



Versión 1



Guía Nacional para la Publicación de Datos Abiertos



COSTA RICA
GOBIERNO DEL BICENTENARIO
2018 - 2022



**DATOS
ABIERTOS**
COSTA RICA

Mayo 2020



Contenido

a	Alcance	5
b	Marco normativo	5
c	¿Qué son los datos abiertos?	5
d	Principios de los datos abiertos	6
e	¿Por qué abrir los datos?	7
f	Niveles de apertura	8
g	Proceso de apertura	9
o	Preparación institucional	10
1	Identificación de la demanda con consulta a actores del ecosistema	14
2	Evaluación y estado de los datos	15
3	Priorización de los conjuntos de datos a liberar	17
4	Análisis de confidencialidad de la información	19
5	Limpieza de los conjuntos de datos	21
6	Metadatos	22
7	Formato abierto, neutral, interoperable y georeferenciado	24
8	Asignación de la licencia de uso y reutilización	26
9	Publicación de los datos	27
10	Comunicación y promoción de los datos	28
h	Sostenibilidad	31
i	Continuidad de la iniciativa	31
j	Monitoreo y evaluación	33
k	Comisión Nacional de Datos Abiertos	35
l	Glosario	36
m	Fuentes de referencia	38



Índice de figuras

Figura 1: Niveles de apertura de datos	8
Figura 2: Niveles de apertura de datos (Esquema de desarrollo de 5 estrellas)	8
Figura 3: Proceso de Apertura de datos	9
Figura 4: Conformación de equipo de apertura de datos	12
Figura 5: Evaluación del estado de los datos	15
Figura 6: Estadios	16
Figura 7: Criterios para facilitar la evaluación y priorización de datos	18
Figura 8: Formatos abiertos según tipo de documento e información contenida	24
Figura 9: Licencias <i>Open Data Commons</i>	26
Figura 10: Actividades para promoción y comunicación	29
Figura 11: Actividades para promoción y comunicación	30
Figura 12: Plantilla para proceso de actualización de datos	31
Figura 13: Plantilla para proyectos institucionales en apertura de datos	32
Figura 14: Propuesta para medición y monitoreo de acciones en apertura de datos	34



La Guía Nacional para la Publicación de Datos Abiertos es el instrumento que define las pautas a seguir por las instituciones públicas y por los equipos institucionales de datos abiertos a fin de cumplir lo establecido en el Decreto N.º 40199-MP Apertura de Datos Públicos. En dicha regulación, se indican las etapas que conforman el proceso de apertura de datos. Asimismo, se establece que las especificaciones técnicas y los estándares para ejecutar de forma normalizada el proceso serán definidos en esta guía.

● ● ● ● ● ● ● ● ● ● ● ●

a Alcance

La Guía Nacional para la Publicación de Datos Abiertos es el instrumento que orientará a las instituciones del sector público costarricense para el cumplimiento de lo establecido en el Decreto 40199-MP Apertura de Datos Públicos. Esta guía tiene la finalidad de establecer las especificaciones técnicas y estándares para la ejecución del proceso descrito en el decreto, de forma normalizada en las instituciones públicas.

b Marco normativo

En Costa Rica se cuenta con algunas herramientas legales que guían la apertura de datos públicos, las cuales se mencionan a continuación:

- ◆ **-Constitución Política de Costa Rica. Artículos 27 y 30**
- ◆ **-Ley N° 7975 de Información no Divulgada**
- ◆ **-Ley N° 8968 Protección de la Persona Frente al Tratamiento de sus Datos Personales**
- ◆ **-Ley N° 9694 Sistema de Estadística Nacional**
- ◆ **-Decreto Ejecutivo N° 40199-MP Apertura de Datos Públicos**
- ◆ **-Decreto Ejecutivo N° 40200-MP-MEIC-MC Transparencia y Acceso a la información Pública**
- ◆ **-Directriz Ejecutiva N° 074-MP Apertura de Datos Públicos**
- ◆ **-Directriz Ejecutiva N° 073-MP-MEIC-MC Transparencia y Acceso a la información Pública**



c ¿Qué son los datos abiertos?

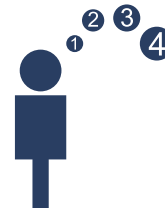
El Decreto 40199-MP Apertura de Datos Públicos define el dato abierto como: “dato disponible en línea, sin procesar, en formato abierto, neutral e interoperable; que permite su uso y reuso, disponible para su descarga en forma completa, sin costo, ni requisitos de registro y procesable en computadora”.



d Principios de los datos abiertos

El Decreto Ejecutivo 40199-MP Apertura de Datos Públicos cita los siguientes principios:

- 1 **Abiertos por defecto:** todos los datos públicos serán dispuestos en formatos abiertos estructurados, con la capacidad de ser procesados e interpretados por equipos electrónicos de manera automática. Deberán realizarse las acciones necesarias para que esta condición se dé desde el momento en que los datos son generados o recolectados.
- 2 **Oportunos y exhaustivos:** los datos públicos abiertos se pondrán a disposición en forma oportuna, completa, precisa y exhaustiva. Además, se mantendrán versiones históricas de los datos para garantizar su permanencia.
- 3 **Accesibles y utilizables:** los datos públicos abiertos deben ser fácilmente visibles y accesibles, poniéndose a disposición de forma gratuita y sin barreras.
- 4 **Comparables, neutrales e interoperables:** los datos públicos abiertos deben estar disponibles en su fuente primaria, con el más alto grado posible de desagregación, o hacer referencia a la base principal cuando se disponga de forma agregada; deben ser fáciles de comparar dentro y entre sectores, a través de localizaciones geográficas y del tiempo. Además, deben ser presentados en formatos estructurados y armonizados que faciliten su comparación, intercambio, trazabilidad y reutilización efectiva.
- 5 **Mecanismo para fortalecer la gobernanza y la participación ciudadana:** los datos públicos abiertos deberán fortalecer la confianza en las instituciones públicas y la gobernanza, entendida esta última como la gestión institucional y su relación con los actores sociales. De manera tal que provean un fundamento para la transparencia y la rendición de cuentas que mejore la toma de decisiones, incremente la demanda y prestación de servicios públicos; así como la formulación y evaluación de programas y políticas para satisfacer las necesidades de la sociedad.
- 6 **Fomento del desarrollo incluyente y la innovación:** la disponibilidad de datos abiertos debe estimular la creatividad y la innovación, generando procesos y espacios para identificar desafíos sociales y económicos, así como el monitoreo de programas de desarrollo sostenible con participación de todos los sectores involucrados.
- 7 **Respeto de los derechos fundamentales:** la recolección, publicación y uso de datos de gobierno deberá respetar la normativa y los principios generales en materia de derechos humanos expresados en tratados internacionales y en la Constitución Política.



e ¿Por qué abrir los datos?

Incorporar la publicación de datos abiertos como parte del flujo de trabajo institucional genera diferentes resultados. Por un lado, quienes publican, hace que toman conciencia del estado de situación e importancia del activo de datos. A la vez, el esfuerzo por mejorar la calidad de los datos facilita el consumo y genera nuevos usos.

Por otro lado, para los diferentes actores de la sociedad que articulan con el gobierno y para los órganos y entes públicos entre sí, se da la oportunidad de producir nuevo valor agregado sobre los datos generados y combinarlos con otras fuentes, ya sean propias o de terceros.

En este contexto, algunos beneficios de los datos abiertos son:

- Fomentar la toma de decisiones basada en evidencias para los sectores público, privado y social.
- Facilitar la prestación de servicios públicos más rápida, innovadora y eficiente.
- Propiciar una mayor transparencia, rendición de cuentas e integridad pública.
- Permitir el empoderamiento ciudadano, enriqueciendo el ejercicio de derecho de acceso a la información pública y fomentando la participación del público.
- Fomentar el espíritu emprendedor e innovador.
- Facilitar la investigación y generación de conocimiento.



f Niveles de apertura

Existen diferentes escalas y parámetros para definir el nivel o grado de apertura de los datos. Tim Berners-Lee, creador de la web y pionero de los datos enlazados (Linked Data), sugirió un esquema de clasificación a partir de 5 estrellas, cada una de las cuales representa un nivel o escalón de desarrollo en la apertura de los datos, del más simple al más complejo.

