

# GUÍA DE DATOS ABIERTOS EN COLOMBIA

**PARA:**

- Abrirse a la innovación.
- Empoderar al ciudadano.
- Medir el impacto de las políticas.
- La transparencia y el control social.
- Mejorar la eficiencia y eficacia del Estado.
- Mejorar o crear productos y modelos de negocio.



MINTIC

vive digital  
para la gente

TODOS POR UN  
NUEVO PAÍS  
PAZ EQUIDAD EDUCACIÓN

# VERSIONES

## VERSIÓN 1.

12/2016

### **Guía para la apertura de datos en Colombia**

Dirigida a entidades del Estado para el desarrollo de procesos de apertura de datos públicos.

## VERSIÓN 2.

06/2016

### **Guía de datos abiertos en Colombia**

Dirigida a las entidades sujeto de aplicación de la Ley 1712 de 2014 de Transparencia y Acceso a la Información Pública, para la aplicación de orientaciones y buenas prácticas en el desarrollo de estrategias de apertura y reuso de datos abiertos.

## VERSIÓN 3.

06/2016

### **Guía de datos abiertos en Colombia**

Modificaciones de diagramación

# 00

# INTRODUCCIÓN

Las formas tradicionales de Gobierno se han ido transformando hacia gobiernos abiertos en donde la sociedad tiene un rol mucho más activo y colaborador, en dónde un problema de interés público ya no es solo asunto de los funcionarios públicos sino que también la ciudadanía puede aportar en su solución, y en donde los actores de Gobierno son mucho más receptivos e interactúan con la comunidad. En este nuevo modelo de gobierno juegan un rol muy importante: la transparencia, la colaboración y la participación y por esto es fundamental contar con acceso a la información pública.

En Colombia, el acceso a la información pública es un derecho fundamental que cumple al menos tres funciones esenciales:

- Garantizar la participación democrática y el ejercicio de los derechos políticos.
- Ser un instrumento fundamental para el ejercicio de otros derechos constitucionales, ya que permite conocer las condiciones necesarias para su realización.
- Garantizar la transparencia de la gestión pública, y por lo tanto, constituirse en un mecanismo de control ciudadano de la actividad estatal.

Los datos abiertos hacen parte de la información pública y representan un gran potencial para el desarrollo de procesos de Gobierno Abierto (*Open Government*)<sup>1</sup>, en la medida en que permiten reutilizar información de diversas fuentes, sin restricciones de uso y en formatos de fácil lectura y análisis, de tal manera que se habilitan nuevas formas de tomar decisiones, de resolver problemas públicos, de entender determinados comportamientos y hasta generar valor no solo social sino económico a través de la innovación.

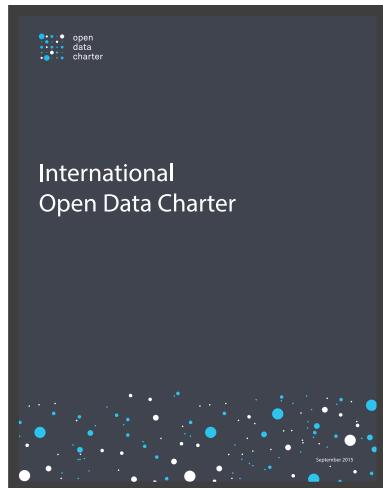
El objetivo del presente documento, es proporcionar orientaciones y buenas prácticas para el desarrollo de estrategias de apertura y reuso de datos abiertos, que estén orientadas a la generación de valor en lo económico, social, político, cultural, ambiental, y en general, en los distintos ámbitos de la sociedad. Este documento comprende las siguientes temáticas que pueden navegarse desde el menú superior:

---

<sup>1</sup> Conjunto de políticas que buscan fomentar la construcción de un Estado más transparente, participativo y colaborativo en los asuntos públicos.

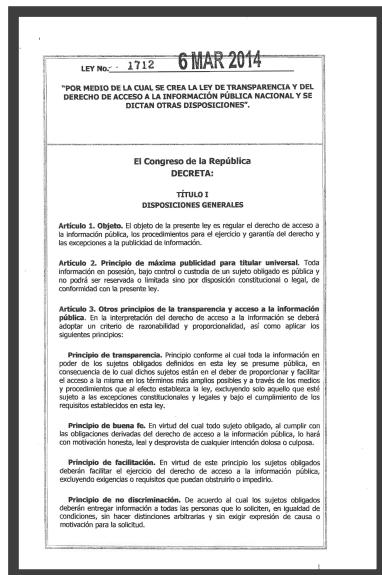
## 01

# ¿QUÉ SON LOS DATOS ABIERTOS?



## La **Carta Internacional de Datos**

Abiertos define los datos abiertos como “datos digitales que son puestos a disposición con las características técnicas y jurídicas necesarias para que puedan ser usados, reutilizados y redistribuidos libremente por cualquier persona, en cualquier momento y en cualquier lugar”. A partir de ello, los datos abiertos son considerados como uno de los factores que impulsan la transformación global favorecida por la tecnología y los medios digitales y permiten a los gobiernos, ciudadanos y organizaciones de la sociedad civil y del sector privado tomar mejores decisiones informadas.



Por otra parte, el **Banco Mundial** en el documento de diagnóstico y recomendaciones para la implementación de un programa de datos abiertos realizado para Colombia en el año 2015<sup>2</sup>, define datos abiertos como los “datos publicados en formatos legibles por máquinas, que están disponibles públicamente bajo una licencia abierta que asegura que pueden ser utilizados, reutilizados o distribuidos libremente por cualquier persona, para cualquier propósito legal”.

En Colombia la **Ley de Transparencia y Acceso a la Información**, define los datos abiertos como “todos aquellos datos primarios o sin procesar, que se encuentran en formatos estándar e interoperables que facilitan su acceso y reutilización, los cuales están bajo la custodia de las entidades públicas o privadas que cumplen con funciones públicas y que son puestos a disposición de cualquier ciudadano, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos”<sup>3</sup>.

<sup>2</sup> Open Data Readiness Assessment (ODRA). Estudio realizado en mayo – Septiembre de 2015. World Bank Group.

<sup>3</sup> Ley 1712 de 2014. Literal J, artículo 6. Definiciones.

**02**

# ¿CUÁLES PRINCIPIOS RIGEN LOS DATOS ABIERTOS?

## PRIMARIOS

Obtenerse en la fuente de origen, con el más alto nivel de detalle posible, no en forma agregada ni modificada.

## PROCESABLES POR MÁQUINAS

Encontrarse en formatos que permitan el procesamiento automático.

## LICENCIADOS DE FORMA ABIERTA

Los conjuntos de datos publicados deben contar con términos de uso y licenciamiento abierto<sup>4</sup>.

