



Grupo de Trabajo
de Datos Abiertos
de Gobierno



Guía de apertura y publicación de Datos **Abiertos** de Gobierno

Guía de Implementación Técnica de Datos Abiertos de Gobierno

Estrategia de Apertura, Publicación y Difusión

Septiembre 2020



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, dirigida a los equipos de datos abiertos de las instituciones públicas, reúne conceptos básicos de datos abiertos de gobierno, los pasos a seguir para abrir y publicar los datos, así como también ofrece referencias a contenidos más específicos para el lector que quiera profundizar en el tema.



Esta guía es publicada bajo Licencia Creative Commons Attributions 4.0 Internacional (CC BY SA 4.0) : 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.
Toda obra derivada de ésta deberá ser generada con estas mismas condiciones.

Versión 2.0 realizada por el Grupo de Trabajo de Datos Abiertos de Gobierno - GTDA como actualización a la versión 1.0 de 2016, realizada por la Autoridad para la Innovación Gubernamental (AIG) de la República de Panamá.

EQUIPO EDITORIAL

Katuska Hull Hurtado (GTDA)
Jorge Troyano Carracedo (GTDA)
Ulises Calvo (GTDA)
Diego Santamaría (GTDA)
Raisa Urribarri (GTDA)
Juan Pablo Rodríguez (ANTAI)
Rolando Becerra (AIG)
Luis Del Vasto (AIG)

Agradecimiento especial a EHC Group por edición gráfica.



República de Panamá
Autoridad Nacional para la Innovación Gubernamental (AIG)
Autoridad Nacional de Transparencia y Acceso a la Información (ANTAI)
Grupo de Trabajo de Datos Abiertos de Gobierno

Mediante la Resolución No. ANTAI-DS-PPTDAG-012-2019 del 29 de abril de 2019, se constituyó y reconoció a los miembros del Grupo de Trabajo de Datos Abiertos de Gobierno, para el periodo de 2019-2021,

Dentro de las funciones principales del Grupo de Trabajo de Datos Abiertos de Gobierno están: Elaborar el Plan de Acción Nacional de Datos Abiertos; realizar el diagnóstico y la hoja de ruta para el mejoramiento continuo de la PPTDAG; colaborar con las instituciones públicas en la identificación de los conjuntos de datos de valor generados en el ejercicio de sus funciones, que se publicarán e integrarán gradualmente al portal de Datos Abiertos; involucrar a los actores del “Ecosistema de Datos Abiertos” que considere necesario para el desarrollo de objetivos específicos; mantener informada a toda la comunidad de datos abiertos en Panamá sobre las actividades de sensibilización, capacitación y oportunidades de proyectos disponibles; promover los principios y usos de datos abiertos; deliberar acerca de cómo mejorar las actividades de implementación del Plan de Acción Nacional de Datos Abiertos.

El grupo colabora de manera activa en la actualización de la guía de apertura de datos dando como resultado este documento.

Mgtra. Elsa Fernández Aguilar
Directora General de la ANTAI

Ing. Luis Oliva
Administrador General de la AIG

GRUPO DE TRABAJO DE DATOS ABIERTO

Juan Pablo Rodríguez (Director de Acceso a la Información - ANTAI)

Luis del Vasto (Director de Ciencias de la Información - AIG)

Rolando Becerra (Coordinador de Datos Abiertos-Administrador del PDA - AIG)

Katuska Hull Hurtado (Secretaria GTDA)

Diego Santamaría (Subsecretario GTDA)

Ulises Calvo Poveda (GTDA)

Jorge A. Troyano Carracedo (GTDA)

Santiago Beltrán (GTDA)

Russel Bean (GTDA)

Raisa Urribarri (GTDA)

Nivia Castellon (GTDA)



“Los datos abiertos son la máxima expresión de la rendición de cuentas y transparencia institucional a la que debe aspirar toda sociedad, y abren un universo de oportunidades a los ciudadanos para crear nuevos productos y aplicaciones de información en beneficio de la sociedad”.

GTDA



Contenido

	Introducción	7
	1. Potencial y Beneficio de los Datos Abiertos de Gobierno	9
	2. Característica de los Datos Abiertos	10
	3. Estructura de los Datos Abiertos	11
	3.1. Tabular.	11
	3.2. Jerárquica.	12
	3.3. En Red.	15
	4. Plan institucional como estrategia para la apertura de datos	16
	4.1. Delegación del Oficial de Información	16
	4.1.1. Funciones	
	4.1.2. Conformación del equipo de trabajo	
	4.1.3. Funciones del equipo de trabajo	
	4.2. Generación de inventario de datos institucionales	19
	4.3. Publicación de Datos de la Institución	20
	4.3.1. Datos de gobierno que pueden ser publicados	20
	4.3.2. Selección del conjunto de datos a liberar	21
	4.3.3. Evaluación del estado del conjunto de datos	22
	a. Facilidad de publicación de datos	
	b. Calidad de los datos	
	4.3.4. Formatos utilizados para publicar datos abiertos	37
	a. Grados de apertura	
	b. Cómo convertir un archivo a formato abierto	
	4.3.5. Aplicar Licencia Abierta a los conjuntos de datos a publicar	47
	4.3.6. Inclusión de metadatos en los conjuntos de datos a publicar	49
	4.3.7. Asignación de URL O URI a los conjuntos de datos de 4 y 5 estrellas	50
	4.3.8. Carga de datos al sitio institucional y portal nacional	52
	a. Actualización y liberación de nuevos conjuntos de datos	
	4.3.9. Limpieza del conjunto de datos	56
	a. Errores frecuentes	
	b. Herramientas recomendadas para limpieza de datos	
	4.3.10. Divulgación del plan de apertura de Datos Abiertos	58
	4.3.11. Promoción en el uso de Datos Abiertos	59
	4.4. Monitoreo de publicación de datos	61
	4.4.1. A nivel nacional	
	4.4.2. A nivel Internacional	
	5. El futuro del portal nacional de datos abiertos	66
	6. Glosario de términos	68
	7. Referencias	70

Introducción

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

El hecho que los ciudadanos puedan acceder a todos estos datos producidos desde 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 a partir de los datos abiertos.

“Los datos abiertos de gobierno, son los datos recolectados y/o producidos por las instituciones públicas, puestos a disposición de los ciudadanos, en formato abierto para que puedan ser reutilizados y redistribuidos por cualquier persona, para cualquier propósito, incluyendo el re-uso comercial, libre de costos y sin restricciones, con el fin de posibilitar la lectura, el seguimiento y combinación con otras fuentes de información para generar nuevos servicios de valor”.¹

La apertura y publicación de datos abiertos requiere de una estructura operativa propia dentro de las instituciones gubernamentales, por la necesidad de ejecutar las directrices generales que lleven a formalizar las actividades de publicación de datos abiertos, tanto en el sitio web de la institución, como en el portal nacional de datos abiertos (PDA).

Estas directrices se encuentran fundamentadas tanto en la Ley 33 de 25 de abril de 2013 que crea la Autoridad Nacional de Transparencia y Acceso a la Información (ANTAI), al determinar la obligatoriedad de una unidad de enlace, en la figura del Oficial de Información, entre cada institución y la ANTAI, que vele, entre otras cosas, por la ejecución de la Política Pública de Transparencia de Datos Abiertos de Gobierno, establecida en el Decreto Ejecutivo 511 de 24 de

¹ Definición oficial para la República de Panamá, (ver artículo 2 de la Resolución DS-3513-2018 ANTAI)

noviembre de 2017; así como en la Resolución de la ANTAI No. DS-3513-2018 que la desarrolla. La designación del Oficial de Información debe darse con fundamento y bajo los parámetros que determina la Resolución No. de 167 de 13 de enero de 2017 de Carrera Administrativa.

Este documento propone un mecanismo para organizar, dentro de cada institución, los componentes que participarán en los procesos de apertura de datos, además de los procedimientos para la identificación, clasificación, tratamiento y publicación de los datos abiertos de gobierno. Como componente final y de gran importancia está la promoción de la apertura de datos en los formatos correspondientes. Al respecto se propone una serie de actividades para divulgar la apertura de datos en la institución y su publicación en el sitio web institucional y el portal nacional de datos abiertos de Panamá.

Es responsabilidad de los directivos de cada institución que se ejecuten dichas directrices, entendiendo la importancia y trascendencia que, en temas de gobernanza de datos, transparencia, interoperabilidad gubernamental, mejoramiento de los servicios públicos, y rendición de cuentas, tienen los datos abiertos, tanto a nivel nacional como internacional.

1. Potencial y Beneficio 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 con base en las necesidades de la comunidad y a mejorar los servicios públicos, en apoyo a los objetivos de desarrollo, generación de conocimiento y buena gobernanza;
- Abrir datos respetando 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 simplificar la información, para hacerla más comprensible a la comunidad a través de artículos, infografías, y visualizaciones de datos, entre otros; esto puede igualmente realizarse a través del uso de herramientas como software especializados que recaban y analizan grandes cantidades de datos.



2. Característica de los Datos Abiertos²



Accesibles

Deben estar disponibles en internet, para dar cabida a la más amplia gama de usuarios y de usos. Esto significa considerar opciones de acceso a discapacitados. No deben existir restricciones tecnológicas que impidan el acceso.



Libre uso - Licencia

Se publicarán datos que puedan ser usados, reusados y redistribuidos por cualquier usuario incluso con fines comerciales. Las instituciones deberán velar porque no se publiquen datos sujetos a ningún derecho de autor, patente, marca o regulación, restricciones de privacidad, seguridad y protección de datos, según la legislación nacional. Las entidades indicarán la licencia más apropiada que aplique a los datos liberados, como por ejemplo la licencia CC BY 4.0, que requiere citar la fuente de origen como único requisito o la licencia CC Zero que no lo requiere, lo que permite la reutilización de los mismos.



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.



No discriminatorios

Serán accesibles sin restricciones de registro previo o solicitud de identificación por parte de los usuarios.



Primarios

Proviene de la fuente de origen con el máximo nivel de granularidad posible, no modificados y sin procesar.



No propietarios

Deben estar disponibles en un formato, en el que ninguna entidad o marca tenga el control exclusivo sobre ellos, así como facilitar su acceso gratuito.



Oportunos

Serán actualizados periódicamente, conforme se generen o cambien.



Legibles por máquinas

Deberán estar estructurados, total o parcialmente, para ser procesados e interpretados por equipos electrónicos de manera automática.



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.

² <https://opengovdata.org/>

3. Estructura de los Datos Abiertos

No todos los datos se pueden expresar adecuadamente en formato de una hoja de cálculo, por tal motivo debe tenerse en cuenta los diferentes formatos de archivos que pueden permitir el libre acceso y reutilización de los datos.

Existen tres estructuras a las que debemos prestar especial atención: la tabular, la jerárquica y la estructura en red.

3.1. Estructura Tabular

La tabular es la estructura más común para la publicación de los datos. Los datos se organizan en filas y columnas y en ellos se enumeran valores secuenciales. Los datos se basan en entradas separadas que no están vinculadas entre ellas. Una estructura de archivo tabular es un formato como el CSV.

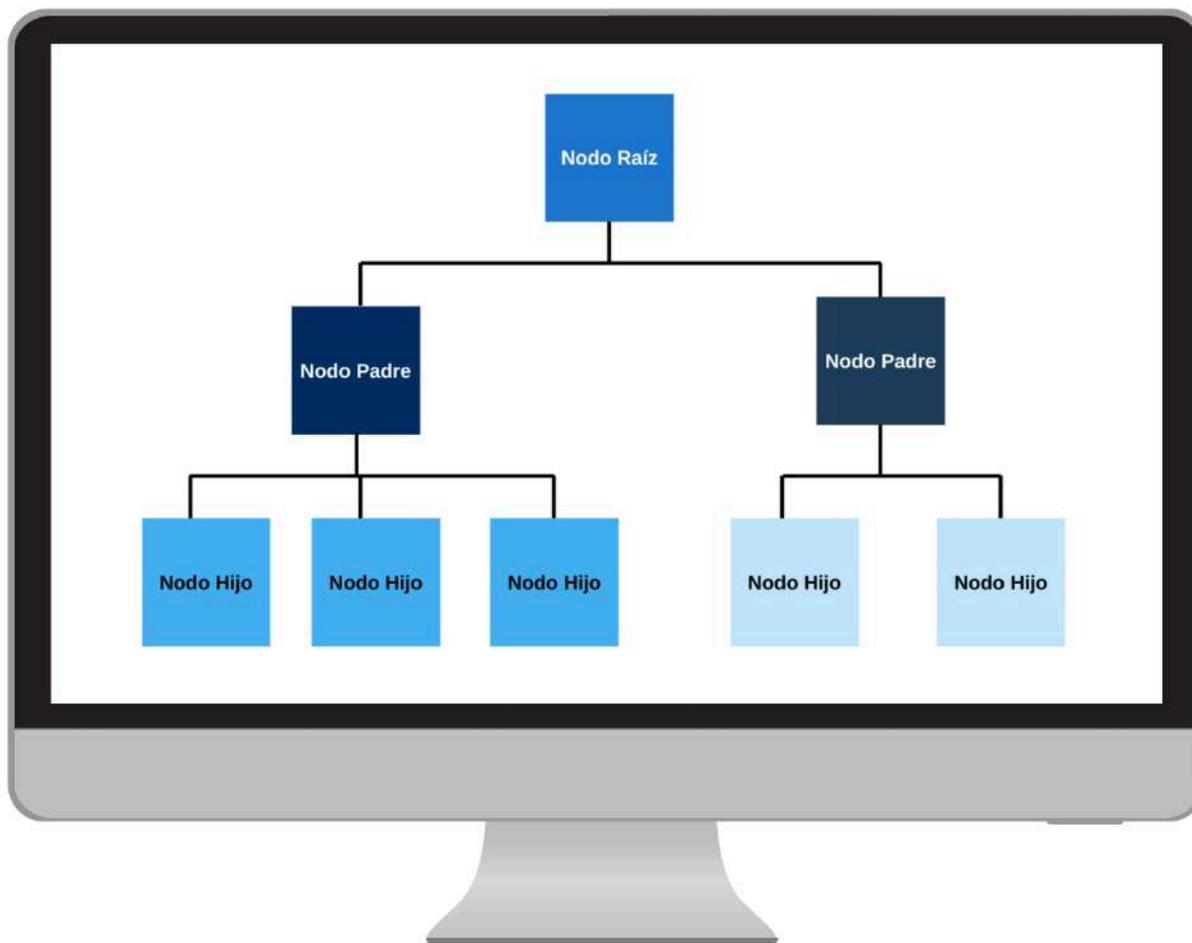
Ver ejemplo a continuación

CONTRATO O COMPRA	FECHA	DESCRIPCIÓN	PROVEEDOR	PRECIO ADJUDICADO B/.
42000324568	14-ENERO	Suministro e instalación de 150 topes para estacionamientos SERTV	Les Palais, S.A.	10,816.00
4200035231	27-ENERO	Suministro, instalación y reparación de cerca perimetral	GLP GROUP, S.A.	19,000.00
4200032561	08-FEBRERO	Suministro, transporte al sitio, instalación, puesta en servicio de un sistema radiante	BTESA América, S.A.	67,990.00
4200034596	09-FEBRERO	Suministro, instalación e integración de enlaces de microondas en banda licenciada	GTC, Supplies, S.A.	117,484.00
SERV-003-2021	09-FEBRERO	Suministro, instalación e integración de enlaces de frecuencia de celulares	Bitcom, S.A.	58,500.00
42000330253	03-MARZO	Servicio de mantenimiento preventivo y correctivo de librería robótica de cintas	Bitcom, S.A.	7,035.00
SERV-001-2021	03-MARZO	Servicio de fumigación trimestral del edificio principal de SERTV	César Abdiel, Araúz	24,888.64
42000332583	12-MARZO	Suministro, instalación y configuración de sistema de edición para producción de imagen	DATA SERVE, S.A.	225,860.70
SERV-0042021	16-MARZO	Servicio de modificación de sistema de iluminación del edificio principal de SERTV	Rio Pison	36,010.00
SERV-002-2021	29-MARZO	Mantenimiento, actualización y soporte para el software de gestión de activos de tecnología informática	Nextlifesoft	8,400.00
SERV-021-2020	30-MARZO	Mantenimiento preventivo y correctivo del sistema UPS	Tecnología Aplicada, S.A.	166,212.00
TOTAL				723,215.34

