

Guía para el uso y aprovechamiento de Datos Abiertos en la Auditoría General de la República

Tabla de contenido

Introducción.....	3
1. Datos abiertos	5
1.1 Principios.....	6
2. Beneficios de los Datos Abiertos.....	7
2.1 Transparencia y Control Social	7
2.2 Mejoramiento o creación de productos, servicios y negocios innovadores.....	7
2.3 Pronosticar y prevenir fenómenos	7
2.4 Generar nuevo conocimiento.....	8
3. Marco de Gobernanza	8
3.1 Marco Normativo y Lineamientos.....	8
3.1.1 Marco Normativo.....	8
3.2 Ecosistema de Actores	9
3.3 Herramientas.....	10
3.4 Monitoreo.....	10
4. Ciclo de Datos Abiertos.....	10
4.1 Establecer el plan de apertura.....	11
4.1.1 Identificar.....	11
4.1.2 Analizar.....	14
4.1.3 Priorizar y Programar.....	15
4.2 Estructurar y publicar.....	17
4.2.1 Documentar	17
4.2.2 Estructurar.....	17
4.2.3 Publicar.....	18
4.3 Comunicar y promover	20
4.3.1 Dar a conocer	20
4.3.2 Vincular diferentes actores	20
4.3.3 Consolidar y posicionar.....	20
4.4 Monitorear la calidad y el uso	21
4.4.1 Medir calidad.....	21
4.4.2 Medir usos.....	22
4.4.3 Medir impacto.....	22
Glosario.....	23
ANEXO 1 PLAN DE APERTURA DE DATOS.....	25

Introducción

El papel que juega la ciudadanía en la construcción de Estado, es cada vez más visible. Esta visibilidad debe interpretarse como una señal de la efectividad de las políticas de participación ciudadana y de las estrategias para luchar contra la corrupción que se han venido adoptando en Colombia. El vehículo primordial para que ello suceda ha sido el acceso a la información.

Como lo señala Naser y otros (2020)¹ *“El gobierno abierto refiere a un cambio de paradigma de gestión pública que atraviesa a las instituciones de todos los poderes del Estado, en tanto que el mismo sostiene que la relación con la ciudadanía debe ser el centro del accionar de la toma de decisiones del quehacer público.”*

En palabras de Bárcena, Alicia (2020)² *“El gobierno abierto aboga por un aumento significativo de la transparencia, el acceso a la información y la rendición de cuentas, lo cual, a la luz de la pandemia del COVID-19, es crucial para garantizar que los recursos destinados a la respuesta y recuperación se usen de manera eficiente y eficaz.”*

El acceso a la información por parte de la ciudadanía, más allá del nivel de consulta, es fundamental para lograr una construcción colaborativa Estado – Ciudadanía, de los temas de interés de las entidades públicas.

Como lo dice el Ministerio de las Tecnologías de la Información y las Comunicaciones MinTIC, 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.”*

La Auditoría General de la República comparte las anteriores premisas, y como parte de su propósito misional, la participación ciudadana cobra un papel muy importante, en la lucha contra la corrupción y en la transformación del control fiscal. Por ello, ha considerado necesario adaptar el documento *Guía del uso y aprovechamiento de los datos abiertos* de MinTIC a las características institucionales, para que sirva de instrumento de orientación de primera mano.

¹ Naser, Fideleff y Tognoli (2020) Tomado de: <https://biblioguias.cepal.org/EstadoAbierto/concepto>

² Ibidem.

Esta guía se enmarca en la Política de Gobierno Digital y busca contribuir a la materialización del principio de “toma de decisiones basadas en datos a partir del aumento en el uso y aprovechamiento de la información”.

Este documento está basado en la Guía para el uso y aprovechamiento de Datos Abiertos en Colombia, elaborado por MinTIC y su objetivo es proporcionar 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 AGR. Se desarrolla en cuatro capítulos que muestran la línea entre disponer y usar datos abiertos.

Empieza por describir los datos abiertos en el contexto del marco legal colombiano y se presentan ejemplos sobre los datos abiertos que publica la AGR; 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.

Finalmente, se describen 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 la AGR: 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”.

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. A continuación, se señalan los conjuntos de los datos abiertos de la AGR, disponibles en el Portal de Datos del Estado Colombiano:

- Consolidado de los Hallazgos del Plan General de Auditorías.
- Consolidado de la cantidad de derechos de petición atendidos por las dependencias de la Auditoría General de la República.
- Consolidado de la cantidad de Contralorías Territoriales vigiladas por la Auditoría General de la República.
- Consolidado de la cantidad de los Sujetos y Puntos de Control.

