



POLÍTICA NACIONAL DE DATOS ABIERTOS

GUIA DE IMPLEMENTACION

Secretaría de Participación Ciudadana, Transparencia y Anticorrupción de la
Presidencia



PRESENTACIÓN

Los Datos Abiertos son una herramienta que facilita la obtención de objetivos para múltiples sectores. Para poner los datos en acción y detonar la generación de valor, es importante socializar la publicación y promover el uso de los Datos Abiertos. Por lo anterior, las Instituciones deberán utilizar los canales de difusión pertinentes para dar a conocer los planes, acciones y cumplimiento con las disposiciones de la Política Nacional de Datos Abiertos.

Para conocer sobre la importancia y los principios de dato Abiertos, refiérase al documento de la Política Nacional de Datos Abiertos (URL). Este documento está orientado a facilitar la implementación de dicha política.

Pasos para implementar la Política Nacional de Datos Abiertos en su Institución

Actividad	Lineamientos de la PNDA
1. Nombrar al enlace y al administrador de Datos Abiertos institucionales	1- Marco Administrativo
2. Crear la Comisión Institucional de Datos Abiertos (CIDEA)	
3. Crear inventario Institucional de datos	2 - Inventario Institucional
4. Priorizar datos a publicar	5 - Apertura de los Datos
5. Crear plan de publicación	8- Usar Datos Abiertos 9- Gestión del cambio 10- Planificación de Recursos
6- Preparar Datos	4 - Datos Interoperables
7- Publicar datos	6- Identificar Demanda de Datos
8- Comunicar y promover el uso .	7- Promover uso de Datos Abiertos

1. Definir enlace y administrador institucional de datos abiertos.

El titular de cada Institución designará a dos servidores públicos responsables de dar cumplimiento al Política Nacional de Datos Abiertos, el Enlace Institucional y el Administrador



Guía de Implementación de la Política Nacional de Datos Abiertos

de Datos Abiertos.

El "Enlace Institucional" debe ser un miembro de la alta gerencia y actuará en representación directa del titular de la institución, por lo que contará con poder de decisión en los temas relacionados a la PNDA y será responsable de coordinar las acciones para cumplir con la Política de Datos Abiertos.

El "Administrador de Datos Abiertos" será el Director de Tecnologías de la Información de la Institución, o equivalente, y será responsable de coordinar la creación de un Inventario Institucional de Datos e implementar los procesos tecnológicos para que los datos abiertos se publiquen en el sitio institucional y se vean reflejados en datos.gov.sv.

Sobre estas designaciones, el Titular de la institución deberá enviar un oficio con las designaciones al Secretario Técnico y de Planificación de la Presidencia, siguiendo la plantilla No. 1 que se encuentra en el portal datos.gov.sv

2. Crear la Comisión Institucional de Datos Abiertos (CIDEA)

La Secretaría Técnica y de Planificación de la Presidencia podrá hacer recomendaciones sobre la designación de estas áreas cuando las Instituciones cuenten con áreas administrativas estratégicas encargadas específicamente de planeación y administración de datos.

El Enlace y el Administrador deberán coordinar la creación de la CIDEA bajo el lineamiento que manda a cumplir la Política de Datos Abiertos y la presente Guía, con las Unidades Administrativas pertinentes al interior de su institución, Además del Enlace y Administrador de Datos, esta CIDEA de trabajo debe incluir al menos a:

- **Dirección de Planificación**
- **Oficial de Archivo**
- **Oficina de Información de Respuesta** o Unidad de Acceso a Información Pública, según sea el caso.

El titular deberá facilitar los espacios y tiempos necesarios para que la CIDEA elabore el plan de implementación de la Política para toda la institución.

La asignación y roles del personal de la institución dentro de la CIDEA serán:

Rol	Funciones
Enlace de Datos Abiertos	<ul style="list-style-type: none"> - Garantizar la elaboración del plan de implementación de la PNDA de su institución. - Asegurarse de que las recomendaciones de esta Guía sean atendidos por las áreas responsables. -Facilitar proceso de apertura de datos, priorizar y divulgar su publicación.



