



# GUÍA PARA LA APERTURA DE DATOS EN COLOMBIA

# GUÍA PARA LA APERTURA DE DATOS EN COLOMBIA

Ministerio de Tecnologías de la Información y las Comunicaciones  
Programa de Gobierno en línea  
Coordinación de Investigación, Políticas y Evaluación

Derechos de Autor  
Ministerio de Tecnologías de la Información y las Comunicaciones  
Programa Gobierno en línea



Ministerio de Tecnologías de la  
Información y las Comunicaciones  
República de Colombia



## TABLA DE CONTENIDO

1.	INTRODUCCIÓN A LA GUÍA .....	4
2.	¿CÓMO USAR ESTA GUÍA? .....	8
	Preparación para la apertura de datos .....	8
	¿Cómo puedo identificar mi equipo de trabajo? .....	9
	¿Cómo elaboro mi plan de trabajo? .....	12
	¿Qué debo tener en cuenta en mis reuniones individuales con los grupos de trabajo? .....	13
3.	¿Cómo realizo la identificación de la información? .....	14
	¿Qué es la identificación de la información? .....	14
	¿Qué tengo que hacer para realizar la identificación de la información? .....	15
	¿Para qué me sirve realizar el inventario de información de mi entidad?.....	18
4.	¿Cómo realizo el análisis de la información? .....	19
	¿Cómo ejecuto la revisión de la información bajo los criterios jurídicos? .....	20
	¿Cómo identifico la información publicable? .....	23
	¿Para qué me sirve realizar el análisis de la información?.....	24
	¿Cómo realizo la publicación del inventario de información en el sitio web de la entidad? .....	24
5.	¿Cómo realizo la priorización de los datos? .....	28
	¿Qué se requiere para la priorización de datos?.....	28
	¿Cómo diligencio la matriz para la priorización de información a publicar? .....	30
	Instrumento de priorización de información .....	32
	¿Para qué me sirve realizar una priorización de mis datos?.....	34
6.	¿Cómo realizo la documentación de los datos? .....	35
	¿Cómo se realiza la identificación de estándares/formatos propuestos de publicación para la Descripción de Metadatos? .....	35
	¿Cuál es la Estructuración de los formatos de la Ficha de Descripción de Metadatos?.....	40
	Instrumento de documentación de los conjuntos de datos .....	41

¿Para qué me sirve realizar una documentación de la ficha de los conjuntos de datos?.....	42
7. ¿Cómo realizo la estructuración, cargue y publicación de los conjuntos de datos?.....	43
¿Qué debo tener en cuenta antes de la publicación? .....	43
Estructuración de los datos a cargar.....	48
¿Cuáles son los mecanismos de publicación de los conjuntos de datos?.....	50
9. REFERENCIAS.....	55
10. ANEXOS .....	58
Para profundizar.....	58
Siglas y acrónimos .....	64

## 1. INTRODUCCIÓN A LA GUÍA

Esta guía está dirigida a los servidores públicos, interesados en adquirir conocimientos y herramientas prácticas que permitan a las entidades del Estado implementar procesos de apertura de datos públicos y de esta manera facilitar el desarrollo de servicios de valor agregado por parte de terceros. Se ha estructurado en torno a una serie de etapas que permiten desarrollar un proceso progresivo y guiado hasta conseguir la publicación de datos en formato abierto.

Con esta guía, se espera contribuir al entendimiento, la apropiación y la ejecución de acciones concretas que permitan a las entidades hacer un mejor aprovechamiento de la información pública y, a su vez, facilitar la colaboración de terceros que apoyen el desarrollo y provisión de servicios para el cumplimiento de los objetivos de cada entidades y del Estado en su conjunto.

### **Definición de Datos Abiertos**

Datos Abiertos son todos aquellos datos primarios, sin procesar, en formatos estándar, estructurados e interoperables que facilitan su acceso y permiten su reutilización, los cuales están bajo la custodia de las entidades públicas y que pueden ser obtenidos y ofrecidos sin reserva alguna, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos.

### **Características de los Datos Abiertos**

Uno de los máximos exponentes de este movimiento es Tim Berners Lee, reconocido como el creador de la Web, quien bajo el lema “raw data now!” lidera la exigencia ciudadana y empresarial según la cual se debe hacer un esfuerzo especial y urgente por liberar los datos y vincularlos en una nueva Web de los Datos.

En Octubre de 2007, treinta defensores del movimiento datos abiertos, se reunieron en Sebastopol, California, y definieron los ocho principios que deben regular la apertura de los datos de un Gobierno. El objetivo de dicho trabajo consistía en desarrollar una comprensión robusta de por qué los datos de un Gobierno son esenciales para la democracia. Según este grupo de trabajo, “con el hecho de adoptar los ocho principios de datos abiertos, los Gobiernos del mundo pueden transformarse en Gobiernos más efectivos, transparentes y relevantes para nuestras vidas.”<sup>1</sup>. De acuerdo con lo anterior,

---

<sup>1</sup> Op.cit. “The 8 principles of Open Government Data”: <http://www.opengovdata.org/home/8principles>, (consultado el 12 de diciembre de 2011).

los datos del Gobierno son considerados abiertos si son publicados siguiendo los siguientes principios o características:

- **Completos:** Todos los datos públicos deben estar disponibles. Los datos públicos no contemplan datos privados ni limitaciones de seguridad o privilegios.
- **Primarios:** Los datos deben ser recolectados en la fuente de origen, con el nivel de granularidad más alto posible, no en forma agregada ni modificada.
- **Oportunos:** Los datos deben estar disponibles tan rápido como sea necesario para garantizar el valor de los mismos.
- **Accesibles:** Los datos deben estar disponibles para el rango más amplio de usuarios y para el rango más amplio de propósitos.
- **Procesables por máquinas:** Los datos deben estar estructurados razonablemente para permitir un procesamiento automático.
- **No discriminatorios:** Los datos deben estar disponibles para cualquiera persona, sin requerir un registro.
- **No propietarios:** Los datos deben estar disponibles en un formato sobre el cual ninguna entidad tiene un control exclusivo.
- **Libres de licencias:** Los datos no deben estar sujetos a ningún derecho de autor, patente, marca registrada o regulaciones de acuerdo de secreto. Se podrán permitir restricciones razonables de privacidad, seguridad o privilegios.

Más recientemente, a mediados de 2010, Berners-Lee estableció una clasificación de cinco categorías para puntuar la calidad de los datos abiertos<sup>2</sup>. Esta categorización se ha convertido en un principio técnico a partir del cual diseñar una primera ruta de trabajo en la construcción progresiva de plataformas de datos abiertos. La Figura 1 resume los mencionados 5 niveles de estrellas.

**Figura 1: Modelo de las 5 Estrellas**

- ★ **Publicar** los datos **en la web**, independientemente de su formato, bajo licencia abierta
- ★★ **Publicar datos estructurados** y en formatos procesables (Excel en lugar de imágenes, p.e.)
- ★★★ **Ibidem anterior + datos en formato no propietario** (CSV en lugar de Excel, p.e.)
- ★★★★ **Ibidem anteriores + usar estándares abiertos** de W3C (RDF y SPARQL)
- ★★★★★ **Ibidem anteriores + enlaza los datos** con otras fuentes de datos para proveer contexto (LOD)

Fuente: <http://inkdroid.org/journal/2010/06/04/the-5-stars-of-open-linked-data/>

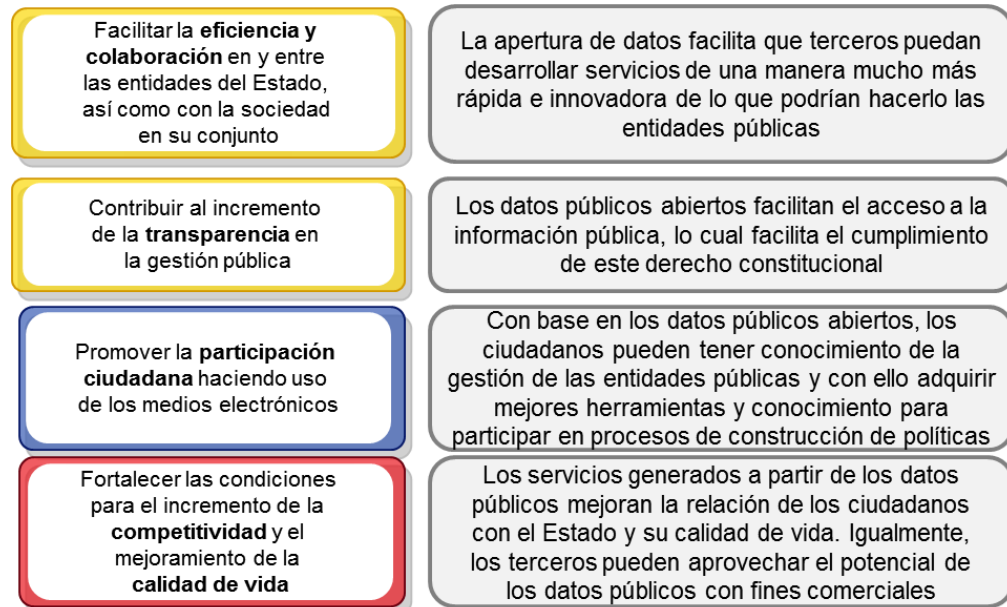
<sup>2</sup> Ver exposición y video en <http://inkdroid.org/journal/2010/06/04/the-5-stars-of-open-linked-data/> (consultado el 21 de septiembre de 2011).

## Datos Abiertos y la Estrategia de Gobierno en línea

La apertura de datos busca un mayor acceso a la información pública para diversos propósitos comerciales y no comerciales que facilitan la transparencia, la participación y la colaboración. Igualmente, busca la generación de servicios de valor agregado a la sociedad en general, a través del desarrollo de aplicaciones y servicios realizados por las comunidades de desarrollo, la industria de tecnologías de información, la academia o cualquier ciudadano.

De esta forma, mediante la apertura de datos se apoya el cumplimiento de los objetivos de la Estrategia de Gobierno en línea en Colombia así:

**Figura 2: Relación entre los objetivos de la Estrategia de Gobierno en línea y los Datos Abiertos**



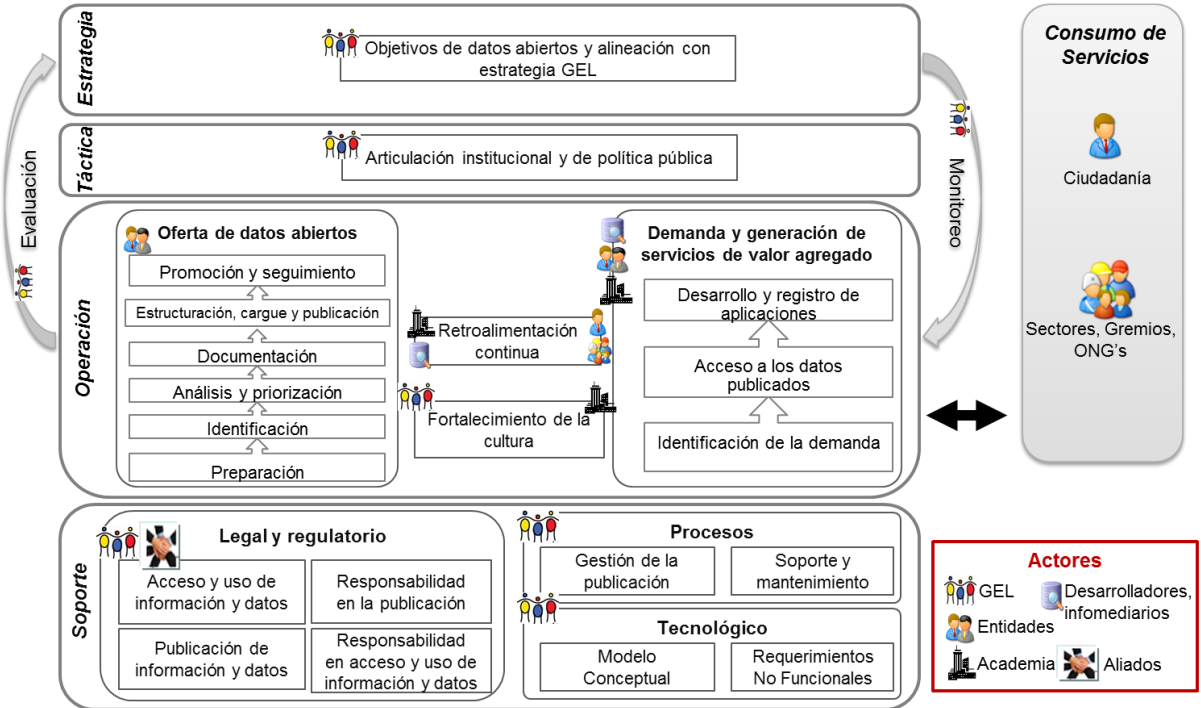
## Modelo de Madurez de Datos Abiertos

El modelo de datos abiertos en Colombia, busca cumplir los siguientes objetivos generales:

- Generar progreso económico en el país por medio de la información suministrada de forma masiva, permanente y fiable, a ciudadanos y empresas, para que sea utilizada o transformada para la generación de servicios de gobierno a la sociedad en general.

- Fomentar la transparencia del Estado por medio de la información y los datos puestos al alcance del ciudadano y las empresas.

**Figura 3: Modelo de datos abiertos**



En el marco del modelo anterior, esta guía contribuye al desarrollo e implementación del componente relacionado con la oferta de datos abiertos por parte de las entidades públicas.

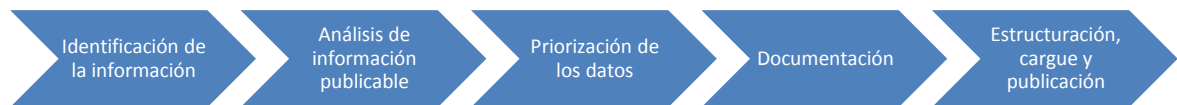


## 2. ¿CÓMO USAR ESTA GUÍA?

La presente guía presenta un proceso ordenado de apertura de datos el cual puede ser desarrollado por cualquier entidad. Este proceso incluye a su vez una cadena de valor, cuyos componentes se desarrollarán a lo largo del documento, con miras a facilitar un mejor entendimiento de la guía.

El proceso de apertura de datos de las entidades contempla las siguientes etapas:

Figura 4 - Metodología de apertura de datos



A continuación se describen los pasos de la metodología para la publicación de datos abiertos:

- **Identificación de la información:** En esta fase se busca identificar y caracterizar la información con que cuentan las entidades.
- **Análisis de información publicable:** En esta fase se clasifica la información que es publicable y la que no lo es, de acuerdo con el marco Jurídico establecido.
- **Priorización de los datos:** En esta fase se definen aquellos datos con los cuales se implementará el proceso de apertura, de acuerdo con un análisis estratégico. Al finalizar se realiza la definición del inventario a publicar.
- **Documentación:** En esta fase se describen los datos a publicar, con el fin de hacerlos comprensibles y facilitar su búsqueda y uso.
- **Estructuración, cargue y publicación:** En esta fase se identifican los diferentes procesos de cargue y publicación de los datos en la plataforma tecnológica que permitirá el posterior acceso a los datos abiertos a los ciudadanos y a la sociedad en general.

### Preparación para la apertura de datos

Antes de ejecutar las fases sugeridas para el proceso de apertura de datos, se recomienda a las entidades tener en cuenta las siguientes actividades, las cuales se requieren durante todo el proceso:

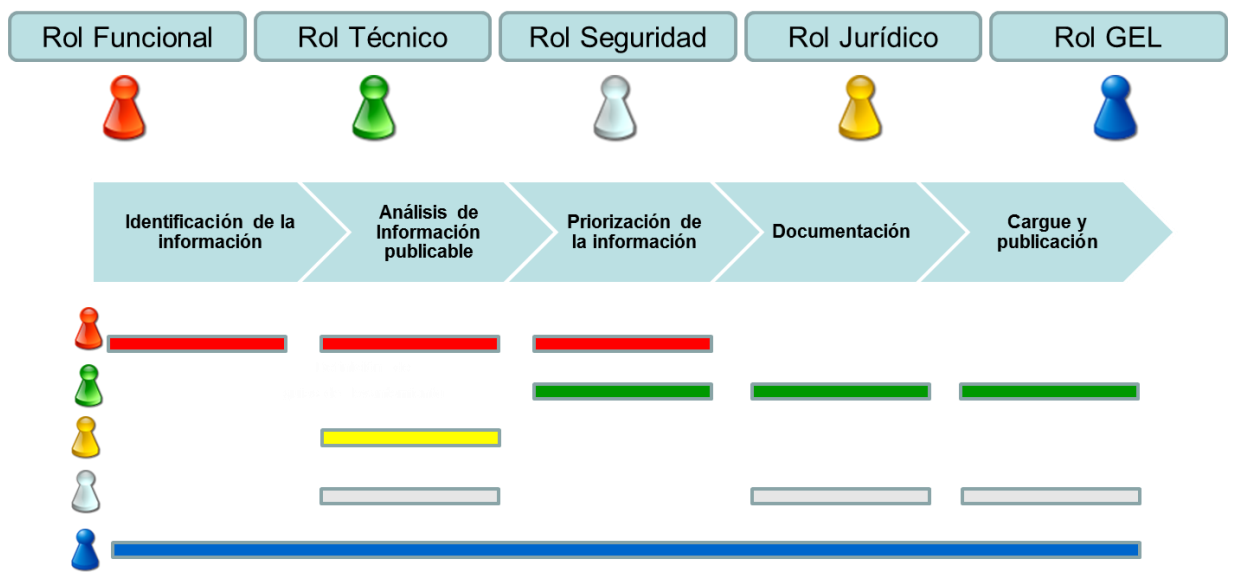
- Identificar el equipo de trabajo.
- Elaborar un plan de trabajo detallado.