## ACCESIBLES

Estar disponibles para el rango más amplio de usuarios y para el rango más amplio de propósitos.

## NO PROPIETARIOS

Estar disponibles en un formato sobre el cual ninguna entidad tenga control exclusivo.

## NO DISCRIMINADOS

Estar disponibles para cualquiera persona, sin requerir registro o autenticación.

## COMPLETOS

Reflejar la totalidad del tema y contener el mayor detalle posible, garantizando que la información suministrada sea suficiente y consistente y que no contenga datos nulos.

## OPORTUNOS Y ACTUALIZADOS

Estar disponible tan rápido como sea necesario para garantizar su valor y mantener una frecuencia de actualización que garantice la utilidad del dato.

<sup>4</sup> Ver anexo 1. Uso de licencias abiertas.



LICENCIA ABIERTA ES AQUELLA  
QUE PERMITE DE FORMA LIBRE  
Y SIN RESTRICCIONES: **EL USO**  
**EXTRACCIÓN ADAPTACIÓN**  
**TRANSFORMACIÓN** COPIA  
**DISTRIBUCIÓN REDISTRIBUCIÓN**  
**REUTILIZACIÓN**  
**MODIFICACIÓN** COMPILEACIÓN  
**DIFUSIÓN DE LOS DATOS**  
**PUBLICADOS**

Tales principios se encuentran en concordancia con la Carta Internacional de Datos Abiertos, lanzada en la Cumbre de la Alianza por el Gobierno Abierto en Octubre de 2015. Consultela en: <http://opendatacharter.net>.

03

# ¿POR QUÉ ES IMPORTANTE ABRIR Y REUTILIZAR DATOS?



Los datos abiertos pueden generar valor en diferentes áreas, dentro de las que se encuentran:

## 1. TRANSPARENCIA Y CONTROL SOCIAL

Se busca que la información de la gestión pública esté completamente visible y que permita realizar acciones de control social. Un ejemplo de esta aplicabilidad de los datos es la realizada por proyectos como *¿Where does my money go?*, desarrollado por el Reino Unido, en donde se exponen los datos de cómo el Gobierno está invirtiendo el dinero de los impuestos.

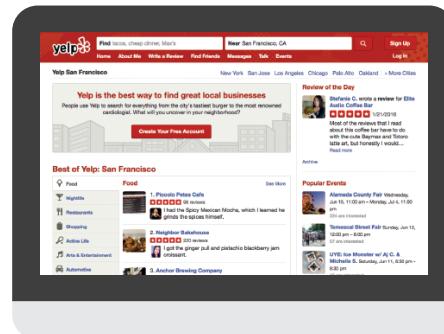


**¿WHERE DOES MY MONEY GO?**  
<http://app.wheredoesmymoneygo.org>

## 2. EMPODERAMIENTO

Los datos abiertos pueden ayudar a que las personas tengan un rol más activo en la sociedad. Un ejemplo de este empoderamiento es el caso de San Francisco (EUA), en donde el Gobierno publica los resultados de las inspecciones sanitarias en los restaurantes de la ciudad y a partir de dicha información, los ciudadanos deciden dónde comer y los restaurantes se ven sujetos a mejorar sus prácticas de manipulación de alimentos. Al final el beneficio se observa en la mejora de la sanidad pública de la ciudad y en la generación de otras plataformas que usan esta información.

Otro ejemplo de empoderamiento es el caso de los hospitales en Londres, quienes publican cifras sobre sus tasas de éxito y fracaso en la atención, procedimientos y personal médico, con lo cual los pacientes pueden decidir el hospital o lugar más conveniente para su caso y los médicos pueden aprender de sus colegas mediante el acceso a información específica de procedimientos médicos.



**YELP**  
<http://www.yelp.com>



**ZAGAT**  
<http://www.zagat.com>



**HOSPITAL DATA**  
<http://www.telegraph.co.uk/news/health/expat-health/9479874/Publish-hospital-data-and-save-thousands-of-lives.html>

### 3. MEJORAMIENTO O CREACIÓN DE PRODUCTOS, SERVICIOS Y MODELOS DE NEGOCIO INNOVADORES

Los datos abiertos tienen un gran potencial para generar valor económico a partir de su uso, pues además de generar beneficios a nivel de transparencia y rendición de cuentas, tienen la posibilidad de ser usados para mejorar o crear nuevas empresas o productos. Por ejemplo, servicios como Mapumental en el Reino Unido o Mapnificent en Alemania, permiten encontrar lugares donde vivir tomando en cuenta la duración del viaje hasta el trabajo, precios de viviendas y belleza de la zona.

El sitio danés Husetsweb ayuda a encontrar maneras de mejorar la eficiencia energética en la vivienda, incluyendo planificación financiera e información sobre constructores. Está basado en información catastral y sobre subsidios del gobierno, así como en el registro de comercio local.



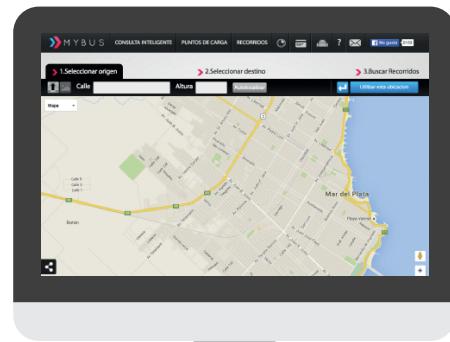
**MAPNIFICENT**  
<http://www.mapnificent.net>



**HUSETSWEB**  
<http://www.husetsweb.dk>

## 4. MEJORAMIENTO EN LA EFICIENCIA Y EFICACIA DE LOS SERVICIOS OFRECIDOS POR EL ESTADO

Los datos abiertos son de gran valor también para el Gobierno, en la medida en que pueden mejorar la eficiencia en la prestación de los servicios de las entidades. En Holanda, el Ministerio de Educación publicó en Internet todos los datos relacionados con educación para su reutilización. A partir de ello, disminuyó el número de preguntas de los ciudadanos, reduciendo el volumen de trabajo. Otro ejemplo es el caso de MyBus, un sitio web desarrollado por programadores



**MYBUS**  
<http://www.mybus.com.ar>

de la Ciudad de Mar del Plata Argentina, que ayuda a encontrar la mejor forma para moverse en la ciudad.