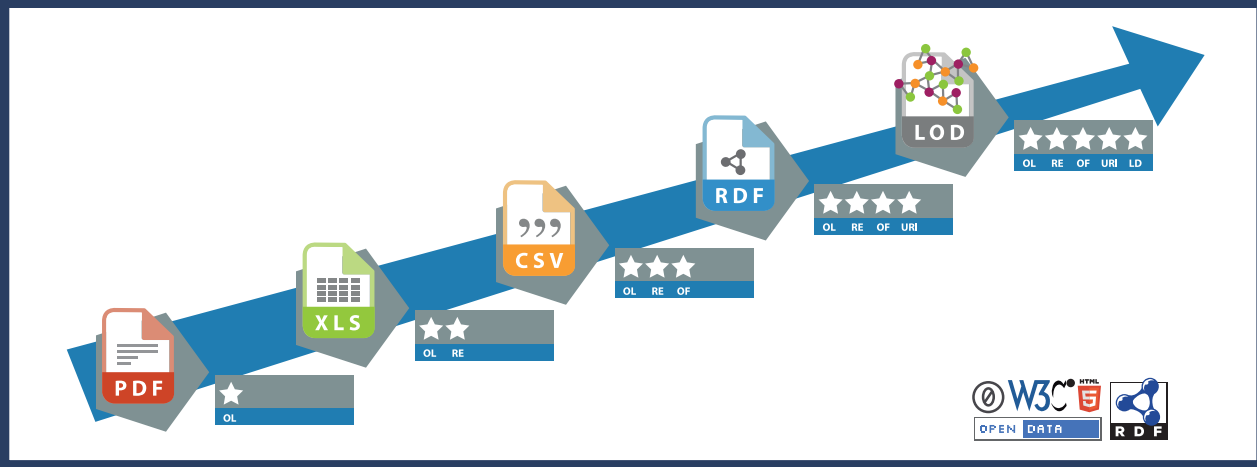
Figura 1: Niveles de apertura de datos

 OL	Disponibles en la web, bajo licencia abierta. Ejemplo: Datos en .pdf disponibles en la web.
 OL RE	Lo anterior, en forma de datos estructurados, en un formato que pueda ser interpretado o procesado por máquinas. Ejemplo: Datos en un archivo .xls
 OL RE OF	Lo anterior, y en un formato de archivo libre. Ejemplo: Datos en un archivo .ods .csv
 OL RE OF URI	Lo anterior, y utilizando estándares recomendados por W3C (RDF-SPARQL). Ejemplo: Datos con una URI o URL (identificador web)
 OL RE OF URI LD	Lo anterior, más el establecimiento de vínculos con los datos de otros y de esa forma proveerles contexto.

Fuente: Adaptación de <https://5stardata.info/en/>

Este cuadro tiene la funcionalidad de diagnosticar el nivel de apertura de datos con los que cuenta cada institución y provee grados alcanzables para llegar a niveles cada vez más refinados de apertura de datos. Asimismo, esta guía recomienda un camino gradual hacia la producción nativa de datos abiertos enlazados (nivel cinco estrellas).

Figura 2: Niveles de apertura de datos (Esquema de desarrollo de 5 estrellas, Tim Berners-Lee)



Fuente: Adaptación de <https://5stardata.info/en/>

g Proceso de apertura

La apertura de datos demanda ser asumida como un nuevo servicio o proceso institucional, que se incorpore de manera permanente en la estrategia y operación de la organización. Cada institución debe asegurar las condiciones organizacionales que permitan ubicar y desarrollar la apertura de datos de forma continua, apoyada con un sistema de evaluación y mejora permanente que potencie su alcance y funcionamiento.

El proceso de apertura consiste en la planeación, según las instancias habituales de la institución, para lograr la definición de objetivos y las actividades para alcanzar su cumplimiento.

El proceso de apertura de datos debe abordarse como una labor en la que todas las unidades generadoras de datos, de manera articulada en diferentes áreas, desde los aspectos legales, tecnológicos, de capacitación y generación de documentación aporten al proceso.

Figura 3: Proceso de apertura de datos



Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.

A continuación se detallan con precisión los pasos a seguir para el proceso de apertura de datos que permitirá que las acciones que contemplen en el Plan de Acción para la Apertura de Datos, se enmarquen en la planificación institucional y puedan vincularse a instrumentos existentes, tales como planes operativos, planes estratégicos de departamentos de tecnología, comunicación, programas internos e interinstitucionales (ej.: Índice de Transparencia Institucional) y de los mecanismos de participación ciudadana existentes.

0 Preparación institucional

Para que el proceso de apertura de datos se afiance y se mantenga, es necesario conformar un equipo multidisciplinario, dinámico, capaz de ajustarse e innovar con el entorno. En esta etapa se realiza el diagnóstico de las condiciones de la gestión de datos de la institución para el proceso de apertura, se define la normativa, estructura organizacional y el planteamiento estratégico indispensables para su desarrollo. Incluye los siguientes elementos, que deben ponerse en práctica. **Tomando en cuenta las particularidades de cada contexto institucional y se desarrolla con las siguientes actividades:**

Diagnóstico de capacidad institucional

Este ejercicio inicial permite conocer en detalle las condiciones de la institución en todos aquellos aspectos que se relacionan con un proceso de apertura de datos, tales como:

- A Identificar iniciativas existentes:** documentar si existen iniciativas institucionales en materia de datos abiertos y gobierno abierto, así como la existencia de grupos de trabajo, planes de trabajo, proyectos u otras iniciativas existentes.
- B Identificar recursos disponibles:** considerar si actualmente la institución cuenta con recursos humanos, tecnológicos y económicos disponibles para ejecutar la estrategia de datos abiertos.
- C Levantar el inventario de información:** listado de bases de datos de información pública, listado de datos abiertos en proceso de apertura y lista de datos abiertos ya publicados.
- D Identificar marco normativo específico:** identificar si la institución dispone de un marco normativo específico, convenios, oficios o acuerdos en materia de datos abiertos.
- E Cuantificar la capacitación en datos abiertos:** identificar la cantidad de personas funcionarias que han recibido capacitación y conocen del tema de datos abiertos. Mapear esta información en las distintas áreas de la institución, incluyendo los distintos niveles institucionales.
- F Identificar los puntos de ingreso de demandas de información:** listar los diferentes espacios donde ingresan las solicitudes de información, por ejemplo: contralorías de servicio, correo electrónico, oficios, sistemas digitales, entre otros.

Instrumentos para el respaldo y formalización del proceso de apertura:

A Identificado en el diagnóstico y el marco normativo existente, cada institución deberá proceder a la adopción de aquellos instrumentos que le sean de su competencia y a generar los específicos que sean requeridos según la naturaleza del ente.

B En consecuencia, corresponderá a cada institución realizar los actos necesarios y emitir las directrices internas para sustentar y dar continuidad al desarrollo del proceso de apertura, tales como acuerdos de junta directiva, consejo de administración y municipales.

Equipo

Como servicio público, las iniciativas de datos abiertos experimentan los mismos retos de la administración pública pues necesitan de una dirección central que lo gestione pero también de los mecanismos, para llevar el cambio cultural a toda la organización. Por esto es necesario insistir en un alto liderazgo para su conducción.

Si se desea el éxito y la permanencia en el tiempo de la iniciativa, las unidades o dependencias generadoras de los datos deben estar bajo una coordinación técnica centralizada y transversal designada por la persona jerarca de la institución, delegando en dicha coordinación suficiente autoridad para garantizar el cumplimiento de las directrices que se emitan.

