

Álvarez García, S; Gértrudix Barrio, M; Rajas Fernández, M

La construcción colaborativa de bancos de datos abiertos como instrumento de empoderamiento ciudadano

Revista Latina de Comunicación Social, núm. 69, enero-diciembre, 2014, pp. 661-683

Laboratorio de Tecnologías de la Información y Nuevos Análisis de Comunicación Social
Canarias, España

Disponible en: <http://www.redalyc.org/articulo.oa?id=81931771032>



Revista Latina de Comunicación Social,

ISSN (Versión electrónica): 1138-5820

jpablos@ull.es

Laboratorio de Tecnologías de la Información y

Nuevos Análisis de Comunicación Social

España

Cómo citar este artículo / Referencia normalizada



S Álvarez García, M Gértrudix Barrio, M Rajas Fernández (2014): “La construcción colaborativa de bancos de datos abiertos como instrumento de empoderamiento ciudadano”. *Revista Latina de Comunicación Social*, 69, pp. 661 a 683.



http://www.revistalatinacs.org/069/paper/1029_UR/32es.html


DOI: [10.4185/RLCS-2014-1029](https://doi.org/10.4185/RLCS-2014-1029)

La construcción colaborativa de bancos de datos abiertos como instrumento de empoderamiento ciudadano

The collaborative construction of open databases as tools for citizen empowerment

S Álvarez García [[CV](#)] [ [ORCID](#)] [ [GS](#)] Profesor del Departamento de Comunicación 2 y Ciencias Sociales - Universidad Rey Juan Carlos (URJC) - sergio.alvarez@urjc.es

M Gértrudix Barrio [[CV](#)] [ [ORCID](#)] [ [GS](#)] Profesor del Departamento de Comunicación 2 y Ciencias Sociales - Universidad Rey Juan Carlos (URJC) - manuel.gertrudix@urjc.es

M Rajas Fernández [[CV](#)] [ [ORCID](#)] [ [GS](#)] Profesor del Departamento de Comunicación 2 y Ciencias Sociales - Universidad Rey Juan Carlos (URJC) - mario.rajas@urjc.es

Abstracts

[ES] **Introducción:** En el contexto actual de medios y redes sociales digitales, cuando la ciudadanía pasa de consumidora a ‘prosumidora’ de información, comienza a demandar transparencia a los poderes públicos y a tomar conciencia de su potencialidad colectiva para cambiar el mundo en el que vive, compra y vota. **Metodología:** Este artículo recurre a la técnica del estudio de casos para analizar portales de datos en abierto e identificar los elementos tecnológicos y estratégicos que facilitan una reutilización y redifusión colaborativas de la información para la toma de decisiones y la construcción social de la realidad por los ciudadanos. **Resultados y conclusiones:** Puede hablarse de un ‘movimiento’ de iniciativas basadas en datos abiertos, que liberan información pública de temática fundamentalmente financiera y política; que la organizan atendiendo principalmente a criterios geopolíticos; y que la ponen a disposición de los ciudadanos mediante estándares tecnológicos y modelos de participación que señalan el camino de los cambios.

[EN] **Introduction:** In the current context of digital media and social networks, in which citizens have gone from being consumers to ‘prosumers’ of information, they begin to demand transparency from the government and to become aware of their collective potential to change the world in which they live, buy and vote. **Method:** This article uses the case study technique to analyse open data

portals and identify the technological and strategic elements that make it easier for citizens to collaboratively re-use and retransmit information that allows them to make informed decisions and construct their social reality. **Results and conclusions:** There is a ‘movement’ of open data initiatives, which are mostly dedicated to release public financial and political information; to organise it mainly according to geopolitical criteria; and make it available to citizens through technological standards and participation models that show the way to achieve change.

Keywords

[ES] Banco de datos; empoderamiento ciudadano; datos abiertos; tecnologías emergentes.

[EN] Database; citizen empowerment; open data; emerging technologies.

Contents

[ES] 1. Introducción. 1.1. Las bases de un movimiento: participación ciudadana, periodismo ciudadano y *crowdsourcing*. 1.2. Más allá de la comunicación programada: los ciudadanos frente a los datos. 1.3. Los bancos de datos abiertos vinculados en la Red. 2. Metodología. 2.1. Estrategias metodológicas. 2.2. Muestra de análisis. 2.3. Técnicas e instrumentos de investigación. 3. Resultados. 3.1. Liberando datos financieros. 3.2. *Open Data* para el desarrollo regional. 3.3. Concentración y vinculación de datos abiertos. 3.4. Datos abiertos para resetear el periodismo. 4. Conclusiones y discusión. 4.1. El ‘movimiento’ de datos abiertos para el empoderamiento. 4.2. Agenda y encuadre para la construcción social abierta de la realidad. 4.3. Tecnología y estrategias para la colaboración y la apertura: en busca del Ciudadano *OpenData*. 5. Notas. 6. Bibliografía.

[EN] 1. Introduction. 1.1. The foundations of a movement: citizen participation, citizen journalism and crowdsourcing. 1.2. Beyond programmed communication: citizens and open data. 1.3. Open databases on the Web. 2. Method. 2.1. Methodological strategies. 2.2. Sample of analysis. 2.3. Research techniques and instruments. 3. Results. 3.1. Opening financial data. 3.2. Open data for regional development. 3.3. Concentration and linking of open data. 3.4. Open data to restart journalism. 4. Conclusions and discussion. 4.1. The ‘movement’ of citizen empowerment through open data. 4.2. Agenda and framing for the open social construction of reality. 4.3. Technology and strategies for collaboration and openness: in search of the open data citizen. 5. Notes. 6. List of references.

Traducción de **Cruz Alberto Martínez-Arcos** (Doctor en Medios y Comunicaciones por la Universidad de Londres, 2012, y Profesor en la Universidad Autónoma de Tamaulipas, México)

1. Introducción

El artículo muestra los resultados iniciales de un estudio que evalúa el fenómeno de la participación ciudadana en la construcción de bancos de datos abiertos como instrumento primigenio de la producción de información basada en análisis, tratamiento y visualización de datos, y cómo las acciones generativas de estos productos en la Red son una fórmula diferencial de empoderamiento al servicio de la ciudadanía.

Los ciudadanos tienen hoy la capacidad de participar en la construcción de la información de una forma inequívocamente distinta a como lo venían haciendo hasta ahora. Uno de los aspectos más sugerentes que ofrece el actual escenario de medios sociales es la capacidad de la ciudadanía para intervenir activamente en el flujo comunicativo en sus diferentes fases: la constatación de los

acontecimientos, la producción de piezas informativas, el filtrado y gestión de información o la redifusión de esta.

1.1. Las bases de un movimiento: participación, periodismo y *crowdsourcing*

El periodismo ciudadano implica la participación activa de los ciudadanos de forma no profesional en el mundo del periodismo (Bowman y Willis, 2003; Gillmor, 2004; Meso, 2006; Armentia, 2008; Sampetro, 2009; López, 2009; Gómez y Méndez, 2013). Para algunos autores (Nip, 2006; Rosenberry y St. John, 2010) los orígenes del periodismo ciudadano o periodismo participativo están vinculados con el denominado periodismo cívico (*civic journalism*), del que recoge la necesidad de dar entrada al público en los medios, participación ahora facilitada por internet y las herramientas 2.0. (García de Madariaga, 2006). Sin embargo, algunos autores ponen en tela de juicio el término vinculado a las aportaciones ciudadanas, por su dependencia respecto a la agenda de los propios medios, su falta de periodicidad y la escasez de fuentes contrastadas (Domingo, 2008; García y Capón, 2004; Varela, 2005; Orihuela y Cambronero, 2006; Almirón, 2006; Reich, 2008; Lacy, 2010; Rost, 2010).

La web 2.0. ha favorecido el desarrollo de nuevos instrumentos de participación ciudadana como el *crowdsourcing* (Howe, 2006), un modelo de colaboración basado en la participación masiva de voluntarios que nace vinculado al desarrollo *open source* (Torvalds, 1992) y que se corresponde con el ‘modelo bazar’ de producción descrito por Eric S. Raymond (1997), cuyas raíces están en la economía del don y del prestigio (Mauss, 1925) actualizada por el *engagement* y la *gamificación* (Takahashi, 2010). Para los defensores de este modelo de creación colaborativa o co-creación, el *crowdsourcing* responde a la idea del consumidor como ‘prosumidor’ (Toffler, 1981; Friedman, 2005) y a que la colaboración masiva y especializada en la Red favorece la calidad de los resultados, porque condiciona positivamente las aportaciones individuales (Efecto Hawthorne).

Desde la teoría crítica de la co-creación (Ritzer y Jungerson, 2010; Cova, Dali y Zwick, 2011), este modelo responde a un nuevo concepto de externalización (*outsourcing*) no remunerado que aprovecha el potencial del software social y que en la práctica no supone un cambio real en el rol de los consumidores (Humphreys y Grayson, 2008).

En el ámbito periodístico, el *crowdsourcing* consiste en la recolección por parte de periodistas profesionales de material informativo aportado por un gran número de ciudadanos (Fondevila Gascón, 2013), de manera que las aportaciones ciudadanas, a través del uso de herramientas colaborativas en la Red (Niles, 2007), se convierten en fuentes para la práctica periodística (Gillmor, 2010).

