

GUÍA

Guía para apertura de datos en Colpensiones

Proceso: Investigación Desarrollo y Analítica

1. OBJETIVO

Definir el plan de apertura de datos abiertos de Colpensiones, que es aplicable por los servidores públicos, trabajadores en misión, academia, sociedad civil, desarrolladores de aplicaciones, gremios de periodistas, entidades públicas, etc., las cuales hacen parte del ecosistema de actores de datos abiertos con el fin de suministrar y publicar de manera transparente y eficiente, en los diferentes medios definidos, los datos generados desde la entidad hacia las partes interesadas para su utilización y transformación de manera que faciliten el desarrollo de servicios de valor agregado por parte de terceros.

2. ALCANCE

Inicia con la identificación de los datos susceptibles a ser publicados, continúa con el análisis de información aplicable, posteriormente la priorización de información, estructurar y publicar, comunicar y promover, y termina con el monitoreo de la calidad y uso.

3. DEFINICIONES

- 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 Abiertos: Son 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.
- 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.
- **Open Data (Datos Abiertos):** "Filosofía y práctica que persigue que determinados datos de los Gobiernos estén disponibles de forma libre a todo el mundo, sin restricciones de copyright, patentes u otros mecanismos de control."
- **Oferta de datos abiertos:** La oferta de datos es el elemento principal de la operación por medio del cual se publica la información de las entidades públicas y las privadas con funciones públicas, del orden nacional y territorial, en formatos estructurados y reutilizables bajo los parámetros de calidad definidos más adelante.
- **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.

4. DESCRIPCIÓN DEL DOCUMENTO

- **Plan de Apertura de datos:**

El plan de apertura de datos abiertos incluye las siguientes actividades:

- **Identificar:**

Busca realizar un levantamiento de la información pública que administre Colpensiones 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 tuvieron en cuenta los siguientes datos:

MACROPROCESO / PROCESO: Direccionamiento Estratégico/Investigación, Desarrollo y Analítica	CÓDIGO: DIE-IDA-GUI-001	VERSIÓN: 02	PÁGINA 2 de 8
---	----------------------------	----------------	------------------

- **Número de pensionados en Colpensiones desagregado por departamento y municipio:** Esta información contiene la cantidad acumulada de pensionados en Colpensiones desagregado por Departamento y Municipio.
- **Número de afiliados a Colpensiones desagregado por departamento y municipio:** Esta información contiene la cantidad acumulada de afiliados en Colpensiones desagregado por Departamento y Municipio.
- **Cantidad de pensionados de Colpensiones por tipo de pensión:** Cantidad de personas pensionadas en Colpensiones por tipo de pensión
- **Esquema de publicación de información Colpensiones:** El Esquema de Publicación de Información es el instrumento del que disponen los sujetos obligados para informar, de forma ordenada, a la ciudadanía, interesados y usuarios, sobre la información publicada y que publicará, conforme al principio de divulgación proactiva de la información previsto en el artículo 3° de la Ley 1712 de 2014, y sobre los medios a través de los cuales se puede acceder a la misma.
- **Número de pensionados de Colpensiones por rango salarial:** Número de Pensionados por Colpensiones clasificados por rango salarial en escala del salario mínimo del año legal vigente.
- **Cantidad de pensionados de Colpensiones por rango de edad:** Cantidad de pensionados de Colpensiones clasificados por rango de edad, género, valor de la pensión en cantidad de salarios mínimos legales vigentes para el año 2019 (corte diciembre) y por tipo de pensión.
- **Índice de Información Clasificada y Reservada Colpensiones:** Es el inventario de la información pública generada, obtenida, adquirida o controlada por Colpensiones, que ha sido calificada como clasificada o reservada.
- **Registro de Activos de Información Colpensiones:** Contiene el inventario de la información pública que la Administradora Colombiana de Pensiones Colpensiones, genera, adquiere, transforma y controla.