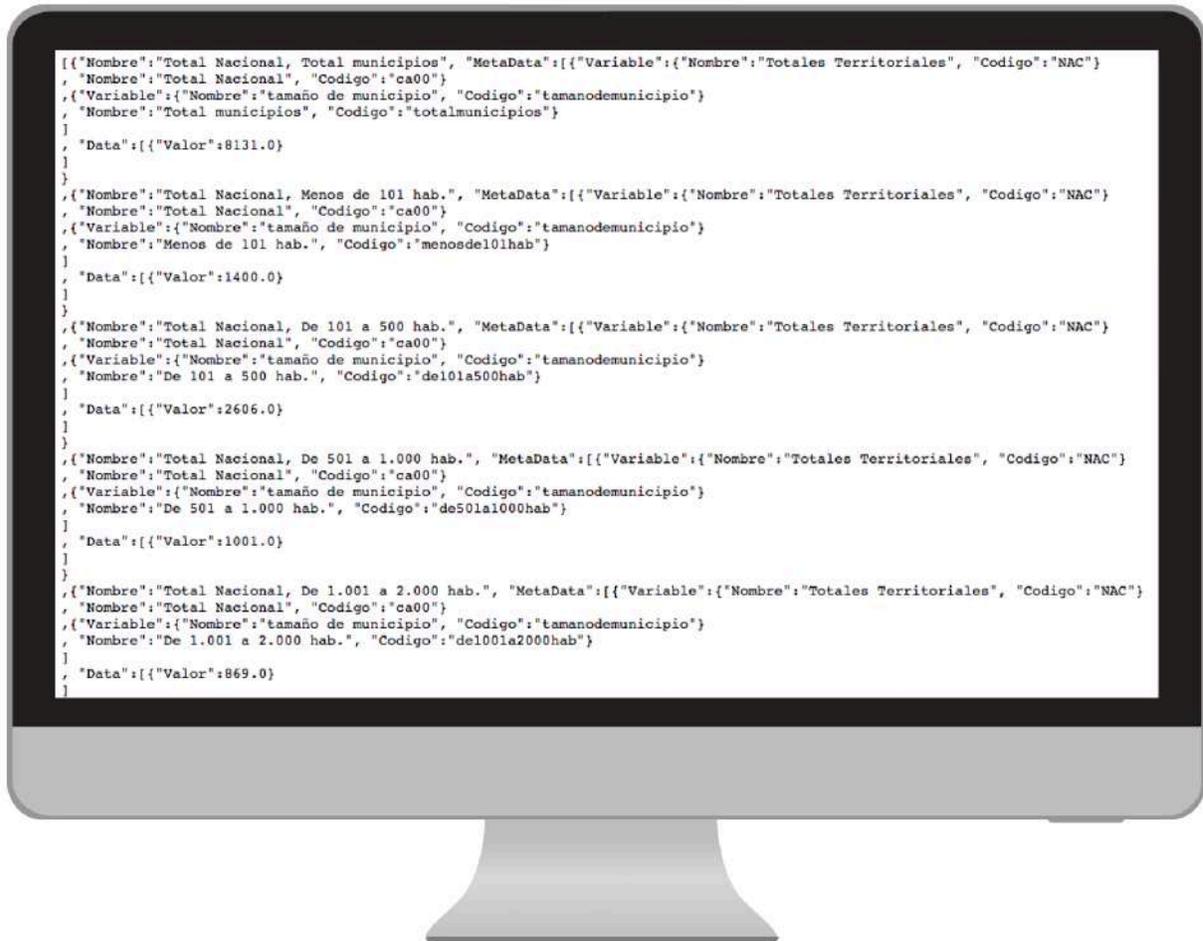
3.2. Estructura Jerárquica

Los datos jerárquicos muestran las relaciones que existen entre los puntos de datos, como un árbol genealógico, datos geoespaciales o los municipios de un país. Cuando el conjunto de datos depende de la relación entre los puntos de datos y sigue una estructura en la que los puntos de datos están vinculados en árboles verticales, una estructura jerárquica de datos en un formato como JSON o XML es ideal.

Ejemplo de estructura jerárquica



Ejemplo de XML describiendo la estructura jerárquica municipal



Ejemplo de datos geoespaciales

UNIDAD EJECUTORA	LATITUD	LONGITUD
Centro Especializado de Toxicología	8.956295	-79.541686
Policlínica Especializada de Calle 17	8.956546	-79.541530
Policlínica Especializada de Calidonia	8.962381	-79.539256
Hospital general	8.980846	-79.535691
Consulta Externa Especializada	8.980662	-79.534698
Policlínica Especializada de San Francisco	8.997185	-79.511221
Policlínica Especializada de Bethania	9.021664	-79.525864
ULAPS de San Cristobal	9.028056	-79.492178
Policlínica de San Miguelito	9.037875	-79.495753
ULAPS del Hipodromo	9.036982	-79.467214
Policlínica Especializada de Juan Díaz	9.055207	-79.434789
CAPPS de Plaza Tocumen	9.060664	-79.422784
CAPPS de Los Nogales	9.077122	-79.411916
CAPPS de Pedregal	9.063430	-79.431989
CAPPS de Torrijos Carter	9.077475	-79.491826
Policlínica de Santa Librada	9.072028	-79.505415
ULAPS de Las Cumbres	9.079699	-79.527307
Hospital San Judas Tadeo	9.038630	-79.482013

3.3. Estructura en Red

Los datos estructurados en red permiten que existan relaciones entre cualquier combinación de elementos en cualquier dirección. Un buen ejemplo de una estructura de datos en red es una red social. Piense en su red de amigos y sus amigos en Facebook, o en los contactos de primer, segundo y tercer grado de LinkedIn.

La mayoría de estos datos se actualizan con tanta frecuencia que las descargas del archivo serían demasiado pesadas para la mayor parte de los usuarios, por lo que no es recomendable la publicación como un archivo descargable. Este tipo de estructura de datos abiertos puede publicarse a través de una interfaz de máquina, también conocida como interfaz de programación de aplicaciones (API, por su sigla en inglés)

Ver ejemplo a continuación



4. Plan Institucional como estrategia para la apertura de datos.

4.1. Delegación del Oficial de Información

El Oficial de Información, tal como esta dispuesto en el artículo 6 numeral 8, y el artículo 8 numeral 1 de la Ley 33 de 25 de abril de 2013 que crea la Autoridad Nacional de Transparencia y Acceso a la Información (ANTAI), tiene la responsabilidad de liderar la apertura de datos. Sin embargo, la responsabilidad institucional de publicarlos en el Portal DAP y las consecuencias de que esto no se cumpla, recaen en igual medida tanto en el Representante Legal de la institución como en el Oficial de Información. De ahí la conveniencia de que exista un responsable de dar seguimiento a las actividades a desarrollar para la apertura de datos y de que todos los funcionarios con nivel directivo o de jefatura, colaboren con el equipo de trabajo de Datos Abiertos de la institución, para así cumplir con los objetivos de publicación en el Portal DAP.

4.1.1. Funciones del oficial de información en la apertura de datos

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

- Determinar, en colaboración con el Grupo de Trabajo de Datos Abiertos de Panamá - GTDA, los conjuntos de datos de valor generados por la Institución.

4.1.2. Conformación del Equipo de Trabajo

El Oficial de Información junto con el equipo de trabajo que convoque, serán los responsables de implementar la política pública de datos abiertos de gobierno en toda la institución.

El equipo de trabajo será la estructura operativa encargada de supervisar el levantamiento y adecuación de los datos de la institución, y además actuará como un consejo institucional para desarrollar e implementar la política de datos abiertos.

El tamaño del equipo corresponderá a la dimensión y ámbitos de ejecución de la institución y deberá estar conformado como mínimo por:

1. El Oficial de Información que, como ya indicamos, lleva el rol de liderar la apertura de datos dentro de la institución, desde la identificación hasta la publicación de la información.
2. Al menos un (1) representante del Departamento de Informática o del área de TI de la institución, como responsable del aspecto técnico de la publicación de datos abiertos en los formatos correctos.
3. Al menos un (1) representante del Departamento de Asesoría 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, así como el cumplimiento de las normativas especiales como la de Protección de Datos Personales;
4. Al menos un (1) representante de cada una de las unidades, proyectos y áreas que generan y/o administran información; es decir realizan el rol de identificación de información.

Cada persona designada deberá tener un suplente que asistirá a las reuniones del equipo y dará seguimiento a las tareas asignadas al principal, en ausencia de éste. Es oportuno aclarar que las designaciones al equipo de trabajo, tanto del principal, como del suplente, son permanentes, es decir, las personas designadas para formar parte del equipo de trabajo serán los representantes

Roles del Equipo de Trabajo

	Identificación de Datos	Análisis de Información	Priorización de Información	Levantamiento del conjunto de Datos	Carga en el portal Nacional
 ROL LÍDER	<input checked="" type="checkbox"/>				
 ROL TÉCNICO			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
 ROL LEGAL		<input checked="" type="checkbox"/>			
 ROL FUNCIONAL	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		

permanentes de su área de trabajo.

4.1.3. Funciones del Equipo de Trabajo

Como estructura operativa, el equipo de trabajo de datos abiertos de la institución, debe trabajar en función de los roles identificados por su especialidad, en actividades ya identificadas que van desde la identificación de los datos de valor a publicar, la revisión de la información para el cumplimiento de los controles legales de los contenidos, así como el levantamiento de los conjuntos de datos y su carga en el portal. Para el cumplimiento de sus funciones es importante tener una ruta de trabajo que comprenda:

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



4.2. Generación de Inventario de Datos Institucionales

Es importante hacer un inventario y ordenar la información con la que cuenta la institución; para esto se crea el inventario de datos institucional, que incluye 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 Oficial de Información 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. Este inventario de datos será conformado 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. Este inventario será realizado con la colaboración del Grupo de Trabajo de Datos Abiertos de Gobierno (GTDA) conforme lo estipulado en el artículo 13 de la Resolución de la ANTAI No. DS- 3513-2018.³

El inventario debe salir de los departamentos depurado para seleccionar la información y posteriormente poder liberar en formato de datos abiertos de acuerdo al interés social para su reutilización. Debe priorizarse los datos en rangos de importancia para que la publicación se realice en ese orden

Modelo de Plantilla para Inventario de Datos

Responsable (persona que genera el conjunto de datos o es responsable del servicio de información de donde sale el mismo)	
Nombre del conjunto de datos o servicio de información	
Temáticas que lo componen	
Descripción del servicio o conjunto de Datos	
Reserva legal de la información (describir si contienen datos privados, sensibles, confidenciales, de seguridad nacional, relativos a datos personales)	

³ <https://www.antai.gob.pa/wp-content/uploads/2018/02/Resoluci%C3%B3n-No.-DS-3513-2018-de-17-de-enero-de-2018.pdf>

Plataforma tecnológica, aplicación o programa en que están contenidos los datos	
Periodo de actualización del conjunto de datos o servicios de información	
Estado del conjunto de datos (revisado, depurado, filtrado)	

4.3. **Publicación de Datos 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.

4.3.1. **Datos de Gobierno que pueden ser publicados**

Los datos públicos de gobierno que pueden ser publicados con formatos 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”*; y que no se encuentre dentro de las reservas legales a la información como datos de carácter personal, ni de seguridad nacional, con derecho de autor, o sujetos a restricciones específicas. Por ejemplo, la información contenida en el presupuesto público, ejecución del presupuesto, estadísticas, datos del censo, compras públicas, desempeño/ eficiencia de los servicios, datos del transporte, datos judiciales, policíacos no sujetos a reservas legales, datos de ambiente, energía, comercio e industria, educación, servicios públicos y otros.

Al realizar el inventario que será publicado se le debe dar importancia a la apertura de nuevos datos, que sirvan para el desarrollo económico, social, investigativo, asegurándose de que los mismos tienen alta demanda dentro de la comunidad. Asimismo, la serie de conjuntos de datos estarán fechados y deberán mantenerse en series anuales, de tal forma que se puedan bajar datos completos masivamente para los análisis y comparaciones estadísticas.

Al utilizar diversos tipos de 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. En cumplimiento de la política pública de datos abiertos del gobierno de Panamá (DAGP) es obligatorio 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/o 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>.

4.3.2. Selección de conjunto de datos a liberar

Para seleccionar el conjunto de datos a liberar se considerarán las necesidades de los usuarios, no tiene sentido publicar datos abiertos si no hay nadie que esté preparado o motivado para generar nuevo valor a partir de ellos. Piense en los problemas que existen en el país y de qué manera sus datos se relacionan con ellos en el sentido de ofrecerle una herramienta que ayude a crear soluciones.

Puede identificar los datos que ya tiene 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. Se recomienda priorizar aquellos que tengan el mejor equilibrio entre sí, teniendo en consideración los siguientes criterios:

1. Conjuntos de datos para los cuales existe una alta demanda o importancia estratégica.
2. Disponibles para publicación (Disponibilidad y buena calidad de los datos para publicación), o que ya estén publicados en las páginas web de la institución o en algún otro lado, pero no en formato abierto.
3. Conjuntos de datos requeridos por la ley o normativa.
4. Datos con requerimientos específicos o llamados como prioridades nacionales.
5. Datos que se le solicitan habitualmente.
6. Información existente en la oficina que puede ser de interés para la población, información que se le brinda habitualmente a la prensa o que se publica en medios gráficos.
7. Información que se les brinda a otros organismos o instituciones que puede ser de interés general o particular de otra institución.
8. Datos contenidos en informes de gestión de la institución.

4.3.3. Evaluación del estado del conjunto de datos

Esta guía contempla dos subdivisiones para la evaluación del estado de los datos: facilidad de publicación de datos y calidad de los datos. La primera se refiere al estado de los datos para su publicación, qué conjuntos de datos requieren de poco esfuerzo o trabajo previo para ser publicados, debido a su estado, y la segunda se refiere a la capacidad o característica de los datos que aportan valor o satisfacen la necesidad de reutilización con respecto a otros datos.

a. Facilidad de publicación de datos

La manera más fácil de evaluar el estado del conjunto de datos que se quiere publicar, es mediante la revisión de criterios de controles previamente establecido. A continuación presentamos una plantilla modelo que facilita la evaluación de los datos tomando en consideración indicadores como la forma de captura, la estructura en la que se encuentran, el formato, el criterio legal de publicación, y su actualización. Cada uno de éstos criterios de evaluación dan una puntuación de 1 a 3, en donde a mayor puntuación, mayor el cumplimiento del conjunto de datos en el criterio de facilidad de publicación.

Modelo de Plantilla

INDICADOR	NIVEL	VALOR	TOTAL
1. CAPTURA	No se recolectan datos	1	
	Se recolectan pero no están sistematizados	2	
	Se recolectan y están sistematizados	3	
2. ESTRUCTURA	No estructurados	1	
	Semiestructurados	2	
	Estructurados	3	
3. FORMATO	Físico	1	
	Digital no estructurado	2	
	Digital estructurado	3	
4. LEGAL	Información de acceso restringido/confidencial	1	
	Información que requiere anonimización (de carácter personal)	2	
	Información de Acceso Público	3	
5. ACTUALIZACIÓN	Desactualizado	1	
	Actualizado al periodo anterior	2	
	Actualizado	3	
PUNTUACIÓN	Máximo puntaje	15	

Calificación de la facilidad de publicación

Valores de 1 y 5. Implica que es un conjunto de datos que técnicamente requiere de mayor tratamiento para lograr ser publicado en formato abierto.

