

2016

# Apertura y Publicación de Datos Abiertos de Panamá


*<https://www.datosabiertos.gob.pa>*



## ***De datos públicos a datos abiertos***

*Los organismos estatales recogen, producen, reproducen y difunden datos para cumplir con su función pública. Algunos de estos datos ya están publicados o debieran estarlo. Incorporar la publicación de **datos públicos en formatos abiertos**, abre la puerta a la posibilidad que los mismos sean reutilizados en nuevos proyectos, que puedan combinarse con otras fuentes de datos y generar nuevas aplicaciones desarrolladas por el gobierno, la sociedad civil, organizaciones, empresa privada y ciudadanos en general.*

*Esta guía brinda una introducción al tema de datos abiertos. La misma reúne conceptos básicos de datos abiertos de gobierno, los pasos a seguir para abrir y publicar los datos. También se ofrecen referencias a contenidos más específicos para el lector que quiera profundizar en el tema.*



***“Más allá de la simplificación de los trámites – gobierno electrónico, y la transparencia – apertura parcial en cumplimiento de la ley, los datos abiertos reflejan una efectiva visibilidad y rendición de cuentas institucionales, abriendo todo un universo a los ciudadanos, ávidos de aprovechar esta oportunidad para crear nuevos productos de información y aplicaciones en favor de la sociedad”.***

**BUENAS PRÁCTICAS** Versión 1.0 – 2016

*Este documento ha sido adaptado (2016) por la Autoridad Nacional para la Innovación Gubernamental (AIG) de la República de Panamá, con el respaldo de la Agencia para el Desarrollo del Gobierno de Gestión Electrónica y la Sociedad de la Información y el Conocimiento (AGESIC), de la República Oriental de Uruguay.*

*Usted es libre de copiar, distribuir, comunicar y difundir públicamente este documento así como hacer obras derivadas, siempre y cuando tengan en cuenta citar la obra de forma específica y no utilizar esta obra para fines comerciales. Toda obra derivada de esta deberá ser generada con estas mismas condiciones.*

# CONTENIDO

INTRODUCCIÓN.....	3
I. POTENCIAL Y BENEFICIOS DE LOS DATOS ABIERTOS DE GOBIERNO .....	4
II. CARACTERÍSTICAS DE LOS DATOS ABIERTOS <sup>1</sup> .....	4
III. PLAN INSTITUCIONAL COMO ESTRATEGIA PARA LA APERTURA DE DATOS.....	5
A. Delegación del líder institucional de datos abierto .....	5
<i>Funciones del líder de datos abiertos</i> .....	5
B. Conformación del equipo de trabajo .....	5
<i>Funciones del equipo de trabajo</i> .....	6
C. Generación del inventario de datos institucional.....	6
D. Publicación de los datos abiertos de la institución .....	6
1. Datos de gobierno que pueden ser publicados .....	6
2. Formatos utilizados para publicar datos abiertos .....	7
3. Convertir un archivo a formato abierto.....	9
-Desde un archivo en Excel (.XLS a .CSV) .....	9
-Desde un gestor de bases de datos .....	11
4. Aplicar una licencia abierta a los conjuntos de datos a publicar .....	13
5. Incluir los metadatos de los conjuntos de datos .....	14
6. Asignar URL o URI a los conjuntos de datos de 4 y 5 estrellas .....	14
7. Subir los datos abiertos al sitio institucional y portal nacional .....	15
a. Crear conjuntos de datos – para introducir los metadatos del conjunto de datos .....	16
b. Agregar datos - para subir el fichero o enlace del conjunto de datos al portal.....	17
8. Programar actualizaciones y liberación de nuevos conjuntos de datos .....	17
IV. DIVULGAR LA PUBLICACIÓN DE DATOS ABIERTOS .....	18
V. GLOSARIO DE TÉRMINOS .....	19

## INTRODUCCIÓN

Las instituciones públicas del estado recopilan y producen grandes volúmenes de datos relacionados con variados temas: energía, clima, economía, salud, medio ambiente, agricultura, defensa, seguridad pública, sociedad, cultura, presupuestos, entre muchos otros.

El hecho de que los ciudadanos puedan acceder a todos los datos de las instituciones públicas en forma simple y con formatos diferentes impacta positivamente en la sociedad y tiene entre otras, dos ventajas importantes: la primera es que perfecciona la democracia debido a los mayores niveles de participación ciudadana en el quehacer gubernamental; la segunda es que influye positivamente en la economía al permitir la reutilización y redistribución de los datos estimulando la innovación y el emprendimiento, para crear obras derivadas de información dirigidas a la comunidad a partir de los datos abiertos.

*“Los datos abiertos son aquellos que se publican en formato primario, que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier ciudadano sin restricciones, y que se encuentran sujetos, cuando más, al requerimiento de reconocimiento del origen y de compartirlos de la misma manera en que aparecen”.*

La apertura y publicación de datos abiertos requiere de una organización especial en las instituciones gubernamentales, por la necesidad de establecer las directrices generales que lleven a formalizar estas actividades de publicación de datos abiertos en el sitio Web de la institución y el portal nacional de datos abiertos.

Este documento propone un mecanismo para organizar, dentro de la institución, los componentes que participan en los procesos de apertura de datos, además de los procedimientos para la identificación, tratamiento y publicación de los datos abiertos de gobierno.

Como componente final y de gran importancia está la difusión de la liberación de datos en formato abierto. Al respecto, se propone una serie de actividades de divulgación acerca de la apertura de datos en la institución y su colocación en el sitio Web institucional y en el portal nacional.

## I. POTENCIAL Y BENEFICIOS DE LOS DATOS ABIERTOS DE GOBIERNO

Los datos generados por las instituciones públicas al ser puestos a disposición de la sociedad en calidad de Datos Abiertos, incrementan su potencial de uso, de reutilización y de redistribución, generando impacto en el sentido de:

- ❖ Fomentar la participación ciudadana y fortalecer la democracia;
- ❖ Incrementar la transparencia y rendición de cuentas;
- ❖ Impulsar el crecimiento económico, fortalecer la competitividad y promover la innovación;
- ❖ Estimular una mayor eficiencia gubernamental, lo que obliga a diseñar políticas públicas en base a las necesidades de la comunidad y a mejorar los servicios públicos, en apoyo a los objetivos de desarrollo, generación de conocimiento y buena gobernanza;
- ❖ Emitir datos en el marco del respeto a la privacidad y protección de la información personal;
- ❖ Permitir la interoperabilidad entre organizaciones y sistemas. La habilidad de integrar componentes es esencial para construir sistemas complejos, grandes y ricos en información.
- ❖ Fomentar la investigación y el periodismo de datos, lo que permite recabar y analizar grandes cantidades de datos mediante software especializado para hacer comprensible la información a la comunidad a través de artículos, infografías, y visualizaciones de datos, entre otros.