Un aspecto que ha cobrado especial significación con la emergencia de la comunicación local desde la dimensión de los sistemas de datos abiertos, a través de iniciativas como la inglesa *Openly Local*, las innovaciones en el campo del Periodismo Líquido (Deuze, 2008) o la *reportería* distribuida, colaborativa y de código abierto (Briggs, 2007).

En cuanto a las motivaciones que incentivan estos modelos de participación abierta y desinteresada, y que responden al “principio de intercambio o la donación diferida” (Ortega y Rodríguez, 2011: 26), han sido profusamente estudiadas desde diferentes ámbitos como el antropológico (Maus, 1925; Malinowski, 2001), el económico (Polanyi, 1989; Tapscott y Williams, 2007), el biológico (Pennisi, 2009) o el cultural (Kelty, 2008).

1.2. Más allá de la comunicación programada: los ciudadanos frente a los datos

El fenómeno de los datos abiertos profundiza en las posibilidades sociales y comunitarias que permiten, desde hace ya una década, los servicios y contenidos digitales abiertos (García y Gértrudix 2011; Álvarez y Gértrudix, 2011). La capacidad de modificación, reutilización, combinación, comentario, recomendación, selección, registro, y cualquier otra operación de re-construcción o remixación por adicción, supresión, yuxtaposición, o combinación (Sonvilla-Weiss, 2010) que ya era posible con los documentos, en el nuevo escenario *Linked Data* (Heart y Bizer, 2011) y de web semántica es viable también con los datos, lo que abre un inusitado campo de oportunidades en la elaboración de contenidos para los medios, pero también pone en manos de los ciudadanos una posibilidades hasta ahora desconocidas.

Tradicionalmente, los ciudadanos –los individuos masa o alineados programados de Freytas (2009)– han tenido un acceso mediatizado a los datos a través de los medios de comunicación de masas, y casi siempre a través de documentos derivados de estos. La carencia de una auténtica cultura de la transparencia y de rendición de cuentas (Schultz, 1998; Eberwein, Flenger, Lauk y Leppik, 2011; Howard, 2012), la ausencia de una regulación clara hasta fechas recientes, y la imposibilidad de disponer de tecnologías adecuadas, ha impedido que se hayan tenido oportunidades ciertas de disponer de forma directa y sencilla acceso a los *datasets* y a los microdatos que resultan de interés para los asuntos públicos.

Sin embargo, en los últimos años estamos asistiendo a un cambio de conciencia y actitud ciudadana, concretado en un proceso cultural de empoderamiento colectivo sobre la capacidad que el acceso, registro, tratamiento y procesado de los datos en abierto otorga para el ejercicio de una acción ciudadana y política más directa, más informada, más crítica y, en consecuencia, más comprometida. Ello ha dado carta de naturaleza al nacimiento de distintas iniciativas, más o menos organizadas a través de distintas fórmulas en la Red, que no sólo ejercen presión sobre los responsables de la Administración o de las corporaciones para que hagan accesibles los datos de una forma adecuada para su tratamiento, sino que además cuentan con mecanismos y capacidad cooperativa suficiente para trazar aquellas partes de los mapas de datos que se no se ofrecen o directamente se ocultan.

Este concepto de ‘dato disponible en abierto’ –que lo es bajo los conceptos de fuente, producción, mediación, distribución, recepción y cooperación, *feedback* e interrelación– supone un cambio de cultura organizativa que se deriva en cambios en las lógicas comunicativas, pero también en la comprensión del fenómeno y de la intervención ciudadana como un elemento que genera valor añadido –cadena de valor bajo el paraguas del sector *Infomediario*–. Pero se trata también de un instrumento de calidad democrática que se proyecta en nuevas aperturas, pues evidencia las necesidades, los problemas, las posibilidades o las oportunidades, y, sobre todo, porque facilita una intervención más directa del ciudadano en los asuntos públicos, le empodera para ejercer un control continuado de la acción política y empresarial, y contribuye a alimentar un marco necesario de transparencia y rendición de cuentas.

Este proceso de toma de conciencia requiere un conjunto de pasos que quedan ilustrados a través de los 10 principios enumerados por la Coalición Pro Acceso [1]. El primer paso del empoderamiento consiste en tomar conciencia de la existencia del objeto y del derecho del sujeto sobre él, y en este planteamiento equivaldría a los tres primeros principios y al último que establecen el “derecho de acceso a la información como derecho fundamental, su cobertura a todas las entidades públicas, a todos los poderes del Estado y a todas aquellas entidades privadas que realicen funciones públicas, su ámbito de aplicación a toda la información elaborada, recibida o en posesión de las entidades

públicas, sin importar cómo esté almacenada”, y a la exigible independencia de los órganos que deben garantizar el derecho de acceso.

El segundo, recogido en los principios 4 al 8, afecta al conocimiento de los derechos de las condiciones de acceso de los ciudadanos, y a cómo exigirlos. El tercero, indicado en el noveno principio, refleja tanto el conocimiento del derecho como del deber, en cuanto que como ciudadanos también formamos, o podemos formar parte, de organismos que están sujetos a esa transparencia. En el fondo, se trata de una dimensión de fortalecimiento de la cultura de la transparencia y del derecho a saber. Y ello implica, a su vez, una concienciación sobre la naturaleza y características de un servicio público basado en la rendición de cuentas, no porque los ciudadanos exijan esto, sino porque ha de formar parte del modelo y de las prácticas comunes de ese ejercicio de la actividad pública.

1.3. Los bancos de datos abiertos vinculados en la Red

En el ámbito del uso de los datos abiertos como instrumento de empoderamiento, la estandarización es una demanda de igual o mayor importancia que la que se realiza desde otros sectores y actividades. La International Organization for Standardization (ISO) enriquece el concepto destacando que el fin último de los estándares es conseguir un nivel de ordenamiento óptimo en los contextos económico, tecnológico o político, a través de procesos que orientan hacia usos comunes y repetidos ante problemas reales o potenciales. En concreto, las normas que la organización considera de vital importancia son las referidas a “la calidad, la ecología, la seguridad, la economía, la fiabilidad, la compatibilidad, la interoperabilidad, la eficiencia y la eficacia”, por su capacidad para facilitar el comercio, pero también para “propagar el conocimiento, compartir los avances tecnológicos y las buenas prácticas de gestión” (ISO, 2011).

Estas características son necesariamente exigibles a los datos en abierto, y en concreto a la gestión del conocimiento que les acompaña en su camino hacia la (re)construcción social de la realidad por la ciudadanía. Los procesos y normas de estandarización son fundamentales para garantizar la calidad de las aportaciones a las bases y colecciones de datos; para facilitar y agilizar el tratamiento y el análisis de la información; para permitir la recuperación de ese conocimiento en condiciones óptimas; para afianzar los cimientos de la explotación de los datos en bruto; y para dotar a esos datos del hilo suficiente y oportuno en el tejido de la vinculación con otros datos, interpretaciones o productos informativos.

Los datos abiertos vinculados reúnen estas demandas y potencialidades para relacionar la estandarización con el empoderamiento: se trata de datos brutos estructurados y publicados en la Red que son interoperables y, por lo tanto, pueden interconectarse entre sí para ofrecer una mejor experiencia de usuario en su gestión integral. Como indican Bizer y otros, “La meta de *Linked Data* es permitir a la gente compartir datos estructurados en la web tan fácilmente como se comparten documentos hoy” (2011).

Para ello es necesario que cumplan determinados requisitos: a) Estar publicados bajo un modelo de licencia abierta; b) Los enlaces vinculantes de los datos deben estar representados de la forma más simple utilizando RDF (*Resource Description Framework*); c) Deben tener un localizador único (URI); d) Deben estar publicados en línea bajo el protocolo ‘http’.

Las ventajas que ofrecen los datos abiertos vinculados son, entre otras: a) Mejorar notablemente la experiencia de búsqueda de los usuarios al eliminar las ambigüedades semánticas de una búsqueda y b) Ofrecer la posibilidad de generar e implementar aplicaciones basadas en datos propios o de otros.

El primer elemento sujeto a estas prescripciones es el formato de los datos. Como aclara el Portal de Gobierno Abierto de la Comunidad Foral de Navarra [2], “la información se publica en formatos de datos estructurados para facilitar que pueda ser utilizada de forma automática por los lenguajes de programación”, lo que significa la multiplicación de opciones y posibilidades de acceso, tratamiento, difusión y reutilización. La estructura es la clave en la diferenciación y optimización de la gestión de los datos, como se puede apreciar en el acercamiento a los formatos más extendidos en este ámbito: representación de datos y fórmulas en formato de tabla (CSV, XLS/XSLX, ODS); intercambio de datos basado en JavaScript (JSON); almacenamiento de datos espaciales (SHP); intercambio de información estructurada entre plataformas diferentes (XML), etcétera.

