



El futuro digital  
es de todos

Gobierno  
de Colombia  
MinTIC

## Guía para el uso y aprovechamiento de Datos Abiertos en Colombia

Ministerio de Tecnologías de la Información y las Comunicaciones  
Viceministerio de Economía Digital  
Dirección de Gobierno Digital

Sylvia Constaín Rengifo  
Ministra de las Tecnologías de la Información y las Comunicaciones

Germán Rueda Jimenez  
Viceministro de Economía Digital

Equipo de trabajo  
Claudia Patricia Pico – Directora Gobierno Digital  
Luisa Fernanda Medina  
Ana Carolina Escobar  
Carolina Matamoros  
Nubia Santofimio  
Adriana Vargas Tamayo  
Claudia Milena Rodríguez  
Jose Ricardo Aponte  
Carlos Julio León Caicedo

Versión	Observaciones
Versión 1 Diciembre 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 Junio 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 reúso de datos abiertos.
Versión 3 Junio 2016	Guía de datos abiertos en Colombia - modificaciones de diagramación
Versión 4 Enero 2019	Actualización temática de la Guía para el uso y aprovechamiento de Datos Abiertos en Colombia
Versión 5 Mayo 2019	Guía para el uso y aprovechamiento de Datos Abiertos en Colombia - modificaciones de diagramación
Versión 6 Septiembre 2019	Actualización proceso Ciclo de Datos alineado al reto de máxima velocidad.

Comentarios, sugerencias o correcciones pueden ser enviadas al correo electrónico: [gobiernodigital@mintic.gov.co](mailto:gobiernodigital@mintic.gov.co)

Guía para el uso y aprovechamiento de Datos Abiertos en Colombia



Esta Guía para para el uso y aprovechamiento de Datos Abiertos en Colombia de la Dirección de Gobierno Digital se encuentra bajo una [Licencia Creative Commons Atribución 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/).



## Tabla de contenido

<b>Introducción</b> .....	<b>4</b>
<b>1. Datos abiertos</b> .....	<b>5</b>
1.1 Principios.....	7
<b>2. Beneficios de los Datos Abiertos</b> .....	<b>8</b>
2.1 Transparencia y Control Social .....	8
2.2 Mejoramiento o creación de productos, servicios y negocios innovadores.....	9
2.3 Pronosticar y prevenir fenómenos .....	9
2.4 Generar nuevo conocimiento .....	10
<b>3. Marco de Gobernanza</b> .....	<b>11</b>
3.1 Marco Normativo y Lineamientos .....	11
3.1.1 Marco Normativo .....	11
3.1.2 Lineamientos .....	12
3.2 Ecosistema de Actores.....	13
3.3 Esquema de incentivos.....	15
3.4 Monitoreo .....	15
<b>4. Ciclo de Datos Abiertos</b> .....	<b>17</b>
4.1 Establecer el plan de apertura .....	17
4.1.1 Identificar.....	18
4.1.2 Analizar .....	22
4.1.3 Priorizar.....	23
4.2 Estructurar y publicar .....	25
4.2.1 Documentar.....	25
4.2.2 Estructurar .....	27
4.2.3 Publicar .....	29
4.3 Comunicar y promover .....	30
4.3.1 Dar a conocer.....	31
4.3.2 Vincular diferentes actores .....	31
4.3.3 Consolidar y posicionar .....	32
4.4 Monitorear la calidad y el uso .....	33
4.4.1 Medir calidad.....	33
4.4.2 Medir usos.....	34
4.4.3 Medir impacto .....	35
<b>Glosario</b> .....	<b>37</b>
<b>ANEXO I: Índices Internacionales y marco normativo que fijan prioridades de publicación</b> .....	<b>38</b>



# Introducción

Las formas tradicionales de Gobierno se han ido transformando hacia el concepto de Estado Abierto en donde la sociedad tiene un rol más activo, y en el que un problema de interés general ya no es solo asunto de los funcionarios públicos sino de la ciudadanía, quien además de aportar nuevos diagnósticos puede solucionarlos de manera colaborativa. En este nuevo modelo de interacción ciudadano - Estado juegan un rol indispensable la disponibilidad y el acceso a la información pública como facilitadores de la colaboración, la participación y el control social.

En Colombia, el acceso a la información pública es un derecho consagrado en la Constitución de 1991 y los datos abiertos son una de las maneras más efectivas de hacer realidad este derecho. Desde el gobierno colombiano se promueve que la información creada, transformada y usada por la administración pública pertenezca a la sociedad; es decir, que el contenido pueda ser consultado, usado y compartido por todos los colombianos, y se encuentre siempre a su disposición, con el fin de que se creen productos o servicios de valor para dar solución a problemáticas públicas.

Esta guía se enmarca en la Política de Gobierno Digital y contribuye al logro del propósito de “toma de decisiones basadas en datos a partir del aumento en el uso y aprovechamiento de la información”, que aporten a la optimización y desarrollo de servicios, políticas, normas, planes, programas, proyectos de interés público o privado, adelantado por ciudadanos, empresas y el mismo Estado.

El presente documento proporciona orientaciones y buenas prácticas para el desarrollo de estrategias de apertura y uso de datos, encaminadas a la generación de valor económico y social en los distintos ámbitos de la sociedad. Este documento se desarrolla en cuatro apartados que le permiten al lector avanzar en su tarea de disponer y usar datos abiertos.

En un primer apartado se describen los datos abiertos en el contexto del marco legal colombiano y se presentan ejemplos sobre los datos abiertos de las distintas entidades del sector público; seguido por el recuento de los principios establecidos en la Carta Internacional de Datos Abiertos donde se establecen las buenas prácticas para la liberación de datos abiertos gubernamentales y los beneficios de usar datos abiertos en torno a temas de transparencia y control social, mejoramiento de servicios públicos, prevención de fenómenos y generación de nuevo conocimiento.

El tercer apartado describe de manera detallada los cinco componentes claves del marco de gobernanza de la Iniciativa Datos Abiertos en Colombia: a) marco normativo, b) actores, c) herramientas, d) incentivos y e) monitoreo; y finalmente, el conjunto de pasos y actividades que permiten gestionar el ciclo de vida de los datos abiertos en las entidades públicas: establecer el plan de apertura, estructurar y publicar los datos, comunicar y promover el uso de datos y monitorear la calidad y el uso.



# 1. Datos abiertos

Los datos abiertos son información pública dispuesta en formatos que permiten su uso y reutilización bajo licencia abierta y sin restricciones legales para su aprovechamiento. En Colombia, la Ley 1712 de 2014 sobre Transparencia y Acceso a la Información Pública Nacional, define los datos abiertos en el numeral sexto 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”.

De este modo, la Ley establece la obligatoriedad de las entidades públicas de “divulgar datos abiertos”, teniendo en cuenta las excepciones de acceso a la información, asociadas a información clasificada y reservada establecidas en su título tercero, artículos 18 y 19<sup>1</sup>. A continuación, se señalan algunos ejemplos de datos abiertos de entidades públicas disponibles en el Portal de Datos del Estado Colombiano:

- Los [resultados de las elecciones](#) dispuestos por la Registraduría Nacional del Estado Civil, con información detallada de las últimas jornadas electorales, tanto presidenciales, como territoriales y legislativas.
- Las [contrataciones públicas](#) dispuestas por la Agencia de Contratación - Colombia Compra Eficiente, donde se disponen todos los gastos en contratación que realiza el estado, permitiendo saber a quién le compra el Estado, cuanto gastó y en qué gasto.
- Los resultados sobre calidad educativa tales como las [pruebas Saber](#) dispuestas por el Instituto Colombiano para la Educación Superior – ICFES.
- Información de los [códigos únicos de medicamentos](#) para los registros sanitarios vigentes dispuestos por el Ministerio de Salud.

Los datos abiertos tienen como principal característica que están publicados en sitios web de manera gratuita, accesibles al público, dispuestos en formatos que permiten su uso, reutilización y aprovechamiento sin restricciones legales y bajo licencia abierta<sup>2</sup>. Algunos de los formatos de datos abiertos más comunes son:

- CSV (Valores Separados por Coma): es un tipo de documento de texto plano para representar datos tabulados en columnas separadas por comas (o punto y coma) y filas separadas por saltos de línea.
- XML (Lenguaje Etiquetado Extensible): es el estándar utilizado para el intercambio de información estructurada entre diferentes plataformas.

---

<sup>1</sup> Ley 1712 de 2014 de Transparencia y del Derecho de Acceso a la Información Pública. Recuperado el 2 del 4 de 2019, en [https://www.mintic.gov.co/portal/604/articles-7147\\_documento.pdf](https://www.mintic.gov.co/portal/604/articles-7147_documento.pdf).

<sup>2</sup> La licencia abierta de bases de datos es un acuerdo de licencia copyleft diseñado para permitir a los usuarios compartir, modificar y usar libremente una base de datos, manteniendo la misma libertad para los demás.

- RDF (Infraestructura para Descripción de Recursos): es un lenguaje de propósito general para representar información en la web, es una de las tecnologías esenciales para la web semántica.
- RSS (Sindicación Realmente Sencilla): es un formato basado en XML para la ordenación de contenidos de páginas web actualizados frecuentemente (noticias, blogs o podcasts).
- JSON (Notación de objetos Javascript): es un formato simple y ligero para el intercambio de datos, de fácil lectura e interpretación por las personas y sencillo para ser analizado y generado por las máquinas.
- ODF (Formato de Documento Abierto para Aplicaciones Ofimáticas): es un formato de fichero estándar para el almacenamiento de documentos ofimáticos.
- WMS (Servicio Web de Mapas): es un servicio definido por el OGC (Open Geospatial Consortium) que produce mapas de datos referenciados espacialmente, de forma dinámica a partir de información geográfica.
- WFS (Servicio Web de características): es un servicio estándar, que ofrece una interfaz de comunicación que permite interactuar con los mapas servidos por el estándar WMS.

Por otro lado, la definición del tipo de licencia bajo el cual se usarán estos datos abiertos se convierte en una de las principales actividades. Las licencias estándar en el entorno internacional se han constituido en un elemento de confianza digital en el uso de datos abiertos, ya que están concebidas como una forma de protección de los derechos de autor y de apoyo al dominio público de las obras protegidas, fomentando el acceso a la información pública, orientaciones sobre licencias pueden ser identificadas en el instructivo para la definición de licencias de datos, elaborado por la Unidad Administrativa Especial de Catastro Distrital - IDECA<sup>3</sup>.