Valores de 6 y 10. Implica que es un conjunto de datos que técnicamente requiere de mejoras menores para ser publicado en formato abierto.

Valores de 11 y 15. Implica que es un conjunto de datos que técnicamente cuenta con buenas condiciones para ser publicado en formato abierto.

b. Calidad de los datos

Con la finalidad de focalizarnos en la reutilización de los datos, se toman en consideración seis (6) indicadores para evaluar su calidad, a saber: estructura, completitud, importancia, desglose, transparencia de evaluación y actualización.

La buena calidad ayudará a que el conjunto de datos que sea publicado pueda ser reutilizado, ya sea para una investigación académica, interpretación del mercado, rendición de cuentas, información complementaria en un negocio y hasta la creación de emprendimientos y aplicaciones tecnológicas.

- **Estructura:** Un conjunto de datos puede encontrarse en cuanto a su estructura, de tres maneras:
 - Estructurado
 - Semi estructurado
 - No estructurado

Estructurado: cuando tanto su contenido como su formato son legibles por máquinas como por seres humanos. En caso de que la estructura de los datos sea tabular deberá tomar en cuentas las siguientes consideraciones:

1. Cada variable es una columna.
2. Cada observación o registro es una fila.
3. Cada intersección de fila y columna es una celda.
4. Cada conjunto de observaciones es una tabla.
5. Debe tener una línea única de cabecera
6. Debe tener registro único por fila
7. El crecimiento del registro de datos debe ser vertical

Semi estructurados: conjuntos de datos que no son fácilmente legibles por máquinas, pero son entendibles por humanos; es decir, que no cumplen con todas las consideraciones o características expresadas anteriormente.

No estructurados: conjuntos de datos en los que el contenido no es fácilmente legible ni por los humanos ni por las máquinas, ya que como norma no se encuentran ordenados en forma de tabla.

A continuación, veremos varios ejemplos, en los cuales podemos distinguir la estructura de los datos presentados y su cumplimiento o no

con las variables tabulares.

Ejemplo No. 1 Datos no Estructurados

Tomando como ejemplo un texto que aparece en la página web del Servicio Nacional de Migración (cuadro: Solicitudes Admitidas a Trámite), que explica las solicitudes admitidas a trámite para personas en calidad de refugiados, se puede notar que no tiene la estructura detallada más arriba, ya que las variables no están establecidas como tablas, sino que se presentan como párrafos completos de textos.

Debemos tener presente que un texto sólo debe incluirse en un conjunto de datos abiertos, incorporado al pie de página, para dar una explicación de un cuadro tabular siempre y cuando sea en el formato PDF o XLSX, pero no debe ser incluido en el formato CSV, ya que este formato es para conjuntos de datos estructurados y ordenados.



Como podemos observar de la información anterior, y haciendo anotaciones a las consideraciones de datos tabulares:

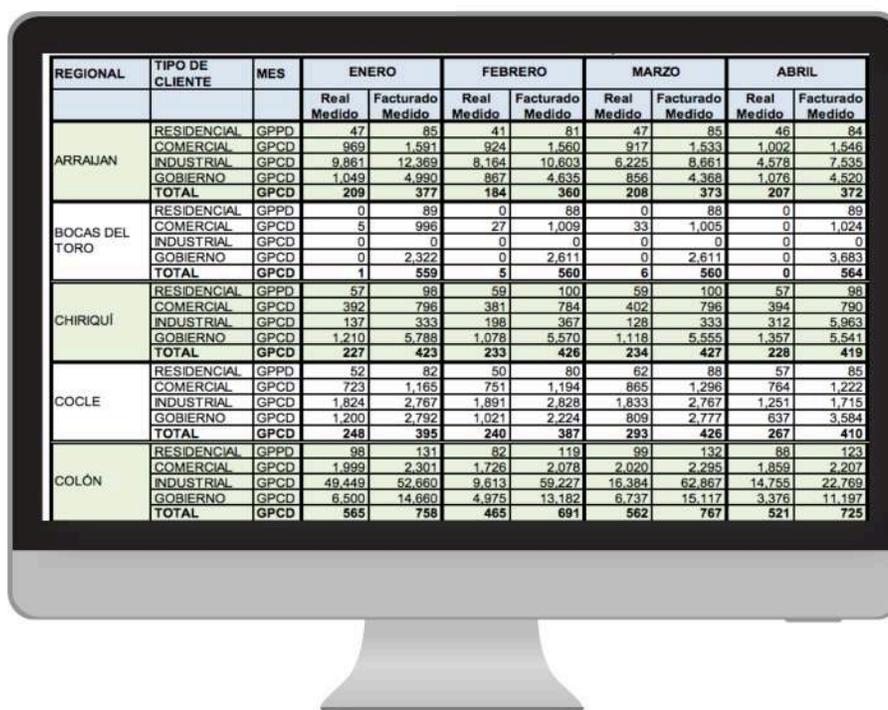
1. Las variables no están expresadas en columna.
2. Cada observación no es un registro único por filas ya que no está ordenado en ningún cuadro.
3. No existen celdas, ni líneas de cabeceras.

4. No existe ninguna tabla.
5. No es posible un crecimiento ordenado de los datos

Ejemplo No. 2 Datos

Semi-estructurados

A continuación presentamos un ejemplo, en donde los datos son fácilmente legibles para los humanos, pero su lectura resulta difícil para las máquinas debido a que no cumplen todas las consideraciones mencionadas en la estructura ideal de datos.



REGIONAL	TIPO DE CLIENTE	MES	ENERO		FEBRERO		MARZO		ABRIL	
			Real Medido	Facturado Medido						
ARRAJUAN	RESIDENCIAL	GPPD	47	85	41	81	47	85	46	84
	COMERCIAL	GPCD	969	1,591	924	1,560	917	1,533	1,002	1,546
	INDUSTRIAL	GPCD	9,861	12,369	8,164	10,603	6,225	8,661	4,578	7,535
	GOBIERNO	GPCD	1,049	4,990	867	4,635	856	4,368	1,076	4,520
	TOTAL	GPCD	209	377	184	360	208	373	207	372
BOCAS DEL TORO	RESIDENCIAL	GPPD	0	89	0	88	0	88	0	89
	COMERCIAL	GPCD	5	996	27	1,009	33	1,005	0	1,024
	INDUSTRIAL	GPCD	0	0	0	0	0	0	0	0
	GOBIERNO	GPCD	0	2,322	0	2,611	0	2,611	0	3,683
	TOTAL	GPCD	1	559	5	560	6	560	0	564
CHIRIQUÍ	RESIDENCIAL	GPPD	57	98	59	100	59	100	57	98
	COMERCIAL	GPCD	392	796	381	784	402	796	394	790
	INDUSTRIAL	GPCD	137	333	198	367	128	333	312	5,963
	GOBIERNO	GPCD	1,210	5,788	1,078	5,570	1,118	5,555	1,357	5,541
	TOTAL	GPCD	227	423	233	426	234	427	228	419
COCLE	RESIDENCIAL	GPPD	52	82	50	80	62	88	57	85
	COMERCIAL	GPCD	723	1,165	751	1,194	865	1,296	764	1,222
	INDUSTRIAL	GPCD	1,824	2,767	1,891	2,828	1,833	2,767	1,251	1,715
	GOBIERNO	GPCD	1,200	2,792	1,021	2,224	809	2,777	637	3,584
	TOTAL	GPCD	248	395	240	387	293	426	267	410
COLÓN	RESIDENCIAL	GPPD	98	131	82	119	99	132	88	123
	COMERCIAL	GPCD	1,999	2,301	1,726	2,078	2,020	2,295	1,859	2,207
	INDUSTRIAL	GPCD	49,449	52,660	9,613	59,227	16,384	62,867	14,755	22,769
	GOBIERNO	GPCD	6,500	14,660	4,975	13,182	6,737	15,117	3,376	11,197
	TOTAL	GPCD	565	758	465	691	562	767	521	725

Sobre las consideraciones de los datos tabulares expuestos podemos anotar:

1. Cada variable es una columna.
2. Cada observación no es un registro por filas, ya que las observaciones de tiempo expresada en meses están como columnas, y estas deben ser expresadas en fechas exactas en filas, en un formato como dd/ mm/aa.
3. Cada intersección de fila y columna no es una celda ya que algunas están celdas combinadas.
4. Cada conjunto de observación es una tabla.
5. No tiene una línea única de cabecera.
6. No tiene un registro único por filas ya que están en celdas combinadas

- justo dónde están los nombres de las localidades geográficas.
- El crecimiento del registro no es vertical.

Ejemplo No. 3 Datos

Estructurados

El siguiente ejemplo cumple todas las consideraciones de datos estructurados por su orden y organización.

FECHA	No. DE ACTA	AGENTE ECONÓMICO	PROVINCIA	CORREGIMIENTO	ACTIVIDAD
02-ene-20	897,472	ZAZ FFOOD STORE	Panamá	SAN FRANCISCO	Minisuper
02-ene-20	897,599	MINI MAEKET CENTRO PARK	Panamá	BELLA VISTA	Minisuper
02-ene-20	897,600	MINI MARKET CENTRO PARK	Panamá	BELLA VISTA	Minisuper
03-ene-20	897,540	MINI SUPER AMADOR, S.A	Panamá	ANCÓN	Minisuper
03-ene-20	897,629	PRICE SMART PANAMÁ	Panamá	BELLA VISTA	Supermercados
06-ene-20	897,335	COMISARIATO Y BODEGA HAU	Panamá	AMELIA DENIS DE ICAZA	Minisuper
06-ene-20	897,423	MR AREPAS METROMALL	Panamá	JUAN DIAZ	Restaurantes / fondas
06-ene-20	897,541	TACO BELL	Panamá	JOSE DOMINGO ESPINAR	Restaurantes / fondas
06-ene-20	897,630	RIBSSTATE METRO MALL	Panamá	RUFINA ALFARO	Restaurantes / fondas
07-ene-20	897,337	CAFETERÍA CAKE BOON PLACE	Panamá	JUAN DIAZ	Panaderías / refresquerías
07-ene-20	897,520	SAN FELIPE NERE PUESTO 18	Panamá	SAN FELIPE	Tiendas, abarroterías,bancos,kioscos, etc
07-ene-20	897,522	ATHANASIOU MULTIPLAZA	Panamá	SAN FRANCISCO	Restaurantes / fondas
07-ene-20	897,568	FARMACIA ARROCHA, S.A.	Panamá	PUEBLO NUEVO	Farmacias
07-ene-20	897,631	PUESTO DE LEGUMBRES	Panamá	SAN FELIPE	Tiendas, abarroterías,bancos,kioscos, etc
07-ene-20	897,678	MERCADO SAN FELIPE NERI LOCAL	Panamá	SAN FELIPE	Tiendas, abarroterías,bancos,kioscos, etc
08-ene-20	897,523	GRUPO ONCOLOGOS MÉDICOS	Panamá	PUEBLO NUEVO	Clinicas
08-ene-20	897,680	RESTAURANTE DIABLICOS	Panamá	SAN FELIPE	Restaurantes / fondas
10-ene-20	897,569	FONDA DOMINICOS PANAMÁ	Panamá	VERACRUZ	Restaurantes / fondas
14-ene-20	897,654	MINI SUPER LA ECONÓMICA	Panamá	CALIDONIA	Minisuper
14-ene-20	897,704	RESTAURANTE EL TRAPICHE	Panamá	BELLA VISTA	Restaurantes / fondas
14-ene-20	897,705	MEDIC FEET	Panamá	BELLA VISTA	Salón de belleza
16-ene-20	897,638	CENTRO MÉDICO CERRO BATEA	Panamá	BELISARIO FRIAS	Clinicas
16-ene-20	897,640	KIOSCO CRISTINA	Panamá	ANCÓN	Tiendas, abarroterías,bancos,kioscos, etc
16-ene-20	897,658	CLÍNICA MÉDICA LA ESPERANZA,SA	Panamá	BELISARIO FRIAS	Clinicas
16-ene-20	897,659	KIOSCO N° 5	Panamá	ANCÓN	Tiendas, abarroterías,bancos,kioscos, etc
17-ene-20	897,643	KFC-SUCURSAL N° 20	Panamá	ANCÓN	Restaurantes / fondas
20-ene-20	897,543	ABARROTERÍA OXCAR	Panamá	PUEBLO NUEVO	Tiendas, abarroterías,bancos,kioscos, etc

Sobre las consideraciones expuestas acerca de los datos tabulares podemos anotar que la tabla cumple con las condiciones de los datos estructurados ya que:

- Cada variable es una columna, no presenta celdas combinadas.
- Cada observación es un registro por filas; esto se nota claramente por que se observa la fecha específica de la observación y un registro exacto del formulario utilizado en este caso un número de acta.
- Cada intersección de columna y fila es una celda, no presenta celdas combinadas.
- Cada conjunto de observación es una tabla.

5. Tiene una línea única de cabecera.
 6. Tiene un registro único por filas.
 7. El crecimiento del registro es vertical.
- **Compleitud:** El conjunto de datos se considerará completo si contiene todos sus atributos y propiedades, si es único, es decir, que no se repite dentro del portal, y si son primarios. Por su parte, los datos se consideran incompletos si tienen mínimas cantidades de filas (menos de 10), o cuando están expresados como consolidados de información, ya que cuentan
c o n
pocas posibilidades de ser reutilizados.

Ejemplo No. 4 Datos Incompletos

Tal como se indicó arriba, se consideran los datos incompletos cuando un cuadro contiene menos de 10 filas, y la información es escasa para su reutilización, así mismo su la información se presentan como datos consolidados, es decir constituye una muestra muy general de información, tal como podemos apreciar en el ejemplo presentado a continuación.

REGIÓN	GASTOS TOTALES EN BALBOAS	AGUA PRODUCIDA EN MILLARES DE GALONES
TOTAL	80,060	97,018
BOCAS DEL TORO	304	338
COCLÉ	2,294	3,059
COLÓN	10,568	11,497
CHIRIQUÍ	4,487	12,284
PANAMÁ ESTE Y DARIÉN	1,427	593
HERRERA	2,311	2,914
LOS SANTOS	1,529	1,108
PANAMA METRO	49,329	58,000
VERAGUAS	2,455	2,232
PANAMÁ OESTE	6,156	4,993

Ejemplo No. 5 Datos Completos

El siguiente ejemplo muestra un conjunto de datos completos, ya que cumple con los criterios detallados, como suficiente cantidad de filas que aportan abundancia de contenido con el cual trabajar, así como estructuras ideales al contar con regustos únicos por filas y columnas, además que el contenido puede crecer verticalmente.