A la estructura se añade la potencialidad para la representación que tienen los datos, es decir, para recibir cargas de significado y protagonizar por tanto el devenir de la llamada web semántica: los microformatos –definidos y estandarizados por la W3C– y los microdatos –personalizables– explican a las máquinas cómo pueden trabajar con los datos, cómo vincularlos, cómo hacerlos, en definitiva y en la práctica, interoperables, accesibles, comprensibles; y finalmente, capaces de articularse como piezas de construcción colaborativa de conocimiento.

Directamente relacionada con estas funciones está la redifusión o sindicación web, que permite multiplicar factorialmente la publicación de información en la Red, mediante la presentación de datos procedentes de una misma fuente en los espacios suscritos a los contenidos y actualizaciones de dicha fuente, espacios habitualmente vinculados a través de los medios y redes sociales; a su vez, estos espacios pueden agregar información procedente de otros emisores, enriqueciendo y facilitando el contraste y la comparación de datos y su interpretación.

Todas estas operaciones de publicación, redifusión y reutilización se regulan a través de un marco legal al que, en el caso de los datos en abierto, se incorporan iniciativas específicas de licenciamiento que responden –como difícilmente puede hacerlo el *copyright* de todos los derechos reservados– a los objetivos definidos en relación con el empoderamiento de la ciudadanía a través de la participación en la construcción de la información y en el flujo comunicativo.

En el ámbito de la publicación de contenido, el modelo que más se ajusta al objetivo de hacer extensiva la información basada en datos en abierto a la mayor parte posible de la ciudadanía, es el de las licencias Creative Commons [3], permiten gestionar derechos relacionados con la comunicación pública: el reconocimiento de la autoría; el uso comercial y la producción de obras derivadas.

En el caso de los datos propiamente dichos, la iniciativa Open Data Commons [4] desarrolla tres tipos de licencias básicas, siempre referidas específicamente a datos y a bases de datos, y que responden en mayor o menor medida a los requisitos de redifusión, uso y reutilización mencionados: todas ellas permiten la copia o distribución, la producción de trabajos a partir de los datos y la adaptación o modificación de las bases de datos, sin más restricciones, como la Public Domain Dedication and License (PDDL); con la única obligación de atribuir la información generada como la Attribution License (ODC-By); y con la finalidad de extender en el tiempo el carácter abierto de los datos originales, que solo pueden ser tratados si los productos o adaptaciones se publican y difunden bajo la misma licencia, la Open Database License (ODC-ODbL).

2. Metodología

2.1. Estrategias metodológicas

El objetivo de este estudio es identificar en qué medida y bajo qué manifestaciones los bancos de datos en abierto pueden sentar las bases para la construcción colaborativa de la realidad social y para la instrumentación de estas construcciones en el empoderamiento ciudadano.

Este objetivo se instrumentaliza a través de varias preguntas de investigación, que constituyen el factor determinante del enfoque, la técnica y la muestra a seleccionar:

- ¿Existen características comunes suficientes en las iniciativas para el empoderamiento ciudadano basadas en datos en abierto como para hablar de una marcada tendencia o de un ‘movimiento’?
- Si los medios de comunicación de masas han venido determinando la construcción social de la realidad mediante la agenda-setting y el framing o encuadre, ¿cómo se configura temáticamente la información basada en datos abiertos para intentar jugar ese papel?
- ¿Qué peso tiene la tecnología en la conformación de entornos y plataformas digitales que faciliten el trabajo colaborativo con datos en abierto y la difusión de los propios datos y de los productos informativos basados en ellos a una masa crítica de ciudadanos?
- ¿Qué bases estratégicas de publicación es posible articular desde las iniciativas de bases de datos colaborativas para el empoderamiento ciudadano, con el objetivo de potenciar la eficacia de los procesos de información para la toma de decisiones por la ciudadanía?

Para responder a estas preguntas de investigación es obligado concentrar la mirada en los bancos de datos abiertos que se han venido identificando en la fase inicial de las investigaciones con las que se relaciona el presente artículo. Esos proyectos e iniciativas constituyen el objeto formal de estudio y a ellos se aplican las consideraciones metodológicas referidas a continuación.

Siguiendo la identificación de paradigmas metodológicos de Bermejo Berros (2014: 333), la fase de la investigación sobre datos abiertos y empoderamiento ciudadano de la que se ocupa este artículo responde a las características de la ‘perspectiva interpretativa’: se entiende que las comunidades de usuarios “construyen una realidad social a través de los procesos de comunicación”, una “construcción social intersubjetiva que genera espacios de intercambio simbólicos” y que “se interesa más por entender en profundidad los procesos de comunicación que por la función que éstos pueden tener para modificar conductas”.

Cabe matizar que, aunque el análisis se realiza desde un enfoque sistémico, el empoderamiento no es interpretable como una modificación de la conducta de la ciudadanía, y que lo que persigue este análisis no es identificar las estrategias o técnicas de comunicación eficaz para provocar determinados efectos en los usuarios de portales de datos abiertos, sino comprender qué modelos instrumentales articulan la información para ponerla al servicio de un conjunto de conductas y actitudes dadas.

El procedimiento metodológico seguido es el propio de la perspectiva interpretativa ya comentada: se trata de una investigación empírico-inductiva, que desde un claro enfoque cualitativo realiza la

observación de los sistemas y procesos de construcción colaborativa de bancos de datos abiertos, cuyos resultados son sometidos a un razonamiento inductivo por parte de los autores.

2.2. Muestra de análisis

Tanto el tipo de objetivos que plantea la investigación –con la identificación de procesos y elementos sistémicos de la construcción colaborativa como meta destacada–, como la técnica y el enfoque de las herramientas empleadas –observación y análisis cualitativos de portales de iniciativas de apertura de datos públicos en la Red–, como las operaciones para la interpretación de los resultados –razonamiento inductivo– conducen sin equívocos a la pertinencia de un muestreo no probabilístico, más concretamente de conveniencia. A partir de estas consideraciones, los criterios para la selección de los casos a analizar en esta fase de la investigación se resumen en la pertinencia de iniciativas no gubernamentales que destaquen por su declarada intención de poner a disposición de los ciudadanos espacios para la recopilación e intercambio de datos de interés político y económico que no proporcionan de forma abierta las administraciones o gobiernos correspondientes.

Dado que el objetivo último es identificar los procesos y sistemas que convergen en la construcción colaborativa de bancos de datos abiertos, no se ha estimado conveniente incorporar a la muestra casos que podrían responder a una cobertura geográfica o temática completas. No obstante, en la revisión del estado de la cuestión se ha observado un número elevado de este tipo de iniciativas, con una doble finalidad: metodológica, en tanto en cuanto ha permitido identificar los criterios óptimos de configuración de la muestra; y analítica, con un enriquecimiento del capítulo de conclusiones a partir de la referencia y relación de estos ejemplos con los resultados del estudio de los casos seleccionados.

Esta amplia observación de iniciativas sirve para constatar la naturaleza emergente y heterogénea del objeto de esta investigación, lo que introduce una nueva demanda en la concreción de la muestra específica de análisis: la necesidad de contar con casos a los que se pueda considerar exploradores, precursores, multiplicadores... en el ámbito del estudio; y de sumar los suficientes para no dejar fuera del examen ninguno de los rasgos principales de la heterogeneidad del fenómeno que se estudia.

Teniendo en cuenta todas estas consideraciones y condiciones, se ha realizado una selección de ocho iniciativas específicas de construcción colaborativa de bancos de datos abiertos relacionados con temas recurrentes del empoderamiento ciudadano. Tres de ellas (Open Corporates [5], Open Spending [6], Open Charities [7]) pertenecen a la iniciativa general Open Company Data Index para la apertura de datos de empresas y corporaciones, un proyecto que ha facilitado la generación de numerosos productos derivados, entre los que destacan por su difusión los dedicados a presentar el gasto al que se destinan los impuestos en diferentes países.

Open Data Latinoamérica [8] y Africa Open Data [9] constituyen modelos de colaboración ceñidos a una región geográfica, y representan a un tipo de iniciativa que se ha multiplicado para atender a zonas regionales de extensión diversa. Otros dos casos (Get the Data [10] y Data Hub [11]) destacan por ser los primeros en explorar la concentración o la construcción de foros de intercambio de datos procedentes de otros proyectos abiertos.

Y completa la muestra Poderopedia [12], una iniciativa que aglutina casos específicos como los de Chile, Venezuela o Colombia; que es pionera en el impulso del protagonismo del periodismo en la construcción colaborativa de información para el empoderamiento ciudadano; y que tiene como

objetivo principal proporcionar las bases para la creación de proyectos de transparencia y rendición de cuentas en otros países del mundo.

2.3. Técnicas e instrumentos de investigación

El estudio de caso se revela como la técnica óptima para abordar el análisis de estas ocho iniciativas que componen la muestra de análisis, fundamentalmente por su capacidad para, como afirma Yin (en Wimmer y Dominick, 1989: 160), “investigar un fenómeno actual dentro de su contexto de vida real, y en el que las fronteras entre el fenómeno y su contexto no quedan claramente delimitadas”. McDonald y Walker (en Albert, 2006: 216) entienden esta técnica como “un examen de un caso en acción”, es decir, como proceso, como actividad metodológica en permanente actualización –sobre todo en objetos abiertos y transformativos como los analizados en esta investigación–. También destaca el estudio de casos por su valor en la búsqueda de relaciones entre factores y elementos, de tendencias o caminos que se van trazando, más que de hechos o sucesos que poder únicamente confirmar.