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. Para el licenciamiento de los datos abiertos una alternativa es 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 abierta Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)<sup>4</sup> permite hacer uso, transformación, distribución, redistribución, reutilización, compilación, extracción, copia, difusión, modificación y/o adaptación de los datos abiertos, citando siempre la fuente.

---

<sup>3</sup> Unidad Administrativa Especial de Catastro Distrital – IDECA. Instructivo para la definición de licencias de datos. Recuperado el 22 de 4 de 2019, en <https://www.ideca.gov.co/sites/default/files/documentacion/instructivolicencadatos.pdf>.

<sup>4</sup> Commons, C. (s.f.). Attribution-ShareAlike 4.0 International (CC BY-SA 4.0). Recuperado el 22 de 4 de 2019, en <https://creativecommons.org/licenses/by-sa/4.0/deed.es>





La adopción de este tipo de licencia reconoce la producción o generación de los datos por la entidad publicadora, señalando que ésta no será responsable de la utilización o transformación de los mismos datos, ni tampoco de 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 manifestación regularmente la adopta la entidad en la plataforma de datos donde los dispone, en Colombia se encuentra en el Portal de Datos del Estado Colombiano.

## 1.1 Principios

Los principios y buenas prácticas para la liberación de datos abiertos han sido establecidos en la Carta Internacional de Datos Abiertos, adoptada por Colombia desde 2016 como instrumento orientador en la generación y uso de datos.

Los adherentes a la Carta Internacional de Datos Abiertos reconocen que los gobiernos y otras organizaciones del sector público poseen grandes cantidades de datos que pueden ser de interés para los ciudadanos, y que esos datos son un recurso reutilizable. En consecuencia, se acordó seguir los principios que sientan las bases para el acceso a los datos y para su publicación y uso. Además de las características básicas de los datos abiertos, los principios actuales que promueve la Carta son:

Ilustración 1. Principios de los Datos Abiertos



Fuente: Elaboración Propia

## 2. Beneficios de los Datos Abiertos

La efectiva reutilización de los Datos Abiertos puede generar múltiples beneficios que pueden ser vistos desde perspectivas económicas y sociales donde se pueden desarrollar procesos que lleven a la generación de valor en áreas como la transparencia y el control social, el mejoramiento o creación de productos o servicios, el mejoramiento en la eficiencia y eficacia de los servicios ofrecidos por el Estado, la medición predictiva del impacto de políticas, y la toma de decisiones en todos los niveles, a continuación se destacan algunos ejemplos de estos beneficios:

### 2.1 Transparencia y Control Social

Disponer la información pública del quehacer de las entidades en datos abiertos, permite que la sociedad civil, las entidades públicas y otros grupos de interés como academia y periodistas, realicen procesos de investigación y control social a sus gobernantes en torno a temas de manejo de dineros públicos, cumplimiento de promesas de campaña, inversión de recursos, entre otros. Algunos ejemplos de esto son:



El proyecto “La Letra Menuda de la Contratación”, de la [plataforma de periodismo de investigación y de datos](#), usó datos abiertos de compras públicas (más de 6 millones de registros disponibles en [www.datos.gov.co](http://www.datos.gov.co) dispuestos por Colombia Compra Eficiente) y analizó cuáles de estos procesos de contratación pueden ser susceptibles de corrupción. Ver proyecto en: <http://especiales.datasketch.co/contratos-colombia/>



Otra iniciativa de transparencia es “Elecciones y contratos”, solución tecnológica desarrollada a partir de un proceso de construcción colectiva entre [sociedad civil y entidades públicas](#)<sup>5</sup> que permitió conocer datos acerca de los candidatos a cargos de elección popular y sus financiadores. Esta solución aportó al control social y permitió la realización de cruces de información entre datos oficiales de financiación de campañas electorales suministrados por Cuentas Claras y datos de contratación pública suministrados por Colombia Compra Eficiente. Ver iniciativa en: <http://elecciones-contratos.monitorciudadano.co/>

---

<sup>5</sup> Entidades participantes: Transparencia por Colombia, la Corporación Somos Más, Datasketch y el Centro de Innovación Pública Digital de MinTIC.



## 2.2 Mejoramiento o creación de productos, servicios y negocios innovadores

La publicación y uso de datos abiertos ayuda a las entidades públicas a mejorar la prestación de servicios a los ciudadanos, así como a optimizar procesos internos. Disponer y analizar grandes cantidades de datos puede permitir:

- Conocer cuál es el tiempo de espera de un ciudadano para acceder a un servicio y proponer mejoras para optimizarlo.
- Saber qué trámites son los más demandados y mejorar la forma en que éstos se realizan.
- Entender en qué zonas del país o la ciudad un incidente o suceso es más recurrente y adelantar acciones para monitorear, controlar, reducir o aumentar los sucesos.



Un ejemplo de cómo los datos abiertos ayudan a mejorar los servicios de las **entidades públicas** es la solución móvil “Clic Salud” del Ministerio de Salud, la cual utiliza datos abiertos de las Instituciones Prestadoras de Salud (IPS), las Entidades Prestadoras de Salud (EPS) y los precios de los medicamentos en el país, con el fin de mostrar a los ciudadanos más información en tiempo sobre calidad en las EPS y comparar los precios de un mismo medicamento con base en su principio activo, para mejores decisiones a la hora de comprar, ahorrando dinero y tiempo. Enlace de descarga de la aplicación móvil:

<https://play.google.com/store/apps/details?id=com.minsalud.clicsalud&hl=es> 419

## 2.3 Pronosticar y prevenir fenómenos

El análisis de los datos permite pronosticar y prevenir fenómenos, mejorando la toma de decisiones. La analítica de datos estructurados y no estructurados ha sido muy útil en aspectos tales como prevención del crimen, pronóstico del tiempo y el análisis de enfermedades, pues a partir de la información existente ha sido posible crear o mejorar las políticas públicas para reducir el impacto negativo que presentan estos fenómenos.

Los datos abiertos permiten a los distintos actores del ecosistema (gobernantes, ciudadanos, periodistas) mapear el estado de un sector o tema y mejorar, por ejemplo, la cobertura en salud, el estado de las vías en una ciudad y la asignación de los recursos en una entidad, entre otros.



El proyecto **Aclímate Colombia**, plataforma desarrollada en el marco de una asociación intersectorial liderada por el Centro Internacional de Agricultura Tropical (CIAT), una organización de la **sociedad civil**, con grupos de la industria del **sector privado** y **entidades**

[públicas](#) que aprovecha una diversidad de fuentes de datos, para ayudar a los agricultores a entender cómo enfrentar mejor los patrones climáticos.

A partir de uso y análisis de datos agrícolas y ambientales este proyecto permitió:

- Identificar los mayores períodos secos que afectan la temporada de crecimiento y generan retrasos en la siembra.
- Evitar pérdidas significativas por siembras prematuras y condenadas al fracaso (1.800 hectáreas).
- Identificar municipios con mayor riesgo asociado a los periodos de sequía.
- Incorporar el uso de la información en la toma de decisiones sobre cuándo sembrar, qué productos sembrar.

## 2.4 Generar nuevo conocimiento

Los datos abiertos son insumo fundamental para la generación de nuevo conocimiento. Actores como la academia o el periodismo pueden sacar provecho de la disposición de datos de alta calidad. Artículos académicos, periodísticos e incluso estudios de mercado, pueden realizarse a partir del uso de datos abiertos.



[La mina que fulmina](#), es un completo análisis periodístico que, con base en datos sobre víctimas de minas antipersona en Colombia dispuestos en el Portal de Datos del Estado Colombiano [datos.gov.co](#), que detalla la situación de civiles, fuerza pública y guerrilleros en el país. Este desarrollo fue realizado por estudiantes de periodismo del Politécnico Grancolombiano en el marco del [concurso Datos a la U 2018](#) del Ministerio TIC.

## 3. Marco de Gobernanza

La Gobernanza de la Iniciativa Datos Abiertos en Colombia está respaldada por el nuevo enfoque de la política de Gobierno Digital, que busca la reconstrucción de los lazos de confianza entre el ciudadano y el Estado. Proporciona elementos claves que permitan estimular la liberación, un mejor acceso, un mayor intercambio, reutilización e impacto. En Colombia este marco de gobernanza se compone de 5 elementos claves:

Ilustración 2. Marco de Gobernanza de Iniciativa Datos Abiertos Colombia



Fuente: Elaboración Propia

### 3.1 Marco Normativo y Lineamientos

#### 3.1.1 Marco Normativo

En Colombia, la Constitución Política consagra el acceso a la información como derecho fundamental en los artículos 20 y 74, reconociéndolo como un derecho habilitante de otros derechos. Partiendo de esta premisa, Colombia adopta un marco legal que garantiza la transparencia y el derecho de acceso a la información pública a los ciudadanos en la [Ley 1712 de 2014](#)<sup>6</sup> que establece procedimientos para garantizar este derecho y se obliga a las entidades –en el marco de la seguridad y privacidad de la información respetando los datos reservados y clasificados- a publicar o liberar datos abiertos.

Como instrumentos normativos adicionales, el [Decreto 1078 de 2015](#) que desarrolla la Política de Gobierno Digital, con el propósito de toma de decisiones a partir de datos, así como el [Decreto 1081 de 2015](#) define pautas relacionadas con la forma en que las entidades públicas deben implementar lo definido en la ley y así mismo, [la Resolución 3564 de 2015 del Ministerio TIC](#) establece, entre otros aspectos, los estándares para la publicación de información, incluyendo las condiciones técnicas para la apertura de los datos abiertos en Colombia.

