

# Periodismo de datos: transferencia de resultados para la construcción de historias en el ecosistema de proximidad

María Valcárcel-Tabernero  
Universidade de Santiago de Compostela

José Sixto-García  
Universidade de Santiago de Compostela

## *Data journalism: transfer of results for the construction of stories in the local ecosystem*

### RESUMEN ABSTRACT

Esta investigación propone el desarrollo de un modelo de construcción de historias periodísticas en el marco del periodismo de proximidad a través de la selección, recolección y análisis de datos de la ciudad de Lugo, orientado a reposicionar el rol profesional del periodista como integrante de un equipo de carácter transdisciplinar mediante el uso de herramientas *big data*. Se plantean nuevas formas de producir, preservar y difundir información a través de la creación de un modelo innovador que utiliza técnicas y herramientas pertenecientes al ecosistema *big data* y se implementa el modelo con un estudio de caso. Las conclusiones se basan en la pertinencia, innovación y eficacia del modelo propuesto con el fin de recuperar la confianza en la profesión en el contexto de una ciudad abierta, inteligente e interactiva y en el entorno de un cosmopolitismo comunicativo donde el concepto de dignidad humana impregna todos los aspectos de la investigación de cara a contribuir a un lenguaje común y una mirada abierta como rasgos inherentes a la sociedad.

*This research proposes the development of a model for the construction of journalistic stories in the framework of local journalism through the selection, collection and analysis of data from the city of Lugo, aimed at repositioning the professional role of journalist as a member of a transdisciplinary team through the use of big data tools. New ways of producing, preserving and disseminating information are proposed through the creation of an innovative model that uses techniques and tools belonging to the big data ecosystem and the model is implemented with a case study. The conclusions are based on the relevance, innovation and effectiveness of the proposed model in order to regain confidence in the profession in the context of an open, intelligent and interactive city and in the environment of a communicative cosmopolitanism where the concept of human dignity permeates all aspects of research in order to contribute to a common language and an open look as inherent features of society.*

### PALABRAS CLAVE KEYWORDS

Periodismo de proximidad; Big data; Ciudad inteligente; Periodismo de datos; Redefinición periodística; Cosmopolitismo comunicativo.

*Local journalism; Big data; Smart city; Data journalism; Journalistic redefinition; communicative cosmopolitanism.*

## Periodisme de dades: transferència de resultats per a la construcció d'històries en un ecosistema de proximitat

### RESUM

Aquesta recerca proposa el desenvolupament d'un model de construcció d'històries periodístiques en el marc del periodisme de proximitat a través de la selecció, la recollida i l'anàlisi de dades de la ciutat de Lugo, amb la voluntat de reposicionar el rol professional del periodista com a integrant d'un equip de caràcter transdisciplinar, mitjançant l'ús d'eines de *big data*. Es plantegen noves maneres de produir, preservar i difondre informació mitjançant la creació d'un model innovador que utilitza tècniques i eines pertanyents a l'ecosistema *big data* i s'implementa el model amb un estudi de cas. Les conclusions es basen en la pertinència, innovació i eficàcia del model proposat per tal de recuperar la confiança en la professió en el context d'una ciutat oberta, intel·ligent i interactiva i en l'entorn d'un cosmopolitisme comunicatiu on el concepte de dignitat humana impregna tots els aspectes de la investigació de cara a contribuir a un llenguatge comú i una mirada oberta com a trets inherents de la societat.

### PARAULES CLAU

Periodisme de proximitat; Big data; Ciutat intel·ligent; Periodisme de dades; Redefinició periodística; Cosmopolitisme comunicatiu.

## 1. Introducció

El ecosistema híbrid de Lugo, enmarcado en el periodismo de proximidad e inmerso en el paradigma comunicativo de la convergencia, está llevando a cabo una adecuación tardía al ciberespacio y sus dinámicas, sin acabar de definir su rol en el nuevo entorno. Este hecho ha dado lugar a una progresiva pérdida de identidad que ha generado un desajuste importante a la hora de conectar con una audiencia nativa digital y sus funcionamientos para producir, consumir y difundir contenidos. Esto supone una encrucijada comunicativa no resuelta todavía (López-García, 2012) en el territorio de los medios lucenses, que siguen perpetuando un modelo de coexistencia matricial y digital consistente en la replicación de estructuras.

En clave autonómica, se han registrado dos tendencias en los cibermedios gallegos según el estudio del Ecosistema de Cibermedios en España, coordinado por Toural-Bran y López-García (2019) y las investigaciones sobre la evolución de los cibermedios hiperlocales de López-García, Negreira-Rey (2016); López-García, Negreira-Rey y Rodríguez-Vázquez (2016); Negreira-Rey y López-García (2021) y Negreira-Rey (2017, 2021, 2022). Por una parte, se crean nativos digitales, de carácter renovador con aprovechamiento sobre todo de redes sociales y participación ciudadana, en la mayoría de los casos, "en clave galleguista" (Negreira-Rey, 2017) y, por otra parte, se replican los matriciales en la esfera digital evitando iniciativas innovadoras que puedan poner en riesgo su rentabilidad (López, Limia, Toural y Pereira, 2012). El mayor número de hiperlocales se concentran en las provincias de A Coruña y Pontevedra, mientras que Lugo y Ourense se mantienen en una posición significativamente retrasada, con un 18,6% y un 7,4%, respectivamente (Negreira-Rey, 2021).

La hibridación comunicativa en Lugo existe en la medida en que existen medios tradicionales y medios digitales, matriciales y nativos, compartiendo un mismo territorio y, por tanto, combinando modos de producción y difusión de noticias. Sin embargo, estos procesos se llevan a cabo de una manera repetitiva, con introducción lenta de transformaciones que impliquen el aprovechamiento de las herramientas digitales en un contexto de hipertextualidad, interactividad, multimedialidad y nuevas narrativas. En este sentido, se pone de manifiesto una ausencia de modelos a seguir en el marco de proximidad lucense que consigan abrir vías de investigación y conceptualizaciones innovadoras.

Esta investigación pretende establecer un nexo relacional entre lo hiperlocal (Nygren et al., 2018; Negreira-Rey et al., 2021) y las nuevas herramientas computacionales y narrativas en el ámbito lucense, que derivará en un nuevo modelo de proximidad, en fase de prototipado, enfocado en el periodismo de datos. En esta línea, las iniciativas llevadas a cabo en el ámbito europeo indican que existe un campo de investigación

que marca un proceso abierto y aún en desarrollo (Arias-Robles y López-López, 2020; Stalph, Hahn y Liewehr, 2023).

### 1.1. Objetivos de la investigación

El objetivo general es el siguiente: (OG) Crear un modelo de construcción de historias periodísticas a través de la selección, recolección y análisis de datos masivos de la ciudad de Lugo, utilizando herramientas de big data que se podrán preservar, expandir y difundir por medio de una plataforma digital abierta, interactiva e innovadora. Los objetivos específicos son:

- **OE1.** Crear un archivo de datos de la ciudad Lugo (dataset de acceso abierto).
- **OE2.** Detectar problemas sociales y culturales de la localidad a partir del dataset.
- **OE3.** Potenciar una relación constante con los públicos cocreadores de contenidos.
- **OE4.** Establecer canales de participación y cocreación.
- **OE5.** Desarrollar una red creativa de historias periodísticas en constante retroalimentación.
- **OE6.** Construir un repositorio de historias periodísticas alojadas en la web.
- **OE7.** Reposicionar el rol del periodista como integrante de un equipo transdisciplinar en el que convergen disciplinas, técnicas y herramientas propias de las humanidades digitales.
- **OE8.** Contribuir al desarrollo individual y colectivo, tanto de la comunidad académica, como de la población en general.

## 2. Estado de la cuestión

### 2.1. Big data y periodismo de datos

La irrupción de lo digital en los procesos comunicativos ha abierto el panorama investigador al incorporar el pensamiento computacional a la información, así como el uso de herramientas propias que permiten el acceso a una gran cantidad de datos, su análisis, su cuantificación y visualización, y el establecimiento de relaciones, patrones e interpretaciones (Revuelta-Bayod, 2018). Asimismo, la interoperabilidad de los procesos permite su preservación y compatibilidad con otras aplicaciones y navegadores, lo que facilita, en buena medida, su proyección.