## II. CARACTERÍSTICAS DE LOS DATOS ABIERTOS <sup>1</sup>

1. **Accesibles:** Disponibles en línea y de fácil acceso;
2. **Integrales:** Deberán ser completos y sin procesar, en la medida de lo posible deben contener el tema que describen en detalle y con los metadatos necesarios;
3. **Gratuitos:** Se obtendrán sin entregar a cambio contraprestación alguna, tomando en cuenta lo establecido en la normativa nacional de protección de datos;
4. **No discriminatorios:** Serán accesibles sin restricciones de acceso para los usuarios;
5. **De libre uso:** las entidades indicarán la licencia más apropiada que aplique a los datos liberados, como ejemplo la licencia CC BY 4.0, que requiere citar la fuente de origen como único requisito;
6. **Legibles por máquinas:** Deberán estar estructurados, total o parcialmente, para ser procesados e interpretados por equipos electrónicos de manera automática;
7. **Primarios:** Provenirán de la fuente de origen con el máximo nivel de granularidad posible, no modificados y sin procesar;
8. **Oportunos:** Serán actualizados periódicamente, conforme se generen o cambien;
9. **Permanentes:** Se deberán conservar en el tiempo, para lo cual, las versiones históricas relevantes para uso público, se mantendrán disponibles a través de identificadores adecuados para tal efecto.

---

<sup>1</sup> <http://www.opengovdata.org/home/8principles>

### III. PLAN INSTITUCIONAL COMO ESTRATEGIA PARA LA APERTURA DE DATOS

#### A. Delegación del líder institucional de datos abiertos

Es necesario en primera instancia designar al *líder de datos abiertos*, quien será el responsable institucional de la apertura de datos. La designación, como norma, es realizada por la autoridad de mayor jerarquía en la institución. Designar un líder es necesario debido a la organización a desarrollar para la apertura de datos y su posterior seguimiento. Adicionalmente, el líder de datos abiertos deberá ser el enlace con otras instituciones para coordinar acciones conjuntas de apertura de datos.

##### *Funciones del líder de datos abiertos*

1. Convocar a los actores de su institución que tomarán parte en la apertura de datos y conformar el equipo de trabajo;
2. Coordinar acciones con los integrantes del equipo de trabajo y vigilar la gestión del mismo en la correcta implementación de las directrices para la apertura y publicación de datos;
3. Conformar el plan, junto con el equipo de trabajo, de la apertura de datos, definiendo las actividades y tareas, igualmente, delegando funciones a los miembros del equipo;
4. Velar por la aplicación de la norma legal vigente en la publicación de datos personales, confidenciales o con derecho de autor (operación de filtrado);
5. Implementar los procesos para que los datos se publiquen en el sitio institucional y en el portal nacional, y mantenerlos actualizados con la periodicidad programada;
6. Ser el enlace institucional con la Autoridad Nacional de Innovación Gubernamental para la coordinación y actualización del portal nacional de datos abiertos <https://www.datosabiertos.gob.pa>;
7. Llevar el control del cumplimiento del plan institucional para la apertura de datos, de acuerdo al avance de las actividades y tareas realizadas por el grupo de trabajo.

#### B. Conformación del equipo de trabajo

El equipo de trabajo actuará como un consejo institucional para implementar la política de datos abiertos en toda la institución.

##### *El equipo deberá estar formado como mínimo por:*

1. El líder de datos abiertos;
2. La Dirección de Informática;
3. Los departamentos, unidades, proyectos y áreas que generan y/o administran información;
4. El asesor legal como apoyo para analizar, asesorar y orientar sobre los datos publicados, en el marco de la reserva legal de los documentos que la tienen.



### ***Funciones del equipo de trabajo***

1. Acordar compromisos para el desarrollo de las fases de apertura de datos con todas las áreas involucradas de la entidad;
2. Planificar las actividades, tareas y realizar las estimaciones de tiempo necesarias para las mismas;
3. Dirigir y dar seguimiento a la publicación de datos abiertos en el sitio web institucional y portal nacional;
4. Publicitar la apertura de datos, de tal forma que esto sea conocido por la mayor cantidad de personas en la comunidad;

### **C. Generación del inventario de datos institucional**

Es importante ordenar la información con la que cuenta la institución. Para esto se crea el inventario de datos institucional incluyendo toda la información que se produce a partir de las bases de datos y otros sistemas de manejo de datos con información estadística, financiera, investigativa, administrativa, legal, de geolocalización, de proyectos, entre otros. El líder de datos abiertos solicitará a los responsables de generar y/o administrar datos, el llenado de la plantilla del inventario de datos institucional para todos los conjuntos de datos o servicios de información que generan y/o administran.

El inventario de datos será conformado de manera inicial con el insumo que proporcionen las unidades administrativas y áreas responsables de generar y/o administrar datos y posteriormente se irán añadiendo conjuntos de datos nuevos. El equipo se encargará entonces de consolidar la información de las plantillas en un solo listado.

#### ***Plantilla para el inventario de datos institucional***

Responsable del conjunto de datos o servicio de información	Nombre del conjunto de datos o servicio de información	Temas que lo componen	Descripción del servicio o conjunto de datos	Contiene datos privados, confidenciales o de seguridad: - describir	Plataforma, tecnología, aplicación o programa en que están los datos	Período de actualización del conjunto de datos o servicio de información	Estado del conjunto de datos (revisados, depurados, filtrados)
---	--	-----------------------	--	--	--	--	--

### **D. Publicación de los datos abiertos de la institución**

Una vez que el inventario esté listo el siguiente paso es decidir en qué orden serán publicados los datos abiertos, tanto en el sitio web de la institución, como en el portal nacional, teniendo como guía el potencial de los mismos en generar impacto en la comunidad. Los datos que generan gran impacto son aquellos que tienen mayor valor para la comunidad.