<sup>6</sup> Ley 1712 de 2014 de Transparencia y del Derecho de Acceso a la Información Pública. (6 de 03 de 2014). Recuperado el 2 de 4 de 2019, de [https://www.mintic.gov.co/portal/604/articles-7147\\_documento.pdf](https://www.mintic.gov.co/portal/604/articles-7147_documento.pdf)

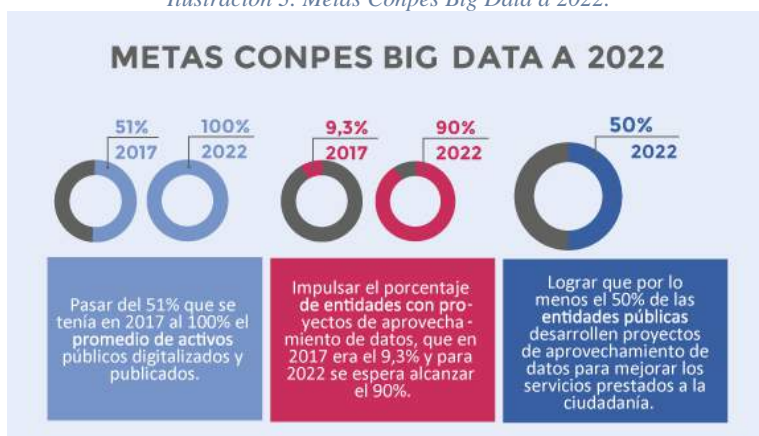
Complementario a estas disposiciones, el Gobierno nacional aprobó el documento [Conpes 3920 de 2018](#) Política Nacional de Explotación de Datos (Big Data) y con ello definió las condiciones habilitantes para aumentar el aprovechamiento de datos públicos y garantizar que estos sean gestionados como activos capaces de generar valor social y económico.

El principal objetivo de esta política es aumentar el aprovechamiento de datos en Colombia, mediante el desarrollo de las condiciones de apertura y uso de activos que generen valor social y económico. La política aborda cuatro ejes estructurales sobre los cuales el país debe trabajar y realizar acciones que permitan la explotación y el aprovechamiento de datos:

1. Generación de datos digitales: de calidad y con completitud.
2. Cultura de datos: para propiciar que haya demanda de datos y conciencia en la generación de valor.
3. Capital humano para la explotación de datos: personas capaces de producir, procesar y reutilizar datos para resolver problemáticas.
4. Marco jurídico, ético e institucional: para habilitar la generación de valor, y proteger de los ciudadanos en el contexto de disponibilidad y explotación masiva de datos.

Con la implementación de esta política se espera aumentar el promedio de activos públicos digitalizados y publicados y que las entidades públicas desarrollen proyectos de aprovechamiento de datos para mejorar los servicios prestados a la ciudadanía.

*Ilustración 3. Metas Conpes Big Data a 2022.*



Fuente: Elaboración Propia

### 3.1.2 Lineamientos

En desarrollo del marco normativo, los lineamientos son aquellas directrices o disposiciones que deben ser implementadas por las entidades públicas para el desarrollo de la Política de Gobierno Digital y en consecuencia la apertura y aprovechamiento de datos abiertos. Estos lineamientos se desarrollan a través de estándares, guías, recomendaciones o buenas prácticas que buscan servir de instrumento, para que los diferentes actores del ecosistema de datos implementen sus planes de apertura y uso.

En Colombia se han generado diferentes lineamientos desde el Ministerio TIC y otras entidades líderes de política, anexos a esta Guía que establecen directrices específicas sobre la gestión y el uso de los datos, entre los cuales se destacan:

- [Guía de Estándares de Calidad e Interoperabilidad](#): proporciona orientaciones y buenas prácticas para el desarrollo de estrategias de apertura de datos que cumplan con los principales estándares de calidad e interoperabilidad de datos abiertos.
- [Marco de arquitectura empresarial](#): es el principal instrumento para implementar la Arquitectura TI de Colombia y habilitar la Política de Gobierno Digital.
- [Marco de Interoperabilidad](#): determina los lineamientos y recomendaciones para el intercambio eficiente de información entre entidades del Estado.
- [Información](#): es el estándar definido por el Estado colombiano para intercambiar información entre organizaciones, facilitando el entendimiento de los involucrados en los procesos de intercambio de información.
- [Guía de anonimización](#): tiene como propósito orientar a las entidades integrantes del Sistema Estadístico Nacional sobre el proceso de anonimización de bases de datos que provienen de registros administrativos y de operaciones estadísticas.
- [Lineamientos técnicos para la producción y Gestión de información geográfica](#): presenta un marco general de normas técnicas que permitan consolidar bajo una estructura estándar la información geográfica de las entidades del Estado.

## 3.2 Ecosistema de Actores

En términos de gobernabilidad, el Ministerio TIC es el principal responsable del diseño, la implementación y la promoción de la política de Datos Abiertos en Colombia, y trabaja de manera coordinada con las entidades líderes de la política de Transparencia y Acceso a la Información Pública como la Secretaría de Transparencia de la Presidencia de la República, el Departamento Administrativo de la Función Pública - DAFP, el Departamento Nacional de Estadística - DANE, el Departamento Nacional de Planeación -DNP, el Archivo General de la Nación y con el apoyo de la Consejería Presidencial para la Innovación y la Transformación Digital.

Además de las entidades públicas, el ecosistema de actores incluye a la academia, desarrolladores, periodistas, sociedad civil y organismos multilaterales como la Organización de Estados Americanos (OEA), el Banco Mundial y la Organización para la Cooperación y el Desarrollo Económicos (OCDE) que han participado en Colombia en el impulso, promoción y desarrollo de la iniciativa.

Ilustración 4. Ecosistema de Actores de Datos Abiertos.



Fuente: Elaboración Propia

### 3.3 Herramientas

El Ministerio TIC ha dispuesto el Portal de Datos del Estado Colombiano <http://www.datos.gov.co/>, como el espacio digital para la disposición de la información pública estructurada del país y el punto de encuentro del ecosistema para su uso y aprovechamiento. Éste es operado y administrado por la Dirección de Gobierno Digital y reúne las herramientas que se proveen a los ciudadanos y entidades para publicar, acceder y usar los datos abiertos de Gobierno.

A enero de 2019, más de 1.100 entidades públicas de todo el país han dispuesto alrededor de 10.000 conjuntos de datos en temáticas tales como salud, educación, agricultura, movilidad y seguridad. Algunas de las principales funcionalidades de este portal son:

- **Sección participa:** permite a cualquier ciudadano hacer solicitudes de apertura de datos a las entidades públicas, las cuales son gestionadas a través del Ministerio TIC, y las respuestas son publicadas en un conjunto de datos en el portal. Ver Sección participa en <https://datos.gov.co/nominate>
- **Sección de herramientas para desarrolladores:** el portal expone una interfaz de aplicaciones (API) que permite consumir los conjuntos de datos automáticamente para que los desarrolladores pueden hacer uso de los datos para sus aplicaciones. Ver sección para desarrolladores en: <https://herramientas.datos.gov.co/es/content/desarrollar-usando-los-datos>
- **Sección de usos:** permite a los ciudadanos y entidades publicar y conocer las soluciones desarrolladas a partir de datos abiertos tales como aplicaciones web, móviles, sitios web de periodismo de datos, proyectos académicos y visualizaciones. Ver sección de usos en <https://herramientas.datos.gov.co/es/usos>



- **Sección de Calidad:** permite a los usuarios la generación de reportes de calidad de datos, permitiendo conocer la cantidad de errores más comunes por entidad, las categorías de datos con más errores y las tendencias de estos, esta información es utilizada para hacer seguimiento de la publicación y tomar decisiones para la mejora de estos. Ver sección de reportes en <https://herramientas.datos.gov.co/es/calidaddedatos>
- **Federación de portales externos de datos abiertos:** Esta funcionalidad ofrece la posibilidad de integrar de manera federada otros portales de datos abiertos de entidades públicas, por ejemplo, el Portal de la Alcaldía de Bogotá <http://datosabiertos.bogota.gov.co/>.

### 3.3 Esquema de incentivos

Como uno de los mecanismos para dinamizar el ecosistema de datos se han diseñado incentivos para la disposición y uso de datos abiertos, algunos de estos son:

- Capacitación y asesoría especializada presencial y virtual de datos abiertos, a nivel nacional y territorial para entidades públicas, desarrolladores, empresarios y estudiantes.
- Cursos certificados para servidores públicos que promuevan la apertura y el uso.
- Concursos de innovación, colaboración y aprovechamiento.
- Financiación de proyectos específicos.

Entre los programas más destacados de la iniciativa Datos Abiertos se encuentran:

- **Datos a la U:** dirigido a estudiantes y docentes, busca promover el uso de datos en la academia, visibilizar los trabajos realizados por los equipos académicos de cualquier nivel de formación y generar reconocimiento. Ver sitio web del concurso: <https://datosalau.datos.gov.co/686/w3-propertyvalue-40781.html>
- **Máxima Velocidad:** Estrategia de aceleración en la que se plantean retos de apertura, uso y calidad de los datos abiertos con el fin de dinamizar la generación de datos abiertos de valor en las entidades públicas. Ver sitio web de la estrategia: <http://www.maximavelocidad.gov.co/647/w3-channel.html>
- **Eventos de innovación:** maratones de desarrollo, DataJams, ejercicios de visualización, [calidatones](#) y retos de control social, aperturando y aprovechando datos. Ver eventos en <https://herramientas.datos.gov.co/es/blog/>

### 3.4 Monitoreo

Lo que no se mide no se puede mejorar. En tal sentido es preciso que los publicadores de datos establezcan mecanismos para medir la calidad y el uso de estos. Ministerio TIC cuenta con algunas alternativas para monitorear la apertura y el uso de los datos abiertos tales como:

- **FURAG:** el Formulario Único Reporte de Avances de la Gestión, mide anualmente a las entidades del orden nacional y territorial el avance en las políticas de desarrollo administrativo, incluidos los indicadores de proceso y resultado relacionados con la

apertura y uso de datos abiertos en el marco de la Política de Gobierno Digital y Acceso a la Información Pública. Los resultados de la medición están disponibles en: <http://estrategia.gobiernoenlinea.gov.co/623/w3-propertyvalue-7651.html>

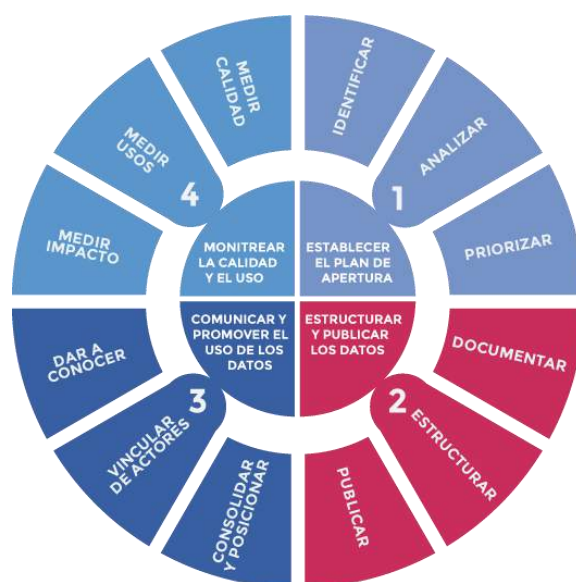
- **Sello de excelencia:** instrumento que permite garantizar las condiciones de calidad necesarias para acceder a los servicios digitales en forma confiable. El sello de excelencia establece tres niveles de calidad en datos abiertos y cada nivel establece unos requisitos mínimos que se deben cumplir. Ver requisitos en: [http://estrategia.gobiernoenlinea.gov.co/623/articles-17464\\_recurso\\_32.pdf](http://estrategia.gobiernoenlinea.gov.co/623/articles-17464_recurso_32.pdf)  
El corazón del Sello de Excelencia es la comunidad virtual, integrada por academia, industria TI, ciudadanos, entidades públicas y organizaciones internacionales. Esta comunidad interactúa bajo un esquema de colaboración, promueve procesos de innovación en las entidades participantes, además de verificar y validar datos abiertos. Información sobre el proceso de certificación en: [www.sellodeexcelencia.gov.co/](http://www.sellodeexcelencia.gov.co/)
- **Índices internacionales:** Colombia se evalúa en diferentes mediciones internacionales. Estos índices han permitido analizar las tendencias mundiales en términos de datos abiertos, comparar resultados y estrategias con otros países de la región y evaluar la evolución de la iniciativa en términos de apertura y uso, así como el impacto esperado que tienen los datos abiertos en el mundo.

