyCoast no incluían contratos de arrendamiento relevantes, y, por tanto, el equipo implementó un filtro, excluyendo municipios (es decir, Paros, Mykonos, Thira y Patmos) para garantizar que los resultados fueran representativos de la realidad. Es evidente que todo el proceso (gestión, limpieza y filtrado) ha sido bastante complejo y ha implicado el uso de programación.

Aunque no hay evidencias acerca de si las visualizaciones se crearon con herramientas específicas, ya que no está claro que emplean diseños basados en plantillas ni incluyen elementos específicos de la aplicación (marcas de agua, fuentes de marca, componentes de interfaz de usuario reconocibles), sí tienen un estilo consistente y preciso, y utilizan esquemas de color simples (Figura 1). Además, los tipos de visualizaciones, los mapas térmicos (el primero muestra la concentración de concesiones y el segundo el costo por metro cuadrado, ya que el tamaño del punto y el color se relacionan con un mayor costo) son muy especializados y es muy difícil crearlos con herramientas estándar de análisis de datos (DataWrapper y Flourish). En general, existe una alta probabilidad de que las visualizaciones del artículo 1 se crearan mediante programación. También es probable que otras etapas del proyecto de periodismo de datos (limpieza, comprensión y validación de datos) se completaran utilizando métodos similares.

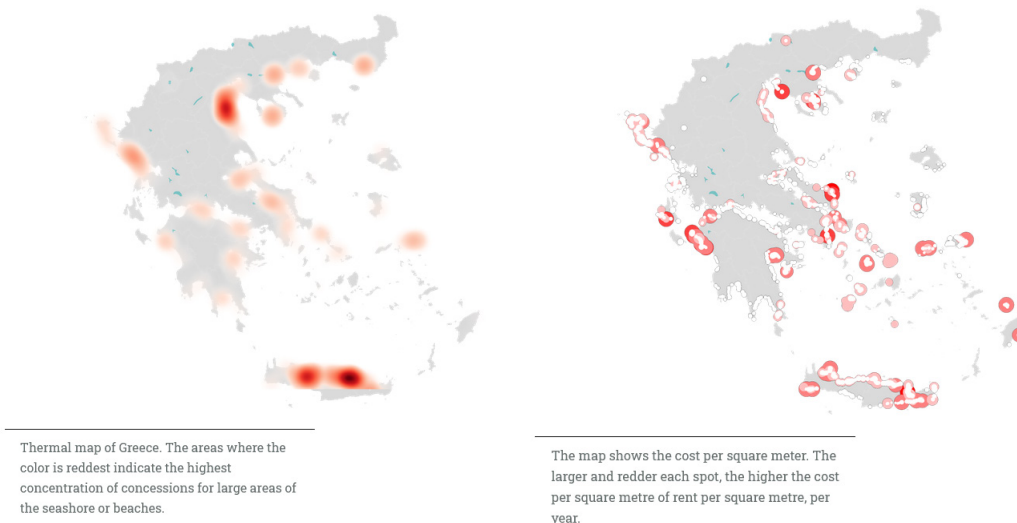


Figura 1. Dos visualizaciones estáticas que muestran (a) la mayor concentración de concesiones de grandes áreas de costa o playas y (b) el coste de arrendamiento por metro cuadrado (<https://lab.imedd.org/en/analysame-ta-dedomena-gia-tis-ellinikes-paralies-tou-my-coast/>)