- Realizar reuniones de seguimiento con los grupos de trabajo.

### ¿Cómo puedo identificar mi equipo de trabajo?

La apertura de datos implica una serie de análisis y acciones que incluyen lo tecnológico, lo jurídico y lo funcional y por ello, un requisito imprescindible para adelantar dicho proceso, es contar con el apoyo de un grupo de trabajo que lidere y soporte las acciones requeridas. Por tanto, la identificación de los grupos de trabajo se fundamenta en los siguientes cinco roles necesarios dentro de un proceso de apertura de datos:

**Figura 5: Roles y actividades necesarios para implementar la apertura de datos**



De acuerdo con la ilustración anterior, la identificación de los grupos de trabajo debe partir de los siguientes roles o competencias:



**Rol funcional:** La principal tarea es la de facilitar el acceso a la información de la entidad y analizar los mejores usos de la misma, de manera que se convierta en un activo para la organización.

**Actividades:** Identificación de la información, análisis de información publicable, priorización de la información.

**Perfil:** Personal a cargo de procesos de negocio de la entidad que tiene el conocimiento de los mismos y de la información que se produce.



**Rol técnico:** La principal tarea es la de facilitar y orientar el acceso a los sistemas y repositorios de información de la entidad y liderar la publicación de los datos en formatos abiertos.

**Actividades:** Priorización de la información, documentación, estructuración, cargue y publicación.

**Perfil:** Personal técnico, preferiblemente ingenieros de sistemas con conocimientos en estándares abiertos, identificación de información, extracción y manejo de bases de datos. Conocimiento en los requerimientos para la publicación y manejo del portal web de la entidad.



**Rol seguridad:** La principal tarea del rol de seguridad es la de garantizar un manejo adecuado de los riesgos y las políticas de seguridad de la entidad para la protección de la información.

**Actividades:** Análisis de la información publicable, documentación, estructuración, cargue y publicación.

**Perfil:** Personal técnico, preferiblemente ingenieros de sistemas con conocimientos en aseguramiento de la calidad en la información e identificación de riesgos de seguridad.



**Rol jurídico:** El rol principal es el de servir como apoyo para analizar, asesorar, conceptuar y orientar sobre los datos que son susceptibles de poner a disposición de cualquier persona, sin que esto implique la vulneración de los derechos fundamentales de los individuos y el incumplimiento de la normatividad, en cuanto a respetar la reserva legal que tienen algunos datos.

**Actividades:** Análisis de los datos que se pueden publicar.

**Perfil:** Abogado o profesional con conocimientos en el marco jurídico de los procesos de la entidad y en temas relacionados con la transparencia y el acceso a la información pública, y con conocimiento en la reserva legal y protección que tienen algunos datos e información, en virtud de la Ley.



**Rol GEL:** La principal actividad es la de actuar como un canal de comunicación entre Gobierno en línea y la entidad, para la toma de decisiones. Por otro lado, debe estar presente en la toma de las decisiones en cada uno de los procesos de la apertura de datos en lo concerniente a la entidad. El rol también tiene la responsabilidad de coordinar acciones al interior de la entidad. Es el dinamizador.

**Actividades:** Identificación de la información, análisis de información publicable, priorización de la información, documentación, estructuración, cargue y publicación.

**Perfil:** CIO o servidor público delegado por la entidad con habilidades en coordinación y dinamización de equipos de trabajo.

Las entidades podrán utilizar la siguiente documentación que permitirá un mayor control y orden en la ejecución de las actividades para la apertura de datos:

- Apuntes de la reunión.
- Actas de toma de decisión.
- Planilla de asistencia.
- Informes de seguimiento.

Una vez identificados los grupos y elaborada la documentación, se debe consensuar con todos los roles involucrados en el proceso de apertura de datos un calendario o cronograma de sesiones de trabajo con el propósito de sincronizar las sinergias de los diferentes grupos de trabajo.

#### RECUERDE:

1. Una **elección adecuada** del equipo de trabajo puede minimizar el tiempo utilizado en cada una de las fases y avanzar más rápidamente.
2. Lo **importante** no son los cargos sino **los roles**. Dos o más de los roles establecidos anteriormente pueden ser ejercidos por una sola persona.

## ¿Cómo elaboro mi plan de trabajo?

La fase de planificación incluye la definición de las tareas a realizar, los tiempos, los productos que se obtendrán y el esfuerzo que se destinará para las diferentes tareas. Esto implica establecer de la forma más precisa estimaciones de las tareas de la apertura de datos y determinar los recursos necesarios.

Los objetivos de esta etapa son los siguientes:

- Planificar y documentar estimaciones de tiempo y actividades necesarias para la identificación del inventario de datos.
- Acordar compromisos para el desarrollo de la fase inicial con todas las personas y áreas involucradas de la entidad.
- Organizar todos los recursos que sean necesarios para la apertura de los primeros conjuntos de datos.

### Establecimiento de estimaciones

Engloba el conjunto de actividades que van encaminadas a predecir la cantidad de información del proceso de apertura, el tiempo, el equipo y el capital humano requeridos. Este proceso proporciona una visión clara de las tareas que han de realizarse durante el desarrollo del proceso y los recursos necesarios.

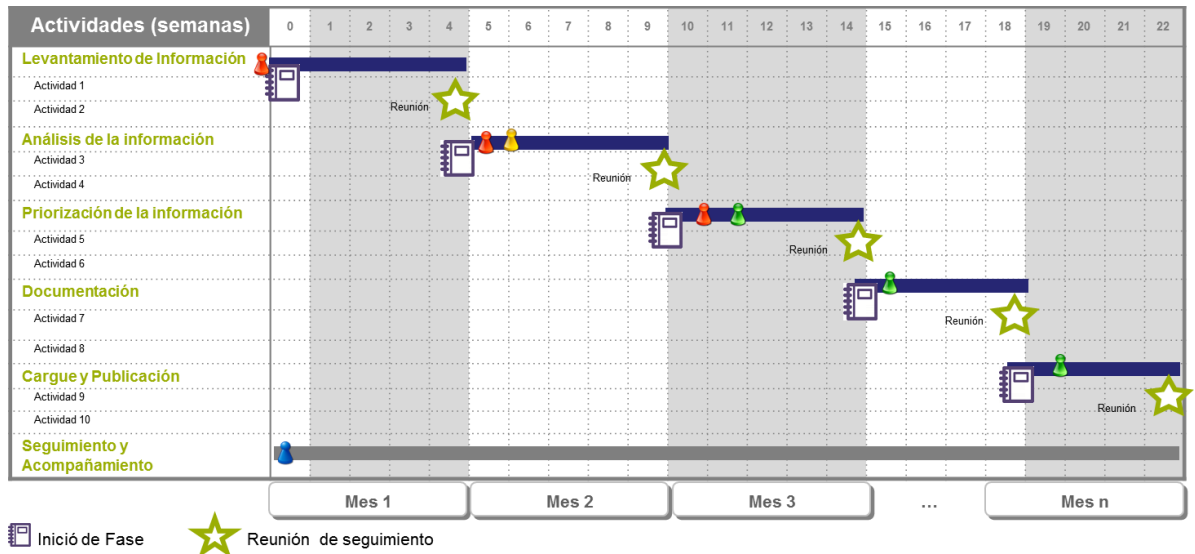
### Planificación del proyecto de apertura de datos

Engloba la organización de las tareas en orden lógico, la identificación de dependencias entre las mismas, la identificación de hitos principales y el establecimiento del calendario para su desarrollo, lo que proporcionará una línea base que servirá durante todo el seguimiento y control del proyecto. Además de la identificación de riesgos que pueden afectar al proyecto.

Implica básicamente cubrir cinco necesidades:

- Prever todas las tareas que deben realizarse.
- Proporcionar una línea base para el control del progreso del proyecto.
- Asignar tareas a los miembros del equipo.
- Asegurar que cada miembro del proyecto sepa cuáles son sus responsabilidades.
- Identificar los riesgos que puedan afectar al desempeño del proyecto.

Figura 6: Plan de trabajo



**Nota:** La anterior es una ilustración de ejemplo que puede servir para la elaboración de un plan de trabajo. Se recomienda utilizar una herramienta de planificación de trabajo, por ejemplo MS Project; en su defecto, si en la entidad no posee licenciamiento, puede realizar el seguimiento en MS Excel.

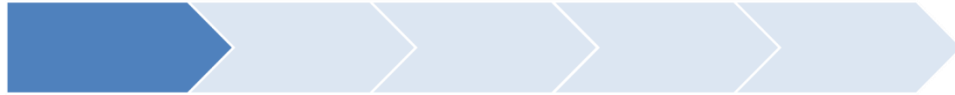
### ¿Qué debo tener en cuenta en mis reuniones individuales con los grupos de trabajo?

Como producto de las sesiones individuales de trabajo se debe obtener:

- El inventario de datos consensuado.
- Análisis y visto bueno de los datos a abrir con base en un análisis jurídico.
- Priorización de los datos a publicar, en función del impacto a la ciudadanía y al esfuerzo para publicarlo.
- Definición de estándares/formatos a utilizar y estructuración de los datos.
- Documentación de los procesos de apertura de datos.
- Elección el mecanismo conveniente para la publicación.
- Publicación de los primeros conjuntos de datos.

**RECUERDE:**

1. Después de realizar la identificación del equipo de trabajo, **elabore un plan de trabajo detallado.**
2. **Coordine reuniones** de trabajo individuales.



### 3. ¿Cómo realizo la identificación de la información?

- ❖ El **objetivo** de esta etapa es reunir toda la información sobre la cual se basan las fases siguientes del proceso de apertura de datos.
- ❖ El **resultado** esperado para esta etapa es el inventario de información de la entidad para su posterior publicación.
- ❖ **Perfiles requeridos:** Rol funcional y Rol GEL.
- ❖ **Actividades a desarrollar:** Identificación de la información, elaboración de la plantilla de identificación de la información y seguimiento de actividades.
- ❖ **Recursos a utilizar:** Plantilla de identificación de información.
- ❖ **Otras herramientas de apoyo o referencias:** N/A

#### ¿Qué es la identificación de la información?

La identificación de la información busca realizar un levantamiento de la información que administra la entidad y que puede llegar a convertirse en datos abiertos.

Para efectos de realizar la identificación de la información y de conformidad con lo establecido en la normatividad, se tendrá en cuenta la siguiente definición:

- **Información pública:** Es la agrupación ordenada de datos, que permite otorgarles una utilidad y uso en determinado contexto, y que se genera a partir del desarrollo de actividades para el funcionamiento del Estado, es decir de los registros periódicos de las actividades misionales de las entidades o como consecuencia del ejercicio de funciones de rutina en el Estado.

La entidad deberá identificar la información que administra mediante la descripción de los siguientes ítems:

- **Nombre de información:** Nombre que agrupa y describe de manera general las cualidades y funciones de la información.
- **Descripción:** Define de qué se trata la información y brinda una descripción a alto nivel de su contenido.

- **Área responsable de la información:** Nombre del área responsable de la generación y administración de la información.
- **Tipología de Información:** Identifica el tipo de Información, la cual ha sido establecida en los siguientes tipos:
  - Económica y comercial (financiera, de empresas, estadísticas económicas, tales como las referidas a inflación, desempleo).
  - Ambiental (datos meteorológicos, datos hidrográficos, uso de tierras).
  - Agrícola y pesquera (cosechas, explotaciones, utilización de recursos, producción, precios, insumos).
  - Social (demografía, actitudes, salud, enfermedades, datos censales).
  - Seguridad (criminalidad y condenas, decisiones judiciales).
  - Política (comunicados de prensa, actas de las administraciones centrales y territoriales).
  - Científica (investigaciones).
  - Geográfica (mapas, direcciones, fotografías del territorio, datos geológicos e hidrográficos, datos topográficos).
  - Transporte y tráfico (vías, congestión vehicular, obras, desvíos, cámaras en carretera, accidentes).
- **Ámbito geográfico:** Identifica el ámbito geográfico de la información (Nacional, Departamental, Municipal).
- **Idioma:** Establece el idioma en el que se encuentra la información.
- **Fuente Primaria:** Identifica si la información es el resultado de procesos de consolidación, análisis o transformaciones; o si por el contrario esta no ha sido procesada y puede ser considerada información de fuente primaria.
- **Evidencia de solicitud:** Identifica si hay evidencia de solicitud de dicha información y se identifica al demandante de la misma.
- **Tipo de Información/Origen:** Establece si la información está soportada por un documento físico, por un medio digital o un sistema, o por algún otro tipo de formato audio visual.
- **Frecuencia de generación de información:** Identifica la frecuencia con que se genera la información en la entidad, o la frecuencia con que la información cambia y es actualizada.
- **Formato:** Identifica el formato en el cual se encuentra la información identificada.

## ¿Qué tengo que hacer para realizar la identificación de la información?

Para realizar la identificación de la información (y para cada una de las etapas de esta guía) la entidad podrá apoyarse en el documento de Excel adjunto (Plantilla de Apertura de



Datos), el cual servirá como instrumento de apoyo para la ejecución de las distintas etapas para la apertura de datos.

**Figura 7: Hoja de Identificación de información**

Ministerio de Tecnologías de la Información y las Comunicaciones		Plantilla de identificación de información		Gobierno en línea Colombia	
<b>Información de la entidad</b>					
Entidad:	Ministerio de Educación Nacional				
Sector:	Educación Nacional				
Nombre del funcionario(s):	xxxx xxxx				
Área(s) responsable(s):	OFICINA DE TECNOLOGIA Y SISTEMAS DE INFORMACIÓN				
<b>Información de contacto del funcionario(s)</b>					
Teléfono:	311xxxxxxx				
E-mail:	xxxxxxxx@mineducacion.gov.co				
Nombre de la información	Descripción	Área responsable de la información	Tipología de Información	Ámbito geográfico	Idioma
Instituciones de Educación superior	identificación, nombre y caracterización de las Instituciones de Educación Superior (Universidades, Instituciones Técnicas y	Oficina de tecnología y sistemas de Información	Educación	Nacional	Español

**Proceso de identificación de información de la entidad:**

La entidad debe verificar si dentro del marco de su política de seguridad de la información, ha hecho un trabajo de levantamiento del *inventario de activos de información*<sup>3</sup> el cual puede ser mapeado con la información susceptible de ser abierta. En el caso en que no se haya realizado dicho inventario, se debe consultar con el rol de seguridad de la información propuesto para conocer el estado de avance de la entidad.

En todo caso la entidad debe establecer las fuentes de información y el contexto que desea levantar teniendo en cuenta los *procesos de la entidad*, clasificados de la siguiente manera:

<sup>3</sup> Por activos de información se entiende al conjunto de elementos con valor informativo que son propiedad de una empresa, entidad o individuo, y que reflejan su actividad. Por lo que respecta a los activos de información que poseen los organismos públicos, existen cuatro categorías: Publicaciones oficiales, Bases de datos públicas, Repositorios digitales, Archivos y registros administrativos.

Figura 8: Clasificación de Procesos al interior de la entidad



Para cada tipo de proceso, la entidad debe tener en cuenta la información que está siendo generada:

- *Procesos Misionales:* En los procesos misionales la entidad debe verificar la información relacionada con los trámites, servicios y procesos que soportan el “día a día” de su trabajo, de acuerdo con sus funciones y competencias. La entidad debe incluir acá la información particular relacionada con los actores involucrados en dichos trámites y servicios así como los sujetos y dominios de negocio soportados.
- *Procesos Estratégicos:* La entidad deberá revisar dentro de su esquema de procesos, aquellos encaminados en la mejora continua y en el cumplimiento de sus objetivos estratégicos, que incluya información por ejemplo relacionada con la prestación de los servicios de la entidad, la atención a los usuarios, etc.
- *Procesos de Apoyo:* La entidad debe revisar la información que es generada y que soporta los procesos de apoyo relacionados con la contratación, el recurso humano, la información financiera y contable.
- *Procesos de Evaluación:* Se debe revisar la información que soporte la toma de decisiones y los procesos de monitoreo y evaluación dentro de la entidad.