#### **1. Datos de gobierno que pueden ser publicados**

Los datos públicos de gobierno que pueden ser abiertos, son aquellos que se definen con el término “Información” en la ley de acceso a la información pública (ley 6 de 2002): “Todo tipo de datos contenidos en cualquier medio, documento o registro impreso, óptico, electrónico, químico, físico o biológico”. El punto clave a destacar es que la apertura de datos

no se centra en datos personales ni de seguridad nacional, con derecho de autor, o sujetos a restricciones específicas. Como ejemplo tenemos el presupuesto público, ejecución del presupuesto, estadísticas, datos del censo, compras públicas, desempeño/eficiencia de los servicios, datos del transporte, datos judiciales, policivos, ambiente, energía, comercio e industria, y otros.

Empezando con lo más sencillo, seleccione datos que ya tiene y que son fáciles de publicar asegurándose que los mismos tienen alta demanda dentro de la comunidad. Prepare series de conjuntos de datos, por ejemplo series anuales, de tal forma que se puedan bajar datos completos masivamente para los análisis.

Al utilizar muchos formatos abiertos, se amplía la gama de usuarios que pueden aprovechar los datos, ya que estos formatos permiten que sean interpretados por la mayoría de los sistemas computacionales. Es altamente recomendable que los datos se publiquen en formatos estructurados como el CSV. Sin embargo, para dar respuesta a toda esa gama de usuarios, sectores y casos de uso, se debe considerar la publicación de los mismos conjuntos en diferentes formatos estructurados de estándar abierto convenientes al tipo de usuario, como lo son XML, JSON, RDF, GeoJSON, KML, GML, RSS, WMS, HTML, DBF, incluyendo formatos como SHP y XLSX/ODS.

Los conjuntos de datos deben estar disponibles para descarga directa con sus URLs únicos y sin registro de acceso mediante contraseñas o interfaces de usuario, ya que estos datos serán procesados por computadoras, no sólo por personas. Para lo anterior, se recomienda cargar los conjuntos de datos abiertos en el servidor web de la institución en la sección destinada a mostrar la información institucional de datos abiertos, para que paulatinamente se vayan añadiendo a los datos que están en formato tradicional. Para facilitar las descargas de los datos abiertos, los archivos muy grandes se deben segmentar y comprimir (utilizando formatos abiertos de compresión como gzip). Los URLs de los conjuntos de datos publicados en el sitio Web de la institución, serán utilizados posteriormente para referenciar estos conjuntos de datos en el Portal de Datos Abiertos <https://www.datosabiertos.gob.pa>

## 2. Formatos utilizados para publicar datos abiertos

Tim Berners-Lee<sup>2</sup> propuso una categorización del grado de apertura de los datos basado en estrellas, refiriéndose a qué tan abiertos y usables son los datos que puede ofrecer una institución:

**1 estrella:** Los datos deben estar disponibles en la web en cualquier formato y bajo una licencia abierta para que sean considerados como datos abiertos. Cualquier archivo disponible en formato .PDF, .DOC, .ODT, .TXT, entre otros, y que esté bajo una licencia abierta se considera dato abierto de 1 estrella.

**2 estrellas:** Los datos deben estar estructurados y en un formato que pueda ser interpretado o procesado por máquinas; deberán estar disponibles en la Web y tener una licencia abierta de uso. Como ejemplo están los datos en formato .XLS, XLSX, organizados en una tabla, sin contraseña.

---

<sup>2</sup><http://5stardata.info/es/>



**3 estrellas:** Similar al anterior pero en un formato no propietario. El nivel 3 estrellas implica los requerimientos de los niveles 1 y 2 menos los formatos, que en este caso deben ser CSV o XML. En estos casos se pueden manipular los datos de cualquier forma sin limitación de características o de uso de algún tipo de software en particular para ello. Son fáciles de publicar.

**4 estrellas:** El conjunto de datos abiertos debe cumplir con las condiciones de las estrellas previas y además utilizar los estándares recomendados por W3C que son los formatos RDF y SPARQL. En este nivel es necesario usar URIs (que es una dirección web de un dato que sirve para enlazarlo con otros datos) para identificar cosas y propiedades, de manera que se pueda apuntar a datos. Los datos se pueden vincular desde cualquier otro sitio (Web o local); se pueden marcar como favoritos.

**5 estrellas:** Son los conjuntos de datos vinculados con los de otras personas, habiéndosele dotado de contexto. En la práctica, es la información en la web que a su vez apunta a otras fuentes de datos. Por ejemplo, si se publica información sobre un libro dentro del conjunto de datos, enlazar descripciones del libro que pertenezcan a otros publicadores de datos abiertos.

Los niveles 4 y 5 son considerados formas avanzadas de publicación de datos en el sentido que no se limitan solo a presentar dato. En la mayoría de los casos, se espera iniciar la liberación de datos de gobierno utilizando como mínimo el nivel de 3 estrellas. Una vez afianzado el proceso de apertura, tanto organizativa como operativamente, se espera que la institución incluya los niveles 4 y 5 estrellas en su gestión de liberación de datos, para posibilitar un uso más amplio de la información y mayores oportunidades de re-uso.

Para profundizar en el tema de formatos, a continuación se listan enlaces con más información:

- **RDF, SPARQL y Linked Data**

Guías y tutoriales.

<http://linkeddata.org/guides-and-tutorials>

<http://www.w3.org/TR/rdf-sparql-query/>

- **Comunidad de Web Semántica**

Uno de los principales lugares para obtener acceso a información y recursos sobre web semántica es el sitio de esta comunidad.

[http://semanticweb.org/wiki/Main\\_Page](http://semanticweb.org/wiki/Main_Page)

- **Para desarrolladores PHP**

Graphite es una librería para PHP construida sobre RC2 y facilita el desarrollo de aplicaciones que utilicen datos en formato RDF. Su nueva funcionalidad de “descripción de recursos” la hace una excelente forma de proveer a un sitio web datos desde SPARQL.