Guía de Implementación de la Política Nacional de Datos Abiertos

	<ul style="list-style-type: none"> - Ser el principal punto de contacto con la Secretaría Técnica y de Planificación de la Presidencia, representando a su institución en todas las actividades a las que sea convocado
Administrador de Datos Abiertos	<p>Coordinar con Unidades Generadoras de datos para: Crear el Inventario de Datos y asegurar la conversión de conjuntos de datos a formatos abiertos e interoperables.</p> <ul style="list-style-type: none"> -Asegurar la calidad de los datos, su consistencia, seguridad y disponibilidad para ser consultados desde su web institucional y del portal nacional de Datos Abiertos (datos.gob.sv) - Crear plan de publicación de datos y asegurar su cumplimiento - Coordinar con la OIR que entre la información a publicar se encuentre la de mayor demanda del público.
Dirección de Planificación	<p>Identificar los mecanismos institucionales adecuados para la recopilación y ordenamiento de la información que se produce institucionalmente. –</p> <p>Incorporar en el plan institucional el plan de publicación de acuerdo a las prioridades que se identifiquen, vinculándolo a las estrategias del Plan Quinquenal.</p>
Oficial de Información y Respuesta	<ul style="list-style-type: none"> - Apoyar la priorización de datos a ser publicados. Preparar un reporte anual con los conjuntos de datos más solicitados por la ciudadanía, para ser publicados como Datos Abiertos. - Llevar estadísticas de la descarga de bases de datos y hacer encuestas o consultas sobre otros datos que puedan interesar al público. - Velar por la aplicación de la normativa vigente en materia de datos personales, reservados o confidenciales en relación a los Datos Abiertos de la Institución
Oficial de Archivo	<ul style="list-style-type: none"> - Apoyar en el ordenamiento de datos a ser publicado y en la revisión de los metadatos relacionados. - Velar por la aplicación de la normativa vigente en materia de gestión documental física y electrónica
PARTICIPACIONES DE APOYO AL CIDEA:	
Unidades Generadoras o Gestores de datos	<p>Participar eventualmente de las reuniones a las que sean convocados por la CIDEA, a fin de que aseguren la calidad mínima mínima de los datos y prepararlos para su publicación como Datos Abiertos (publicación de los procesos de creación y metadatos), siguiendo las recomendaciones de esta guía</p> <ul style="list-style-type: none"> - Participar en la elaboración del plan de implementación de la PNDA y apoyar en la publicación de Datos Abiertos según las prioridades que se identifiquen y de los datos que vayan a ser abordados.



Guía de Implementación de la Política Nacional de Datos Abiertos

Dirección de comunicaciones	<ul style="list-style-type: none"> - Diseñar y coordinar la ejecución de la estrategia de promoción y difusión de los datos que están a disposición ciudadana. - Publicitar el inventario institucional de los datos y del portal de datos abiertos
-----------------------------	---

Consideración: Si la institución por su naturaleza o tamaño no cuenta con suficiente personal o las unidades enunciadas, el enlace institucional y el enlace administrativo serán quienes asumirán la implementación de la política en su totalidad.

3. Crear inventario Institucional de Datos

El Administrador de Datos deberá solicitar a todas las Unidades Administrativas el llenado del **Catálogo de información reutilizable** para cada conjunto de datos que generan o administran. Cada unidad institucional deberá reportar al Administrador de Datos Abiertos los datos que generan, la prioridad en este caso son los datos primarios y misionales, es decir aquellos que por competencia produce la institución y que nadie más produce para lo cual la institución deberá analizar su ley de creación, servicios al ciudadano y Plan Estratégico. En dicho catálogo se incluye información sobre la clasificación de los datos, de acuerdo a la Ley de Acceso a Información Pública.

El Administrador junto al CIDAE será responsable de consolidar la información y filtrarla para elaborar el **Inventario Institucional de Datos** y seleccionar aquellas que conformarán el Catálogo Gubernamental de Datos Abiertos en datos.gob.sv.

Este ejercicio también servirá para identificar duplicidades en la generación de datos dentro de la institución o la ausencia de algunos (para determinar su elaboración). Si es el caso, se deberán tomar acciones para la designación de una sola área responsable para la generación (o consolidación) de un conjunto de datos o definir a la responsable de los nuevos.

Este inventario deberá ser revisado semestralmente y enviado al enlace de la Secretaría Técnica y de Planificación de la Presidencia que haya sido designado para ello.