Las actividades relacionadas con el levantamiento de información a partir de los procesos, deben ser realizadas por el responsable y conocedor de los mismos dentro de la entidad. Para el caso de nuevos procesos, el líder del proceso deberá entregar al grupo de apertura de datos toda la información necesaria junto con la Plantilla de Apertura de Datos para que estos sean integrados al inventario de la entidad y para que se realice el posterior análisis de la misma.

## ¿Para qué me sirve realizar el inventario de información de mi entidad?

Realizar una identificación del inventario de información trae consigo los siguientes beneficios:

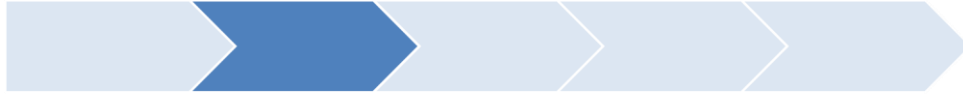
- Aumenta el conocimiento que la entidad tiene de sus activos de información.
- Ofrece elementos que permiten mejorar la comunicación entre la entidad y sus usuarios consumidores de información.
- Ayuda a determinar las necesidades de los usuarios eficientemente para que puedan ser satisfechas cuando soliciten algún tipo de información.
- Permite aumentar la percepción de confianza que los usuarios pueden tener de la entidad y del Estado en general.
- Genera transparencia administrativa que conlleva además la no discriminación entre empresas y particulares a la hora del acceso a la información.

### RECUERDE:

1. **Identifique** la información que está bajo la responsabilidad de la entidad.
2. **Diligencie** el formato en Excel que acompaña la guía para la correcta identificación de la información.

### ¡IMPORTANTE!

LA IDENTIFICACIÓN DE LA INFORMACIÓN ES LA BASE PARA LA ELABORACIÓN DEL INVENTARIO DE INFORMACIÓN DE DATOS ABIERTOS.



#### 4. ¿Cómo realizo el análisis de la información?

- ❖ El **objetivo** de esta etapa es analizar la información identificada en la primera fase del proceso, para separar la información publicable y que no contemple ninguna restricción normativa bajo la Ley Colombiana.
- ❖ El **resultado** esperado para esta etapa es tener el inventario de información publicado en la página Web de la entidad o en el catálogo de datos del estado colombiano, identificando la información publicable, la no publicable y las razones de reserva de la misma.
- ❖ **Perfiles requeridos:** Rol funcional, Rol Seguridad, Rol Jurídico y Rol GEL.
- ❖ **Actividades a desarrollar:** Identificar la normatividad y acuerdos de confidencialidad de la información al interior de la entidad, para analizar la información de acuerdo con los criterios jurídicos para la apertura de datos y verificar con base a la ley, cual información se puede publicar.
- ❖ **Recursos a utilizar:** Plantilla de análisis de información, Inventario de información, sitio web de la entidad, Catálogo de datos Abiertos del Estado Colombiano.
- ❖ **Otras herramientas de apoyo o referencias:** Normatividad vigente: Leyes, Decretos, Acuerdos.

Con base en el marco normativo vigente, se pueden considerar los siguientes tipos de información:

- **Datos e información publicable:** Aquella que debe ponerse a disposición de cualquier persona, por quienes estén obligados, por no ser de aquella que la Ley se ha dado el carácter de reservada.
- **Datos e información no publicable:** Aquella a la que la Ley le ha otorgado el carácter de reservada y protección legal y por tal motivo no puede ser publicada.
- **Datos e información personal semiprivada:** Datos e información personal que no es de dominio público, pero que ha sido obtenida u ofrecida por orden de una autoridad administrativa en el cumplimiento de sus funciones o en el marco de los principios de administración de datos personales. Esta información puede ser o no sujeta a reserva por su titular.

Tanto la información publicable, como la que no es posible publicar, obedecen a la protección legal que se le da a los datos y a la información.

En cuanto a la información semiprivada, esta corresponde a aquella que teniendo protección legal, es posible, para una entidad pública, ponerla a disposición de otra entidad pública, pero solamente para efectos de las tareas misionales que, en virtud de la ley, se le han otorgado a la entidad solicitante de la información.

A manera de ejemplo ilustrativo, para orientar lo que se entiende por información semiprivada, están las bases de datos con información personal, recolectadas por el SISBEN, o por la Registraduría. Ellas compilan información personal, que tiene el carácter de reservada, pero en virtud de las funciones legales de cada entidad pública, pueden entregarse, bajo estrictos protocolos de seguridad, para garantizar su reserva y divulgación no autorizada al público en general.

Para la entrega de datos semiprivados, de acuerdo con el ejemplo contenido en el párrafo anterior, la entidad debe:

- Verificar que se trata de datos semiprivados conforme a la clasificación y orientación contenida en esta guía.
- Verificar que el solicitante es una entidad pública y que el funcionario responsable, lo es de acuerdo con sus funciones legalmente asignadas mediante el manual de funciones de la entidad.
- Verificar que la información solicitada, corresponde al ejercicio de las funciones misionales de la entidad (Por ejemplo, no sería lógico que el Ministerio de Agricultura, solicite la base de datos de los pacientes portadores de VIH, pues esa temática no le corresponde dentro de sus áreas misionales)
- Asegurarse que la entidad que recibe la información, garantizará la confidencialidad y reserva de la misma, para lo cual se recomienda suscribir un documento de compromiso de confidencialidad de la entidad receptora de la información.
- Entregar la información con todos los protocolos y encriptaciones que garanticen su protección y no divulgación a terceros.

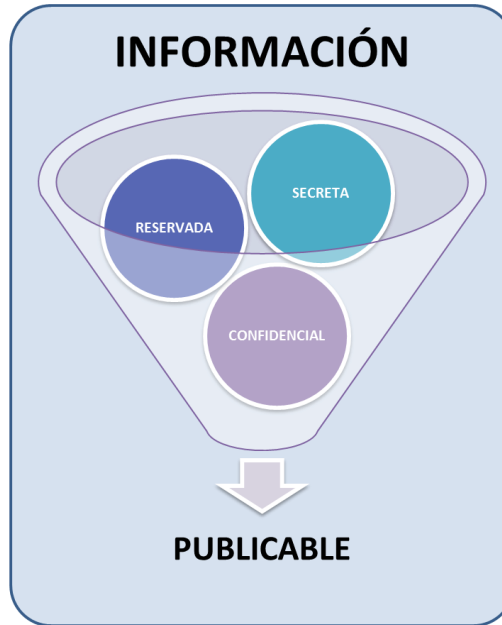
En todos los casos, es necesario que la entidad, antes de entregar o publicar los datos o la información, efectúe la clasificación y análisis de la misma, para verificar que información pueden publicar, y cual no, conforme a la normatividad vigente.

### ¿Cómo ejecuto la revisión de la información bajo los criterios jurídicos?

Para clasificar la información se recomienda a la entidad utilizar el método de separación. Este método permite filtrar por diferentes criterios y separar la información que para la implementación de datos abiertos no es información publicable, utilizando el marco normativo enunciado en esta guía. La información que afecte la defensa nacional, la

seguridad del Estado, la protección de la seguridad pública, la relacionada con actuaciones sometidas por una norma al deber de reserva, secreto o confidencialidad, entre otras, no constituye información publicable.

**Figura 9: Método de separación**



La información comprendida en las categorías relacionadas en la Tabla 1, no es publicable.

**Tabla 1: Tipos de información no publicables**

Tipo de Información	Aspectos Relacionados
<b>Información secreta:</b> Es toda aquella información que tiene establecido tal carácter en razón de leyes especiales en el marco jurídico colombiano, que vulnere, afecte o atente contra los siguientes aspectos relacionados.	<b>Actuaciones políticas:</b> Se refiere a los expedientes que contengan información sobre actuaciones del Estado en el ejercicio de sus competencias constitucionales no sujetas a derecho administrativo.
	<b>Defensa nacional y seguridad del Estado:</b> Quedan excluidos los expedientes que contengan información sobre la defensa nacional o sobre la seguridad interior o exterior del Estado. La Ley 52 de 1990 establece que tiene carácter de reservado cierta información del Consejo Nacional de Seguridad.
	<b>Política monetaria:</b> Los expedientes relativos a actuaciones administrativas derivadas de la política monetaria; estos pueden definirse como política económica dirigida a controlar las variaciones en la cantidad o costo para asegurar y mantener la estabilidad económica de Colombia. El Decreto 624 de 1989 determina el carácter de reservado de cierta información que aparezca en las declaraciones tributarias.
	<b>Investigación de delitos:</b> Según el Decreto 300 de 1997 le da el carácter de reservada a cierta información del Instituto Nacional Penitenciario y Carcelario (INPEC). Se excluyen los expedientes tramitados para la

Tipo de Información	Aspectos Relacionados
	<p>investigación de los delitos cuando pudiesen colocarse en peligro la protección de los derechos y libertades de terceros o las investigaciones que se estén realizando.</p> <p><b>Material clasificado:</b> En la Ley 766 de 2002, por medio de la cual se aprueba la “Convención sobre asistencia en caso de accidente nuclear o emergencia radiológica” se protege la confidencialidad de los datos en caso de emergencia nuclear o radiológica.</p> <p><b>Información registral:</b> Según lo establecido en la Ley 95 de 1985 se contempla como carácter de reservados a ciertos documentos que reposen en la Registraduría Nacional del Estado Civil, relacionados con temas como los de la identidad de los ciudadanos, datos biográficos, específicos el acceso a los datos contenidos en el Registro Civil, entre otros.</p> <p>Toda información que, con base en el análisis de riesgo haya sido clasificada como secreta o un tipo que cumpla con la definición de información secreta presentada en este documento.</p>
<p><b>Información reservada:</b> Es toda aquella información que tiene establecido tal carácter en razón de leyes especiales en el marco jurídico colombiano, que vulnere, afecte o atente contra el objetivo de garantizar el buen funcionamiento de los poderes públicos y, en general, preservar la eficacia de las decisiones que tomen las administraciones públicas. Por otro lado no se permite la publicación de la información, si contempla alguno de los siguientes aspectos relacionados.</p>	<p><b>Atenta con relaciones internacionales:</b> Todas aquellas obligaciones y disposiciones de los acuerdos internacionales sobre protección de los derechos de propiedad intelectual, que estén vigentes y afecten directamente a una entidad.</p> <p><b>Pone en riesgo la vida, la dignidad, la seguridad o la salud de las personas:</b> Toda aquella información que contengan datos referentes a la intimidad de las personas estará reservado a éstas, existiendo la posibilidad de ejercer un control sobre la misma. La Ley 1266 de 2008 le da un tratamiento especial a aquellos datos de información personal que no tengan el carácter de públicos almacenada en bases de datos para preservar el derecho Constitucional al Habeas Data.</p> <p><b>Propiedad industrial y derecho de la competencia:</b> Todas aquellas materias protegidas por el secreto comercial o industrial, evitando así el acceso a informaciones confidenciales que puedan obrar en poder de la Administración y garanticen el derecho a la libre competencia.</p> <p><b>Desproteger descubrimientos científicos, tecnológicos o culturales desarrollados:</b> La información conservada por instituciones educativas y de investigación, tales como centros escolares, universidades, archivos nacionales, bibliotecas, centros de investigación, entre otros, que puedan vulnerar el derecho de patentes de dichos temas.</p> <p>Toda información que, con base en el análisis de riesgo haya sido clasificada como reservada o un tipo que cumpla con la definición de información reservada.</p>
<p><b>Información confidencial:</b> Es toda aquella información que tiene establecido tal carácter en razón de leyes especiales en el marco Jurídico colombiano, que vulnere, afecte o atente contra los aspectos mencionados en la columna “Aspectos Relacionados” de</p>	<p><b>Estrategias de negocio, competitividad, expansión, entre otros:</b> La información reportada, relativa a secretos empresariales u otro respecto de la cual exista norma legal de reserva o confidencialidad tendrá carácter reservado frente a terceros y solamente podrá ser conocida por una entidad.</p> <p><b>Historias clínicas, ubicación geográfica específica, entre otros:</b> La Ley 23 de 1981 establece como carácter de reservado a la historia clínica, estableciendo como necesario la autorización del paciente para que dicha información sea otorgada a un tercero. Toda información que se considere que pudiese ubicar geográficamente, la posición específica de los ciudadanos (GPS) y vulnere su derecho de confidencialidad de la información personal, y atente</p>

Tipo de Información	Aspectos Relacionados
este ítem.	contra su seguridad no será considerada publicable.
	<b>Secreto profesional:</b> En la Ley 43 de 1990 se le da el carácter de reservada a la información suministrada al contador, se aplica la figura del secreto profesional.
	Toda información que, con base en el análisis de riesgo haya sido clasificada como confidencial o un tipo que cumpla con la definición de información confidencial presentada en este documento.

## ¿Cómo identifico la información publicable?

**Nota:** Para esta segunda tarea se tendrá como insumo el archivo adjunto en formato Excel para la apertura de datos que acompaña esta guía, en la pestaña de análisis de información.

En este sentido, los aspectos a tener en cuenta son los siguientes:

- El Rol Jurídico debe Identificar junto con su equipo jurídico la normativa de la entidad que regula la información previamente levantada en la etapa inicial para identificar la información confidencial, reservada y secreta.
- Para el correcto diligenciamiento de la plantilla de análisis se deben responder a los diferentes criterios del método de separación especificando a cuáles afecta, atenta o entra en contraposición en cada uno de ellos. Para ello responda **SI**, en el caso que la información afecte totalmente, **NO** en el caso que no afecte y **Parcialmente** en el caso que un subconjunto de la información no cumpla con los criterios anteriormente mencionados. Adicionalmente diligencie la columna de soporte jurídico teniendo en cuenta el ejercicio de separación realizado.



**Figura 10: Hoja de Análisis de información**

Ministerio de Tecnologías de la Información y las Comunicaciones República de Colombia www.mintic.gov.co		Plantilla de Análisis de apertura de información		Gobierno en línea Colombia	
<b>Información de la entidad</b>					
Entidad:	Ministerio de Educación Nacional				
Sector:	Educación Nacional				
Nombre del funcionario(s):	XXXXXXXXXX				
Área(s) responsable(s):	OFICINA DE TECNOLOGÍA Y SISTEMAS DE INFORMACIÓN				
<b>Información de contacto del funcionario(s)</b>					
Teléfono:	311XXXXXXXXXX				
E-mail:	XXXXXXXXXX@mineducacion.gov.co				
Nombre de la información	Descripción	Área responsable de la información	Tipología de información	Ámbito geográfico	Idioma
Instituciones de Educación superior	Instituciones de Educación Superior (Un	Oficina de tecnología y sistemas de Informació	Educación	Nacional	Español
Programas de Educación superior	ón de los programas que poseen la	Oficina de tecnología y sistemas de Informació	Educación	Nacional	Español
Directorio Único de Establecimientos Educativos	cterización de los establecimientos	Oficina de tecnología y sistemas de Informació	Educación	Nacional	Español

**Sugerencia:** El diligenciamiento de la plantilla puede ser adaptado por las entidades con el propósito de tener algún elemento o instrumento adicional que permita realizar el análisis jurídico.

### ¿Para qué me sirve realizar el análisis de la información?

Realizar un análisis de información trae consigo los siguientes beneficios:

- Ofrece una clasificación que ayuda a la identificación de niveles de seguridad de la información.
- Aumenta el conocimiento de la normatividad que apoya a la administración de la información.
- Garantiza la publicación bajo un análisis legal para su libre uso público sin repercusión en la publicación sobre la entidad.
- Permite identificar las limitaciones de tipo legal para publicar la información que administra.
- Permite tener la certeza legal sobre la información que se puede publicar.

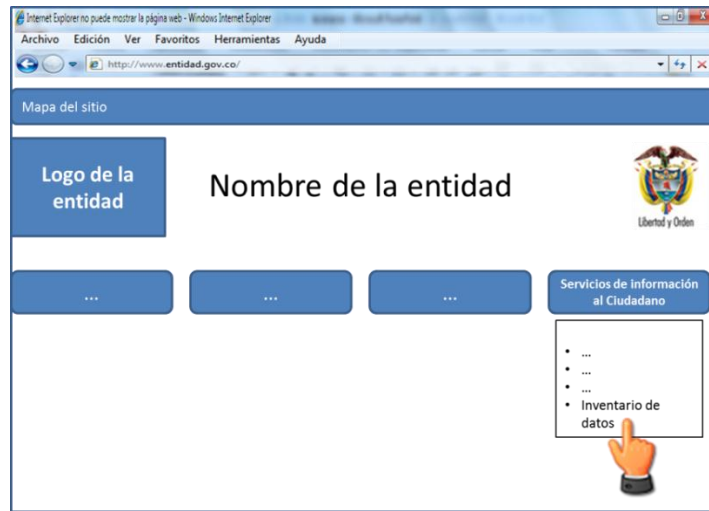
### ¿Cómo realizo la publicación del inventario de información en el sitio web de la entidad?