<http://graphite.ecs.soton.ac.uk/>

- **Para desarrolladores Java**

<http://jena.apache.org/>

<http://www.openrdf.org/>

- **Para desarrolladores .NET**

<http://www.brightstardb.com/>

<http://code.google.com/p/linqtordf>

- **Implementaciones SPARQL**

<http://www.w3.org/wiki/SparqlImplementations>

- **Protégé**

Es un editor de ontologías gratuito y de código abierto.

<http://protege.stanford.edu>

Sin importar la clasificación, en todos los casos, el primer paso es seleccionar el conjunto de datos para liberar. Por ello hay que comenzar a identificar los datos que ya tenemos publicados en el sitio web institucional, los que no se publican pero están organizados, o los que son de fácil transformación. Tome en cuenta que los conjuntos de datos deben ser completos, primarios, actualizados y anonimizados. Priorice respondiendo las siguientes preguntas:

- ¿Qué información le solicitan habitualmente y qué exige un procesamiento de datos?
- ¿Qué información existe en su oficina con datos que pueden ser de interés para la población y grupos de interés, generando gran impacto?
- ¿Qué información se brinda habitualmente a la prensa, o se publica en medios gráficos que puede ser reutilizada?
- ¿Qué información se brinda a otros organismos que puede ser de interés general?

Una vez que haya identificado la información candidata para publicar, debe seleccionar la que se encuentre en condiciones de ser publicada. Tenga en cuenta aspectos legales, de completitud, capacidad para mantenerla actualizada, formato, elimine los datos personales si existen (anonimización), etc. Si piensa publicar datos utilizando formatos de 1 estrella (.PDF, .DOC, .ODT, .TXT), cerciórese que no llevan contraseñas contra lectura, impresión, copia, descarga, etc.

### **3. Convertir un archivo a formato abierto**

Si tiene planeado publicar datos con más de 1 estrella, en este punto mostraremos 2 casos en que se transforma el formato del conjunto de datos a publicar. Esto posibilita mejorar la clasificación del dato abierto. Por ejemplo, si se está en 1 o 2 estrellas podría pasar a 2 o 3 estrellas respectivamente.

La conversión se puede hacer desde un archivo o desde un gestor de base de datos; comenzaremos por la primera.

#### ***-Desde un archivo en Excel (.XLS a .CSV)***

A modo de ejemplo tomamos información relacionada con un censo de hogares en formato Excel y realizaremos la transformación a formato abierto CSV.

CUADRO 1								
Población por departamento de residencia habitual, según área y sexo. Total País								
Departamento	Total	Hombres	Mujeres	Ignorado (1)	Área Urbana			
					Total	Hombres	Mujeres	Ignorado (1)
Total	3.286.314	1.577.725	1.708.481	108	3.110.701	1.478.967	1.631.626	108
Montevideo	1.319.108	613.990	705.014	104	1.305.082	604.965	700.013	104
Resto País	1.967.206	963.735	1.003.467	4	1.805.619	874.002	931.613	4

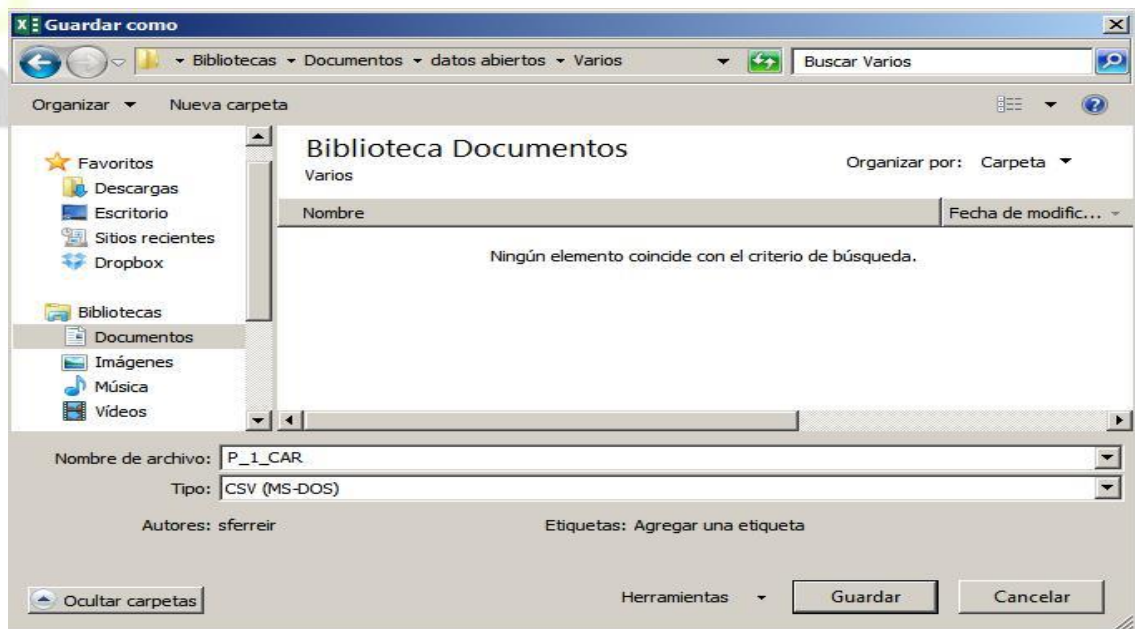
Previo a la conversión es necesario depurar la tabla de todos los elementos que pueden obstaculizar su uso de manera automatizada; estos elementos son las líneas, los sombreados, caracteres en negritas, itálicas, colores, los títulos, encabezados, y cualquier otro tipo de formateo. Dejamos solamente los nombres de las columnas, filas y los datos numéricos.