Desde esta perspectiva, el diseño de la investigación se concreta a través del desarrollo del modelo de análisis DEPUC (DEscripción, Participación, Uso y Características), que se despliega a través de los siguientes criterios e instrucciones de la plantilla para recogida de información y análisis de los casos seleccionados:

Tabla 1. Instrumento de recogida de información de los casos de la muestra de análisis

Criterios	Iniciativa Open Data
URL	URL principal del proyecto o iniciativa
Temática y objetivos	Breve descripción de los fines y sentido de la iniciativa y de la temática o foco de interés de su trabajo de apertura de datos
Soporte	Entidades que apoyan financiera o institucionalmente la iniciativa, distinguiendo en la posible entre Entidades financiadoras y Entidades de Apoyo
Ámbito	Alcance geopolítico de la iniciativa (Mundial, Regional -zona-, Nacional o Local)
Modelo de participación	Identificar el sistema de procesos y normas que regulan la aportación de datos a la iniciativa
Modelo de explotación	Identificar el sistema de procesos y términos de uso de los datos y productos derivados publicados
Organización de datos	Cómo se publican y se da acceso a los datos en abierto que proporciona la iniciativa a través de su portal o plataforma en Internet. Indicar si se presentan los siguientes instrumentos relacionados: catálogo; buscador; colecciones; datos georreferenciados
Formato de datos	Formato en el que se da acceso a los usuarios a los datos que proporciona la iniciativa. Indicar si se proporcionan datos en los siguientes formatos: CSV, XLS, JSON, XML, RDF, HTML, PDF, DOC u otros que pudieran identificarse
Redifusión social	Instrumentos, canales y redes de difusión de los datos en abierto y/o de contenidos y productos –visualizaciones, reorganización...– generados a partir de ellos. Indicar si se habilitan opciones para compartir datos e información relacionada en los siguientes canales, medios o redes sociales: Facebook, Twitter, Google+, LinkedIn u otros que pudieran identificarse
Reutilización	Herramientas o aplicaciones que permiten al usuario reutilizar los datos que proporciona la iniciativa en la consecución de nuevas bases o colecciones de datos, o de contenidos o productos

	generados a partir de ellos. Indicar específicamente si se identifican estos dos servicios: embebido en sitio web; API (aplicación de tratamiento de datos) y, en este último caso, comentar lo que se estime necesario
Documentación	Identificar si la iniciativa ofrece documentación específica para el uso de las aplicaciones y servicios de acceso y tratamiento de los datos, en cualquiera de las áreas relacionadas. Si se estima procedente, indicar el/los formato/s de la documentación proporcionada
Licencia de la oferta de Datos	Identificar el tipo de licencia que especifica al usuario si puede publicar –y en qué condiciones hacerlo– los datos proporcionados o adaptaciones y productos derivados o generados a partir de ellos. Observar específicamente la declaración sobre las siguientes licencias: ODbL (Open Data Commons Open Database License); PDDL (Open Data Commons Public Domain Dedication and License); ODC-By (Open Data Commons Attribution License); Creative Commons –y sus múltiples variantes–

3. Resultados

El estudio de los casos seleccionados como corpus de análisis aporta información suficiente para poder responder a las preguntas de investigación planteadas. En los siguientes párrafos se resume el conjunto de resultados de la observación detallada de estos portales de información basada en datos abiertos, según su naturaleza y enfoque particulares.

3.1. Liberando datos financieros

OPEN CORPORATES (<http://opencorporates.com/>) es un repositorio de datos con información de empresas de todo el mundo, tanto de entidades y corporaciones como de personas que trabajan o están vinculadas a ellas. De temática financiera, afirman a modo de eslogan que su objetivo es “hacer una cosa sencilla (aunque grande): tener una URL por cada compañía que haya en el mundo”. A principios de junio de 2014, ya había conseguido información de más de 70 millones de corporaciones.

Es una iniciativa de Chris Taggart y Rob McKinnon, colaboradores de diversos proyectos de apertura de datos públicos en el Reino Unido y directores de la empresa de tecnologías de la información y actividades informáticas Chrion Ltd., que apoya financiera y legalmente el proyecto. En productos específicos cuenta con el soporte de instituciones como el World Bank Institute [13].

Contemplan varios grados de participación en la aportación de datos al proyecto: desde informar por correo electrónico de la localización de registros oficiales de empresas para que Open Corporates consiga los datos a través de su colaboración con ScraperWiki [14], hasta el registro como colaborador y la extracción e incorporación de datos mediante herramientas y servidores específicos.

La información se presenta al usuario a través de un buscador sencillo y de filtros avanzados por países o sector de actividad de las empresas. Puede accederse de forma gratuita y bajo petición a las bases de datos, siempre que no se vayan a utilizar con propósitos comerciales. Los datos se ajustan a los estándares habituales de reutilización y redifusión y pueden ser tratados con una API de desarrollo propio que es compatible con la herramienta Google Refine.

La política de redifusión y apertura queda patente en la publicación en el blog del proyecto de los productos y resultados basados en sus datos, en las facilidades para compartirlos a través de Facebook, Twitter, Google+ y LinkedIn, y en la adopción de la licencia ODC Open Database License (ODbL). Desarrollan propuestas y proyectos complementarios para acercar la información a los

ciudadanos mediante interfaces amigables y comprensibles, como el portal de seguimiento de corporaciones OPENLEIs [15], el mapa interactivo Open Corporate Network Data [16], o el índice Open Company Data Index [17] –financiado por el mencionado World Bank Institute–.

OPEN SPENDING (<http://openspending.org/>) es un servicio en línea de la Open Knowledge Foundation [18] que monitoriza y almacena las operaciones financieras de los gobiernos en el mundo bajo el lema “¡Es nuestro dinero!”.

La construcción colaborativa es el principal motor de la iniciativa, que se define a sí misma como “proyecto impulsado por una comunidad [19] para crear una base de datos abierta y gratuita de las transacciones financieras públicas a lo largo del mundo” y cuyos miembros eligen al Grupo de Dirección de entre personas de diversas instituciones del ámbito de los datos abiertos.

Ofrece a los usuarios un amplio catálogo de *datasets* [20] clasificados por países e idiomas y distintas herramientas con las que poder almacenar nuevos *dataset* [21], visualizar los existentes [22] o tratarlos para crear nuevas aplicaciones a través de su API abierta [23]. Dedicamos amplios recursos a documentar, guiar e incluso a formar a los usuarios en la generación de productos informativos multimedia e interactivos basados en sus datos.

Publica los resultados y productos generados por la comunidad a través de su blog y de los propios sitios de los proyectos desarrollados a partir de sus datos, que se ajustan a los estándares mínimos exigibles; desde esos espacios se difunde por medios y redes sociales. Su invitación y promoción del *crowdsourcing* es una constante en los esfuerzos de esta iniciativa, que se complementa con los usos abiertos y con la vocación de construcción colaborativa que permite y fomenta la licencia ODC Open Database License (ODbL).

OPEN CHARITIES (<http://opencharities.org/>) es un portal que parte de los datos no reutilizables del portal web de la Charity Commission –sobre instituciones de caridad– del Reino Unido y los organiza en una base de datos abiertos a los que puede acceder cualquier ciudadano. El proyecto está desarrollado por Pushroad, una pequeña empresa dedicada a la innovación web y a la apertura de datos públicos, y liderado por Chris Taggart –miembro del Local Public Data Panel, desarrollador de OpenlyLocal.com y responsable de Open Corporates–.

Se trata de una iniciativa de dimensiones reducidas, aunque valiosa por su enfoque local y su alto nivel de concreción. Se centran en la apertura de información gubernamental no reutilizable, por lo que la colaboración del usuario en el proyecto se reduce a solicitar actualización de información. No recogen ni publican productos generados a partir de las fuentes por terceros, se limitan a ofrecer acceso a los datos a través de tres herramientas: la base de datos completa, los listados de instituciones –filtrados por nuevas incorporaciones, mayores ingresos y mayores gastos– y un buscador.

Utilizan los formatos de datos estandarizados habituales; apuestan por la redifusión a través de fuentes para la sindicación de contenidos (RSS) y de la opción ‘Share this’ para compartirlos a través de un amplio número de servicios, medios y redes sociales; y cuentan con una API que permite tratar y reutilizar los datos. Todo ello apoyado en la licencia la licencia ODC Open Database License (ODbL).

3.2. Open Data para el desarrollo regional

OPEN DATA LATINOAMÉRICA es una plataforma que detecta sets de datos de gobiernos e instituciones, y libera información recolectada en encuentros, *workshops* y *hackatones* a través de las actividades desarrolladas por los capítulos regionales de organizaciones como Hacks/Hackers, International Center for Journalists, PinLatam o Poderopedia. Está conectado temática e institucionalmente con el programa de Desarrollo Global del World Bank Institute. Es significativa la importancia que dan a los criterios de organización y presentación de datos: geográficos –por país y mapas georeferenciados–, educación, finanzas públicas, salud y economía.

