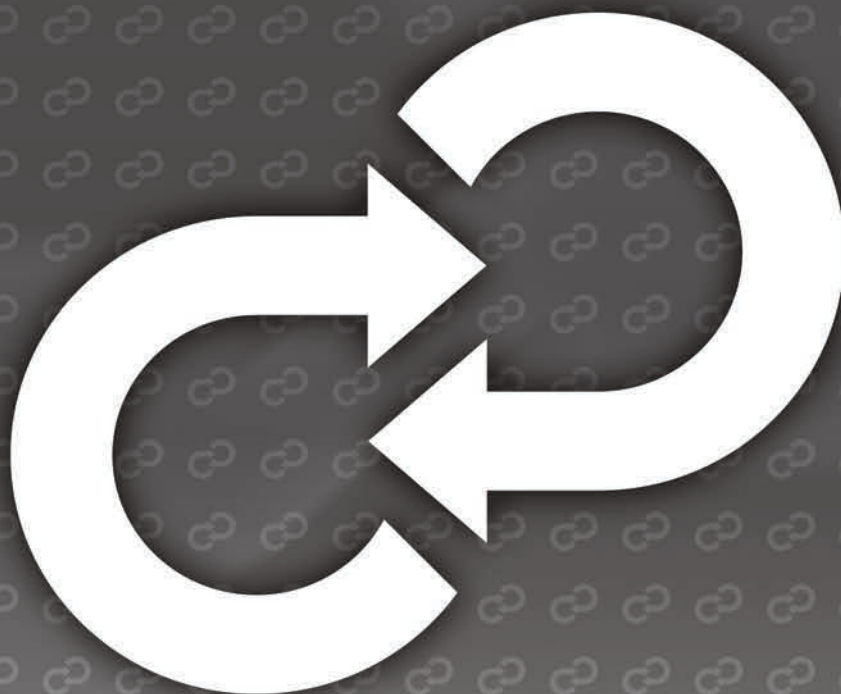


Guía rápida de publicación en datos.gub.uy



GUIA RAPIDA DE PUBLICACION EN

datos.gub.uy

Publicación en el Catálogo Nacional de Datos Abiertos

Versión 1.0 – 2012

Este documento ha sido elaborado por AGESIC (Agencia para el Desarrollo del Gobierno de Gestión Electrónica y la Sociedad de la Información y el Conocimiento)

Usted es libre de copiar, distribuir, comunicar y difundir públicamente este documento así como hacer obras derivadas, siempre y cuando tengan en cuenta citar la obra de forma específica y no utilizar esta obra para fines comerciales. Toda obra derivada de esta deberá ser generada con estas mismas condiciones.

De datos públicos a datos abiertos.

*Los organismos estatales del sector público recogen, producen, reproducen y difunden datos para cumplir con su función pública. Algunos de estos datos ya están publicados o debieran estarlo. Incorporar la publicación de **datos públicos en formatos abiertos** abre la puerta a la posibilidad de que los mismos sean reutilizados en nuevos proyectos, que puedan combinarse con otras fuentes de datos y generar nuevas aplicaciones desarrolladas por el gobierno, por la sociedad civil, organizaciones, empresas o ciudadanos en general.*

Actualmente no es posible acceder, desde un único punto, a los distintos sitios Web y recursos del Estado que ofrecen información pública. Además esta información es publicada siguiendo pautas particulares de cada organismo. Es por eso que la posibilidad de contar con un Catálogo Nacional de Datos Abiertos que presente la información de forma centralizada, otorgándole mayor visibilidad, facilidad de uso y acceso es un beneficio tanto para el Estado como para los ciudadanos.

Este documento fue pensado como una guía de referencia rápida para conocer el proceso de publicación de conjuntos de datos en el catálogo nacional de datos abiertos datos.gub.uy. La misma reúne conceptos básicos de datos abiertos de gobierno y los pasos a seguir a la hora de catalogar.

CAPÍTULO I

Introducción

Introducción

¿Qué son los Datos Abiertos?

Datos Abiertos son aquellos datos que están disponibles libremente para su utilización, reutilización y redistribución.

Los Datos Abiertos deben cumplir con las siguientes premisas.

- **Disponibilidad y acceso:** Los datos deben estar disponibles un costo razonable y de una forma conveniente y modificable.
- **Reutilización y redistribución:** El formato de los datos permitirá su reutilización, redistribución e integración a otros conjuntos de datos.
- **Participación universal:** Todas las personas deben poder utilizar, reutilizar y redistribuir la información sin restricciones.