El **Inventario institucional de Datos** es la base para crear el **Catálogo Gubernamental Datos Abiertos** y para su llenado se realizarán los siguientes pasos:

Actividad	Descripción	Duración
1. Identificar temas sustantivos	<p>La identificación de los temas sustantivos se elabora a partir de un ejercicio de revisión de los siguientes seis elementos:</p> <ol style="list-style-type: none"> 1. Información derivada de los procesos misionales de la Institución según su ley de creación. Esta 	4 horas



Guía de Implementación de la Política Nacional de Datos Abiertos

	<p>debe estar incluida dentro de las principales funciones que desempeña la Unidad Generadora de Datos (marco jurídico de la Institución, plan estratégico, manuales de procedimientos).</p> <ol style="list-style-type: none"> 2. Información utilizada para diseño y seguimiento de programas presupuestarios o políticas públicas ejecutadas por la institución. 3. Información para la atención a tratados, acuerdos o compromisos internacionales. 4. Información para la atención a solicitudes presidenciales, legislativas o sectoriales. 5. Información utilizada para la difusión periódica en publicaciones oficiales o académicas, y 6. Información recurrente en solicitudes ciudadanas. 	
<p>2. Definir los conjuntos de datos</p>	<p>Cada uno de los temas identificados en el paso anterior está relacionado con al menos un conjunto de datos, que a su vez está relacionado con el quehacer de la Unidad Generadora de Datos correspondiente. Estos pueden ser, entre otros, datos sobre la situación histórica, actual o proyectada:</p> <ol style="list-style-type: none"> 1. del sector donde se desempeña la función, 2. de la infraestructura utilizada, 3. del ejercicio presupuestal asociado, 4. del personal dedicado, 5. de indicadores de la operación, o 6. de la regulación aplicable. <p>En el Catálogo de información reutilizable, estos Grupos de Datos se deben registrar bajo la Columna X, titulada 'nombre del grupo', preferentemente con 'unidades de medida' y 'tema'.</p> <p>Por ejemplo:</p> <ul style="list-style-type: none"> - viviendas con servicio de agua potable, - habitantes beneficiados por el programa escuela solidaria, - monto ejecutado en proyectos de infraestructura de telecomunicaciones 	<p>4h</p>
<p>3. Definir los recursos de Datos</p>	<p>Cada Conjunto de Datos pueden ser publicado en diversos formatos para consumo del público ya sea por variaciones en su estructura (ej.: cartográfico, tabular), temporalidad (ej.: años, meses, días), medio (ej.: CSV, API), unidades de medida (población, viviendas), etc. Cada una de estas variantes es conocida como Recurso</p>	<p>16h</p>



Guía de Implementación de la Política Nacional de Datos Abiertos

	<p>de Datos.</p> <p>Es importante asegurarse de que el Recurso de Datos considere un solo concepto, en un solo ámbito geográfico y dentro de un solo periodo de tiempo, sin mezclar datos de diversos tipos.</p> <p>Cada Recurso de Datos se debe registrar en la Columna C 'nombre del recurso' como 'formato', 'unidad', 'materia', 'detalle geográfico' y 'período'</p> <p>Por ejemplo:</p> <ul style="list-style-type: none"> - mapa de viviendas con cobertura de alcantarillado por municipio 2014, - tabla de habitantes beneficiados por departamento en julio 2013 - monto ejecutado en proyectos de infraestructura de telecomunicaciones <p>Adicionalmente, se incluye una breve descripción aclaratoria del Recurso de Datos en la Columna X '¿de qué es?'</p>	
<p>4- Identificar formato y responsables</p>	<p>Es necesario identificar el formato de los Recursos de Datos (base de datos, Excel, etc.) y se registra en la Columna G '¿en qué plataforma, tecnológica, programa o sistema se albergan?'</p> <p>De igual forma, se debe registrar el cargo del funcionario designado como Responsable para su actualización en la Columna A, 'responsable', siendo la persona que ostenta la jefatura de la unidad generadora.</p> <p>Una vez identificadas las fuentes de datos y sus responsables, cada Unidad Generadora registra su listado en la plantilla del Inventario de Datos para su entrega al Administrador de Datos.</p> <p>Partiendo de la definición de Información Pública prevista en la LAIP y de la regulación aplicable a la función específica, la Unidad Generadora verifica en la Columna X '¿tiene datos confidenciales o reservados?' de ser necesario, en la Columna X 'razón por la cuál es confidencial o reservada los datos' debe referir el fundamento jurídico, con base en la LAIP, que así lo acredita.</p> <p>Por último, en la Columna X 'fecha estimada de publicación en datos.gob.sv (en caso de ser datos públicos) (AAAA-MM)' la Unidad Generadora determina una fecha estimada de publicación de cada Recurso de Datos.</p>	<p>2 horas</p>



