



GUÍA PARA LA APERTURA Y PUBLICACIÓN DE DATOS ABIERTOS DE GOBIERNO

BUENAS PRÁCTICAS



GUÍA PARA LA APERTURA Y PUBLICACIÓN DATOS ABIERTOS DE GOBIERNO

BUENAS PRÁCTICAS

Versión 2.0 - 2018

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 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 que los mismos sean reutilizados en nuevos proyectos, que puedan combinarse con otras fuentes de datos y que puedan generar nuevas aplicaciones desarrolladas por el gobierno, por la sociedad civil, organizaciones, empresas o ciudadanos en general.

Esta guía brinda introducción al tema de datos abiertos, buenas prácticas para la publicación, normativa vigente y como implementarla. También se ofrecen referencias a contenidos más específicos para el lector que quiera profundizar en el tema.

Tabla de contenido

Introducción	7
¿Qué son los datos abiertos?	8
¿Qué datos se pueden abrir?	9
Capítulo I - Apertura de Datos	10
Clasificación de los Datos Abiertos	11
1 Estrella	11
2 Estrellas	12
3 Estrellas	12
4 Estrellas	12
5 Estrellas	13
Procedimiento para la apertura de Datos	13
Precondiciones necesarias	13
Reglas básicas para la apertura de datos	13
Primeros pasos para abrir datos	15
Publicación en el catálogo nacional de datos abiertos	22
Buenas prácticas para la publicación	23
Organización de la publicación	23
Descripción de la publicación	25
Visibilidad de los datos	28
Capítulo III - Normativa de Datos Abiertos	29
Artículo 82 de Ley N° 19.355	30

Decreto reglamentario	30
Condiciones establecidas para la publicación de datos abiertos	31
Formatos	31
Localización	31
Licencia de Datos Abiertos - Uruguay	32
Implementación de la normativa de datos abiertos	33
Capítulo IV - Glosario de términos	34
Anexo I - Normativa asociada	37
Anexo II - Modo de presentación de proyecto para cumplimiento de normativa	42

Capítulo I

Introducción

¿Qué son los Datos Abiertos?

Los **Datos Abiertos** son aquellos datos que pueden ser utilizados y distribuidos libremente por cualquier persona u organización, sin otra limitación que la atribución y la normativa nacional vigente. Esto significa que un dato abierto puede ser utilizado por todas las personas y con cualquier objetivo, ya sea crear un emprendimiento, una investigación científica o periodística, crear nuevos servicios o mejorar servicios existentes, etc; y con la sola limitante de respetar la normativa vigente y a lo sumo indicar la procedencia de los datos que están siendo utilizados.

Para que una publicación de datos abiertos sea apropiada deberá cumplir con los siguientes principios:

1. Los datos deben ser completos

Todos los datos estarán disponibles. Los únicos datos que pueden omitirse en una publicación son aquellos que incumplen alguna normativa vigente. Esto significa que si un dato no incumple ninguna norma debe ser publicado aunque no se le considere en principio de alguna utilidad.

2. Los datos deben ser primarios

Los datos se publican tal como se extraen de la fuente de origen, con el mayor nivel posible de granularidad y no se aceptarán 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. Un dato publicado fuera del período en el cual se lo considera de interés pierde relevancia; por lo tanto mantener las publicaciones actualizadas y en tiempo es fundamental.

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 previo para su descarga.

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 licencias restrictivas

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. No obstante es importante que las licencias que se asocian a las publicaciones de datos abiertos sean licencias que habilitan la utilización de los mismos libremente.

¿Qué datos se pueden abrir?

Se puede abrir toda la Información pública que se genera o está en posesión de cualquier organismo público, sea o no estatal, salvo las excepciones establecidas por la Ley N° 18.381 de Acceso a la Información Pública. A modo de ejemplo, algunas de las excepciones son los datos personales, datos que pueden afectar la seguridad nacional, datos que estén sujetos a otra normativa como la del Secreto Bancario, entre otros.

Capítulo II

Apertura de datos

Clasificación de los datos abiertos

Tim Berners-Lee¹ propuso una categorización del grado de apertura de los datos basado en estrellas, respecto a qué tan abiertos y usables son los datos que puede ofrecer una institución

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.

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

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

4 estrellas: Todo lo anterior más la utilización de estándares recomendados por W3C (RDF - SPARQL).

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

1 ESTRELLA

Según la clasificación antes descrita, cualquier dato que esté disponible en la Web en cualquier formato bajo una licencia abierta, es un dato abierto de categoría 1 estrella.