## 5. MEDICIÓN PREDICTIVA DEL IMPACTO DE POLÍTICAS

Los datos abiertos también pueden ser usados para prever situaciones futuras a partir del desarrollo de políticas o proyectos de impacto social. Un ejemplo es el caso de la reforma educativa en Chile, que ha sido desarrollada a partir de datos abiertos provistos por el Gobierno y mediante los que se han medido aspectos del sistema educativo como falta de equidad espacial o cumplimiento de estándares mínimos de aprendizaje y se han podido predecir y prevenir fenómenos



**BIG DATA PARA LA TOMA DE DECISIONES EN EDUCACIÓN**  
<http://http://yogobierno.org/big-data-para-la-toma-de-decisiones-en-educacion/>

de alto impacto social como la deserción o el abandono escolar.

## 6. NUEVOS CONOCIMIENTOS A PARTIR DE FUENTES DE DATOS COMBINADAS Y GENERACIÓN DE PATRONES EN GRANDES VOLÚMENES DE DATOS

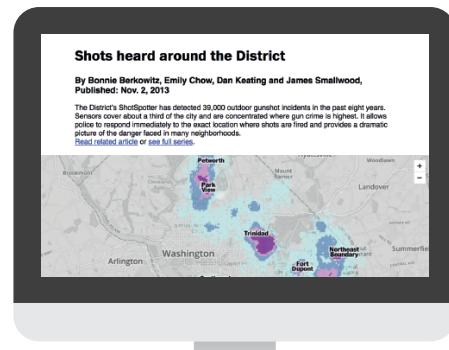
La posibilidad de contar con grandes volúmenes de datos abiertos en formato digital, permite que estos puedan ser procesados fácilmente y generar visualizaciones para descubrir tendencias o identificar nuevas perspectivas de un problema. Este uso de los datos se presenta en áreas como el periodismo de datos, en donde se analizan situaciones o se generan historias a partir de la combinación, análisis, y visualización de datos abiertos.

En Colombia a partir de datos abiertos publicados por el Instituto Nacional de Salud – INS sobre calidad del agua potable para el consumo humano, el periódico El Tiempo generó un análisis sobre municipios del país en donde se consume la mejor y la peor agua del País. Otro ejemplo es uno de los artículos del periódico Washington Post, en donde se cruzaron datos de asaltos con armas de fuego y grabaciones de tiroteos, para



### ¿SABE USTED DÓNDE SE TOMA LA MEJOR Y LA PEOR AGUA DEL PAÍS?

<http://www.eltiempo.com/multimedia/especiales/calidad-del-agua-en-colombia/16555634/1>



### SHOTS HEARD AROUND THE DISTRICT

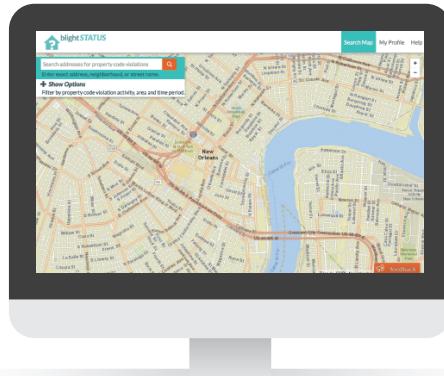
<http://www.washingtonpost.com/wp-srv/special/local/dc-shot-spotter/>

explicar en profundidad el problema de la criminalidad en algunas zonas de la ciudad.

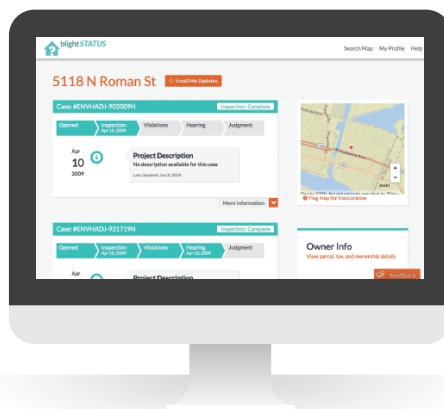
## 7. TOMA DE DECISIONES

El análisis de los datos abiertos facilita la toma de decisiones no sólo por parte del Estado para la definición de políticas, sino por parte de ciudadanos y empresas, en materia de salud, educación, inversión, medio ambiente, etc.

En Nueva Orleans, el problema de las propiedades que fueron abandonadas después del huracán Katrina, se ha logrado resolver a partir del análisis de datos abiertos, pues han permitido acelerar el proceso de declaración legal de una propiedad abandonada y la definición de otros usos. En este caso la ciudad ha transformado más de 13.000 de estas propiedades y a través del sitio web de BlightStatus, las personas pueden comprobar el progreso de transformación de las propiedades en su vecindario.



**BLIGHTSTATUS**  
<http://blightstatus.nola.gov>



**BLIGHTSTATUS**  
<http://http://blightstatus.nola.gov/addresses/195374>

## INFRAESTRUCTURA INSTITUCIONAL DE DATOS

Otros casos en el uso de datos abiertos son:

### CIUDATOS

<http://ciudatos.com>

Es la plataforma de datos abiertos de la Red de Ciudades Cómo Vamos, que proporciona información sobre cambios en la calidad de vida de 11 ciudades, para consulta y análisis de todos los ciudadanos.

<http://www.humboldt.org.co/es/servicios/infraestructura-institucional-de-datos>

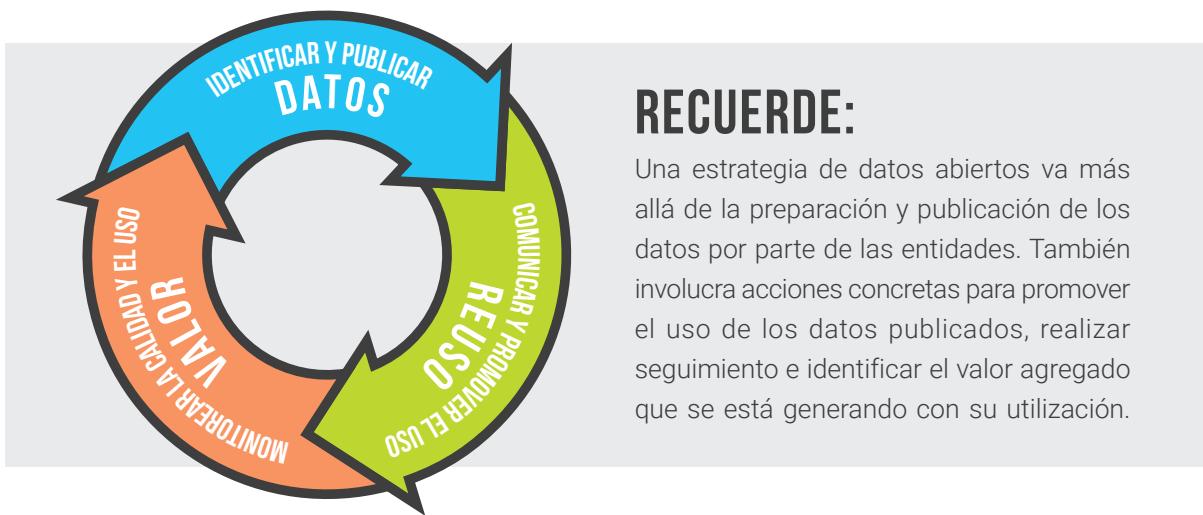
Es el catalogador de información geográfica y biológica del Instituto Alexander Von Humboldt, que ofrece acceso a información generada por el Instituto y por estudiantes, investigadores, entidades estatales y ONG. Esta información interopera con otros sistemas de información, como la del Sistema de Información Ambiental, la Infraestructura Colombiana de Datos, y la Infraestructura Colombiana de Datos Espaciales.