Guía de Implementación de la Política Nacional de Datos Abiertos

	Para realizar este paso se puede descargar el documento <i>Inventario Institucional de Datos</i> disponible en datos.gob.sv/plantilla/inventario.xls	
5- Verificar y Aprobar Inventario	<p>Una vez recibidos los Inventarios de Datos de cada Unidad Generadora, el Administrador de Datos de la dependencia verifica que esté completa cada plantilla e integra el Inventario Institucional de Datos, ordenando e identificando claramente los temas misionales a los que pertenecen los diferentes Conjuntos de Datos.</p> <p>Una vez integrado el inventario, la CIDEA deberá revisar que efectivamente el Inventario sea representativo y cumple con la intención fundamental de uso, que los temas identificados sean exhaustivos y sustantivos, en correspondencia a las funciones principales de la Institución (según lo establecido en el paso 1). Esta sesión deberá ser documentada mediante una minuta debidamente firmada por los miembros de la CIDEA.</p>	4 h
6- Presentar Inventario	<p>Presentar el Inventario Institucional de Datos (ver formato de Hoja de Cálculo disponible en datos.gob.sv/plantilla/inventario.xls) junto con el acuerdo de la sesión de la CIDEA donde se formaliza el Inventario a la STPP.</p> <p>Es importante considerar que el Inventario deberá aprobarse dentro de los 70 días posteriores a la publicación de la Guía de Implementación de la Política de Datos Abiertos, por lo que se recomienda enviarlo junto el acuerdo para su aprobación con al menos 10 días antes de esta fecha.</p>	

4. Creación del Catálogo Gubernamental de datos abiertos

(Tiempo máximo para el cumplimiento *100 días*)

La priorización de datos es importante para identificar qué datos tienen mayor potencial de generar impacto y por lo tanto se deben empezar a publicar como Datos Abiertos. Esta priorización deberá hacerse con base en los datos que tengan más valor para la sociedad y los objetivos del gobierno.

La priorización deberá incluir:

1. Todos los conjuntos de datos considerados como Información de Interés Nacional.
2. Todos aquellos conjuntos de datos de carácter público que sean publicados en los sitios



Guía de Implementación de la Política Nacional de Datos Abiertos

y otras plataformas institucionales.

3. Los datos demandados por la ciudadanía mediante las solicitudes de acceso a la información más recurrentes.
4. Datos sobre los programas de la Institución que ayuden a cumplir con los objetivos del Plan Quinquenal u otros compromisos internacionales que el Gobierno ha firmado; así como con las acciones relevantes de la Institución.

Actividades a realizar por la institución

- a. Solicitar a la Oficina de Información y Respuesta un reporte de las solicitudes de acceso a la información pública más recurrentes para que sean incluidas en la primera publicación.
- b. Sesionar la Comisión Institucional de Datos Abiertos para priorizar los datos del Inventario Institucional de Datos según los objetivos de la Institución y el Plan Quinquenal.

La Secretaria Técnica de la Presidencia podrá hacer recomendaciones sobre la priorización de los datos cuando las entidades o dependencias cuenten con información que demuestre el cumplimiento de los objetivos del Plan Quinquenal de Desarrollo, inclusive cuando la misma esté relacionada objetivos y metas de otras instituciones.

4.1 Plan de Publicación

(Tiempo máximo para el cumplimiento *100 días*)

Este Plan será publicado anualmente, a más tardar el segundo viernes de febrero, con el objetivo de dar a conocer las acciones que serán realizadas por cada Institución en materia de datos abiertos. Como uno de sus componentes principales, el Plan de Publicación contendrá el calendario de publicación anual de los datos abiertos de la Institución, esto con el objetivo de señalar al público las acciones y fechas propuestas para la publicación de Datos Abiertos.

Actividades a realizar por la institución:

- a. Preparar el calendario de publicación de acuerdo al formato disponible en datos.gob.sv/plantilla/calendario.xls
- b. Publicar el Plan de Apertura Institucional como parte del catálogo de datos de la Institución, siguiendo las instrucciones del Paso X de esta guía.

5- Preparar los conjuntos de datos

Es importante estandarizar la manera en que los datos serán utilizados y consumidos por el público en general y el propio gobierno. Al exportar los conjuntos de datos a formatos estructurados, se facilita el consumo e interpretación de la información por terceros, así como por máquinas.

Al utilizar formatos abiertos, se extiende la audiencia de usuarios que pueden aprovechar los datos, ya que los formatos abiertos tienen su especificación o estándar abierto, que permite que sean interpretados por la mayoría de los sistemas computacionales.



Guía de Implementación de la Política Nacional de Datos Abiertos

Esta Guía pide que los datos se publiquen en formatos estructurados CSV. Sin embargo, para atender la diversidad de audiencias, sectores y casos de uso, se fomentará la publicación de los mismos conjuntos en diferentes **formatos estructurados de estándar abierto que resultan convenientes de acuerdo a la naturaleza de cada conjunto de datos**, como lo son XML, JSON, RDF, GeoJSON, KML, DBF, e inclusive propietarios como SHP y XLSX cuando la OIR establezca que este es el formato más solicitado por la ciudadanía. En general, se recomienda la publicación de datos en formatos altamente estructurados y abiertos; ver tabla 1 "Referencia de formatos, su nivel de apertura y estructura".