La colaboración en el proyecto se articula a través de una comunidad con tres tipos de perfiles de usuario básico: los que pertenecen simplemente a la comunidad, los que descargan *datasets* y los que aportan *datasets*. Para perfiles de consumo, el portal permite explorar el catálogo de conjuntos de datos a partir de los productos informativos destacados, de la organización por país, tipo de recurso o tema; o bien utilizando el buscador avanzado de la plataforma.

Para perfiles de reutilización o redifusión, se proporcionan opciones de descarga en formatos de datos estándar; de compartir en redes sociales, de trabajar a través de una API propia o de la inserción directa en una hoja de cálculo de Google. La vocación de servicio público se pone de manifiesto con intensidad en la configuración de términos de uso bajo la licencia Open Data Commons Public Domain Dedication and License (PDDL).

AFRICA OPEN DATA es una plataforma que aspira a convertirse en el mayor repositorio de datos sobre gobierno, sociedad civil, empresas y donaciones en África. Apoyada por las organizaciones Code for Africa, Open Knowledge Foundation y The Open Institute, es parte de la iniciativa de creación de un ecosistema impulsado por los datos en abierto UJUZI [24], financiada por la Iniciativa de Medios de África (IAM), el Instituto del Banco Mundial (WBI) y Google.

Declara una organización de *datasets* para su uso por ciudadanos, medios de comunicación, activistas, gobiernos y sociedad civil. En la práctica de la plataforma, se contempla un conjunto concreto de perfiles que permite detectar la preocupación por la flexibilidad en la construcción colaborativa de este repositorio: usuarios libres, comentaristas moderados –sus comentarios se supervisan y pueden ser publicados íntegramente, editados o no publicados– y contribuidores –aportan principalmente *datasets* procedentes de localizaciones o de liberaciones de datos, que son además valorados y recomendados por los usuarios de la comunidad–.

Todas estas operaciones se realizan en un entorno de interacción y trabajo que clasifica los datos por países y temas, que ajusta los formatos de datos a los estándares mínimos, y que permite la redifusión a través de los medios y redes sociales más extendidos. Su base tecnológica es CKAN [25], una plataforma de software libre que permite construir y gestionar portales de datos, y que proporciona una API para la explotación y reutilización de datos abiertos. En el aspecto legal, la relación de los datos con los ciudadanos se regula mediante la licencia Open Data Commons Attribution (ODC-BY).

3.3. Concentración y vinculación de datos abiertos

GET THE DATA (<http://getthedata.org>) es una iniciativa particular de Rufus Pollock and Tony Hirst que aborda el problema de la ubicación de los datos abiertos desde una perspectiva radicalmente social: son los ciudadanos que participan en la plataforma –con formato de foro

avanzado— quienes preguntan y responden sobre dónde y cómo pueden acceder a los datos que les interesan. Es una comunidad en el sentido más puro y simple del término: cualquiera puede preguntar y responder, editar las preguntas y respuestas de otros, y no hay más moderador que una/o mismo; la única regulación es el sistema de ‘karma’ habitual en los foros de usuarios más extendidos de la Red, y que otorga puntos de autoridad en función de la cantidad de aportaciones y de cómo valora la comunidad la calidad de las mismas.

La temática y la organización de los datos vienen dados por el volumen y la identificación de tópicos de los hilos de comentarios, preguntas y respuestas. El único formato exigido para los datos que se referencian es el enlace bajo el protocolo http; las únicas herramientas para la reutilización son las que puedan recomendar los participantes en la comunidad; las guías para abordar esa reutilización, la generación de productos informativos a partir de los datos referenciados o la redifusión de unos y otros son responsabilidad de la actividad de los usuarios del sistema. Todo ello con el marco legal de la licencia para contenidos Creative Commons Attribution-ShareAlike 3.0.

DATA HUB (<http://datahub.io/>) es otro de los proyectos que no se centra directamente en liberar datos relacionados con una determinada temática, objetivo o zona geográfica. Sus objetivos se sitúan en el tramo final del proceso de comunicación basado en datos abiertos para, como sintetiza su eslogan, ofrecer “el camino fácil para obtener, usar y compartir datos”. El *software* CKAN es la base de esta “plataforma gratuita y poderosa de gestión de datos” de la Open Knowledge Foundation.

Como localizador y concentrador de *datasets* procedentes de otros portales y fuentes, es una suerte de compendio de características y estándares en este ámbito, principalmente en lo que se refiere a formatos de datos abiertos, pero también en las vías de redifusión social o de licenciamiento.

Los esfuerzos y, por tanto, el principal aprendizaje que puede extraerse de esta iniciativa, se localizan en la organización del vasto conjunto de datos y colecciones que se concentran en el espacio público de su plataforma, y más concretamente en el filtrado: organizaciones, etiquetas, formatos de datos... Y su principal activo, la configuración de ‘grupos’ de *datasets* por el usuario que se registre en la plataforma, lo que conforma a su vez un espacio reconstruido y reorganizado a partir de las aportaciones y comentarios realizados desde los intereses específicos de la comunidad.

3.4. Datos abiertos para resetear el periodismo

PODEROPEDIA (www.poderopedia.org) se define a sí misma como una plataforma colaborativa cuyo objetivo es ayudar a entender las relaciones entre personas, empresas y organizaciones que ocupan la agenda de los medios de comunicación de masas y que a través de visualizaciones y otros productos informativos multimedia e interactivos, acercan al ciudadano un mundo tradicionalmente alejado y con demasiadas sombras.

Se interesa naturalmente por el poder, por quienes lo ejercen y por cómo lo hacen, con un enfoque eminentemente profesional que se sitúa en los últimos eslabones de la cadena del empoderamiento con datos abiertos, en los que se establece un diálogo directo con el ciudadano.

Resulta especialmente significativa la declaración de intenciones de la organización que, apoyada por la Knight Foundation y Start-Up Chile, promueve el proyecto: “Estamos reseteando el Periodismo. Promovemos la transparencia, la web abierta y la innovación digital en el Periodismo y los medios. Fomentamos el aprendizaje y trabajamos por el interés público. Usamos información y tecnología para redefinir el futuro de las noticias y cambiar el mundo”.

A pesar de la cercanía a la tarea periodística, la participación en el proyecto se realiza a través de varios niveles que incluyen, además de a periodistas, a programadores, diseñadores y colaboradores ciudadanos. Y aunque inicialmente desarrollan su actividad con carácter nacional –en Chile–, ofrecen soporte y apoyo a quien desee crear un ‘Capítulo’ nacional de Poderopedia en otros países.

Más preocupados por una presentación de la información clara y didáctica –la documentación sobre cómo usar e interpretar los contenidos y productos de Poderopedia es prolija y enriquecida en lo que a optimización de elementos multimedia se refiere–, los datos son el alimento de productos más elaborados, por lo que la atención a formatos estandarizados queda en un segundo plano. No así las posibilidades de redifusión y de viralización del contenido, entre las que destacan la sindicación o suscripción a las novedades del portal o la adopción de una licencia del ámbito de los contenidos multimedia, la Creative Commons 3.0 Atribución-NoComercial- SinObrasDerivadas, que permite además combinar una oferta gratuita con la puesta a disposición de servicios preferenciales con coste adicional.

4. Conclusiones y discusión

Lo observado en la revisión de estado de la cuestión y, sobre todo, en el análisis de casos, permite extraer las siguientes conclusiones, de carácter interpretativo y válidas para identificar tendencias y fenómenos en proceso.

4.1. El ‘movimiento’ de datos abiertos para el empoderamiento

Se ha detectado un número suficiente de tendencias y manifestaciones innovadoras comunes como para hablar de un ‘movimiento’ en torno a objetivos e intereses a los que orientar la apertura de datos y la construcción colaborativa de información basada en ellos, fundamentado en la convicción de que ello supone un impulso de regeneración democrática –aspecto ya apuntado en estudios previos (Borger, Hoof, Costera & Sanders, 2013)– y en la necesaria reacción frente a los modelos de una sociedad “programada” (Freitas, 2009) o “McDonalizada” (Ritzer, 1996).

Comenzando precisamente por esos fines, y tras explorar a su vez lo que declaran al respecto las instituciones que promueven, financian o apoyan los proyectos analizados, se observa una firme creencia en el poder del conocimiento colectivo como motor de cambio. Este aspecto aglutinador parece esencial a la vista de las dificultades apuntadas por estudios previos como los de Reich (2008) o Hurrell (2012) sobre la capacidad de pervivencia temporal de estas acciones ciudadanas.

La Open Knowledge Foundation –responsable de iniciativas como Open Spending, The Data Hub o Africa Open Data– lo resume a modo de eslogan: “Observa cómo los datos pueden cambiar el mundo, un mundo donde el conocimiento crea el poder de la mayoría, no de unos pocos”. El World Bank Institute –apoyo directo o indirecto de cinco de los ocho proyectos analizados– sintetiza su actividad en tres puntos –“conocimientos de libre acceso, gestión en colaboración y soluciones innovadoras”– que constituyen la fórmula óptima para hacer frente a los desafíos del desarrollo en un “entorno de constante cambio”.