Los datos abiertos tienen como principal característica que están publicados en la página web de la AGR, de manera gratuita, accesibles al público, dispuestos en formatos que permiten su uso, reutilización y aprovechamiento sin restricciones legales y bajo licencia abierta.

La licencia bajo la cual se usan los datos abiertos, que publica la Entidad en el portal datosabiertos.gov.co, está definida por dicho portal.

La licencia abierta Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) permite hacer uso, transformación, distribución, redistribución, reutilización, compilación, extracción, copia, difusión, modificación y/o adaptación de los datos abiertos, citando siempre la fuente.

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:

- **Abiertos por defecto.** La información debe estar disponible para consulta ciudadana desde el momento de su generación.
- **Comparables e interoperables.** Pueden ser comparados desde distintos sectores, territorios y periodos de tiempo.
- **Oportunos y completos.** Los datos abiertos son relevantes solo si aportan valor para el usuario de la información.
- **Mejorar la gobernanza y la participación ciudadana.** Abrir datos fortalece la gobernanza y favorece la transparencia.
- **Accesibles y utilizables.** Son gratuitos y bajo licencia abierta facilitando los datos a la mayor cantidad de usuarios y máquinas.
- **Apoyar el desarrollo, lo incluyente y la innovación.** El uso de datos abiertos permite la construcción de nuevo conocimiento que puede ser utilizado en la generación de valor social y económico.

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 productos generados por la AGR, 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 la AGR 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 interés relacionados con el control fiscal.

2.2 Mejoramiento o creación de productos, servicios y negocios innovadores

La publicación y uso de datos abiertos debe ayudar a la AGR a mejorar la respuesta 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 una respuesta y proponer mejoras para optimizarlo.
- Saber qué temas 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.

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 la prevención de la corrupción, pues a partir de la información existente es posible crear o mejorar las políticas públicas para reducir el impacto negativo que presenta este fenómeno.

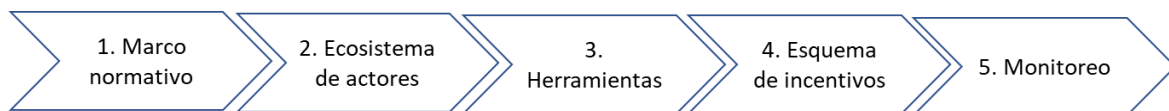
Los datos abiertos permiten a los distintos actores del ecosistema (gobernantes, ciudadanos, periodistas) mapear el estado de un sector o tema y mejorar, por ejemplo, la cobertura del control fiscal, los procesos auditados, o los temas de mayor inquietud para la ciudadanía, entre otros.

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.

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:



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 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.

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:

- Generación de datos digitales: de calidad y con completitud.
- Cultura de datos: para propiciar que haya demanda de datos y conciencia en la generación de valor.
- Capital humano para la explotación de datos: personas capaces de producir, procesar y reutilizar datos para resolver problemáticas.
- 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.

3.1.2 Lineamientos

En desarrollo del marco normativo, los lineamientos son aquellas directrices o disposiciones que deben ser implementadas por la AGR para el desarrollo de la Política de Gobierno Digital y en consecuencia la apertura y aprovechamiento de datos abiertos.

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.

En la AGR, la Oficina de Planeación, en conjunto con el Grupo de Sistemas es la dependencia encargada de liderar la política de Gobierno Digital y de transmitir los lineamientos de la política de Datos Abiertos y trabaja articuladamente con la Auditoría Auxiliar en función de la política de transparencia, acceso a la información y lucha contra la corrupción, al igual que con la Secretaría General como líder de la política de gestión documental y con la Auditoría Delegada para la Vigilancia de la Gestión Fiscal, cuyos procesos derivan en conjuntos de datos que son publicados periódicamente.

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.