POSICIÓN	NOMBRES	APELLIDOS	CEDULA	SEXO	SALARIO BRUTO EN (B/.)	GASTOS DE REPRESENTACIÓN EN (B/.)	TOTAL EN (B/.)
1	Andrea Gisela	Chin Gonzalez	8-314-56	F	6,000.00	2,000.00	8,000.00
2	Omar Agustín	López Arrien	4-876-352-	M	4,500.00	1,000.00	5,500.00
4	Juan Pablo	Lambrano Paz	7-546-21345	M	3,500.00	1,000.00	4,500.00
11	Agustín Rafael	Andrade Ruiz	PE-18-254	M	850.00	0.00	850.00
90	Remigia	Gonzalez Chong	4-329-1001	F	800.00	0.00	800.00
18	Carla Maria	Salcedo Vos	3-986-8354	F	1,800.00	1,000.00	2,800.00
46	Dorca Angelina	De La Rosa	2-674-9087	F	3,000.00	1,500.00	4,500.00
54	José María	Martínez Ruiz	8-4237-735	M	2,000.00	1,000.00	3,000.00
12	Javier Alexander	Ramos Quiroz	8-905-1234	M	2,000.00	1,000.00	3,000.00
14	Arturo	Coronado Luz	8-6479-9035	M	750.00	0.00	750.00
19	Diego Armando	Hurtado Perez	7-904-873	M	850.00	0.00	850.00
36	Regina	Camacho	8-563-10	F	800.00	0.00	800.00
56	Juan Antonio	Cabeza Rendon	8-6534-8739	M	950.00	0.00	950.00
87	Maria Antonieta	Díaz Pérez	8-7649-93	F	1,300.00	1,000.00	2,300.00
22	Ana Victoria	Murillo Rios	8-124-9083	F	1,200.00	1,000.00	2,200.00

- **Importancia:** En el ámbito de importancia, se distinguen los conjuntos de datos de acuerdo con su prioridad:
 - i. **Prioridad 1:** Conjuntos de datos que se consideren de una importancia alta, al contemplar los siguientes criterios:
 - Todo conjunto de datos que se refiera directamente a la misión, visión de la institución pública y a las funciones principales de cada departamento.
 - Datos con requerimientos específicos o llamados prioridad nacional.
 - Datos producto del cumplimiento de los compromisos del plan de acción de gobierno abierto vigente.
 - Datos solicitados en el inventario levantado por el Grupo de Trabajo de Datos Abiertos Panamá.
 - ii. **Prioridad 2:** Conjunto de datos que se consideren de una importancia media, al contener las siguientes características:
 - Datos que se consideren de reutilización media o poco frecuente
 - Todos los datos relativos al manejo del presupuesto del estado, como planillas, gastos públicos, ya que su publicación ya es obligatoria conforme a la ley en la sección de transparencia en su página web.
 - Datos que solo tengan el objetivo de consulta, curiosidad o de señalamiento de algún logro.
 - iii. **Prioridad 3:** Conjunto de datos que se consideren de una importancia baja:
 - Información de una reutilización nula o mínima.
 - Datos sin ningún tipo de relevancia.

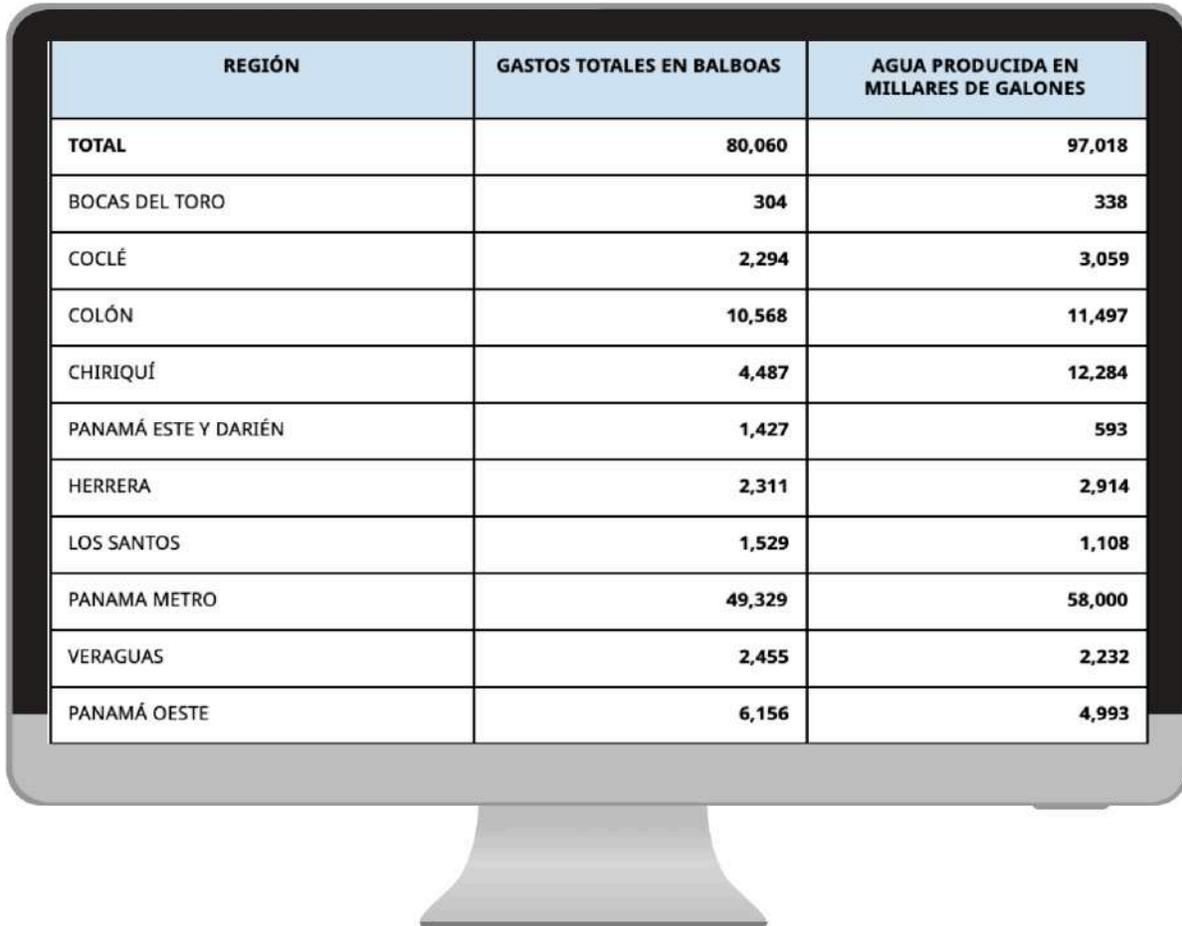
- **Desglose:** Al igual que en el indicador anterior, se distinguen tres niveles, en virtud de la segregación que tenga la información:
 - i. **Nivel de desglose 1:** La información está generalizada y se puede ampliar o desagregar más, por ejemplo, en el caso que se publiquen datos que contenga información relativa a la división política, el publicar sólo en variable de provincias, y no así en circunscripciones más pequeñas o específicas, como corregimientos, hacen que la información sea calificada en nivel

de desglose 1.

En el siguiente ejemplo vemos cómo la división política está solo expresada por provincias, lo cual no aporta mayor profundidad de información.

Ejemplo No. 6 Nivel

de Desglose 1

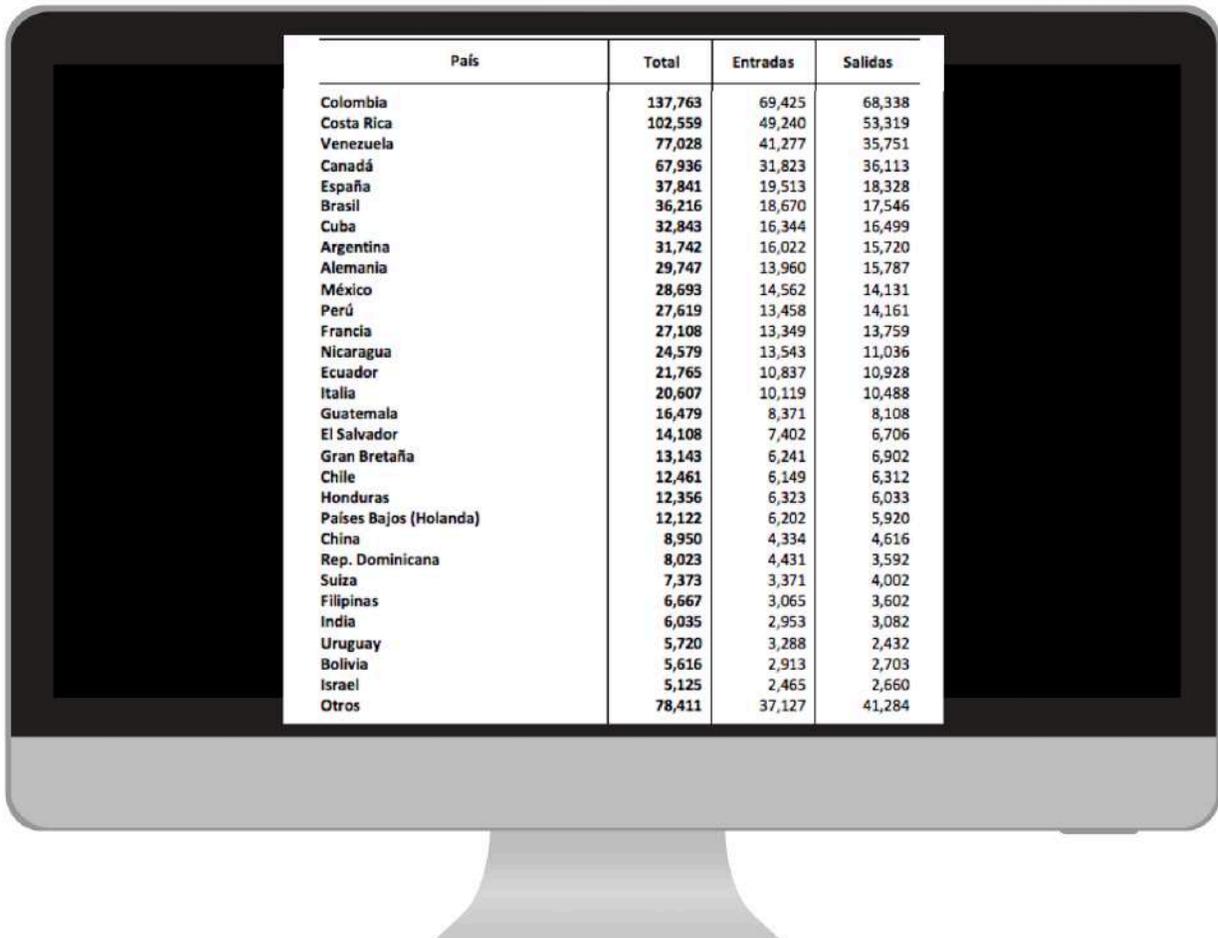


REGIÓN	GASTOS TOTALES EN BALBOAS	AGUA PRODUCIDA EN MILLARES DE GALONES
TOTAL	80,060	97,018
BOCAS DEL TORO	304	338
COCLÉ	2,294	3,059
COLÓN	10,568	11,497
CHIRIQUÍ	4,487	12,284
PANAMÁ ESTE Y DARIÉN	1,427	593
HERRERA	2,311	2,914
LOS SANTOS	1,529	1,108
PANAMA METRO	49,329	58,000
VERAGUAS	2,455	2,232
PANAMÁ OESTE	6,156	4,993

- ii. **Nivel de desglose 2:** En este nivel la información está medianamente desagregada, pero puede ser seccionada mejor. Siguiendo el ejemplo de la división política, se entiende en desglose nivel 2, el conjunto de datos que sólo está desglosado por provincia y distrito, pero que no llega a su estado más granulado. Otro ejemplo sería la información contenida en el presupuesto de gastos, en donde se califica en nivel 2 de desglose el conjunto de datos que tenga una división de gastos segregada, pero la misma sigue expresándose en un consolidado, por ejemplo, si hablamos del gasto de materiales y suministros, el desglose debe indicar si son textiles, alimentos, papelería, entre otros. En ninguna fila debe existir variables que indiquen sólo

consolidados.

Ejemplo No. 7 Nivel de Desglose 2



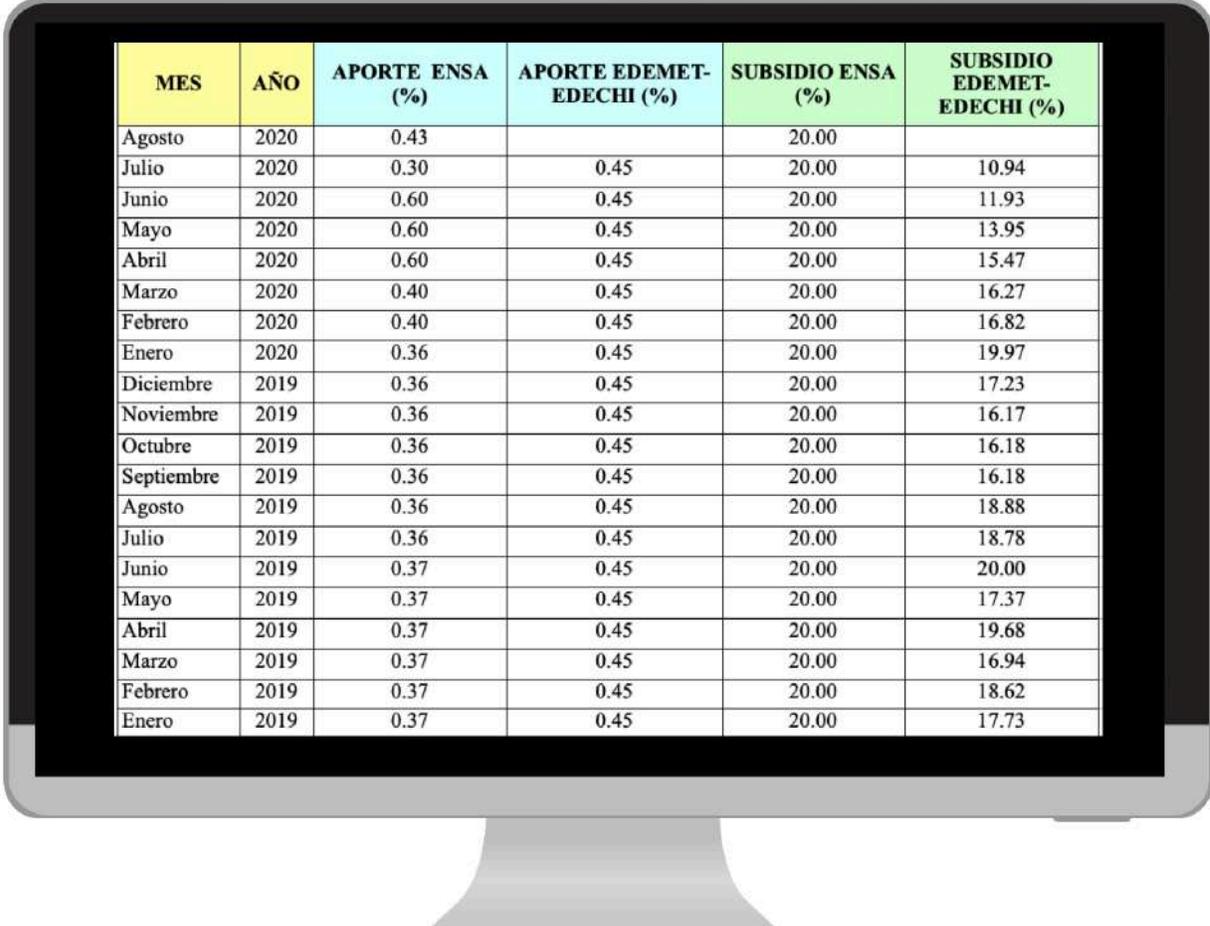
País	Total	Entradas	Salidas
Colombia	137,763	69,425	68,338
Costa Rica	102,559	49,240	53,319
Venezuela	77,028	41,277	35,751
Canadá	67,936	31,823	36,113
España	37,841	19,513	18,328
Brasil	36,216	18,670	17,546
Cuba	32,843	16,344	16,499
Argentina	31,742	16,022	15,720
Alemania	29,747	13,960	15,787
México	28,693	14,562	14,131
Perú	27,619	13,458	14,161
Francia	27,108	13,349	13,759
Nicaragua	24,579	13,543	11,036
Ecuador	21,765	10,837	10,928
Italia	20,607	10,119	10,488
Guatemala	16,479	8,371	8,108
El Salvador	14,108	7,402	6,706
Gran Bretaña	13,143	6,241	6,902
Chile	12,461	6,149	6,312
Honduras	12,356	6,323	6,033
Países Bajos (Holanda)	12,122	6,202	5,920
China	8,950	4,334	4,616
Rep. Dominicana	8,023	4,431	3,592
Suiza	7,373	3,371	4,002
Filipinas	6,667	3,065	3,602
India	6,035	2,953	3,082
Uruguay	5,720	3,288	2,432
Bolivia	5,616	2,913	2,703
Israel	5,125	2,465	2,660
Otros	78,411	37,127	41,284

En el ejemplo anterior se nos presenta un cuadro de movimiento migratorio en Panamá en donde podemos apreciar que el desglose está a nivel de países, lo que nos podría indicar que su nivel de desglose es el más granulado posible, excepto por el hecho de que, en la última fila, en lugar de tener como dato un país nos encontramos con el dato “otros” que nos indica que es un conjunto de países, lo que representa una línea de consolidados, colocándolo en el nivel de desglose 2. Para que estos datos sean de mayor nivel de desglose deberían estar granulados hasta nivel de país, sin excepciones.