En este camino de transformación, la información tiene el papel protagonista y, concretada en el fenómeno del OpenData, constituye el otro punto irrenunciable del ‘movimiento’: la liberación de los datos como derecho fundamental.

Siguiendo con los principios de las organizaciones que desarrollan los proyectos analizados, resultan especialmente expresivas las sentencias como “Un mundo donde los datos nos hacen libres” (Open

Knowledge Foundation); “Creemos que la democracia prospera cuando las personas y las comunidades están informadas y comprometidas” (Knight Foundation); “El gobierno abierto asegura el acceso de los ciudadanos a la información, datos, procesos gubernamentales, involucra a los gobiernos de forma más eficaz en la voluntad y capacidad para responder a los ciudadanos y trabajar en colaboración para resolver los problemas difíciles de gobernanza” (World Bank Institute).

El último rasgo que cabe destacar en la caracterización de este ‘movimiento’ es el de la naturaleza no lucrativa de quienes financian, apoyan o desarrollan los proyectos de datos en abierto para el empoderamiento. Fundaciones y particulares que piden gobiernos transparentes y abiertos, aunque no necesariamente desde posiciones o a través de vías no gubernamentales.

De hecho, la organización con mayor protagonismo en los casos analizados, el World Bank Institute, cuenta actualmente con 188 países miembros y su máximo órgano político –la Junta de Gobernadores– está compuesto generalmente por los Ministros de Finanzas o de Desarrollo de esos países. No obstante, autores como Howard (2012) han mostrado también la capacidad económica que este tipo de proyectos basados en datos pueden llegar a obtener por sí mismos.

4.2. Agenda y encuadre para la construcción social abierta de la realidad

En los párrafos anteriores se ha constatado que las iniciativas de liberación de datos para el empoderamiento buscan el cambio, y que ese cambio lo protagonicen los ciudadanos con sus decisiones informadas. Ahora bien, en este sentido cabe hacerse dos preguntas básicas: ¿sobre qué temas se informa a los ciudadanos desde estos proyectos? y ¿cómo organizan esa información?

En la práctica, se trata de identificar cuáles son las bases que estas iniciativas ofrecen al ciudadano para construir la realidad social, y en qué difieren o coinciden con las aplicaciones básicas de la *agenda-setting* –establecimiento de agenda temática– y el *framing* –encuadre, pautas para la interpretación de los mensajes– que se atribuyen a los medios de comunicación de masas, y que en el marco de estas prácticas, discuten autores como Bakker y Pantii (2009) o Rost (2010).

En la agenda de los portales de datos en abierto predominan los temas financieros, y específicamente, dos tipos de información relacionada con ellos: cómo son las empresas, organizaciones, gobiernos... ‘tras el telón’ –quiénes las dirigen o gestionan, dónde se ubican, dónde tributan, a qué se dedican, con quiénes se asocian o hacen negocio...–; y cómo es el gasto que se realiza –cuánto, en qué, cuándo, por qué...–.

La pregunta “Where does my money go?” es, en la práctica, un tipo de producto informativo basado en datos en abierto que se ha exportado desde su origen en el Reino Unido a otros países y con títulos aún más descriptivos de sus fines, como la visualización interactiva ¿Dónde van mis impuestos? [26] basada en los Presupuestos Generales del Gobierno de España en los últimos años, o la versión japonesa Where does my Money Go? [27] o la macedonia, How much does Macedonia cost us? [28].

Como ya se ha destacado en párrafos anteriores, el World Bank Institute es una de las instituciones más activas en el soporte y la financiación de proyectos de datos abiertos: el sello de sus fines se transmite con especial impronta a los proyectos de ámbito regional, en los que se incorporan temas con mayor contextualización y orientados al desarrollo –el portal Africa Open Data incorpora *datasets* sobre donaciones, gobierno o sociedad civil; Open Data Latinoamérica suma secciones de educación y de salud a las de carácter económico–.

La asociación temática de los datos se realiza de forma directa por parte del receptor o del usuario, sin intervención de los responsables de su publicación en cualquiera de los niveles de decisión que se encuentren. Es en la organización de la información donde radica el potencial subjetivo de estos emisores que parecen renunciar a su voz en favor de la que surja de los mismos datos, donde se produce –intencionalmente o no, con mayor o menor intensidad– el encuadre que orienta la lectura y el consumo de los datos por parte del usuario.

En la organización de los casos analizados en esta investigación, el criterio del “país” –como zona geográfica específica, como gobierno– es el más empleado, aunque no solo en aquellos proyectos en los que el usuario juega un papel pasivo en la arquitectura de estos espacios, sino también en los que la comunidad de participantes determina las secciones o las categorías de los datos. Unos y otros coinciden aquí en la asociación más extendida del concepto de transparencia, exigible en primer término a los gobiernos, comparable en niveles, leyes o aplicación entre países.

El potencial encuadre del emisor sobrevuela con mayor agilidad proyectos como Poderopedia, en los que la preocupación por explicar y hacer comprensibles los datos a los ciudadanos determina una oferta de productos en los que la intervención de –en este caso– los periodistas en la información es necesariamente mayor que cuando se ha de responder simplemente a criterios explícitos, reducidos y claros.

En resumen, las dos grandes áreas temáticas que se abordan son la economía y la política, con algunas ramificaciones que se expresan y matizan a través del modo en que se organizan y presentan los datos, pero que se resumen en una palabra, sin ambigüedades y en mayúsculas: ‘DINERO’.

Basta con invertir el tópico de que “el dinero mueve el mundo” para definir qué se selecciona y cómo se hace llegar la información basada en datos en abierto a los usuarios de estos portales: “el mundo mueve el dinero”. De la aplicación a esta sentencia de las 5 *W*’s que originan las preguntas clásicas del Periodismo (*Who, What, Where, When, Why*) resulta el esquema básico de agenda y encuadre que ponen en marcha las iniciativas analizadas en este estudio: quién mueve el dinero, qué dinero se mueve, y dónde, cuándo y por qué –o matizando, cómo o para qué– se mueve.

El lema del portal analizado Open Spending –que publica *datasets* de operaciones financieras de los gobiernos– parece responder a un ‘por qué’ con su lema “¡Es nuestro dinero!”; su promotora, la Open Knowledge Foundation, aclara la relación con la política cuando justifica su labor encaminada a que los ciudadanos tomemos “decisiones informadas sobre cómo vivimos, qué compramos y quién obtiene nuestro voto”. Una tendencia con la que se alinean otras iniciativas similares como Follow the Money [29]. Agenda y encuadre, por tanto, con la transparencia como motor de arranque y con el empoderamiento como principio –y como medio, y como fin– regulador.

4.3. Tecnología y estrategias para la colaboración y la apertura: en busca del Ciudadano *OpenData*

La tecnología aplicada a los datos y las estrategias que ponen en juego proyectos de *OpenData* como los analizados en este estudio, conforman una base que potencialmente determina la capacidad de generación de un entorno propicio para la construcción colaborativa de la realidad, en sus versiones tanto pasiva como activa.

Salvo excepciones –la comunidad Get the Data con sus enlaces o Poderopedia con su producción de noticias y visualizaciones–, en las iniciativas analizadas predomina el uso de formatos estandarizados

que permiten compartir y reutilizar –y por tanto construir en colaboración– datos y *datasets*: tras el omnipresente lenguaje de marcado HTML, los más frecuentes son JSON, XML, RDF y CSV, basados en fuentes abiertas; pero no hay limitaciones en el recurso a formatos propietarios como XLS (Microsoft Excel) o DOC (Microsoft Word), que en muchos ámbitos –también en el de los datos– pueden considerarse estándares de facto.

Este es uno de los aspectos más interesantes que revela el estudio: la fuerza que tienen los procesos de estandarización que se promueven a partir de estas y otras iniciativas como, por ejemplo, el desarrollo del sistema Global Legal Entity Identifier System (OPENLEIs). Este sistema es una iniciativa soportada por Open Corporates cuyo objetivo es facilitar un identificador único que permita reconocer de forma inequívoca la identidad corporativa de las empresas y, con ello, reconocer la traza de una entidad que opera en los mercados financieros y que puede tomar distintos nombres en función del país en el que opera, las abreviaciones con las que aparece, o que directamente cambia de nombre con el paso del tiempo.

El esfuerzo por configurar un entorno para trabajar con los datos abiertos que se ofrecen queda patente en la puesta a disposición –en todos los casos estudiados– de APIs que permiten el tratamiento de *datasets*. Estas APIs responden a tres modelos de desarrollo: producidas ad hoc para la iniciativa o portal concretos; adaptadas desde código de fuentes abiertas –como la que ofrece la plataforma CKAN–; o mixtas, con vínculos o compatibilidades con otras herramientas, como las de Google, Refine o Maps, o integración directa con Google Drive, etc.

Una conclusión similar puede extraerse de los términos de uso de los *datasets*, concretados en la política de licenciamiento asumida: los portales analizados regulan sus bases de datos y sus contenidos con licencias abiertas, alejadas del *copyright*-todos los derechos reservados. El tipo específico de licencia ‘Commons’ viene determinado en primer término por la información que se regula –Open Data Commons para los datos; Creative Commons para los contenidos, generalmente productos basados en esos datos–, si bien en última instancia los matices expresan estrategias de mayor y menor grado de disposición a la reutilización o a la preservación del espíritu y la filosofía abiertos.