Conforme se ha ido desarrollando la tecnología, también lo ha hecho la investigación en torno al big data en cuanto a la

profundización en el estudio de la inteligencia artificial (Meso Ayerdi et al., 2023), el aprendizaje automático, las redes neuronales artificiales o el aprendizaje profundo (Gualda, 2022). De igual modo, la democratización de la ciencia y la posibilidad de un conocimiento colectivo (Lévy, 2004) han permitido enfrentarse a una serie de desafíos que suponen facilitar el manejo de herramientas inmersas en el ecosistema digital y permitir procesar enormes cantidades de datos (Trillo-Domínguez y De-Moya-Anegón, F., 2022). La innovación (García-Avilés, 2021) es fundamental a la hora de considerar un nuevo modelo comunicativo que lleva aparejados cambios culturales y sociales.

El periodismo de datos, con unas características bien definidas y en constante evolución, como son la hipertextualidad, interactividad, multimedialidad y el uso de herramientas digitales (Vizoso, 2022) está capacitado para poner en práctica el aprendizaje compartido, la interoperabilidad de datos, técnicas y herramientas, con unas audiencias participativas y dispuestas a integrar procesos de cocreación informativa (Sixto-García et al., 2020) en el marco del periodismo de proximidad, funcionando en red (López-García et al., 2016) y con carácter transdisciplinar.

Existe una tendencia creciente a relacionar el periodismo de datos con la proximidad (Ferrerías-Rodríguez, 2016; Flores-Vivar, 2018; Arias-Robles y Carvajal, 2022), en contraposición con una línea mayoritaria agrupada en las iniciativas de grandes medios (Appelgren y Nygren, 2019; Linden, 2019). Esta investigación se inscribe en ella.

### 2.2. La concepción de lo hiperlocal: ventajas y riesgos

Existe una preocupación creciente para situar a la audiencia en una posición central (García-Orosa et al., 2020) y fomentar la búsqueda de temas que atraigan su interés (Negreira-Rey, 2022). Resulta también interesante señalar la asociación de lo hiperlocal en ciberperiodismo con el periodismo constructivo que busca generar un impacto social positivo y que puede contribuir a disminuir la desafección ciudadana y, en consecuencia, el decrecimiento de la exposición a los procesos de desinformación (Serrano-Puche, 2020).

El alto grado de proximidad en cuanto a producción de contenidos genera, en principio, una mayor identificación con el público. Sin embargo, esta afirmación puede llevarnos a percepciones erróneas que hay que tener en cuenta: - ¿Esa cercanía representa la realidad del público que habita ese espacio concreto o sirve de potenciador de comunidades ideologizadas? - ¿El exceso de lo hiperlocal no puede iniciar una nueva tendencia a la hiperespecialización del conocimiento?

### 2.3. Cómo integrar el periodismo de big data en la ciudad inteligente

Se han retomado para esta investigación las consideraciones conceptuales de Richard Sennet (2019) referentes a la ciudad. El autor hace hincapié en una idea de ciudad con dos significados: - Ville: la ciudad en su conjunto, la ciudad construida. - Cité: la naturaleza de la vida de un barrio, los sentimientos de la gente en sus interrelaciones, su apego al lugar. Un tipo de conciencia.

Sennet propone una conceptualización de ciudades abiertas y cerradas afirmando que las ciudades abiertas se rigen por una ambigüedad general con espacios ordenados y, sin embargo, en las ciudades cerradas, basadas en la experiencia de la libre fricción, las diferencias terminan eliminándose o ignorándose si no están alineadas con nuestros deseos. La pregunta que va a enmarcar nuestro modelo sería la siguiente: ¿Existe la posibilidad de contribuir al conocimiento y desarrollo de una ciudad a partir de historias periodísticas generadas por medio de big data? El modelo propuesto en esta investigación apunta en esa dirección. Y debería aplicarse en una ciudad abierta y vivida. En una cité.

En el mismo sentido, este concepto de ciudad vivida no puede concebirse sin una reflexión sobre el cosmopolitismo en el mundo actual que obliga a pensar en los discursos, en su producción y distribución, y en el nuevo espacio comunicativo en el que se mueven. El espacio comunicativo se ha multiplicado de tal manera que no solo es necesario hablar de un mundo global sino de un ecosistema, con un cambio de paradigma, en el que se ponen en cuestión tanto la identidad como la verdad. Por eso es importante fijarse en las narrativas, en las maneras de construir los lenguajes y los códigos con lo que intentamos entender la realidad. Es ahí donde interviene la mirada. Si queremos crear una mirada cosmopolita, hemos de hablar un lenguaje común, inclusivo y justo. Esta investigación pone el foco en el ejercicio del periodismo desde esta perspectiva ética.

### 3. Metodología

La metodología se centra en la presentación de un modelo en fase de prototipado enfocado a la apertura de vías de desarrollo del ciberperiodismo lucense y sustentado en la construcción de historias periodísticas a través de la selección, recolección y análisis de datos masivos de la ciudad de Lugo, utilizando herramientas del ecosistema big data. Esta investigación se concibe como un proceso escalable cuya primera etapa, comprendida en este estudio, consiste en la ideación del modelo a implementar, su aplicación a un caso de estudio de ámbito local para ejemplificar su funcionamiento, y desarrollo de una plataforma digital que establezca las bases de la

participación ciudadana y los canales de cocreación y sirva de repositorio de historias periodísticas de Lugo.

La integración de la participación ciudadana se materializará en la plataforma digital, con canales de acceso abiertos para la subida de datos a través de un buzón de datos y la creación de una historia en función de los datos ya disponibles a través de un registro de usuarios/as y una estructura de selección que funciona como un e-commerce: elección de datos en lugar de productos que se recogen en una carpeta llamada expediente (carrito de la compra) y permanece abierta hasta su validación y conversión en historia abierta.

La validación de las historias estará a cargo del equipo liderado por periodistas que realizan una labor estructurada en varias fases: fact checking de datos subidos por usuarios/as, comprobación de fuentes ciudadanas, comprobación de visualizaciones en función de dos criterios: rigor informativo y extensión contextual en el desarrollo de historias periodísticas.

La tipología de datos utilizada comprende los dominios de las ciencias sociales y las humanidades (Schreibman, Siemens y Unsworth, 2015) al considerar la investigación como un proceso de coexistencia transdisciplinar:

- *Spatial Humanities* (Humanidades Espaciales): mapeo, la visualización geotemporal y la narración locativa.
- *Graphical Approaches to the Digital Humanities* (Aproximaciones gráficas para las Humanidades Digitales): herramientas de visualización.
- *Text Analysis and Visualization* (Análisis y visualización de texto): textos destinados a responder a ciertas preguntas de investigación.

Se establecieron los siguientes criterios para la conformación de un dataset de la ciudad de Lugo:

- Posibilidad de trazar contextos y construir historias a partir de información relevante sobre la ciudad.
- Rigor periodístico en el tratamiento e interpretación de datos y en la construcción de historias.
- Oportunidad para pensar la ciudad desde la cocreación y aporte colectivo de datos.
- Archivo open access, abierto a la investigación, difusión y ampliación de conocimiento sobre la ciudad de Lugo.
- Oportunidad para desarrollar un periodismo de big data colaborativo en la ciudad de Lugo.

Se procedió, en esta primera etapa, a la ideación de un modelo comunicativo en el marco del periodismo de proximidad a través de dos procesos: el proceso de descubrimiento del conocimiento (KDD): *Knowledge Discovery in Databases*

(Maimon y Rokach, 2006) y el proceso de construcción de historias (Modelo 5W), y en una segunda fase se realizó la implementación de ambos a un estudio de caso: Suicidios en Lugo.

Los criterios de selección del estudio de caso responden al nivel de importancia en el ámbito del periodismo local y a la constatación de una problemática de primer orden en el territorio autonómico, siendo Lugo, en función de sus cifras, la primera provincia en sufrir la situación.

### 3.1. Proceso de descubrimiento de conocimiento

1. Comprensión del dominio en el que se aplica el proceso, entendido como un espacio híbrido y transdisciplinar.

2. Selección y adición de los datos atendiendo al tipo de fuentes, a su carácter, a su localización y a su formato. Los criterios principales aplicados a la recogida de datos fueron dos, por un lado, la validez de estos para la implementación de las herramientas del ecosistema big data y, por otro, su significación en cuanto a la aproximación al tema y ampliación contextual del tema. Estos últimos datos sirvieron para generar el dataset del estudio de caso.

- Fuentes documentales: Prensa local: noticias sobre suicidios en Lugo (*El Progreso*) - Redes Sociales (Twitter) - Libros (Manual sobre el suicidio editado por Serfuja S.L. y "Las penas del joven Werther" de Goethe).