- iii. **Nivel de desglose 3:** En este formato la información se encuentra totalmente desagregada y seccionada con el mayor grado de granularidad posible. Continuando con el ejemplo de la división política se entiende en el nivel 3 el conjunto de datos que está desglosado por provincia, distrito y corregimiento, y cualquier otra división sectorial más precisa o específica. Para el caso del conjunto de datos relativos al presupuesto de gastos, el mismo debe tener una división estrictamente segregada, por gastos de productos, materiales y suministros, y todas las cuentas que contenga el presupuesto. Otro caso muy común es el desglose relativo a los años, en donde para considerarse en el nivel más elevado, debe aportarse información para comparar por varios años consecutivos.

A continuación presentamos un ejemplo de conjunto de datos que muestra una información expresada no sólo por años sucesivos, que permite hacer una comparación periódica, sino que incluye los valores de porcentaje (%) de aporte y porcentaje (%) de subsidio de cada una de las empresas, lo que lo hace más completo y exacto.

Ejemplo No. 8 Nivel de Desglose 3



MES	AÑO	APORTE ENSA (%)	APORTE EDEMET-EDECHI (%)	SUBSIDIO ENSA (%)	SUBSIDIO EDEMET-EDECHI (%)
Agosto	2020	0.43		20.00	
Julio	2020	0.30	0.45	20.00	10.94
Junio	2020	0.60	0.45	20.00	11.93
Mayo	2020	0.60	0.45	20.00	13.95
Abril	2020	0.60	0.45	20.00	15.47
Marzo	2020	0.40	0.45	20.00	16.27
Febrero	2020	0.40	0.45	20.00	16.82
Enero	2020	0.36	0.45	20.00	19.97
Diciembre	2019	0.36	0.45	20.00	17.23
Noviembre	2019	0.36	0.45	20.00	16.17
Octubre	2019	0.36	0.45	20.00	16.18
Septiembre	2019	0.36	0.45	20.00	16.18
Agosto	2019	0.36	0.45	20.00	18.88
Julio	2019	0.36	0.45	20.00	18.78
Junio	2019	0.37	0.45	20.00	20.00
Mayo	2019	0.37	0.45	20.00	17.37
Abril	2019	0.37	0.45	20.00	19.68
Marzo	2019	0.37	0.45	20.00	16.94
Febrero	2019	0.37	0.45	20.00	18.62
Enero	2019	0.37	0.45	20.00	17.73

- Transparencia de la evaluación:** Las evaluaciones individuales de cada

conjunto de datos, en cuanto a la facilidad de publicación y la calidad, deben ser tomadas en consideración por el oficial de información, al momento de hacer la selección del conjunto de datos a publicar. Éste constituye un insumo interno del equipo de datos abiertos de la institución, antes de la publicación. Como uno de los elementos de la transparencia, se hace necesaria la inclusión en el cuadro de información que se publica junto al conjunto de datos, la frecuencia de actualización de datos, la cual debe insertarse en la descripción del conjunto que encontramos en el portal nacional de datos abiertos en el apartado “información adicional”. A continuación, presentamos un ejemplo del contenido completo que debe incorporarse en el cuadro de información adicional del conjunto de datos, al publicar en el portal nacional de datos abiertos.

Modelo de cuadro de información de transparencia del conjunto de datos.

Título	ASEP - volumen de agua comprada del IDAAN a la ACP y APSA, hasta octubre 2019
Versión	1.0
Autor	Departamento de Aguas Potables y Aguas Servidas
Email del Autor	Departamento de Aguas Potables y Aguas Servidas (poner link al correo electrónico)
Mantenedor	Juan Pérez
Email del mantenedor	xxxxx@asep.gob.pa

Es recomendable que la institución cuente con un correo específico datosabiertos@dominio de la institución, para que la información o comunicación del usuario del dato, al momento de ponerse en comunicación con el mantenedor, llegue directo al oficial de información o a alguien del equipo de datos abiertos asignado a esta tarea.

- Actualización:** Que un dato se encuentre actualizado es uno de los indicadores de calidad del dato de mayor importancia, toda vez que el mismo es el resultado de un proceso de revisión continua de lo que se publica. Que un dato esté actualizado implica lograr que se encuentre lo más al día posible, a la realidad presente o a la adaptación de los tiempos que corren con respecto a la información. Las instituciones deben cumplir con los períodos de refrescamiento de sus conjuntos de datos en las fechas en las que se han comprometido y anunciado.

A continuación, presentamos un cuadro en el cual podemos evaluar la calidad de los datos, fundamentado en cada uno de los indicadores revisados con anterioridad.

Modelo de cuadro de evaluación de la actualización del conjunto de datos.

INDICADOR	NIVEL	VALOR	TOTAL
1. ESTRUCTURA	No estructurados	1	
	Semiestructurados	2	
	Estructurados	3	
2. COMPLETITUD	Carece de atributos o se repite en otro conjunto de datos	1	
	Cuenta con la mayoría de los atributos	2	
	Cuenta con todos los atributos	3	
3. IMPORTANCIA	Prioridad 1	1	
	Prioridad 2	2	
	Prioridad 3	3	
4. DESGLOSE	Nivel 1	1	
	Nivel 2	2	
	Nivel 3	3	
5. TRANSPARENCIA EN LA EVALUACIÓN Frecuencia de actualización Puntaje facilidad de publicación Puntaje calidad de publicación	Se publica uno de los tres factores de evaluación	1	
	Se publica dos de los tres factores de evaluación	2	
	Se publican los tres factores de evaluación	3	
6. ACTUALIZACIÓN	Desactualizado	1	
	Actualizado al periodo anterior	2	
	Actualizado	3	
PUNTUACIÓN	Máximo puntaje	18	

Calificación de evaluación en actualización

Valor entre 1 y 6. Implica que es un conjunto de datos con una mala calidad.

Valor entre 7 y 12. Implica que es un conjunto de datos con una media calidad.

Valor entre 13 y 18. Implica que es un conjunto de datos con una buena calidad.

4.3.4. Formatos utilizados para publicar Datos Abiertos

a. Grados de apertura

Tim Berners-Lee⁴ 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:



⁴<http://5stardata.info/es/>

Una (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 que esté disponible en formato .PDF, .DOC, .ODT, .TXT, entre otros, y que esté bajo una licencia abierta, se considera dato abierto de 1 estrella. Bajo esta categoría de una estrella la información es difícil de manipular.

Dos (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.

Tres (3) Estrellas: Similar al anterior, pero en un formato no propietario. El nivel 3 estrellas implica la publicación de los archivos de los niveles 1 y 2, más el formato no propietario que en este caso debe ser CSV o XML. En estos dos formatos 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 y se prestan para el intercambio entre máquinas.

Cuatro (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 la 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) y se pueden marcar como favoritos.

Cinco (5) Estrellas: Son los conjuntos de datos vinculados con los de otras instituciones, habiéndose dotado de contexto. En la práctica, es la información contenida en la web que a su vez re direcciona 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.

Para profundizar en el tema de formatos, a continuación, se listan enlaces que nos brindan mayor información, incluyendo guías y tutoriales:

- **RDF, SPARQL y Linked Data**

El mismo comprende Guías y tutoriales

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

- **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>

b. Cómo convertir un archivo a Formato Abierto

En este acápite mostraremos cómo convertir el formato de un conjunto de datos para publicarlo en un mayor grado de apertura, lo que posibilita mejorar la calificación del dato abierto. Por ejemplo, si se encuentra en un formato con evaluación 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.

1. Desde un archivo en Excel a CSV

Los archivos en formato **CSV**, por sus siglas en inglés “*comma-separated values*” son un tipo de documento presentado en formato abierto sencillo para representar datos en forma de tabla, en donde las columnas se separan por comas y las filas por saltos de línea. Para ejemplificar cómo hacer la conversión de un documento en formato Excel a formato abierto CSV, tomaremos un cuadro de información relacionada al presupuesto de una institución. La siguiente imagen representa el conjunto de datos sin depurar.

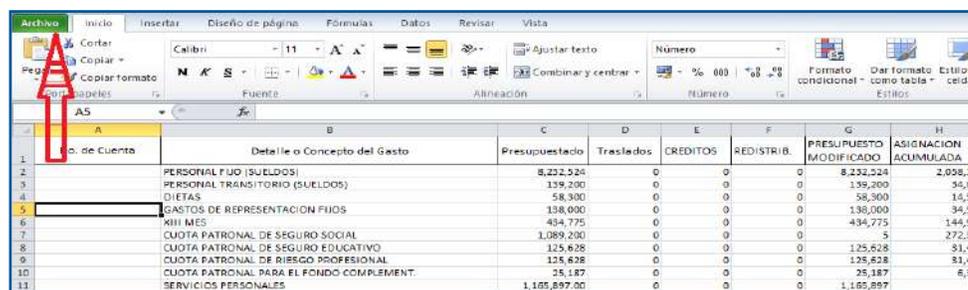
NOMBRE DE LA INSTITUCIÓN												
REPORTE POR TIPO DE PRESUPUESTO Y GRUPO DE GASTO												
DETALLE	PRESUPUESTO LEY	TRASLADOS	CREDITOS	REDISTRIB.	PRESUPUESTO MODIFICADO	ASIGNACION ACUMULADA	Mes - Año			COMPROMISOS		
							EJECUTADO					
							VALOR ABSOLUTO	% ANUAL	% MENSUAL	VALOR ABSOLUTO	% ANUAL	% MENSUAL
TOTAL	73.300.000	0	0	2.681.204	73.300.000	13.983.778	7.455.025,74	10,2	53,3	8.183.259,80	11,2	58,5
FUNCIONAMIENTO	69.302.904	0	0	883.154	69.302.904	12.133.715	6.510.363,97	9,4	53,7	6.999.002,52	10,1	57,7
SERVICIOS PERSONALES	10.694.074	0	0	173.706	10.523.345	1.934.539	1.833.354,21	17,4	94,8	1.833.354,21	17,4	94,8
SUELDOS	8.252.989	0	0	0	8.252.989	1.376.698	1.296.490,50	15,7	94,2	1.296.490,50	15,7	94,2
1 PERSONAL FIJO (SUELDOS)	8.106.589	0	0	0	8.106.589	1.351.098	1.275.250,50	15,7	94,4	1.275.250,50	15,7	94,4
2 PERSONAL TRANSITORIO (SUELDOS)	146.400	0	0	0	146.400	25.600	21.240,00	14,5	83,0	21.240,00	14,5	83,0
DIETAS	20.000	0	0	1.600	20.000	4.000	2.700,00	13,5	67,5	2.700,00	13,5	67,5
GASTOS DE REPRESENTACION FIJOS	138.000	0	0	0	138.000	23.000	19.000,00	13,8	82,6	19.000,00	13,8	82,6
XIII MES	404.530	0	0	0	404.530	134.845	131.055,04	32,4	97,2	131.055,04	32,4	97,2
CONTRIBUCIONES A LA SEGURIDAD SOCIAL	1.343.428	0	0	0	1.343.428	223.890	214.426,69	16,0	95,8	214.426,69	16,0	95,8
1 CUOTA PATRONAL DE SEGURO SOCIAL	1.071.307	0	0	0	1.071.307	178.546	175.601,67	16,4	98,4	175.601,67	16,4	98,4
2 CUOTA PATRONAL DE SEGURO EDUCATIVO	123.690	0	0	0	123.690	20.608	19.486,14	15,8	94,6	19.486,14	15,8	94,6
3 CUOTA PATRONAL DE RIESGO PROFESIONAL	123.690	0	0	0	123.690	20.616	15.686,06	12,7	76,1	15.686,06	12,7	76,1
4 CUOTA PATRONAL PARA EL FONDO COMPLEMENTARIO	24.741	0	0	0	24.741	4.120	3.652,82	14,8	88,7	3.652,82	14,8	88,7
OTROS SERVICIOS PERSONALES	555.127	-170.731	0	172.106	364.396	1.375	0,00	0,0	0,0	0,00	0,0	0,0

Previo a la conversión es necesario depurar la tabla de todos los elementos que puedan obstaculizar su uso de manera automatizada; estos elementos son las líneas y columnas en blanco, los sombreados, caracteres en negritas, itálicas, colores, los títulos, encabezados, celdas combinadas y cualquier otro tipo de formateo. Las filas o columnas con totales o porcentajes tampoco son requeridas para el formato CSV, solo los datos crudos. La siguiente tabla es el resultado del trabajo de depuración de la tabla anterior.

	A	B	C	D	E	F	G	H	I	J	K
1	Detalle	Presupuestado	Traslados	CREDITOS	REDISTRIB.	PRESUPUESTO MODIFICADO	ASIGNACION ACUMULADA	EJECUTADO (ABS)	COMPROMISOS	SALDO DEL DEVENGADO	PAGADO ORIGINAL
2	PERSONAL FIJO (SUELDOS)	8,232,524	0	0	0	8,232,524	2,058,132	1,909,981.07	1,909,981.07	0.00	1,909,981.07
3	PERSONAL TRANSITORIO (SUELDOS)	139,200	0	0	0	139,200	34,800	32,591.67	32,591.67	0.00	32,591.67
4	DIETAS	58,300	0	0	0	58,300	14,578	2,400.00	2,400.00	0.00	2,400.00
5	GASTOS DE REPRESENTACION FIJOS	138,000	0	0	0	138,000	34,500	31,000.00	31,000.00	0.00	31,000.00
6	XIII MES	434,775	0	0	0	434,775	144,926	130,526.65	130,526.65	0.00	130,526.65
7	CUOTA PATRONAL DE SEGURO SOCIAL	1,089,200	0	0	0	5	272,308	254,439.99	254,439.99	0.00	254,439.99
8	CUOTA PATRONAL DE SEGURO EDUCATIVO	125,628	0	0	0	125,628	31,418	28,954.51	28,954.51	0.00	28,954.51
9	CUOTA PATRONAL DE RIESGO PROFESIONAL	125,628	0	0	0	125,628	31,415	23,347.20	23,347.20	0.00	23,347.20
10	CUOTA PATRONAL PARA EL FONDO COMPLEMENT.	25,187	0	0	0	25,187	6,306	5,372.43	5,372.43	0.00	5,372.43
11	SERVICIOS PERSONALES	1,165,897.00	0	0	0	1,165,897	0	0.00	0.00	0.00	0.00
12	SUELDOS (PLANILLA)	20,000.00	0	0	0	20,000	2,000	0.00	0.00	0.00	0.00

Luego de la depuración de la tabla, pasamos al paso de generar el formato CSV mediante la instrucción “guardar como” y buscamos el formato CSV (delimitado por comas), como se muestra en los siguientes pasos:

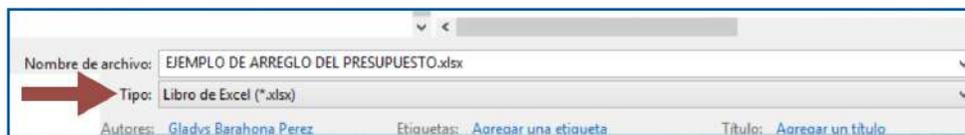
1. Damos click en “Archivo”