Atención especial requieren los modelos de participación que se configuran para regular la actividad de los usuarios y de los grupos en las plataformas de consumo y reutilización de datos abiertos. En los casos analizados se observa que este es el punto de mayores divergencias entre proyectos que persiguen los mismos objetivos o fines muy similares. Lo que sí comparten –salvo Get the Data, cuya naturaleza esencial es la de los foros de expertos que abundan en internet– es la orientación profesional que tiñe los servicios que facilitan el trabajo con los *open data* para generar productos –visualizaciones y organización para hacer más amigables y comprensibles los datos–: aunque lo que se declara es que cualquier usuario puede hacerlo, y se proporciona documentación suficiente y con carácter didáctico, en la práctica estas herramientas parecen destinadas fundamentalmente a programadores y a periodistas, por tanto son estos perfiles profesionales los garantes de la construcción colaborativa de información basada en datos en abierto.

De más facilidades se impregnan las herramientas para la redifusión de la información, sean datos o productos basados en ellos: Twitter, Facebook, Google Plus y LinkedIn dominan el panorama de opciones de redifusión del contenido de los portales, y en general es sencillo ‘republicar’, referenciar, compartir... mediante funciones para integrar, referenciar, enlazar, recomendar... en espacios web o plataformas, medios y redes sociales. Hay bases en estas estrategias, no tanto para la construcción colaborativa de productos informativos basados en datos en abierto por los ciudadanos,

como para la construcción colaborativa de una realidad social a través de vías alternativas a la comunicación de masas.

Ahora bien, lo que en cierta medida garantizan estos procesos es un empoderamiento ‘individual’, es decir, el de quienes consumen productos basados en datos en abierto, sea directamente de portales específicos, sea gracias a lo que recibe de las redes sociales. Pero, ¿qué hay del empoderamiento ‘colectivo’? Más concretamente, ¿qué pasa con los que no ‘frecuentan’ estos canales y redes? Quizá el papel de los medios de comunicación de masas siga siendo fundamental para que estos productos lleguen a una masa crítica –en el doble sentido de la palabra– capaz de cambiar –en algunos sistemas, con el voto por ejemplo– el orden o estado de las cosas... No en vano, algunos de los debates que generan actualmente los contenidos generados por los usuarios pivotan sobre la solvencia y credibilidad de las fuentes (Meso, 2013) o sobre el carácter complementario que puedan alcanzar (Reich, 2008).

La creciente demanda social de una ampliación de las esferas de acceso a la información pública (Howe, 2006), ha generado un proceso de apertura y creación de una cultura *open* en numerosas instituciones y organizaciones. Y este proceso, está poniendo en manos de los ciudadanos, pero especialmente de los medios y de la nueva generación de comunicadores, un instrumental insospechado para el ejercicio informativo y de construcción de conocimiento de lo que sucede, pero también de intervención en la esfera pública.

El periodista, como el científico, debe aprovechar la oportunidad que le ofrece este nuevo acceso a la información básica y tomarlo como un valor ético, moral, de contribución social. Un proceso de investigación que, siguiendo la teoría de Merton (1965), debe encaminarse a una constante depuración de errores; la ética científica como soporte de una ética de la comunicación (Himanen, 2002), en una adaptación a las importantes transformaciones que se han introducido en el siglo XXI en la tarea de informar (Ferrerías Rodríguez, 2013: 126) y que se ven ampliadas por esta incorporación de los ciudadanos a este proceso activo.

Esta depuración de errores, junto al libre acceso o la transparencia de la información son los elementos esenciales que conforman ese ‘escepticismo organizado’ que debe fundamentar la acción informadora que surge del tratamiento de los datos. Es preciso un diálogo constante y crítico con estos, con sus contradicciones; pero también con los otros, con el colectivo de informadores y la ciudadanía, pues sus resultados deben conformarse colectivamente y “sus fallos e imperfecciones son detectados y gradualmente depurados mediante la crítica del conjunto de la comunidad” (Himanen, 2002: 88).

Dependerá, pues, de la masa crítica –paradójicamente, de cuán masivo sea el entorno por el que circulan los datos en abierto y los productos de información basados en ellos– el poder hablar de un nuevo y serio competidor en la construcción de la realidad social por los medios de comunicación, un competidor que opondrá, a la figura del consumidor y del votante dirigidos, el modelo del ciudadano *opendata*, informado en plenitud y firmemente decidido a decidir.

*Esta investigación ha sido desarrollada dentro de los siguientes proyectos financiados por las siguientes entidades:

- Proyecto “OpenDataCitizen. Ciudadanía Digital y Open *Data Access*: empoderamiento ciudadano a través de los medios sociales en el entorno digital” ([Referencia CSO2012-30756](#)). Ministerio de Economía y Competitividad de España durante el periodo 2013-2015, dentro del Plan Nacional I+D+i.
- Proyecto “Contenidos digitales basados en datos abiertos y construcción participativa de la realidad social” ([Referencia PRIN13_CSJ18](#)). Universidad Rey Juan Carlos y Oficina de Cooperación Universitaria.

5. Notas

- [1] <http://www.proacceso.org/>
- [2] <http://www.gobiernoabierto.navarra.es/es/open-data/datos/formatos>
- [3] <http://es.creativecommons.org/blog/licencias/>
- [4] <http://opendatacommons.org/licenses/>
- [5] <http://opencorporates.com/>
- [6] <http://openspending.org/>
- [7] <http://opencharities.org/>
- [8] <http://www.opendatalatinoamerica.org/home/>
- [9] <http://africaopendata.org/>
- [10] <http://getthedata.org>
- [11] <http://datahub.io/>
- [12] www.poderopedia.org
- [13] <http://wbi.worldbank.org/>
- [14] <https://scraperwiki.com/>
- [15] <http://openleis.com/>
- [16] <https://opencorporates.com/viz/financial/index.html>
- [17] <http://registries.opencorporates.com/>
- [18] <https://okfn.org/>
- [19] <http://community.openspending.org/>
- [20] <http://openspending.org/datasets>
- [21] <http://openspending.org/datasets/new>
- [22] <http://apps.openspending.org/maps/>
- [23] <http://blog.openspending.org/help/api/>
- [24] <http://ujuzi.org/>
- [25] <http://ckan.org/>
- [26] <http://dondevanmisimpuestos.es/>
- [27] <http://spending.jp/>
- [28] www.mkbudget.org/index
- [29] <http://beta.followthemoney.org/>

6. Bibliografía

Albert Gómez, M.J. (2006). *La investigación educativa. Claves teóricas*. Madrid: McGraw-Hill.

Almiron Roig, N. (2006). Los valores del periodismo en la convergencia digital: *civic journalism* y quinto poder. *Revista Latina de Comunicación Social*, 61 (La Laguna, Tenerife). Recuperado el 14 de junio de 2014, de <http://www.ull.es/publicaciones/latina/200609Almiron.htm>

Álvarez, S.; Gertrudix, M. (2011). Contenidos digitales abiertos y participación en la Sociedad Digital. *Enl@ce. Revista Venezolana de Información, Tecnología y Conocimiento*, año 8 (2), pp. 79-93.

Armentia, J.I. (2008). La evolución del periodismo participativo en Internet en *Estudios de periodística XIV: Periodismo ciudadano, posibilidades y riesgos para el discurso informativo*. Salamanca: X Congreso de la Sociedad Española de Periodística.

Bakker, P.; Pantti, M. (2009). Beyond news: user-generated content on Dutch media websites. *Future of journalism conference*, Cardiff Univ. Recuperado el 09 de junio de 2014 de <http://www.caerdydd.ac.uk/jomec/resources/foj2009/foj2009-Bakker-Pantti.pdf>

Bermejo Berros, J. (2014). Evolución de los paradigmas, metodologías y campos de la comunicación en *Revista Latina de Comunicación Social* durante la década 2004-2013. *Revista Latina de Comunicación Social*, 69, pp. 330 a 353. http://www.revistalatinacs.org/069/paper/1014_UVa/17b.html DOI: 10.4185/RLCS-2014-1014

Briggs, M. (2007). *Periodismo 2.0. Una guía de alfabetización digital para sobrevivir y prosperar en la era de la información*. Austin: Universidad de Texas.

Bizer, C.; Cyganiak, R.; Heath, T. (2011). How to Publish Linked Data on the Web. Universidad de Mannheim. Retrieved June 18, 2013, from <http://wifo5-03.informatik.uni-mannheim.de/bizer/pub/LinkedDataTutorial/>

Borger, M.; Hoof, A.; Costera, I.; Sanders, J. (2013) Constructing participatory journalism as a scholarly object. *Digital Journalism*, 1(1), 117-134, DOI: 10.1080/21670811.2012.740267

Bowman, S.; Willis, C. (2003). *We Media*. Reston: The Media Center.

Cairo, A. (2011). *El arte funcional: Infografía y visualización de información*. Madrid: Alamut.

