

Los portales latinoamericanos de datos abiertos: nuevas herramientas para el acceso y la reutilización de la información gubernamental

SILVIA COBO SERRANO
Universidad Complutense de Madrid, España

INTRODUCCIÓN

Cada vez más, las instituciones gubernamentales son conscientes de la importancia social de la apertura de datos y de la disposición de éstos en formatos accesibles, de tal manera que se garantice la transparencia y el derecho de acceso de los ciudadanos a la información gubernamental. En consecuencia, esta apertura, con independencia de la estructura administrativa de cada país, permite la reutilización de los conjuntos de datos públicos, también denominados *datasets*; es decir, facilita “[...] el uso de documentos que obran en poder de organismos del sector público por personas físicas o jurídicas con fines comerciales o no comerciales distintos del propósito inicial que tenían esos documentos en la misión de servicio público para la que se produjeron.”¹ El alcance de esta reutilización fue extendido a bibliotecas (incluidas las universitarias), archivos

y museos europeos, lo que permite crear nuevos servicios a partir de datos públicos con valor añadido en todos los sectores sociales.² Ejemplos de esta reutilización son las numerosas aplicaciones desarrolladas recientemente, tales como conocer en qué se invierte el dinero público, contratación pública, tráfico de los medios de transporte terrestre en tiempo real, horario de los metros, búsqueda de farmacias, alertas sísmicas o incluso la evolución de los nombres de las personas en las últimas décadas.

Por otro lado, la puesta a disposición de los datos gubernamentales en formatos abiertos y por medios electrónicos no ha sido suficiente para satisfacer las necesidades informativas de la ciudadanía, ya que la dispersión de estos conjuntos de datos en los websites institucionales seguía siendo una problemática constante cuya resolución requería de nuevos esfuerzos gubernamentales. En pocos años, este problema se ha resuelto mediante la creación de portales nacionales de datos abiertos; iniciativas públicas de los gobiernos, tanto de los Estados Miembros de la Unión Europea como de América Latina, que permiten la centralización de los activos de información o conjuntos de datos en un único portal. En este sentido, son destacables y paradigmáticos los portales de Reino Unido, Estados Unidos, Australia o Nueva Zelanda así como los estudios comparativos de carácter cualitativo de estos portales de datos abiertos en Europa.

En este sentido, uno de los trabajos más significativos sobre los portales europeos de datos abiertos es la publicación de Ramos Simón *et al.*, en diciembre de 2012, cuyo objetivo fundamental era “[...] identificar los portales *open data* puestos en marcha por los gobiernos de los países miembros de la Unión Europea, describir las licencias de acceso y uso de la información pública y analizar las ca-

racterísticas de los recursos de información desde el punto de vista técnico.”³

En el ámbito latinoamericano, son destacables las iniciativas *Open Data Latinoamérica*⁴ y *Open Data for Development in Latin America and the Caribbean*.⁵ La primera de ellas es un catálogo de datos que permite el acceso a los datasets mediante la utilización del propio catálogo, el servicio de búsquedas y el uso de las colecciones destacadas para, posteriormente, poder reutilizarlos a partir de la explotación de los mismos. Respecto al proyecto OD4D –encabezado por el W3C Brasil y la Comisión Económica para América Latina y el Caribe (CEPAL), y apoyado económicamente por el Centro Internacional de Investigación y Desarrollo de Canadá (IDRC)–, la iniciativa pretende la sensibilización social sobre los beneficios e importancia de los datos abiertos, así como apoyar la “[...] apertura de datos sustentables y durables.” Finalmente, es destacable la celebración, en junio de 2013, de la *Conferencia Regional de Datos Abiertos para América Latina y el Caribe*, en la que se abordaron cuestiones fundamentales como el valor social y económico de los datos públicos, los desafíos gubernamentales, aspectos técnicos, periodismo de datos y la apertura de datos en el ámbito educativo, sanitario y medioambiental, entre otros.

A tenor de lo anteriormente mencionado, el objetivo del presente trabajo es analizar los portales nacionales de datos abiertos más significativos o representativos en América Latina de forma genérica, de tal manera que se pueda determinar el estado o situación actual de los mismos a partir de una serie de parámetros de estudio.

METODOLOGÍA

Dada la extensión geográfica de América Latina y en esta primera aproximación a la temática, en primer lugar se procedió a la identificación de los países con portales de datos abiertos (es decir, nuevas herramientas al servicio de la sociedad para favorecer la transparencia gubernamental y el acceso a datos útiles para la actividad profesional, así como para el desarrollo de nuevas aplicaciones) a partir de la búsqueda de estos nuevos instrumentos en 22 países de Latinoamérica: Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, Haití, Honduras, Jamaica, México, Nicaragua, Panamá, Paraguay, Perú, Puerto Rico, República Dominicana, Uruguay y, finalmente, Venezuela.