El proceso de apertura de datos deberá abordarse como un trabajo en el que varias áreas están involucradas en diferentes dependencias de trabajo, desde los aspectos legales hasta los tecnológicos, pasando por la capacitación y la generación de documentación que aporte al proceso.

Llegar a alcanzar tal nivel organizacional, también es un proceso que depende de las capacidades de cada institución. Como punto de partida para el avance en la implementación del proceso de apertura de datos, se establece la necesidad de desarrollar dos acciones:

Coordinar:

La recepción centralizada de propuestas de liberación de datos a partir de demanda, valoración preliminar de las necesidades para llevar a cabo la propuesta de apertura de datos, coordinación con las instancias internas elaboración del plan de acción, seguimiento y evaluación del proceso y emisión de recomendaciones de mejora.

Implementar:

La persona encargada del proceso técnico de apertura de los datos (limpieza, estandarización, asignación de formato abierto, publicación), también tendrá a su cargo algunas tareas de la fase de comunicación y promoción, tales como la generación de visualizaciones o la realización de actividades para incentivar el uso de los datos publicados.



Es importante tener en cuenta que si bien esta es una labor que requiere de un fuerte componente tecnológico, el equipo debe tener un carácter diverso en cuanto a disciplinas y contemplar la necesidad de incorporar personas de áreas como estadística, derecho, diseño gráfico, archivística y comunicación.

Idealmente se recomienda que cada institución conforme un equipo de trabajo integrado por personas funcionarias que estén asociadas en su quehacer, en el ámbito de las tecnologías de la información, análisis de información, asesoría legal y archivística (gestión de documentos) y comunicación entre otros.

Considerando que actualmente no todas las instituciones cuentan con un equipo interdisciplinario de datos abiertos, se recomienda crear una figura a lo interno de la institución que permita impulsar el desarrollo de la apertura de datos. Las instituciones que ya cuentan con un equipo, deben verificar que se estén asumiendo todas las tareas y realizar los ajustes necesarios.

Para ello, se propone que se conforme, como mínimo, de la siguiente manera:

Rol	Tarea
Coordinador	Supervisa e incentiva las acciones para el logro de la apertura de datos
Analista	Realiza el inventario de datos, los prioriza, crea datasets y los publica a través del portal de datos abiertos
Técnico	Brinda el soporte técnico especializado en la preparación de las bases de datos, la transformación en formatos abiertos y su publicación en un portal de datos abiertos (carga, portal, APIs)
Comunicador	Establece las estrategias y acciones de promoción y difusión de las estrategias, publicaciones y portales de datos abiertos a través de los medios que cuenta la institución

Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.

Adicional a esto, el equipo debe contar con una red de enlaces facilitadores de datos, que provean la información necesaria y la ejecución de tareas como el inventario de datos, priorización y documentación de datasets, y participar activamente en las reuniones de trabajo.

Los aspectos legales asociados a la apertura de datos, en el cumplimiento de la normativa nacional, institucional, protección de datos personales y normas conexas, debe ser apoyada en las unidades o departamentos de asesoría jurídica propias de la institución.



Etapas del proceso de apertura de datos:



1 Identificación de la demanda con consulta a actores del ecosistema

Identificación de la demanda

Un proceso de apertura de datos es más efectivo cuando se mantiene cercanía con la población, se les escucha para conocer sus necesidades e intereses. Esto provoca que además de entregar información útil se pueda esperar que haya consumidores efectivos de los datos publicados.

En general, la apertura de datos no debe de entenderse únicamente como un sinónimo de transparencia o la rendición de cuentas. Es la oportunidad que gobiernos e instituciones tienen para poner en manos de la ciudadanía insumos para la creación de conocimiento de ahí que tenga alta prioridad acercar conjuntos de datos que sean demandados por las audiencias sin que se cuestione su propósito o motivación para solicitarlos. También es necesario resaltar, que existe demanda interna e interinstitucional, de manera que se planifiquen acciones con los actores de esos sectores.

En ese sentido, la identificación de la demanda puede realizarse usando varios formatos para conocer a las audiencias interesadas en datos abiertos, por ejemplo, sondeos, grupos focales, uso de indicadores, entre otros.



Recomendaciones:

- a. Involucre personas usuarias desde el principio del proceso y mantenga contacto con ellas.**
- b. Dialogue con los entes y las personas usuarias actuales y potenciales de los datos (ya sean ciudadanía, empresas, organizaciones no gubernamentales o desarrolladores).**
- c. Tenga en cuenta que muchos datos no alcanzarán directamente a las personas usuarias, sino que esto ocurrirá a través de intermediarios, que se apoderan de los datos y los transforman para entonces volver a presentarlos al público.**
- d. Cree espacios donde la ciudadanía haga recomendaciones.**
- e. Acérquese a eventos o actividades de consulta con personas activistas, representantes de organizaciones, empresarias, directores, cívicas u otras asociadas a la temática.**

Antes de definir las acciones a seguir, se evalúa al interior del equipo encargado del proceso el resultado de los ejercicios. Al finalizar la revisión se deberá elaborar en conjunto con todas las personas integrantes del grupo de trabajo, una lista de bases de datos que se podrían adecuar al proyecto y priorizar los esfuerzos en los que fueron más mencionados por las personas usuarias debido a su demanda y valor.

2

Evaluación y estado de los datos

Después de identificar la demanda y tomando en cuenta los hallazgos del diagnóstico, se debe verificar el estado de los datos y su calidad. Para ello, podemos verificar que cumpla con algunos puntos. Para identificar el estado de cada uno de los conjuntos de datos, se facilita la siguiente tabla:

Figura 5: Evaluación del estado de los datos			
Componente	Nivel	Valor	Calificación
Captura	No se recolectan datos	1	
	Se recolectan datos pero no están sistematizados	2	
	Se recolectan y están sistematizados	3	
Estructura	No estructurados	1	
	Semi estructurados	2	
	Estructurados	3	
Formato	Físico	1	
	Digital (no estructurados)	2	
	Digital (estructurados)	3	
Legal	Información de acceso restringido	1	
	Información que requiere de anonimización	2	
	Información de acceso público	3	
Actualización	Desactualizado	1	
	Actualizado al periodo anterior	2	
	Actualizado	3	
Nivel de desagregación	Procesada (estadísticas)	1	
	Desagregada pero requiere tratamiento	2	
	Desagregada, consciente y limpia	3	
Importancia estratégica y sectorial	Baja	1	
	Media	2	
	Alta	3	
Demanda	Baja	1	
	Media	2	
	Alta	3	
		TOTAL:	

Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.

Una vez aplicada esta evaluación a cada conjunto de datos podemos obtener los siguientes estadios de cada uno de ellos:

Figura 6: Estadios	
Valor entre 1 y 8:	Implica que es un conjunto de datos que técnicamente requiere de mayor tratamiento para lograr ser publicado en formato abierto. Por lo cual obtiene una prioridad de apertura baja.
Valor entre 9 y 16:	Implica que es un conjunto de datos que técnicamente requiere de mejoras menores para ser publicado en formato abierto. Por lo cual obtiene una prioridad de apertura media.
Valor entre 17 y 24:	Implica que es un conjunto de datos que técnicamente cuenta con buenas condiciones para ser publicado en formato abierto. Por lo cual obtiene una prioridad de apertura alta.

Al aplicar mejoras técnicas a cada uno de los conjuntos, es posible volver aplicar la evaluación y generar un nuevo criterio sobre el estado de los datos. Con ello se puede incrementar la cantidad de conjuntos de datos a publicar en formato abierto.

3 Priorización de los conjuntos de datos a liberar

Como ya se ha señalado en el apartado anterior, la apertura de datos es un proceso dinámico y continuo basado inicialmente en la demanda. Sin embargo, también influye la misión y aspiraciones propias de la institución, objetivos políticos y coyuntura nacional-local.