En diciembre de 2007 la organización Opengovdata publicó ocho principios¹ que deben ser considerados por los gobiernos a la hora de publicar datos. Estos principios no indican solamente cuáles datos deberán ser públicos, sino que especifican las condiciones para que los datos públicos sean, además, abiertos.

1. Los datos deben ser completos

Todos los datos estarán disponibles. Los datos se almacenan electrónicamente incluyendo documentos, bases de datos, transcripciones, grabaciones de audio y video, imágenes, etc. Los datos públicos son datos que no están sujetos a las limitaciones vigentes de privacidad, seguridad o privilegio, que se rige por otras leyes.

2. Los datos deben ser primarios

Los datos se publican tal como se extraen de la fuente de origen, con el mejor nivel posible de granularidad, no en formas agregadas o modificadas.

3. Los datos deben ser oportunos

Los datos se hacen disponibles tan pronto como sea necesario para preservar el valor de los mismos.

1 <http://www.opengovdata.org/home/8principles>

4. Los datos deben ser accesibles

Los datos están disponibles para la gama más amplia de usuarios y para la gama más amplia de propósitos.

5. Los datos deben ser procesables por máquina

Los datos estarán razonablemente estructurados para permitir el procesamiento automático de los mismos.

6. El acceso no debe ser discriminatorio

Los datos están disponibles para cualquier persona, sin necesidad de registro.

7. El formato de los datos no debe ser propietario

Los datos están disponibles en un formato sobre el cual ninguna entidad tiene control exclusivo.

8. Los datos deben ser libres de licencia

Los datos no están sujetos a ningún derecho de autor, patentes, marcas o reglamento de secreto comercial. Restricciones razonables de privacidad, seguridad y de privilegios se pueden permitir ya que se rigen por otras leyes.

¿Qué datos de gobierno podrían ser abiertos?

Los datos públicos de gobierno establecidos en la Ley N° 18.381 de Acceso a la Información Pública. El punto clave a destacar es que la apertura de datos se centra en los datos no personales que no están sujetos a restricciones específicas.

Reglas básicas para iniciar la apertura de datos

- Cumplir con la mayor cantidad de los ocho principios nombrados anteriormente.
- Abrir conjuntos de datos pequeños y sencillos; incluso pueden ser parte de un conjunto de datos más amplio.
- Elegir información que no insuma demasiados recursos para su elaboración, utilizar información que esté siendo distribuida en otros formatos y hacerla disponible en formatos abiertos, es la forma más sencilla de comenzar con la apertura de datos.
- Al seleccionar un conjunto de datos tener en cuenta el público objetivo. A veces la información importante para la organización no lo es para los reutilizadores o posibles consumidores.
- Observar que se está haciendo en otros organismos similares.
- Aplicar una licencia abierta. En caso de no asignar una licencia abierta al conjunto de datos, se asume que el conjunto de datos se adapta a la licencia de datos abiertos del catálogo.

Catálogo de datos abiertos

¿Qué es?

Es un portal que busca facilitar el acceso de los ciudadanos a los datos públicos, fomentar su reutilización y el desarrollo de nuevos servicios y productos.

Pretende incentivar a los organismos a publicar sus datos en formatos abiertos y ofrecer espacios de participación y de colaboración ciudadana. La herramienta utilizada para gestionar el catálogo es CKAN² y actualmente es utilizada por varios países en iniciativas de similares características.

² <http://ckan.org>

¿Qué beneficios brinda?

- Permite acceder, desde un único punto, a los distintos sitios Web y recursos del Estado que ofrecen información pública.
- Presenta la información de forma centralizada, brindando mayor visibilidad, facilidad de uso y acceso a la misma.
- Fomenta la innovación y posibilita la utilización de datos abiertos en el desarrollo de aplicaciones por empresas u organizaciones sociales.
- Promueve la participación ciudadana.
- Mejora aspectos de transparencia en los procesos de gobierno.
- Facilita la interoperabilidad entre diferentes organismos públicos.

¿Cómo se puede acceder?

Para navegar se podrá acceder al Catálogo Nacional de Datos Abiertos de Gobierno a través de datos.gub.uy o directamente desde catalogodatos.gub.uy

Para poder añadir conjuntos de datos a dicho catálogo, el organismo necesita solicitar previamente un usuario. Podrá hacerlo enviando el pedido a datosabiertos@agesic.gub.uy proporcionando la siguiente información.