Tabla 1: Referencia de formatos, su nivel de apertura y estructura:

Formato	Tipo de datos	Apertura	Estructura	Notas	Ejemplo
CSV (Obligatorio)	Tabulares	Alta	Alta	Es un formato simple y universal, que estructura los datos en filas y columnas separadas por un carácter definido.	Población total por cada departamento. Número total de beneficiarios de un programa social por departamento.
XLSX	Tabulares	Media	Alta	La estructura es un XML con una especificación abierta, lo cual permite que cualquier software pueda interpretarlo.	Número total de accidentes de carretera. Número total de doctores y enfermeras en cada hospital.
XLS	Tabulares	Baja	Alta	Es un archivo binario y de formato propietario que restringe la utilización de este archivo por software de terceros.	Número total de escuelas primarias y secundarias en cada municipio.
DBF	Tabulares	Alta	Alta	Es un archivo que estructura los datos de forma simple. Aunque es uno de los archivos que contiene un SHP, es comúnmente utilizado o de forma individual para	Producción de petróleo y gas por entidad. Valor de las exportaciones de crudo por



Guía de Implementación de la Política Nacional de Datos Abiertos

				publicar datos tabulares.	cada destino geográfico.
XML	Estructurados	Alta	Alta	Lenguaje de marcado desarrollado por la W3C para almacenar datos en forma legible y estructurada, que propone un estándar para el intercambio de información entre diferentes plataformas.	Número total de
RDF	Estructurados	Alta	Alta	Los archivos con esta terminación, generalmente refieren a RDF/XML, que es la sintaxis XML. Otra sintaxis común es JSON-LD, que serializa RDF utilizando JSON.	
KML	Espaciales(vectorial)	Alta	Alta	Es un estándar abierto aprobado por la OGC y connotación XML. Comúnmente son agrupados y comprimidos en formato KMZ.	Información geoespacial de los Municipios Información geoespacial de zonas urbanas y rurales
WFS	Espaciales (vectorial)	Alta	Alta	Es una interface estándar aprobada por la OGC para realizar peticiones de entrega datos en notación GML(basada a su vez en XML).	
ODT	Texto	Alta	Baja	Formato con notación XML, del estándar abierto	Total de familias beneficiadas por localidad.



Guía de Implementación de la Política Nacional de Datos Abiertos

				OpenDocument	Volumen de las exportaciones por semestre. Información sobre unidades médicas y recursos en salud. Número total de afiliados al sistema del Seguro Social.
DOCX	Texto	Media	Baja	Ver notas del formato XLSX.	
PDF	Texto	Media	Baja	Este formato contiene los datos de manera desestructurada.	

Para cada tipo de conjunto de datos; sea tabular, vectorial, estructurado o de texto, existen una serie de formatos adecuados que se podrán elegir para su publicación, esto con el objetivo de poner los conjuntos de datos a disposición del público en los formatos más útiles.

El formato CSV será obligatorio para todos los conjuntos de datos, se podrán publicar formatos adicionales para cada conjunto, según los usos del público en general; por ejemplo, en el caso de datos tabulares hacer disponible la versión XLS de archivo junto con el CSV. Esto para asegurar la mayor inclusión posible.

Es importante asegurar la calidad del archivo exportado, para lo anterior, las Instituciones encargadas de la publicación deberán considerar lo siguiente:

5.1. Promover la interoperabilidad.

- 5.1.1. Evitar el uso de caracteres fuera del rango de ASCII6 (como los acentos) en el nombre del archivo y URL.
- 5.1.2 Utilizar codificaciones universales, como UTF-87, y en caso de utilizar otra codificación como ISO 8859-18 (latin-1), informar a los usuarios en los metadatos que acompañan el conjunto de datos.
- 5.1.3 Estructurar los campos de tiempo (fechas, horas y rangos temporales) utilizando ISO 8601 (p. ej. 2015-02-22).
- 5.1.4 Apegarse a las recomendaciones de la Dirección General de Estadística y Censos cuando se hable de regiones geopolíticas del país. P. ej., evitar "SS", y utilizar "02", "San Salvador", o "SS" uniformemente.

5.2 Hacer entendible los datos a los usuarios.

- 5.2.1. Describir claramente los nombres de columnas, p. ej. "Fecha de solicitud" en lugar de "fech_sol". Se deberá acompañar el conjunto de datos con un diccionario de datos (ver ACTIVIDAD C: Documentar los conjuntos de datos).