Departamento	Total	Hombres	Mujeres	Ignorado	Total Urbano	Hombres Urbano	Mujeres Urbano	Ignorado
Total	3.286.314	1.577.725	1.708.481	108	3.110.701	1.478.967	1.631.626	108
Montevideo	1.319.108	613.990	705.014	104	1.305.082	604.965	700.013	104
Resto País	1.967.206	963.735	1.003.467	4	1.805.619	874.002	931.613	4
Artigas	73.378	36.079	37.299	-	69.854	34.102	35.752	-

Luego guardamos el archivo como CSV (MS-DOS) y abrimos el CSV con un editor de texto; veremos lo siguiente:

```
Departamento;Total;;Hombres;Mujeres;Ignorado ;;Total Urbano;Hombres Urbano;
Total;3.286.314;;1.577.725;1.708.481;108;;3.110.701;1.478.967;1.631.626;108
Montevideo;1.319.108;;613.990;705.014;104;;1.305.082;604.965;700.013;104;;1
Resto País;1.967.206;;963.735;1.003.467;4;;1.805.619;874.002;931.613;4;;161
Artigas;73.378;;36.079;37.299;-;;69.854;34.102;35.752;-;;3.524;1.977;1.547;
```

Convertir un archivo de Excel a CSV es muy sencillo (lo mismo se aplica a Open Office o Libre Office), simplemente en la opción “Guardar como” del documento se debe seleccionar el tipo de archivo que queremos. Ver la figura de abajo.



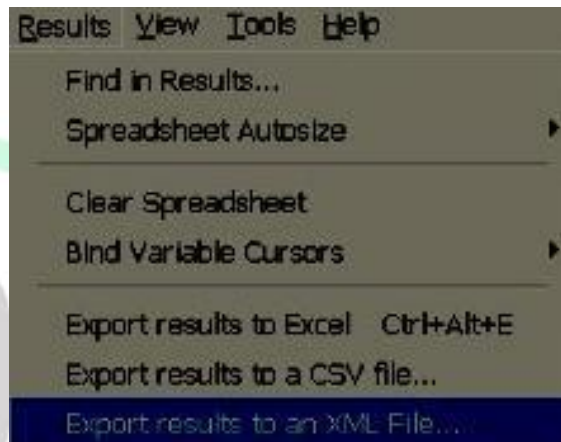
### ***-Desde un gestor de bases de datos***

Para la extracción de datos desde bases de datos al formato XML, se muestran ejemplos.

#### **Ejemplo 1: exportación de datos desde Oracle al formato XML**

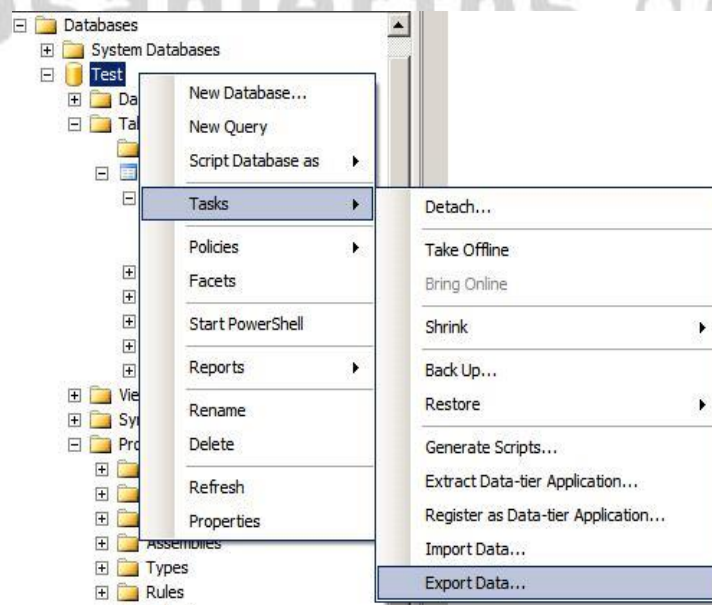
Oracle no cuenta con una interfaz gráfica que permita realizar una exportación de datos a XML directamente, es necesario crear un procedimiento que lea desde una estructura de tabla o vista y genere el documento XML para esos datos.

Sin embargo existen otras herramientas que permiten realizar la exportación, por ejemplo PL/SQL Developer permite realizar la exportación directa de datos a formato XML y CSV desde un menú como muestra la siguiente imagen:



#### **Ejemplo 2: exportación de datos al formato XML mediante MSSQL Server**

MSSQL Server cuenta con un asistente para la exportación de datos muy sencillo de utilizar, permite exportar a diferentes formatos como CSV o Excel. La siguiente es una imagen del menú mencionado:





En su versión 2008 R2 no es posible la exportación directa desde la consola de administración a un formato XML. Para exportar datos a formato XML es necesario realizar cierta programación de manera de lograr el formato deseado. Para ello el desarrollo de procedimientos es la manera indicada donde además se pueden utilizar funciones específicas para la generación de este tipo de documentos.

La versión 2012 de SQL Server provee un comando para la exportación de datos en formato XML que facilita la solución. El comando es el siguiente:

```
bcp tabla_o_vista format nul -f nombre_de_archivo.xml -x
```

### Ejemplo 3: exportación de datos al formato XML mediante MySQL

Utilizando como ejemplo a MySQL podemos fácilmente exportar los datos de una tabla a formato XML válido. Desde la línea de comando se debe ejecutar lo siguiente:

```
>mysqldump --xml test test --user=root --password=mysql
```

Tabla “test” exportada a XML:

Idcampo1	Campo2
1	registro 1
2	registro 2
3	registro 3
4	registro 4

La salida que se obtendrá será la estructura de la tabla, sus campos y sus datos. A continuación se transcribe la salida mencionada.

```
<?xml version="1.0"?>
<mysqldump xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
<database name="test">
<table_structure name="test">
<field Field="idcampo1" Type="int(11)" Null="NO" Key="PRI" Extra="" Comment=""
/>
<field Field="campo2" Type="varchar(50)" Null="NO" Key="" Extra="" Comment="" />
<key Table="test" Non_unique="0" Key_name="PRIMARY" Seq_in_index="1"
Column_name="idcampo1" Collation="A" Cardinality="2" Null=""
Index_type="BTREE" Comment="" Index_comment="" />
<options Name="test" Engine="InnoDB" Version="10" Row_format="Compact" Rows="4"
Avg_row_length="4096" Data_length="16384" Max_data_length="0" Index_length="0"
Data_free="10485760" Create_time="2012-08-22 21:24:11"
Collation="latin1_swedish_ci" Create_options="" Comment="" />
</table_structure>
<table_data name="test">
<row>
<field name="idcampo1">1</field>
<field name="campo2">registro 1</field>
```



```

</row>
<row>
<field name="idcampo1">2</field>
<field name="campo2">registro 2</field>
</row>
<row>
<field name="idcampo1">3</field>
<field name="campo2">registro 3</field>
</row>
<row>
<field name="idcampo1">4</field>
<field name="campo2">registro 4</field>
</row>
</table_data>
</database>
</mysqldump>

```

#### 4. Aplicar una licencia abierta a los conjuntos de datos a publicar

La licencia a aplicar a un conjunto de datos es el documento que expresa formal y explícitamente las condiciones bajo las cuales podrán ser utilizados los datos disponibles en un sitio Web.