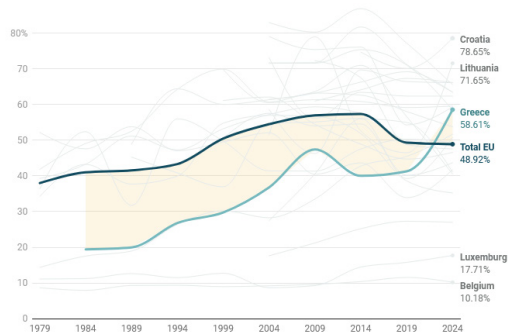
El artículo 2 analiza la cuestión de la participación en las elecciones europeas en Grecia, concretamente, presenta datos que indican que la abstención en dichas elecciones de 2024 ha aumentado en 17,3 puntos porcentuales en comparación con las europeas anteriores de 2019, aunque la abstención se mantuvo estable en la Unión Europea en su conjunto. El artículo incluye una visualización estática y dos interactivas (Figura 2) e integra otra sobre la cuestión de la abstención creada mediante una investigación transfronteriza de EDJNet⁵. Estas visualizaciones se crearon utilizando DataWrapper y Flourish y el artículo incluye una opción para que los lectores descarguen los conjuntos de datos utilizados, incluyendo además cuatro archivos txt y dos pequeños archivos CSV con los datos utilizados para la creación de las visualizaciones de DataWrapper (una opción integrada en las visualizaciones). Como afirman los autores del artículo, los datos se recopilaban desde varias fuentes oficiales online y se extrajeron de archivos PDF (disponibles online). Si bien el *scraping* puede implicar el uso de programación, el tamaño moderado de los datos y el uso de DataWrapper para las visualizaciones nos llevan a suponer la no utilización de programación en este artículo. Es posible que se hayan empleado diversas herramientas que permiten la extracción de tablas de archivos PDF, así como para la limpieza, la comprensión y la validación de datos y también es probable que se empleara una aplicación de hoja de cálculo, apto para la gestión de conjuntos pequeños de datos.

Cabe resaltar que, si bien es evidente que las visualizaciones se crearon con la ayuda de aplicaciones específicas, los conjuntos de datos en archivos CSV y TXT también se han podido trabajar en Python y R (en el caso de R, el proceso es algo más complejo).

⁵ <https://abstencao.divergente.pt/es/home>

Abstention from European elections by year

For the first time, Greece is among the countries with the highest abstention rates from the 2024 European elections: more people abstained than participated in the elections on June 9.



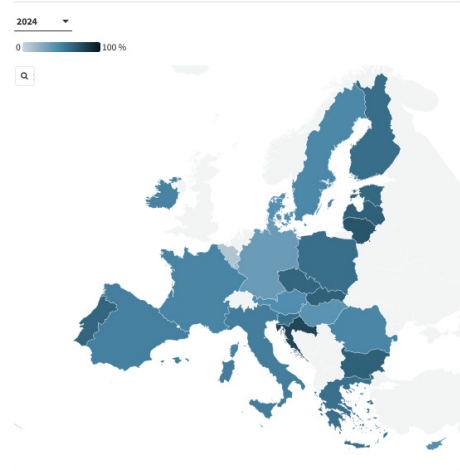
Note: For countries that joined the EU after 1979, abstention rates from European elections held in these member states in the year of their accession are not shown. IMEDD Lab calculated abstention rates based on the available European Parliament data (June 14, 2024) on participation in European elections by country.

Chart: Kelly Kiki, Chrysoula Marinou/IMEDD Lab - Source: European Parliament - Get the data - Download image - Created with Datawrapper

Abstention from the European elections

Croatia (78,65%) and Lithuania (71,65%) have the highest abstention rates in 2024. The abstention rate in Greece reached 58,61% - up from 41,31% in 2019

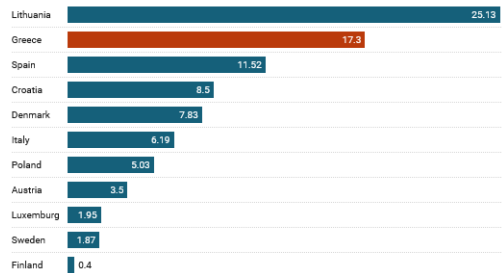
By hovering your mouse over the map, you can see the percentage (%) of abstention in each country in each selection year



Data source: European Parliament - Map: Chrysoula Marinou, Kelly Kiki /IMEDD Lab
 Note: IMEDD Lab calculated the abstention rates based on the available data from the European Parliament (14 June 2024) on country participation in the European elections.

What was the increase in abstention by country in 2024

Percentage point difference compared to 2019



Note: Only countries where abstention from the 2024 European elections increased are included. The abstention rates were calculated by IMEDD Lab based on available data from the European Parliament (June 14, 2024) on participation in the European elections by country.