Una vez se ha identificado los datos abiertos de la entidad, se establecen las fuentes de información que generan dichos conjuntos de datos y así determinar si la generación de estos se realiza de manera manual o automática.

- a. **Analizar:** En esta etapa se examina 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.

MACROPROCESO / PROCESO: Direccionamiento Estratégico/Investigación, Desarrollo y Analítica	CÓDIGO: DIE-IDA-GUI-001	VERSIÓN: 02	PÁGINA 3 de 8
---	----------------------------	----------------	------------------

El siguiente paso consiste en anonimizar los datos el cual 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, es necesario consultar con la gerencia de prevención del fraude. Esto con el fin de determinar 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.

- b. Priorizar:** 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. Por lo anterior, se recomienda tener en cuenta los siguientes criterios a la hora de realizar la priorización de datos abiertos que serán incluidos en el Plan de Apertura:

➤ Criterios de impacto:

- Información que contribuye al crecimiento económico
- Información que puede generar negocio de inmediato
- Área de impacto
- Demanda de los datos

➤ CRITERIOS DE DIFICULTAD

- Esfuerzo requerido para publicar
- Fuente de datos
- Calidad de la información

Conjuntos de Datos	Información que contribuye al crecimiento económico	Información que puede generar negocio de inmediato	Área de impacto	Demanda de los datos	Esfuerzo requerido para publicar	Fuente de datos	Calidad de la información
Número de pensionados de Colpensiones por rango salarial	Contribuye al Sector	Genera valor agregado	Regional / Nacional / Todos los sectores y Coberturas	Demanda por parte de la ciudadanía a través de la solicitud de peticiones	No requiere esfuerzo de desarrollo	Servidor de datos	Certificada
Número de pensionados en Colpensiones desagregado por departamento y municipio	Contribuye al Sector	Genera valor agregado	Regional / Nacional / Todos los sectores y Coberturas	Demanda por parte de la ciudadanía a través de la solicitud de peticiones	No requiere esfuerzo de desarrollo	Servidor de datos	Certificada
Cantidad de pensionados de Colpensiones por tipo de pensión	Contribuye al Sector	Genera valor agregado	Regional / Nacional / Todos los sectores y Coberturas	Demanda por parte de la ciudadanía a través de la solicitud de peticiones	No requiere esfuerzo de desarrollo	Servidor de datos	Certificada
Cantidad de pensionados de Colpensiones por rango de edad	Contribuye al Sector	Genera valor agregado	Regional / Nacional / Todos los sectores y Coberturas	Demanda por parte de la ciudadanía a través de la solicitud de peticiones	No requiere esfuerzo de desarrollo	Servidor de datos	Certificada
Esquema de publicación de información Colpensiones	Contribuye al Sector	Genera valor agregado	Regional / Nacional / Todos los sectores y Coberturas	Demanda por parte de la ciudadanía a través de la solicitud de peticiones	No requiere esfuerzo de desarrollo	Servidor de datos	Certificada
Índice de Información Clasificada y Reservada Colpensiones	Contribuye al Sector	Genera valor agregado	Regional / Nacional / Todos los sectores y Coberturas	Demanda por parte de la ciudadanía a través de la solicitud de peticiones	No requiere esfuerzo de desarrollo	Servidor de datos	Certificada
Registro de Activos de Información Colpensiones	Contribuye al Sector	Genera valor agregado	Regional / Nacional / Todos los sectores y Coberturas	Demanda por parte de la ciudadanía a través de la solicitud de peticiones	No requiere esfuerzo de desarrollo	Servidor de datos	Certificada
Número de Afiliados en Colpensiones desagregado por departamento y municipio	Contribuye al Sector	Genera valor agregado	Regional / Nacional / Todos los sectores y Coberturas	Demanda por parte de la ciudadanía a través de la solicitud de peticiones	No requiere esfuerzo de desarrollo	Servidor de datos	Certificada

c. Estructurar y publicar:

Este esquema estará compuesto por:

a. Documentar

La documentación de los datos consiste en definir los metadatos que tendrá cada uno de los conjuntos de los datos a publicar.