5.3. Habilitar la legibilidad por máquinas.

- 5.3.1. Eliminar cálculos adicionales en formatos tabulares, p. ej. una celda "Total" con la suma de una de las columnas.
- 5.3.2. Remover imágenes, logotipos, celdas unidas, y cualquier otra decoración de los archivos; para facilitar su procesamiento, por lo que deberán ser planos.



Guía de Implementación de la Política Nacional de Datos Abiertos

5.3.3. Los campos numéricos, incluyendo los monetarios y magnitudes, deben permanecer en un formato numérico de tipo entero o flotante. Es decir, evitar agregar símbolos monetarios o de unidades de medición, p. ej. "200 m²" o "\$20", haciendo explícitas las unidades en el nombre de la columna, o en una segunda columna. P. ej. "Monto en Dólares", "Distancia en KM".

5.4. Asegurar la consistencia y calidad de los datos.

5.4.1. Evitar tener valores de diferentes tipos (como texto y número) para una columna en diferentes registros o filas.

5.4.2. Evitar información duplicada.

5.4.3. Eliminar redundancia en datos.

5.4.4. Asegurar la representación de variables (unificar datos).

5.4.5. Evitar errores sintácticos.

5.5. Asegurar la normalización de los datos.

5.5.1. Apegarse a las normas y catálogos nacionales oficiales o en su defecto a mejores prácticas nacionales o internacionales para la estructura de datos estadísticos y geográficos.

5.5.2. En caso de que los datos representan dimensiones físicas, deberán corresponder al Sistema Internacional de Unidades¹².

Ejemplos de un conjunto de datos sin errores:

Fecha	Departamento	Código Oficial	Turistas Nacionales	Turistas Extranjeros
2013-01	Santa Ana	01	45729	6193
2013-01	San Miguel	02	238338	73386

Ejemplos de conjuntos de datos con errores:

		Turistas	
Fecha	Departamento	Nacionales	Extranjeros
Ene 2013	Santa Ana	45,729	6,193
01/2013	SM	238,338	73,386
* pie de página			

Actividades a realizar por parte de las Instituciones:

1. Exportar los datos a archivos estáticos tabulares.
2. Convertir los archivos a formatos abiertos CSV.
3. Identificar si se requiere estructurar algún formato adicional que facilite su lectura y uso de acuerdo al tipo de conjunto de dato.



Guía de Implementación de la Política Nacional de Datos Abiertos

- Asegurar la validez del formato y estructura del archivo.

6- Publicar Datos

Los conjuntos de datos deben estar disponibles para descarga directa en un URL único y evitando cualquier barrera de acceso, como el uso de contraseñas o interfaces de usuario. Debemos considerar que estos datos serán consumidos por máquinas, no sólo por personas.

Cada conjunto de datos debe estar disponible para ser descargado individualmente mediante un hipervínculo en el sistema de la Institución. Para lo anterior, se recomienda cargar los conjuntos de Datos Abiertos en un servidor web, ya sea en el mismo portal en línea con el que cuenta la Institución o en otros servidores de archivos propios o mediante servicios en la nube contratados con tercero.

Para facilitar la descarga, se podrán segmentar y comprimir los archivos mayores a 100 MB (utilizando formatos abiertos de compresión como gzip).

Estos hipervínculos serán utilizadas posteriormente para referenciar los conjuntos de datos en el Catálogo Gubernamental de Datos Abiertos en datos.gob.sv

6.1: Documentar los conjuntos de datos de acuerdo al estándar DCAT

Utilizar datos no debería ser confuso para el usuario. Es importante explicar de dónde vienen los datos y compartir referencias adicionales para entender a fondo el origen de éstos. Existen estándares internacionales que sugieren descripciones de uso común y facilitan la documentación de datos (Dublin Core, etc.).

Todas estas descripciones se capturan en los metadatos con formato DCAT, conforme a la plantilla que se encuentra en la sección correspondiente del portal de internet datos.gob.sv, que permiten identificar los datos abiertos de una organización, describirlos, y hacerlos públicos en datos.gob.sv. Esta tarea de documentación debe ser liderada por el Administrador de Datos Abiertos.

La Ficha descriptiva de conjunto de datos (disponible en datos.gob.sv/plantillas/ficha.xls) contiene la información que deberá ser llenada de acuerdo a las siguientes descripciones para cada metadato:

Metadatos (descriptores) del conjunto de datos (*dcat:Dataset*):

ds:identifier	Identificador único del conjunto de datos, utilizado para agrupar recursos dentro de éste, p. ej. "rezago-social", "unidades_medicas", "adquisiciones". Utilizar caracteres ASCII (p. ej. sin acentos).
ds:title	Título descriptivo del conjunto de datos, p. ej. "Programa de fomento a la agricultura", "Vuelos comerciales".