Chart: Kelly Kiki, Chrysoula Marinou/IMEDD Lab - Source: European Parliament - Get the data - Download image - Created with Datawrapper

Figura 2. Visualizaciones (estáticas e interactivas) de diversos datos sobre la abstención en las elecciones europeas (<https://lab.imedd.org/en/more-people-abstained-from-voting-than-participated-in-the-european-elections/>).

4. Discusión

Según los resultados de los dos estudios de caso examinados en la sección anterior, esta sección trata sobre las preguntas de investigación del estudio.

Con respecto a la **PI1**, preguntando si la codificación se considera una necesidad para practicar el periodismo de datos, podemos responder que la codificación no es una necesidad absoluta y, en muchos casos, el enfoque basado en herramientas puede proporcionar resultados suficientes. El Artículo 2 (que se presentó en la Sección 3), ha dejado claro que el enfoque basado en herramientas se puede emplear con buenos resultados. Heravi y Lorenz (2020) informaron que, aunque el 69% de los periodistas de datos utilizan alguna forma de codificación en su trabajo, muchos otros emplean principalmente software de hojas de cálculo y herramientas de análisis visual. Según Borges-Rey existen diferencias importantes entre los distintos medios de comunicación (2016). Por lo general, las grandes organizaciones de medios emplean equipos que incluyen profesionales con habilidades técnicas en codificación y trabajo con datos, mientras que, en las organizaciones medianas y pequeñas, están más a favor del enfoque basado en herramientas. Otras empresas de comunicación crean equipos compuestos de periodistas con conocimientos en codificación y reporteros de investigación tradicionales que trabajan con datos, promoviendo de este modo la codificación como algo complementario y no esencial (Parasie y Dagiral, 2013). Por otro lado, Young y Hermida encontraron un aumento del uso de las herramientas *no code* y *low code* (la programación sin código) en las salas de redacción (2015), que pueden hacer posible el análisis de datos complejos para periodistas sin habilidades de codificación. Parasie y Dagiral también consideran el uso de la programación en el periodismo de datos como una evolución y no como un requisito previo (2013). No obstante, la mayoría de los periodistas de datos muestran un gran interés en aprender técnicas de programación para así ser más productivos en el campo (Heravi, 2018).

En conclusión, si bien la programación no puede considerarse un requisito indispensable para ejercer el periodismo de datos, sí se está convirtiendo en una habilidad importante y en algunos casos donde los proyectos de periodismo de datos involucran *big data*, es el único enfoque viable.

En referencia a la **PI2**, concerniente a las características de un proyecto de periodismo de datos que influyen en la decisión de utilizar codificación, se pueden extraer algunas conclusiones de los dos estudios de caso en relación con la creación de visualizaciones. Concretamente, donde se hayan empleado visualizaciones específicas imposibles de crear con la ayuda de las herramientas de análisis de datos disponibles, la única opción es usar codificación. Esto significa que, si se requieren visualizaciones como las del artículo 1, donde se deben marcar áreas específicas en un mapa, es bastante difícil o incluso imposible encontrar herramientas de análisis de datos que puedan crear dicha visualización. Sin embargo, en el artículo 2, las visualizaciones no son tan especializadas (gráfico de líneas, gráfico de barras y mapa coroplético) y, por lo tanto, se han podido emplear herramientas comunes de análisis de datos, aunque este tipo de visualizaciones también se pueden crear mediante codificación, ya que ambos lenguajes de programación (Python y R) ofrecen amplias bibliotecas y paquetes.

Sobre las demás etapas de desarrollo de los proyectos de periodismo de datos en nuestros estudios de caso, no existen indicaciones sobre que enfoque se ha empleado, pero se pueden hacer algunas especulaciones. Concretamente, en el Artículo 1, los datos se encuentran en un único archivo PDF, y aunque de tamaño considerable, es manejable con aplicaciones de hojas de cálculo comunes (Microsoft Excel y Google Spreadsheet). Pero, dado que es probable que las visualizaciones se crearan mediante