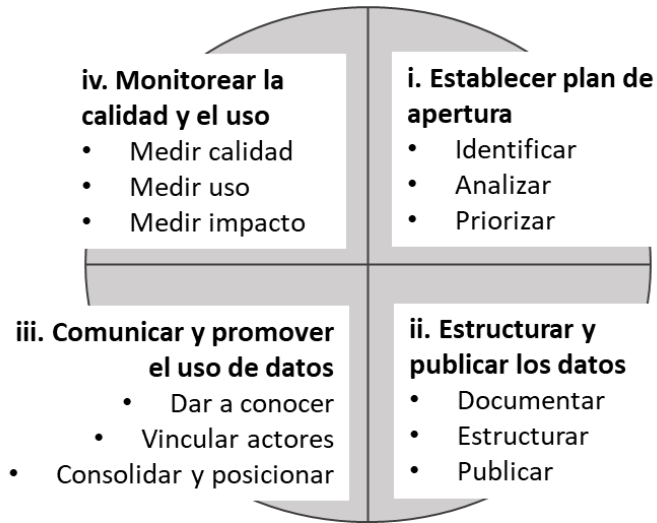
3.4 Monitoreo

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.

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.



4.1 Establecer el plan de apertura

A partir del Marco de Gobernanza de datos abiertos y de acuerdo con el marco normativo vigente la AGR adelantó una estrategia para el proceso de apertura, identificación de los actores que hacen parte del ecosistema, usar las herramientas disponibles para apertura, calidad y uso de sus datos abiertos y generar estrategias de monitoreo a los datos abiertos.

4.1.1 Identificar.

En este paso se identificó 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 AGR en cada una de las fases del ciclo de datos, desde el establecimiento del plan hasta el monitoreo y calidad de uso, se definieron e identificaron los actores con los que la AGR tiene contacto y que son considerados como de interés por el rol que cumplen en el ecosistema (usuario, regulador, proveedor, etc.)

Como ejemplo a continuación se encuentra un listado parcial de tipos de actores identificados del ecosistema en el que se desenvuelve la AGR:

- Consejo Nacional de Contralores

- SINACOF
- Contraloría General de la República
- OLACEFS
- Contralorías Territoriales
- Veedurías ciudadanas
- Academia / Universidades

A modo de recomendación la identificación de estos actores se facilita al identificar los roles que se llevan a cabo en el ecosistema. Finalmente, para cada tipo de actor ya pueden identificarse a los actores asociados a cada uno.

- Contralorías Territoriales – Proveedor – Usuario
- Consejo Nacional de Contralores – Usuario
- SINACOF – Usuario
- Contraloría General de la República – Proveedor – Usuario
- OLACEFS – Usuario
- Veedurías ciudadanas – Usuario
- Academia / Universidades – Usuario
- Ciudadanía interesada en el control fiscal – Proveedor – Usuario

4.1.1.2 Identificar los activos de información

Con la identificación de los activos de información se realizó un levantamiento de la información pública que administra la AGR y que se convierte en datos abiertos. Para efectos de realizar la identificación de esta información y de conformidad con lo establecido en la normatividad vigente, se tuvieron en cuenta durante esta fase los siguientes datos:

- Los datos de los registros de activos de información oficiales generados por la AGR clasificados como dato abierto.
- Los datos públicos generados o administrados por los Sistemas de Información (SIA) de la AGR.
- Las solicitudes de acceso a la información.

En la AGR se han identificado como activos de información:

- Consolidado de los Hallazgos del Plan General de Auditorías.
- Consolidado de la cantidad de derechos de petición atendidos por las dependencias de la Auditoría General de la República.
- Consolidado de la cantidad de Contralorías Territoriales vigiladas por la AGR.
- Consolidado de la cantidad de los Sujetos y Puntos de Control.

Consulte: https://www.auditoria.gov.co/atencion_al_ciudadano/datos-abiertos

Para la AGR 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.

Una vez identificados los datos abiertos de la AGR, se definieron las fuentes de información que generan dichos conjuntos de datos y se determinó 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).

En la AGR, la fuente de información para cada conjunto de datos es manual.

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 surtirse sobre los activos de información que tiene identificados la AGR (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: Auditoría Delegada para la Vigilancia de la Gestión Fiscal.
- Identificar si la información se genera de forma manual y/o automática: Manual.
- Identificar la periodicidad de la actualización de esta información: Anual los datos relacionados con el Proceso Auditor y para el informe de PQRS.
- Identificar si el activo de información puede ser considerado como un conjunto de datos. En la AGR, todos son conjuntos 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. (Ver anexo 1)

#	Identificado por	Activo de información	Responsable	¿Es un conjunto de datos?	Desarrollo automático/manual	Periodicidad	¿Está en datos abiertos?	Vínculo	¿Puede automatizarse?	Sujeto a federación	Puede ser sometido a calidad de datos

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.