## 04

# ¿CÓMO DESARROLLAR UN PROCESO DE APERTURA Y USO DE DATOS?



Un proceso exitoso para el uso de los datos abiertos requiere comunicación, flexibilidad, retroalimentación e interacción constante entre quienes ofertan y quienes reutilizan los datos abiertos. En este sentido, se pueden evidenciar los siguientes momentos en el proceso de apertura y uso de los datos<sup>5</sup>:



<sup>5</sup> Definido a partir del círculo virtuoso de datos abiertos propuesto por el Banco Mundial para el fortalecimiento de la iniciativa de datos abiertos en Colombia (Datos, reuso y valor) y los momentos establecidos en el Manual de Gobierno en línea para la apertura de datos.



**LOS DATOS ABIERTOS SON VALIOSOS EN LA MEDIDA EN QUE SEAN EXHAUSTIVOS, PRECISOS, DE ALTA CALIDAD Y QUE RESPONDAN A NECESIDADES CONCRETAS DE LOS USUARIOS.**

El papel de las entidades que identifican y publican datos abiertos es fundamental, pues la demanda de datos abiertos viene en constante aumento y es cada vez más exigente. Para lograr un mejor y mayor uso de los datos es importante:

**IDENTIFICAR** cómo está conformada la demanda de los datos abiertos de la entidad, sus necesidades y expectativas. Recuerde que la demanda de datos puede venir de la sociedad civil, el sector privado, los organismos internacionales, la academia y grupos de investigación, las ONG o el sector periodístico, entre otros.

**GARANTIZAR** la transparencia en el proceso de recogida, procesamiento y gestión de los conjuntos de datos<sup>6</sup> a publicar.

**IMPULSAR** y dar a conocer los beneficios del uso de los Datos abiertos.

**GARANTIZAR** que se pueda acceder a los datos abiertos de manera directa y sin restricciones, evitando protocolos y procedimientos innecesarios como convenios, solicitudes por email y luego mediante derecho de petición, incurriendo en demoras y costos adicionales.

**PUBLICAR** los datos periódicamente identificando aquellos que tienen mayor demanda por usuario y/o grupo de interés. Estos pueden ser utilizados para generar tendencias, reportes, informes de seguimiento, entre otros.

**MEJORAR** el proceso y los espacios de atención a la demanda de datos abiertos y de respuesta a sus requerimientos. El objetivo es lograr una relación de colaboración entre la entidad que identifica y publica los datos y quienes hacen uso de los mismos.

**PUBLICAR** datos georreferenciados, lo cual corresponde a la información geográficamente referenciada con el fin de resolver problemas complejos de planificación y gestión.

**IMPULSAR** el uso de datos como una herramienta de mejora en la gestión de la entidad.

<sup>6</sup> Conjunto de datos (Dataset): unidad mínima de información sujeta a carga, publicación, transformación y descarga en la plataforma.

## ¿CÓMO IDENTIFICAR LOS DATOS A PUBLICAR?



Verifique cuáles son las solicitudes de información más recurrentes y conviértalas en dato abierto.



Identifique información que ya está publicada en el sitio web en formato PDF, Word, Excel, JPG o tablas dinámicas y sólo requiere disponerse en formato abierto.



Identifique y publique datos de alto impacto en salud, educación, impuestos, movilidad, seguridad ciudadana, salud pública, atención y reparación a las víctimas y ordenamiento territorial\*.



Publique en formato abierto información que ya se encuentra en estudios, análisis y estadísticas, validando el manejo adecuado de la información confidencial.

\* Estos temas fueron priorizados en la Ruta de la Excelencia en Gobierno en Línea (Mapa de Ruta Decreto 1078 de 2015 Título 9 - Capítulo 1), por lo que debe estar atento a futuras actualizaciones.

## ASPECTOS LEGALES A CONSIDERAR

La Ley 1712 de 2014 de Transparencia y Acceso a la Información, establece que la entidad debe contar con un Registro de Activos de Información<sup>7</sup>, que es el inventario de la información pública que genera, obtiene, adquiere, transforma o controla la entidad. Este registro debe ser publicado en la página web de la entidad y en el portal de datos [www.datos.gov.co](http://www.datos.gov.co) o el que haga sus veces.

Dentro de la información que debe tener el Registro de Activos de Información se encuentra:

- Información relacionada con los trámites, servicios y procesos que soportan el “día a día” del trabajo de la entidad.
- Procesos encaminados en la mejora continua y el cumplimiento de los objetivos estratégicos.
- Información que soporta los procesos relacionados con la contratación, el recurso humano o la información financiera y contable.

Así mismo, la ley también establece la realización de un Índice de Información Clasificada y Reservada<sup>8</sup>, en donde la entidad identifica qué información se debe ser clasificada como tal por motivos de protección de datos personales o por motivos de defensa, seguridad nacional o estabilidad económica del País. Este índice debe ser publicado en el sitio web oficial de la entidad, así como en el portal de datos abiertos [www.datos.gov.co](http://www.datos.gov.co) o el que haga sus veces.

---

<sup>7</sup> Decreto 1081 de 2015 Cap. 5, Sección 1 Art 2.1.1.5.1.1. – Registro de Activos de Información.

<sup>8</sup> Decreto 1081 de 2015 – Gestión de la Información reservada – Art. 2.1.1.4.1. Índice de Información Clasificada y reservada – Art. 2.1.1.5.2.1.

## RECUERDE

Los datos abiertos no son contrarios a la protección de datos personales: si existe información asociada a datos personales pero que puede ser valiosa como dato abierto, adelante acciones de anonimización para eliminar aquella información sensible que afecta a personas u organizaciones y cuya identidad debe protegerse legalmente.

Antes de publicar cualquier información que sea sensible, considere otras fuentes de información disponibles y evalúe si la combinación de éstas pueden presentar algún riesgo. Con esto puede prevenir el efecto mosaico, que ocurre cuando la información de una base de datos por sí sola, no genera un riesgo para la identificación de individuos, pero al combinarse con otra información disponible, puede generar tal riesgo.