programación, existe mucha posibilidad de que otras etapas de desarrollo (por ejemplo, limpieza de datos, comprensión de datos y validación de datos) se completaran mediante codificación. En el caso del Artículo 2, los datos empleados estaban disponibles en cuatro archivos diferentes (con tamaños relativamente pequeños) y dado que los conjuntos de datos provienen de diferentes fuentes y, como se ha indicado, se extrajeron de archivos PDF, existe la posibilidad de que se haya empleado un enfoque de programación. Sin embargo, si consideramos que a) existen diversas herramientas que pueden respaldar este proceso y b) las visualizaciones se crearon mediante aplicaciones de análisis de datos, es más probable que no se haya empleado programación para las necesidades del artículo 2. Otros factores importantes son el tamaño del conjunto de datos para seleccionar un enfoque adecuado (basado en herramientas o en código), así como los requisitos de visualización. También cabe destacar que es poco frecuente emplear un enfoque mixto (basado en herramientas y en código) durante un proyecto.

Además de las características del periodismo de datos que se examinaron en el análisis de caso, existen otros parámetros, derivados de la literatura científica, que pueden afectar la elección del enfoque para desarrollar un proyecto de periodismo en este campo. El interés personal en la tecnología, la exposición previa a conceptos de programación y los beneficios profesionales percibidos son parámetros importantes que influyen en la adopción de la codificación por parte de los periodistas (Appelgren y Nygren, 2014). El contexto organizacional (tamaño de la sala de redacción, recursos disponibles y apoyo institucional) tiene su importancia en la selección del enfoque basado en código (Fink y Anderson, 2015). Así mismo, Joung y Hermida destacaron el papel de la infraestructura técnica y la integración del flujo de trabajo como parámetros positivos para adoptar el enfoque basado en código. Coddington destacó como una razón para no elegir el enfoque de programación (2015) el factor de trabajar con grandes conjuntos de datos y la necesidad de actualizaciones frecuentes.

5. Conclusiones

Este estudio ha analizado el uso de la programación en el periodismo de datos moderno, y sus conclusiones revelan la relación entre la codificación y la práctica del periodismo de datos. En primer lugar, si bien la codificación no es imprescindible para el periodismo de datos, se ha convertido en una técnica valiosa en este campo. El análisis demuestra que tanto los enfoques basados en herramientas como en código pueden funcionar en muchos casos para proyectos de periodismo de datos. La elección entre ambos depende de muchos factores, como la complejidad del proyecto, el tamaño del conjunto de datos, las necesidades de repetición y los requisitos de visualización.

Específicamente, el análisis de caso muestra que la decisión de codificar se debe principalmente al tamaño y la complejidad de los datos que se analizan, y que los conjuntos de datos más grandes se benefician más de un enfoque basado en código. Esta decisión también se debe a la especificidad de los requisitos de visualización, donde las que son únicas o complejas requieren soluciones de codificación. Otros factores incluyen la integración y automatización del flujo de trabajo, los recursos y la infraestructura técnica disponibles, así como la formación y las habilidades técnicas de los periodistas. Lo más importante es destacar que no existe un enfoque único para el periodismo de datos, que la codificación puede proporcionar más flexibilidad y capacidades para el análisis y la visualización de datos complejos, pero que los enfoques tradicionales basados en herramientas siguen siendo viables y funcionan para muchos proyectos, sobre todo aquellos con conjuntos de datos de tamaño moderado y necesidades de visualización estándar. Nuevas herramientas y plataformas, como Google Colab y la integración de Python en

Excel, están derribando gradualmente las barreras técnicas para la codificación en el periodismo. Esto significa que la distinción entre enfoques codificados y no codificados será más tenue a medida que los enfoques híbridos se vuelvan más accesibles para periodistas con menos formación técnica. Estos hallazgos tienen implicaciones significativas para la educación periodística y las prácticas en las salas de redacción, implicando que, si bien las habilidades de programación son útiles, deben considerarse una herramienta complementaria, no imprescindible, en el conjunto de conocimientos del periodista de datos. El enfoque debe centrarse en principios periodísticos sólidos y capacidad narrativa, mientras que las habilidades técnicas deben servir para enriquecer, y no para reemplazar, la esencia de la profesión.