El portal **datosabiertos.gob.pa** utilizará la licencia de la organización Creative Commons, de aplicación general conocida como “Reconocimiento 4.0 Internacional (CC BY 4.0)” para los conjuntos de datos contenidos en el Portal Nacional, cargados allí por las instituciones (ver [http://creativecommons.org/licenses/by/4.0/deed.es\\_ES](http://creativecommons.org/licenses/by/4.0/deed.es_ES) y su texto completo en <http://creativecommons.org/licenses/by/4.0/legalcode>). Es posible que alguna institución quiera, independientemente, aplicar otro tipo de licencia al conjunto de datos, como por ejemplo: “De Dominio Público” (ver [https://creativecommons.org/publicdomain/zero/1.0/deed.es\\_ES](https://creativecommons.org/publicdomain/zero/1.0/deed.es_ES) y su texto completo en <https://creativecommons.org/publicdomain/zero/1.0/legalcode>), lo que es igualmente posible.

Parte del texto que contiene la licencia “Reconocimiento 4.0 Internacional (CC BY 4.0)”, es el siguiente:

#### ***Usted es libre de:***

- ***Compartir*** — copiar y redistribuir el material en cualquier medio o formato
- ***Adaptar*** — remezclar, transformar y crear a partir del material para cualquier finalidad, incluso comercial.
- ***El licenciador no puede revocar estas libertades mientras cumpla con los términos de la licencia.***

***Reconocimiento*** — Debe ***reconocer adecuadamente*** la autoría, proporcionar un enlace a la licencia e ***indicar si se han realizado cambios***. Puede hacerlo de cualquier manera razonable, pero no de una manera que sugiera que tiene el apoyo del licenciador o lo recibe por el uso que hace.

## ¿Cómo se indica la licencia en el Portal Nacional de Datos Abiertos?

En el proceso de carga de datos al portal nacional se ofrece una entrada para la licencia. Puede hacer referencia a la licencia para datos abiertos que se encuentra publicada en el sitio <https://www.datosabiertos.gob.pa>. Por otro lado, se puede hacer uso de la licencia que acostumbra especificar en el sitio de su institución donde carga los datos abiertos, si es el caso.

## 5. Incluir los metadatos de los conjuntos de datos

Para facilitarles a los usuarios finales el uso de los conjuntos de datos publicados es necesario hacer una descripción detallada de lo que se publica, mediante los metadatos. Los metadatos hacen una descripción específica del conjunto de datos. Como ejemplo de metadatos tenemos el *nombre del conjunto de datos*, *una descripción*, *etiquetas*, *información legal*, *la licencia usada*, *la cobertura temporal* y *geográfica*, *fecha de publicación*, *última actualización*. Para el caso de estructuras relacionales u hojas de cálculo las columnas pueden ser los metadatos; para los archivos XML se puede considerar el DTD (definición del tipo de documento), etc. Si se publican conjuntos de datos en archivos de formato **.CSV**, se debe suministrar una descripción de las columnas como la que se muestra en la siguiente tabla.

**-Tabla para descripción de campos**

Nombre de atributo	Descripción	Tipo de dato	Recurso relacionado	Información Adicional
<i>&lt;Descripción exacta de la columna tal cual aparece en los archivos de los conjuntos de datos&gt;. <b>Ejemplo:</b> <b>CodigoProv</b></i>	<i>&lt;Descripción detallada del significado del atributo orientada a las personas que pretendan utilizar el atributo&gt;.  <b>Ejemplo:</b> <b>Provincia de Residencia</b></i>	<i>&lt;Numérico, Texto, Fecha/Hora, Moneda&gt; <b>Ejemplo: Numérico</b></i>	<i>&lt;Recurso que da significado al atributo. Ejemplo: si el atributo es un código tiene como recurso relacionado la tabla o recurso donde se describe el significado de dicho código.&gt;  <b>Ejemplo: CódigoProv (código de provincia) apunta al archivo Provincia.xls</b></i>	<i>&lt;Cualquier información adicional sobre el atributo que complemente la información básica.&gt; <b>Ejemplo:</b> En el archivo Provincia.xls están los nombres de las provincias asociadas a dicho código.  <b>Ejemplo:</b> <b>1- Panamá;</b> <b>2-Colón;</b> <b>3-Coclé....</b></i>

## 6. Asignar URL o URI a los conjuntos de datos de 4 y 5 estrellas

Especifique el URL del conjunto de datos. Si es el caso de 4 y 5 estrellas, especifique el nombre del conjunto de datos en la URI de acceso, de modo que el recurso pueda ser identificado y ubicado fácilmente.

### **Ejemplo de URI: usando una hoja de cálculo**

Retomando el ejemplo ya utilizado previamente (Ver: Convertir un archivo a formato abierto) que contiene información relacionada con el censo de hogares en formato Excel, como se muestra en la siguiente imagen:

CUADRO 1								
Población por departamento de residencia habitual, según área y sexo. Total País								
Departamento	Total	Hombres	Mujeres	Ignorado (1)	Área Urbana			
					Total	Hombres	Mujeres	Ignorado (1)
Total	3.286.314	1.577.725	1.708.481	108	3.110.701	1.478.967	1.631.626	108
Montevideo	1.319.108	613.990	705.014	104	1.305.082	604.965	700.013	104
Resto País	1.967.206	963.735	1.003.467	4	1.805.619	874.002	931.613	4

Y que luego le quitamos filas de títulos y resúmenes:

Departamento	Total	Hombres	Mujeres	Ignorado	Total Urbano	Hombres Urbano	Mujeres Urbano	Ignorado
Total	3.286.314	1.577.725	1.708.481	108	3.110.701	1.478.967	1.631.626	108
Montevideo	1.319.108	613.990	705.014	104	1.305.082	604.965	700.013	104
Resto País	1.967.206	963.735	1.003.467	4	1.805.619	874.002	931.613	4
Artigas	73.378	36.079	37.299	-	69.854	34.102	35.752	-

Podemos en este caso, usar lo que era el título de la hoja de cálculo como parte de la URI y de esa manera será posible identificar fácilmente el archivo.