- Fuentes estadísticas: Datos en formato .csv o .xlsx provenientes de distintas instituciones que contienen información sobre Lugo (Instituto Nacional de Estadística (INE), Instituto Galego de Estadística (IGE), datos estadísticos del Ayuntamiento de Lugo, datos estadísticos del Servicio Galego de Saúde e (SERGAS) e Instituto de Medicina Legal de Galicia (IMELGA), referidos a la población por tramos de edad y sexo, y al número de suicidios por años, tanto en la provincia de Lugo como en la comunidad autónoma gallega, así como al mecanismo de perpetración del suicidio.

- Fuentes testimoniales: Se realizaron dos entrevistas en profundidad, semiestructuradas, a un familiar de una persona suicida que en el estudio hemos llamado María para preservar su anonimato por voluntad expresa y a un médico-forense del IMELGA (Instituto de Medicina Legal de Galicia), José Ramón Montenegro.

- Fuentes de big data procedentes de repositorios en abierto: Kaggle: <https://www.kaggle.com/> y World Bank Open Data: <https://data.worldbank.org/>, de donde se

extrajeron datos sobre suicidios por edad y sexo y por países en el ámbito europeo.

Se desarrollaron algoritmos para la aplicación de las herramientas de software en el modelo, que llevaron a cabo operaciones de análisis cuantitativo y ofrecieron gráficos de dispersión y nubes de palabras. Esto permitió inferir la importancia de las correlaciones entre términos en torno al suicidio para, de este modo, establecer asociaciones clave, detectar tendencias y analizar patrones. Los análisis de data mining se pueden extrapolar a cualquier investigación periodística con este modelo para determinar la recurrencia de términos en un texto relacionado con la historia que se esté construyendo.

En esta investigación se ha optado por recurrir a las aplicaciones de código abierto en la mayoría de los casos abogando por un uso de acceso libre que permita a la ciudadanía realizar las operaciones de data mining sin carga económica. El uso de Excel, en este caso, se debió a la disponibilidad previa del paquete Microsoft por parte de los investigadores, teniendo en cuenta que existe software de manejo de hojas de cálculo en versión libre al que los/as ciudadanos/as cocreadores/as pueden recurrir.

3. Preprocesamiento de los datos. Se procedió a la eliminación de ruido mediante la herramienta Excel para la limpieza de datos y la reducción de categorías, en casos de un número manejable de datos, y el entorno RStudio, a través del lenguaje de programación R, para el tratamiento de mayor cantidad de datos semiestructurados y no estructurados.

Se utilizaron dos textos en formato plano.txt, uno de ellos extraído de la "Guía de apoyo para la prevención del suicidio" promovida por Serfuja S.L., en Lugo, y el otro, la novela "The Sorrows of Young Werther" de Johann Wolfgang von Goethe ("Las penas de joven Werther") recuperado del repositorio digital ProjectGutenberg1 (<https://www.gutenberg.org/ebooks/2527>).

4. Transformación de los datos. Se llevó a cabo una transformación de datos para la realización del posterior análisis cuantitativo y las correspondientes visualizaciones de ambos textos mencionados anteriormente.

5. Minería de datos (Data Mining). Se seleccionaron, dentro del ecosistema de big data, las herramientas necesarias capaces de dar respuesta a nuestras preguntas de investigación y de cumplir los objetivos definidos:

- Excel: software incluido en el paquete Microsoft Office que permite la realización de tareas de creación y manejo de hojas de cálculo con datos numéricos y alfanuméricos.

- R y RStudio: R es un entorno y lenguaje de programación con un enfoque al análisis estadístico. R nació como una reimplementación de software libre del lenguaje S, adicionado con soporte para ámbito estático. Se trata de

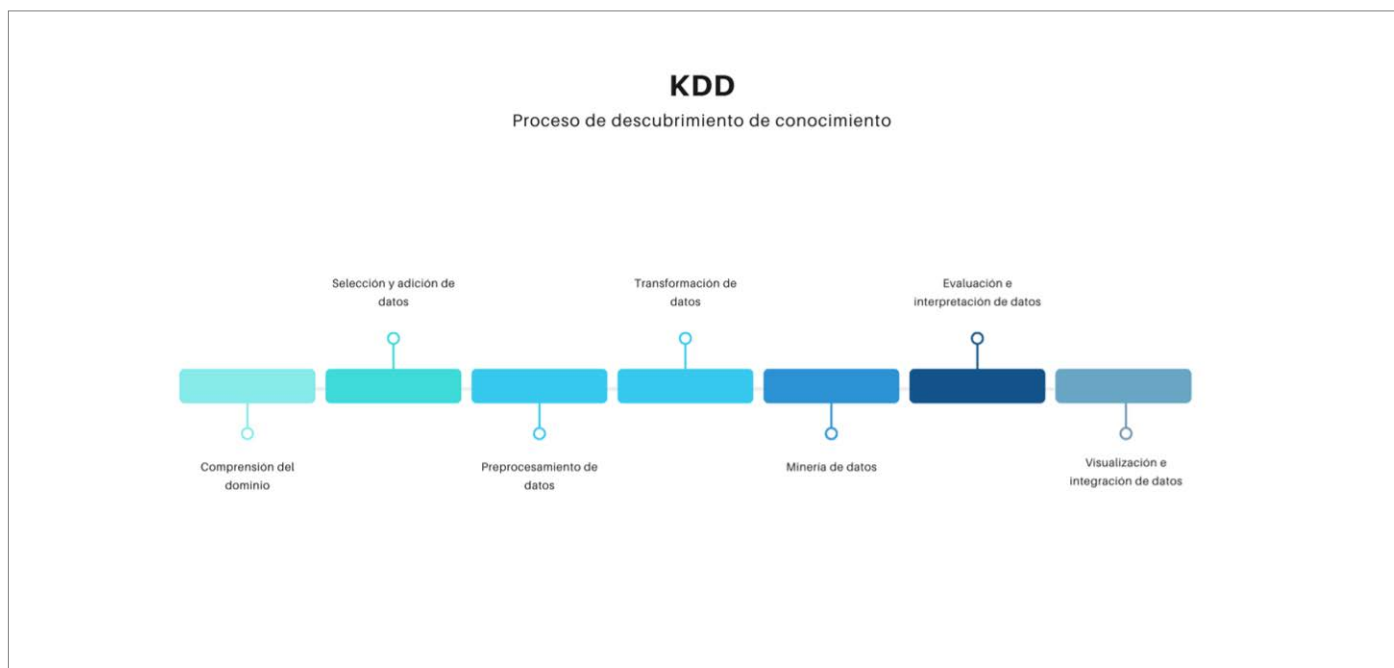


Figura 1. Modelo de descubrimiento del conocimiento basado en Maimon y Rokach (2006). Elaboración propia.

uno de los lenguajes de programación más utilizados en investigación científica, siendo además muy popular en los campos de aprendizaje automático (Machine Learning), minería de datos, investigación biomédica, bioinformática y matemáticas financieras.

- Voyant Tools: Es una aplicación web de código abierto para realizar análisis de corpus de texto. Admite la lectura académica y la interpretación de textos o corpus. Los textos que componen el corpus se pueden crear a partir de documentos online o subidos por los usuarios/as.
- Gephi: aplicación de código abierto que permite su implementación en distintas plataformas.
- Google My Maps: software creado por Google que permite la creación de mapas personalizados con un trabajo en capas y etiquetas descriptivas, facilita una interactividad constante y ofrece un uso intuitivo de cara a las/os usua-

rias/os.

- SocioViz: es una plataforma de análisis de medios sociales impulsado por métricas de análisis de redes sociales que además permite generar archivos en diferentes formatos. Es compatible con la herramienta Gephi, con la que se construyeron visualizaciones de grafos de relaciones.
6. Evaluación e interpretación de los datos. Se redactó la historia periodística de la investigación a través de la técnica del storytelling sobre los suicidios en Lugo.
  7. Visualización e integración. Se generaron las visualizaciones a partir de los datos tratados anteriormente que permitieron contextualizar, definir y explicar el tema de los suicidios en Lugo, con gráficos interactivos, y se creó una página web implementando el modelo y abriéndolo a la cocreación.