Por lo anterior, se recomienda priorizar los conjuntos de datos que la institución pretende publicar primero.

Debe darse prioridad a aquellos conjuntos de datos que tengan el mejor equilibrio entre sí:

- ◆ **Conjuntos de datos para los cuales existe una alta demanda o importancia estratégica**
- ◆ **Disponible para publicación (según estado de los datos) o que ya estén publicados pero no en formato abierto**
- ◆ **Conjuntos de datos requeridos por la normativa (decretos, directrices) o para mejorar la transparencia. Detallados en el artículo 17 del Decreto 40199-MP**
- ◆ **Datos con requerimientos específicos o llamados como prioridades nacionales**



Tener esta valoración aumenta la posibilidad de que los datos se reutilicen inmediatamente después de su publicación, lo cual es importante debido a que genera historias de éxito útiles para obtener soporte para el proceso de abrir más conjuntos de datos.

Además, la selección de datos para su apertura requiere del análisis de criterios de valor para actores internos y externos, costos para publicarlos y mantenerlos y riesgos asociados a su publicación. Estos criterios ayudan a evaluar el impacto que tendrá la apertura de cada conjunto de datos, no sólo para el órgano que la realiza sino también para la sociedad.

En caso de solo tener claridad sobre las prioridades, las siguientes preguntas pueden ayudar con la evaluación mencionada:

Figura 7: Criterios para facilitar la evaluación y priorización de datos

Criterio: Valor	
Subcriterio: Actores	1. ¿Quiénes son o podrían ser las personas usuarias internas y externas? 2. ¿Cuántos son? 3. ¿Para los entes o personas interesadas los datos son de valor a corto o largo plazo? 4. ¿Cuánto valor les reporta la interacción con el recurso? 5. ¿Con qué frecuencia puede ser consumido el recurso?
Subcriterio: Potenciales usos	1. ¿Los datos son utilizados actualmente para cumplir funciones de la institución? 2. ¿Los datos pueden reducir costos? 3. ¿Los datos miden la efectividad de los programas de gobierno? 4. ¿Cuál es el potencial de los datos para reducir los costos de los actores externos al gobierno? 5. ¿Cuál es el potencial de los datos para fomentar la creación de oportunidades de negocio? 6. ¿Cuál es el potencial de los datos para fomentar la innovación? 7. ¿Cuál es el potencial de los datos de crear valor agregado o fomentar el crecimiento económico? 8. ¿Cuál es el potencial de los datos para fomentar nuevas formas de colaboración?
Criterio: Costos	
Subcriterio: Publicación	1. ¿Se requiere un cambio de formato para compartir los datos? 2. ¿Hay definiciones de los datos y contexto suficiente para una comprensión adecuada? 3. ¿Cuál es el tiempo estimado de preparación de los datos para su apertura? 4. ¿Hay aspectos regulatorios importantes a tener en cuenta para compartir los datos (Ej.: privacidad, seguridad, accesibilidad, obligación legal de publicar, etc.)? 5. ¿Se requiere asesoramiento legal para evaluar la apertura de datos?
Subcriterio: Operación y actualización	1. ¿Qué unidad/departamento va a comprometer recursos humanos para compartir los datos? 2. ¿Con cuánta frecuencia se requerirá actualizar los datos? 3. ¿Qué cambios de sistemas deberán llevarse a cabo? 4. ¿Se requieren nuevos recursos tecnológicos? 5. ¿Puede la integridad de los datos ser cuestionada?
Criterio: Riesgos	
Subcriterio: Confiabilidad	1. ¿Pueden los datos inducir a interpretaciones que disminuyan su confianza?
Subcriterio: Legales	1. ¿Los datos pueden comprometer la seguridad pública? 2. ¿Los datos están protegidos por leyes específicas? 3. ¿Existen limitaciones normativas nacionales sobre discrecionalidad de los datos?
Subcriterio: Sociales y políticos	1. ¿Los datos exponen situaciones sensibles de individuos o grupos? Incluso combinados con otros datos existentes 2. ¿Los datos liberados pueden tener consecuencias contraproducentes sobre individuos o grupos de individuos?

Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.



Análisis de confidencialidad de la información

Como parte del proceso de apertura se debe revisar si la información contenida en la base de datos está sujeta a las normas de privacidad y confidencialidad.

Por un lado, la privacidad es un derecho humano internacionalmente reconocido, y como tal abarca el derecho a la protección de datos personales, entendidos estos como cualquier dato relativo a una persona física identificada o identificable.

Por otro lado, la confidencialidad recae en las personas que administran y son responsables de ciertas bases de datos, según se verá, y en toda aquella información que haya sido producida privadamente o que pertenezca a una persona privada.



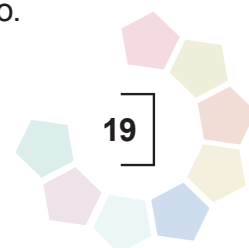
En el caso de datos personales, la Ley N° 8968 Protección de la persona frente al tratamiento de sus datos personales, los clasifica estos en tres tipos: datos personales de acceso irrestricto, datos personales de acceso restringido y datos sensibles. Una definición para cada uno de estos conceptos puede encontrarse en el glosario.

Dicha ley prevé que el tratamiento de datos sensibles se encuentra prohibido, en consecuencia, estos no deberían estar incluidos en bases de datos públicas o privadas, salvo en los casos exceptuados por ley.

En el caso de datos de acceso restringido se debe proceder a emplear técnicas de anonimización que no permitan la identificación de la persona por procesos ulteriores. Esto no es necesario en el caso de datos de acceso irrestricto.

Se deben considerar, adicionalmente, las excepciones y limitaciones a los derechos de privacidad, particularmente los indicados en el artículo 8 de la Ley N° 8968, así como las resoluciones reiteradas de la Sala Constitucional referentes a la información sobre funcionarios públicos, entre otros pero no limitado a que ciertos aspectos tales como clase de labor que ejerce, buen o mal desempeño de su cargo, ventajas o desventajas que el nombramiento supone, y los derechos que en virtud de su cargo obtiene, como por ejemplo el salario, por ser de interés general, son datos públicos y no requieren anonimización (veánse en ese sentido las resoluciones 2013004876 de las 9 horas 5 minutos del 12 de abril del 2013, 2014004037 de las 11 horas y 2 minutos del del 21 de marzo del 2014 y 2015009425 de las 9 horas 5 minutos del 26 de junio del 2015).

Es necesario, considerar que la normativa obliga a los responsables de las bases de datos a adoptar las medidas tecnológicas necesarias para garantizar la seguridad de los datos y evitar su alteración, destrucción accidental o ilícita, pérdida, tratamiento o acceso no autorizado.



Por otro lado, la Ley 7975 de Información no divulgada protege la información privada relacionada con secretos comerciales e industriales. Dicha protección recae sobre la información que sea:

- a) secreta, en el sentido de que no sea generalmente conocida ni fácilmente accesible para las personas introducidas en los círculos donde normalmente se utiliza dicha información,
- b) esté legalmente bajo el control de una persona que haya adoptado medidas razonables y proporcionales para mantenerla secreta, y
- c) tenga un valor comercial.

Dicha protección no abarca aquella información que, al contrario, sea:

- a) del dominio público,
- b) resulte evidente para una persona técnica en la materia o
- c) deba ser divulgada por disposición legal u orden judicial, salvo en los casos exceptuados por ley.

Así las cosas, se debe verificar que la información contenida en las bases de datos no esté cubierta por acuerdos privados de confidencialidad o bien que de ella se desprenda que se han tomado medidas especiales para mantenerla secreta dado su valor comercial.