Los resultados de esta búsqueda inicial permitieron conocer la existencia de numerosos portales de datos abiertos tanto a nivel nacional como estatal o provincial (administración pública de nivel inferior); es decir, muchos de los gobiernos de los distintos niveles de las administraciones públicas publican sus datos de forma uniforme y centralizada mediante la inclusión de los conjuntos de datos en estos portales públicos. De todos ellos, se han seleccionado aquellos que, por un lado, responden a los criterios o parámetros de análisis contemplados en este trabajo (volumen de datasets, accesibilidad, descripción y formatos) y, por otro lado, son representativos para cada país y pioneros con versiones beta en estos momentos. En total, se presentan 10 portales latinoamericanos de datos públicos, tal y como se puede observar en el *Cuadro 1*; de esos portales, únicamente los de México y Perú están relacionados con una estructura administrativa federal y metropolitana, respectivamente.

Los portales latinoamericanos de datos abiertos...

Cuadro 1

Presentación de los portales y catálogos de datos abiertos seleccionados en América Latina

PAÍS	PORTALES DE DATOS ABIERTOS
Argentina	Portal de Datos Público ⁶
Brasil	Portal Brasileiro de Dados Abertos ⁷
Chile	Portal de datos públicos datos.gob.cl ⁸
Colombia	Catálogo de Datos Abiertos ⁹
Costa Rica	Portal de Datos Abiertos ¹⁰
Ecuador	datosabiertos.ec ¹¹
El Salvador	Datos Abiertos El Salvador ¹²
México	Datos Abiertos DF ¹³
Perú	Datos Abiertos de Lima Metropolitana ¹⁴
Uruguay	Catálogo de Datos Abiertos ¹⁵

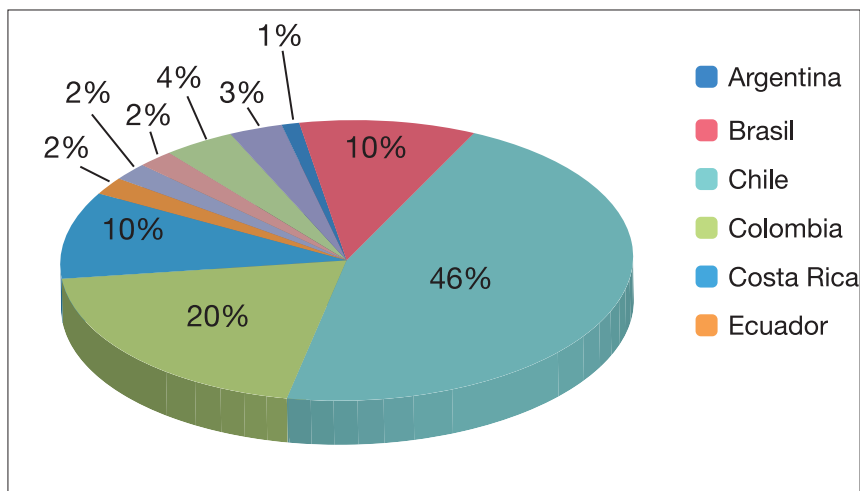
A continuación, se procedió a la identificación de los parámetros de estudio de los portales ya mencionados (volumen de datasets, accesibilidad, descripción y formatos) para, seguidamente, presentar los resultados obtenidos.

VOLUMEN DE DATASETS

El grado de desarrollo de los catálogos de datos abiertos suele estar directamente relacionado con el número de datasets incluidos y campos descriptivos de los mismos. En términos generales, se puede indicar que los catálogos son de reciente creación y su objetivo común es fomentar la transparencia gubernamental, la participación ciudadana y la creación de nuevos servicios tecnológicos a partir de datos públicos. Por ello, son numerosos los países latinoamericanos que forman parte de la Alianza para el Gobierno Abierto (*Open Government Partnership*) desde 2011, iniciativa en la que participan actualmente 63 países del mundo.¹⁶

Respecto al número de datasets incluidos en los catálogos de los portales seleccionados para el presente estudio, éstos se pueden dividir en tres grandes grupos (véase la *Figura 1*). En primer lugar, destaca el conjunto de datos del catálogo chileno, que representa casi el 50% de la muestra al disponer de más de 1.100 datasets. Con bastante diferencia cuantitativa, se identifica un segundo grupo formado por Colombia (aproximadamente 500 datasets), Brasil y Costa Rica (con 240 y 254 conjuntos de datos, respectivamente), que se sitúan entre el 10% y el 20% de la muestra analizada para, finalmente, poder clasificar el tercer grupo de portales cuyos conjuntos de datos se encuentran entre el 1% y el 4% y agrupan a los seis países restantes. En este sentido, se puede decir que son portales cuyos catálogos disponen de escasos datasets, ya que éstos suelen oscilar entre los 20 y los 95, y la suma de todos ellos (seis catálogos) no supera los 350 conjuntos de datos.