[http://www.miservidordedatos.com/resultadosfinales/cuadros/NIVEL\\_NACIONAL/POBLACION\\_POR\\_DEPARTAMENTO\\_RESIDENCIA/segun\\_area\\_ysexo.zip](http://www.miservidordedatos.com/resultadosfinales/cuadros/NIVEL_NACIONAL/POBLACION_POR_DEPARTAMENTO_RESIDENCIA/segun_area_ysexo.zip)

## 7. Subir los datos abiertos al sitio institucional y portal nacional

Una vez que se han preparado los conjuntos de datos, utilizando las mejores prácticas en el arreglo de los mismos y respondiendo las necesidades de los grupos de interés y ciudadanía en general, ponga sus datos abiertos en línea.

Ubique la información liberada en la sección de datos abiertos del sitio Web de su institución, o en la sección donde expone la información para el público en general. El objetivo es adicionar a la información de formato tradicional los datos abiertos, sumando nuevos conjuntos de datos y siguiendo el programa de liberación de datos. Es importante recordar, que si la institución planea en primera instancia liberar información utilizando archivos que ya están expuestos en el sitio Web en formato PDF, de debe eliminar cualquier contraseña y restricción impuesta para uso del archivo (de impresión, copia de texto o descarga). Esto la colocará en nivel de 1 estrella.

Permita que sus datos puedan ser descargados masivamente, en series enteras de tiempo u otro criterio, y si es posible, también a través de APIs; coloque URI a sus conjuntos de datos publicados de 4 y 5 estrellas.


Una vez que los datos abiertos han sido publicados en el sitio Web de su institución se debe entrar al portal nacional <https://www.datosabiertos.gob.pa> para publicar los nuevos conjuntos de datos.

Para ingresar al portal nacional a cada institución se le facilitará usuario y contraseña a través de la Autoridad Nacional de Innovación Gubernamental AIG. La plataforma sobre la cual está el portal nacional es conocida como “CKAN” diseñada por [Open Knowledge Foundation \(OKF\)](http://www.okfn.org/). CKAN está orientada a proveedores de datos siendo una solución completa de software para hacer los datos accesibles y reutilizables al proveer herramientas para publicar, compartir, encontrar y usar los datos (incluyendo almacenamiento de datos y provisión de APIs de datos robustas). Esta plataforma es muy popular y utilizada en muchas organizaciones de distintos países.

Al cargar los datos abiertos de su institución en el portal nacional se genera el catálogo nacional de datos abiertos, que no es otra cosa que la lista de todos los conjuntos de datos cargados por todas las instituciones, con sus metadatos, licencias, datos y apps disponibles.

Como una introducción al portal nacional de datos abiertos a continuación se presentan dos pantallas:

### a. Crear conjuntos de datos – para introducir los metadatos del conjunto de datos



**DAP** Datos Abiertos De Panamá  
datosabiertos.gob.pa

INICIO DATOS TEMÁTICAS ORGANIZACIONES APPS CONTÁCTENOS ACERCA DE

**Datos / Crear Conjunto de datos**

¿Qué son los conjuntos de datos?

Un Conjunto de Datos de CKAN es una colección de recursos de datos (como ficheros), junto con una descripción y otra información, unida a una URL. Los conjuntos de datos son lo que los usuarios ven cuando buscan un dato.

**1 Crear conjunto de datos** **2 Agregar datos**

**ES Título:** eg. Yellow Submarine  
**\* URL:**

**EN Título:** eg. Yellow Submarine

**ES Descripción:** eg. Alguna nota útil sobre los datos  
Puede usar formato Markdown aquí

**EN Descripción:** eg. Alguna nota útil sobre los datos  
Puede usar formato Markdown aquí

**Etiquetas:**

**Licencia:**

License definitions and additional information can be found at <http://opendefinition.org/>

**\* Organización:**

**Visibilidad:**

**Fuente:**

**Versión:**

**Autor:**

**Email del Autor:**

**Mantenedor:**

**Email del Mantenedor:**

La licencia de datos que seleccionó arriba solo aplica para los contenidos de cualquier archivo de recurso que agregue a este conjunto de datos. Al enviar este formulario, usted está de acuerdo en liberar los valores de metadatos que ingrese en el formulario bajo la Licencia Open Database.

**\* Campo requerido**

**b. Agregar datos - para subir el fichero o enlace del conjunto de datos al portal**

Datos / Crear Conjunto de datos

*¿Qué es un recurso?*

Un recurso puede ser cualquier archivo o enlace a un archivo que contiene datos útiles.

1 Crear conjunto de datos      2 Agregar datos

Fichero:

ES Nombre:

EN Nombre:

ES Descripción:   
Puede usar formato Markdown aquí

EN Descripción:   
Puede usar formato Markdown aquí

Formato:

**8. Programar actualizaciones y liberación de nuevos conjuntos de datos**

Este punto indica el compromiso que adquiere la institución frente a los usuarios de la comunidad. La misma se realiza para fijar un calendario de publicación de datos abiertos, que incluye tanto la actualización de los datos publicados, como la publicación de nuevos conjuntos de datos. Para realizar la programación se puede usar la siguiente plantilla:

Líder Institucional de datos abiertos	Título del líder de Datos Abiertos	Correo oficial del líder de DA	Conjuntos de datos que serán publicados	Descripción de los conjuntos de datos	Periodicidad de actualización o publicación	Fecha tentativa de publicación (DD-MM-AAAA)

Se recomienda que la programación tenga un alcance anual y se publique en el sitio web de la institución para orientar a la comunidad.