En materia de datos abiertos, específicamente la Ley 1712 establece que las entidades deben “publicar datos abiertos”, teniendo en cuenta las excepciones de publicar información pública clasificada o información pública reservada. Algunos ejemplos de esta información son los siguientes:

### INFORMACIÓN PÚBLICA CLASIFICADA

- Pone en riesgo la intimidad de las personas.
  - Pone en riesgo la vida, salud o seguridad de las personas.
  - Afecta los secretos comerciales, industriales, profesionales.
  - Afecta los derechos estipulados en el parágrafo del Art. 77/ 1474 de 2011.
- 

**INFORMACIÓN PÚBLICA CUYO ACCESO LIBRE PUEDE CAUSAR DAÑO DE DERECHOS A PERSONAS NATURALES O JURÍDICAS.**

### INFORMACIÓN PÚBLICA RESERVADA

- Defensa y seguridad Nacional y pública.
  - Relaciones internacionales.
  - Debido proceso en los procesos judiciales.
  - Derechos de la infancia y la adolescencia.
  - Estabilidad macroeconómica y financiera del país.
  - Salud pública.
- 

**INFORMACIÓN PÚBLICA CUYO ACCESO LIBRE PUEDE CAUSAR DAÑO A LOS INTERESES PÚBLICOS.**

## ¿CÓMO PUBLICAR DATOS ABIERTOS?

La publicación de datos requiere una preparación previa que tiene como objetivo alistar los datos adecuadamente para que estos puedan ser procesados y reutilizados por terceros. Este alistamiento incluye tres actividades:

DOCUMENTACIÓN

ESTRUCTURACIÓN

CARGA DE DATOS

### DOCUMENTACIÓN

La documentación de los datos consiste en definir los metadatos que tendrá cada uno de los conjuntos de los datos o datasets a publicar. Los metadatos son datos que describen los aspectos básicos de un conjunto de datos y son una herramienta fundamental para organizar, clasificar, relacionar y encontrar los datos necesarios. La cantidad y calidad de metadatos asociados a un conjunto de datos, permite una mejor clasificación y una mejor búsqueda y recuperación de los datos. Algunos metadatos generales son los siguientes:

**Nombre del conjunto de datos:**

Nombre definido para el dataset.

**Descripción:**

Resumen descriptivo del conjunto de datos, en donde se explica su contenido, su contexto y el detalle de lo que se encontrará, de manera que cualquier usuario se encuentre informado sobre el contenido antes de descargar el recurso.

**Entidad:**

Nombre de la entidad o institución que produce o controla la información.

**Categoría:**

Clasificación conceptual básica del dataset en sistemas de categorías o taxonomías disponibles.

**Idioma:**

Idioma de los datos del conjunto de datos.

**Fecha de publicación:**

Fecha en la que se emitió el conjunto de datos.

La cantidad y características de los metadatos, dependerá de la plataforma en donde serán publicados los conjuntos de datos. Para el caso del Portal de Datos del Estado Colombiano, se proporcionan los recursos necesarios para realizar la documentación de los datos, en: [www.datos.gov.co](http://www.datos.gov.co).

## RECUEERDE

La documentación de datos:

- Permite a los usuarios entender el contenido y características de los datos para generar mayor valor en su publicación.
- Permite identificar integración y/o interoperabilidad con otras fuentes de información

## ESTRUCTURACIÓN

Consiste en la preparación del conjunto de datos a publicar en un formato estructurado. Para realizar esta actividad, la entidad debe elegir el formato que considere más adecuado, bajo criterios de menor esfuerzo y costo, pero sin afectar la calidad y accesibilidad de los datos. Los formatos para la publicación de los conjuntos de datos se agrupan en tres tipos:

- **Formatos propietarios:** Son formatos de archivo que requieren herramientas licenciadas para acceder, visualizar o editar los datos. Este tipo de formatos son por ejemplo .DOC, .XLS o .SHP.
- **Formatos libres:** Son formatos de archivo que se pueden crear y manipular para cualquier software, libre de restricciones legales y que aumentan el nivel de reutilización de los datos. Este tipo de formatos son por ejemplo .CSV, .TMX, .XML, .ODF, JSON.

## RECUERDE

La estructuración de datos:

- Busca transformar el dato de formatos no-textuales a formatos que permitan su manipulación y su conexión con otros datos (datos enlazados o linked data).
- Existen diversos formatos para la estructuración de datos, pero a la hora de seleccionarlos se debe tener en cuenta que sean legibles directamente por máquinas, que sean abiertos y permitan un uso sin restricciones y que se encuentren dentro de los más usados por los reutilizadores.
- Debe evitarse el uso de formatos no editables como el PDF.

## CARGA DE DATOS

El cague de datos se realiza a través de su publicación en una plataforma que permita la organización y fácil acceso por parte de quienes van a reutilizar los datos, de manera que se convierta en un sitio de acceso común y que tenga la posibilidad de hacer escalable la información publicada.

### PORTAL DE DATOS DEL ESTADO COLOMBIANO

» Sitio creado por el Ministerio de Tecnologías de la Información y las Comunicaciones, en donde las entidades públicas del orden nacional y territorial publican datos abiertos en temas de su competencia y de interés nacional. Este sitio cuenta con los formatos y procedimientos específicos para realizar la documentación, estructuración y cague y publicación de la información.



<http://www.datos.gov.co>

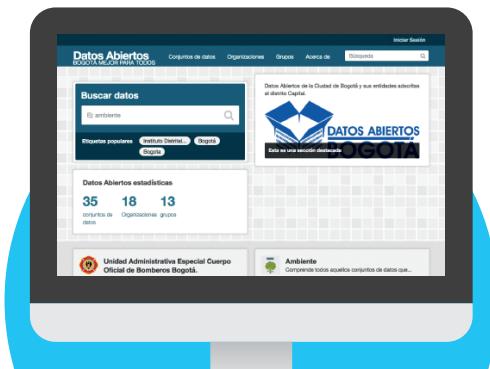
IDENTIFICAR

COMUNICAR

MONITOREAR

Adicional a esta plataforma, en Colombia existen plataformas que han sido desarrolladas por entidades públicas y privadas, sin embargo, las entidades públicas que publiquen sus datos abiertos en sitios diferentes al portal de Datos del Estado Colombiano deberán garantizar que los datos estén debidamente enlazados y referenciados en éste portal.