- Organismo al que pertenece.
- Nombre completo.
- Correo electrónico.

La información para acceder al sitio será enviada por correo electrónico a cada usuario con los siguientes datos:

- La o las URL de conexión.
- Usuario
- Contraseña

Para cada organismo se creará un grupo u organización de manera que los usuarios solo puedan acceder y editar los conjuntos de datos de su organismo.

CAPÍTULO II

Publicación de Datos

Publicación en el catálogo: paso a paso

Clasificación del nivel de apertura de los datos

Existe una clasificación del nivel de apertura de los datos basado en la cantidad de estrellas³.

1 estrella: Los datos deben estar disponibles en la Web en cualquier formato y bajo una licencia abierta, para que sean considerados como datos abiertos.

Ejemplo: .PDF, .JPG

2 estrellas: Los datos deben estar estructurados y en un formato que pueda ser interpretado o procesado por máquinas.

Ejemplo: .XLS

3 estrellas: Igual al anterior pero en un formato no propietario.

Ejemplo: .CSV, XML

4 estrellas: Todo lo anterior más la utilización de estándares recomendados por W3C para identificar cosas.

Ejemplo: RDF + SPARQL

5 estrellas: Todo lo anterior más el establecimiento de vínculos con los datos de otros para proveerles contexto.

Ejemplo: Linked data

³ <http://www.w3.org/DesignIssues/LinkedData.html>

Preparación del conjunto de datos

La preparación del conjunto de datos a publicar exige seguir los siguientes pasos.

- Identificar los archivos de datos o recursos a publicar y que formaran parte del conjunto de datos.
- Generar información de metadatos sobre el conjunto de datos a publicar.
- Asignar una licencia de datos abiertos (opcional)
- Empaquetar los archivos.

Identificación de los datos: Lo principal para comenzar es identificar los datos que tenemos publicados o que son de fácil transformación. Para comenzar este proceso pueden realizarse las siguientes preguntas:

- ¿Qué información le solicitan habitualmente y qué exige un procesamiento de datos?
- ¿Qué información existe en su oficina con datos que pueden ser de interés para la población?
- ¿Qué información se brinda habitualmente a la prensa, o se publica en medios gráficos que puede ser reutilizada?
- ¿Qué información se brinda a otros organismos que puede ser de interés general?

Una vez que haya identificado la información candidata a ser publicada, debe seleccionar la que se encuentre en condiciones de ser publicada. Es necesario considerar aspectos legales, de completitud, capacidad para mantenerla actualizada, formato, etc.

Para facilitar la reutilización de la información se recomienda la publicación en un formato que pueda ser procesado de manera automatizada. Los formatos pueden ser diversos y varían de acuerdo al nivel de apertura al que se quiera llegar.

Para un nivel inicial se aconseja llegar a un nivel de 3 estrellas.

Para ello puede publicar en un formato de texto separado por comas, remover información de títulos, subtítulos, etc. dejando solamente información de nombres de columnas.

Generar Metadatos⁴: En pro de la utilización de los datos por parte de terceros se le debe de proveer de información adicional que describa la información que se publica. Para el caso de estructuras relacionales u hojas de cálculo puede ser la descripción de las columnas; para archivos XML puede ser el DTD, etc. Dependiendo del caso bastaría con asociar un archivo de texto explicando el significado de cada campo o columna. A continuación se sugiere un formato básico para presentar esta información en esta primera fase.

Nombre de atributo	Descripción	Tipo de dato	Recurso relacionado	Información Adicional
<Descripción exacta de la columna tal cual aparece en los archivos de los conjuntos de datos>. Ejemplo: CodigoDepto	<Descripción detallada del significado del atributo orientada a las personas que pretendan utilizar el atributo>. Ejemplo: Departamento de Residencia	<Numérico, Texto, Fecha/Hora, Moneda> Ejemplo: Numérico	<Recurso que da significado al atributo. Ejemplo: si el atributo es un código tiene como recurso relacionado la tabla o recurso donde se describe el significado de dicho código.> Ejemplo: Departamentos.xls	<Cualquier información adicional sobre el atributo que complemente la información básica.> Ejemplo: En el archivo Departamento.xls están los nombres de los departamentos asociados a dicho código. Ejemplo: 1- Montevideo; 2- Soriano; 3- Paysandu....