Tomado del Manual 3.0, Manual para la implementación de la Estrategia Gobierno en línea de la República de Colombia: *Una vez se tengan los conceptos claros y con base en el inventario de información, la entidad debe publicar dicho inventario en su sitio web, en la sección denominada “Servicios de información al Ciudadano”. En dicho inventario debe*

señalarse la información que es reservada de acuerdo con lo establecido en la ley sobre reserva de información.

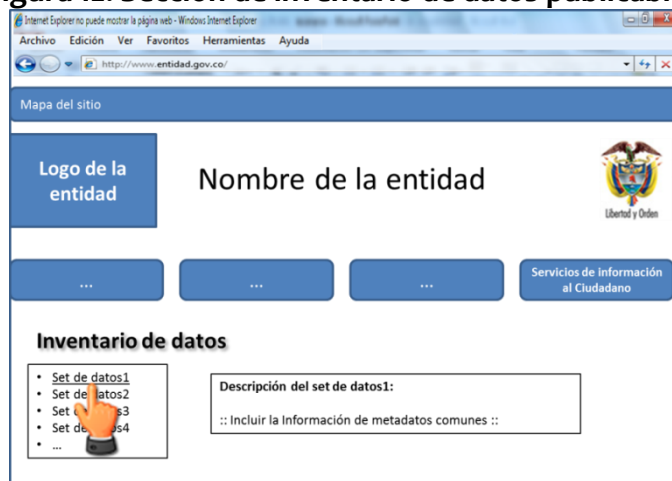
Cumpliendo con lo establecido en el marco de la Estrategia de Gobierno en Línea para el Modelo de Datos abiertos la entidad debe publicar todo el inventario de datos en la sección de “Servicios de información al Ciudadano”, a continuación se muestra una imagen ilustrativa del lugar de publicación:

Figura 11: Sección de publicación del inventario de datos



En la sección de inventario de datos se deben enlistar todo el inventario de datos. Para el inventario de datos publicables se publica el listado junto con su descripción de metadatos común, identificados en los capítulos anteriores:

Figura 12: Sección de inventario de datos publicables

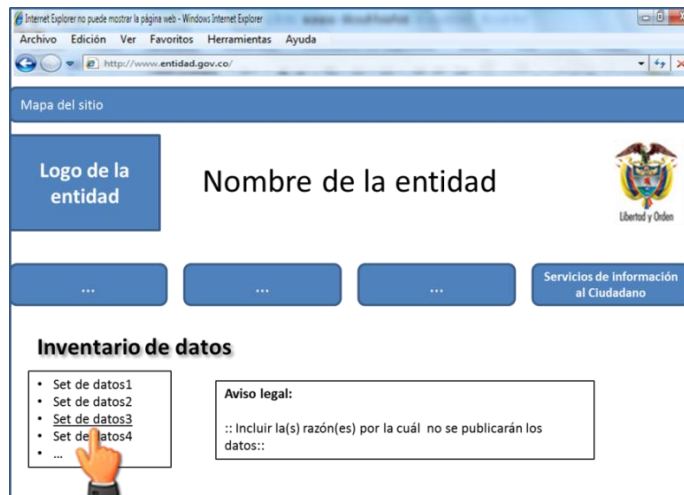


En cuanto al inventario de los datos no publicables, se considera lo siguiente:

Tomado del Manual 3.0 Manual para la implementación de la Estrategia Gobierno en línea de la República de Colombia: *Dado que existen restricciones legales para publicar cierto tipo de información, conviene hacer una revisión jurídica sobre estos aspectos. Por ello, la entidad debe identificar las limitaciones de tipo legal para publicar la información que administra y la forma en que la información de carácter confidencial y personal asociada, será protegida.*

Por lo anterior se deben aclarar los motivos por los cuales la entidad se reserva la publicación de los datos la entidad, con un aviso legal en la descripción del set de datos:

Figura 13: Sección de inventario de datos no publicables



En el capítulo de identificación de la información retomar la ficha en Excel en su pestaña *de Identificación de la información* para cada inventario que se reserva mencionar lo establecido en la columna: Motivos de reserva, para detallar el aviso legal de los datos no publicables.

**Nota:** Utilice cualquier medio de publicación para el inventario, puede presentar el inventario como HTML, Excel, Flash o cualquier tipología que utilice la política de diseño gráfico de su portal. Se debe resaltar en esta sección un link de enlace hacia el portal de datos del estado: [www.datos.gov.co](http://www.datos.gov.co), con el fin que los usuarios puedan tener acceso a todos los datos de las entidades del estado.

**RECUERDE:**

1. **Identifique** la normativa de su entidad que puede ser empleada para realizar el correcto análisis de la información.
2. **Analice** la información utilizando el método de separación para identificar la información secreta, confidencial y reservada.
3. **Diligencie** el formato en Excel de apertura de datos que acompaña la guía para el correcto análisis de la información.
4. **Publique** el inventario de información que incluya la información que es publicable y la que no y el soporte jurídico para su no publicación.

**¡IMPORTANTE!**

EL ANÁLISIS DE LA INFORMACIÓN ES EL FILTRO JURÍDICO QUE CLASIFICA LA INFORMACIÓN PARA SER PUBLICABLE.



## 5. ¿Cómo realizo la priorización de los datos?

- ❖ El **objetivo** de este capítulo es definir los parámetros que van a permitir priorizar los datos a publicar en formato abierto, al igual que establecer el cronograma a corto, mediano y largo plazo, para la publicación.
- ❖ El **resultado** esperado para esta etapa es tener los datos ordenados bajo criterios de alto impacto y de menor dificultad para la entidad.
- ❖ **Perfiles requeridos:** Rol funcional, Rol técnico y Rol GEL.
- ❖ **Actividades a desarrollar:** Identificación de la fuente de información, valoración de los criterios de evaluación, definición de los datos a publicar.
- ❖ **Recursos a utilizar:** Plantilla de priorización de los datos.
- ❖ **Otras herramientas de apoyo o referencias:** N/A

### ¿Qué se requiere para la priorización de datos?

Para la etapa de priorización de los datos se pueden usar los siguientes parámetros con el fin de identificar los primeros conjuntos de datos a publicar.

- **Parámetros a priori:** Esta categoría de parámetros son aplicados a todo el inventario de información y permite definir el cronograma de publicación de los conjuntos de datos. Incluye dos tipos de parámetros a evaluar:
  - **Impacto:** Mide la incidencia y los beneficios agregados de cada uno de los elementos del inventario de información, sobre el ciudadano, sector privado, el Estado, entre otros.  
Ejemplo: La necesidad demandada de información a través de un número mayor de ciudadanos, puede generar mayor impacto; de manera similar, los datos que muestran situaciones a nivel de todo el territorio nacional pueden tener mayor impacto ya que abarca un mayor foco de interés.
  - **Dificultad de implantación:** Permite evaluar el grado de facilidad para el proceso de apertura de datos.  
Ejemplo: Si no se requiere hacer desarrollos para la extracción de los datos, no tiene dificultad o es mínima, mientras que si se requiere diseñar y estructurar

nuevos procesos de extracción y depuración de la información habrá mayor dificultad.

- **Parámetros a posteriori:** Éstos se emplean una vez priorizados los conjuntos de datos. En este sentido, se identifican los siguientes parámetros:
  - **Prelación:** Permite medir la importancia relativa entre catálogos de información, es decir, permite diferenciar aquellos catálogos cuyos criterios de “impacto” y “dificultad de implantación” son similares o iguales.
  - **Voluntad:** Este parámetro agrupa aquellos criterios que según las entidades consideren pueden afectar el proceso de apertura de datos seleccionados.

A continuación se presentan los criterios a evaluar para cada uno de los parámetros:

Figura 14: Criterios de evaluación



Cada parámetro puede tener varios criterios y a cada uno de ellos se le debe asignar una puntuación, dependiendo del número de criterios. Por ejemplo, si se identifican 4 criterios de “impacto”, entonces al criterio más importante se le asigna 4 puntos y al menos importante 1 punto. Siguiendo con este ejemplo y con base en la gráfica anterior, a continuación se detalla el ejercicio de priorización:

criterio	Escala	Peso	Descripción	Calificación
<b>IMPACTO</b>				
Información que contribuye al crecimiento económico	1-4	0,25	Aumentar la productividad de los diferentes sectores productivos del país	(1) No contribuye al Sector en su crecimiento (4) Contribuye ampliamente al sector en su crecimiento
Información que puede generar negocio de inmediato	1-4	0,25	Genera nuevos modelos de negocio que pueden surgir a partir de la publicación del catálogo de datos generados.	(1) No se ha contemplado el valor agregado. (4) Genera valor agregado.
Área de impacto	1-4	0,25	Destino del beneficio de la apertura de datos	(1) No se ha identificado. (2) Internacional. (3) Sector Público. (3) Ciudadanos/Sector Privado. (4) Regional/ Nacional. (4) Todas las anteriores.
Demanda de los datos	1-4	0,25	Impacta en calidad de vida de los ciudadanos	(1) No se ha identificado la demanda de información. (2) Demanda por parte de la ciudadanía a través de la solicitud de peticiones o servicios. (3) Demanda por parte de desarrolladores. (4) Gran demanda a través Consultas propias de la entidad.
<b>DIFICULTAD</b>				
Esfuerzo requerido para publicar	1-4	0,33	Dificultad a nivel de los procesos internos de la entidad para desarrollar la publicación por el equipo de sistemas.	(1) Requiere desarrollo (Construcción de ETL, BI, entre otros) (4) No requiere esfuerzo de desarrollo, se tienen las herramientas para extraer la información de las bases de datos
Fuente de datos	1-4	0,33	Dificultad en los sistemas de información para la extracción de la información, debido a su funcionalidad operativa.	(1) Se encuentra en un servidor con una base de datos de transacciones diaria. (4) Se encuentra en un servidor con una base de datos de reportes e históricos.
Calidad de la información	1-4	0,33	Estado de la información para la publicación de los datos.	(1) No se han realizado procesos de calidad de la información. (2) Calidad Media. (3) Alta calidad. (4) La información está certificada.

### ¿Cómo diligencio la matriz para la priorización de información a publicar?

Luego de obtener el listado oficial de criterios de evaluación del inventario de información, éstos se deben aplicar a toda la lista de datos o información que lo compone,

obteniendo de esta forma una puntuación para el Impacto y otra para la Dificultad de Implantación.

Los criterios relacionados con el parámetro de Voluntad se deben evaluar una vez priorizado todo el inventario de datos. Dicha evaluación se debe realizar aplicando la siguiente fórmula matemática a cada una de las agrupaciones de criterios (Impacto y Dificultad de implantación):

**Figura 15: Fórmula matemática priorización**

$$\begin{array}{c}
 \text{Valoración } CT_1 \\
 \vdots \\
 \text{Valoración } CT_m
 \end{array}
 =
 \begin{array}{c}
 \text{Peso } C_1 \times \text{Puntuación } C_1CT_1 \\
 + \dots + \\
 \text{Peso } C_n \times \text{Puntuación } C_nCT_1
 \end{array}$$
  

$$\begin{array}{c}
 \text{Valoración } CT_m \\
 = \\
 \text{Peso } C_1 \times \text{Puntuación } C_1CT_m \\
 + \dots + \\
 \text{Peso } C_n \times \text{Puntuación } C_nCT_m
 \end{array}$$

- $CT_x$  es la Criterio Total “x”
- $C_y$  es el criterio de priorización “y”
- $C_yCT_x$  es el criterio de priorización “y” aplicado al Criterio Total “x”
- Los criterios van desde  $x = 1 \dots n$  y  $y = 1 \dots m$

La definición de la matriz de priorización permitirá representar los siguientes cuadrantes:

- Quick-Win (QW): En él se ubican aquellos datos cuya publicación tenga un gran impacto y una dificultad de implantación relativamente baja, con lo que la publicación se podrán implementar en el corto plazo.
- Medio Plazo (MP): Este cuadrante representa los datos que, debido a su baja complejidad de implantación y su bajo o medio impacto, se deben implementar en el mediano plazo.
- Largo plazo (LP): En este cuadrante se ubican aquellos datos que tengan un alto impacto, pero cuya implantación se estime excesivamente compleja como para hacerla en el corto plazo.
- No aporta valor (NV): En este cuadrante se ubican aquellos datos que por su escaso impacto y su elevada complejidad, se considera que no aportarían gran valor.



## Instrumento de priorización de información

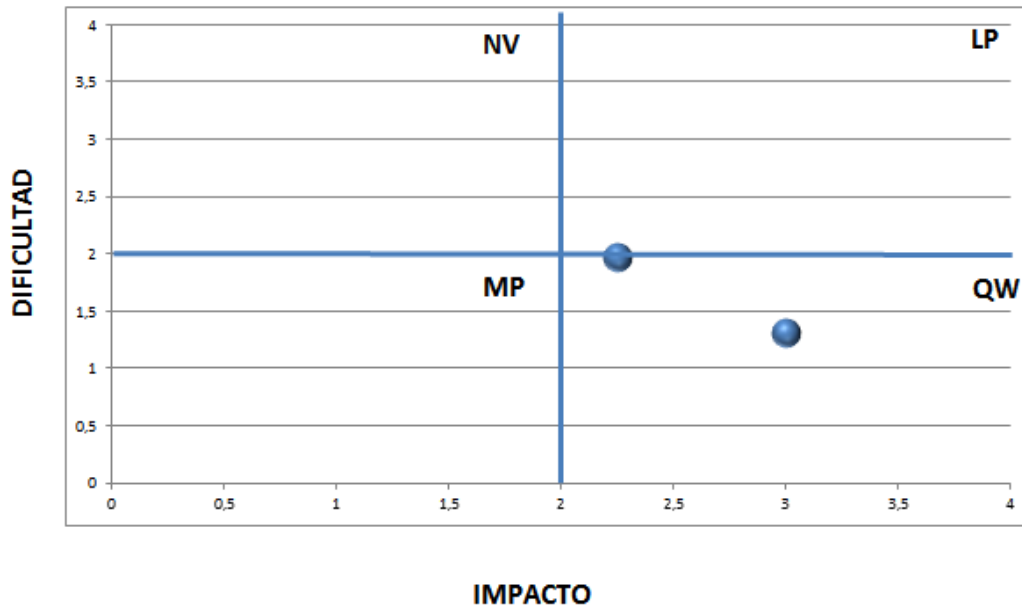
La plantilla que acompaña esta guía realiza el cálculo de los criterios totales luego de realizar la calificación en cada conjunto de datos o de información. Evaluados todos, el siguiente paso debe ser situar dichos conjuntos en una Matriz de Priorización de dos ejes, en la que se definirán los cuadrantes que permiten el plan de acción a corto, mediano y largo plazo, para la publicación de los primeros datos.

En dicha matriz se representan los valores obtenidos de los dos parámetros para cada uno de los conjuntos de datos. La representación de los valores debe ser la siguiente:

- Impacto: El valor obtenido para el impacto de los datos, se debe representar en el “Eje Y” de la matriz.
- Dificultad de implantación: En el “Eje X” de la matriz se deben representar las puntuaciones obtenidas en cuanto a la dificultad de apertura de los datos.

La matriz que se elabora tiene las siguientes características:

**Figura 16: Matriz de Priorización**



Una vez priorizados todos los datos y definidos cuáles de ellos quedan agrupados dentro de los Quick-Win, Medio Plazo, Largo Plazo o No aporta Valor, el siguiente paso debe reflejar dicha priorización en un cronograma de publicación que determine en qué momento se debe realizar la publicación de los datos. Para ello evalúe según los criterios de voluntad y prelación; descritos al inicio de este capítulo, el orden del inventario empezando por los cuadrantes QW, seguido por MP, LP y finalmente NV.

**Figura 17: Cuadro de Definición de los Datos Priorizados**

Nombre del inventario de información	Nivel de prelación
1	QW
2	MP
3	LP
4	NV

## ¿Para qué me sirve realizar una priorización de mis datos?

Realizar una priorización de los datos trae consigo los siguientes beneficios:

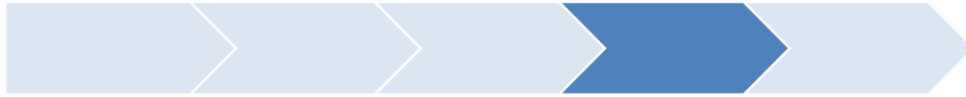
- Ofrece elementos que permiten tomar decisión sobre cuáles son los conjuntos de datos que debo enfocar primero mis esfuerzos en el proceso de apertura.
- Permite a la entidad definir su cronograma a corto mediano y largo plazo en la apertura de datos.
- Identifica el impacto de la apertura de datos sobre el esfuerzo a realizar a nivel de la entidad.

### RECUERDE:

1. **Califique** la información a priorizar.
2. Diligencie el **formato en Excel** que acompaña la guía para la correcta priorización de la información.
3. Defina los conjuntos de datos a publicar.