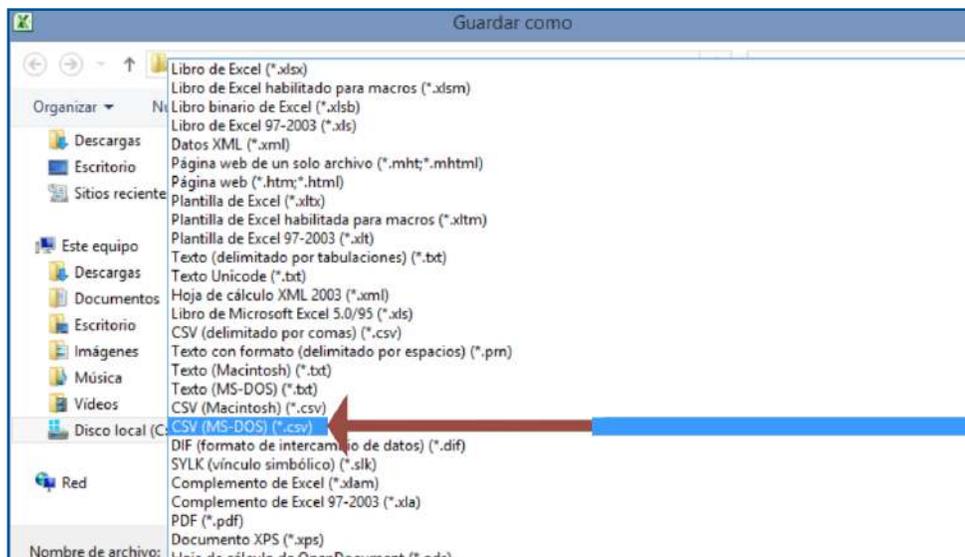
2. Seleccionamos “guardar como”



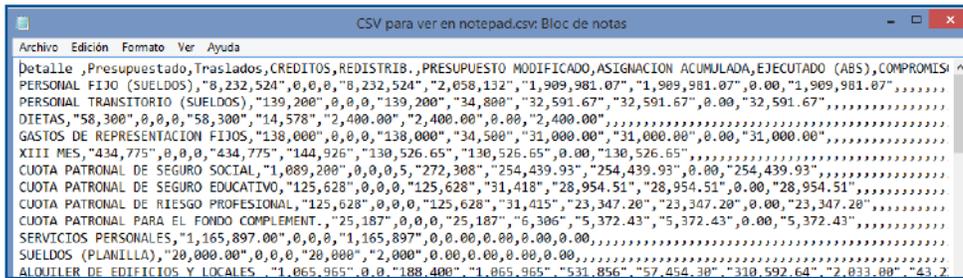
3. Selecciona “tipo”



4. Selecciona CSV (MS-DOS)



Para verificar nuestra conversión abrimos el archivo resultante (CSV) con el bloc de notas (Notepad) y veremos una tabla plana separada por comas o espacios en blanco, como la que se muestra en la siguiente imagen.



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.

2. Desde un Gestor de Bases de Datos

Para gestionar información en formato abierto desde una base de datos es necesario adaptarse a un meta lenguaje que permita el intercambio de información entre diferentes plataformas. El Formato XML es una adaptación del “Standard Generalized Markup Language”, es decir, un lenguaje que permite organizar y etiquetar documentos. No es un lenguaje en sí mismo, sino un sistema que permite definir el lenguaje de acuerdo con las necesidades.

Para la extracción de datos desde bases de datos al formato XML podemos hacerlo de diversas formas, tal como ejemplificamos a continuación:

- **Ejemplo 1: Exportación de datos desde Oracle al formato XML**

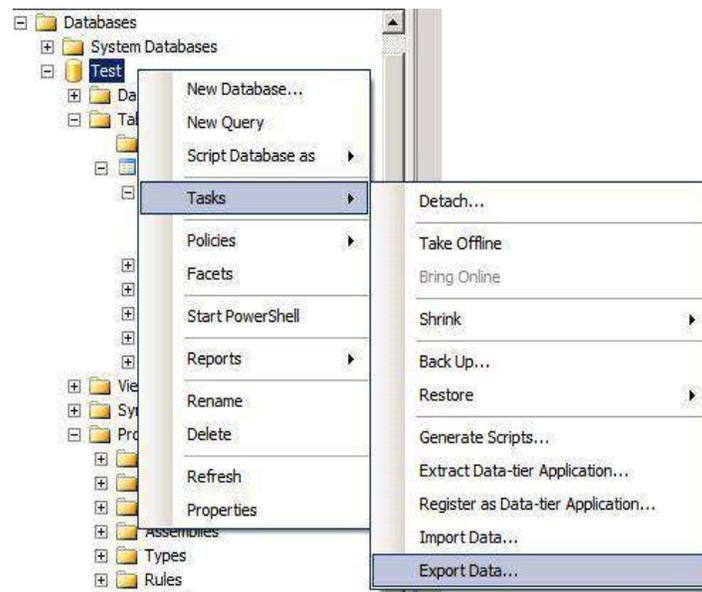
Oracle no cuenta con una interfaz gráfica que permita realizar una exportación de datos a XML directamente, es necesario crear un procedimiento que lea desde una estructura de tabla o vista y genere el documento XML para esos datos.

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:

Results	View	Tools	Help
	Find in Results ...	Spreadsheet Autosize	▶
	Clear Spreadsheet	Blnd Variable Cursors	▶
	Export results to Excel. Ctrl+Alt+E	Export results to a CSV file ...	
	Export results to an XML file ...		

- **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. Se va a administración, en la carpeta base de datos entramos a tareas y seleccionamos exportar datos. La siguiente es una imagen del menú mencionado:



En la 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 para 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	<i>registro 1</i>
2	<i>registro 2</i>
3	<i>registro 3</i>
4	<i>registro 4</i>