No obstante, cabe destacar que este estudio presenta ciertas limitaciones, ya que incluye una pequeña muestra de estudios de caso de un solo medio de comunicación. Además, no es fácil definir con total certeza el enfoque empleado para cada artículo de periodismo de datos sin obtener la información necesaria del propio equipo periodístico que desarrolló dichos proyectos. Sin embargo, el presente artículo debe tratarse como un estudio de concepto, con el objetivo de introducir un debate sobre el papel de la programación en el periodismo de datos moderno.

A medida que surgen nuevas herramientas y tecnologías mediante la realización de encuestas a periodistas especializados, las investigaciones futuras podrían centrarse en el impacto a largo plazo de la alfabetización programática en las salas de redacción, la eficacia de los enfoques híbridos que combinan métodos tradicionales y codificados, y la evolución del periodismo de datos. Esto podría proporcionar a la comunidad científica nuevas perspectivas sobre cómo el periodismo de datos sigue adaptándose y evolucionando con los avances tecnológicos y las cambiantes necesidades de las salas de redacción.

6. Agradecimientos

Este artículo ha sido traducido al inglés por Brian O'halloran, a quien se agradece su trabajo

7. Contribuciones específicas de cada autor

	Nombre y apellidos
Concepción y diseño del trabajo	Andreas Veglis
Metodología	Andreas Veglis
Recogida y análisis de datos	Andreas Veglis
Discusión y conclusiones	Andreas Veglis
Redacción, formato y revisión	Andreas Veglis

8. Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses

9. Referencias bibliográficas

- Angelou, I., Kansaras, V., Kourkouridis, D., & Veglis, A. (2020). Social media followership as a predictor of news website traffic. *Journalism Practice*, 14(6), 730-748. <https://doi.org/10.1080/17512786.2019.1635040>
- Appelgren, E., & Nygren, G. (2014). Data journalism in Sweden: Introducing new methods and genres of journalism into 'old' organizations. *Digital Journalism*, 2(3), 394-405. <https://doi.org/10.1080/21670811.2014.884344>
- Ausserhofer, J., Gutounig, R., Oppermann, M., Matiasek, S., & Goldgruber, E. (2020). The datafication of data journalism scholarship: Focal points, methods, and research propositions for the investigation of data-intensive newswork. *Journalism*, 21(7), 950-973. <https://doi.org/10.1177/1464884917700667>
- Balaji, N., Pai, B. K., Bhat, B., & Praveen, B. (2021, February). Data visualization in splunk and Tableau: a case study demonstration. In *Journal of Physics: Conference Series* (Vol. 1767, No. 1, p. 012008). IOP Publishing. <https://doi.org/10.1088/1742-6596/1767/1/012008>
- Batt, S., Grealis, T., Harmon, O., & Tomolonis, P. (2020). Learning Tableau: A data visualization tool. *The Journal of Economic Education*, 51(3-4), 317-328. <https://doi.org/10.2139/ssrn.3438993>
- Bauzon, J., Murphy, C., & Wahi-Gururaj, S. (2021). Using macros in Microsoft Excel to facilitate cleaning of research data. *Journal of Community Hospital Internal Medicine Perspectives*, 11(5), 653-657. <https://doi.org/10.1080/20009666.2021.1954282>
- Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, 13(4), 544-559. <https://doi.org/10.46743/2160-3715/2008.1573>
- Borges-Rey, E. (2016). Unravelling data journalism: A study of data journalism practice in British newsrooms. *Digital Journalism*, 4(3), 336-352. <https://doi.org/10.1080/17512786.2016.1159921>
- Bounegru, L., Gray, J., Venturini, T., & Mauri, M. (2018). A field guide to 'Fake News' and other information disorders. *A Field Guide to "Fake News" and Other Information Disorders: A Collection of Recipes for Those Who Love to Cook with Digital Methods*, Public Data Lab, Amsterdam (2018). <https://doi.org/10.2139/ssrn.3097666>
- Bradshaw, P. (2017). Computational thinking and the next wave of data journalism. *Online Journalism Blog*. onlinejournalismblog.com/2017/08/03/computational-thinking-data-journalism
- Burke, J. (2023). Why and how to use Google Colab. *TechTarget*. <https://www.techtarget.com/searchenterpriseai/tutorial/Why-and-how-to-use-Google-Colab>
- Cairo, A. (2012). *The Functional Art: An introduction to information graphics and visualization*. New Riders. <https://doi.org/10.12775/tsb.2014.025>
- Card, S. K., Mackinlay, J., & Shneiderman, B. (Eds.). (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann. ISBN: 1558605339