En el siguiente capítulo se explica que implica una licencia abierta, a modo introductorio es un documento que expresa formal y explícitamente las condiciones bajo las cuales podrán ser utilizados los datos abiertos sujetos a esa licencia.

Entonces, a modo de resumen:

Cualquier archivo disponible en sitios Web en cualquier formato .PDF, .DOC, .ODT, .TXT, entre otros; y que esté bajo una licencia abierta será clasificado como dato abierto de 1 estrella.

¹ <http://www.w3.org/People/Berners-Lee/>

2 ESTRELLAS

El siguiente nivel de clasificación, es 2 estrellas. Tiene que cumplir todo lo que implica estar en 1 estrella (publicado y bajo una licencia abierta) y además los datos deben estar estructurados y en un formato que pueda ser interpretado o procesado por máquinas. Sin embargo estos datos dependen de software propietario para poder utilizarlos.

Por ejemplo, datos que se encuentren en formato .XLS o .SHP se clasifican como 2 estrellas.

3 ESTRELLAS

El siguiente nivel de clasificación, 3 estrellas implica lo mismo que los niveles 1 y 2 pero el formato debe ser no propietario. Los siguientes formatos son algunos ejemplos de este nivel:

.CSV - archivo de texto separado por comas

.XML - archivo de texto estructurado mediante marcado

.KML - formato para representar datos geográficos en archivo de texto estructurado

Estos datos son estructurados y no se requiere de software privativo para poder interpretarlos o utilizarlos.

4 ESTRELLAS

Para que un dato abierto se considere de 4 estrellas debe cumplir con las condiciones anteriores (licencia abierta y estructurado en formato no propietario) y además utilizar estándares recomendados por W3C como pueden ser RDF y SPARQL; este nivel utiliza URI para poder referenciar recursos a los que la gente puede acceder.

5 ESTRELLAS

El nivel 5 estrellas cumple con el nivel 4 y además los datos se encuentran enlazados; lo que permite dar contexto a esos datos.

Procedimiento para la apertura de datos

PRECONDICIONES NECESARIAS

Es necesario que el proceso de apertura de datos se incorpore a los procesos naturales de la organización. Como toda temática relativamente nueva implica un cambio cultural que no todas las organizaciones tienen en cuenta al momento de iniciar un proceso de apertura de datos. Por lo tanto, es extremadamente necesario que esto se realice avalado mediante un mandato formal de las autoridades de la organización. También es necesario encontrar un referente que sea capaz de conformar un equipo multidisciplinario que lidere la iniciativa en la organización y tenga el reconocimiento de las autoridades. Por otra parte el proceso se debe apoyar con instancias de sensibilización y capacitación a los funcionarios y ciudadanos para que puedan publicar y hacer uso de estos datos.

REGLAS BÁSICAS PARA LA APERTURA DE DATOS

La publicación de datos requiere conocer las obligaciones que como organismo se tienen en cuanto a cumplir con la normativa vigente sobre publicación de datos abiertos y por otro lado conocer la demanda de información que existe desde la ciudadanía; sociedad civil, academia, medios de prensa, etc.

Teniendo claro la demanda existente de datos se debe comenzar por los datos que sean más sencillos de publicar. Para saber cuáles son los datos más sencillos se pueden seguir las siguientes recomendaciones:

1. Cuando hablamos de datos sencillos son aquellos que se pueden extraer fácilmente de un sistema informático, a un costo relativamente bajo de trabajo y que por su naturaleza no generen disputas o controversias con la opinión pública.
2. Que los datos sean de fácil mantenimiento y que la publicación pueda permanecer actualizada en el tiempo. Algunas publicaciones y estructuras de datos modifican su información frecuentemente y mantenerlas actualizadas

no es un tema menor. Por lo que comenzar con conjuntos de datos que no tengan una tasa alta de cambios facilita el inicio del proceso de publicación.