Las principales mediciones se realizan a través de los índices internacionales de datos abiertos: Open Data Index, Open Data Barometer, OUR Data Index de la OCDE. Ver Anexo I: índices internacionales y marco normativo que fijan prioridades de publicación.

## 4. Ciclo de Datos Abiertos

Establecer un ciclo para el proceso de apertura, mejora y uso de datos abiertos permite que los actores del ecosistema definan acciones a tomar en cada una de las fases estratégicas de este ciclo: construcción de plan de apertura, estructuración, comunicación y monitoreo. La buena implementación y seguimiento a estas actividades redundará en calidad de datos y, por tanto, en usos de impacto. Adicionalmente es un proceso recurrente, que debe presentarse con una vigencia definida, como recomendación este ciclo se planea y ejecuta cada dos años.

Ilustración 5 Ciclo de Vida de los Datos.



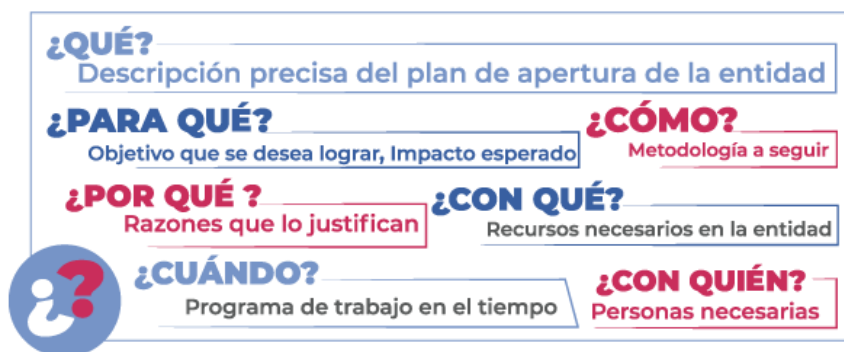
Fuente Elaboración Propia

### 4.1 Establecer el plan de apertura, mejora y uso de datos

A partir del Marco de Gobernanza de datos abiertos y de acuerdo con el marco normativo vigente es posible adelantar una estrategia para adelantar el proceso de apertura, identificar los actores que harán parte del ecosistema incluyendo publicadores, usuarios e incluso reutilizadores, usar las herramientas disponibles para apertura, calidad y uso de sus datos abiertos, diseñar un esquema de incentivos y generar estrategias de monitoreo al valor de los datos abiertos.

En esta fase inicial de planificación se generan las directrices y se define el marco de gobernanza de la iniciativa de datos a planificar, esto incluye la identificación del ecosistema y sus actores, las tareas a realizar, los recursos, los tiempos, los productos que se obtendrán, el esfuerzo que se destinará para las diferentes tareas y el impacto esperado de los datos a abrir. Para ello es importante trazarse unos objetivos claros que respondan a las preguntas claves en un proceso de planeación:

Ilustración 6. Preguntas claves en un plan de apertura de datos.



Fuente Elaboración Propia

El plan de apertura, mejora y uso de datos abiertos incluye las siguientes actividades:



### 4.1.1 Identificar

En este paso es necesario identificar toda la información a tratar y los actores con los que se debe interactuar en el desarrollo del ciclo de datos.

#### 4.1.1.1 Identificar el ecosistema de interacción

Con el fin de contar con el apoyo y soporte del ecosistema en el que interactúa la entidad en cada una de las fases del ciclo de datos, desde el establecimiento del plan hasta el monitoreo y calidad de uso, es necesario definir e identificar los tipos de actores y actores con los que la entidad tiene contacto y que son considerados como de interés por el rol que cumplen en el ecosistema (usuario, regulador, proveedor, etc.)

Para desarrollar esto se recomienda seguir los lineamientos incluidos en la “Guía de innovación abierta por medios electrónicos”. Esta puede ser descargada al seguir el siguiente vínculo: [https://estrategia.gobiernoenlinea.gov.co/623/articles-8250\\_Guiainnovacion.pdf](https://estrategia.gobiernoenlinea.gov.co/623/articles-8250_Guiainnovacion.pdf)

El valor agregado que generará interactuar con el ecosistema será descrito en cada uno de los pasos del ciclo. Por ejemplo, en la de identificación de los activos de información, acudir al ecosistema servirá para consultar qué activos de información son de interés para los actores en el ecosistema.

Como ejemplo a continuación se encuentra el listado parcial de tipos actores identificados del ecosistema en el que se desenvuelve el Ministerio de Salud y Protección Social<sup>7</sup>

- Asociaciones de pacientes
- Asociaciones de usuarios
- Veedurías ciudadanas
- Academia Nacional de Medicina
- Asociaciones y sociedades científicas
- Academia / Universidades
- Organismos No Gubernamentales
- Organismos internacionales (Organización Panamericana de la Salud, Banco Interamericano de Desarrollo, Banco Mundial, etc.)
- Industria de tecnologías de la salud y sus agremiaciones

A modo de recomendación la identificación de estos actores se facilita al identificar los roles que se llevan a cabo en el ecosistema. A partir de los roles (ej. usuarios) pueden identificarse qué tipos de actores desempeñan ese rol (ej. Asociaciones de pacientes). Finalmente, para cada tipo de actor ya pueden identificarse a los actores asociados a cada uno (ej. Banco Mundial es un actor de Organismos Internacionales).

#### **4.1.1.2 Identificar los activos de información**

La identificación de los activos de información busca realizar un levantamiento de la información pública que administra la entidad y que puede llegar a convertirse en datos abiertos. Para efectos de realizar la identificación de esta información y de conformidad con lo establecido en la normatividad vigente, se debe tener en cuenta durante esta fase los siguientes datos:

- Los datos de los registros de activos de información oficiales generados por la entidad clasificados como dato abierto.
- Los datos públicos generados o administrados por los Sistemas de Información de la entidad.
- Los datos públicos utilizados por herramientas digitales, aplicativos web o móviles desarrollados de la entidad.
- Los datos solicitados por la ciudadanía de manera recurrente mediante solicitudes de acceso a la información o los que deriven de ellas; también aquellos que puedan resolver alguna problemática de la entidad.
- Los datos de otras entidades o de dependencias al interior de la misma entidad con el fin de evitar la duplicidad a la hora de publicar datos en el portal. Para ver el inventario

---

<sup>7</sup> Esta información fue tomada del “MANUAL PARA LA PARTICIPACIÓN DE LOS ACTORES DEL SISTEMA GENERAL DE SEGURIDAD SOCIAL EN SALUD EN EL DESARROLLO DE LAS GUÍAS DE ATENCIÓN INTEGRAL – GAI” desarrollado en 2010 por el Ministerio de Salud y está disponible en: <http://gpc.minsalud.gov.co/Documents/MANUAL%20DE%20PARTICIPACION%20GAI.pdf>



de todos los datos abiertos existentes en el Portal de Datos del Estado Colombiano consulte:

<https://www.datos.gov.co/dataset/Vista-Inventario-de-Datos-Portal-Nacional-de-Datos/5vw4-mas9>

Para este momento, es importante vincular a los actores del ecosistema de datos, buscando que, desde la apertura de los datos, se atiendan necesidades específicas de sociedad civil, academia y sector privado, para darle continuidad al ciclo de datos y garantizar la pertinencia y oportunidad de estos. Así mismo consultar con ellos qué necesidades de información tienen, qué información es valiosa, cuál consideran que no es accesible, etc.

De esta manera, se contribuye a promover la reutilización de datos de la entidad y difundir soluciones desarrolladas con datos abiertos. Algunas actividades que se pueden adelantar para vincular a los actores del ecosistema son:

- Desarrollar procesos de co-creación de herramientas tecnológicas o aplicaciones para identificar conjuntos de datos que generarían valor al disponerlos en formato abierto, con participación de desarrolladores y usuarios. Más información sobre el desarrollo de estos ejercicios en la Guía de Innovación Pública, disponible en el siguiente enlace: [http://estrategia.gobiernoenlinea.gov.co/623/articles-8250\\_Guiainnovacion.pdf](http://estrategia.gobiernoenlinea.gov.co/623/articles-8250_Guiainnovacion.pdf).
- Realizar talleres o encuentros con usuarios o grupos de interés para generar propuestas de creación de conjuntos y 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.

Una vez se ha identificado los datos abiertos de la entidad, se debe identificar las fuentes de información que generan dichos conjuntos de datos y determinar si la generación de estos se realiza de manera manual (disponible en hojas de cálculo, por ejemplo) o automática (proviene de un sistema de información).

Si la fuente de los datos es un sistema de información, se debe identificar y priorizar la adecuación y desarrollo de las funcionalidades requeridas en el sistema de información (software), para que de esta manera el proceso de carga de los datos abiertos sea automatizado a través de proceso de extracción, transformación y carga.

#### **4.1.1.2.1 Caracterización de los activos de información identificados**

Para esto es importante considerar como punto de partida la información con la que cuenta la entidad en el cumplimiento de la Ley 1712 de 2014 de Transparencia y Acceso a la Información Pública, puesto que este paso debe surtir sobre los activos de información que tiene identificados la entidad (además de los que pueda producir en respuesta a necesidades del ecosistema).