Licenciamiento de los datos: En el caso de que ya exista información publicada en los sitios Web del organismo, ya sea en formatos propietarios (como pueden ser .PDF, .DOC, .XLS por nombrar los más usuales) o simplemente archivos de texto plano (.TXT), se debe aplicar una licencia de datos abiertos para estar en el nivel de 1 estrella.

En el Portal se ofrece una licencia de datos abiertos de gobierno de Uruguay y es la recomendada por AGESIC. Esta licencia será la que aparecerá por defecto en el momento de la catalogación. Además aparecen otras opciones que podrán ser seleccionadas si el organismo no adhiere a la licencia por defecto.

Empaquetado: Teniendo el archivo de datos en un formato abierto y el archivo de metadatos correspondiente; restaría empaquetar toda la información en forma de archivo comprimido para su publicación..

⁴ http://www.agesic.gub.uy/innovaportal/v/2256/1/agesic/metadatos:_especificacion_tecnica_disponible.html

Publicación en el catálogo

El acceso al catálogo deberá hacerse utilizando cualquiera de las URL que se habilitaron para ello. Esta información se debió recibir previamente por correo electrónico en respuesta al proceso de solicitud de acceso al catálogo junto con el usuario y la contraseña.

Antes de comenzar con la publicación de datos en el catálogo es necesario tener en cuenta la información que deberá ser ingresada obligatoriamente al agregar un nuevo conjunto de datos.

Un conjunto de datos está formado por uno o varios recursos. Los recursos son los archivos de datos, metadatos y licencia que va a contener el conjunto de datos.

Agregar conjunto de datos

Para agregar un conjunto de datos debe ingresar por la opción “agregar un nuevo conjunto de datos”, donde se solicitarán los siguientes metadatos como obligatorios:

- **Título:** nombre genérico que se le dará al conjunto de datos.

Recomendado: Accidentes en rutas nacionales	No recomendado: Accidentes en rutas 2010
---	--

- **URL:** se sugiere en base al título, la misma puede ser modificada en caso de que ya exista.
- **Licencia:** se debe seleccionar una licencia de la lista. Por defecto aparecerá la licencia de datos abiertos de gobierno de Uruguay.
- **Descripción:** breve descripción de la información contenida en el conjunto de datos.
- **Organización/Grupo:** debe seleccionar el grupo u organización al que pertenece el conjunto de datos, para el caso de organización también se exige seleccionar si el conjunto de datos va a ser público o privado.

Una vez finalizado el ingreso de los datos generales para el conjunto de datos, deben agregarse los recursos que lo conformarán.

Agregar recursos

El ingreso de recursos se puede hacer de tres formas:

- Enlazando a un archivo.
- Enlazando a una API.
- Subiendo el archivo.

Las dos primeras opciones son válidas para organismos que tienen sus datos publicados en sus portales y en el catálogo solo crean un enlace a dichos recursos.

Para los organismos que no cuentan con un espacio en sus portales para la publicación de datos, es posible subir los recursos de manera que queden alojados dentro del catálogo.

Para cada recurso se solicitarán los siguientes metadatos:

- **Nombre:** nombre del recurso.

Recomendado: Acc_rutas_nac_2012.xml	No recomendado: Acc_rutas.xml
--	-------------------------------

- **Descripción:** breve descripción de los datos del recurso.
- **URL:** se crea en base a cómo se publica el recurso.
- **Formato:** formato del archivo, .csv, .odt, .dbf, etc.
- **Tipo de recurso:** se debe seleccionar de la lista desplegable.
- **Última modificación:** se debe ingresar la fecha de última modificación del recurso.
- **Frecuencia de actualización:** Debe indicar la frecuencia con que se publicarán los datos. Esta información es muy importante para los reutilizadores y exige un compromiso de parte del organismo publicador. Este campo no estará disponible en las primeras versiones del sitio de pruebas.

Tabla de contenidos

Contenido

Introducción	4
¿Qué son los Datos Abiertos?	4
¿Qué datos de gobierno podrían ser abiertos?	6
Reglas básicas para iniciar la apertura de datos	6
Catálogo de datos abiertos	6
¿Qué es?	6
¿Qué beneficios brinda?	7
¿Cómo se puede acceder?	7
Publicación en el catálogo: paso a paso	11
Clasificación del nivel de apertura de los datos	11
Preparación del conjunto de datos	12
Publicación en el catálogo	14
Agregar conjunto de datos	14
Agregar recursos	15