Who	What	When	Where	Why
Se eligen los/as protagonistas o el/la protagonista de la historia que puede tener un carácter individual o colectivo.	La tipología de la historia que se quiera construir determinará la ruta. En esta primera etapa se fijan dos: Cultura/Sociedad	<ul style="list-style-type: none"> <li>· Dimensión histórica: línea de tiempo que abarcaría el siglo XX hasta la segunda década del siglo XXI.</li> <li>· Dimensión tecnológica: línea de tiempo que abarcaría la década de los 90 del siglo XX hasta la segunda década del siglo XXI.</li> </ul>	Cartografía de Lugo dividida en cuatro zonas principales. Creación de un entorno geolocalizado correspondiente a cuatro zonas diferenciadas de la ciudad: Centro/Barrios/ Periferia/Rural	Visualización y storytelling

Tabla 1. Criterios de las 5W para establecer la ruta de la historia. Fuente: elaboración propia.

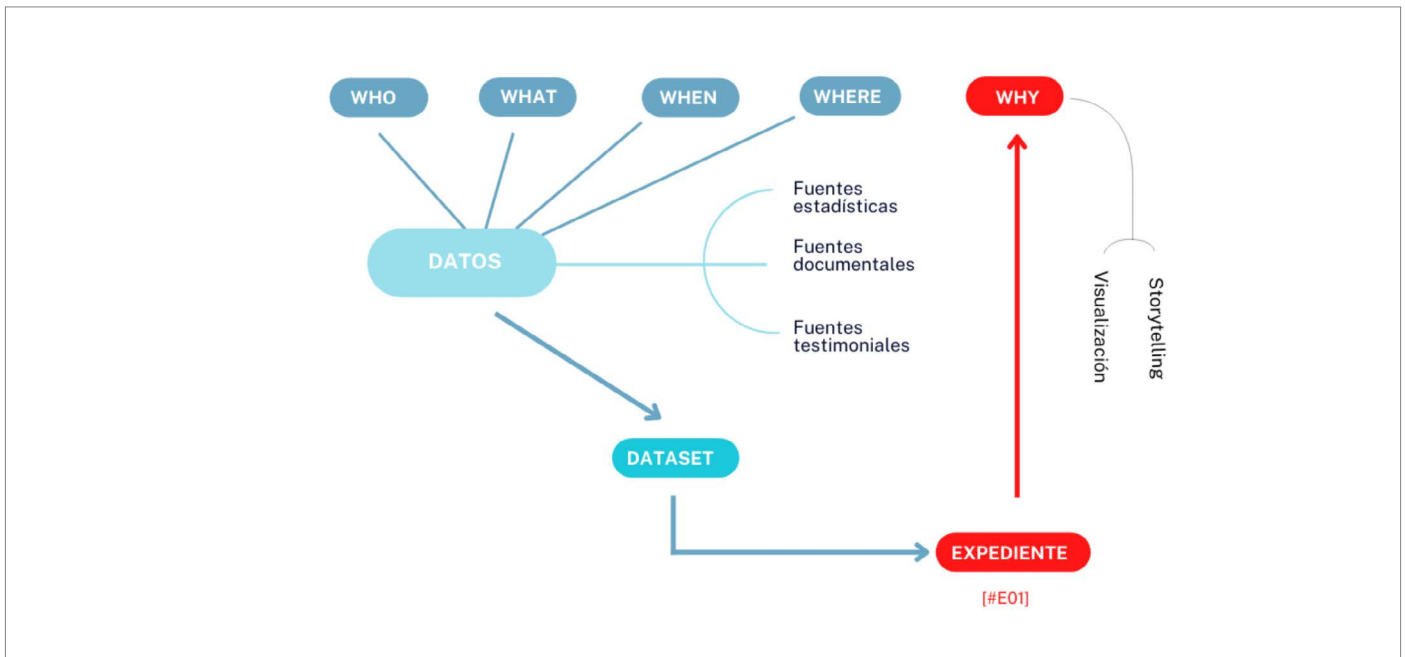


Figura 2. Modelo de construcción de historias periodísticas. Elaboración propia.

### 3.2. Proceso de construcción de historias

La construcción de historias va a venir determinada por la elección de cada usuario/o que interactúa con los datasets de las 4W. En función de las distintas selecciones que vaya realizando, irá estableciendo una ruta determinada, personalizada, pero, a la vez, de acceso abierto. Esa ruta propia va a constituir un archivo de información de ciudad de Lugo, que se guardará con el nombre de "expediente + n". A partir de ese expediente se elaborará la historia a través de visualizaciones y storytelling, que se corresponderá con la W5(Why).

El punto de partida para fijar la ruta será el modelo periodístico de las 5W: Who, What, When, Where, Why (Quién Qué, Cuándo, Dónde, Por qué). De una manera intuitiva, la navegación por la web irá conformando el recorrido (ver figura 2).

Tenemos 4 datasets cuyos datos son susceptibles de ser analizados por separado (cada W con su propio dataset) o interrelacionándolos (los datasets de unas W con otras), estableciendo rutas específicas y personalizadas. Cada archivo de datos se conforma en función de unos criterios establecidos previamente (ver tabla 1).

### 3.3. Aplicación del modelo a un caso de estudio: suicidios en Lugo

El modelo se ha planteado como un proceso con un desarrollo, por lo que se nutre, se enriquece y se expande a medida que se

incorporan datos y se desarrollan historias. Es una propuesta metodológica con un recorrido amplio.

Tras el dataset del estudio de caso: suicidios en Lugo, construido en la fase anterior, se procedió a la construcción a través de la combinación de los datos de las 4W. Los ejes y sus nexos quedaron contextualizados, tanto temporal como espacialmente, pudiendo establecerse un nuevo modelado de datos a partir de las relaciones entre las zonas, los timelines y su casuística.

La visualización de datos (W5) proporcionó un soporte enriquecido que es la clave de la creación de historias. Lev Manovich se refiere a las posibilidades abiertas por las herramientas de representación en el sentido de que "nos permite ver los patrones y estructuras que subyacen en conjuntos de datos aparentemente arbitrarios" (Manovich, 2008, p. 130).

### 3.4. Construcción de la historia

Los datos seleccionados son susceptibles de modelarse en función de los objetivos de cada tema elegido e historia que se va a construir.

Es importante destacar que el número de ejes es flexible. Se seleccionaron, para el caso de estudio, nueve nodos, por el interés de sus posibles combinaciones y la ampliación contextual. El modelado de datos está orientado hacia una extensión y desarrollo que permita ir conformando un mapa profundo y detallado de la ciudad con cada nueva historia.