- Coddington, M. (2015). Clarifying Journalism's Quantitative Turn: A typology for evaluating data journalism, computational journalism, and computer-assisted reporting. *Digital Journalism*, 3(3), 331-348. <https://doi.org/10.1080/21670811.2014.976400>
- Colman, F., Bühlmann, V., O'Donnell, A., & van der Tuin, I. (2018). *Ethics of coding: A report on the algorithmic condition*. https://ualresearchonline.arts.ac.uk/id/eprint/14236/1/Ethics_of_Coding_A_Report_on_the_Algorit.pdf
- Crawley, M. J. (2012). *The R book*. John Wiley & Sons. <https://onlinelibrary.wiley.com/doi/epdf/10.1002/9781118448908>
- Cruz, J. (n.d.). *Why journalists need data and coding skills more than ever*. International Journalists' Network. Available at <https://ijnet.org/en/story/why-journalists-need-data-and-coding-skills-more-ever>
- Few, S. (2013). 35. Data visualization for human perception. *The Encyclopedia of Human-Computer Interaction*, International Design Foundation. https://data.europa.eu/sites/default/files/course/4.1_DataVisualizationForHumanPerception.pdf
- Fink, K., & Anderson, C. W. (2015). Data Journalism in the United States: Beyond the "usual suspects". *Journalism Studies*, 16(4), 467-481. <https://doi.org/10.1080/1461670x.2014.939852>
- Gray, J., Bounegru, L., & Chambers, L. (2012). *The Data Journalism Handbook*. O'Reilly Media. https://datajournalismhandbook.org/uploads/first_book/DataJournalismHandbook-2012.pdf
- Guerrero, H. (2019). *Excel data analysis*. Springer International Publishing. <https://educons.edu.rs/wp-content/uploads/2020/05/2019-Excel-Data-Analysis.pdf>
- Hamilton, J. T. (2016). *Democracy's Detectives: The Economics of Investigative Journalism*. Harvard University Press. <https://doi.org/10.4159/9780674973572-001>
- Hannaford, L. (2015). Computational journalism in the UK newsroom. *Journalism education*, 4(1), 7-21. <https://e-space.mmu.ac.uk/575953/1/Article%20Computational%20Journalism.pdf>
- Heravi, B. R. (2018). Data journalism in 2017: A summary of results from the global data journalism survey. In *Transforming Digital Worlds: 13th International Conference*, iConference 2018, Sheffield, UK, March 25-28, 2018, Proceedings 13 (pp. 107-113). Springer International Publishing. https://doi.org/10.1007/978-3-319-78105-1_13
- Heravi, B. R., & Lorenz, M. (2020). Data journalism practices globally: Skills, education, opportunities, and values. *Journalism and media*, 1(1), 26-40. <https://doi.org/10.3390/journalmedia1010003>
- Kinnestand, S. (2023). Introducing Python in Excel: The Best of Both Worlds for Data Analysis and Visualization, Microsoft 365 Blog. <https://rb.gy/nshr7e>.
- Knight, M. (2015). Data Journalism in the UK: a preliminary analysis of form and content. *Journal of Media Practice*, 16:1, pp. 55-72. <https://doi.org/10.1080/14682753.2015.1015801>
- Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, 34(1), 26-38. <https://doi.org/10.1177/0741932512452794>
- Lewis, S. C., & Westlund, O. (2015). Big Data and Journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism*, 3(3), 447-466. <https://doi.org/10.1080/21670811.2014.976418>