Cova, B.; Dali, D.; Zwick, D. (2011). Critical perspectives on consumers' role as "producers": Broadening the debate on value co-creation in marketing processes. *Marketing Theory*, 11 (3), p. 231-241.

Deuze, M. (2008). The changing context of news work: Liquid journalism for a monitorial citizenry. *International Journal of Communication*, 2(18).

Domingo, D. (2008). Interactivity in the daily routines of online newsrooms: dealing with an uncomfortable myth. *Journal of Computer-Mediated Communication*, 13(3), pp. 680–704.

Eberwein, T.; Flenger, S.; Lauk, E.; Leppik, T. [Eds] (2011). *Mapping Media Accountability in Europe and Beyond*. Colonia: Halem.

Ferreras Rodríguez, E. M. (2013). Aproximación teórica al perfil profesional del 'Periodista de datos'. *Icono 14*, 11(2), pp. 115-140. DOI: 10.7195/ri14.v11i2.573

Freytas, M. (2009). Miro la televisión, luego existo. *TodosChile*. Recuperado el 18 de mayo de 2014 de <http://xurl.es/rwyju>

Friedman, T. (2005). *La Tierra es plana. Breve historia del mundo globalizado del siglo XXI*. España: mr ediciones.

Fondevila Gascón, J. F. (2013). Periodismo ciudadano y *cloud journalism*: un flujo necesario en la sociedad de la banda ancha. *Revista Comunicación y Hombre*, 9.

García de Madariaga, J.M. (2006). Del periodismo cívico al participativo: nuevos medios, viejas inquietudes. *Revista ZER*.

García, F.; Gertrudix, M. (2011). Naturaleza y características de los servicios y los contenidos digitales abiertos. *CIC Cuadernos de Información y Comunicación*, 16, pp. 115-138.

Gillmor, D. (2004). *We the Media*. California: O'Reilly Media.

Gillmor, D. (2010). *Mediactive*. California: Dan Gillmor.

Gracia, B.; Capón, J. (2004). Las bitácoras o weblogs y la lógica del campo informativo. Un análisis comparativo con la agenda mediática tradicional. *Estudios de mensaje periodístico*, 10, pp. 113-128.

Gómez, J.M.; Méndez-Muros, S. (2013). El periodismo de cercanía en el bien común de la humanidad” en Diezhandino, M.P.; Sandoval, M.T. *Sociedad Española de Periodística, XVIII Congreso Internacional: Los nuevos desafíos del oficio del periodismo*.

ISO (2011). *ISO in brief: International Standards for a sustainable world*. Ginebra: ISO Central Secretariat.

Himanen, P. (2002). *La ética del hacker y el espíritu de la era de la información*. Madrid: Destino.

Howard, A. (2012). *Data for the Public Good*. California: O'Reilly Media.

Howe, J. (2006). The Rise of Crowdsourcing. *Wired*, 14(6).

Humphreys, A.; Grayson, K. (2008). The Intersecting Roles of Consumer and Producer: A Critical Perspective on Co-production, Co-creation and Prosumption. *Sociology Compass*, 2(3), pp. 963-980.

Hurrell, S. (2012). Is the role of journalists as gatekeepers of information being negated by the growth of blogging and citizen journalism? *Media Context Essays*. Recuperado el 08 de junio de 2014 de <http://stephaniehurrell.files.wordpress.com/2012/05/mc-essay.pdf>

Kelty, C.M. (2008). *Two Bits. The Cultural Significance of Free Software*. Durham: Duke University Press.

Lacy, S.; Duffy, M.; Riffe, D.; Thorson, E.; Fleming, K. (2010). Citizen journalism web sites complement newspapers. *Newspaper Research Journal*, 31(2), pp. 34-46.

López, X. (2009). Nuevos informadores para un periodismo más dialogante. En VV.AA *Periodismo ciudadano: posibilidades y riesgos para el discurso informativo*. X Congreso de la Sociedad Española de Periodística. Salamanca: Universidad Pontificia de Salamanca

Mauss, M. (1925). Essai sur le Don. Forme et raison de l'échange dans les sociétés archaïques. *L'année sociologique, 1923-1924*, 1.

Malinowski, B. (2001). *Los argonautas del Pacífico Occidental*. Barcelona: Península.

Merton, R. (1965). *Teoría y estructuras sociales*. México DF: Fondo de Cultura Económica.

Meso, K. (2006). *Introducción al ciberperiodismo: Breve acercamiento al estudio del Periodismo en Internet*. Bilbao: Universidad del País Vasco.

Meso Ayerdi, K. (2013). Periodismo y audiencias: inquietudes sobre los contenidos generados por los usuarios. *Cuadernos.info*, 33, 63-73. DOI: 10.7764/cdi.33.515

Niles, R. (2007). A Journalist's Guide to Crowdsourcing. Recuperado el 22 de mayo de 2014 de <http://www.ojr.org/ojr/stories/070731niles/>

Nip, J.Y. (2006). Exploring the second phase of public journalism. *Journalism Studies*, 7(2).

Orihuela, J.L.; Cambroner, A. (2006). *La revolución de los blogs: cuando las bitácoras se convirtieron en el medio de comunicación de la gente*. Madrid: La esfera de los libros.

Ortega, F.; Rodríguez, J. (2011) *El potlach digital. Wikipedia y el triunfo del procomún y el conocimiento compartido*. Madrid: Cátedra.

Pennisi, E. (2009). On the origins of cooperation. *Science*, 325 (5945), pp. 1196-1199.

Polanyi, K. (1989). *La gran transformación. Crítica del liberalismo económico*. Madrid: Ediciones de La Piqueta.

Raymond, E. (1997). *The Cathedral & the Bazaar*. California: O'Reilly Media.

Reich, Z. (2008). How citizens create news stories. *Journalism Studies*, 9(5), pp. 739-758. doi: 10.1080/14616700802207748

Ritzer, G. (1996). *The McDonaldization of Society*. California: Pine Forge Press.

Ritzer, G.; Jurgenson, N. (2010). Production, Consumption, Prosumption The nature of capitalism in the age of the digital 'prosumer'. *Journal of consumer culture*, 10(1), pp. 13-36.

Rosenberry, J.; St John III, B. (2010). *Public journalism 2.0. The promise and reality of a citizen*. Londres: Routledge.

Rost, A. (2010). La participación en el periodismo digital. Muchas preguntas y algunas posibles respuestas. En F. Irigay, D. Ceballos y M. Manna (Coords.). *Periodismo digital en un paradigma en transición* (pp. 96-109). Rosario, Argentina: Universidad Nacional de Rosario.

Sampedro, V. (2009). Periodismo ciudadano, precariedad laboral y depauperación de la esfera pública. En Pérez Herrero, P. Rivas Nieto y Gelado Marcos, R. (Coordinadores) *Estudios de Periodística XIV – posibilidades y riesgos para el discurso informativo*. Salamanca: Ed. Universidad Pontificia de Salamanca.

Shultz, J. (1998). *Reviving the fourth state*. Cambridge: University Press.

Sonvilla-Weiss, S. (2010). Introduction: Mashups, remix practices and the recombination of existing digital content. *Mashup cultures*, 8-23. Viena: Springer.

Takahashi, D. (2010). Gamification gets its own conference. En *Venture Beat*. [Online]. Available: <http://venturebeat.com/2010/09/30/gamification-gets-its-own-conference>.

Tapscott, D.; Williams, A.D. (2007). *Wikinomics. La nueva economía de las multitudes inteligentes*. Barcelona: Paidós.

Toffler, A. (1981). *The third wave*. New York: Bantam books.

Torvalds, L. (1992). Release notes for Linux v.0.12. Recuperado el 12 de junio de 2014 de <https://www.kernel.org/pub/linux/kernel/Historic/old-versions/RELNOTES-0.12>

Wimmer, R. D.; Dominick, J. R. (1996). *La investigación científica de los medios de comunicación, una introducción a sus métodos*. Barcelona: Bosch.

Artículos relacionados

Aitamurto, T.; Sirkkunen, E.; Lehtonen, P. (2011). Trends in Data Journalism. *Espoo, VTT*.

Henninger, M. (2013). Data-driven journalism. *Reassessing Journalism*, 157.

Huijboom, N.; Van den Broek, T. (2011). Open data: an international comparison of strategies. *European journal of ePractice*, 12(1), pp. 1-13.

Jönsson, A.M.; Örnebring, H. (2011). User-generated Content and the News: empowerment of citizens or interactive illusion? *Journalism Practice*, 5(2), pp. 127-144.

Cómo citar este artículo / Referencia normalizada

S Álvarez García, M Gértudix Barrio, M Rajas Fernández (2014): “La construcción colaborativa de bancos de datos abiertos como instrumento de empoderamiento ciudadano”. *Revista Latina de Comunicación Social*, 69, pp. 661 a 683.

http://www.revistalatinacs.org/069/paper/1029_UR/32es.html

DOI: [10.4185/RLCS-2014-1029](https://doi.org/10.4185/RLCS-2014-1029)

Artículo recibido el 10 de septiembre de 2014. Aceptado el 19 de octubre. Publicado el 31 de octubre.