## IV. DIVULGAR LA PUBLICACIÓN DE DATOS ABIERTOS

La institución debe divulgar acerca de su publicación de datos abiertos y el programa de actualización de los mismos. Esto se puede realizar con las herramientas que ya existen: el sitio web institucional (página de datos abiertos y sección de noticias, por ejemplo) y el correo electrónico. Es importante que cada vez que sean cargados nuevos conjuntos de datos se promocionen con ambas herramientas. Otras herramientas importantes para difundir la apertura de datos son las redes sociales las cuales pueden ser utilizadas para envío de alertas de liberación de nuevos conjuntos de datos a usuarios seguidores, además de promocionar su uso y re-uso. Las redes sociales son muy útiles para publicar información de forma rápida, y llevar mensajes cortos y puntuales sobre datos publicados recientemente.

Al utilizar el correo electrónico como medio de difusión es recomendable identificar previamente a usuarios y sus áreas de interés, lo que posibilitará estructurar listas de acuerdo a los perfiles de interés de estos. Sin embargo, la promoción a través de estas listas requiere de una planeación e investigación en la comunidad, instituciones de gobierno, empresa privada, organizaciones de la sociedad civil, etc., de tal forma, que se pueda conformar efectivamente las mismas. Los mensajes en los correos deben ser explicativos e incluir contenidos sobre los conjuntos de datos liberados.

Para la obtención de los mejores resultados en la divulgación del uso y re-uso de los datos abiertos, es necesario realizar una planificación y programación de actividades incluyendo tanto a los grupos de interés y comunidad en general, como a las unidades generadoras de los datos abiertos en las instituciones, en la medida de lo posible. Por ejemplo, al realizar actividades de consulta ciudadana se facilita la interacción entre usuarios y generadores de datos abiertos con el objeto de intercambiar ideas, opiniones y sugerencias, lo que redundará en el mejoramiento de los servicios de información por parte de las instituciones de gobierno, al tomar en cuenta de manera directa la opinión de la comunidad.

Otro método que puede ser utilizado para generar y mantener el interés de los grupos y comunidad en general en el uso y re-uso de los datos abiertos de gobierno, es publicar boletines con la descripción de los datos que serán liberados de acuerdo a la programación institucional de apertura de datos. Los eventos especiales y actividades con la comunidad son ocasiones apropiadas para promocionar los datos abiertos. En ellos se pueden hacer presentaciones sobre uso y re-uso de los datos abiertos, como visualizaciones de experiencias en otros países, enseñándoles el valor agregado que se puede sacar de ellos diseñando aplicaciones en temas especializados.

## V. GLOSARIO DE TÉRMINOS

**ANONIMIZACIÓN** - El proceso de adaptación de datos de manera que los individuos no pueden ser identificados a partir de ella.

**API**- (*Application Programming Interface*) Una manera en que los programas informáticos se hablan entre ellos. Puede ser entendido en términos de cómo un programador envía instrucciones entre los programas.

**CSV** – Formato de archivo, donde cada valor es delimitado por una coma o punto y coma; es una forma sencilla de publicar información pero necesariamente se debe publicar también información sobre los datos que contiene (metadatos – información de columnas, etc.). De otra manera esta información será difícilmente interpretada.

**DBF** – Formato de archivos utilizado por dBase, generalmente es el formato del archivo de base de datos para computadores personales.

**DBMS** – Database Management System; sistema de gestión de bases de datos.

**DOC** – Formato de archivo creado por la herramienta de procesamiento de textos de Microsoft.

**DTD** – Describe el formato de los datos (Data Type Definition) en archivos XML, HTML, etc.

**JSON** – JavaScript Object Notation es un lenguaje muy sencillo de automatizar su procesamiento. Si bien no está tan expandido como XML su filosofía similar y su fácil interpretación lo hacen un lenguaje muy apropiado para el intercambio de información.

**KML** – Keyhole Markup Language, formato de archivo utilizado para la representación de datos geográficos en tres dimensiones, basado en XML.

**LINKED DATA (Datos Relacionados)** – Forma parte de la web semántica, a diferencia de la web actual donde los documentos de hipertexto se relacionan a diferentes puntos en el HTML, esta nueva forma enlaza los datos con cosas arbitrarias que se describen en formato RDF.

**METADATOS** – Datos que describen otros datos. En este caso los metadatos describen el conjunto de datos publicados en el portal nacional en detalle.

**ODS** – Formato de archivo de Open Document, es un estándar abierto para documentos de hojas de cálculo.

**ONTOLOGÍA** – En ciencias de la información y ciencias de la computación se entiende por ontología a la representación del conocimiento como un conjunto de conceptos pertenecientes al mismo dominio y las relaciones entre estos conceptos.

**OWL** – Es una recomendación de W3C y fue diseñado para proveer un formato común para el procesamiento de la información en la web, es un lenguaje para ser procesado por computadoras y no está diseñado para mostrar información a los usuarios. Es parte de la web semántica.

**PDF** – Formato portable de documentos, estándar para la publicación de documentos en la web.

**RDF** – Es un marco para la descripción de recursos en la web. El formato RDF permite mediante a un documento XML describir recursos publicados en Internet.

**SHP** – Formato de archivo informático de datos espaciales desarrollado por la empresa ESRI comercializadora de productos ArcGIS y ArcInfo.

**SPARQL** – Motor de consultas para la web semántica.

**TXT** – Formato de archivo de texto plano.

**URI** – Identificador único de recurso, permite identificar únicamente un recurso, pagina, documento, dirección de correo electrónico, etc.

**VOCABULARIOS**– Es un conjunto de términos acordados y agrupados en un espacio de nombres que sirven para describir cosas.

**XLS** – Formato de archivo creado por Microsoft para su herramienta de hoja de cálculo.

**XML** – Formato extensible de marcas, lenguaje que permite la definición de etiquetas o marcas derivado del SGML.

**W3C** – Consorcio internacional que produce recomendaciones para la WWW.

**WEB SEMÁNTICA** – La web semántica es una visión futura de la web donde la información publicada debe tener un significado exacto, esta información debe ser entendida y procesada por computadoras. “Es una web de datos que pueden ser procesados directa o indirectamente por maquinas” según la definición de Tim Berners-Lee.

**WEB SERVICES** – En el caso que la información sea extraída de una fuente de información en constante modificación se recomienda la utilización de servicios web que expongan dicha información en tiempo real.



**datosabiertos.gob.pa**