3. Que las publicaciones sean oportunas también es importante pensando que pueden existir emprendimientos basados en estos datos, por lo que el compromiso del organismo con los usuarios es fundamental para la sostenibilidad tanto de la iniciativa de datos como de los emprendimientos. A modo de ejemplo, de poco sirve publicar los horarios de ómnibus de invierno, si la publicación se hace en los meses de verano con seis meses de atraso. Puede que los datos estén perfectos pero de nada sirven para un emprendedor que debe dar los horarios de ómnibus en tiempo real y menos aún para el usuario de una aplicación.
4. Es recomendable que las publicaciones de datos abiertos se encuentren centralizadas en catálogos de datos abiertos que facilitan su uso y disponibilidad. Además de dar la infraestructura necesaria, los catálogos exigen que cada publicación cuente además con información sobre lo que allí se publica. Esta información se la conoce como metadatos y es de utilidad para que los usuarios puedan entender correctamente los datos y conocer el alcance de la publicación.
5. Otra cuestión importante a la hora de publicar datos abiertos es que el principal consumidor de estos datos no es el ciudadano común, sino que son intermediarios que tomarán esos datos para generar “algo”. Por lo que los datos deben estar en las mejores condiciones de calidad posible, esto no significa que deben ser perfectos pero si en las mejores condiciones que el sistema actual permite. Los errores que puedan surgir se pueden corregir en publicaciones futuras y seguramente sean los usuarios quienes comiencen a reportar algunos de estos errores de calidad de datos.
6. Se puede comenzar con los datos que ya están publicados en todo o en parte en el sitio web del organismo. Con esto nos aseguramos al menos que lo que se va a publicar no tiene restricciones normativas. Además de que sabemos que existe una fuente de datos desde donde se extrajeron para su publicación.
7. En cuanto a mejorar la rendición de cuentas; se puede comenzar por la información que difunde el organismo. Apoyar la publicación de noticias con datos abiertos es una forma de transparentar el proceso de ejecución de las políticas públicas; y dar a conocer los datos es importante para la evaluación de la gestión del organismo.

PRIMEROS PASOS PARA ABRIR DATOS

A continuación se describe el paso a paso para quien quiere abrir datos y no sabe por dónde empezar:

1. Seleccionar el conjunto de datos a abrir
2. Convertir el formato del archivo a un formato abierto
3. Aplicar una licencia abierta
4. Incluir los metadatos
5. Publicar

1. Seleccionar los datos a abrir

Sin importar la clasificación, en todos los casos, el primer paso es seleccionar los datos que se quieren publicar como datos abiertos.

Desde el punto de vista interno a la organización la opción más sencilla es comenzar por lo que ya se tiene publicado o que es de fácil transformación. Para comenzar este proceso puede realizar las siguientes preguntas:

- ¿Qué información solicitan habitualmente y que exige un procesamiento de datos?
- ¿Qué información existe en la organización 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 intercambia con 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. Tenga en cuenta aspectos legales, de completitud, capacidad para mantenerla actualizada, formato, etc.



Ayuda:

Se recomienda conocer la demanda existencia.

Cuando el proceso de apertura esté en niveles más avanzados de maduración se pueden realizar instancias de diálogo con la ciudadanía que permitan conocer cuál es el interés existente en la población por los datos del organismo. Es importante tener en cuenta a diferentes actores como son la academia, los periodistas y la sociedad civil organizada. En estas instancias de diálogo se plantean

requerimientos por parte de los participantes y desde el organismo se trata de buscar soluciones a estos problemas, muchas de las cuales pueden estar acompañadas por la publicación de datos abiertos. También es posible abrir canales de comunicación utilizando otros medios diferentes al presencial como pueden ser consultas en línea, redes sociales, etc. Este relevamiento de solicitudes de datos se debe documentar para luego evaluar la viabilidad de su publicación.

2. Convertir los datos a un formato abierto

Los formatos en Uruguay para la publicación de datos abiertos son los que cumplen con el nivel de tres estrellas y superiores según la escala de Tim Berners-Lee vista en el capítulo II.

Existen dos tipos de formatos adecuados para la publicación de datos abiertos:

- Formatos Generales. Son formatos que pueden especializarse para usos más específicos. Los dos más habituales son CSV, XML y RDF.
- Formatos Específicos. Son formatos basados en los formatos generales, pero que ya están especializados para un uso determinado.

Algunos ejemplos son:

- ✓ Open Fiscal Data Package → Publicación de presupuesto
- ✓ Open Contracting → Publicación de compras y contrataciones

Unificación de formatos

La conversión de datos no solo implica convertir el archivo a otro formato, también hay que tener en cuenta que los datos contenidos cumplan estándares de codificación de manera de facilitar su uso.

A modo de ejemplo, un campo que debe contener fechas no podría contener distintos formatos.

Incorrecto	Correcto
31/12/01	01/08/2002
31-12-01	20/02/2005
31 dic. 2001	01/01/1000

Para los campos que no tienen datos, es conveniente indicar que el campo es vacío agregando un indicador de que no existe el dato, evitando dejar la celda vacía (vacío, nulo, no aplica) en caso de que sea un campo numérico debería indicarse con cero y en caso de fechas se podría tomar 01/01/1000 u otra fecha conocida.