Algunas plataformas complementarias de publicación de datos abiertos son:



<http://datosabiertos.bogota.gov.co>

## « PORTAL DE DATOS DE BOGOTÁ

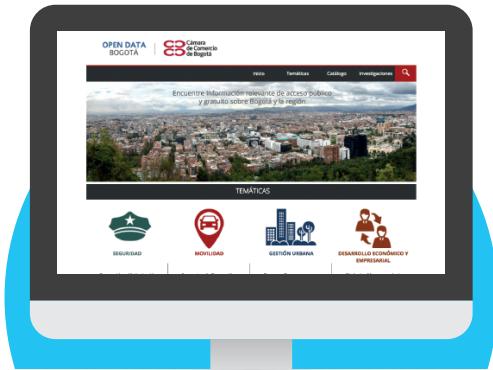
Contiene datos abiertos de la Ciudad y de las entidades adscritas al Distrito Capital.



<http://ciudatos.com>

## « CIUDATOS

Es el sitio de datos abiertos desarrollado por la Red de Ciudades Cómo Vamos, entidad de la sociedad civil de Medellín.



<http://opendatabogota.ccb.org.co>

## « ODATA

portal de la Cámara de Comercio Bogotá que pone a disposición de los empresarios de Colombia datos sociales, económicos y de movilidad, entre otros, para la toma de decisiones.



<http://datos.tulua.gov.co>

## « DATOS ABIERTOS DEL MUNICIPIO DE TULUÁ

es el sitio de la Alcaldía de Tuluá, en donde se publican datos abiertos en temas de salud, educación, agricultura, seguridad, turismo, presupuesto y finanzas del municipio.

## Otros sistemas de publicación:

Son lenguajes para consultar archivos, como el SPARQL, WMS, WFS, WS-API, entre otros.

## TENGA EN CUENTA

Sobre el Portal de Datos del Estado Colombiano:

La Ley 1712 de 2014 y el Decreto 103 de 2015 (compilado en el Decreto 1081 de 2015), establecen que los sujetos obligados de estas normas deben publicar en el Portal de Datos del Estado Colombiano:

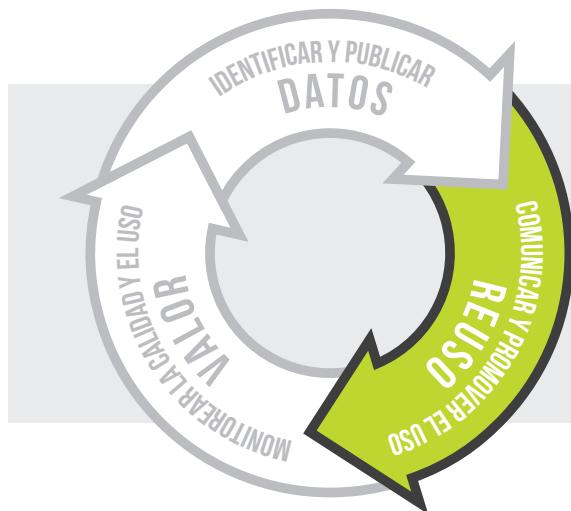
- El registro de activos de información.
- El índice de información clasificada y reservada.

Las entidades que publiquen sus datos abiertos en sitios diferentes al portal de Datos del Estado Colombiano [www.datos.gov.co](http://www.datos.gov.co) o el que haga sus veces, deben garantizar que dichos datos queden enlazados o relacionados en el Portal de Datos, mediante estrategias de bases de datos federadas<sup>9</sup>. Para mayor información, puede dirigir sus inquietudes al correo: [datosabiertos@mintic.gov.co](mailto:datosabiertos@mintic.gov.co)

---

<sup>9</sup> Ver definición en el glosario.

- IDENTIFICAR
- COMUNICAR
- MONITOREAR



EL ROL DE LAS ENTIDADES DEL ESTADO EN EL USO DE DATOS ABIERTOS NO TERMINA CON SU PUBLICACIÓN. TAMBIÉN DEBEN DESEMPEÑAR UN PAPEL ACTIVO PARA APOYAR UNA REUTILIZACIÓN EFICAZ E INNOVADORA DE LOS MISMOS.

## ¿CÓMO COMUNICAR Y PROMOVER EL USO DE LOS DATOS ABIERTOS?

Una vez se cuenta con la publicación de datos abiertos, es necesario desarrollar un plan de comunicación para:



### Fomentar la oferta:

Facilitar la apertura de datos.



### Fomentar la demanda:

Facilitar la reutilización y consumo de Datos Abiertos.



### Comunicar los resultados:

Sensibilizar sobre el uso de datos abiertos y sus logros y beneficios.

Además de la labor de comunicación al exterior de la entidad, se debe realizar una labor informativa dirigida a los empleados, a fin de transformar el manejo tradicional de los datos públicos en un manejo con mayor apertura y orientado a que los datos sean lo más ricos y completos posible.

A partir de tales objetivos, el plan de comunicaciones se podría plantear sobre los siguientes momentos:



### DAR A CONOCER

### VINCULACIÓN DE LOS AGENTES

### CONSOLIDACIÓN Y POSICIONAMIENTO



## DAR A CONOCER

Contiene el conjunto de comunicaciones internas y externas de la entidad, para sensibilizar sobre el uso de datos abiertos y sus beneficios, como mejora en temas de transparencia y control social, toma de decisiones, desarrollo de aplicaciones, mejora de servicios, periodismo de datos, entre otros.

Al respecto la entidad puede desarrollar estrategias de comunicación interna y externa así como hacer uso de redes sociales para difundir estos mensajes. Algunas actividades específicas para dar a conocer los usos y beneficios de los datos abiertos pueden ser:

- Generar mensajes internos vía intranet o e-mailing.
- Realizar encuentros o reuniones internas.
- Desarrollar campañas informativas con medios de comunicación para dar a conocer beneficios con el uso de los datos abiertos de la entidad.
- Realizar campañas de comunicación a través de Urna de Cristal para incentivar el uso de datos de la entidad y difundir soluciones desarrolladas con datos abiertos de la entidad.

 IDENTIFICAR COMUNICAR MONITOREAR

## VINCULACIÓN DE LOS AGENTES

Esta etapa tiene como objetivo adelantar actividades para vincular a diferentes actores que promuevan la reutilización de datos de la entidad, difundir soluciones desarrolladas con datos abiertos y desarrollar actividades para identificar nuevas posibilidades con el uso de los datos.

Algunas actividades que se pueden adelantar para vincular a los agentes pueden ser:

- Desarrollar procesos de co-creación de herramientas tecnológicas o aplicaciones para impulsar la reutilización de datos, con participación de desarrolladores y usuarios<sup>10</sup>.
- Realizar talleres o encuentros con usuarios o grupos de interés para generar propuestas de usos creativos con los datos abiertos.
- Desarrollar planes o estrategias de incentivos para promover la participación de los empleados, usuarios y grupos de interés en el desarrollo de las actividades que impulsen la identificación y uso de datos abiertos de la entidad.