- a) Verificar los siguientes aspectos de los activos de información identificados.



- Identificar la dependencia responsable del desarrollo y actualización de información
- Identificar si la información se genera de forma manual y/o automática
- Identificar la periodicidad de la actualización de esta información
- Identificar si el activo de información puede ser considerado como un conjunto de datos

b) Diligenciar las siguientes preguntas a cada uno de los conjuntos de datos públicos identificados en el inventario de activos de información de la entidad:

- ¿Este conjunto de datos es publicado en Datos Abiertos?
- ¿Este conjunto de datos podría ser sometido a un proceso de automatización en su actualización?
- ¿Este conjunto de datos está sujeto a un proceso de federación con otras fuentes de información?
- ¿Este conjunto de datos podría ser sometido a un proceso de calidad de datos y mejora de información?

Se recomienda que esta información se diligencie en una matriz en la que cada registro (fila) es un activo de información y las variables (columnas) son las preguntas relacionadas en la parte superior.

Ejemplo:

**Identificación y Caracterización de los activos de información para el Plan de apertura, mejora y uso de Datos de la Entidad XXXX**  
 Descripción: Identificación y caracterización de los activos de información a ser dispuestos o mejorados como conjuntos de datos abiertos en [www.datos.gov.co](http://www.datos.gov.co)

#	Identificado por:	Activo de Información	Responsable	Es un Conjunto de Datos (SI/NO)	Desarrollo Automático/ Manual de la información	Periodicidad de Actualización	Está en Datos Abiertos (SI/NO)	Vinculo Datos	¿Podría ser automatizada su actualización? (Ya está automatizado, SI, NO)	Sujeto a procesos de Federación (SI/NO)	¿Podría ser sometido a un proceso de calidad de datos y mejora de la información? (SI/NO)
1	Entidad X	Ejemplo 1	Dependencia Y	SI	Manual	Anual	NO	N/A	SI	NO	SI
2	Entidad X	Ejemplo 2	Dependencia X	NO	Automático	Mensual	SI	<a href="http://www.datos.gov.co">www.datos.gov.co</a>	Ya está automatizado	NO	NO
...	...	...	...	...	...	...	...	...	...	...	...
N	Entidad X	Ejemplo N	Grupo Z	SI	Manual	N/A	NO	N/A	NO	SI	SI
N+1	Academia - UNAL	Información Y	Grupo Z	SI	Manual	N/A	NO	N/A	NO	SI	SI



### TENER EN CUENTA

El papel de las entidades que identifican y publican datos abiertos es fundamental, pues se requiere un aumento de la oferta de datos abiertos. Para lograr un mejor y mayor uso de los datos es importante identificar cómo está conformada la demanda de los datos abiertos, sus necesidades y expectativas y generar la oferta en atención a esta demanda. 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 periodismo, entre otros.



## 4.1.2 Analizar

El objetivo de esta etapa es analizar la información identificada en la primera fase del proceso, para determinar la información publicable y validar que no tenga ninguna restricción normativa bajo la Ley Colombiana al igual que para definir las acciones a realizar en los activos de información publicables con el propósito de aperturar, mejorar o usar dicha información.

Los datos abiertos no son contrarios a la protección de datos personales: si existe información asociada a datos personales que puede ser valiosa como dato abierto, se debe adelantar un proceso de anonimización para eliminar aquella información sensible que afecta la privacidad de personas u organizaciones.

Anonimizar los datos es un aspecto fundamental y obligatorio antes de la publicación de datos personales sensibles. La anonimización es un proceso técnico que hace que la información de identificación personal no sea visible, logrando que no sean identificables los individuos a los que pertenece la información. Este proceso permite ocultar o eliminar los datos o características que identifique plena o parcialmente a las personas, organizaciones o características individuales de la fuente de información de manera irreversible, de tal forma que impidan identificar a una persona u organización.

Antes de publicar cualquier información que sea sensible, se deben considerar otras fuentes de información disponibles y es necesario evaluar si la combinación de éstas puede 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 de 2014 de Transparencia y del derecho de acceso a la información pública establece que las entidades deben “publicar datos abiertos”, teniendo en cuenta las excepciones de publicar información pública clasificada o información pública reservada, por lo que en esta fase se hace necesario analizar los datos identificados en el marco de la legislación vigente:

Ilustración 7. Información clasificada y reservada.



Fuente Elaboración Propia

### 4.1.2.1 Determinar las acciones a realizar

Una vez definido el conjunto de activos de información publicables que son responsabilidad de la entidad, es necesario definir qué acciones realizará la entidad respecto a los mismos para poder programarlos dentro de su respectivo plan de apertura, mejora y uso de datos.

La recomendación es que dentro del listado de los activos de información identificados en el paso anterior es vincular a cada activo de información la(s) acción(es) que se registrarán a ser realizadas en el plan de datos de la entidad. Información que también puede usar como insumo la interacción con el ecosistema, los resultados de una consulta pública o de acciones de co-creación.

Acciones:

- Abrir:** disponer el conjunto en Datos Abiertos [www.datos.gov.co](http://www.datos.gov.co)
- Automatizar:** procesos para que el cargue automatizado de la información
- Actualizar:** actualización de los registros de la información conforme a la periodicidad del conjunto de datos.
- Mejorar:** acciones de mejora y calidad de datos.
- Usar:** identificación de usos de datos para generar valor agregado a la información disponible.

### 4.1.3 Priorizar y Programar

La priorización de datos de valor es importante para que los esfuerzos de gobernanza, apertura y publicación estén enfocados en la provisión de datos de utilidad, valor y con potencial de crear impacto en la ciudadanía y de solucionar problemáticas públicas. Adicionalmente se recomienda acudir al ecosistema para averiguar, bien sea por consulta

pública o actividades de co-creación, los criterios de priorización que son identificados por los actores de interés de la entidad. Por lo anterior, se recomienda tener en cuenta lo siguiente a la hora de realizar la priorización de datos abiertos que serán incluidos en el Plan de Apertura:

- Los datos que permiten dar cumplimiento a metas institucionales de transparencia, participación, control social y colaboración.
- Los datos que contribuyen a un propósito de política pública específico y que hagan parte de las categorías establecidas por las mejores prácticas a nivel nacional e internacional en la materia (Ver Anexo I: índices internacionales y marco normativo que fijan prioridades de publicación), y aquellos solicitados directamente por otras entidades públicas.
- Los datos que favorecen al cumplimiento de las metas trazadas en el [Plan Nacional de Desarrollo](#) y los planes departamentales y municipales.
- Los datos que aportan al cumplimiento de los [Objetivos de Desarrollo Sostenible](#).
- Los datos abiertos de gran tamaño, alta concurrencia de usuarios y de permanente actualización, esto con el fin de que se planee la construcción de procesos automáticos de extracción, transformación y carga de datos (ETL) desde los sistemas de información de la entidad al Portal de Datos del Estado Colombiano [www.datos.gov.co](http://www.datos.gov.co).
- Los datos que han sido solicitados por los ciudadanos en ejercicios virtuales o presenciales de consulta pública.
- Los datos han sido caracterizados como prioritarios por uno o varios actores de interés del ecosistema.

#### 4.1.3.1 Programar

Para el desarrollo del plan adicionalmente es necesario definir un responsable y una fecha de implementación, este plan debe tener vigencia durante dos años, por lo que las acciones a realizar deben estar programadas en dicho periodo de tiempo (24 meses). Con el fin de determinar la fecha de implementación de las acciones determinadas para cierto activo de información o conjunto de datos se recomienda realizar ya haber desarrollado priorización de estos.

Para el desarrollo de este paso se documenta por activo de información o conjunto de datos, quién lo identificó, si es considerado como prioritario, la acción a realizar, el responsable junto con su correo de contacto, y la fecha de implementación.

Cuando esta información se encuentre diligenciada y se cuente con acciones, responsables y fechas para realizar acciones de apertura o mejora de datos se considerará que la entidad a diligenciado en su totalidad su plan de apertura, mejora y uso de datos.

### Plan de apertura, mejora y uso de datos de la Entidad XXXX

Descripción: Listado de acciones, tiempos y responsables para la apertura, actualización, automatización y uso de los conjuntos de datos de la entidad.

#	Identificado por:	Conjunto de Datos	Prioritario (SI / NO)	Acción(es) a tomar (Ninguna, abrir, automatizar, actualizar, mejorar y usar)	Responsable (Sólo para prioritarios)	Correo Responsable (Sólo para prioritarios)	Fecha Implementación (Sólo para prioritarios)
1	Entidad X	Ejemplo 1	NO	Ninguna	N/A	N/A	N/A
2	Entidad X	Ejemplo 2	SI	Actualizar	Servidor Público X	personax@entidad.gov.co	30/04/2020
...	...	...	...	...	...	...	...
N	Entidad X	Ejemplo N	NO	Ninguna	N/A	N/A	N/A
N+1	Actor 1	Información X	NO	Aperturar	N/A	N/A	N/A
N+2	Actor 2	Información Z	SI	Automatizar y Ajustar en Calidad	Servidor Público Y	personay@entidad.gov.co	1/08/2021



## RESULTADO

Como resultado de esta fase se tendrá el Plan de Apertura, Mejora y Uso de Datos de la entidad con el inventario de datos priorizados a publicar.

## 4.2 Estructurar y publicar

Luego de contar con el inventario de datos abiertos priorizados en el Plan de Apertura generado de la fase anterior, se requiere realizar una preparación para publicar los datos abiertos, la cual consiste en alistar los datos para que éstos puedan ser procesados y reutilizados por terceros. Este alistamiento incluye tres actividades:

Documentar

Estructurar

Publicar

### 4.2.1 Documentar

La documentación de los datos consiste en definir los metadatos que tendrá cada uno de los conjuntos de los datos a publicar. Los metadatos son características que describen los aspectos básicos de un conjunto de datos y son una herramienta fundamental para organizar, clasificar, relacionar y encontrar datos. 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 ejemplos de metadatos son los siguientes:

- **Nombre del conjunto de datos:** título que describe de manera corta el conjunto de datos. Este debe ser suficientemente descriptivo para el ciudadano por lo cual se recomienda usar lenguaje claro y eliminar las siglas, caracteres especiales o abreviaturas del nombre del conjunto.
- **Descripción:** resumen descriptivo del conjunto de datos, en donde se explica el contenido, procedencia, contexto en tiempo y lugar, así como 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 conjunto de datos en sistemas de categorías o taxonomías disponibles tales como Agricultura, Transporte, Salud, Educación, etc.
- **Licencia:** licencia bajo la cual se publica el conjunto de datos. Ver tipos de licencias abiertas en <https://co.creativecommons.org/> y en <https://www.ideca.gov.co/sites/default/files/documentacion/instructivolicencadatos.pdf>.
- **Fecha de Publicación:** fecha en la que se publicó el conjunto de datos.
- **Periodicidad:** indica la frecuencia de actualización del conjunto de datos.