### ¡IMPORTANTE!

UNA VEZ DEFINIDA LA MATRIZ DE PRIORIZACIÓN SE APLICA EL PARÁMETRO DE VOLUNTAD QUE ESTABLECE EL ORDEN DE LA DEFINICIÓN DEL INVENTARIO DE INFORMACIÓN



## 6. ¿Cómo realizo la documentación de los datos?

- ❖ El **objetivo** de este capítulo es describir los datos identificados en la priorización de manera que sea comprensible para todos los usuarios.
- ❖ El **resultado** esperado de esta etapa es la ficha con la descripción de los metadatos comunes de cada de los conjuntos de datos a publicar.
- ❖ **Perfiles requeridos:** Rol técnico, Rol Seguridad y Rol GEL.
- ❖ **Actividades a desarrollar:** Realizar una identificación de estándares/formatos propuestos de publicación (Ficha Descripción de Metadatos), describir los metadatos comunes y específicos de los conjuntos de datos a publicar.
- ❖ **Recursos a utilizar:** Plantilla de documentación de datos.
- ❖ **Otras herramientas de apoyo o referencias:** N/A

### ¿Cómo se realiza la identificación de estándares/formatos propuestos de publicación para la Descripción de Metadatos?

Un metadato no es más que un dato estructurado sobre la información, o sea, información sobre información, o de forma más simple, datos sobre datos. Otras definiciones los establecen como “descripciones estructuradas y opcionales que están disponibles de forma pública para ayudar a localizar objetos”, o como “datos estructurados y codificados que describen características de instancias conteniendo informaciones para ayudar a identificar, descubrir, valorar y administrar las instancias descritas”<sup>4</sup>.

El objetivo básico de las etiquetas de documentación es aportar al catálogo de datos la meta-data subjetiva del propietario de la información (entidad) y que posiblemente no se encuentra en el contenido de los datos; todo en beneficio de las búsquedas en el Catálogo de Datos.

---

<sup>4</sup> Op. cit. “NC-RISP-01 Manual Práctico de Reutilización”. Proyecto Aporta.

**Ejemplo:** Un contenido puede estar destinado a ciudadanos, siendo este un criterio que aporta la entidad (subjeto) ya que puede que no se mencione la palabra “ciudadano” en dicho contenido. En este caso, el autor catalogará el contenido con la etiqueta “ciudadano” del eje “destinatarios”.

Bajo estas condiciones, la función que caracteriza a los metadatos es la de localizar, identificar y describir recursos, legibles e interpretables por máquina.

A continuación se enumeran los objetivos activamente perseguidos, durante el diseño del modelo de datos para la publicación de información del sector público<sup>5</sup>:

**Figura 18: Objetivos de la catalogación los metadatos**



- **Interoperabilidad con otros conjuntos de metadatos**  
La interoperabilidad de los metadatos permite la interacción entre los mismos y permite las búsquedas indexadas de los mismos.
- **Extensibilidad**  
La extensibilidad es fundamental en cualquier conjunto de metadatos. El diseño de conjuntos de metadatos debe ser capaz de incluir un núcleo central con conceptos básicos comunes al que se puedan sumar posteriormente (sin destruir la lógica del conjunto) elementos que cumplan necesidades descriptivas más específicas. La idea

<sup>5</sup> NC-RISP-01\_Manual\_Practico\_de\_Reutilizacion-Proyecto Aporta

es que una aplicación que “conozca” los metadatos básicos sea capaz de utilizarlos sin perder demasiada interoperabilidad.

- **Modularidad**

La modularidad de los metadatos debe permitir combinaciones sintácticas (a nivel de lo que se busca literalmente) y semánticas (a nivel del significado de lo que se busca) con otros metadatos.

- **Principio de utilidad**

Se han identificado tres aspectos claves desde el punto de vista de la utilidad de las descripciones. La descripción de un activo de información cualquiera debe cumplir con tres requisitos:

- Permitir el descubrimiento de activos de información. Esto implica la capacidad de utilizar mecanismos de búsqueda de activos de información, así como estructuras de navegación categorizada (directorios y listados).
- Permitir la identificación y en menor medida la localización directa, de los activos de información.
- Valoración de los recursos. Desde el punto de vista de la puesta a disposición de los recursos informativos, proporcionar al responsable de reutilizar la información la capacidad de valorar la pertinencia de un determinado activo es clave si se persigue hacerlo de manera “adecuada”.

- **Principio de simplicidad**

La simplicidad debe perseguirse tanto de cara al proceso de descripción de activos de información que tendrá lugar en el ámbito de los generadores de información publicable, como en el proceso de recuperación e interacción con las descripciones que deberá afrontar el ciudadano.

Desde el punto de vista de la descripción de la información es fundamental no establecer la obligatoriedad de campos descriptivos para los que frecuentemente no se posee una información completa. En el mismo sentido, el establecimiento de varios niveles descriptivos ayuda a la interoperabilidad con el resto de los datos cuando una descripción muy detallada no sea posible.

La Ficha de descripción de metadatos está compuesta por los siguientes tipos de metadatos:

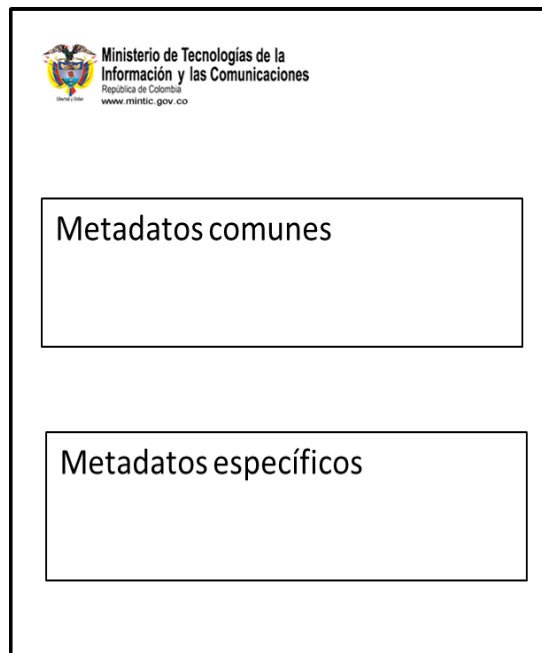
- **Metadatos comunes:** Aquellos que tienen todos los contenidos, independientemente de su naturaleza. Por ejemplo, todas los documentos,

eventos, informaciones, ayudas, etc. tienen asociado un nombre, una descripción, una tipología y en general una serie de metadatos comunes.

- **Metadatos específicos:** Aquellos propios del contenido en función de su naturaleza (tipología de contenido). Por ejemplo, las publicaciones tienen unos metadatos particulares como la fecha de la publicación del titular. Los procedimientos administrativos, a su vez, tienen unos metadatos particulares como el objeto, la vigencia o el lugar de presentación.

Para la estructuración de la Ficha Descripción de Metadatos se incorporan los objetivos de la categorización y la tipología de los metadatos, como se muestran en la siguiente figura:

**Figura 19: Estructura de la Ficha del Catálogo de Información de la entidad**



La información para la apertura se puede encontrar en los siguientes formatos<sup>6</sup>:

<sup>6</sup> <http://dadesobertes.gencat.cat/es/formats.html#anc02>

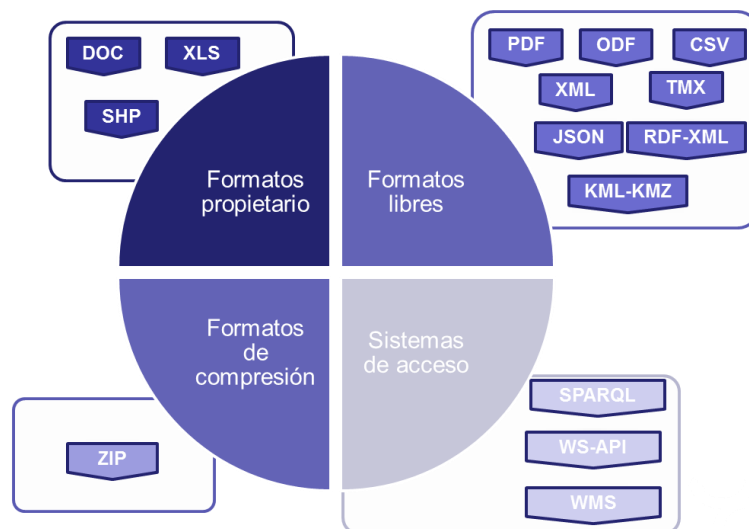
Figura 20: Formatos de apertura de datos



- **Formatos propietarios:** Son formatos de archivo que requieren herramientas para acceder, visualizar o editar la información que no son públicas.
- **Formatos libres:** Son formatos de archivo que se pueden crear y manipular para cualquier software, libre de restricciones legales.
- **Sistemas de acceso:** Lenguajes para consultar archivos.
- **Formatos de compresión:** Son formatos de almacenamiento que comprimen la información a través de algoritmos.

A continuación se presentan ejemplos de cada uno de los tipos de formatos para la publicación de la información:

Figura 21: Ejemplos de formatos de conjuntos de datos





La definición de cada uno de los formatos de la anterior gráfica se puede encontrar en el apartado: **Anexos**, Ver sección **para profundizar**.

## ¿Cuál es la Estructuración de los formatos de la Ficha de Descripción de Metadatos?

Los elementos que componen la Ficha de Descripción de Metadatos deben dar respuesta clara a los siguientes interrogantes:

- Quién: Descripción de los protagonistas objeto de los conjuntos de datos.
- Qué: La descripción del contenido de los conjuntos de datos.
- Cuándo: La descripción de las variables de tiempo asociadas a los conjuntos de datos.
- Dónde: Las referencias de localización de los datos.
- Porqué: La motivación para abrir dichos datos.
- Cómo: El procedimiento para la descarga de los datos.

A ellas habría que añadirle, en un contexto de la Administración Pública, otros elementos identificativos como:

- Referencias Institucionales: Referencias de autores o de las instituciones objeto o protagonistas de los datos.
- Organismos productores: Entidades que suministran los datos.
- Localización: El lugar y modo de acceso a los datos, links, FTP, entre otros.
- Accesibilidad: La durabilidad y conservación temporal de los datos.
- Referencias Normativas: Las referencias de tipo legal que puedan desprenderse o sobre las que estén basadas los datos.

A continuación se define la estructura de los metadatos de la ficha de Descripción de Metadatos:

**Tabla 2: Estructura de los Metadatos**

Tipología	Metadato	Descripción	Tipos de datos
Metadato común	Título	Nombre concreto del dataset	Texto
	Identificador	Identificador único del dataset en la base de datos	Texto (URI), numérico (ISBN, DOI, número local)
	Descripción	Descripción/resumen del dataset	Texto
	Categoría	Nombre(s) de categoría(s) o grupo(s) al cual pertenece el dataset	Texto/lista de texto
	Autor	Entidad y dependencia que origina el dataset	Texto
	Idioma	Idioma de los datos del dataset	Texto

Tipología	Metadato	Descripción	Tipos de datos	
	Audiencia	Audiencia objetivo del dataset	Texto	
	Palabras clave	Palabras clave para la ser identificados en la búsqueda	Texto separado por comas	
	Ámbito Geográfico	Identifica la catalogación de la información a nivel: Nacional, Departamental, Municipal, Distrital.	Texto/lista de texto	
Metadato específico	Fecha Publicación	Fecha de la carga del dataset en el catálogo de datos	Fecha (AAAA- MM_DD)	
	Fecha Última Actualización	Fecha de última actualización del dataset en el catálogo de datos	Fecha (AAAA- MM_DD)	
	Frecuencia Actualización	Frecuencia con la que se actualizan los datos del dataset en el catálogo de datos	Texto	
	Fecha Inicio	Fecha desde la cual se tienen en cuenta datos para el dataset	Fecha (AAAA- MM_DD)	
	Fecha Fin	Fecha hasta la cual se tienen en cuenta datos para el dataset	Fecha (AAAA- MM_DD)	
	Número de Versión	Número de la última versión del dataset en el catálogo de datos	Texto/numérico	
	URL Descarga	URL en el que el dataset puede ser descargado	Texto	
	Formato Descarga	Formatos en los que se encuentra disponible el dataset	Texto/lista de texto	
	Tamaño Descarga	Tamaño del dataset en Kb	Numérico	
	URL Esquema	URL de acceso al esquema, diccionario de datos o descripción de los datos y la estructura del dataset.	Texto	
	URL Documentación	URL a documentación adicional sobre el dataset	Texto	
	URL Programa	URL a sitio web de la entidad que describe los programas o aplicaciones (hechos o por hacer) relacionados con el dataset	Texto	
	Etiquetas	Palabras clave del contenido del dataset	Lista de texto	
	Licencia	Tipo de licencia	Texto	
	URL Licencia	URL de enlace a la descripción y detalles de la licencia	Texto	
	Nombre Responsable	Nombre de la persona que administra el dataset	Texto	
	Email Responsable	Email de la persona que administra el dataset	Texto	
	<b>Adicionar si es un inventario geográfico</b>			
	Cobertura de DataSet	Delimitador de coordenadas según la NTC 4611	Numérico	
	Metadatos adicionales	Describir metadatos adicionales según la NTC 4611	Texto	

## Instrumento de documentación de los conjuntos de datos

Se recomienda diligenciar la tabla que se presenta a continuación y que se encuentra en el archivo de Excel anexo, la cual contiene la información requerida para cada uno de los aspectos:

Figura 22: Ficha de la documentación de los conjuntos de datos

Ministerio de Tecnologías de la Información y las Comunicaciones República de Colombia www.mintic.gov.co		Documentación del Inventario de Datos		Gobierno en línea Colombia	
<b>Información de Metadatos Comunes</b>					
<b>Título:</b>					Texto
<b>Identificador:</b>					Texto (URI), numérico (ISBN, DOI, número local)
<b>Descripción:</b>					Texto
<b>Categoría:</b>					Texto/Lista de texto
<b>Autor:</b>					Texto
<b>Idioma:</b>					Texto/Lista de texto
<b>Audiencia:</b>					Texto/Lista de texto
<b>Palabras Clave</b>					Texto separado por comas
<b>Ámbito Geográfico:</b>					Texto/Lista de texto
<b>Información de Metadatos Específicos</b>					
<b>Fecha Publicación:</b>					Fecha (DD-MM-AAAA)
<b>Fecha Última Actualización:</b>					Fecha (DD-MM-AAAA)
<b>Frecuencia Actualización:</b>					Texto
<b>Fecha Inicio:</b>					Fecha (DD-MM-AAAA)
<b>Fecha Fin:</b>					Fecha (DD-MM-AAAA)

## ¿Para qué me sirve realizar una documentación de la ficha de los conjuntos de datos?

Realizar una documentación de los conjuntos de datos trae consigo los siguientes beneficios:

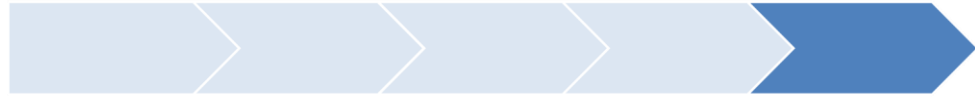
- Permite entender a cualquier lector el contenido de los datos.
- Permite a la entidad saber las características y la ubicación de sus datos.
- Identifica al conjunto de datos para que pueda destacarse más fácilmente en las búsquedas en el Catálogo de datos del Estado Colombiano.

### RECUERDE:

1. **Identifique** los conjuntos de datos documentar, a partir de la priorización realizada.
2. **Documente los metadatos Comunes y/o Específicos**
3. Diligencie la **Plantilla para la Descripción de los Metadatos** en el **formato de Excel** que acompaña la guía para la correcta documentación de los datos.

### ¡IMPORTANTE!

LA DOCUMENTACIÓN PUEDE ESTAR DESTINADA A LOS CIUDADANOS, POR LO CUAL DEBE SER CLARA Y DE FÁCIL ENTENDIMIENTO. UNA CORRECTA CATALOGACIÓN DE LA INFORMACIÓN BENEFICIA LAS BÚSQUEDAS DEL PORTAL DE DATOS ABIERTOS DE COLOMBIA.



## 7. ¿Cómo realizo la estructuración, cargue y publicación de los conjuntos de datos?

- ❖ El **objetivo** de este capítulo es definir los lineamientos para el cargue y publicación de los conjuntos de datos identificados en la etapa de priorización.
- ❖ El **resultado** esperado para esta etapa es la publicación de los primeros conjuntos de datos.
- ❖ **Perfiles requeridos:** Rol técnico, Rol Seguridad y Rol GEL.
- ❖ **Actividades a desarrollar:** Elección del mecanismo de publicación de los datos, publicación de los primeros conjuntos de datos en el Catálogo de Datos del Estado Colombiano.
- ❖ **Recursos a utilizar:** Portal del Catálogo de Datos del Estado Colombiano ([www.datos.gov.co](http://www.datos.gov.co))
- ❖ **Otras herramientas de apoyo o referencias:** N/A

A continuación se presentan los pasos que debe seguir la entidad si para publicar los primeros conjuntos de datos abiertos:

### ¿Qué debo tener en cuenta antes de la publicación?

Se ha definido un aviso legal que incorpora los principales aspectos legales a tener en cuenta al realizar la publicación de los datos. A continuación se presenta dicho aviso, que estará presente en el Catálogo de Datos Abiertos del Estado Colombiano ([www.datos.gov.co](http://www.datos.gov.co)) y que acompañará la publicación de los datos de la entidad:

#### Aviso legal

##### CONDICIONES GENERALES DE ACCESO Y USO

Este aviso legal contiene las **condiciones generales** que regulan el acceso y el uso del presente sitio web, del que es titular el Ministerio de Tecnologías de la Información y las Comunicaciones -MINTIC. Las entidades proveedoras de los datos publicados en este sitio web son titulares de los mismos.