### RECUERDE

Hay un gran interés tanto en el acceso como en la reutilización de datos por parte de muchas organizaciones y personas de la sociedad civil: Existen varios referentes en materia de demanda de datos abiertos, como aquellos que son grandes reutilizadores, saben hacer visualizaciones, promulgan los beneficios de los datos abiertos o se relacionan con referentes internacionales. Todos ellos aportan en la generación de valor a partir del uso de datos abiertos.

---

<sup>10</sup> Para el desarrollo de esta actividad se puede tomar como referencia la Guía para Desarrollar Ejercicios de Innovación Abierta, en el sitio web del Centro de Innovación: <http://centrodeinnovacion.gobiernoenlinea.gov.co/es/caja-herramientas/ejercicios-de-innovacion-abierta>

IDENTIFICAR

COMUNICAR

MONITOREAR

## CONSOLIDACIÓN Y POSICIONAMIENTO

En esta etapa se busca que la entidad incorpore su estrategia de datos abiertos como parte integral de sus planes estratégicos o planes de acción, de manera que la apertura y uso de datos pasen de ser actividades puntuales o iniciativas coyunturales, a iniciativas estructurales que cuentan con el conocimiento y respaldo del nivel directivo, con la asignación de equipos responsables de liderar las actividades para identificar, publicar, promover y hacer seguimiento al uso de los datos abiertos y con la asignación de recursos para adelantar estas actividades.

En esta etapa se espera que de forma permanente la entidad desarrolle jornadas de sensibilización con usuarios internos y externos sobre los beneficios de los datos abiertos, realice campañas de comunicación sobre soluciones desarrolladas a partir de datos abiertos, desarrolle procesos de co-creación para impulsar la reutilización, establezca contacto con medios de comunicación para difundir logros o beneficios obtenidos con el uso de datos abiertos de la entidad, entre otros.

### CINCO CLAVES

**1** Promueva la reutilización de la información en formato abierto no solo a usuarios externos, sino al **interior de la entidad**.

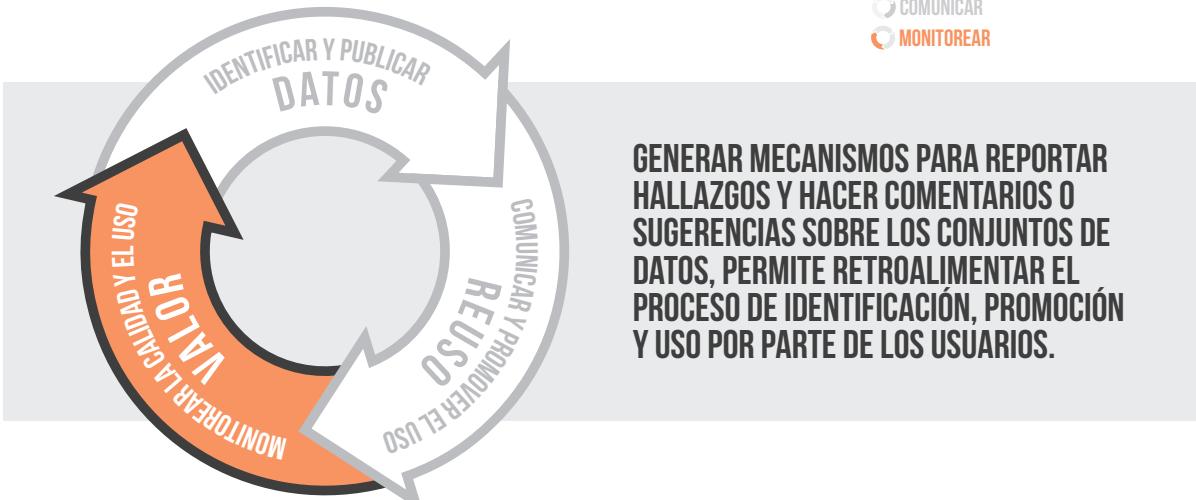
**2** Seleccione los **canales de difusión** más apropiados según el público potencial para el uso de los datos abiertos.

**3** Mencione a los usuarios de manera explícita, los **posibles usos** de los conjuntos de datos de la entidad. Por ejemplo en la toma de decisiones, la mejora de los servicios de la entidad, usos en la academia y en la investigación, etc.

**4** Mantenga actualizados los datos publicados. El interés en el uso de los mismos se mantendrá si los datos son pertinentes y oportunos.

**5** Contemple **incentivos** que pueda otorgar para el uso de los datos abiertos

- IDENTIFICAR
- COMUNICAR
- MONITOREAR



El uso de los datos abiertos es un aspecto estrechamente ligado a su calidad. En la medida en que se garantice que los datos abiertos publicados son completos, oportunos y consistentes, es posible promover su uso y por lo tanto, generar valor a partir de ello.

Para monitorear el uso de los datos abiertos, es posible utilizar algunos indicadores de tipo cuantitativo que permiten determinar su demanda y nivel de uso. Algunos de estos pueden ser:

- Número de solicitudes de apertura de datos.
- Número de dependencias o áreas de la entidad que abren datos.
- Porcentaje de conjuntos de datos abiertos estratégicos publicados.
- Número de conjuntos de datos descargados Vs. número de conjuntos de datos publicados.
- Número de aplicaciones generadas a partir de datos abiertos.
- Número de publicaciones que hicieron uso de los conjuntos de datos abiertos (papers, artículos, noticias, libros, etc.).

Por otra parte, las actividades de monitoreo también deben orientarse a identificar el valor generado para ciudadanos, usuarios y grupos de interés a partir de la publicación de datos abiertos, con lo cual se apunta al fortalecimiento de las acciones en materia de gobierno abierto. En este sentido, el valor público puede ser identificado en los siguientes ámbitos<sup>11</sup>:

- **Financiero:** Impacto actual o futuro de ingresos, valor de activos, pasivos o cualquier otro aspecto relacionado con la riqueza y el riesgo.
- **Político:** Impacto en una persona o un grupo de influencia o partidos políticos como producto de la acción del gobierno o su política.
- **Social:** Impacto en las relaciones con la comunidad o familias, en la movilidad social, estatus o identidad.
- **Estratégico:** Impacto en personas o grupos económicos relevantes en sus objetivos y recursos para la innovación o el planeamiento.
- **Ideológico:** Impacto en las creencias, en la moral o en los compromisos éticos en la sociedad.
- **Legitimidad y Respeto:** Impacto a nivel de la confianza, integridad y legitimidad de entidades públicas y privadas.

---

<sup>11</sup> Tomado de “Open Goverment and E-Government Democratic Challenges from a public value perspective”.