Por último, es importante resaltar que los procesos de apertura de datos y de protección de datos son complementarios, y no se puede garantizar una correcta apertura sin una adecuada protección de tales datos. En caso de incertidumbre sobre si algún conjunto de datos incumple con los estándares y normas de protección de datos personales, se recomienda buscar asesoramiento, particularmente con las asesorías jurídicas de la institución, así como recurrir a la Agencia de Protección de Datos de los Habitantes.



A fin de profundizar más en detalle el proceso de anonimización se recomienda consultar las normas y resoluciones emitidas por la Agencia de Protección de Datos de los Habitantes (PRODHAB), así como las otras herramientas anexas a esta guía.



5 Limpieza de los conjuntos de datos

La limpieza de datos es el proceso mediante el cual se busca:

1

Comprender qué datos son de carácter e interés público y pueden ser publicados sin restricciones y cuáles están sujetos a restricciones de privacidad, seguridad o derechos de autor.

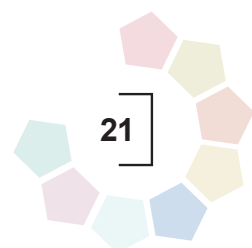
2

Normalizar y estandarizar la información a publicar.



Para profundizar más en el tema de limpieza de datos se recomienda verificar la Guía Quartz: Limpieza de datos, creada por *School of Data*¹, en la cual se puede verificar cómo resolver cualquier problema asociado a problemas técnicos dentro de la base de datos. Existen para este caso algunas herramientas informáticas dirigidas a ello, como es el caso de OpenRefine que nos permite actuar sobre los datos desorganizados.

¹ <https://schoolofdata.org>



6 Metadatos

Los metadatos describen los aspectos básicos de un conjunto de datos y son una herramienta fundamental para comprender la información.

Asimismo, los metadatos proporcionan información adicional que facilita a los consumidores entender mejor el significado de los datos, su estructura y aclarar otros asuntos, como derechos y términos de licencia, la organización de los datos y su calidad, los métodos de acceso y el calendario de actualizaciones de los conjuntos de datos. También, pueden ser de diferentes tipos, pueden clasificarse en diferentes taxonomías, con diferentes criterios de agrupación.

Por ejemplo, una taxonomía específica podría definir tres tipos de metadatos según características descriptivas, estructurales y administrativas. Una taxonomía diferente podría definir tipos de metadatos con un esquema según las tareas donde se usan los metadatos, como es el caso de reutilización y descubrimiento .

Estos campos pueden variar en detalle, desde descripciones muy básicas como explicar el tema general de la base de datos, hasta proveer detalle semántico para permitir un alto grado de legibilidad para máquinas. Lograr esto último incrementa en gran manera la apertura y utilidad de los datos.

Un ejemplo de ello, son los siguientes:

- ◆ **Nombre, tipo de datos y descripción de los campos:** ¿qué significa cada campo? ¿qué datos puedo encontrar en esa columna? ¿qué dicen y qué no dicen esos datos, cómo debo leerlos?
- ◆ **Palabras clave:** clasifican a un *dataset* como perteneciente a un conjunto de tópicos.
- ◆ **Tema:** clasifican a un *dataset* como perteneciente a un determinado tema, dentro de una jerarquía temática.
- ◆ **Fecha de publicación:** ¿cuándo se publicó por primera vez este *dataset*?
- ◆ **Fecha de última modificación:** ¿cuándo se actualizó por última vez este *dataset*?
- ◆ **Frecuencia de actualización:** ¿cada cuánto se actualiza este *dataset*?
- ◆ **URL de descarga:** ¿cómo dispongo de los datos, desde dónde puedo descargarlos?

La cantidad y calidad de metadatos asociados a un conjunto de datos, permite tanto una mejor clasificación como una mejor búsqueda y recuperación de los datos.

Algunos metadatos generales y obligatorios son los siguientes:

- 
- El título del conjunto de datos.
 - Una descripción del conjunto de datos.
 - Las palabras clave que describen el conjunto de datos.
 - La fecha de publicación del conjunto de datos.
 - La entidad responsable (editor) de hacer que el conjunto de datos esté disponible.
 - El punto de contacto para el conjunto de datos.
 - La cobertura espacial/geográfica del conjunto de datos.
 - El período temporal que cubre el conjunto de datos y la frecuencia de actualización.
 - La versión del conjunto de datos.
 - La fecha de la última modificación del conjunto de datos.
 - Los temas / categorías cubiertos por un conjunto de datos.
 - Licencia de uso asignada.
 - Recurso (enlace de distribución, tamaño, formatos, descripción, fecha de los metadatos).
 - Metadatos estructurales: descripción de las variables contenidas en el conjunto de datos.
 - Documentos (título, enlace de distribución, descripción, tamaño. formato)

7

Formato abierto, neutral, interoperable y georeferenciado

Para garantizar acceso universal e inclusión, así como una garantía innata de la transparencia de los datos, es necesario que los conjuntos de datos se encuentren en formatos abiertos. Esto garantiza que dichos datos sean neutrales e interoperables, así como georeferenciados cuando corresponda. Es decir, un formato de archivo que no suponga una limitación de acceso ya sea por la exclusividad de *software* en un sistema operativo, por restricciones legales, por limitaciones impuestas por el *software* con un fin en particular (educación, sector privado, sector público o desarrollo científico) así como por la carencia técnica del manejo interno de los datos, promoviendo así la utilización de los conjuntos de datos a grandes escalas. Los formatos abiertos garantizan la capacidad de integrarse y entrelazarse con otros datos sin modificar su individualidad e integridad.

En este caso, cuando se habla de datos abiertos, se promueve la generación y publicación de datos estructurados. Estos buscan transformar el dato de formatos textuales a formatos que permitan su manipulación y su conexión con otros datos (datos enlazados o *linked data*).

Los formatos libres corresponden a formatos de archivo que se pueden crear y manipular por distintos programas de *software*, disponibles en diferentes plataformas y que su implementación concreta en una plataforma específica no representa restricciones aumentando el nivel de reutilización de los datos.

Se puede disponer de los distintos formatos abiertos según cada tipo de documento e información contenida, por ejemplo:

Figura 8: Formatos abiertos según tipo de documento e información contenida

Texto	Hojas de cálculo	Geodatos y mapas
odt txt html odp rdf csv	ods * cvs * Como alternativa al XLS	gpkg * gpx GeoJSON GeoCVS * GeoPackage

Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.

Se recomienda evitar el uso de formatos de tipo propietario, es decir, que solo tienen un único proveedor de servicios, que no son auditables o representan una barrera para acceder, usar o visualizar los datos. Por ejemplo: .doc, .xls, .shp. Si bien es conocido que algunos de estos formatos son de uso común en la población, se recomienda manipular los datos desde un inicio en un formato abierto para que eso no conlleve pérdida en los metadatos de los datos o genere redundancia o duplicación de esfuerzos al tener que convertir manualmente los datos (ya sea por proceso automático o secuencialmente) para su uso o procesamiento, generando así obsolescencia en el uso y transparencia de los datos.

Asimismo, de acuerdo con el Informe *The State of Open Data 2019*² elaborado por el *Open Data for Development (OD4D)*, se estima que cerca del 80% de los datos de gobierno tienen un componente espacial o de tipo geográfico. Cuando los datos a publicar tienen este componente, se recomienda añadir y vincular el dato espacial al *dataset* a publicar, o bien generar un nuevo *dataset* de tipo espacial.

Por medio de colecciones de imágenes de datos satelitales, datos vectoriales espaciales y geodatabases se puede ofrecer este tipo de información para publicar o bien utilizar plataformas libres para la publicación y generación de capas de información, como por ejemplo las plataformas de mapeo libre y colaborativo.

Además mediante servicios de publicación de mapas en web, se permite la exploración visual de la información de una manera rápida y sencilla mediante una aplicación. Estos servicios deben permitir como mínimo mostrar, navegar y descargar los datos, así como acercarse o alejarse para concretar o ampliar el campo de visión, además de mostrar los atributos de los datos geográficos.