La utilización de este sitio web implica la **aceptación** expresa y plena de las condiciones generales de acceso y uso de la versión publicada, en el instante en que el usuario acceda al mismo, sin perjuicio de las circunstancias particulares que pudieran aplicarse a algunos contenidos o servicios en concreto.

Para efectos del presente aviso legal, además de las definiciones legales y reglamentarias, se adoptan y aceptan las siguientes definiciones:

**Datos públicos:** Son todos aquellos datos primarios, sin procesar, en formatos estándar que facilitan su acceso y permiten su reutilización, los cuales están bajo la custodia de las entidades públicas y que pueden ser obtenidos y ofrecidos sin reserva alguna, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos.

**Información pública.** Es la agrupación ordenada de datos públicos, que permite otorgarle a los datos una utilidad y uso en determinado contexto, y que se genera a partir del desarrollo de actividades para el funcionamiento del Estado, es decir de los registros periódicos de las actividades misionales de las entidades, o como consecuencia del ejercicio de funciones de rutina en el Estado.

#### **Obligatoriedad de las condiciones generales**

Las presentes condiciones generales vinculan a cualquier persona natural o jurídica que haga uso y/o transformación de los datos sometidos a estas.

#### **Autorización de uso y/o transformación de los datos y cesión no exclusiva de derechos de propiedad intelectual**

Para efectos de las presentes condiciones, se entiende por uso y/o transformación autorizada de los datos, las actividades tales como: extracción, copia, difusión, modificación y adaptación de los datos.

Por dato, debe entenderse la descripción de hechos, situaciones, sucesos o valores, representados mediante símbolos físicos o electrónicos.

Por datos públicos, debe entenderse como todos aquellos datos primarios, sin procesar, en formatos estándar que facilitan su acceso y permiten su reutilización, los cuales están bajo la custodia de las entidades públicas y que pueden ser obtenidos y ofrecidos sin reserva alguna, de forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos.

Por información, debe entenderse la agrupación ordenada de datos, que permite otorgarle a los datos una utilidad y uso en determinado contexto.

Por información pública, debe entenderse la agrupación ordenada de datos, que permite otorgarle a los datos una utilidad y uso en determinado contexto, y que se genera a partir del desarrollo de actividades para el funcionamiento del Estado, es decir de los registros periódicos de las actividades misionales de las entidades o como consecuencia del ejercicio de funciones de rutina en el Estado.

De igual manera, se debe tener en cuenta la definición de mensaje de datos, contenida en el artículo 2º literal a) de la Ley 527 de 1999, que a su tenor prescribe lo siguiente: “(...) a) *Mensaje de datos. La información generada, enviada, recibida, almacenada o comunicada por medios electrónicos, ópticos o similares, como pudieran ser, entre otros, el Intercambio Electrónico de Datos (EDI), Internet, el correo electrónico, el telegrama, el télex o el telefax; (...)*”

Los datos y la información publicados en este sitio web son públicos y por esta razón se permite el uso y/o transformación autorizada de los mismos, por parte de todas las personas que accedan este sitio web, con el fin de que estos puedan hacer aplicaciones por parte de terceros y contenidos de su propia creación. Las aplicaciones y contenidos tendrán la protección de los derechos de autor y las condiciones de uso de su titular.

Sin perjuicio de la protección contenida de las normas de derechos de autor de que gozan las aplicaciones desarrolladas por terceros y de las condiciones de uso que ellas dispongan, los datos siempre serán de acceso al público y de propiedad del Estado.

### **Condiciones de uso y/o transformación de los datos y la información**

Son de aplicación las siguientes condiciones generales para el uso y/o transformación de los datos sometidos a dichas condiciones:

- Se debe tener en cuenta que tanto el diseño como el contenido de esta página web están protegidos por las normas colombianas e internacionales sobre el derecho de autor, en cabeza del Ministerio de Tecnologías de la Información y las Comunicaciones –MINTIC y/o de los respectivos autores de dichos datos, según el caso.
- Está prohibido desnaturalizar el sentido de los datos.
- Debe citarse la fuente de los datos objeto del uso y/o transformación.
- Debe mencionarse la fecha de la última actualización de los datos objeto del uso y/o transformación, siempre y cuando estuviera incluida en el dato original.
- Deben conservarse, no alterarse ni suprimirse los metadatos sobre la fecha de actualización y las condiciones de uso y/o transformación aplicables.

### **Exclusión de responsabilidad**

La utilización de los conjuntos de datos se realizará por parte de los usuarios o agentes del uso y/o transformación, bajo su propia cuenta y riesgo, correspondiéndoles en exclusiva a ellos responder frente a terceros por daños o perjuicios de cualquier índole que pudieran derivarse de dicha utilización.

La entidad productora de los datos publicados no será responsable de la utilización que de sus datos hagan las personas que transformen y/o usen dichos datos, ni tampoco de los daños sufridos o pérdidas económicas que, de forma directa o indirecta, produzcan o puedan producir perjuicios económicos, materiales o sobre datos y aplicaciones, provocados por el uso y/o transformación de los datos.

### **Protección de datos de carácter personal**

El Ministerio de Tecnologías de la Información y las Comunicaciones -MINTIC- es titular de la base de datos de contacto generada con los datos de carácter personal suministrados por los usuarios de este sitio web.

El usuario de este sitio web autoriza el tratamiento automatizado de los datos personales suministrados, necesarios para la prestación de los servicios ofrecidos por el sitio web, respetando siempre la legislación colombiana sobre protección de los datos, en especial el artículo 15 de la Constitución Política de Colombia, relativo “... *al derecho a la intimidad, al buen nombre, al derecho que tienen todas las personas a actualizar y rectificar las informaciones que se hayan recogido sobre ellas, en bancos de datos y en archivos de entidades públicas y privadas...*”; así como también, con la normatividad colombiana que resulta aplicable a la reserva y protección de datos personales. Los datos personales recibidos por el Ministerio de Tecnologías de la Información y las Comunicaciones – MINTIC- no podrán ser entregados a terceros, con fines comerciales privados, a título de venta, cesión, alquiler a terceros o cualquier otra forma.

Los datos personales de los usuarios recibidos por el Ministerio de Tecnologías de la Información y las Comunicaciones –MINTIC son almacenados en bases de datos de propiedad de El Ministerio de Tecnologías de la Información y las Comunicaciones – MINTIC-, quien asume las medidas de índole técnica, organizativa y de seguridad que garanticen la confidencialidad e integridad de los datos de acuerdo con lo establecido en la Ley 527 de agosto 18 de 1999 por la cual, “*se define y reglamenta el acceso y uso de los mensajes de datos, del comercio electrónico y de las firmas digitales, y se establecen las entidades de certificación y se dictan otras disposiciones*”, y demás legislación aplicable.

El usuario del sitio web responderá, en cualquier caso, por la veracidad de los datos facilitados, reservándose el Ministerio de Tecnologías de la Información y las Comunicaciones –MINTIC- el derecho a excluir de los servicios registrados a todo usuario que haya facilitado datos falsos, sin perjuicio de las demás acciones que procedan en Derecho.

El usuario del sitio web podrá ejercer en cualquier momento los **derechos** de acceso, rectificación, cancelación u oposición con respecto a los datos por él suministrados, a través del correo electrónico que se establezca en el sitio web, o mediante los mecanismos que para tal efecto disponga la entidad.

En cualquier momento el usuario puede dar de baja los datos por él suministrados, de acuerdo a los mecanismos que se establezcan por la entidad.

### **Modificaciones**

El Ministerio de Tecnologías de la Información y las Comunicaciones -MINTIC-, se reserva el derecho a modificar en cualquier momento las presentes condiciones de uso, así como cualquier otra condición particular.

### **Hiperenlaces**

Los hiperenlaces contenidos en el sitio web pueden dirigir a páginas web de terceros o sitios vinculados. El Ministerio de Tecnologías de la Información y las Comunicaciones – MINTIC-, no asume ninguna responsabilidad por el contenido, información o servicios que pudieran aparecer en dichos sitios, que tendrán exclusivamente carácter informativo y que en ningún caso implican relación alguna entre el Ministerio de Tecnologías de la Información y las Comunicaciones -MINTIC- y las personas o entidades titulares de los contenidos o de los sitios web donde tal información se encuentre.

Así mismo, el Ministerio de Tecnologías de la Información y las Comunicaciones -MINTIC-, no es responsable de los cambios o actualizaciones de los sitios vinculados, ya que, la inclusión de éstos, no implica su aprobación, por parte del Ministerio de Tecnologías de la Información y las Comunicaciones –MINTIC-, ni asociación con sus operadores.

### **Marcas**

La denominación Ministerio de Tecnologías de la Información y las Comunicaciones, su logotipo y el nombre de algunos de los productos o servicios ofrecidos por el mismo, son marcas registradas por el Ministerio de Tecnologías de la Información y las Comunicaciones –MINTIC-, las cuales se encuentran protegidas por las normas sobre





propiedad industrial y sólo pueden ser explotadas o aprovechadas por dicho Ministerio. Los comentarios y opiniones publicados en esta página pertenecen a sus respectivos autores y, por ende, no comprometen necesariamente a la Entidad, a menos que así se indique expresamente.

## Estructuración de los datos a cargar

La actividad de estructuración de los datos consiste en la preparación del archivo con el conjunto de datos a cargar en un formato estructurado. Para realizar dicha estructuración la entidad deberá tener en cuenta las siguientes actividades:

- 1. Elegir el formato para la publicación de los datos:** La entidad tendrá la autonomía para elegir el formato más adecuado para la publicación de sus datos. Lo anterior bajo criterios de menor esfuerzo y costo pero sin afectar la calidad de la información.  
Los formatos permitidos en la iniciativa de datos abiertos para publicar los datos son:  
XLS, ODF, CSV, XML, GEL-XML, JSON, TXT, RDF-XML, KML-KMZ, SHP, TMX, ZIP.
- 2. Diligenciamiento del encabezado del archivo a cargar:**
  - Para cada conjunto de datos, además de la especificación del Encabezado, se deberá diligenciar la Ficha de Descripción de Metadatos de Metadatos para su categorización y perfilamiento.
  - El archivo de datos debe ser un archivo estructurado con el mismo orden definido en la pestaña Encabezado del Dataset. Los campos deben ir separados por comas y el contenido delimitado con comillas dobles: Ej. "VALOR\_1", "VALOR\_2", "VALOR\_3", "VALOR\_4".
  - Los valores de los campos deberán cumplir el estándar unicode UTF-16. \*Evitar la inclusión de caracteres especiales.
  - Los valores de los campos de tipo fecha se deberán diligenciar con el formato DD-MM-AAAA. Si incluye hora, ésta deberá ser HH:mm:ss.
  - Los valores de campos numéricos con fracciones deberán diligenciarse con . (punto) como separador de decimales.

**Figura 23: Ficha de Estructuración**

 Ministerio de Tecnologías de la Información y las Comunicaciones República de Colombia www.mtisc.gov.co	Plantilla para la Descripción de los Encabezados				
Consideraciones para el diligenciamiento del archivo de datos del Dataset a cargar					
Para cada dataset, además de la especificación del Encabezado, se deberá diligenciar la ficha/pestaña de Metadatos para su categorización y perfilamiento.					
El archivo de datos debe ser un archivo CSV con la misma estructura y orden definidos en esta pestaña Encabezado del Dataset. Los campos deben ir separados por comas y el contenido delimitado con comillas dobles: Ej. "VALOR_1","VALOR_2","VALOR_3","VALOR_4".					
Los valores de los campos deberán cumplir el estándar unicode UTF-16. <i>Evitar la inclusión de caracteres especiales.</i>					
Los valores de los campos de tipo fecha se deberán diligenciar con el formato DD-MM-AAAA. Si incluye hora, ésta deberá ser HH:mm:ss					
Los valores de campos numéricos con fracciones deberán diligenciarse con . (punto) como separador de decimales					
ENCABEZADO DEL DATASET					
Orden del campo	Nombre del campo	Tipo de dato	Longitud	Acepta Nulo	Descripción del campo
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					

**Nota:** Si dentro de la estructuración de los campos a cargar no se encuentra alguno relacionado con el lenguaje de intercambio GEL-XML, se deberá seguir con el proceso de carga de la información e iniciar el procedimiento de incorporación de campos en el lenguaje descritos en el documento LENGUAJE COMÚN DE INTERCAMBIO DE INFORMACIÓN ARQUITECTURA DE DATOS<sup>7</sup>.

3. **Generación de archivo estructurado:** Luego del diligenciamiento del encabezado de los datos, la entidad identifica y ejecuta el mecanismo más adecuado para la generación del archivo con los datos estructurados. A continuación se plantean unos posibles escenarios para esta actividad:
  - a. *Utilizar el administrador de la base de datos:* Los datos almacenados en una base de datos son consultados a través de una funcionalidad de búsqueda (o consultas) y la generación de reportes bajo la estructura definida en el numeral 2.
  - b. *Programación de un script para la extracción de datos:* Consiste en el desarrollo de un programa para la extracción de la información requerida, de acuerdo a los requerimientos de estructura definidos en el numeral 2. La entidad podrá utilizar los frameworks y lenguajes de programación que considere (teniendo en cuenta la plataforma tecnológica que tenga)
  - c. *Exportar a partir de un sistema de información:* generación de reportes estructurados sobre algún sistema de información existente en la entidad, configurable bajo la estructura del numeral 2.

<sup>7</sup> Portal de lenguaje GEL-XML (<http://lenguaje.intranet.gov.co>), en el documento de arquitectura (GEL-DR-MA-003\_Arquitectura\_Datos\_V5.0.pdf)

## ¿Cuáles son los mecanismos de publicación de los conjuntos de datos?

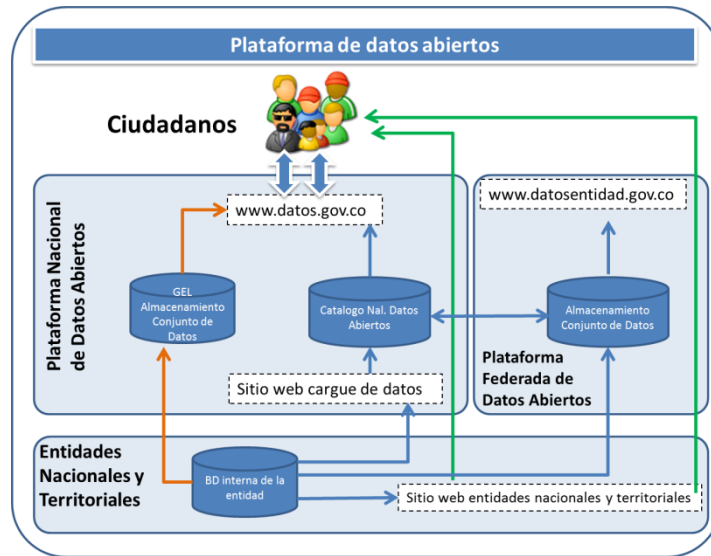
A continuación se detallan los diferentes mecanismos de publicación de los datos abiertos que estarán disponibles para las entidades de acuerdo con la madurez tecnológica de la misma y las características de los primeros datos a publicar. Para la publicación de los datos se han definido cuatro escenarios de publicación, y cada entidad, con el apoyo del Departamento de Sistemas, definirá el mejor mecanismo para la publicación de los datos.

**Figura 24: Escenarios de publicación de datos**



La herramienta principal para la oferta de datos son los catálogos de datos, los cuales consisten en uno o varios portales Web que permiten a los ciudadanos y empresas acceder a los datos abiertos (entre otras funcionalidades). En Colombia se contará con un Catálogo de Datos Abiertos del Estado Colombiano ([www.datos.gov.co](http://www.datos.gov.co)) que centralizará el registro y búsqueda de todos los datos abiertos de las entidades públicas del país; sin embargo éste no restringirá la creación de otros catálogos temáticos en el país que puedan ser desarrollados por distintos sectores o entidades del orden territorial, tal y como lo refleja la siguiente gráfica:

**Figura 25: Plataforma de arquitectura de datos abiertos**



Como se puede observar, [www.datos.gov.co](http://www.datos.gov.co) será el portal Web que centralice la búsqueda y acceso a los datos de nivel nacional y contará con un catálogo de datos nacional que mantendrá todas las referencias a los datos abiertos del país. Las plataformas federadas de datos abiertos consultarán el catálogo de datos abiertos nacional y mantendrán ahí sus referencias, sin importar que las URL de acceso sean distintas y que puedan almacenar conjuntos de datos distintos. Las entidades nacionales y territoriales podrán elegir entre distintos medios de publicación de sus datos abiertos y podrán elegir la plataforma nacional o las federadas para la publicación de los mismos.