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 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:

- a) **Abrir:** disponer el conjunto en Datos Abiertos www.datos.gov.co
- b) **Automatizar:** procesos para el cargue automatizado de la información
- c) **Actualizar:** actualización de los registros de la información conforme a la periodicidad del conjunto de datos.
- d) **Mejorar:** acciones de mejora y calidad de datos.
- e) **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 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.
- 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, que en el caso de la AGR es la Auditoría Delegada para la Vigilancia de la Gestión Fiscal.

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 haber desarrollado la 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 ha diligenciado en su totalidad su plan de apertura, mejora y uso de datos.

En el anexo 1 se podrá consultar el plan de apertura de datos propuesto para la AGR.

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.

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.

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.

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.

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 de la guía de estándares de Calidad e Interoperabilidad de los Datos Abiertos del Gobierno de Colombia.

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 una 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, 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 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, 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, 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.

Dado que la AGR 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.

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.

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:

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 descargas, visitas y visualizaciones hechas a partir de la información publicada por cada entidad en www.datos.gov.co/.

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.

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).

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.

La AGR de acuerdo con el Marco de Referencia de Arquitectura Empresarial del Estado, debe estructurar e implementar modelos de gobierno de datos, donde una de las aristas es la calidad de los mismos. La AGR debe adelantar acciones y establecer controles, que 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:

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/nominat> 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-Por-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 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 que tiene como objetivo involucrar a quienes desarrollan productos (apps, páginas web, modelos predictivos, visualizaciones, periodismo de datos) para que divulguen su desarrollo.

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.

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 abrir, 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.

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 sobre el 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 empleado por determinados usuarios para lograr sus propósitos con eficacia, eficiencia y satisfacción en un contexto de uso específico.

ANEXO 1 PLAN DE APERTURA DE DATOS

1. Identificar el ecosistema de interacción

ACTORES - ROL

Contralorías Territoriales – Usuario – Proveedor

Contraloría General de la República – Usuario – Regulador – Proveedor

Veedurías Ciudadanas – Usuario – Proveedor

OLACEFS – Usuario

Academia / Universidades – Usuario

Ciudadanía interesada en control fiscal – Proveedor – Usuario

Identificar los activos de información

¿QUÉ SE PUEDE PUBLICAR?

Consolidado de los Hallazgos del Plan General de Auditorías.

Consolidado de la cantidad de derechos de petición atendidos por las dependencias de la Auditoría General de la República.

Consolidado de la cantidad de Contralorías Territoriales vigiladas por la AGR.

Consolidado de la cantidad de los Sujetos y Puntos de Control.

Caracterización de los activos de información identificados

Responsable del desarrollo y actualización de la información	Generación de la información (manual y/o automática)	Periodicidad de la actualización de esta información	¿Puede ser considerado como un conjunto de datos? (SI/NO)	¿Es Datos Abiertos?	¿Podría ser sometido a un proceso de automatización en su actualización? (SI/NO)	¿Está sujeto a un proceso de federación con otras fuentes de información? (SI/NO)	¿Podría ser sometido a un proceso de calidad de datos y mejora de información? (SI/NO)
Auditoría Delegada para la Vigilancia de la Gestión Fiscal	Manual	Anual	SI	SI	SI (*)	NO	NO
Auditoría Delegada para la Vigilancia de la Gestión Fiscal	Manual/Automática	Anual	SI	SI	SI (*)	NO	NO
Auditoría Delegada para la Vigilancia de la Gestión Fiscal	Manual	Anual	SI	SI	SI (*)	NO	NO
Auditoría Delegada para la Vigilancia de la Gestión Fiscal	Manual	Anual	SI	SI	SI (*)	NO	NO

(*) Esta respuesta está sujeta a los efectos del proyecto de automatización e inteligencia de negocios que se adelanta.

Analizar (Acciones a adelantar)

Abrir: disponer el conjunto en Datos Abiertos www.datos.gov.co	Automatizar: procesos para 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 de calidad de datos.	Usar: identificación de usos de datos para generar valor agregado a la información disponible.
X		X		
X		X		
X		X		
X		X		

Programar (Responsables y fechas de acciones)

Responsable	Fecha de ejecución de acciones
Pedro Pablo Cruz Rojas – Profesional Universitario 01	15/01/2022
Pedro Pablo Cruz Rojas – Profesional Universitario 01	15/01/2022
Pedro Pablo Cruz Rojas – Profesional Universitario 01	15/01/2022
Pedro Pablo Cruz Rojas – Profesional Universitario 01	15/01/2022