Algunas indicaciones para la apertura son:

- 1 **Aquellos datos que hagan referencia a espacios geográficos, deben contener identificadores geográficos o espaciales (georreferenciación). Asimismo, se recomienda utilizar los estándares del *Open Geospatial Consortium (OGC)*³.** 
- 2 **Cuando sea posible, generar una interfaz de programación de aplicaciones (API por sus siglas en inglés) facilita tanto la labor de la publicación como la del consumo de los datos al tener un punto centralizado y estandarizado de acceso a los datos. Esto facilita la interconexión e interacción funcional o la capacidad para intercambiar información con otros sistemas a través de sus interfaces.**

² <https://www.stateofopendata.od4d.net>

³ <https://www.ogc.org>



Asignación de la licencia de uso y reutilización

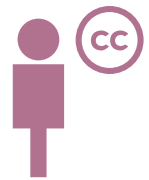
Es importante que la información publicada sea de libre uso, es decir que no tenga barreras para su uso, reuso y distribución. Esto define las condiciones bajo las cuales la institución encargada de los datos hace entrega de estos a las personas usuarias para su utilización.

Sin la licencia, los datos no pueden ser considerados "abiertos", dado que no existen garantías de que la utilización esté resguardada por los principios de la apertura de datos por sí mismos. En ese sentido, un metadato obligatorio que debe indicarse es la licencia de uso asignada al conjunto de datos. Para ello existen herramientas de uso libre que definen los alcances de la apertura.

De acuerdo a lo anterior, la fundación *Open Knowledge International* desarrolló las primeras licencias específicas para datos abiertos, las licencias *Open Data Commons*, que se pueden utilizar según se describen a continuación:

Figura 9: Licencias *Open Data Commons*

- ◆ *Open Data Commons Public Domain Dedication and Licence (PDDL)*, que permite difundir, reutilizar o adaptar los datos sin restricción alguna.
- ◆ *Open Data Commons Attribution License (ODC-BY)*, al igual que la licencia de Creative Commons Reconocimiento, se exige la referencia a la autoría o fuente de los datos para la reutilización de la información.
- ◆ *Open Data Commons Open Database License (ODbL)* que permite la reutilización de los datos siempre que se reconozca la autoría de la información original; se mantenga la misma licencia en las obras derivadas las cuales puede restringir su uso si, además, se distribuye una versión sin dichas restricciones.
- ◆ *Creative Commons Attribution 4.0 (CC-BY-4.0)* esta licencia posibilita no solo copiar y redistribuir el material en cualquier medio o formato, sino también adaptar y transformarlo para cualquier propósito, incluso comercial.



Una breve guía para utilizar una licencia para datos abiertos puede ser encontrada en el sitio *Open Data Commons*: <http://opendatacommons.org/guide/>

Sea para un portal o para un conjunto de datos en particular, se debe especificar el tipo de licencia que le aplica. Para esto, la recomendación es utilizar una que permita la reutilización de los datos, siempre que se reconozca la autoría de la información original y se mantenga la misma licencia en las obras derivadas. Una licencia que cumple con las recomendaciones anteriores es la *Open Data Commons Open Database License (ODbL)*, sin embargo dependiendo del tipo de base de datos a liberar puede utilizarse alguna de las anteriores.

9 Publicación de los datos

La publicación de datos se lleva a cabo a través de su carga en una plataforma que le permita a la institución un fácil acceso por parte de quienes van a reutilizar los datos, de manera que se convierta en un sitio de acceso común y cuente con la posibilidad de hacer escalable la información publicada.



El inciso b. del artículo 11 del decreto 40199, indica que todos los catálogos de datos de las instituciones deberán ser publicados en un portal nacional y mantiene la libertad para que las instituciones publiquen en portales institucionales.

En este aspecto se motiva a las instituciones a tomar una posición proactiva en la generación y publicación de conjuntos de datos de valor público desde sus propios espacios web.





Comunicación y promoción de los datos

El éxito de un proyecto de apertura depende en mucho de las personas que consumen la información. Si los datos no se usan, difunden y generan valor público, la apertura se convierte en un ejercicio poco efectivo.

Es fundamental que los datos puedan ser encontrados con facilidad y que exista una comunidad que activamente los utilice.

Quienes lleven a cabo los procesos de apertura deben generar mecanismos de comunicación, participación y realimentación, así como espacios de uso y reuso con la sociedad y quienes componen el ecosistema de datos.

El enfoque debe ser coherente, estar alineado con la visión institucional y debe estar incorporado claramente en los planes y programas de comunicación y promoción institucional, de modo que se visualice desde las primeras etapas de construcción.

Comunicación

Este componente contiene el conjunto de comunicaciones internas y externas de la entidad, para sensibilizar y promover el uso de datos abiertos y sus impactos en la sociedad. Asimismo, la vinculación en temas de transparencia y toma de decisiones, desarrollo de aplicaciones, periodismo de datos, mejora de servicios, entre otros.

Al respecto la institución puede desarrollar estrategias de comunicación tanto interna como externa, así como hacer uso de redes sociales para difundir estos mensajes.

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

- ◆ Generar mensajes internos o externos vía mensajería instantánea y correo electrónico.
- ◆ Realizar encuentros o reuniones internas y externas.
- ◆ Desarrollar campañas informativas con medios de comunicación para dar a conocer beneficios con el uso de los datos abiertos de la entidad.
- ◆ Realizar campañas de comunicación a través de redes sociales, así como comunicados específicos cada vez que se logra publicar un conjunto de datos.



Promoción del uso

Cuando abrimos los datos, es necesario construir comunidades que hagan uso de la información. Por ejemplo, desarrolladores o visualizadores son esenciales en la primera etapa de apertura de datos, ya que a través de ellos se pueden crear grandes herramientas que, a su vez, lleguen al público en general. Junto a esto, existen organizaciones, comunidades y especialistas en el uso y explotación de los datos abiertos, como por ejemplo investigadores (as), periodistas de datos y comunidades de desarrollo informático.

A fin de mantener activo este ecosistema de uso, se proponen las siguientes actividades:

Figura 10: Actividades para promoción y comunicación

Iniciativa: Ideatones	
¿En qué consiste?	Sesiones colectivas, de preferencia con participantes multidisciplinarios, en donde se analizan datos sobre una problemática en particular con tal de generar ideas que contribuyan en su solución; idealmente, las ideas con mayor potencial de impacto y viabilidad de ejecución deben ser apoyadas, posteriormente para profundizar en análisis, realizar prototipos y ponerlas en práctica en la vida cotidiana.
Resultado esperado	Las ideas generadas pueden abarcar desde propuestas de políticas públicas, intervenciones sociales hasta potenciales aplicaciones tecnológicas.
Impacto en la comunidad	Convivencia, análisis y aprendizaje mutuo entre especialistas multidisciplinarios (analistas de datos, visualizadores, expertos temáticos, programadores, trabajadores sociales, funcionarios públicos, etc.); generación de alianzas multidisciplinarias en relación a una problemática específica; involucramiento de análisis basado en datos entre sectores interesados e inclusive especializados sin uso de datos.
Iniciativa: Mapatones	
¿En qué consiste?	Son sesiones colectivas de levantamiento de datos geoespaciales.
Resultado esperado	La recopilación de datos geoespaciales relevantes para ciertas comunidades o temas que previamente eran inaccesibles.
Impacto en la comunidad	Generan el involucramiento de individuos con diversas capacidades técnicas; los participantes incrementan su solidaridad con la causa para la que los datos son generados.
Iniciativa: Expediciones de datos	
¿En qué consiste?	Sesiones colectivas de preferencia multidisciplinarias en donde grupos de hasta 8 personas analizan datos relacionados con alguna problemática o situación de su interés con tal de llegar a conclusiones sobre la misma.
Resultado esperado	Análisis detallado basado en datos sobre un tema, situación o problemática; aprendizaje de tácticas y herramientas para el análisis de datos.
Impacto en la comunidad	Socialización de conjuntos de datos, iniciativas y vinculación entre comunidades y perfiles.