Figura 1
Estimación porcentual de los datasets existentes en los catálogos de los portales abiertos de América Latina.



ACCESO Y DESCRIPCIÓN DE LOS DATASETS

Respecto al acceso a la información, en todos los catálogos se utilizan sistemas de búsqueda simple (introducción de términos en lenguaje natural) que pueden ser combinados con la selección de otros filtros de búsqueda, como por ejemplo formatos u organismos. Aunque no todos los filtros que se relacionan a continuación aparecen en todos los portales, se puede indicar que los catálogos de datos abiertos de los portales gubernamentales de América Latina seleccionados para el presente estudio permiten filtrar la búsqueda por: entidad (hace referencia al organismo productor del conjunto de datos; otros términos utilizados para referirse a este filtro son *Autor* u *Organizaciones*), categorías temáticas, etiquetas, formatos, grupos de conjuntos, licencias y recursos (*Figura 2*).

Figura 2

Ejemplo de filtros de búsqueda en el Portal de Datos Abiertos de Costa Rica

Seleccione... Ir

Filtrar recursos

Por entidad
Mostrar todas

Por tipo
Mostrar todos

Por categoría
Mostrar todas

Aplicar filtros

Titulo

Bares y Restaurantes en Palmares	Municipalidad de Palmar
Bares y Restaurantes en Palmares	Municipalidad de Palmar
Supermercados y Licorerías en Palmares	Municipalidad de Palmar
Supermercados y Licorerías en Palmares	Municipalidad de Palmar
Stios para hospedarse en Palmares	Municipalidad de Palmar
Gasolineras en Palmares	Municipalidad de Palmares Eco
Gasolineras en Palmares	Municipalidad de Palmares Eco
Entidades Financieras en Palmares	Municipalidad de Palmar
Entidades Financieras en Palmares	Municipalidad de Palmar
Farmacias de Palmares	Municipalidad de Palmares Gestio
Farmacias de Palmares	Municipalidad de Palmares Gestio
Restaurantes de comida rapida en Palmares	Municipalidad de Palmar

Fuente: <http://datosabiertos.gob.go.cr>

La información gubernamental y el acceso a la información...

En lo que a campos descriptivos de los registros se refiere, existe cierta variación entre los mismos, si bien es cierto que los portales de Argentina, Ecuador y Uruguay presentan grandes similitudes en la estructura y, por tanto, en los campos de descripción, ya que utilizan la misma plataforma tecnológica para la creación de sus portales de datos abiertos. Esta similitud, aunque con otra plataforma tecnológica, también se presenta en los catálogos de Costa Rica y Perú.

Con independencia de esta circunstancia y siendo conscientes de la variación numérica de los campos descriptivos de los catálogos, se puede indicar que los campos con mayor frecuencia de aparición en los registros de los catálogos son:

- Título.
- Descripción.
- Autor / Entidad / Institución.
- Mantenedor (normalmente, suele ser el autor).
- Etiqueta.
- Categoría.
- Datos y recursos para descargar.
- Url.
- Fecha de publicación.
- Fecha de última actualización.
- Alcance geográfico / Cobertura espacial.
- Alcance temporal / Cobertura temporal.
- Licencia.

Con menor frecuencia de aparición, otros campos identificados son:

- Documentación asociada.
- Formato fuente (indicación del tipo de archivo desde el que se han recolectado los datos).
- Visitas al datasets.

FORMATOS

Otro de los parámetros de análisis fue el formato de los archivos. En este sentido, el estudio pretende la identificación de los formatos disponibles en los diez portales latinoamericanos de datos abiertos, ya que un análisis de la adecuación de los mismos para la descarga y posterior reutilización de los conjuntos de datos requeriría un estudio de mayor profundidad, el cual no es objeto de este trabajo introductorio a la materia.