A continuación se detallan los cuatro escenarios:

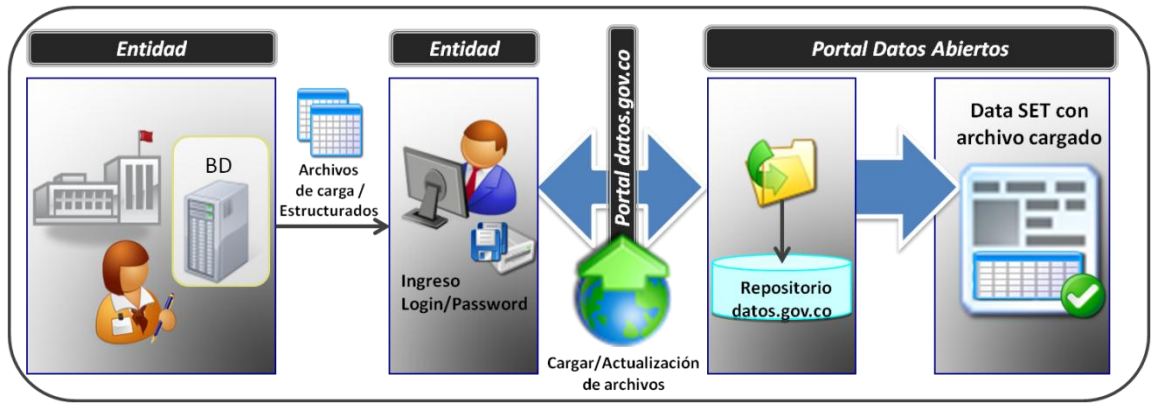


**Descripción:** La entidad carga el set de datos al Catálogo de Datos ([www.datos.gov.co](http://www.datos.gov.co)) y realiza la documentación del mismo manualmente. Se crea una copia del set de datos en [www.datos.gov.co](http://www.datos.gov.co).

**Criterios de utilización:** Información con frecuencia de actualización anual o semestral. La entidad no posee la tecnología para asumir la responsabilidad del alojamiento del set de datos (o no desea asumir la responsabilidad).

**Criterios tecnológicos:** La entidad podrá utilizar este mecanismo cuando considere que no cuenta con un servidor Web que soporte y aloje los datos a cargar. Así mismo la entidad debe contar como mínimo con conexión a internet que permita el acceso a [www.datos.gov.co](http://www.datos.gov.co) o al portal de datos abiertos que la entidad desee.

Figura 26: Escenario - Carga de archivo de set de datos

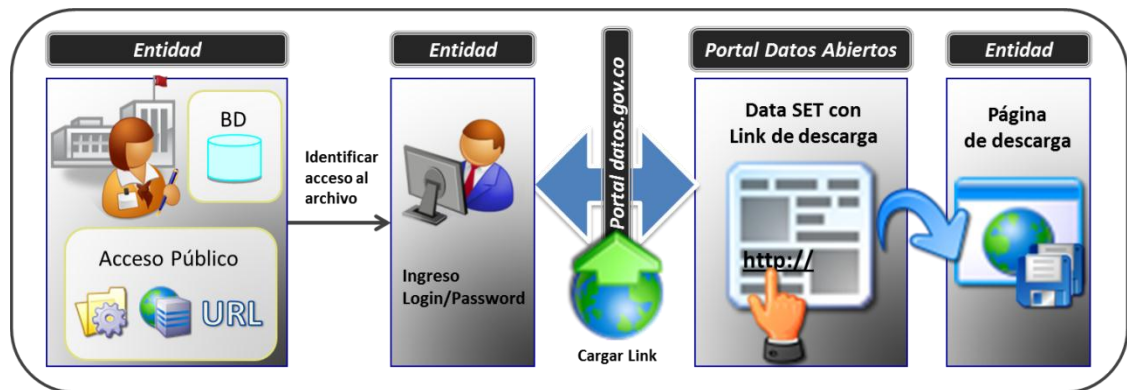


**Descripción:** La entidad carga el set de datos en una URL pública, manual o automáticamente. Se crea una referencia a la URL en el Catálogo de Datos ([www.datos.gov.co](http://www.datos.gov.co)).

**Criterios de utilización:** Información con frecuencia de actualización anual o semestral. Si la carga es automática puede tener una mayor frecuencia. La entidad asume la responsabilidad del alojamiento y la actualización del set de datos.

**Criterios tecnológicos:** La entidad puede utilizar este mecanismo cuando cuenta con un servidor Web que soporte y aloje los datos a cargar. En este caso la entidad debe contar como mínimo con conexión a internet que permita el acceso a [www.datos.gov.co](http://www.datos.gov.co) o al portal de datos abiertos que la entidad desee y el servidor Web mencionado con salida a internet.

Figura 27: Escenario – Referencia de archivo de conjuntos de datos



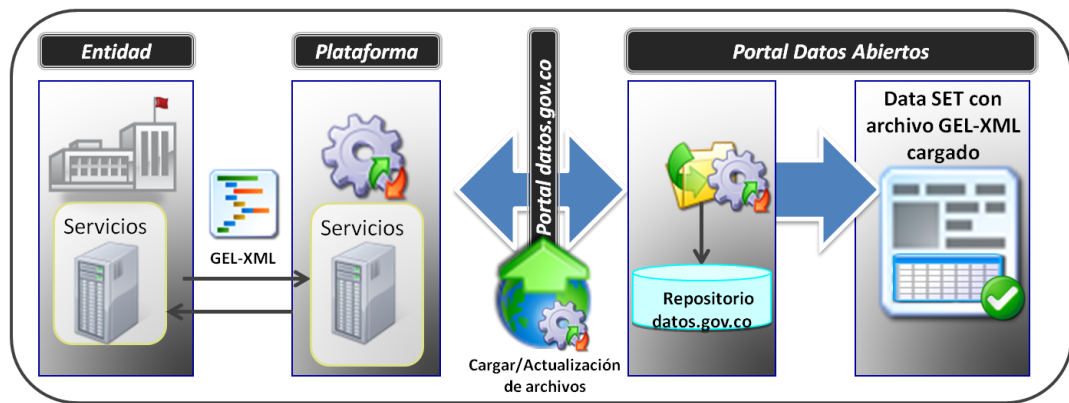
Transferencia de datos

**Descripción:** Se realiza una transferencia automática del archivo de set de datos a un repositorio central para su publicación.

**Criterios de utilización:** Información con frecuencia de actualización semanal o diaria. Se cuenta con el conocimiento y la tecnología para realizar desarrollos a medida que permitan la transferencia de los mismos.

**Criterios tecnológicos:** La entidad puede utilizar este mecanismo cuando cuenta con el conocimiento para realizar desarrollo de software que permita la transferencia automática de los conjuntos de datos. En este caso la entidad debe contar como mínimo con conexión a internet que permita el acceso a [www.datos.gov.co](http://www.datos.gov.co) o al portal de datos abiertos que la entidad desee, y además debe disponer de las herramientas automáticas para la generación de los archivos de los conjuntos de datos y la transferencia automática de los mismos.

**Figura 28: Escenario – Transferencia de datos**



Servicio de información en línea

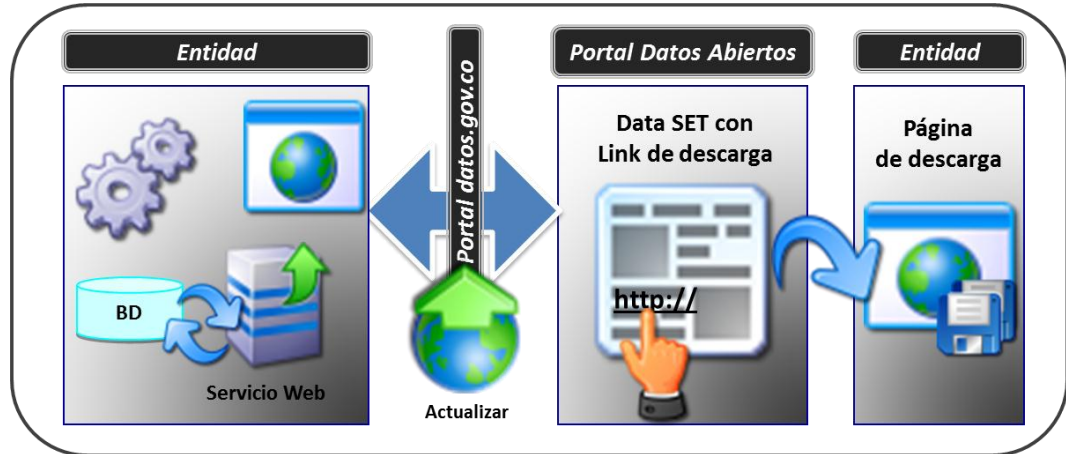
**Descripción:** La entidad realiza el desarrollo de un servicio de consulta de la información en línea y referencia el mismo por medio de una URL en el Catálogo de Datos ([www.datos.gov.co](http://www.datos.gov.co)).

**Criterios de utilización:** Información con frecuencia de actualización diaria o menor. La entidad asume la responsabilidad de la disponibilidad del servicio publicado.

**Criterios tecnológicos:** La entidad puede utilizar este mecanismo cuando considera que la transferencia de archivos puede ser muy lenta y además cuenta con un servidor Web que garantice la disponibilidad del servicio publicado. En este caso la entidad debe contar como mínimo con conexión a internet que permita el acceso a

www.datos.gov.co o al portal de datos abiertos que la entidad desee y el servidor Web mencionado con salida a internet.

Figura 29: Escenario – Servicio de información en línea



**RECUERDE:**

1. **Escoja** junto al Departamento de Sistema el mejor **mecanismo de publicación** de sus primeros conjuntos de datos.
2. Publique sus primeros datos en [www.datos.gov.co](http://www.datos.gov.co).
3. Luego de los primeros datos publicados, puede realizar el mismo procedimiento señalado en esta guía para la publicación de todos aquellos datos que son susceptibles de abrir.

**¡IMPORTANTE!**

LA PUBLICACIÓN DE LOS PRIMEROS DATOS DE LA ENTIDAD ES EL RESULTADO DEL MODELO INCREMENTAL DE DATOS ABIERTOS EN SU FASE INICIAL. SIN EMBARGO, ESTA GUÍA SERVIRÁ EN ADELANTE PARA PUBLICAR TODOS LOS DATOS QUE DESEE UNA ENTIDAD.

## 9. REFERENCIAS

- Apertura de datos públicos (Open Data): Una oportunidad para las AAPP y para la sociedad. Oesía Tecnología Sinfónica, Abril 2011
- Aporta. (Junio de 2011). *Estudio de caracterización del sector infomediario*. (Ministerios PTyAD, MITYC) Consultado el 8 de Septiembre de 2011, de <http://www.aporta.es/web/guest/estudioRISP2011>
- Australia, G. (s.f.). *australia.gov.au*. Consultado el 8 de Septiembre de 2011, de <http://australia.gov.au>
- Berners-Lee, T. (2009). *Raw Data Now!*(TED) Consultado el 8 de Septiembre de 2011, de [http://www.youtube.com/watch?v=OM6XIIcm\\_qo](http://www.youtube.com/watch?v=OM6XIIcm_qo)
- Biurrun, F. (25 de Noviembre de 2010). *Un monopolio estatal más, la distribución de sentencias*. (LegalToday.com) Consultado el 16 de Septiembre de 2011, de <http://www.legaltoday.com/opinion/articulos-de-opinion/un-monopolio-estatal-mas-la-distribucion-de-sentencias>
- Lurismatica, *Reutilización de la información pública y difusión del patrimonio cultural gestionado por la diputación foral de gipuzkoa*, Gipuzkoa 2.0
- Ministerio de Relaciones exteriores de Uruguay, Presentación: Acceso a la información pública.pdf, Dra. Laura Brunet 2009.
- Obama, B. (2009). *Transparency and Open Government*. (La Casa Blanca (EEUU)) Consultado el 8 de Septiembre de 2011, de [http://www.whitehouse.gov/the\\_press\\_office/TransparencyandOpenGovernment/](http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment/)
- Observatorio-Legislativo. (Noviembre de 2008). *Derecho a la Información*. Consultado el 2011 de Septiembre de 2011, de Boletín n° 113: [http://www.icpcolombia.org/archivos/observatorio/boletin\\_113](http://www.icpcolombia.org/archivos/observatorio/boletin_113)
- OGWG. (8 de Diciembre de 2007). *8 Principles of Open Government Data*. (Open Government Working Group) Consultado el 8 de Septiembre de 2011, de <http://www.opengovdata.org/home/8principles>
- Pérez, E. M. (s.f.). *Medio Ambiente & Derecho, Revista electrónica de derecho ambiental*. Consultado el 8 de Septiembre de 2011, de EL DERECHO DE ACCESO A LA INFORMACIÓN EN EUROPA Y AMÉRICA LATINA: UN ENFOQUE CONSTITUCIONAL: <http://huespedes.cica.es/aliens/gimadus/10/DERECHOACCESO.htm>
- Portal de Datos Abiertos del Gobierno Vasco, Opendata Euskadi: <http://www.irekia.euskadi.net/>  
<http://opendata.euskadi.net/>



- Proyecto de Apertura de Datos Públicos del Ayuntamiento de Barcelona “OpenDataBCN”, Barcelona, 22 de noviembre de 2010, Ayuntamiento de Barcelona.
- Shadbolt, N. (21 de Enero de 2011). *A year of data.gov.uk*. Consultado el 21 de Septiembre de 2011, de The Guardian:  
<http://www.guardian.co.uk/news/datablog/2011/jan/21/data-gov-nigel-shadbolt-government>
- Shane, D. (14 de Julio de 2010). *Private sector reluctant to join open data movement*. (Information Age) Consultado el 16 de Septiembre de 2011
- Staromiejski, E. (Octubre de 2009). *SIGeDA, a case study for approaching eGRIMS impact on Private Industry*. (IFIP, IWEI (Enterprise Interoperability). Materiales bajo petición: rsanchis@cigip.upv.es) Obtenido de eGRIMS (eGovernment Records and Information Management Systems):  
<http://www.remplanet.eu/web/consortium/workspaces/wp7/t7.4/events/2nd-ifip-wg5.8-workshop-on-enterprise-interoperability-iwei-2009>
- Staromiejski, E. (10 de Enero de 2011). *El papel de la gestión de la documentación y archivo en la Reutilización de la Información del Sector Público*. (EAPC (Escuela de Administración Pública de Catalunya) y ESAGED (Escuela Superior de Archivística, Universidad Autónoma de Barcelona)). Consultado el 16 de Septiembre de 2011, de <http://www20.gencat.cat/portal/site/eapc/menuitem.ca54cfbb17b4abf5272a63a7b0coe1a0/?vgnnextoid=04ef4785df29d210VgnVCM2000009b0c1e0aRCRD&vgnnextchannel=04ef4785df29d210VgnVCM2000009b0c1e0aRCRD&vgnnextfmt=default>
- Staromiejski, E. (14 de Septiembre de 2011). *Ponencia: "Más datos y menos metadatos (o de la triste historia de la no estructuración de la información) "*. (SOCINFO, Sociedad de la Información). Consultado el 16 de Septiembre de 2011, de <http://www.socinfo.es/art-socinfo/2150-socinfo-celebro-qadministracion-electronica-en-cataluna-ivq>:  
<http://www.socinfo.es/art-socinfo/2150-socinfo-celebro-qadministracion-electronica-en-cataluna-ivq>
- Trevijano, E. (Diciembre de 2003). *Constitución española: sinopsis artículo 105*. (Congreso del Estado español). Consultado el 8 de Septiembre de 2011, de <http://www.congreso.es/consti/constitucion/indice/sinopsis/sinopsis.jsp?art=105&tipo=2>
- USA, G. (s.f.). *Office of E-Government & Information Technology*. (La Casa Blanca) Consultado el 8 de Septiembre de 2011, de <http://www.whitehouse.gov/omb/e-gov/>

- Wikipedia. (1999). *Ley Organica de Proteccion de Datos*. Consultado el 8 de Septiembre de 2011, de [http://es.wikipedia.org/wiki/Ley\\_Org%C3%A1nica\\_de\\_Protecci%C3%B3n\\_de\\_Datos\\_de\\_Car%C3%A1cter\\_Personal\\_de\\_Espa%C3%B1a](http://es.wikipedia.org/wiki/Ley_Org%C3%A1nica_de_Protecci%C3%B3n_de_Datos_de_Car%C3%A1cter_Personal_de_Espa%C3%B1a)
- Wikipedia. (s.f.). *Transformational Government*. (Wikipedia) Consultado el 2011 de Septiembre de 2011
- World Wide Web W3C: Publishing Open Government Data
- <http://www.w3.org/TR/gov-data/>

## Para profundizar

### ZIP

Formato de almacenamiento sin pérdida que se puede utilizar tanto con programas propiedad como libres. Los archivos tienen formato .zip

Para mayor información consulta: [www.pkware.com/support/zip-app-note/](http://www.pkware.com/support/zip-app-note/)

**Formatos libres**<sup>8</sup> Son formatos de archivo que se pueden crear y manipular para cualquier software, libre de restricciones legales.