Figura 11: Actividades para promoción y comunicación

Iniciativa: Hackatones	
¿En qué consiste?	Espacios de hasta 48 horas en donde participantes (de preferencia multidisciplinarios) realizan soluciones tecnológicas basadas en datos; comúnmente se otorgan premios para apoyar que las soluciones ganadoras puedan ser terminadas y llevadas a usuarios finales. En un escenario óptimo, los equipos ganadores son incubados y apoyados por organizaciones clave para la ejecución del proyecto.
Resultado esperado	Visualizaciones, plataformas de información, sitios web, aplicaciones móviles, servicios web de datos, dispositivos mecatrónicos para generación de datos, etc.; aprendizaje entre los participantes en torno a trabajo en equipo, acceso y procesamiento de datos, uso de plataformas tecnológicas, interfaces de usuario y definición de productos basados en datos.
Impacto en la comunidad	Aumento de colaboración, aprendizaje mutuo y sinergias entre miembros de equipos; competencia creativa, tecnológica y analítica entre equipos de participantes; alianzas posteriores al evento para realizar otros proyectos.
Iniciativa: Mediatones	
¿En qué consiste?	Espacios de hasta 48 horas en donde participantes (de preferencia multidisciplinarios) realizan proyectos comunicacionales basadas en datos ya sea en periodismo, activismo o comunicación social; comúnmente se otorgan premios para apoyar que los proyectos ganadores puedan ser terminados, publicados alcanzando a sus audiencias meta. En un escenario óptimo, los equipos ganadores son incubados por especialistas en comunicación, periodismo, datos y tecnología y se cuenta con el compromiso por parte de medios y organizaciones para la publicación de proyectos.
Resultado esperado	Reportajes periodísticos, visualizaciones, campañas de incidencia y acciones de comunicación social; aprendizaje entre los participantes en torno a trabajo en equipo, acceso, procesamiento y visualización de datos, uso de plataformas tecnológicas, interfaces de usuario, diseño gráfico, diseño de la información y definición de productos basados en datos.
Impacto en la comunidad	Aumento de colaboración, aprendizaje mutuo y sinergias entre miembros de equipos; competencia creativa, tecnológica y analítica entre equipos de participantes; alianzas posteriores al evento para realizar otros proyectos; aumento de la cultura de datos en organizaciones de la sociedad civil, colectivos sociales y periodismo.

Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.

h Sostenibilidad

Un aspecto importante a tomar en cuenta luego de publicar datos es el mantenimiento de históricos en el portal y la actualización de estos datos de acuerdo a su periodicidad de recolección. Por ello se propone utilizar la ficha detallada a continuación, para asegurar un efectivo proceso de actualización.

Figura 12: Plantilla para proceso de actualización de datos

Título del conjunto de datos	Descripción del conjunto de datos	Fecha de publicación del conjunto de datos	Responsable de la publicación del conjunto de datos	Período temporal que cubre el conjunto de datos	Frecuencia de actualización del conjunto de datos

Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.

Con esto disponemos de una herramienta que nos determine la periodicidad y el procedimiento de actualización de los datos en el portal. Esto es un elemento que mejora la gestión de la institución tanto a lo interno como a lo externo.

i Continuidad de la iniciativa

Con el objetivo de orientar acciones posteriores al proceso de apertura de datos públicos, se debe construir una ruta de trabajo para definir proyectos y acciones necesarias para mejorar la facilitación de datos abiertos a la ciudadanía.

Para facilitar este proceso se diseñó la siguiente plantilla para la construcción de compromisos institucionales de apertura de datos. Un compromiso se presenta como un proyecto con objetivos y metas concretas que aporta a que las instituciones avancen en uno o más de los principios de los datos abiertos.

Las siguientes tablas plantean elementos que resulta importante considerar en la propuesta de planificación de un proyecto de apertura de datos. Sugerimos completar las tablas para cada uno de los proyectos o compromisos que defina la institución. El conjunto de proyectos llevará a la formulación del plan de acción institucional.

Figura 13: Plantilla para proyectos institucionales en apertura de datos		
Nombre de la institución:		
Departamento responsable		
Persona responsable		
Cargo y departamento		
Correo electrónico		
Otros participantes y sociedades	Gobierno	
	Sector privado, sociedad civil, otros	
Estatus - problemática a solucionar que aborda el compromiso		
Objetivo principal		
Descripción del compromiso (140 caracteres)		
Hito:		Fecha de Inicio / Fecha de finalización:

Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.

j Monitoreo y evaluación

Las tareas de monitorear y evaluar son importantes porque permiten a las partes interesadas:

- ◆ Comprender el progreso de la iniciativa
- ◆ Aprender de la experiencia y generar nuevas capacidades
- ◆ Identificar oportunidades de mejora
- ◆ Mejorar la responsabilidad de la puesta en marcha y el uso de los recursos
- ◆ Revisar y probar los supuestos que se identificaron al inicio de la iniciativa
- ◆ Mejorar procesos y políticas
- ◆ Comprender mejor las relaciones entre la publicación y reutilización de datos
- ◆ Garantizar una memoria institucional
- ◆ Generar evidencia para obtener apoyo para continuar con la ejecución



Por esta razón, es importante definir diferentes conjuntos de indicadores. Cada objetivo y compromiso podría requerir su propio conjunto de indicadores cualitativos y cuantitativos para garantizar que la puesta en marcha de un plan se mueva en la dirección correcta. Los indicadores que sean definidos también deben responder a las diferentes etapas de la implementación para tomar acciones correctivas según sea necesario.



Se recomienda que el conjunto de indicadores de monitoreo y evaluación se construya de forma colaborativa y se acuerde con las partes interesadas. Esto permitirá un mayor apoyo para las acciones de ejecución, evaluación y respuesta. Del mismo modo, las partes interesadas también deberán participar en el proceso de evaluación y aceptar asumir la responsabilidad de corregir lo que se ha acordado como necesario.

Incluso cuando exista una alta participación de los interesados en el monitoreo y evaluación, la retroalimentación que pueda provenir de otras partes (es decir, actores externos al proceso) no debe limitarse. Esta retroalimentación puede jugar un papel importante al realizar el ejercicio de monitoreo y evaluación. Por lo tanto, dar y recibir comentarios constructivos debe ser alentado y aceptado.

Al seleccionar los diferentes indicadores, es importante tener en cuenta que existen indicadores para diferentes propósitos. Algunos de ellos apuntan a medir entradas; otros miden actividades.

La combinación de indicadores de todas las siguientes categorías ayuda a realizar una evaluación integral en todas las etapas:

Figura 14: Propuesta para medición y monitoreo de acciones en apertura de datos

Tipo de indicador	Según recursos	Según actividades	Según productos	Según resultados	Según impacto
Responder a la pregunta	¿Qué recursos son requeridos?	¿Qué actividades son requeridas?	¿Qué productos se obtienen del proyecto?	¿Cuáles son los beneficios de corto-mediano plazo?	¿Cómo contribuye el proyecto a los objetivos estratégicos?
El indicador permite	Evaluar si los recursos asignados son utilizados de acuerdo con lo definido	Evaluar las actividades planificadas que son efectivamente ejecutadas	Evaluar que los recursos y las actividades generan los productos esperados. Si no es así, es necesario hacer cambios	Evaluar si los productos generan los resultados esperados. Si no es así, es necesario hacer cambios	Evaluar si los resultados contribuyen a alcanzar los objetivos estratégicos. Si no es así, es necesario hacer cambios
Ejemplo	Estimación de presupuesto para un taller de capacitación en datos abiertos	Organizar el taller para el desarrollo de capacidades de datos abiertos	Número de funcionarios públicos capacitados	Número de conjuntos de datos abiertos tras consultas a la ciudadanía	Mayor apertura de datos públicos que generan valor económico

Fuente: Comisión Nacional de Datos Abiertos, Guía Nacional para la Publicación de Datos Abiertos.