La cantidad y características de los metadatos, dependerá de la plataforma en donde serán publicados los conjuntos de datos. El Portal de Datos del Estado Colombiano <http://www.datos.gov.co/> usa el estándar de metadatos **DCAT**<sup>8</sup> para describir los conjuntos de datos dentro de los catálogos publicados, lo que aumenta el potencial de ser encontrados en búsquedas en la web, y ayuda a que aplicaciones externas puedan consumir fácilmente los conjuntos de datos disponibles.

Para el caso de los datos geográficos frecuentemente se encuentran documentados a través de metadatos geográficos, que siguen referentes internacionales como la ISO19115 – “Geographic information – Metadata”. Este estándar contiene los elementos necesarios para facilitar el descubrimiento de datos y su acceso. Contiene además los elementos mínimos del metadato propuesto por DCAT y [www.datos.gov.co](http://www.datos.gov.co/)<sup>9</sup>

Adicionalmente, es necesario documentar cada dato a publicar en el conjunto de datos, esto se realiza mediante el estándar de [Lenguaje Común de Intercambio de Datos](#), que permite representar el concepto teniendo en cuenta sus relaciones, validaciones y estructura.

---

<sup>8</sup> DCAT (Data Catalog Vocabulary) es un vocabulario RDF diseñado para facilitar la interoperabilidad entre los catálogos de datos publicados en la Web.

<sup>9</sup> En Colombia, la ISO 19115 es la más adoptada y en la Infraestructura Colombiana de Datos Espaciales se pueden encontrar documentos y guías que facilitan su uso, a través del Catálogo Nacional de Metadatos, basado en un software libre que puede ser empleado por toda la comunidad nacional. <http://metadatos.igac.gov.co/geonetwork/srv/spa/catalog.search;jsessionid=BC0D138C0B1A671D9C0D2AB02D6A8914?node=srv#/home>





## ¡RECORDAR!

La documentación de datos:

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

### 4.2.2 Estructurar

Consiste en representar los datos para que se dispongan en formatos procesables por máquina y abiertos, asegurando que los datos estén disponibles de modo que puedan ser utilizados por computadoras y usuarios. La estructuración de datos busca transformar el dato de formatos no-textuales a formatos que permitan su manipulación y/o su conexión con otros datos (datos enlazados).

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, permitan un uso sin restricciones y se encuentren dentro de los más usados por los reutilizadores. No deben utilizarse formatos cerrados o no editables como por ejemplo información en formato PDF.

Los datos se desactualizan con el tiempo, por ello, las herramientas y procesos de calidad de datos actúan sobre los datos para corregir su deterioro. Existen diversos tipos de formatos para publicar datos abiertos, por lo cual, se recomienda seguir los siguientes formatos de publicación:

- **Datos tabulares:** Se recomienda el formato CSV, en los cuales las cadenas de caracteres que conforman los datos se acomodan en filas separadas por saltos de línea y columnas separadas por comas.
- **Datos estructurados:** Se recomienda el uso de los formatos JSON o XML cuya especificación se encuentra disponible abiertamente.
- **Datos espaciales:** Se recomienda el uso de los formatos GeoPackage, SHP, GeoJSON, WMS, WFS, KML, GeoTIFF

A fin de que los datos sean más útiles, estos deben ser fáciles de encontrar y comparar por diferentes variables como sectores, localización geográfica, fechas, entre otros, por lo cual, se deben emprender acciones para mejorarlos tales como:

1. La interoperabilidad de los datos publicados mediante su estandarización

Para estandarizar los datos abiertos las áreas generadoras o administradoras de Datos deberán aplicar:

- [El Marco de Interoperabilidad](#) que le permitirá cumplir con los estándares establecidos para el intercambio de datos abiertos entre distintas entidades públicas.
  - [El Lenguaje Común de Intercambio de información](#) que es el estándar definido para intercambiar información entre las entidades, facilitando el entendimiento de los involucrados en los procesos de intercambio de información.
  - [Lineamientos técnicos para la producción y Gestión de información geográfica](#) para los datos abiertos geográficos, los cuales frecuentemente se encuentran en formatos que solo pueden ser legibles por software de Sistemas de Información Geográfica. Sin embargo, sin importar el tipo de licencia del software en el que se estructuraron los datos geográficos, estos deben ser dispuestos en formatos interoperables que permitan que cualquier usuario pueda hacer uso de ellos en un software libre.
2. La usabilidad de los datos, mejorando la calidad, legibilidad y lógica de publicación de los conjuntos y recursos de datos

Para incrementar la usabilidad de los Datos Abiertos, es necesario tener en cuenta que:

- Los datos abiertos de alta calidad son una condición previa para analizarlos, reutilizarlos y garantizar su valor.
- Esta característica de los conjuntos de datos determina qué tanto se puede confiar en ellos para reutilizarlos y transformarlos en productos, servicios o nuevas aplicaciones.
- Es importante evaluar la calidad de los datos para verificar su exactitud, confiabilidad y aptitud para el uso. Para tal fin se puede hacer uso del [guía de estándares de Calidad e Interoperabilidad de los Datos Abiertos del Gobierno de Colombia](#).
- En el caso de datos geográficos, se debe monitorear o evaluar su calidad antes de ser publicados. Para tal fin se puede hacer uso del estándar internacional ISO 19157 – “Geographic information - Data Quality” o de los instrumentos en idioma español que en Colombia existen para su adopción en [www.icde.org.co](http://www.icde.org.co)<sup>10</sup>. Las evaluaciones de calidad más recomendadas son:
  - a) Exactitud posicional (absoluta o relativa): evalúa que los datos de ubicación o posición coincidan con la realidad, teniendo en cuenta un sistema de referencia por coordenadas.
  - b) Totalidad: disponer de la totalidad de información espacialmente generada, atendiendo a eventuales restricciones en alguno(s) de sus atributos por las excepciones que dispone la Ley 1712 de 2014.
  - c) Consistencia de formato/Usabilidad: se mide el nivel de uso del paquete de datos geográficos a publicar, la funcionalidad y capacidad de accesibilidad a través de los portales que empleen de visitantes, es decir, que se encuentren en formatos interoperables.

---

<sup>10</sup> Algunos recursos se encuentran en:

Plantilla: [http://www.icde.org.co/sites/default/files/Plantilla\\_Calidad\\_ISO19157.pdf](http://www.icde.org.co/sites/default/files/Plantilla_Calidad_ISO19157.pdf)

Guía: [http://www.icde.org.co/sites/default/files/Gu%C3%ADa\\_Evaluaci%C3%B3n\\_Calidad\\_V.1.1.pdf](http://www.icde.org.co/sites/default/files/Gu%C3%ADa_Evaluaci%C3%B3n_Calidad_V.1.1.pdf)

### 4.2.3 Publicar

Este paso o actividad consiste en poner a disposición de los usuarios el conjunto de datos abiertos a través del cargue de datos en una plataforma que permita la organización y fácil consulta 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.

El **Portal de Datos del Estado Colombiano** [www.datos.gov.co](http://www.datos.gov.co), es el sitio creado por el Ministerio TIC, en donde las entidades públicas del orden nacional y territorial deben publicar los datos abiertos de interés nacional en temas de su competencia. Este sitio cuenta con los formatos y procedimientos específicos para realizar la documentación estructuración, cargue y publicación de la información.

Algunas herramientas para facilitar la publicación de datos abiertos son las siguientes:

- **Automatización de carga de datos:** este proceso se realiza a través de procesos de extracción, transformación y carga denominados ETL, por sus siglas en inglés, que permitirán ejecutar automáticamente el proceso de carga de los datos abiertos desde los sistemas de información de las entidades públicas a [www.datos.gov.co](http://www.datos.gov.co) con la periodicidad deseada y previamente configurada. Estos procesos se recomiendan cuando se deben cargar grandes volúmenes de datos con una alta periodicidad (tiempo real, diaria, semanal) y que se requiera dejar sistematizado el proceso de actualización de datos.
- **Federación de portales de datos abiertos:** este proceso consiste en referenciar o enlazar conjuntos de datos de portales externos de datos abiertos, de tal manera que se puedan acceder desde [www.datos.gov.co](http://www.datos.gov.co), unificando el acceso a los datos e integrando las iniciativas locales de datos abiertos del sector público. Para llevar a cabo este proceso es necesario que el portal de datos externo cuente con el estándar de metadatos DCAT y se realice el proceso que se describe en el [manual de federación](#).
- **Carga manual de datos:** Este proceso se realiza a través de la herramienta de carga dispuesta en [www.datos.gov.co](http://www.datos.gov.co), el ofrece tres opciones de carga
  1. Carga tradicional desde un archivo en los formatos .csv, .tsv, .xls, .xlsx, .zip (shapefile), .json (GeoJSON), .geojson, .kml, .kmz, buscando el archivo fuente en su computadora.
  2. Importar conjunto de datos tabular o geoespacial desde una URL que está alojada externamente para tipos de archivo en formatos .csv, .tsv, .xls, .xlsx, .zip (shapefile), .json (GeoJSON), .geojson, .kml, .kmz
  3. Enlace a una fuente de datos externa: permite vincular a una fuente de datos externa para referenciar varios archivos a un conjunto de datos y vincular múltiples conjuntos de datos al origen de datos externo general.

Si su entidad reporta información a sistemas nacionales (Por ejemplo, SECOP, SIIF, SIMAT) no es necesario que publique esta información a título propio, dado que estas entidades consolidadoras están obligadas a publicar toda la información completa y consolidada del país.



## RESULTADO

Como resultado de esta fase se tendrán los datos abiertos estructurados bajo estándares de calidad, se habrá definido un esquema de publicación automático o manual de datos y se publicarán los datos abiertos de la entidad de acuerdo con lo establecido en el Plan de Apertura de datos.

### 4.3 Comunicar y promover

Los datos no tienen valor si no se usan. En tal sentido, los productores deben informar a sus audiencias sobre la publicación de los datos y establecer mecanismos que permitan identificar y difundir los productos, servicios, artículos periodísticos, investigaciones académicas y en general cualquier producto que use datos.