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>
```

4.3.5. Aplicar una licencia abierta a los conjuntos de datos a publicar

Tomando en consideración que el objetivo de la publicación de los datos en formato abierto es la reutilización por cualquier persona y para cualquier propósito, se hace necesario establecer de forma clara las condiciones que deben regir la utilización de los conjuntos de datos publicados por cada institución; estas condiciones están determinadas por las licencias de uso, mediante las cuales se dan a conocer a los usuarios las condiciones en materia de propiedad intelectual bajo las cuales podrán ser utilizados los datos disponibles en un sitio Web.

Sin una licencia de uso, los datos no podrían ser realmente utilizados con las características de datos abiertos, y su reutilización por emprendedores e innovadores sería imposible.

Existen licencias ya desarrolladas para la apertura de datos que pueden ser utilizadas libremente y son las licencias “*Creative Commons*”. Estas licencias permiten que quienes crean un determinado producto (en este caso, un conjunto de datos) puedan retener los derechos intelectuales sobre los mismos, sin que esto impida la copia, distribución y distintas formas de utilización.⁵

Para efectos del Portal DAP www.datosabiertos.gob.pa se utilizarán dos tipos de licencia de la organización Creative Commons: **Reconocimiento 4.0 Internacional (CC BY 4.0)** y **CC0 o CCZero** para los conjuntos de datos que sean publicados.

A continuación, se explica de forma breve estas dos licencias:

- **Licencia Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0):**⁶ Permite copiar y redistribuir el material en cualquier medio o formato, así como realizar adaptaciones, re mezclar, transformar y construir sobre el material para cualquier propósito, incluso comercial. Estas libertades de uso se otorgan bajo los siguientes términos:
 - En cuanto a la atribución, debe otorgarse el crédito correspondiente, proporcionar un enlace a la licencia e indicar si se realizaron cambios.

⁵ <https://biblioguias.cepal.org/c.php?g=495473&p=4398203>

⁶ <https://creativecommons.org/licenses/by/4.0/>

- Sin restricciones adicionales, es decir, no pueden aplicarse términos legales o medidas tecnológicas que restrinjan legalmente que otros hagan con los datos cualquier cosa que permita la licencia.

Se pueden revisar los términos completos y códigos legales de la licencia en el siguiente link: <https://creativecommons.org/licenses/by/4.0/legalcode>

- **Licencia Creative Commons CC0 1.0 Universal (CC0 1.0) Dedicación de dominio público⁷:** Permite a los científicos, educadores, artistas y otros creadores y propietarios de contenido protegido por derechos de autor o base de datos, renunciar a esos intereses y derechos en sus obras y, por lo tanto, colocarlas de la manera más completa posible para el dominio público, de modo que otros puedan construir libremente, mejorar y reutilizar las obras para cualquier propósito sin restricciones bajo derechos de autor o leyes de bases de datos. Puede copiar, modificar, distribuir y realizar el trabajo, incluso con fines comerciales.

Se pueden revisar los términos completos y códigos legales de la licencia en el siguiente link: <https://creativecommons.org/publicdomain/zero/1.0/legalcode>

A diferencia de las licencias de CC que permiten a los titulares de los derechos de autor elegir entre una variedad de permisos mientras retienen sus derechos de autor, CC0 permite otra opción más: la de optar por no recibir protección de derechos de autor y bases de datos, y los derechos exclusivos otorgados automáticamente a los creadores: el " sin derechos reservados " alternativa a otras licencias.

A continuación observamos la licencia en el portal nacional de datos abiertos.

The image shows a screenshot of the ASEP (Autoridad Nacional de los Servicios Públicos) portal. On the left, there is a sidebar with the ASEP logo and social media links. The main content area displays a Creative Commons Attribution license. To the right, there is a table titled 'Concesiones en trámite' with the following data:

Campo	Valor
Título	ASEP, Concesiones en trámite
Descripción	Concesiones de generación de energía eléctrica utilizando la energía hidráulica.
Fuente	https://www.asep.gob.pa/?page_id=12866
Versión	1.0
Autor	Unidad de Gestión Ambiental
Email del Autor	Unidad de Gestión Ambiental
Mantenedor	Michele E. Moreno
Email del Mantenedor	Michele E. Moreno

⁷ <https://creativecommons.org/publicdomain/zero/1.0/>

4.3.6. Inclusión de metadatos en los conjuntos de datos a publicar

En términos generales los metadatos son aquellos datos que describen el contenido de los archivos, o la información contenida en ellos. Son de gran importancia en el proceso de gestión y gobernanza de los datos, ya que facilitan su búsqueda y análisis, así como su estandarización.

Para la mejor gestión del conjunto de datos, es necesario que se describa su contenido, información y otras características, a través de los metadatos.

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 cuales hacen una descripción específica del conjunto de datos, describiendo su contenido, información y otras características, como por ejemplo la indicación del nombre del conjunto de datos, una descripción, etiquetas, información legal, la licencia usada, la cobertura temporal y geográfica, fecha de publicación, última actualización. Específicamente para el caso del Portal de Datos Abiertos de Panamá, los metadatos están divididos en metadatos del conjunto de datos, y metadatos de los archivos, en los que se detalla la fecha de publicación y actualización.

A continuación, un ejemplo de los metadatos, de un conjunto de datos existente en el portal nacional de datos abiertos.

Información Adicional

Campo	Valor
Título	ASEP, Concesiones en trámite
Descripción	Concesiones de generación de energía eléctrica utilizando la energía hidráulica.
Fuente	https://www.asep.gob.pa/?page_id=12866
Versión	1.0
Autor	Unidad de Gestión Ambiental
Email del Autor	Unidad de Gestión Ambiental
Mantenedor	Michele E. Moreno
Email del Mantenedor	Michele E. Moreno

Información adicional

Campo	Valor
Última actualización	Septiembre 3, 2020
Creado	Septiembre 9, 2019
Formato	XLSX
Licencia	Creative Commons Attribution
Nombre	Concesiones en Trámite
Descripción	Concesiones de generación de energía eléctrica utilizando la energía hidráulica.

Para el caso en que se requiera determinar la estructura interna de los esquemas de datos, como en el caso de publicaciones en formatos CSV, donde se debe suministrar una descripción del esquema, por ejemplo, suministre una tabla de descripción de campos. Ver ejemplo de cuadro que pudiera agregarse a los metadatos presentados:

Nombre del Atributo	Descripción	Tipo de Dato	Recurso Relacionado	Información Adicional
Descripción exacta de la columna tal cual aparece en los archivos de los conjuntos de datos. Ejemplo: Código Prov	Descripción detallada del significado del atributo orientada a las personas que pretenden utilizar el atributo. Ejemplo: Prov de Residencia	Numérico, texto, fecha, hora, moneda, Ejemplo: Numérico	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. Ejemplo: Código Prov (código de provincia) apunta al archivo Provincia.xls	Cualquier información adicional sobre el atributo que complemente la información básica. Ejemplo en el archivo provincia.xls están los nombres de las provincias asociadas a dicho código. Ejemplo: <ol style="list-style-type: none"> 1. Panamá 2. Colón 3. Coelé ...

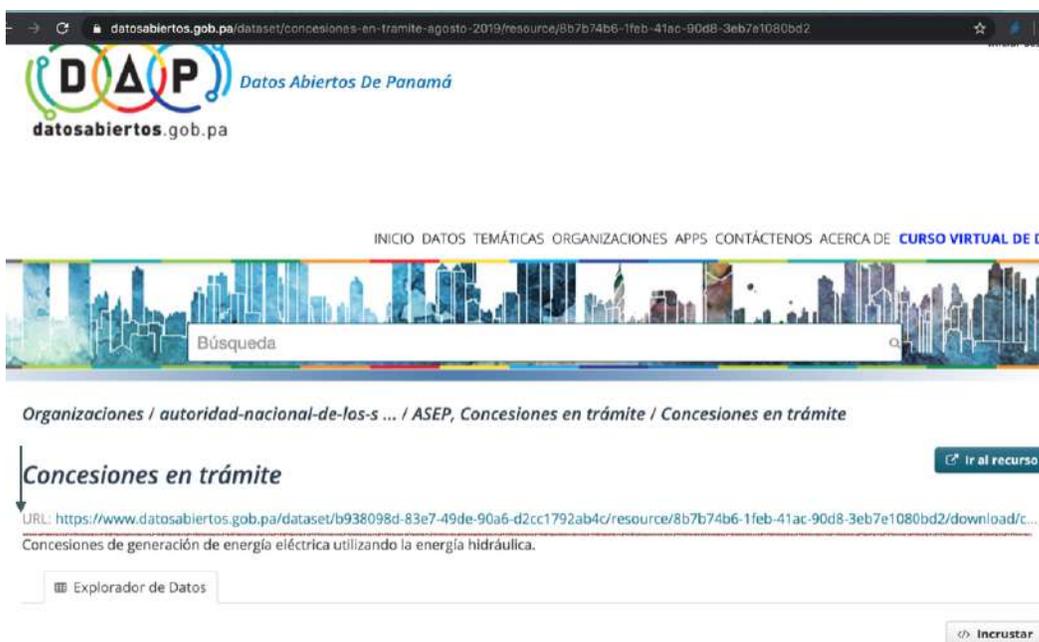
4.3.7. Asignación de URL o URI a los conjuntos de datos de 4 y 5 Estrellas

Con la finalidad que el recurso (dato-información) pueda ser identificado y ubicado fácilmente por los usuarios del portal, es necesario que se especifique el “Localizador de Recurso Uniforme o URL” (por sus siglas en inglés) del conjunto de datos.

Si es el caso de conjuntos de datos catalogados como de 4 y 5 estrellas, se debe especificar el nombre del conjunto de datos conocido como Identificador de Recurso Uniforme (URI por sus siglas en inglés) de acceso.

En ambos casos se corresponden a la cadena de caracteres que identifican los recursos en la red de forma inequívoca. La diferencia entre ambos es que los URI son fijos, y los URL pueden variar en el tiempo.

Ver ejemplo de URL en la fig. abajo que ejemplifica un conjunto de datos existentes en el portal nacional:



Para el caso de la conversión de datos en formato abierto, en donde se hace una depuración de la información contenida en las tablas, podemos usar el título de la hoja de cálculo como parte de la URI y de esa manera será posible identificar fácilmente el archivo.

Por ejemplo:

http://www.miservidordedatos.com/resultadosfinales/cuadros/PRESUPUESTO_INSTITUCIONAL_MES_AÑO

4.3.8. Carga de datos al sitio institucional y al portal nacional

Una vez que se han preparado los conjuntos de datos, utilizando las mejores prácticas de selección del conjunto de datos, evaluación del estado del conjunto de datos, estandarización de formato, y demás acciones encaminadas a responder a las necesidades de los grupos de interés y ciudadanía en general, debemos poner los datos abiertos en línea.

En primer lugar, se debe ubicar 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 agregar 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 publicados en su sitio web en el formato PDF, 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.

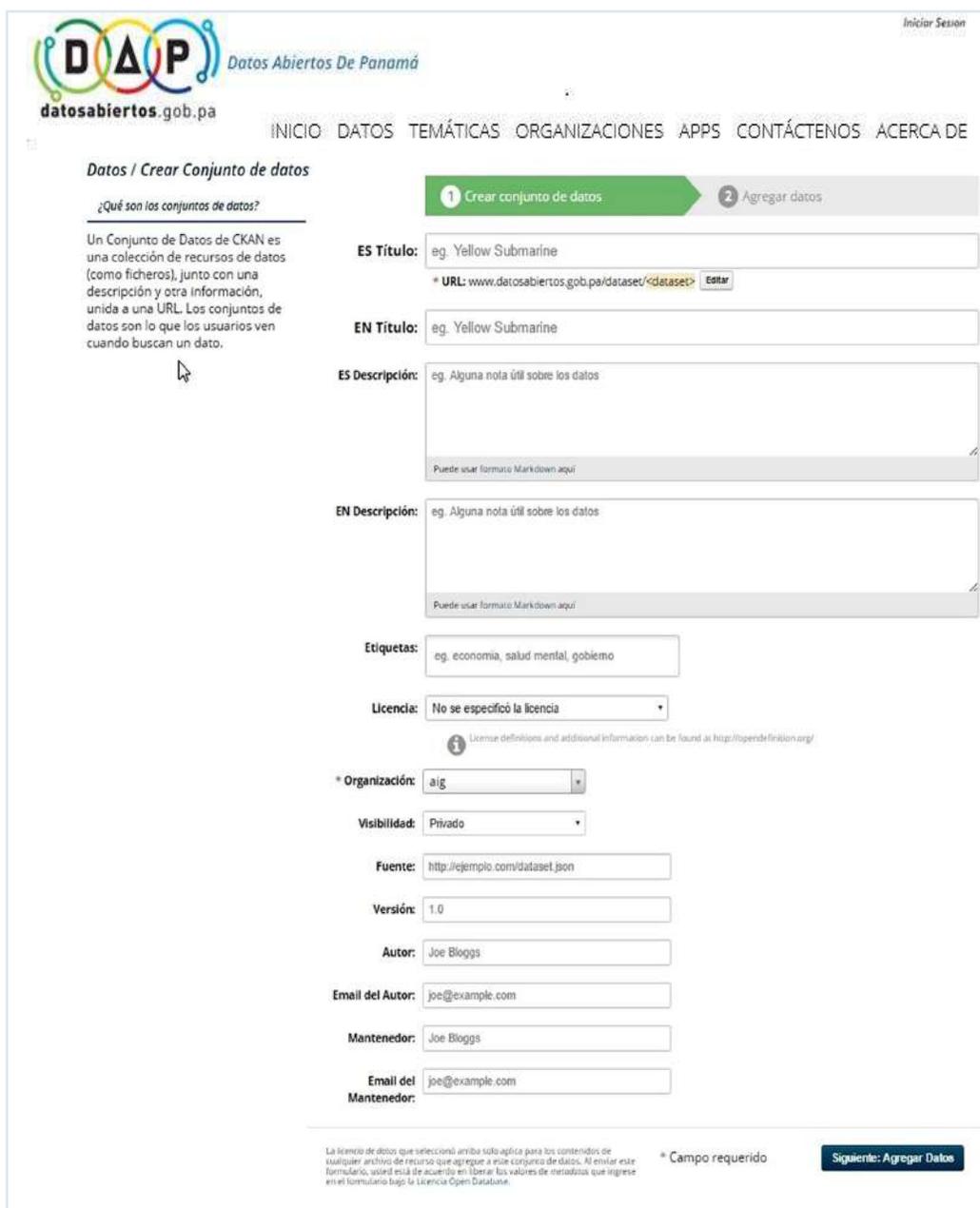
Una vez que los datos abiertos han sido publicados en el sitio web de su institución se debe entrar al portal nacional de datos abiertos de Panamá (DAP) <https://www.datosabiertos.gob.pa> para publicar los nuevos conjuntos de datos.

Para ingresar al portal DAP 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)⁸. 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.

⁸ <https://okfn.org/>

A continuación, se presenta un ejemplo del formulario de captura de los metadatos del conjunto de datos (el proceso de publicación de datos abiertos se explica en la **“Guía de Publicación en el Portal DAP”**, documento de insumo que es entregada a la institución al momento que se le asigna el usuario y contraseña para ingreso al portal).

Modelo de formulario de publicación de metadatos de un 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: www.datosabiertos.gob.pa/dataset/<dataset>

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: eg. economia, salud mental, gobierno

Licencia: No se especificó la licencia
License definitions and additional information can be found at <http://opendefinition.org/>

* **Organización:** aig

Visibilidad: Privado

Fuente: <http://ejemplo.com/dataset.json>

Versión: 1.0

Autor: Joe Bloggs

Email del Autor: joe@example.com

Mantenedor: Joe Bloggs

Email del Mantenedor: joe@example.com

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 ingresó en el formulario bajo la Licencia Open Database.

* Campo requerido

Modelo de formulario de publicación de metadatos de los archivos

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:

a. Actualización y liberación de nuevos conjuntos de datos

Es imprescindible mantener actualizado el portal DAP acorde a la realidad de publicación de cada institución para que los datos publicados constituyan un insumo de valor para los usuarios.

Este punto indica el compromiso que adquiere la institución frente a los usuarios de la comunidad al determinar un programa de actualización y liberación, fijando un calendario de publicación de datos abiertos que incluye igualmente la publicación de nuevos conjuntos de datos. Se recomienda que la programación tenga un alcance anual y se publique en el sitio web de la institución para orientar a la comunidad.

A continuación, presentamos una plantilla, que puede ayudar para la programación recomendada:

Modelo de plantilla de actualización

Institución	ASEP
Oficial de Información	Nombre del Oficial de Información
Correo oficial del Oficial de Información	xxxxxxxx@asep.gob.pa
Conjunto de datos	Volumen de Agua comprada al IDAAN y a la ACP y APSA periodo xxxx
Descripción de los Datos	Volumen de agua en millones de galones comprada por el IDAAN a la Autoridad del Canal de Panamá (ACP) y Aguas de Panamá, S.A. (APSA) distribuida en Panamá Metro
Versión	Actualización a xxxxxx /Nuevo conjunto de datos
Periodicidad de la Actualización / Publicación	Tremiestral
Fecha tentativa de publicación.	DD-MM-AA

4.3.9. Limpieza del conjunto de datos

Es fundamental que los datos estén limpios y sin errores para trabajar con ellos y generar nuevo valor. Los datos limpios son coherentes, no contienen duplicaciones y están listos para que una máquina los procese. Cuando están limpios resulta fácil combinarlos con otros conjuntos de datos y obtener mucha más información.

a. Errores frecuentes

Existe una serie de errores típicos que se cometen al momento de publicar los datos, a los que deberíamos estar atentos en cualquier conjunto de datos con el que estemos trabajando:

1. **Formatos erróneos:** por ejemplo, un conjunto de datos en donde las fechas pueden estar escritas de forma incoherente.
2. **Variación de la escritura:** Generalmente se intenta ahorrar tiempo a la hora de introducir datos abreviando términos. Cuando estas abreviaturas no son coherentes, se pueden generar errores en el conjunto de datos. Las diferencias en el uso de las mayúsculas, los espacios y los géneros de los adjetivos también pueden dar lugar a este tipo de errores.
3. **Duplicación de registros:** Un registro duplicado ocurre cuando el mismo dato se ha introducido más de una vez. Los registros duplicados suelen producirse cuando se combinan conjuntos de datos o cuando no se sabe que ya existía una entrada.
4. **Redundancia de los datos:** Los datos redundantes son cualquier información que no es relevante para trabajar con el conjunto de datos. Normalmente, un conjunto de datos se crea para un fin concreto que requiere detalles, y alguno de estos datos es posible que no se necesiten. Un caso típico de datos redundantes suelen ser las filas con cantidades totales. Otras veces se combinan columnas de datos o se replican para facilitar su legibilidad.
5. **Mezcla de las escalas numéricas:** Los valores numéricos de los conjuntos de datos suelen usar distintas escalas para que resulte más fácil leerlos. En los conjuntos de datos de los presupuestos, por ejemplo, las unidades suelen estar expresadas en millones, así 1,200,000 suele convertirse en 1,2 M. Sin

embargo, las cantidades más pequeñas como 800,000 continúan escribiéndose en su forma extendida. Para una máquina esto significa que leen la cifra extendida como 1,2... lo que genera errores.

6. Mezcla de rangos: Los datos a veces se miden en rangos, como la edad o el salario. Para que una máquina pueda entender estos rangos es importante separar los valores altos de los bajos.

b. Herramientas recomendadas para limpieza de datos

1. Open Refine:⁹

Open Refine es una herramienta de software diseñada para tratar datos que no se han limpiado. La herramienta consiste en un navegador basado en columnas que le permite solucionar errores en todo un conjunto de datos abiertos mediante una sola acción.

Entre los errores que pueden solucionarse se encuentran:

- Formatos de fecha
- Variación de la escritura
- Duplicación de registros
- Redundancia de los datos
- Mezcla de las escalas numéricas
- Mezcla de rangos

2. Programas de hojas de cálculo:

También encontramos programas de hojas de cálculos, con los cuales es más fácil solucionar algunos errores como los que detallamos a continuación:

- Errores ortográficos
- Redundancia de los datos
- Comprobación de las cifras

⁹ openrefine.org

4.3.10. Divulgación del plan de apertura de Datos Abiertos

Cada institución debe divulgar su plan de publicación de datos abiertos y el programa de actualización. Esto se puede realizar con herramientas ya existentes como, por ejemplo, en el sitio web institucional (página de datos abiertos y sección de noticias, por ejemplo) y en 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.

4.3.11. Promoción en el uso de Datos Abiertos

Cuando abrimos los datos es necesario construir comunidades que hagan uso de la información, por ejemplo, desarrolladores o visualizadores son esenciales en la primera etapa de apertura de datos, ya que a través de ellos se pueden crear grandes herramientas que, a su vez, lleguen al público en general.

Junto a esto, existen organizaciones, comunidades y especialistas en el uso y explotación de los datos abiertos, por ejemplo, los investigadores, periodistas de datos, ONG y comunidades de desarrolladores informáticos.

Por eso, con el fin de mantener activo este ecosistema de uso, se proponen las siguientes actividades:

Nombre de la Iniciativa	Descripción	Resultados	Impacto en la comunidad
IDEATONES	Sesiones colectivas, de preferencia con personas participantes multidisciplinas, en donde se analizan datos sobre una problemática en particular con tal de generar ideas que contribuyan en su solución; idealmente, las ideas con mayor potencial de impacto y viabilidad de implementación deben ser apoyadas, posteriormente, para profundizar en análisis, realizar prototipos e implementarlas en la vida cotidiana.	Las ideas generadas pueden abarcar desde propuestas de políticas públicas, intervenciones sociales hasta potenciales aplicaciones tecnológicas.	Convivencia, análisis y aprendizaje mutuo entre especialistas (analistas de datos, visualizadores, expertos en temáticas, programadores, trabajadoras sociales, funcionarias públicas, etc.); generación de alianzas multidisciplinas en relación con una problemática específica; involucramiento de análisis basado en datos entre sectores interesados e inclusive especializados sin uso de datos.