Comisión Nacional de Datos Abiertos

Ministerio de Comunicación	Geannina Sojo Navarro Sigrid Segura Artavia
Ministerio de Ciencia, Tecnología y Telecomunicaciones	Diego Vargas Pérez Vivian Aguilar Aguilar
Ministerio de Planificación Nacional y Política Económica	Yazmín Montoya Jiménez Iván Acuña Chaverri
Dirección General del Archivo Nacional	Ivannia Valverde Guevara Denise Calvo López
Instituto Nacional de Estadística y Censos	Ronny Umaña Palma Pilar Ramos Vargas
Sociedad civil	Susana Soto González
Consejo Nacional de Rectores	Dennis Víquez Ruíz Marco Monge Vílchez
Sector privado	Guillermo Rodríguez Zúñiga Simón Echavarría Palacio
Sector privado	David Zamora Barrantes Jorge Umaña Cubillo

Autoridades

Presidente de la República de Costa Rica
Carlos Alvarado Quesada

Ministra de Comunicación
Agustín Castro Solano

Contacto

datosabiertos@presidencia.go.cr

Coordinación

Coordinadora de Gobierno Abierto
Geannina Sojo Navarro

Asesora en Datos Abiertos
Sigrid Segura Artavia

Diagramación: Federico Espinosa Porras

Esta obra está sujeta a la licencia
Reconocimiento-NoComercial-Compartir
Igual 4.0 Internacional de Creative
Commons.
Para ver una copia de esta licencia, visite
<http://creativecommons.org/licenses/by-nc-sa/4.0/>.



I Glosario

Armonización: aplicación de estructuras comunes, previamente definidas y acordadas para facilitar la integración de conjuntos de datos.

Bases de datos: conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso.

Catálogo de datos abiertos: el inventario de los conjuntos de datos puestos a disposición en un portal de datos abiertos disponible en la web.

Conjunto de datos: la serie de datos estructurados, vinculados entre sí y agrupados dentro de una misma unidad temática, de forma que puedan ser procesados apropiadamente para obtener información.

Dato: representación simbólica (numérica, alfabética, algorítmica, espacial, etc.) de un atributo o variable cuantitativa o cualitativa.

Dato público: cualquier tipo de dato de interés público que sea generado o resguardado por quien ejerza una función o potestad pública y que no tenga su acceso restringido por ley.

Dato abierto: dato disponible en línea, sin procesar, en formato abierto, neutral e interoperable, que permite su uso y reuso, disponible para su descarga en forma completa, sin costo ni requisitos de registro y procesable en computadora. Deberá cumplir con las siguientes características:

Completo: todo dato público debe estar disponible sin más excepciones que las establecidas por la ley.

Primario: el dato público debe ser puesto a disposición desde la fuente, con el mayor nivel posible de desagregación y sin modificar.

Actualizado y oportuno: el dato público debe ser puesto a disposición en forma pronta, según su periodo de generación para preservar su valor de reuso.

Accesible: el dato público debe encontrarse fácilmente disponible para todas las personas y para cualquier uso no prohibido por ley.

Procesable automáticamente: el dato público debe estar estructurado para permitir el procesamiento automático por computadoras.

No discriminatorio: el dato público debe estar disponible para cualquier persona sin requisitos de registro previo.

No propietario: el dato público debe estar disponible en formatos neutrales e interoperables.

De libre uso y reuso (conocido como licencia de uso y reuso de datos): el dato público debe estar libre de cualquier restricción de derechos de autor, propiedad intelectual o industrial, o de información comercial no divulgada. Puede tener restricciones razonables de privacidad o por niveles de usuario, acorde a la ley.

Datos Personales: cualquier dato relativo a una persona física identificada o identificable.

Datos personales de acceso irrestricto: los contenidos en bases de datos públicas de acceso general, según dispongan leyes especiales y de conformidad con la finalidad para la cual estos datos fueron recabados.

Datos personales de acceso restringido: los que, aun formando parte de registros de acceso al público, no son de acceso irrestricto por ser de interés solo para su titular o para la Administración Pública.

Datos sensibles: información relativa al fuero íntimo de la persona, como por ejemplo los que revelen origen racial, opiniones políticas, convicciones religiosas o espirituales, condición socioeconómica, información biomédica o genética, vida y orientación sexual, entre otros.

Ecosistema: Conjunto de actores institucionales, del sector empresarial, la academia y la ciudadanía que interactúan en los procesos de apertura y uso de los datos abiertos.

Formato abierto, neutral e interoperable: Es el formato de archivo que no supone una dificultad de acceso, que está libre de cualquier restricción legal o económica para su utilización y que tiene la capacidad de integrarse y entrelazarse con otros datos.

Guía Nacional para la Publicación de Datos Abiertos: documento de orientación para la ejecución de los procesos de apertura de datos.

Interoperabilidad tecnológica: es la capacidad de generar la interconexión e interacción funcional o la capacidad para intercambiar información con otros sistemas a través de sus interfaces.

Licencia para uso y reuso de datos abiertos: documento que especifica el derecho de uso de los datos públicos abiertos. Establece y garantiza condiciones para el acceso, la redistribución, el reuso, la ausencia de restricciones tecnológicas, la atribución de origen, la no discriminación de grupos o personas y la integridad de los datos.

Metadatos: son los datos estructurados y actualizados que describen el contexto y las características de atribución, contenido, captura, procesamiento, calidad, condición, acceso, distribución u otras de un conjunto de datos; que faciliten su búsqueda, identificación y uso.

Plan de acción de datos abiertos: documento que reúne las acciones concretas y compromisos para la implementación de los procesos de apertura de datos y la promoción de su uso.

Portal de datos abiertos: Plataforma digital, neutral e interoperable que sirven para almacenar, compartir y conectar los catálogos de datos y todos los recursos destinados a facilitar su uso, tales como: aplicaciones, herramientas de visualización, simulaciones, normativas o guías.



Fuentes de referencia

Asamblea Legislativa. (2011). Ley de Protección de la Persona frente al tratamiento de sus datos personales. Ley No. 8968. Recuperado de http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=70975&nValor3=85989&strTipM=TC

Banco Mundial. (2019). Caja de Herramientas de Datos Abiertos de Gobierno. Recuperado de <http://opendatatoolkit.worldbank.org>

Open Data for Development. (2019). Sector Geospatial, State of Open Data. Recuperado de <https://stateofopendata.od4d.net>

Open Knowledge Foundation. (s.f). Open Data Commons. Recuperado de <https://opendatacommons.org>

Open Knowledge Foundation. (s.f). Open Data Handbook. Recuperado de <http://opendatahandbook.org>

Open Knowledge Foundation. (s.f). Open Definition. Recuperado de <http://opendefinition.org>

Poder Ejecutivo. (2017). Decreto Ejecutivo No. 40199-MP Apertura de Datos Públicos. Recuperado de http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=84004&nValor3=108193&strTipM=TC

Poder Ejecutivo. (2017). Directriz Ejecutiva No. 074-MP Sobre la apertura de Datos Abiertos. Recuperado de http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=84012&nValor3=108205&strTipM=TC

The International Open Data Charter. (2015). Open Data Charter. Recuperado de https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf The World Wide Web Consortium (W3C). (2017). Data on the Web Best Practices. Recuperado de <https://www.w3.org/TR/dwbp/>

United Nations. (2017). Guide on Lessons for Open Government Data Action Planning for Sustainable Development. Recuperado de <https://www.un.org/development/desa/capacity-development/2018/01/11/guide-on-lessons-for-open-government-data-action-planning-for-sustainable-development/>

datosabiertos.go.cr/guia



**Guía Nacional para la
Publicación de Datos
Abiertos**