Ilustración 8. Comunicar y promover el uso de datos abiertos.



Fuente Elaboración Propia

Además de la labor de comunicación al exterior de la entidad, se debe realizar una labor informativa dirigida a los servidores públicos, 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

Vincular actores

Dar a conocer



## ¡RECORDAR!

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.

### 4.3.1 Dar a conocer

Esta actividad está orientada a desarrollar un conjunto de comunicaciones internas y externas de la entidad, para sensibilizar sobre el uso de datos abiertos y sus beneficios, como la mejora en temas de transparencia y control social, toma de decisiones, desarrollo de aplicaciones, mejora de servicios, periodismo de datos, entre otros. Estas estrategias de uso y divulgación pueden ser formuladas a partir del análisis de estadísticas sobre descargue, visitas y visualizaciones hechas a partir de la información publicada por cada entidad en [www.datos.gov.co/](http://www.datos.gov.co/).

Algunas actividades específicas para dar a conocer los usos y beneficios de los datos abiertos pueden ser:

- **Incorporar en las estrategias de comunicación** de la entidad la promoción de los conjuntos de datos que publica la organización. Esto se puede hacer a través de comunicados de prensa, mensajes internos, mensajes en redes sociales, campañas informativas, etc.
- **Elaborar contenidos (infografías, artículos, videos)** a partir de datos abiertos y promoverlos desde la entidad. También crear contenidos que sensibilicen a los reutilizadores y ciudadanos sobre la importancia de los datos abiertos.

Un ejemplo de este tipo de estrategias es el Programa de televisión “El Dato” desarrollado por el Ministerio TIC en convenio con Canal 13, el cual está dirigido a ciudadanos para enseñarles sobre el impacto de los datos abiertos ¿Qué son?, ¿Para qué sirven?, ¿Cómo se usan? y ¿Cómo están cambiando las vidas de los colombianos y de comunidades en todo el mundo?, este tipo de recursos están disponibles para ser reutilizados por las entidades en sus estrategias de comunicación: <https://canaltrece.com.co/programas/el-dato-datos-abiertos-open-data/>

### 4.3.2 Vincular diferentes actores

Esta actividad tiene como objetivo adelantar proyectos para vincular a diferentes actores del ecosistema de datos que promuevan la reutilización de la información disponible, difundir soluciones desarrolladas con datos abiertos e implementar iniciativas para identificar nuevas posibilidades con el uso de estos.

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

- Realizar **maratones de desarrollo** y otros ejercicios de uso de datos, que incentiven entre el ecosistema el aprovechamiento de esta información. Algunos ejemplos de este tipo de ejercicios son:
  - **DataJam de Seguridad Ciudadana 2019:** evento organizado por USAID Colombia en alianza con entidades públicas y privadas, cuyo objetivo fue generar soluciones innovadoras a un problema social, relacionado con seguridad y convivencia ciudadana basados en datos abiertos publicados en el [www.datos.gov.co](http://www.datos.gov.co) y fuentes de datos no estructuradas como redes sociales.
  - **Rally de Datos:** ejercicio que se lleva a cabo anualmente en el día internacional de los datos abiertos, organizado por entidades públicas y de sociedad civil que busca promover el control social y la participación ciudadana. En su versión 2018, se enfocaron en veeduría de las obras públicas en ejecución, tomando como punto de partida datos abiertos del Ministerio de Hacienda y Crédito Público.
- Promover proyectos que usen tecnologías emergentes para desarrollar productos o servicios a partir de datos abiertos. Como orientación en este tema puede hacer uso de los siguientes estudios realizados sobre uso de tecnologías emergentes dispuestos en la sección de herramientas de [www.datos.gov.co](http://www.datos.gov.co):
  - Soluciones para resolver necesidades en Gobierno Digital a través del uso de la Inteligencia Artificial.
  - Estudio sobre el uso de la analítica de datos en el Gobierno Digital.
  - Herramientas libres para la analítica de datos.

### 4.3.3 Consolidar y posicionar

Esta actividad busca que la entidad incorpore su plan de datos abiertos como parte integral de los 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 su ecosistema de interacción, en especial 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. Así mismo dependiendo de cada público es importante diferenciar las estrategias de divulgación y orientación de la información de interés de cada uno de forma que estos se acerquen, promocionen y promuevan la(s) iniciativa(s).

Se recomienda hacer uso de las herramientas disponibles en [www.datos.gov.co](http://www.datos.gov.co) para promocionar las soluciones creadas a partir de datos abiertos. Desde el Ministerio TIC se ha dispuesto en este portal un espacio denominado “Usos” en el que cualquier ciudadano puede



publicar un producto realizado a partir de datos abiertos Ver <https://herramientas.datos.gov.co/es/usos>



## RESULTADO

Como resultado de esta fase se tendrán las estrategias de comunicación, difusión y posicionamiento de los datos abiertos de la entidad, la vinculación de los actores que harán uso de los datos abiertos de la entidad y la divulgación de los usos de los datos abiertos publicados.

## 4.4 Monitorear la calidad y el uso

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.

### ¡RECORDAR!

Lo que no se mide no se puede mejorar, es importante generar mecanismos para reportar hallazgos y hacer comentarios o sugerencias sobre los conjuntos de datos; lo anterior permite retroalimentar el proceso de identificación, promoción y uso por parte de los usuarios.

Las entidades de acuerdo con el [Marco de Referencia de Arquitectura Empresarial](#) del Estado deben estructurar e implementar modelos de gobierno de datos, donde una de las aristas es la calidad de los mismos. Las entidades deben adelantar acciones y establecer controles, que les permitan garantizar la calidad de los datos desde la fuente o recolección de los datos, para mantener altos estándares de calidad en el ciclo de vida de estos.

Para monitorear la calidad y el uso de los datos abiertos, se deben tener en cuenta las siguientes acciones de seguimiento:

Medir calidad

Medir usos

Medir impacto

### 4.4.1 Medir calidad

Para monitorear la calidad de los datos abiertos publicados, es posible utilizar algunos indicadores de tipo cuantitativo que permiten determinar su calidad.

Algunos de estos pueden ser:



- **Número de solicitudes de apertura de datos:** identificar las PQRDS relacionadas con acceso a la información y analizar en la sección participa del Portal de Datos <https://datos.gov.co/nominate> las solicitudes ciudadanas sobre sus temas, disponible para descargar de solicitudes, en el siguiente enlace: <https://www.datos.gov.co/Participacion-ciudadana/Datos-Abiertos-Solicitados-Para-Los-Ciudadanos-En-L/9it7-3gvu>
- **Número de dependencias o áreas estratégicas de la entidad que abren datos:** es importante vincular a las dependencias de la entidad que contengan información estratégica y de alto valor e impacto al momento de la publicación.
- **Porcentaje de conjuntos de datos abiertos estratégicos publicados:** luego del plan de apertura se debe hacer seguimiento a las metas de apertura establecidas de datos estratégicos de la entidad.
- **Número de conjuntos de datos descargados vs. número de conjuntos de datos publicados:** para evaluar este indicador recomendamos realizar una vista filtrada del inventario de conjuntos de datos en [www.datos.gov.co](http://www.datos.gov.co) que le permita visualizar solo los conjuntos de datos de la entidad y validar las descargas por parte de los ciudadanos.
- **Número de conjuntos de datos visitados vs. descargados:** este indicador permitirá identificar aquellos conjuntos que son visitados con frecuencia pero que el ciudadano decide no descargar, para revisar si el conjunto tiene temas de calidad por mejorar o comentarios ciudadanos solicitando mejoras puntuales.

#### 4.4.2 Medir usos

Con el fin de mapear el impacto que está teniendo el uso y aprovechamiento de datos abiertos, se puso en marcha una [sección de usos](#) de datos abiertos, dispuesta en [www.datos.gov.co](http://www.datos.gov.co) que tiene como objetivo involucrar a quienes desarrollan productos (apps, páginas web, modelos predictivos, visualizaciones, periodismo de datos) para que divulguen su desarrollo<sup>11</sup>.

Este espacio tiene como objetivo:

- Visibilizar y dar mayor alcance a los productos a partir de datos.
- Recopilar buenas prácticas en uso y aprovechamiento de datos.
- Crear comunidad entre quienes usan y aprovechan los datos.
- Identificar cuáles son los sectores /temáticas y tipos de productos que más se realizan.

Para monitorear los usos de los datos abiertos, es posible utilizar algunos indicadores de tipo cuantitativo que permiten determinar los productos desarrollados a partir de sus datos abiertos tales como:

---

<sup>11</sup> Ver conjunto de datos de usos para descarga: <https://www.datos.gov.co/Ciencia-Tecnologia-e-Innovacion/Repositorio-Seccion-de-Usos-Datos-Abiertos/4t9r-7v7i/data>

- **Número de soluciones web o móviles generadas a partir de datos abiertos:** este indicador es uno de los más importantes porque permite identificar la usabilidad de los datos publicados y si se cumple con el impacto esperado al publicarlos, por lo cual, si el conjunto de datos no tiene asociados usos, ni descargas es un indicador importante que el dato no es útil y debe mejorarse o cambiarse.
- **Número de proyectos que hicieron uso de los conjuntos de datos abiertos:** entre estos pueden ser artículos de investigación y periodísticos, noticias, libros, visualizaciones, etc.
- **Número de proyectos de emprendimiento o nuevos negocios** generados a partir del uso de datos abiertos.

### 4.4.3 Medir impacto

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 Estado Abierto. En este sentido se recomienda:

- Realizar mediciones periódicas para todos los indicadores con el propósito de identificar tendencias sobre los datos de mayor impacto y tomar decisiones sobre acciones a seguir en el plan de apertura.
- Apoyar a las áreas estratégicas o entidades aliadas en el desarrollo de sus capacidades para recolectar, gestionar, analizar, medir y definir los datos más estratégicos a aperturar, estableciendo responsabilidades claras en la gobernanza de la información y evitando duplicidad de datos publicados en una misma entidad o sector.
- Adaptar la estrategia de comunicación dirigida a los actores del ecosistema que permita **mapear los distintos usos** que se le están dando a los datos de la entidad, y medir si se cumplió con el impacto esperado.
- Introducir de forma gradual más **parámetros de medición impulsados por ciudadanos**, que representen las experiencias reales del usuario para pasar de una perspectiva institucional a una ciudadana sobre los resultados e impactos de los datos abiertos.
- Adoptar un **componente de caso de éxito con datos** que permita apoyar una cultura de monitoreo y gestión del desempeño basado en datos. Dicho componente debe incluir un perfil de proyecto, alineación estratégica y normativa, análisis costo-beneficio, puesta en marcha de los servicios y evaluación y gestión de riesgos.
- Adelantar, financiar o participar en estudios que permitan **cuantificar el valor económico de los datos abiertos** en el sector al que pertenece la entidad.
- **Establecer alianzas** con entidades públicas y privadas, así como organizaciones de la sociedad civil, academia, ONG para construir proyectos de explotación de datos que den solución a una problemática pública de alto impacto.