MAPATONES	Son sesiones colectivas de levantamiento de datos geoespaciales.	La recopilación de datos geoespaciales relevantes para ciertas comunidades o temas que previamente eran inaccesibles.	Generan el involucramiento de personas con diversas capacidades técnicas quienes incrementan su solidaridad con la causa de que los datos sean generados.

<p>EXPEDICIÓN DE DATOS</p>	<p>Sesiones colectivas de preferencia multidisciplinas en donde grupos de hasta 8 personas analizan datos relacionados con alguna problemática o situación de su interés con tal de llegar a conclusiones sobre la misma.</p>	<p>Análisis detallado basado en datos sobre un tema, situación o problemática; aprendizaje de tácticas y herramientas para el análisis de datos.</p>	<p>Socialización de conjuntos de datos, iniciativas y vinculación entre comunidades y perfiles.</p>
<p>HACKATONES</p>	<p>Espacios de hasta 48 horas en donde las personas participantes (de preferencia multidisciplinas) desarrollan soluciones tecnológicas basadas en datos; comúnmente se otorgan premios para apoyar que las soluciones ganadoras puedan ser terminadas y llevadas a usuarios finales; en un escenario óptimo, los equipos ganadores son incubados y apoyados por organizaciones clave para la implementación del proyecto.</p>	<p>Visualizaciones, plataformas de información, sitios web, aplicaciones móviles, servicios web de datos, dispositivos mecánicos para generación de datos, etc.; aprendizaje entre las personas participantes en torno a trabajo en equipo, acceso y procesamiento de datos, uso de plataformas tecnológicas, interfaces de persona usuaria y definición de productos basados en datos.</p>	<p>Aumento de colaboración, aprendizaje mutuo y sinergias entre las personas del equipo; competencia creativa, tecnológica y analítica entre equipos de personas participantes; alianzas posteriores al evento para realizar otros proyectos.</p>
<p>MEDIATONES</p>	<p>Espacios de hasta 48 horas en donde personas participantes (de preferencia multidisciplinas) realizan proyectos comunicacionales basados en datos ya sea en periodismo, activismo o comunicación social; comúnmente se otorgan premios para apoyar que los proyectos ganadores puedan ser terminados, publicados, alcanzando a sus audiencias meta; en un escenario óptimo, los equipos ganadores son incubados por especialistas en comunicación, periodismo, datos y tecnología y se cuenta con el compromiso por parte de medios y organizaciones para la publicación de proyectos.</p>	<p>Reportajes periodísticos, visualizaciones, campañas de incidencia y acciones de comunicación social; aprendizaje entre las personas participantes en torno a trabajo en equipo, acceso, procesamiento y visualización de datos, uso de plataformas tecnológicas, interfaces de persona usuaria, diseño gráfico, diseño de la información y definición de productos basados en datos.</p>	<p>Aumento de colaboración, aprendizaje mutuo y sinergias entre las personas del equipo; competencia creativa, tecnológica y analítica entre equipos de participantes; alianzas posteriores al evento para realizar otros proyectos; aumento de la cultura de datos en organizaciones de la sociedad civil, colectivos sociales y periodismo.</p>

4.4. Monitoreo de publicación de datos

4.4.1. A nivel Nacional (Realizado por el Grupo de Trabajo de Datos Abiertos-GTDA)

El monitoreo se realiza para verificar la calidad de los conjuntos de datos con respecto a su potencial reutilización, cantidad, si cuentan con licencia, en qué formato se publicó, entre otros aspectos. Se realizará periódicamente por el Grupo de Trabajo de Datos Abiertos (GTDA) conforme a las funciones que detalla la Política de Datos Abiertos de Gobierno, y será publicado en las páginas web de la ANTAI, AIG y divulgado en los diferentes medios de comunicación.

A continuación, presentamos dos ejemplos del monitoreo de calidad de conjuntos de datos actualmente publicados en el Portal de Datos Abiertos de Panamá, tal como debe operar el monitoreo realizado por el GTDA, con su respectiva retroalimentación.

Ejemplo 1¹⁰:

REGION	2017			
	PRODUCCION	FACTURACION	% DE PERDIDA O AGUA NO CONTABILIZADA	PRODUCCION DIARIA
TOTAL	200,400,016	107,631,496	46.3	549,041.140
ARRAIJAN	14,887,827	7,903,406	46.9	40,788.567
BOCAS DEL TORO	5,394,986	1,823,948	66.2	14,780.784
COCLE	4,621,648	4,167,747	9.8	12,662.049
COLON	17,656,353	7,564,125	57.2	48,373.570
CHIRIQUI	15,932,028	9,172,837	42.4	43,649.392
PANAMA E.Y DARIEN	2,499,850	1,128,178	54.9	6,848.904
HERRERA	3,486,989	3,445,325	1.2	9,553.395
LOS SANTOS	3,012,641	2,839,669	5.7	8,253.811
PANAMA METRO	111,929,274	57,115,340	49.0	306,655.545
PANAMA OESTE	15,529,926	7,919,559	49.0	42,547.742
VERAGUAS	5,448,494	4,551,362	16.5	14,927.381

¹⁰ <https://www.datosabiertos.gob.pa/dataset/asep-analisis-del-volumen-producido-vs-volumen-facturado-de-agua-potable-del-idaan-por-region>

Retroalimentación de la evaluación:

INDICADOR	NIVEL	VALOR	TOTAL
1. ESTRUCTURA	No estructurados	1	
	Semiestructurados	2	
	Estructurados	3	3
2. COMPLETITUD	Carece de atributos o se repite en otro conjunto de datos	1	
	Cuenta con casi todos sus atributos y no se repite en otro conjunto de datos	2	2
	Contiene todos sus atributos	3	
3. IMPORTANCIA	Información de prioridad 1	1	
	Información de prioridad 2	2	
	Información de prioridad 3	3	3
4. DESGLOSE	Nivel de Desglose 1	1	1
	Nivel de Desglose 2	2	
	Nivel de Desglose 3	3	
5. TRANSPARENCIA	Se publica uno de los tres indicadores de la transparencia (2 evaluaciones y frecuencia de la actualización de datos)	1	
	Se publica dos de los tres indicadores de la transparencia (2 evaluaciones y frecuencia de la actualización de datos)	2	
	Se publica tres indicadores de la transparencia (2 evaluaciones y frecuencia de la actualización de datos)	3	3
6. ACTUALIZACIÓN	Desactualizado	1	
	Actualizado al período anterior	2	2
	Actualizado	3	
PUNTUACIÓN	Máximo puntaje	18	14

Este conjunto de datos obtuvo 14 puntos de 18 alcanzando un 78 % de nivel de calidad.

Retroalimentación de la Evaluación:

1. Desglose: Se califica con un nivel de desglose 1 ya que la característica que posee de división política del conjunto de datos solo llega a nivel de

provincias; con algunas áreas geográficas determinadas como Arraiján y Panamá Este.

2. Estructurado: El conjunto de datos está bien estructurado, así que obtiene puntuación 3.
3. Importancia: Se considera información de mayor prioridad nivel 3, ya que el conjunto de datos es relativo a las funciones principales del departamento de la institución y es información que se relaciona con uno de los compromisos de Gobierno Abierto Panamá.
4. Completitud: Entra en un nivel de desglose 2 ya que le faltaba algunos atributos, propiedades o características de un dato completo.
5. Transparencia de evaluación: Consideramos que se publicaron los 2 puntajes de evaluación y la frecuencia de actualización en la información adicional del conjunto de datos.
6. Actualización: Se considera desactualizado ya que el conjunto de datos no llega hasta el último mes completado a la fecha de publicación.

Ejemplo 2¹¹:

Detalle	Presupuesto Ley	Presupuesto Modificado	Asignado Modificado	Compromiso Acumulado	Devengado	Pagado
	(a)	(b)	(c)	(d)	(e)	(f)
Total	3,218,744.00	3,218,744.00	1,601,670.68	1,335,606.93	47,983.83	1,170,893.05
SERVICIOS PERSONALES	2,071,148.00	2,071,148.00	1,027,717.68	908,470.52	22,205.46	886,265.06
Sueldos	1,613,160.00	1,601,611.00	795,031.00	714,954.70	1,006.67	713,948.03
Personal Fijo (Sueldos)	1,559,160.00	1,538,611.00	759,031.00	700,178.02	0.00	700,178.02
Personal Transitorio	0.00	9,000.00	9,000.00	5,250.00	0.00	5,250.00
Personal Contingente	54,000.00	54,000.00	27,000.00	9,526.68	1,006.67	8,520.01
Gastos de Representación	138,000.00	138,000.00	69,000.00	69,000.00	0.00	69,000.00
Gastos de Representación	138,000.00	138,000.00	69,000.00	69,000.00	0.00	69,000.00
Sobretiempo	0.00	1,298.00	1,298.00	1,296.44	373.48	922.96
Sobretiempo	0.00	1,298.00	1,298.00	1,296.44	373.48	922.96
XIII Mes	47,163.00	47,531.00	16,089.68	14,298.21	0.00	14,298.21
XIII Mes	47,163.00	47,531.00	16,089.68	14,298.21	0.00	14,298.21
Contribuciones a la Seguridad Social	272,825.00	275,887.00	139,478.00	102,178.33	20,632.02	81,546.31
Cuota Patronal de Seguro Social	219,589.00	221,267.00	111,472.00	81,719.60	16,429.08	65,290.52
Cuota Patronal de Seguro Educativo	24,198.00	24,403.00	12,307.00	8,892.03	1,833.31	7,058.72
Cuota Patronal de Riesgo Profesional	24,198.00	25,340.00	13,244.00	9,818.28	2,009.31	7,808.97
Cuota Patronal para el Fondo Complementario	4,840.00	4,877.00	2,455.00	1,748.42	360.32	1,388.10
Creditos Reconocidos por Servicios Personales	0.00	6,821.00	6,821.00	6,742.84	193.29	6,549.55
Sueldos	0.00	4,124.00	4,124.00	4,059.67	0.00	4,059.67
Gastos de Representación Fijo	0.00	350.00	350.00	350.00	0.00	350.00
Sobretiempo	0.00	1,327.00	1,327.00	1,326.24	0.00	1,326.24
XIII Mes	0.00	114.00	114.00	113.06	0.00	113.06
Contribuciones a la Seguridad Social	0.00	906.00	906.00	893.87	193.29	700.58
SERVICIOS NO PERSONALES	843,233.00	820,745.00	399,286.00	304,774.25	12,541.87	199,032.29
Alquileres	346,794.00	337,992.00	164,598.00	154,163.81	1,922.79	121,340.36
De Edificios y Locales	290,184.00	290,184.00	145,092.00	145,084.56	1,450.92	119,452.88
De Equipo de Oficina	6,000.00	6,000.00	3,000.00	2,359.35	471.87	1,887.48
Otros Alquileres	50,610.00	41,808.00	16,506.00	6,719.90	0.00	0.00
Servicios Básicos	90,250.00	86,081.00	41,105.00	36,211.93	4,734.91	18,355.10
Agua	1,000.00	500.00	0.00	0.00	0.00	0.00
Aseo	1,000.00	630.00	180.00	180.00	0.00	75.00
Correo	1,000.00	400.00	0.00	0.00	0.00	0.00
Energía Eléctrica	57,770.00	48,315.00	19,429.00	17,712.85	3,668.73	14,044.12
Telecomunicaciones	7,340.00	7,840.00	4,172.00	2,136.35	38.66	547.98
Servicio de Transmisión de datos	13,840.00	20,096.00	13,174.00	12,912.94	1,027.52	2,598.08
Servicio de Telefonía Celular	8,300.00	8,300.00	4,150.00	3,269.79	0.00	1,089.92

¹¹ <https://www.datosabiertos.gob.pa/dataset/tat-ejecucion-presupuestaria-enero-a-junio-2019>

Retroalimentación de la evaluación:

INDICADOR	NIVEL	VALOR	TOTAL
1. ESTRUCTURA	No estructurados	1	
	Semiestructurados	2	
	Estructurados	3	3
8. COMPLETITUD	Carece de atributos o se repite en otro conjunto de datos	1	
	Cuenta con casi todos sus atributos y no se repite en otro conjunto de datos	2	
	Contiene todos sus atributos	3	3
9. IMPORTANCIA	Información de prioridad 1	1	
	Información de prioridad 2	2	
	Información de prioridad 3	3	3
10. DESGLOSE	Nivel de Desglose 1	1	
	Nivel de Desglose 2	2	2
	Nivel de Desglose 3	3	
11. TRANSPARENCIA	Se publica uno de los tres indicadores de la transparencia (2 evaluaciones y frecuencia de la actualización de datos)	1	
	Se publica dos de los tres indicadores de la transparencia (2 evaluaciones y frecuencia de la actualización de datos)	2	
	Se publica tres indicadores de la transparencia (2 evaluaciones y frecuencia de la actualización de datos)	3	3
12. ACTUALIZACIÓN	Desactualizado	1	1
	Actualizado al período anterior	2	
	Actualizado	3	
PUNTUACIÓN	Máximo puntaje	18	15

Este conjunto de datos obtuvo 15 puntos de 18 alcanzando un 83 % de calidad.

1. Desglose: Se califica con un nivel de desglose 2 ya que se tiene un buen desglose de la información, pero no muestra los comprobantes de pago de cada transacción específica.
2. Estructurado: El conjunto de datos está estructurado
3. Importancia: Se considera información de prioridad 3 ya que el conjunto de

datos es relativo a las funciones principales de la institución.

4. Completitud: Se considera que la información está completa ya que posee todos sus atributos.
5. Transparencia de evaluación: Considerados los 2 puntajes de evaluación y la frecuencia de actualización en la información adicional del conjunto de datos.
6. Actualización: Se considera desactualizado ya que el conjunto de datos no llega hasta el último mes completado a la fecha de publicación.

4.4.2. A nivel Internacional

A nivel internacional encontramos varias organizaciones de evaluación de la calidad de los datos publicados por las instituciones gubernamentales, dentro de la política de datos abiertos. Uno de los principales que actualmente evalúa a Panamá es el Open Data Barometer (ODB)¹², el cual constituye una guía de medición global de cómo los gobiernos publican y utilizan datos abiertos para la rendición de cuentas, la innovación y el impacto social.

El mismo es aplicable a los gobiernos que han adoptado la Carta Internacional de Datos Abiertos¹³, firmada por la República de Panamá en el año 2015, y en él se evalúa, además de la madurez en la apertura de datos de cada país, otros criterios como el impacto político, social y económico de los datos, en 15 categorías de datos:

1. Mapas nacionales
2. Datos de Propiedad de la Tierra
3. Datos Detallados del Censo
4. Presupuesto Detallado del Gobierno
5. Datos Detallados del Presupuesto Público
6. Registro de Compañías
7. Legislación
8. Horarios del Transporte público
9. Datos de Comercio Internacional
10. Desempeño del sector salud
11. Datos de Rendimiento de la Educación Primaria o Secundaria
12. Estadísticas de delincuencia
13. Estadísticas ambientales nacionales
14. Resultados de las elecciones nacionales
15. Contratos públicos

¹² https://opendatabarometer.org/?_year=2017&indicator=ODB

¹³ <https://opendatacharter.net/principles-es/>

En su cuarta edición publicada en 2017 Panamá obtuvo una valoración de 30 puntos de 100, resaltando una nula puntuación en el impacto emergente de los datos publicados, lo que evidencia la necesidad de los equipos de trabajo de datos abiertos dentro de cada institución, en publicar conjuntos de datos de valor y calidad para que puedan ser reutilizados por los ciudadanos en general.

La evaluación completa de Panamá puede revisarse en el siguiente link: https://opendatabarometer.org/?_year=2017&indicator=ODB

5. El Futuro del Portal Nacional de Datos Abiertos de Gobierno

El trabajo de los actores activos en el ecosistema de datos abierto a nivel de país, la constante adecuación hacia la inclusión de políticas y estándares internacionales en la materia, así como coadyuvar en la mejora constante a nivel país; esto, aunado a la responsabilidad de los estados, de garantizar el acceso a la información, no solo por temas de datos abiertos, y transparencia, lleva a que anualmente se mantengan planes de mejora a la política de datos abiertos y por ende al portal nacional.

Entre las acciones establecidas a corto plazo para el portal de datos abierto de Panamá encontramos:

1. Mejoras al portal DAP que incluyen, entre otras, la accesibilidad para personas con discapacidad; esto debe ser una responsabilidad estatal como un elemento de prestación mínima de servicios, garantizando que sus páginas web sean accesibles, sin limitaciones algunas por razones de discapacidad. La Autoridad Nacional de Transparencia y Acceso a la Información (ANTAI) está realizando importantes esfuerzos para ir hacia políticas de inclusión dentro de los sitios Web institucionales y que ello se constituya en una herramienta adicional para el monitoreo de la transparencia que ANTAI realiza mensualmente a las instituciones¹⁴; por lo que todas deben trabajar para lograr este objetivo. Otras de las mejoras que se incluirán en el portal lo constituye la creación de un espacio de publicación de las acciones del GTDA, destinado al contacto con sociedad civil, academia, empresa privada y otros; así como la

¹⁴ <https://www.antai.gob.pa/pagina-webs-del-estado-deben-ser-accesibles-a-las-personas-con-discapacidad/>

publicación de las evaluaciones de monitoreo de la calidad de los datos subidos al portal.

2. Aumentar las instituciones que publican datos abiertos.
3. Profundizar la apertura en las instituciones de gobierno con datos de calidad.
4. Proyectar los datos abiertos a las organizaciones de sociedad civil, académicos, empresarios, medios de comunicación con el fin de su uso, re-uso, redistribución y generación de valor económico.

6. 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, en el cual 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 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, página, 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.

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 máquinas” según la definición de Tim Berners-Lee.

WEB SERVICES – En el caso de 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.

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.

7. Referencias

- Programa de formación en línea del portal Europeo de Datos <https://www.europeandataportal.eu/es/training/elearning>
- Guía Nacional para la Apertura y Publicación de Datos Abiertos de Gobierno de Costa Rica <http://www.gobiernoabierto.go.cr/datosabiertos/>
- Guía práctica para la publicación de datos tabulares en archivos CSV <https://datos.gob.es/es/documentacion/guia-practica-para-la-publicacion-de-datos-tabulares-en-archivos-csv>
- Guía para el uso y aprovechamiento de Datos Abiertos en Colombia, versión actualizada 2019 <https://herramientas.datos.gov.co/sites/default/files/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf>
- 8 Principios de los Datos de Gobierno Abierto <https://opengovdata.org/>
- 5 Estrellas de los Datos Abiertos <https://5stardata.info/es/>
- Licencias y Herramientas de Derechos Autorales “Creative Commons” <https://creativecommons.org/licenses/>
- Evaluación de Portal y Datos-Comunidades CEPAL http://comunidades.cepal.org/ilpes/sites/default/files/2018-07/IGA02_Evaluaci%C3%B3n%20Portal%20Datos_0.xlsx
- Estándares de Accesibilidad a la Web para Personas con Discapacidad <https://www.w3.org/WAI/standards-guidelines/w3c-process/#standards>
- Guía de medición del Open Data Barometer <https://opendatabarometer.org>



**Grupo de Trabajo
de Datos Abiertos
de Gobierno**

**República de Panamá
2020**