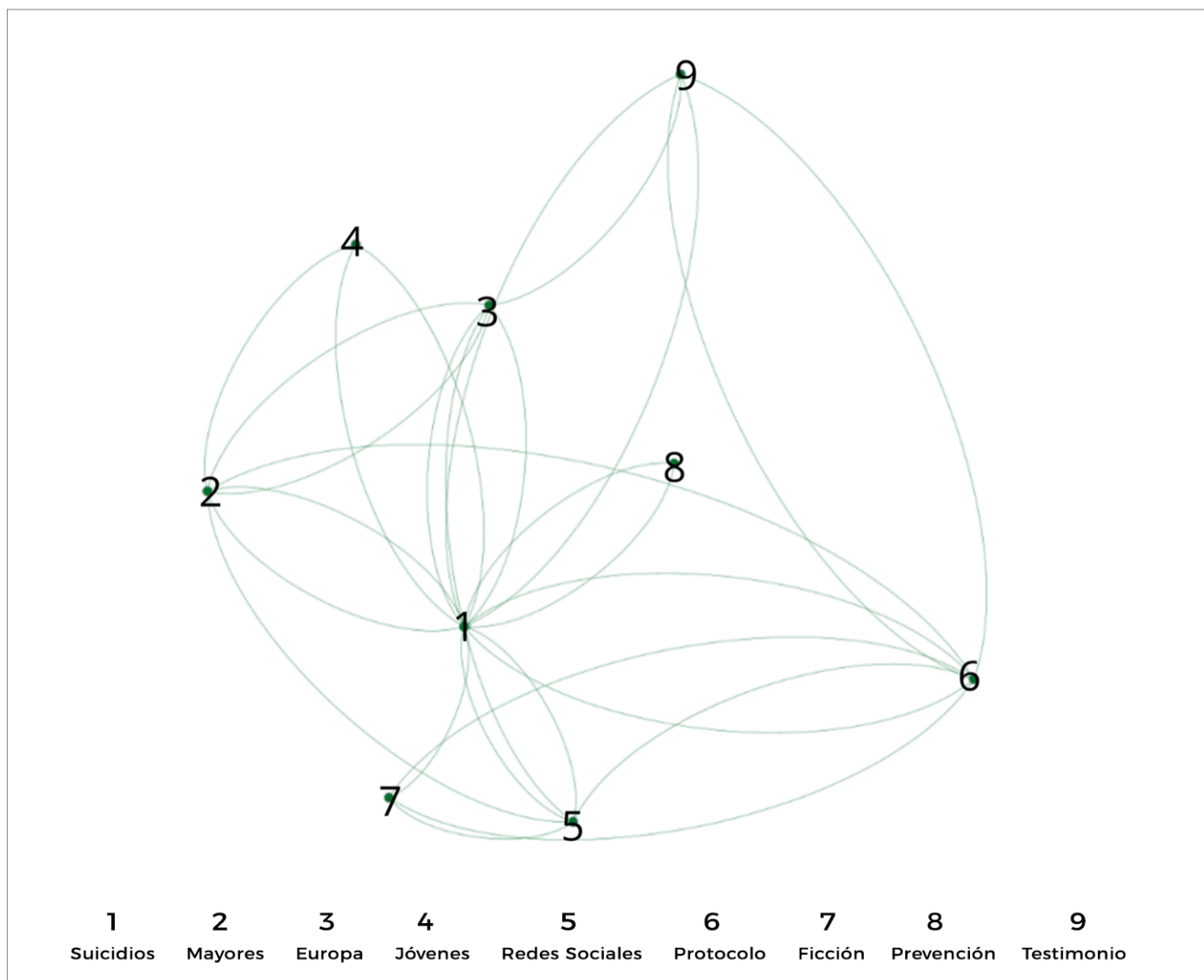


Figura 3. Nodo de relaciones. Ejes, grafos, nodos y aristas. Suicidios en Lugo. Elaboración propia.

Suicidios	Mayores	Europa	Jóvenes	Redes Sociales	Protocolo	Ficción	Prevención	Testimonio
Exposición visual de la situación: Lugo en el contexto gallego.  Preocupación experto ante los datos (audio)	Suicidios por edad y sexo.	Panorama europeo en cuanto a tasa de suicidios por edad y sexo.	Suicidios por edad y sexo.	Twitter(X) #suicidios	Geolocalización zonas	Caso Werther  Realidad	Protocolo explicado por experto (audio)	Narración María (nombre ficticio): experiencia ante el suicidio de su marido en Lugo

Tabla 2. Estructuración del storytelling en torno a los ejes. Elaboración propia.

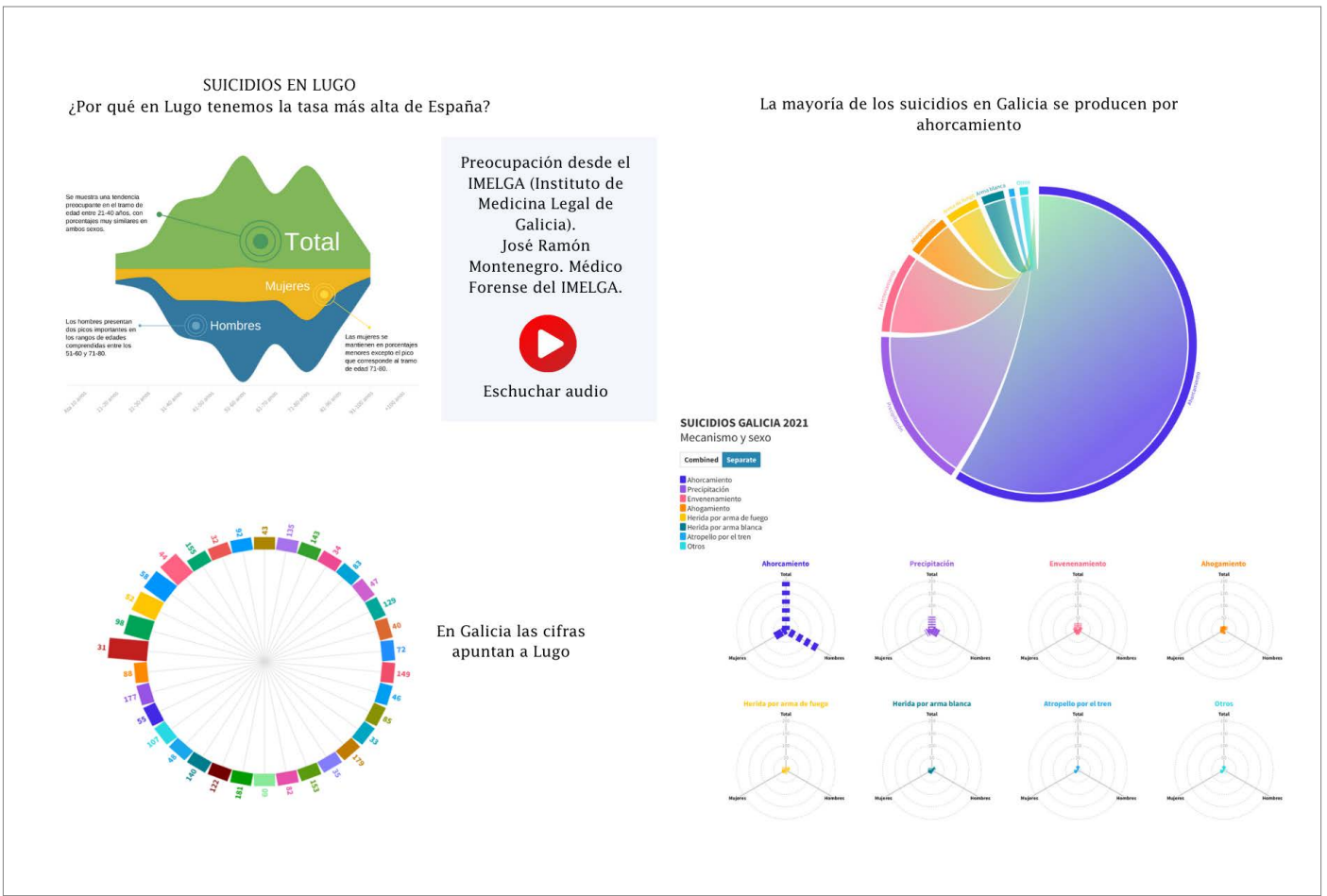


Figura 4. (1). Storytelling de big data sobre suicidios en Lugo. Elaboración propia.



Figura 5. (2). Storytelling de big data sobre suicidios en Lugo. Elaboración propia.