- Lorenz, M. (2010). Data driven journalism: What is there to learn? Presented at *IJ-7 Innovation Journalism Conference*, 7-9 June, Stanford, CA. <https://www.slideshare.net/slideshow/datadriven-journalism-what-is-there-to-learn/4644302>
- Loth, A. (2019). *Visual analytics with Tableau*. John Wiley & Sons. ISBN: 9781119560203
- Lutz, M. (2001). *Programming python*. O'Reilly Media, Inc. ISBN: 9781449398712
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt. ISBN: 0544227751
- Parasie, S., & Dagiral, E. (2013). Data-driven journalism and the public good: “Computer-assisted-reporters” and “programmer-journalists” in Chicago. *New Media & Society*, 15(6), 853-871. <https://doi.org/10.1177/1461444812463345>
- Porlezza, C. (2024). The datafication of digital journalism: A history of everlasting challenges between ethical issues and regulation. *Journalism*, 25(5), 1167-1185. <https://doi.org/10.1177/14648849231190232>
- Protopsaltis, A., Sarigiannidis, P., Margounakis, D., & Lytos, A. (2020, August). Data visualization in internet of things: tools, methodologies, and challenges. In *Proceedings of the 15th international conference on availability, reliability and security* (pp. 1-11). <https://doi.org/10.1145/3407023.3409228>
- Rogers, S. (2008). Turning official figures into understandable graphics, at the press of a button. Inside the Guardian Blog, <https://www.theguardian.com/help/insideguardian/2008/dec/18/unemploymentdata>
- Ruppert, E., Isin, E., & Bigo, D. (2017). Data politics. *Big data & society*, 4(2). DOI: 10.1177/2053951717717749
- Setiyanto, S., & Setiawan, I. (2022). Data science with excel. *International Journal of Computer and Information System (IJCIS)*, 3(3), 104-110. <https://doi.org/10.29040/ijcis.v3i3.79>
- Siapera, E., & Veglis, A. (Eds.). (2012). *The handbook of global online journalism*. John Wiley & Sons. <https://doi.org/10.1002/9781118313978>
- Silverman, C. (2014). Verification Handbook. European Journalism Centre <http://verificationhandbook.com/downloads/verification.handbook.pdf>
- Simon, B. (2021). Coding With Data in the Newsroom, in Bounegru, L., & Gray, J. (eds). *The Data Journalism Handbook: Towards a Critical Data Practice* (p. 124-127). Amsterdam University Press. <https://doi.org/10.1017/9789048542079.019>
- Tellis, W. (1997). Application of a case study methodology. *The qualitative report*, 3(3), 1-19. <https://doi.org/10.46743/2160-3715/1997.2015>
- Tock, K. (2019). Google CoLaboratory as a platform for Python coding with students. *RTSRE Proceedings*, 2(1). <https://doi.org/10.32374/rtsre.2019.013>
- Veglis, A. (2013). Education of Journalists on ICTs: Issues and Opportunities. *Journal of Applied Journalism & Media Studies*, 2(2), 265-279. https://doi.org/10.1386/ajms.2.2.265_1
- Veglis, A., & Bratsas, C. P. (2021). Data Journalism: Definition, Skills, Difficulties, and Perspectives. In *Encyclopedia of Information Science and Technology*, Fifth Edition (pp. 1140-1151). IGI Global. <https://doi.org/10.4018/978-1-7998-3479-3.ch079>

- Veglis, A., & Bratsas, C. (2017a). Journalists in the age of Data Journalism: The case of Greece. *Journal of Applied Journalism & Media Studies*, 6(2), 225-244. https://doi.org/10.1386/ajms.6.2.225_1
- Veglis, A., & Bratsas, C. (2017b). Towards a taxonomy of data journalism. *Journal of Media Critiques*, 3(11), 109-121. <https://doi.org/10.17349/jmc117309>
- Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Sage publications. ISBN 9781452242569
- Young, M. L., Hermida, A., & Fulda, J. (2018). What makes for great data journalism? A content analysis of data journalism awards finalists 2012-2015. *Journalism practice*, 12(1), 115-135. <https://doi.org/10.1080/17512786.2016.1270171>
- Young, M. L., & Hermida, A. (2015). From Mr. and Mrs. Outlier to Central Tendencies: Computational journalism and crime reporting at the Los Angeles Times. *Digital Journalism*, 3(3), 381-397. <https://doi.org/10.1080/21670811.2014.976409>
- Wu, S. (2013). A review on coarse warranty data and analysis. *Reliability Engineering and System*, 114, 1-11. <https://doi.org/10.1016/j.res.2012.12.021>