### PDF (Formato de Documento Portátil):

Documento de formato portable universal que mantiene la apariencia del documento independientemente del sistema operativo que se utilice (multiplataforma). Incluye cualquier combinación de texto, multimedia e hipertexto y además se puede cifrar el contenido y firmarlo digitalmente. Es el estándar ISO, desde 2008, para ficheros contenedores de documentos electrónicos de cara a la conservación de larga duración. Es una especificación que se puede crear, visualizar o modificar con herramientas de software libre. Este formato originalmente era propietario (hasta 2008)

Para mayor información consulta: [www.adobe.com/es/products/acrobat/adobepdf.html](http://www.adobe.com/es/products/acrobat/adobepdf.html)

### ODF (Formato de Documento Abierto para Aplicaciones Ofimáticas OASIS):

Open Document es un formato de fichero estándar para el almacenamiento de contenidos ofimáticos como hojas de cálculo, gráficos y presentaciones. Es el estándar para el intercambio de texto con formato. Las extensiones son:

- Texto: .odt
- Hoja de cálculo: .ods
- Dibujo: .odg
- Gráfico: .odc
- Fórmula matemática: .odf

---

<sup>8</sup> <http://dadesobertes.gencat.cat/es/formats.html#anc02>

- Base de datos: .odb
- Imagen: .odi
- Documento maestro: .odm

Para mayor información consulta: [www.odfalliance.org/](http://www.odfalliance.org/)

### **CSV (Valores separados por coma)**

Tipos de documentos en formato abierto sencillo para representar datos en formato de tabla, en columnas separadas por comas (o punto y coma, donde la coma es el separador decimal) y las filas son saltos de línea. Los campos que tienen una coma, un salto de línea o una comilla doble deben cerrarse entre comillas dobles. No indica un juego concreto de caracteres, ni cómo van situados los bytes, ni el formato para el salto de línea. Las extensiones que se utilizan son .csv y .txt.

Para mayor información consulta: [www.tools.ietf.org/html/rfc4180](http://www.tools.ietf.org/html/rfc4180)

### **XML (Lenguaje de Etiquetado Extensible)**

Es un metalenguaje simple pero estricto, desarrollado por W3C. Desarrolla un papel fundamental en el intercambio de una gran variedad de datos. XML es un formato que permite la interpretación de datos a través de varias aplicaciones. Es una simplificación y adaptación del SGML y permite definir la gramática de lenguajes específicos. En realidad, XML es una manera de definir lenguajes para diferentes necesidades.

Para mayor información consulta: [www.w3.org/standards/xml/core](http://www.w3.org/standards/xml/core)

### **TMX (Translation Memory eXchange)**

Estándar de XML, que es un DTD que sirve para el intercambio de memorias de traducción. Creado por el comité OSCAR (Open Standards for Container/Content Allowing Re-use).

Para mayor información consulta:  
[www.w3.org/2002/02/01-i18n-workshop/LocFormats#TMX](http://www.w3.org/2002/02/01-i18n-workshop/LocFormats#TMX)  
[www.lisa.org/Translation-Memory-eXchange-TMX.34.0.html](http://www.lisa.org/Translation-Memory-eXchange-TMX.34.0.html)

### **JSON (Notación de Objetos de JavaScript)**

Formato ligero de intercambio de datos, fácil de entender para las personas, y ofrece sencillez a las máquinas en la generación e interpretación. Está basado en un subconjunto del lenguaje de programación JavaScript, adecuado para la programación por parte del cliente.

Para mayor información consulta: [www.json.org/json-es.html](http://www.json.org/json-es.html)

### **RDF/XML (Infraestructura para Descripción de Recursos)**

Modelo para la representación de los recursos de la web en expresiones con la forma sujeto-predicado-objeto. El sujeto es el recurso que se describe, el predicado es la propiedad sobre la que se quiere establecer el recurso y el objeto es el valor de la propiedad con el que se establece la relación. La combinación de RDF con otras herramientas permite añadir significado a las páginas y es una de las tecnologías esenciales para la web semántica. Para que sea interpretable, se representa en formato XML.

Para mayor información consulta: [www.w3.org/standards/techs/rdf#w3c\\_all](http://www.w3.org/standards/techs/rdf#w3c_all)

### **KML-KMZ (Lenguaje de Etiquetado de Ojo de Cerradura)**

Gramática XML y formato de archivo para la creación de modelos y almacenamiento de funciones geográficas como puntos, líneas, imágenes, polígonos y modelos que se mostrarán principalmente en aplicaciones de mapas. Se utiliza para compartir lugares e información entre aplicaciones. Es el estándar del Open Geospatial Consortium y se puede utilizar a través de Google Earth. Los archivos KML se distribuyen comprimidos como archivos KMZ.

Para mayor información consulta:  
[www.code.google.com/intl/es/apis/kml/documentation/kmlreference.html](http://www.code.google.com/intl/es/apis/kml/documentation/kmlreference.html)  
[www.opengeospatial.org/standards/kml/](http://www.opengeospatial.org/standards/kml/)

**Formatos propietario**<sup>9</sup> Son formatos de archivo que requieren herramientas que no son públicas.

#### **DOC (Microsoft Office Word):**

Formato cerrado para transferir textos formateados o sin formato. Puede contener textos, imágenes, gráficos y enlaces. Desde la versión 2007 se trabaja con un nuevo formato, docx, que es más avanzado y comprime más el documento.

Para mayor información consulta: [www.office.microsoft.com/es-es/word](http://www.office.microsoft.com/es-es/word)

#### **XLS (Microsoft Office Excel):**

Contiene filas y columnas de celdas y cada una puede incluir datos, que pueden ser palabras, números o fórmulas que tengan datos y resolver ecuaciones dinámicamente. Las hojas de cálculo XLS también pueden contener tablas y gráficos que muestran todas las secciones o de los datos seleccionados.

Para mayor información consulta: [www.office.microsoft.com/es-es/excel/](http://www.office.microsoft.com/es-es/excel/)

#### **SHP (ESRI):**

Shapefile es un formato propietario de datos espaciales que es el estándar para el intercambio de información geográfica entre Sistemas de Información Geográfica (GIS). Es un formato vectorial de almacenamiento digital donde se guarda la localización de elementos geográficos y los atributos asociados a ellos, pero sin capacidad para almacenar información topológica. Lo generan varios archivos, mínimo tres y tiene tres tipos de extensiones: .shp, .shx y .dbf

Para mayor información consulta: [www.esri.com/products/index.html#data\\_panel](http://www.esri.com/products/index.html#data_panel)

#### **Infomediario/a**

Término con el que se hace referencia a una empresa que reutiliza la información de las Administraciones para producir un servicio o una obra con valor añadido. Suele utilizarse en referencia a la Industria Infomediaria. Existen múltiples ejemplos de industrias infomediarias, como aquellas que editan bases de datos de legislación y jurisprudencia a

---

<sup>9</sup> <http://dadesobertes.gencat.cat/es/formats.html#anc02>

partir de las sentencias y normas publicadas en los boletines oficiales, o las que trabajan con información de ámbito Meteorológico para realizar informes más específicos para las Administraciones con el fin de, por ejemplo, detectar riesgos de incendios forestales a través de datos de previsiones e históricos de situación meteorológica en un determinado territorio (por poner solo un par de casos).

### **Injerto**

Inclusión de datos dentro de un subconjunto de datos. Se trata de un subproceso de "agregación de datos".

### **Licencia de información (licencia tipo)**

Contrato mediante el cual una persona recibe de otra el derecho de uso de varios de sus bienes, normalmente de carácter no tangible o intelectual, a cambio del pago o no de una cantidad determinada por el uso de los mismos. Este tipo de contratos, con carácter general, recogen una serie de cláusulas que atienden a cuatro aspectos específicos: derechos sobre la obra, limitaciones y prohibiciones, garantías y responsabilidades, jurisdicción competente y derecho aplicable.

La **licencia de información** es la modalidad que tienen los organismos del Sector Público para autorizar la publicación de documentos públicos de forma no exclusiva y no discriminatoria respecto a otros usuarios.

### **Mediación**

Expresión general utilizada para referirse a cualquier tratamiento de los datos o de la información como paso previo a su apertura. Casos concretos de mediación son "el podado", "un injerto" o "la anonimización".

### **Microformato**

(A veces abreviado con  $\mu$ F o uF) Forma simple de agregar significado semántico a un contenido legible por un humano mientras que para una máquina puede ser sólo texto plano. Según el W3C, son conjuntos de formatos de datos abiertos y simples, desarrollados sobre estándares ya existentes, ampliamente adoptados, incluyendo XHTML (HTML y XML) y CSS.

### **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.

### **Open Government (Gobierno Abierto)**

La mayoría de los proyectos de Gobierno Abierto recientes se ajustan sobre todo a la definición contenida en la Iniciativa Gobierno Abierto (Open Government Initiative) publicada por Barack Obama en el primer día de su mandato (20 de enero de 2009). En esta iniciativa, Obama declara proponerse "la creación de niveles de apertura en el Gobierno sin precedentes. La filosofía de esta iniciativa se basa en la transparencia, la participación y la colaboración". Así reza la definición entrada en Wikipedia:

"El Gobierno abierto es la [doctrina política](#) que sostiene que los temas de [Gobierno](#) y [administración pública](#) deben ser abiertos a todos los niveles posibles en

cuanto a transparencia, esto debe ir unido a la creación de espacios permanentes de [participación ciudadana](#) y colaboración ciudadana”.

#### **Podado** (de datos)

Eliminación de un subconjunto de datos originalmente incluidos en un documento electrónico estructurado (los datos de un elemento complejo dentro de un documento XML, por ejemplo).

#### **RISP** (Acrónimo de “Reutilización de la Información del Sector Público”)

Se trata de un esfuerzo Jurídico y de las Administraciones para dar acceso a los documentos, información y datos de carácter público (es decir no sujetos a ninguna restricción legal) a ciudadanos y empresas para que éstos puedan usarlas como mejor les convenga, pero muy especialmente con la intención de que generen nuevos servicios y productos de valor añadido. Esta iniciativa se impulsó sobre todo a partir del año 2003 con una Directiva europea que todos los Estados Miembros de la Unión están transponiendo.

**Sistemas de acceso**<sup>10</sup> Lenguajes para consultar archivos.

#### **SPARQL (Protocolo Simple y Lenguaje de consulta de RDF)**

Lenguaje estandarizado para la consulta de datos RDF, normalizado por el W3C. Es una recomendación oficial del W3C desde enero de 2008 para el desarrollo de la web semántica.

Para mayor información consulta: [www.w3.org/standards/techs/sparql](http://www.w3.org/standards/techs/sparql)

#### **Web services - API (Interfaz de programación de aplicaciones)**

Son interfaces de programación de aplicaciones (API) o la API de la web que se accede a través de HTTP y se ejecuta en un sistema remoto de alojamiento de los servicios solicitados. Los servicios web son sistemas de software diseñados para apoyar la interacción interoperable máquina a máquina sobre una red. Tiene una interfaz descrita en un formato procesable por una máquina y otros sistemas interactúan con el servicio web de una manera prescrita por su descripción utilizando mensajes SOAP, transmitido a través de HTTP con una serialización XML en conjunto con otras normas relacionadas con la web.

Para mayor información consulta: [www.w3.org/TR/ws-gloss/](http://www.w3.org/TR/ws-gloss/)  
[www.w3.org/standards/techs/wsdl](http://www.w3.org/standards/techs/wsdl)

---

<sup>10</sup> <http://dadesobertes.gencat.cat/es/formats.html#anc02>



## **WMS (Servicio de Mapas Web)**

Formato que produce mapas de datos referenciados espacialmente, de forma dinámica, a partir de información geográfica. Es estándar internacional. Los mapas WMS se generan normalmente en un formato de imagen como PNG, GIF o JPG, y opcionalmente como gráficos vectoriales en formato SVD o WebCGM.

Para mayor información consulta: [www.opengeospatial.org/standards/wms](http://www.opengeospatial.org/standards/wms)

## **Siglas y acrónimos**

AEAT	Agencia Estatal Tributaria Española
AGE	Administración General del Estado
AGIMO	Australian Government Information Management Office
AOC	Administración Oberta de Catalunya
API	Application Programming Interface
App	Abreviación de “Application”, programa informático
APTC	Asociación Pro Transparencia de las Cuentas Públicas
ASEDIE	Asociación multisectorial de la información
ATOM	Formato de sindicación XML usado para la web
BOE	Boletín Oficial del Estado
CAD	Computer-Aided Design
CAGR	Compound Annual Growth Rate
CE	Comisión Europea
CENDOJ	centro documental del Consejo General del Poder Judicial
CEO	Chief Executive Officer
Cfr.	Abreviación latina “Ceferre” que significa “comparar”, “ver”
CHOCOLATE	Citizen-centricity, Harmonisation, Openness, Cross-border collaboration, Organisational change, Legal frameworks, Action plan, Trust, and Engagement
CINTEL	Centro de Investigación de las Telecomunicaciones
CIO	Chief Information Officer
CMS	Content Management System
CSS	Cascade Style Sheet
CSV	Comma Separated Values
CTIC	Acrónimo de la Fundación Centro Tecnológico

DC	Dublin Core
DNP	Dirección Nacional de Planeación
DXF	Drawing Exchange Format
ENI	Esquema Nacional de Interoperabilidad
ENS	Esquema Nacional de Seguridad
epSOS	european patients Smart Open Services
ESADE	Escuela Superior de Administración y Dirección de Empresas
Eurostat	Oficina Europea de Estadística ubicada en Luxemburgo
FICOD	Foro Internacional de Contenidos Digitales
FLOSS	Free Libre Open Source Software
FOAF	Friend Of A Friend
FOI	Freedom of Information Act
GEA	Government Enterprise Architecture
GEL	Gobierno en Línea
GML	Graph Modeling Language
HTML	Hypertext Markup Language
ICT	International Telecommunication Union
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
IOP	Interoperabilidad
JRC-IPS	Joint Research Centre, Institute for Prospective Technological Studies
JSON	JavaScript Object Notation, formato ligero para el intercambio de datos
KML-KMZ	Keyhole Markup Language comprimidos (KMZ)
LOD	Linked Open Data
LOPD	Ley Orgánica de Protección de Datos
MITyC	Ministerio de Industria, Turismo y Comercio
MPTyAP	Ministerio de Política Territorial y Administración Pública
MS	Microsoft
MSIS	Management of Statistical Information Systems
MySQL	Motor de base de datos SQL Open Source
NTI	Norma Técnica de Interoperabilidad
OASIS	Advancing Open Standards for the Information Society
OCDE	Organización para la Cooperación y el Desarrollo Económico
OFK	Open Knowledge Foundation
ONU	Organización de las Naciones Unidas
Op. cit.	Obra citada en las referencias bibliográficas
OSOR.EU	Open Source Observatory and Repository
PAE	Portal de Administración Electrónica
PCI	Plataforma de Colaboración Interadministrativa (Plataforma de Interoperabilidad para todo el Sector Público de Cataluña)
PDF	Portable Document Format

PDF-A	Portable Document Format for Archiving
PEPPOL	Pan-European Public Procurement Online
PIB	Producto Interior Bruto
PICA	Plataforma de Integración y Colaboración Administrativa (Plataforma de Interoperabilidad del Gobierno de Cataluña)
PPP	Purchasing Power Parity
RAI	Registro de Aceptaciones Impagadas
RDF	Resource Description Framework
RDF/N3	Terse RDF Triple Language, formato de serialización para grafos RDF
RISP	Reutilización de la Información del Sector Público
RSS	Really Simple Syndication
SARA	Red privada de comunicaciones que permite transferencias seguras entre todas las Administraciones Públicas del Estado español
SEMIC.EU	_Semantic Interoperability Centre Europe
SETSI	Secretaria de Estado de Telecomunicaciones y para la Sociedad de la Información
SPARQL	SPARQL Protocol and RDF Query Language
SPOCS	Simple Procedures Online for Cross- Border Services
SQL	Structured Query Language
STORK	Secure idenTity acrOss boRders linKed
TIC	Tecnologías de la Información y de las Comunicaciones
UBL	Universal Business Language
UK	United Kingdom (Reino Unido)
UN	United Nations
UNECE	United Nations Economic Commission for Europe
UOC	Universitat Oberta de Catalunya
URL	Uniform Resource Locator
US	United States
W3C	World Wide Web Consortium
WMS	Web Map Service
XLS	Extensión de los ficheros de hojas de cálculo Excel de Microsoft
XML	eXtended Markup Language
XSL	XML Style Sheet
ZIP	Extensión de volumen de ficheros comprimidos