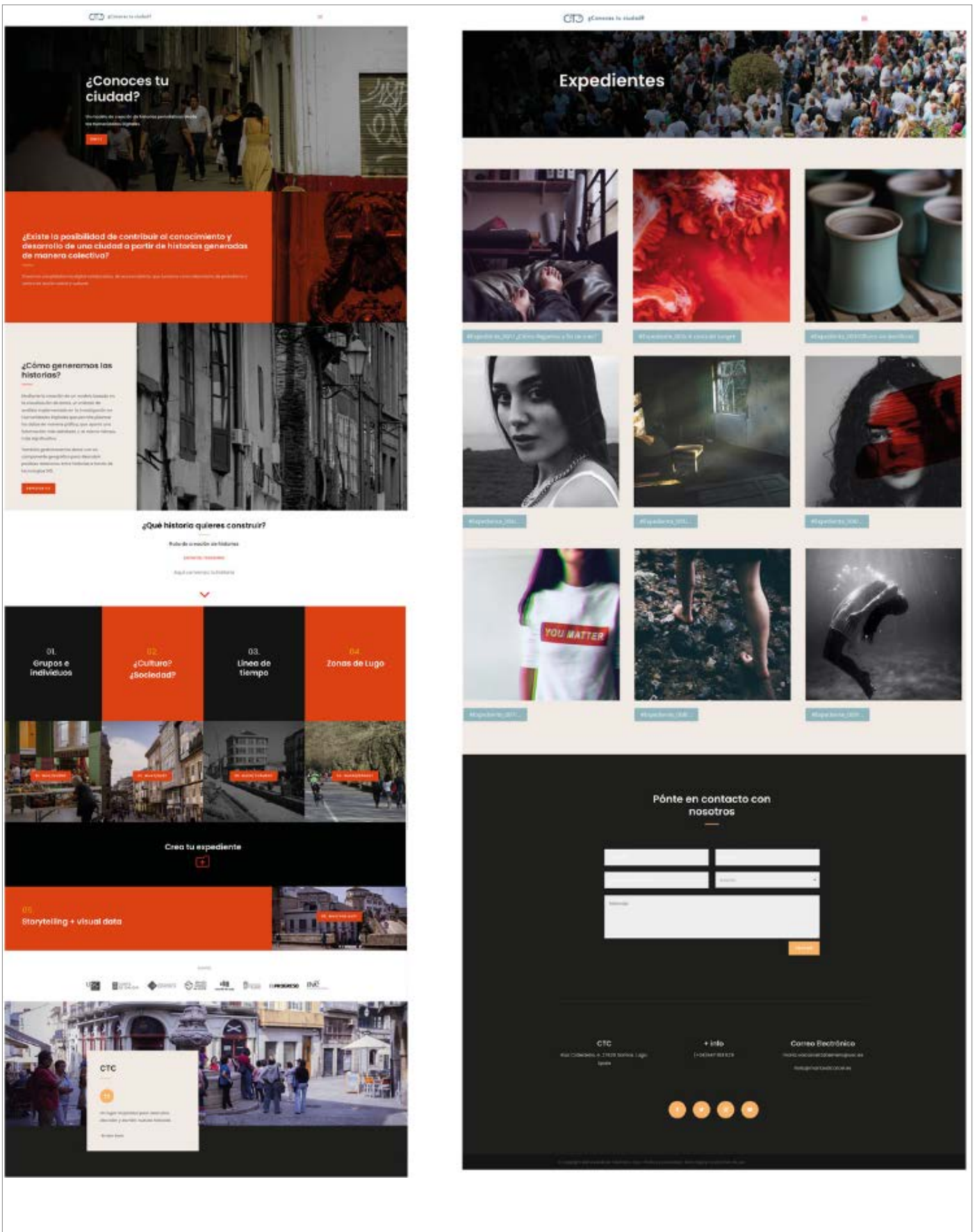


Figura 8. Home de la página web de la investigación. Elaboración propia. (<https://mannersmedia.es>)

### 3.5. Elaboración de los expedientes

Las historias creadas con el modelo propuesto y la técnica del storytelling serán publicadas, difundidas y compartidas a través de una plataforma digital –página web de realización propia– desarrollada con la herramienta Wordpress, un sistema de gestión de contenidos o CMS (Content Management System), de software libre, que posee una interfaz sencilla e intuitiva en su manejo.

Dado que esta investigación está enfocada a la participación y cocreación, la facilidad de uso va a permitir a las/os usuarias/os realizar consultas, aportar datos e información y crear nuevas historias. Cada historia creada se convertirá en un expediente que será clasificado con la etiqueta: #expediente+n. La historia creada en la investigación lleva la etiqueta [#E01].

Con ese expediente –que funciona como una base de datos personalizada– se construye la historia. A medida que se vayan creando más expedientes, la web dispondrá de un repositorio de acceso abierto en el que las historias podrán ser, no sólo visualizadas, sino también, expandidas por otras/os usuarias/os.

De esta forma, se contribuye a crear la Historia (un recopilatorio de historias colectivas) de la ciudad de Lugo, desde el territorio del periodismo de proximidad, con la contribución de sus habitantes a través del desarrollo de sus propios datasets y, por tanto, de sus propias visiones de los diferentes temas.

El expediente de la historia periodística: Suicidios en Lugo es el resultado de la aplicación del modelo propuesto.

## 4. Discusión y conclusiones

Esta investigación posee un carácter limitado ya que aborda la primera etapa de un proyecto de largo recorrido, autónomo y con vocación práctica, por lo tanto, se contemplan etapas posteriores que trascienden los límites del estudio actual, pero que se han tenido en cuenta a la hora de establecer las preguntas y de determinar objetivos. El equipo responsable de la implementación del modelo ha de ser transdisciplinar debido a la convergencia de herramientas y campos de conocimiento humanístico, periodístico y tecnológico, y las historias creadas a partir del modelo necesitan la validación de periodistas para ser publicadas y accesibles. De este modo, la profesión periodística redefine su rol, erigiéndose en líder de equipos dinámicos capaces de interactuar con profesionales de otros ámbitos y con la ciudadanía. En este sentido, confirmamos la relevancia de una investigación de discursos

posteriores a la publicación de las historias creadas (Witschge, 2008), así como a su preservación (Heravi et al., 2021).

El modelo está ideado para una implementación independiente susceptible de funcionar como un medio innovador en el contexto del periodismo de proximidad, aunque también se puede pensar como parte de una redacción ya constituida, a la manera de un lab del propio medio. Los resultados de esta investigación apuntan a que el periodismo de big data, en interrelación constante con las humanidades digitales, puede ofrecer el marco teórico y metodológico para su desarrollo en el ecosistema lucense, y la apuesta por un proyecto innovador, atractivo, eficaz e interactivo, sin perder la profesionalidad que se le presupone – y así debe seguir siendo – a todo/a periodista.

Tomando como espacio de investigación la ciudad de Lugo, este estudio ha planteado, no sólo que es posible sino que es rupturista la creación de un modelo de construcción de historias a través de la selección, recolección y análisis de datos de la ciudad, utilizando herramientas digitales inmersas en el ecosistema de big data, que se podrán preservar, expandir y difundir por medio de una plataforma digital abierta, interactiva e innovadora y que cumple con los rasgos definitorios de los nuevos medios propios del ciberperiodismo determinados por Díaz Noci y Salaverría (2003).

Cuestión de enorme relevancia es el desarrollo del modelo en el ecosistema del big data, fundamentado en la necesidad de búsqueda de información en términos apropiados a la sociedad actual. Los métodos y herramientas seleccionados evidenciaron su idoneidad para la construcción de historias en el territorio de proximidad. En este sentido, el marco analítico planteado por el prototipo puede nutrirse y afianzarse con investigaciones enfocadas a la visualización de datos (Stalph y Heravi, 2021) y su sistematización.

La propuesta de cocreación de las historias, estableciendo canales de participación en una plataforma de acceso abierto, incide en la importancia de contribuir al conocimiento de la ciudad de un modo colectivo, potenciando y dotando de sentido a los conceptos de cultura e inteligencia colectivas.

El acceso al conocimiento de una ciudad, a partir de este modelo, permitirá la evolución de dicha ciudad, de manera que aumentará su dinamismo y la circulación libre de información. Y, además, contribuirá a que su ciudadanía sea consciente de dos cuestiones fundamentales: por un lado, el verdadero valor histórico y social del lugar en el que residen; por otro, su nuevo rol activo como integrantes de una comunidad más democrática además de una nueva forma de consumo (prosumición) propia de la cultura digital del siglo XXI.

Con respecto al cumplimiento de objetivos, la propuesta y el desarrollo del modelo mediante su aplicación al estudio de caso y su traslado a una plataforma web determinan

el cumplimiento del objetivo general (OG) planteado en la investigación. En cuanto a los objetivos específicos: OE1. El desarrollo del caso de estudio (suicidios en Lugo) con el modelo propuesto nos ha permitido iniciar el archivo de datos (dataset) de acceso abierto de la ciudad de Lugo. Como ya se ha explicado anteriormente, esta investigación se ha desarrollado con la vocación de proceso y no con de estudio cerrado. Es de desear, por tanto, que el modelo planteado tenga un recorrido y una evolución futuras, con lo que el dataset se iría ampliando a medida que se construyan historias periódicas sobre la ciudad de Lugo. OE2. La aplicación práctica del modelo propuesto en el caso de los suicidios ha permitido detectar un problema social en la ciudad. OE3. OE4. OE5. OE6. Se ha implementado en la web un canal de participación para la construcción de historias a través de un formulario por pasos que responde al modelo planteado en la investigación, lo que supone la potenciación de una relación con el público cocreador a través de la apertura de un canal de participación y cocreación, y el desarrollo de una red de historias periódicas alojadas en un repositorio web. OE7. La puesta en práctica del modelo planteado en la investigación supone redefinir el rol del/a periodista en el marco de un equipo de trabajo transdisciplinar, en el que las humanidades digitales juegan un papel esencial y complementan el modelo mediante la utilización de técnicas y herramientas que le son propias. OE8. La propuesta del modelo de storytelling de big data aplicado al ecosistema comunicativo de Lugo abren nuevas vías de investigación y sugiere su implementación en otros ámbitos y una ampliación de estudios en esa línea. Del mismo modo, su implementación real contribuye al desarrollo de la población al abrir nuevas vías de aprendizaje y conocimiento.