Guía de Implementación de la Política Nacional de Datos Abiertos

ds:description	Una explicación de los datos, con suficiente detalle para que los usuarios puedan entender si es de su interés, p. ej. "Apoyos otorgados a través del programa Opciones Productivas, desglosado a nivel de localidad".
ds:keyword	Lista de términos clave separados por coma, que facilitarán al usuario la búsqueda del conjunto de datos. Es importante considerar el uso de términos no técnicos. p.ej. "salud, medicinas, compras, agricultura".
ds:modified	Fecha y hora de la última modificación del conjunto de datos; en formato ISO8601 14, p. ej. "2014-05-27T01:42:05-05:00"
ds:contactPoint	Nombre de la persona de contacto que atenderá dudas y comentarios sobre el conjunto de datos.
ds:mbox	Correo electrónico de contacto para atender dudas y comentarios sobre el conjunto de datos.
ds:temporal	La fecha o fechas que cubren los datos, p. ej. "2013", "2010/2012", "2014-01/2014-04". Si es un rango de fechas, deberán ordenarse ascendentemente.
ds:spatial	El espacio geográfico que cubre el conjunto de datos. Puede ser una región, el nombre de un lugar, un polígono o un cuadro delimitador de coordenadas geográficas (bounding box) en GML15. P. ej. "San Miguel", 009, http://www.geonames.org/3583462/departamento-de-san-miguel.html , o "32.71,-112.32 27.99, -118.45".
ds:landingPage	Dirección electrónica para obtener mayor documentación o información sobre el conjunto de datos, como lo puede ser un manual, un sitio web, o un diccionario de datos. Este documento sirve como guía adicional para que el usuario entienda con mayor detalle los datos.
ds:accrualPeriodicity	Frecuencia con la cual el conjunto de datos será publicado o actualizado, p. ej. "mensualmente".

Metadatos (descriptores) del recurso o descargable (*dcat:Distribution*):

ds:identifier	La clave que identifica al conjunto de datos al que pertenece (y bajo el que se agrupa) este recurso. Ver ds:identifier.
rs:title	Título descriptivo del recurso o descargable, p. ej. "Asignación Presupuestaria del 2013", "Apoyos por municipio", "Apoyos por localidad".
rs:description	Ver ds: description. Esta explicación es adicional a la que existe en el conjunto de datos.
rs:downloadURL	Dirección electrónica (enlace) para la descarga del recurso.



Guía de Implementación de la Política Nacional de Datos Abiertos

rs:mediaType	Formato de archivo del recurso a descargar, p. ej. "text/csv", "application/rss+xml". Este campo permite al usuario buscar conjuntos de datos por formato en www.datos.gob.mx .
rs:byteSize	El tamaño en bytes del recurso o descargable, p. ej. 3145728
rs:temporal	Ver ds:temporal
rs:spatial	Ver ds:spatial

El producto final de este paso será un archivo CSV con la documentación de todos los conjuntos de datos a publicar.

Para documentar metadatos usando DCAT debe usarse la Ficha descriptiva de conjunto de datos que se encuentra en la sección correspondiente del portal de internet datos.gob.sv/plantillas/ficha.xls

6.2: Publicar el Catálogo Gubernamental de Datos Abiertos en datos.gob.sv

El administrador de datos abiertos será el encargado de generar y publicar el catálogo generado en el paso anterior.

Al cargar el catálogo de datos de la Institución en formato CSV, se transformará y traducirá automáticamente al estándar DCAT en formato JSON y los conjuntos de datos aparecerán en datos.gob.sv. Además se registrará un histórico de catálogos, y se proveerá de ayuda a los funcionarios responsables.

Las mejores prácticas de la comunidad internacional de Datos Abiertos, apuntan a que un catálogo central de datos abiertos facilita la búsqueda y acceso para los ciudadanos.

Al publicar los datos abiertos en datos.gob.sv, éstos son asociados inmediatamente a los términos de la licencia CC-BY-SA [detalles en datos.gob.sv/licencia], con los cuales se da la certidumbre de que cualquier persona puede usar los datos para cualquier fin legal que se proponga, con el único requisito de citar la fuente de origen y compartirlo de igual forma.

En datos.gob.sv se concentraran una gran variedad de conjuntos de datos de diferentes Instituciones, Gobiernos locales y municipales, así como de órganos autónomos. Al publicar en datos.gob.sv la Institución se suma a la Política de Datos Abiertos oficialmente.