# 05

# GLOSARIO

## ANONIMIZAR

Proceso para remover datos personales de una base de datos, buscando la publicación segura de datos para el reuso.

## ARCHIVO DE DATOS (DATASET)

Un conjunto de datos que forman una unidad significativa y que habitualmente pueden ofrecerse con un fichero.

## BASES DE DATOS FEDERADAS (BDF)

Son vistas unificadas de bases de datos independientes que aparentan ser una sola base de datos, pero son una colección de sistemas de bases de datos independientes, cooperativos, heterogéneos, que son autónomos y que permiten compartir todos o algunos de sus datos. Una BDF aparenta ser una BD normal y corriente, pero no tiene existencia física, es una vista lógica.

## CATÁLOGO DE SERVICIOS DE TI

Este catálogo recoge y describe los servicios de TI que la entidad presta actualmente a las demás entidades y al público en general.

## CONJUNTO DE DATOS (DATASET)

Unidad mínima de información sujeta a carga, publicación, transformación y descarga en la plataforma.

## CSV (VALORES SEPARADOS POR COMA)

Formato abierto y sencillo para representar datos en formato de tabla, en columnas separadas por comas (o punto y coma, donde la coma es el separador decimal) y las filas son saltos de línea. Los campos que tienen una coma, un salto de línea o una comilla doble, deben cerrarse entre comillas dobles. Las extensiones que se utilizan son .csv y .txt.

## DATOS ENLAZADOS O LINKED DATA

Método de publicación de datos estructurados para que puedan ser interconectados y más útiles. Se basa en tecnologías Web estándar, tales como HTTP, RDF y los URL, pero en vez de utilizarlos para servir páginas web para los lectores humanos, las extiende para compartir información de una manera que puede ser leída automáticamente por ordenadores. Esto permite que sean conectados y consultados datos de diferentes fuentes.

## EFFECTO MOSAICO

Combinación de información disponible de bases anonimizadas que permite la identificación de los individuos. Es importante evitar este efecto para proteger los datos de las personas.

## ENTIDAD

Organismos establecidos por la legislación Colombiana, los cuales tienen la facultad de definir inventarios de datos y conjuntos de datos a publicar.

## FORMATOS LIBRES

Son formatos de archivo que se pueden crear y manipular para cualquier software, libre de restricciones legales.

## FORMATOS PROPIETARIOS

Son formatos de archivo que requieren herramientas que no son públicas.

## METADATO

Los metadatos son "datos sobre datos" - es decir, los datos que describen los aspectos básicos de un conjunto de datos, por ejemplo, cuándo se creó el conjunto de datos, cuál es la agencia responsable de la base de datos, el formato de los datos, etc.

## PETI O PETIC

Plan Estratégico de Tecnologías de la Información el cual define la estrategia de TI en la entidad para el periodo correspondiente. En ocasiones este plan está incorporado dentro del Plan Estratégico Institucional.

## REUTILIZACIÓN DE DATOS

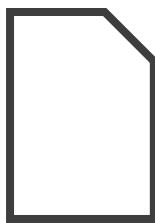
Producto que se elabora a partir de los datos públicos, puede ser una visualización, una aplicación web, un servicio, un cuadro de mandos, una noticia o una información, una gráfica, un dibujo, una gráfica dinámica entre otras cosas.

## REUTILIZADORES DE DATOS

Aquellas personas que con los Datos Abiertos como materia prima elaboran productos o servicios, pueden ser tales como emprendedores, empresas, ONGs, periodistas, hackers cívicos, o cualquier persona que tenga conocimientos del tratamiento y la manipulación de los datos.

## URL

Localizador de recursos uniforme, (sigla en inglés de uniform resource locator), es una secuencia de caracteres, de acuerdo a un formato modélico y estándar, que se usa para nombrar recursos en Internet para su localización o identificación. Referencia.



# ANEXO 1. USO DE LICENCIAS ABIERTAS

Las licencias estándar en el entorno internacional se han constituido en un elemento de seguridad y de generación de confianza en el uso de datos abiertos gubernamentales, ya que están concebidas como una forma de protección de los derechos de autor y de apoyo al dominio público de las obras protegidas por éste, fomentando el acceso a la información pública.

En la actualidad más de 30 países, y diferentes organizaciones internacionales como el Banco Mundial y las Naciones Unidas, hacen uso de licencias estándar para compartir sus trabajos<sup>12</sup>, por ello, para el licenciamiento de los datos abiertos se recomienda usar licencias abiertas tipo Creative Commons (CC), que proporcionan un conjunto estándar de términos y condiciones que los creadores y otros titulares de derechos, pueden utilizar para compartir obras originales de su autoría y cualquier otro material sujeto a derechos de autor y a derechos conexos que se especifican en la licencia pública.

La Licencia Pública CC tiene una vigencia de aplicación igual al plazo de protección de los Derechos de Autor y Derechos Conexos, en donde si el licenciatario incumple las condiciones de la Licencia, éstas terminarán (información tomada de la traducción de la versión 4.0 que actualmente están realizando los integrantes de CC América latina, para mayor información consultar: <http://co.creativecommons.org/>).

---

<sup>12</sup> Para mayor información consultar  
[https://wiki.creativecommons.org/wiki/Government\\_use\\_of\\_Creative\\_Commons](https://wiki.creativecommons.org/wiki/Government_use_of_Creative_Commons)

## □ ANEXO 1. USO DE LICENCIAS ABIERTAS

---

De igual manera, con el fin de definir los términos de uso y condiciones generales para el uso de los datos publicados en los sitios web, es importante definir una cláusula de licencia abierta dentro de los siguientes términos:

Se podrá hacer uso, transformación, distribución, redistribución, reutilización, compilación, extracción, copia, difusión, modificación y/o adaptación de los datos y de la información publicada en este sitio web, citando siempre la fuente de los datos. La entidad productora de los datos no será responsable de la utilización que hagan las personas que transformen y/o usen dichos datos, ni tampoco de los daños sufridos o pérdidas económicas que, de forma directa o indirecta, produzcan o puedan producir perjuicios económicos, materiales o sobre datos y aplicaciones, provocados por el uso y/o transformación de los datos. Esta licencia de uso se rige por la legislación colombiana, independientemente del entorno jurídico del usuario, cualquier disputa que llegue a surgir en la interpretación de estos términos se resolverá bajo el amparo de la Ley colombiana”.

En adición a lo anterior, también se recomienda proteger las bases de datos a través de licencias abiertas estándar o Conformant Licenses, entre las que se encuentran las licencias creative commons attribution 4.0. Este tipo de licencia permite proteger la estructura de las bases de datos, sus contenidos y garantizar que los usuarios de los datos atribuyan la titularidad de sus derechos a los creadores cuando se haga reuso de los mismos.