b. 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. 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

Actualmente en Colpensiones, se publica la información en formato CSV para mayor facilidad en la consulta pública de información.

c. Publicar

En la actualidad Colpensiones lleva a cabo el proceso de publicación a través de la carga manual de datos, proceso que se realiza a través de la herramienta de carga dispuesta en www.datos.gov.co. Esto se realiza por el módulo de carga tradicional desde un archivo en los formatos .csv



d. Comunicar y promover

Una vez se cuenta con la publicación de datos abiertos, se realizan los siguientes pasos:

- Facilitar la apertura de datos
- Facilitar la reutilización y consumo de datos de apertura
- Sensibilizar sobre el uso de los datos abiertos y sus logros y beneficios

e. Monitorear la Calidad y Uso:

En Colpensiones se garantiza que los datos abiertos publicados sean completos, oportunos y consistentes, sea posible promover su uso y, por lo tanto, generar valor a partir de ello. En ese sentido, para monitorear la calidad y el uso de los datos abiertos, se deben tener en cuenta las siguientes acciones de seguimiento:

- Medir calidad: Monitorear la calidad de los datos abiertos publicados.
- Medir usos: a través de las publicaciones en el portal de www.datos.gov.co
- Medir impacto: 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.

f. Actualización de Datos Abiertos:

A continuación, se describen los responsables de la publicación y los periodos de actualización de los datos abiertos que Colpensiones pone a disposición de los ciudadanos y personas interesadas, de la siguiente forma:

Conjuntos de Datos	Área responsable de publicación	Periodicidad
Número de pensionados de Colpensiones por rango salarial	Gerencia de Planeación Institucional	Trimestral
Número de pensionados en Colpensiones desagregado por departamento y municipio	Gerencia de Planeación Institucional	Trimestral
Cantidad de pensionados de Colpensiones por tipo de pensión	Gerencia de Planeación Institucional	Trimestral
Cantidad de pensionados de Colpensiones por rango de edad	Gerencia de Planeación Institucional	Trimestral
Esquema de publicación de información Colpensiones	Gerencia de Planeación Institucional	Anual
Índice de Información Clasificada y Reservada Colpensiones	Gerencia de Planeación Institucional	Anual
Registro de Activos de Información Colpensiones	Gerencia de Planeación Institucional	Anual
Número de Afiliados en Colpensiones desagregado por departamento y municipio	Gerencia de Planeación Institucional	Trimestral

g. Promoción de datos:

La promoción de datos se realizará a través de la página web.

Actividad	Canal Comunicación	Responsable	Periodicidad	Objetivo
Publicar en la página web de Colpensiones una información que promocióne y direcciona al conjunto de datos	Sitio web de la entidad	Gerencia de Planeación Institucional	Anual	Motivar la consulta y uso del conjunto de datos abiertos

CONTROL DE CAMBIOS DEL DOCUMENTO

FECHA	VERSIÓN	MODIFICACIÓN	ELABORÓ	REVISÓ	APROBÓ
23/12/2019	V1	Se Incluye: Objetivo, alcance, del plan de apertura de datos. Definiciones Se modifica el plan de apertura de datos en sus etapas de identificación, análisis, priorización, estructuración y publicación, comunicación y promoción, monitoreo y uso.	Nombre: Henry Rodríguez Hurtado Cargo: Profesional Senior	Nombre: Arelix del Pilar Ibáñez Cargo: director (A) Prospectiva y Estudios	Nombre: Arelix del Pilar Ibáñez Cargo: director (A) Prospectiva y Estudios
10/03/2020	V2	Se Incluye: El periodo de publicación y se precisa el cambio del nombre de la guía.	Nombre: Henry Rodríguez Hurtado Cargo: Profesional Senior	Nombre: Julio Jiménez Cargo: Director Prospectiva y Estudios	Nombre: Julio Jiménez Cargo: Director Prospectiva y Estudios