Por tanto, la posibilidad de contribuir al conocimiento y desarrollo de la ciudad de Lugo quedaría demostrada con la implementación del modelo propuesto en la medida en que se crea una red de participación colectiva en torno a la ciudad y sus problemáticas. El modelo propuesto, su implementación en esta investigación, así como su desarrollo futuro, contribuye a aumentar el conocimiento de la ciudad y la información de su ciudadanía en torno a los asuntos de la ciudad, gracias a la apertura de canales de participación y a la posibilidad de implicación en la construcción de historias. Este supuesto estaría contemplado en la siguiente etapa de la investigación, aún no realizada.

La transdisciplinariedad y el movimiento open data recorren el planteamiento y desarrollo de este trabajo, como se ha demostrado con la utilización de técnicas y herramientas de storytelling y big data, así como con la apuesta decidida por la cocreación. A pesar de que la comprobación definitiva de la eficacia de estas medidas supera los límites de esta investigación, sí es posible avanzar que, con este modelo, se ha abierto un proceso hasta ahora inexistente en Lugo, con capacitación

suficiente para que la participación colectiva se materialice.

Se ha constatado que el periodismo de datos es, en sí mismo, una hibridación de otras fórmulas de ejercer la profesión y que aún no se ha implementado en las iniciativas locales actuales. El uso sistematizado de big data, en cuanto a recogida, selección y tratamiento de datos, así como al uso de herramientas y técnicas propias del pensamiento computacional, permitirán el abordaje de historias que afectan a la ciudadanía de una manera innovadora, participativa y contextualizada, logrando, a su vez, a través de la visualización de datos, una interpretación pormenorizada y un enriquecimiento de las propias historias.

Una ciudad abierta, siguiendo las pautas expuestas en esta investigación, pasa por aplicar modelos como el propuesto, que construyan las bases de desarrollo de la ciudad a partir de los datos y las historias compartidas y una aproximación innovadora a la hora de construirlas.

La conceptualización de ciudad como el espacio para el desarrollo del periodismo de proximidad se ha analizado en base a su significación en el marco del modelo propuesto. La implementación del modelo de storytelling de big data en Lugo requiere, a su vez, de una propuesta de prototipo de ciudad abierta, inteligente, innovadora e interactiva – lo que hemos llamado *ciudad abierta* a lo largo de la investigación –. Esta ciudad sería la apropiada para el planteamiento del cosmopolitismo comunicativo como marco ideológico para la creación del modelo que se propone. Como consecuencia, esta investigación abarca, no únicamente la formulación de modelo en sí, sino también el planteamiento de un entorno propicio para su puesta en práctica.

No se ha podido demostrar, en el marco de esta investigación un rechazo a la innovación porque no hay experiencias en ese sentido que se hayan podido analizar. Este hecho supone un riesgo y una oportunidad.

A la hora de analizar el estudio de caso implementando el modelo propuesto se ha constatado que existe la posibilidad de abrir procesos de democratización de la información a través del aprovechamiento del ecosistema de big data aplicado al periodismo de proximidad. Esto supone la asimilación de la convergencia digital como el camino para emprendimientos innovadores e híbridos. De las interrelaciones en el ecosistema surgen modos nuevos de mirar y, por tanto, nuevos modos de contar historias periódicas.

#### **4.1. Nuevas líneas de investigación**

Esta propuesta permite continuar el desarrollo del modelo planteado, y seguir avanzando en la línea de investigación, lo que supone abordar y seguir explorando técnicas de procesamiento del lenguaje natural (PLN) y técnicas de deep learning para la construcción de determinadas historias, con

la finalidad de que la implementación de dichos recursos proporcione nueva información, así como la generación de nuevo conocimiento.

Abre, además, vías de evolución de la ciudad de Lugo, a partir de las historias construidas y compartidas que, sin dejar de ejercer el rigor periodístico esencial para el ejercicio de la profesión, permite la interacción con la ciudadanía y, como consecuencia, una conexión mayor.

Este modelo contribuye al fomento del valor histórico y social de la ciudad, permitiendo la construcción de una ciudadanía activa y democrática, consciente de la importancia de su participación.

Aporta vías de solución para la recuperación de la confianza en la profesión periodística, planteando un nuevo rol activo – conformado por la alfabetización digital – liderando equipos transdisciplinares que entiendan las nuevas formas de hacer periodismo como procesos, flexibles e innovadores, en los que el periodismo va de la mano de las humanidades digitales, como una manera de aunar la acción comunicativa con la capacidad crítica propia de estas disciplinas.

Para finalizar, se han planteado unos indicadores que han permitido valorar en qué medida o medidas el modelo va a ser capaz de afrontar los retos del periodismo de proximidad en una interrelación transdisciplinar haciendo uso de las herramientas del ecosistema big data:

1. Analizar la pertinencia e integridad de los datos que se utilizarán en la construcción de las historias periodísticas.
2. Evaluar la funcionalidad de la plataforma digital (página web) en cada paso del proceso de descubrimiento de conocimiento a la hora de construir las historias.
3. Verificar y contrastar las fuentes para la construcción de historias para su posterior validación y difusión pública.
4. Detectar posibles dificultades en la implementación de la cocreación de historias.
5. Examinar el impacto de la página web y de las historias

que se visualizan, se expanden y se comparten a través de métricas adecuadas.

6. Evaluar la sostenibilidad del proyecto y la posibilidad de desarrollo de sus siguientes fases, que serían las siguientes:

- Innovación y experimentación de nuevos formatos, lenguajes y modos de construir historias periodísticas en el ecosistema big data.
- Implementación del modelo en otras ciudades que presenten características similares a las de Lugo en el marco del periodismo de proximidad.

## Referencias

- Appelgren, E. y Nygren, G. (2019). HiPPOs (Highest Paid Person's Opinion) in the Swedish media industry on innovation: A study of news media leaders' attitudes towards innovation. *The journal of media innovations*, 5(1), 45-60. <https://doi.org/10.5617/jomi.6503>
- Arias-Robles, F. y Carvajal, M. (2022). Periodistas de datos. Análisis del perfil sociodemográfico del profesional en España y Latinoamérica. *Estudios sobre el Mensaje Periodístico* 28(1), 621-637. <https://dx.doi.org/10.5209/esmp.76395>
- Arias-Robles, F. y López-López, P. (2021) Driving the Closest Information. Local Data Journalism in the UK. *Journalism Practice*, 15(5), 638-650. <https://doi.org/10.1080/17512786.2020.1749109>
- Deuze, M. (2021). On the 'grand narrative' of media and mass communication theory and research: a review. *Profesional de la información*, 30(1), 1-13. <https://doi.org/10.3145/epi.2021.ene.05>
- Díaz-Noci, J. y Salaverría, R. (coords.) (2003). *Manual de redacción ciberperiodística*. Ariel.
- Ferreras-Rodríguez, E.M. (2016). El periodismo de datos en España. *Estudios Sobre El Mensaje Periodístico*, 22(1), 255- 272. [https://doi.org/10.5209/rev\\_ESMP.2016.v22.n1.52594](https://doi.org/10.5209/rev_ESMP.2016.v22.n1.52594)
- Flores-Vivar, J. M. (2018). Algoritmos, aplicaciones y big data, nuevos paradigmas en el proceso de comunicación y de enseñanza-aprendizaje del periodismo de datos. *Revista de Comunicación*, 17(2), 268-291. <https://doi.org/10.26441/RC17.2-2018-A12>
- García-Avilés, J. A. (2021). Artículo de revisión: La investigación sobre

## PUBLICIDAD



<https://www.upf.edu/web/ocm>



Universitat  
Pompeu Fabra  
Barcelona

Departamento  
de Comunicación  
Grupo DigiDoc



El **Observatorio de Cibermedios** es una producción del *Grupo de Investigación en Documentación Digital y Comunicación Interactiva (DigiDoc)* del **Departamento de Comunicación** de la **Universitat Pompeu Fabra**.