El oficial de información de cada institución será el encargado de cargar la información en el portal de datos abiertos.

7. COMUNICAR Y PROMOVER

Cada institución, por medio de la CIDEA, y el Gobierno, por medio de la Dirección de Gobierno Electrónico, deberá trabajar para la creación del ecosistema que genere valor agregado a la publicación de los datos. Los objetivos de esta promoción son:

7.1 Fomentar la oferta por medio de la liberación de la mayor cantidad de datasets publicados.



Guía de Implementación de la Política Nacional de Datos Abiertos

7.2 Fomentar la demanda por medio de la reutilización y consumo de datos que otras instituciones gubernamentales han publicado.

7.3 Comunicar la publicación de nuevos conjuntos de datos así como el uso que se estén dando a datos publicados por otras instituciones.

7.4 La estrategia de promoción deberá estar acompañada de un proceso de información, educación y capacitación a los servidores públicos orientado a promover una cultura de apertura por defecto y que los datos sean lo más rico y completos posibles.

7.5 De la misma forma se deberá promover el uso especializado de los datos por medio de:

- Festivales de creación colaborativa de soluciones a problemáticas públicas
- Premios a aplicaciones especializadas que se deriven del uso del portal datos.gob.sv
- Comités interinstitucionales e intersectoriales de uso y transferencia de datos entre instituciones públicas.
- Convenios con aliados estratégicos para llevar los datos e información a la población que no tiene acceso a internet y recursos tecnológicos

Glosario

ASCII: es un código de caracteres basado en el alfabeto latino, comúnmente utilizados por sistemas computacionales para representar un mínimo conjunto de caracteres derivados del inglés moderno. La mayoría de los sistemas informáticos utilizan y entienden este código.

Catálogo Gubernamental de Datos Abiertos: el inventario único de los conjuntos de datos puestos a disposición de la población, en el portal de internet datos.gob.sv, por las Instituciones.

CIDEA: Comisión institucional de Datos Abiertos

CSV: el tipo de documento en formato abierto sencillo para representar datos en forma de tabla, en las que las columnas se separan por comas (o punto y coma en donde la coma es el separador decimal y las filas por saltos de línea.

DCAT: el Vocabulario de Catálogos de Datos es un estándar internacional diseñado para facilitar la interoperabilidad entre los catálogos de datos publicados en la red; para mayor información consultar <http://www.w3.org/TR/vocab-dcat/>.

Formatos Tabulares: los formatos que presentan los datos en forma de tabla con filas y columnas.

HTML (Lenguaje de marcas de hipertexto): el lenguaje de marcado para la elaboración de páginas web. Define una estructura básica y un código (denominado código HTML) para la definición de contenido de una página web, como texto, imágenes, entre otros.

Instituciones: las entidades, dependencias de la Administración Pública Federal o empresas productivas del Estado.

JSON (JavaScript Object Notation): es un formato ligero para el intercambio de datos que parte de un subconjunto de la notación literal de objetos de JavaScript. Dada su simplicidad, su uso se ha generalizado y popularizado en la Web.

Metadatos: los datos estructurados y actualizados que describen el contexto y las características de contenido, captura, procesamiento, calidad, condición, acceso y distribución



Guía de Implementación de la Política Nacional de Datos Abiertos

de un conjunto de datos, que sirven para facilitar su búsqueda, identificación y uso.

OIR: Oficina de Información y Respuesta, unidad encargada de la implementación y ejecución de la Ley de Acceso a la Información Pública.

PDF: (*portable document format*): el formato de almacenamiento de documentos digitales independiente de plataformas de software o hardware.

TXT: el archivo informático compuesto únicamente por texto sin formato.

URL: es la secuencia de caracteres, de acuerdo a un formato estándar, que se usa para nombrar recursos en Internet para su localización o identificación, como por ejemplo documentos textuales, imágenes, páginas Web. La URL del sitio de datos abiertos es <http://datos.gob.sv>

XLS: la extensión de archivo por defecto del formato propietario de Excel.

XLSX: es la extensión de archivo de Excel, en su versión de formato abierto. Ésta cumple con la característica de seguir el estándar abierto Office Open XML.

XML: (lenguaje de marcas extensible): el lenguaje de marcas utilizado para almacenar datos en forma legible. Permite definir la gramática de lenguajes específicos (de la misma manera que HTML) para estructurar documentos grandes. Da soporte a bases de datos, siendo útil cuando varias aplicaciones deben comunicarse entre sí o integrar información.