## RESULTADO

Como resultado de esta fase se tendrán las estrategias de monitoreo y seguimiento de la calidad y uso de los datos abiertos de la entidad, en esta fase se habrán definido los indicadores que permitirán evidenciar el cumplimiento del plan de apertura trazado y el impacto en el uso de los datos abiertos de la entidad.



# Glosario

**Anonimizar:** Proceso técnico que hace que la información de identificación personal no sea visible, logrando que no sean identificables los individuos a los que pertenece la información. Anonimizar permite ocultar o eliminar los datos o características que identifique plena o parcialmente a las personas, organizaciones o características individuales de la fuente de información de manera irreversible, de tal forma que impidan identificar a una persona u organización.

**Conjunto de datos:** Una unidad mínima de información sujeta a carga, publicación, transformación y descarga en la plataforma.

**Datos enlazados (*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 URI, 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.

**Estándar:** Es un conjunto de características y requisitos que se toman como referencia o modelo y son de uso repetitivo y uniforme. Para que sea un estándar debe haber sido construido a través de consenso y refleja la experiencia y las mejores prácticas en un área en particular.

**Formatos abiertos:** Son formatos de archivo que se pueden crear y manipular por cualquier software, libre de restricciones legales.

**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.

**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, 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.

**Usabilidad:** Significa facilidad de uso, atributo de calidad, que identifica el grado en que un producto puede ser usado por determinados usuarios para lograr sus propósitos con eficacia, eficiencia y satisfacción en un contexto de uso específico.

# ANEXO I

## Índices Internacionales y marco normativo que fijan prioridades de publicación

El marco normativo nacional fija algunos conjuntos de datos que es obligatorio publicar para entidades públicas o entidades mixtas con funciones públicas:

[La ley 1712 de 2014](#) de transparencia y del derecho de acceso a la información Pública en su título II. de la publicidad y del contenido de la información.

*Artículo 11. Información mínima obligatoria respecto a servicios, procedimientos y funcionamiento del sujeto obligado. “Todo sujeto obligado deberá publicar la siguiente información mínima obligatoria de manera proactiva:*

*k) Los sujetos obligados deberán publicar datos abiertos, para lo cual deberán contemplar las excepciones establecidas en el título 3 de la presente ley. Adicionalmente, para las condiciones técnicas de su publicación, se deberán observar los requisitos que establezca el Gobierno Nacional a través del Ministerio de las Tecnologías de la Información y las Comunicaciones o quien haga sus veces.”*

[El Decreto 1081 de 2015](#): En su Capítulo 2 Publicación y divulgación de la información pública – transparencia activa.

*Artículo 38. Publicación del Registro de Activos de Información*

*Artículo 40. Publicación de Índice de Información Clasificada y Reservada*

Existen organizaciones internacionales que establecen los conjuntos de datos deseables a publicar como:

[Open Data Barometer \(ODB\)](#) : Este índice tiene como objetivo medir la prevalencia e impacto de las iniciativas de datos abiertos en 115 países en el mundo, analiza las tendencias mundiales y proporciona datos comparativos sobre países y regiones utilizando una metodología en profundidad que combina datos contextuales, evaluaciones técnicas e indicadores secundarios. Además de la madurez de las iniciativas de cada país, se evalúa el impacto político, social y económico que tienen los datos abiertos 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

En su cuarta versión en el año 2017, Colombia ocupó el puesto 25. Ver resultados de la evaluación en: [https://opendatabarometer.org/country-detail/?\\_year=2017&indicator=ODB&detail=COL](https://opendatabarometer.org/country-detail/?_year=2017&indicator=ODB&detail=COL). Ver Resultados del Reporte Global del ODB 2017 <https://opendatabarometer.org/leadersedition/report/?lang=es> - [executive-summary](#)

**Global Open Data Index(ODI)**: En este índice se evalúa el estado de los datos abiertos en 94 países alrededor del mundo, basándose en la disponibilidad y accesibilidad de datos de cada país en quince categorías que incluyen:

1. Presupuesto de Gobiernos
2. Estadísticas Nacionales
3. Contratación Pública
4. Límites administrativos
5. Proyecto de ley
6. Calidad del aire
7. Mapas nacionales
8. Pronóstico del tiempo
9. Registro de la compañía
10. Resultados de las elecciones
11. Ubicaciones
12. Calidad del agua
13. Gastos gubernamentales
14. Propiedad de la tierra

En su última versión en el año 2016, Colombia ocupó el puesto 14 entre 94 países. Ver resultados de la evaluación <https://index.okfn.org/place/co/>

**Our Data Index**: Esta encuesta que realiza la Organización para la Cooperación y el Desarrollo Económicos (OCDE), mide los esfuerzos realizados por los gobiernos para aumentar la accesibilidad y disponibilidad de los datos gubernamentales, y para fomentar la reutilización para entregar valor. Esto incluye evaluar las acciones tomadas para conocer el ecosistema de datos abiertos y comprometer a los diferentes grupos de usuarios para generar impacto (por ejemplo, valor de los beneficios sociales y económicos, así como una mejor gobernanza). En su última versión en el año 2017, Colombia ocupó el puesto 4. Ver resultados de la evaluación en <http://www.oecd.org/gov/digital-government/open-government-data.html>

Categoría	Conjuntos de datos
<i>Compañías</i>	Registro de compañías
<i>Crimen y justicia</i>	Estadísticas de crimen Seguridad



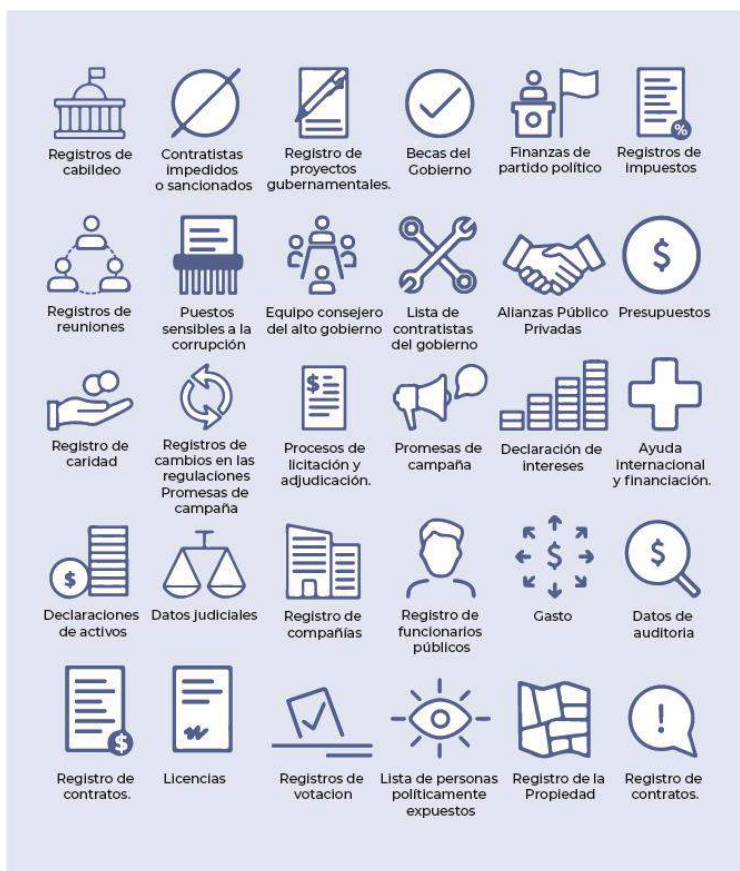
<b>Observación de la tierra</b>	Meteorología / clima Agricultura Bosques Pesca Fauna
<b>Educación</b>	Lista de Escuelas Desempeño de Escuelas Habilidades Digitales
<b>Energía y medio ambiente</b>	Niveles de Polución Consumo de energía
<b>Finanzas y contratos gubernamentales</b>	Gastos de Transacciones Contratos Licitaciones Futuras Presupuestos Locales Presupuesto Nacional Planeado Presupuesto Nacional Gastado
<b>Topografía geoespacial</b>	Códigos postales Mapas Nacionales Mapas Locales
<b>Desarrollo global</b>	Ayudas Seguridad Alimentaria Industrias Extractivas Tierra
<b>Responsabilidad del gobierno y democracia</b>	Puntos Contacto del Gobierno Resultados de Elecciones Legislación y estatutos Salarios Hospitalidad y Regalos
<b>Salud</b>	Datos de prescripción Datos de rendimiento
<b>Ciencia e investigación</b>	Datos del Genoma Investigación y actividad educativa. Resultado de Experimentos
<b>Estadísticas</b>	Estadísticas nacionales Censo Infraestructura Riqueza Habilidades Movilidad social y bienestar. Alojamiento Seguro de salud y prestaciones por desempleo. Sociedad del envejecimiento
<b>Transporte e Infraestructura</b>	Horarios de transporte Penetración de banda ancha Estadísticas de matriculación de vehículos de motor.
<b>Datos de presupuesto abierto</b>	Propiedad benéfica Reuniones de cabildeo Gestión de riesgos Datos de socorro en casos de desastre

[Programa Interamericano de Datos Abiertos contra la corrupción \(PIDA\)](#): Busca fortalecer las políticas de apertura de información, e incrementar la capacidad de los gobiernos y ciudadanos en la prevención y el combate a este flagelo.



Esta iniciativa establece unas líneas de acción clave entre las que se destaca la apertura de información estratégica que permita, a partir de uso, la identificación, prevención de riesgos de corrupción. A continuación, se encuentran los datos abiertos que se deberían tener disponibles en todos los países y/o ciudades son:

*Ilustración 9. Conjuntos de Datos para combatir la corrupción del PIDA.*



Fuente: PIDA. Elaboración Propia



**El futuro digital  
es de todos**

**Gobierno  
de Colombia  
MinTIC**