El Observatorio de Cibermedios (OCM) forma parte del proyecto del Plan Nacional "*Parámetros y estrategias para incrementar la relevancia de los medios y la comunicación digital en la sociedad: curación, visualización y visibilidad (CUVICOM)*". PID2021-1235790B-I00 (MICINN), Ministerio de Ciencia e Innovación ( España).

innovación en periodismo, un campo diverso y pujante (2000-2020). *Profesional de la información*, 30(1), 1-33. <https://doi.org/10.3145/epi.2021.ene.10>

García-Orosa, B., López-García, X. y Vázquez-Herrero, J. (2020). Journalism in digital native media: Beyond technological determinism. *Media and Communication*, 8(2), 5-15. <https://doi.org/10.17645/mac.v8i2.2702>

Gualda, E. (2022). Social big data, sociología y ciencias sociales computacionales. *Empiria: Revista de metodología de ciencias sociales*, (53), 147-177. <https://doi.org/10.5944/empiria.53.2022.32631>

Heravi, B. R., Cassidy, K., Davis, E. y Harrower, N. (2021). Preserving Data Journalism: A Systematic Literature Review. *Journalism Practice*, 16(10), 2083-2105. <https://doi.org/10.1080/17512786.2021.1903972>

Lévy, P. (2004). *Inteligencia Colectiva: por una antropología del ciberespacio*. OPS.

Lindén, C.-G., Tuulonen, H., Bäck, A., Diakopoulos, N., Granroth-Wilding, M., Haapanen, L., Leppänen, L., Melin, M., Moring, T., Munezero, M., Sirén-Heikel, S., Södergård, C. y Toivonen, H. (2019). *News automation. The rewards, risks and realities of 'machine journalism'*. WAN-IFRA.

López-García, X. (2012). Cibermedios locales: en la encrucijada por la indefinición de modelos para tiempos convulsos. *Ámbitos. Revista Internacional de Comunicación* (21), 9-20. <https://www.redalyc.org/articulo.oa?id=16823120001>

López-García, X. (2022). Redefinición de metacompetencias, competencias y habilidades de los periodistas digitales. *Anuario ThinkEPI*, (16). <https://doi.org/10.3145/thinkepi.2022.e16a03>

López-García, X., Negreira-Rey, M. C. y Rodríguez-Vázquez, A. I. (2016). Cibermedios hiperlocales ibéricos: el nacimiento de una nueva red de proximidad. *Cuadernos. Info*, (39), 225-240. <https://doi.org/10.7764/cdi.39.966>

Maimon, O. y Rokach, L. (2005). Introduction to Knowledge Discovery in Databases. En Maimon, O. y Lior Rokach (ed.). *Data Mining and Knowledge Discovery Handbook*, (pp. 1-17). Springer.

Manovich, Lev (2002). What is New Media? Eight Propositions. En N. Wardrip-Fruin y N. Montfort (eds.) *New Media from Borges to HTML, The New Media Reader*. (pp. 1-9). The MIT Press.

Meso-Ayerdi, K., Larrondo-Ureta, A., y Peña-Fernández, S. (2023). Algoritmos, inteligencia artificial y periodismo automatizado en el sistema híbrido de medios. *Textual & Visual Media*, 17(1), 1-6. <https://doi.org/10.56418/txt.17.1.2023.0>

[doi.org/10.56418/txt.17.1.2023.0](https://doi.org/10.56418/txt.17.1.2023.0)

Negreira-Rey, M.-C. (2022). Capítulo 4. Distribución geográfica de los medios nativos digitales. *Espejo De Monografías De Comunicación Social*, (7), 71-81. <https://doi.org/10.52495/c4.emcs.7.p92>

Negreira-Rey, M.-C. y López-García, X. (2021). A Decade of Research on Hyperlocal Media: An International Approach to a New Media Model. *Online Journal of Communication and Media Technologies*, 11(3), e202111. <https://doi.org/10.30935/ojcm/11082>

Nygren, G., Leckner, S. y Tenor, C. (2018). Hyperlocals and Legacy Media: Media Ecologies in Transition. *Nordicom Review*, 30(1), 33-49. <https://doi.org/10.1515/nor-2017-0419>

Revuelta-Bayod, M. J. (2018). "Big Data": crisis y nuevos planteamientos en los flujos de comunicación de la cuarta revolución industrial. *Área Abierta*, 18(2), 309-324. <https://doi.org/10.5209/ARAB.59521>

Schreibman, S., Siemens, R. y Unsworth, J. (eds.). (2015). *A new companion to digital humanities*. John Wiley & Sons.

Sennet, R. (2019). *Construir y habitar*. Anagrama.

Serrano-Puche, J. (2020). Periodismo constructivo: una respuesta a las razones de los usuarios para evitar las noticias. *Cuadernos. Info*, (46), 153- 177. <https://doi.org/10.7764/cdi.46.1714>

Sixto-García, J., López-García, X. y Tournal-Bran, C. (2020). Oportunidades para la cocreación de contenidos en los diarios nativos digitales. *Profesional de la información*, 29(4), 1-19. <https://doi.org/10.3145/epi.2020.jul.26>

Stalphy, F. (2018). Classifying Data Journalism. *Journalism Practice*, 12 (10), 1332-1350. <https://doi.org/10.1080/17512786.2017.1386583>

Stalphy, F. y Heravi, B. (2021): Exploring Data Visualisations: An Analytical Framework Based on Dimensional Components of Data Artefacts in Journalism. *Digital Journalism*, 11(9), 1641-1663 <https://doi.org/10.1080/21670811.2021.1957965>

Stalphy, F., Hahn, O. y Liewehr, D. (2023) Local Data Journalism in Germany: Data-driven Reporting Amidst Local Communities and Authorities. *Journalism Practice*, 17(9), 1882-1901, <https://doi.org/10.1080/17512786.2021.2019089>

Trillo-Domínguez, M. y De-Moya-Aneón, F. (2022). Mapa de la investigación científica de Co- municación en España: frentes de estudio y rankings de autores, publicaciones e instituciones. *Profesional de la Información*, 31(1), e310112. <https://doi.org/10.3145/epi.2022.ene.12>

## PUBLICIDAD





**Máster Universitario Online en Marketing Digital: Buscadores (SEO y SEM) y Usabilidad (UX y UI)**

Solicita información | Próxima edición: noviembre 2024 - julio 2025

Vizoso, Á. (2022). Capítulo 16. Periodistas de datos y visualización de la información. *Espejo de Monografías de Comunicación Social*, (7), 247-256. <https://doi.org/10.52495/c16.emcs.7.p92>

Witschge, T. (2008). Examining online public discourse in context: A mixed method approach. *Javnost-The public*, 15(2), 75-91. <https://javnost-thepublic.org/article/2008/2/5/>

## CV

### María Valcárcel Tabernero

- maria.valcarcel.tabernero@usc.es
- <https://orcid.org/0000-0001-6927-8412>
- Doctora en Comunicación e Información Contemporánea y licenciada en Humanidades por la Universidade de Santiago de Compostela. Ha ampliado sus estudios con el Máster Internacional en Humanidades Digitales, el Máster en Gestión Cultural y el Máster de Escritura para Televisión y Cine. Entre sus líneas de investigación se encuentran el periodismo digital, new media studies y humanidades digitales. Cuenta con experiencia profesional en periodismo cultural, gestión cultural y periodismo corporativo. Actualmente pertenece al grupo de investigación Novos Medios de la Universidade de Santiago de Compostela y es Profesora

Asociada del Departamento de Comunicación de la Universidade de Santiago de Compostela.

### José Sixto García

- jose.sixto@usc.es
- <https://orcid.org/0000-0002-2988-0975>
- Profesor del Departamento de Ciencias de la Comunicación de la Universidade de Santiago de Compostela. Cuenta con más 100 publicaciones, la mayoría en revistas científicas. Entre las más recientes figuran libros como *Gestión profesional de redes sociales. Rutinas y estrategias de Community Management* (2018) o *Fundamentos de marketing digital* (2016), ambos publicados en Comunicación Social, y artículos como "Co-creation in North American and European digital native media: Web, social networks and offline spaces", publicado en *Journalism* (2022). Acredita participación en varios proyectos de investigación de ámbito autonómico, nacional y europeo. Recibió el Premio Drago Latina en 2016 al artículo más citado y ha sido invitado como conferenciante en varias universidades extranjeras. Con respecto a las actividades de transferencia, irigió el Instituto de Medios Sociales (2013-2019) y ha realizado diferentes acciones con el sector empresarial.

## PUBLICIDAD

# VISUALISATIONS AND NARRATIVES IN DIGITAL MEDIA

## Methods and current trends

EDITORS

Pere Freixa, Lluís Codina, Mario Pérez-Montoro & Javier Guallar



**DigiDoc**  
Digital documentation and  
interactive communication  
Research group

Profesional de la  
**información**