En consecuencia, se han identificado (aproximadamente) 25 formatos durante el análisis de los diez catálogos de datos abiertos, entre los que destacan los formatos de carácter ofimático, web y de compresión. Sin lugar a dudas, el formato CSV está presente en los diez catálogos latinoamericanos y es seguido de cerca por el formato de documentos ofimáticos XLS / XLSX; esto es un 70% del total de la muestra analizada. En posiciones inferiores se encuentran los formatos HTML, PDF, ZIP, otros textuales como DOC o TXT, y también formatos para archivos de datos y lenguajes de marcado. En el *Cuadro 2*, se presenta el ranking de los cinco formatos más frecuentes en los catálogos de datos analizados.

Cuadro 2
Top 5 de los formatos más frecuentes en los catálogos de datos

Posición	Formatos	% de aparición en los portales
1	CSV	100%
2	XLS (XLSX)	70%
3	HTML / XML / PDF	40%
4	ZIP	30%
5	DOC (DOC) / TXT / SAV / KML / RAR / JSON / DBF	20%

DISCUSIÓN

En primer lugar, se observa el compromiso de los gobiernos latinoamericanos para fomentar la transparencia y la apertura de datos mediante la creación de portales y la participación en la Alianza del Gobierno Abierto, iniciativa de la que forman parte países latinoamericanos tales como México, Brasil, Colombia, Perú, Chile, Argentina y Uruguay, entre otros.

Asimismo, sería recomendable reducir el volumen desigual de los datasets en los catálogos de los portales con la intención de garantizar el equilibrio entre estas herramientas latinoamericanas de acceso a la información pública.

En general, las búsquedas simples pueden generar ruido y poca precisión en los resultados obtenidos, aunque ciertamente la utilización de varios filtros puede resolver esta problemática, ya que funcionan como una búsqueda avanzada. En consecuencia, se ha de señalar esta circunstancia como una fortaleza que requiere de la actualización permanente de dichos filtros.

Finalmente, se potencia el periodismo de datos a partir de fuentes gubernamentales a un *clic* del ratón y la creación de nuevas aplicaciones tecnológicas que inciden en el bienestar social.

AGRADECIMIENTOS

Al proyecto “Buenas prácticas en el acceso a la información gubernamental”, PAPIIT IN 403113, apoyado por la Dirección General de Asuntos del Personal Académico de la UNAM; al proyecto de investigación “Organización del acceso, uso y reutilización de la información del sector público en España. Hacia la consolidación de una industria de la infor-

mación”, financiado por Plan Nacional de I+D en España (Ref.: CSO2010-17451); al Ministerio de Educación, Cultura y Deporte (España) en apoyo al programa “Formación del Profesorado Universitario”.

REFERENCIAS BIBLIOGRÁFICAS

1. Directiva 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información del sector público (*Diario Oficial de la Unión Europea* 31.12.2003). [s.p.] [en línea], <http://eurex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:ES:PDF>
2. Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013 por la que se modifica la Directiva 2003/98/CE, relativa a la reutilización de la información del sector público (*Diario Oficial de la Unión Europea* 27.6.2013) [en línea], <http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:ES:PDF>
3. L. F. Ramos Simón, R. Arquero Avilés, I Botezan, S. Cobo Serrano, A. Sala Jiménez, R. Sánchez Jiménez, F. del Valle Gastaminza (2012), “De la reutilización de información del sector público a los portales de datos abiertos en Europa”, en *BID: textos universitarios de biblioteconomía i documentació*, núm. 29. [s.p.] [en línea], <http://www.ub.edu/bid/29/ramos2.htm>
4. Open Data Latinoamérica [en línea], <http://www.opendatalatinoamerica.org/>
5. Open Data for Development in Latin America and the Caribbean [en línea], <http://www.od4d.org>
6. Argentina. Portal de Datos Públicos [en línea], <http://datospublicos.gob.ar>

La información gubernamental y el acceso a la información...

7. Brasil. Portal Brasileiro de Dados Abertos [en línea], <http://dados.gov.br>
8. Chile. Portal de Datos Públicos “datos.gob.cl” [en línea], <http://datos.gob.cl>
9. Colombia. Catálogo de Datos Abiertos [en línea], www.datos.gov.co
10. Costa Rica. Portal de Datos Abiertos [en línea], <http://datosabiertos.gob.go.cr>
11. Ecuador. datosabiertos.ec [en línea], <http://datosabiertos.ec>
12. El Salvador. Gobierno Abierto [en línea], <http://www.gobiernoabierto.gob.sv/>
13. México. Datos Abiertos D.F. [en línea], <http://datosabiertos.df.gob.mx>
14. Perú. Datos Abiertos de Lima Metropolitana [en línea], <http://lima.datosabiertos.pe>
15. Uruguay. Catálogo de Datos Abiertos [en línea], <https://catalogodatos.gub.uy>
16. Open Government Partnership [en línea], <http://www.opengovpartnership.org>