Para las direcciones existen muchos formatos posibles por lo que si logramos separar la dirección en la mayor cantidad de campos posibles facilitaremos su correcta interpretación.

Dirección
Andes 1365 esq. 18 de julio, apto 7
Andes 1365, apto 7
Andes 1365 entre 18 y Colonia, apto 7

Esta dirección se puede separar en varios campos

Calle	Numero	Esquina1	Esquina2	Apto
Andes	1365	18 de julio		7
Andes	1365			7
Andes	1365	18	Colonia	7

El proceso de depuración de datos, puede requerir de varias interacciones para estandarizar los datos. Para el caso de las direcciones en la segunda interacción deberíamos estandarizar la columna esquina1.

Calle	Numero	Esquina1	Esquina2	Apto
Andes	1365	18 de julio		7
Andes	1365			7
Andes	1365	18 de julio	Colonia	7

En la tercera interacción deberíamos hacer que todas las direcciones para Andes 1365 queden iguales, completando el campo Esquina2.

Agesic publica en formato abierto datos referenciales de direcciones de manera que todos los conjuntos de datos abiertos utilicen la misma codificación

[Acceder al Catálogo de Datos Abiertos](#)



Ayuda:

Consultar el modelo de direcciones propuesto por Agesic.

3.- Licenciamiento de datos abiertos

Aplicar una licencia de datos abiertos es esencial para que el dato publicado sea clasificado como abierto. La licencia a aplicar a un conjunto de datos es el documento que expresa las condiciones bajo las cuales podrán ser utilizados los datos disponibles en un sitio Web.

Todo conjunto de datos abiertos que sea publicado por un organismo del Estado deberá hacer referencia en la declaración de sus metadatos a la licencia de datos abiertos de Uruguay, mientras que el resto de las organizaciones podrán optar por la licencia de Uruguay u otras licencias abiertas.

Una licencia de datos abiertos debe acompañar siempre a cualquier publicación de datos abiertos.

El texto de la licencia de datos abiertos de Uruguay se encuentra publicado en el sitio datos.gub.uy.

También se transcribe en el Capítulo III de este documento.

Para las organizaciones que publican datos en el Catálogo Nacional de Datos Abiertos y no son del Estado (por ejemplo organizaciones de la sociedad civil, de la academia o empresas privadas) existe la posibilidad de utilizar otra licencia. En este caso, el usuario Catalogador podrá seleccionar otra de las licencias abiertas que dispone el catálogo o publicar junto al archivo de datos abiertos el texto de la licencia correspondiente.

4.- Metadatos

Para facilitar la utilización de los datos por parte de terceros se les debe de proveer de información adicional (metadatos) que describan la información que se publica.

Según el formato en el que se está publicando existen diferentes formas de publicar los metadatos, la siguiente tabla muestra ejemplos para los formatos más comunes:

Formato de los recursos de datos	Metadatos
CSV	Descripción de las columnas en archivo CSV u otro formato. Ver siguiente ejemplo.
XML	XML Schema

JSON

JSON Schema

A la hora de publicar los datos abiertos en un formato abierto, se debe hacer referencia al archivo con los metadatos de los recursos publicados.

Este archivo debe explicar el significado de cada columna, como se muestra en la siguiente imagen:

Nombre de atributo	Descripción	Tipo de dato	Recurso relacionado	Información Adicional
<i><Descripción exacta de la columna tal cual aparece en los archivos de los conjuntos de datos>.</i>	<i><Descripción detallada del significado del atributo orientada a las personas que pretenden utilizar el atributo>.</i>	<i><Numérico, Texto, Fecha/Hora, Moneda></i>	<i><Recurso que da significado o al atributo. Ejemplo: si el atributo es un código tiene como recurso relacionado la tabla o recurso donde se describe el significado o de dicho código.></i>	<i><Cualquier información adicional sobre el atributo que complementa e la información básica.></i>

<i>Ejemplo: CodigoSexo</i>	<i>Código asociado al sexo de la persona</i>	<i>Numérico</i>	https://catalogodatos.gub.uy/datos-referenciales-codigueras/225c6f8-ce7c-468c-8913-aebad840-7dcc/reso-urce/0ea2-3b5a-dd59-4928-a9f1-72f15d1af-786/download/sexo_v1.xml	<i>Agesic proporciona el conjunto de datos referenciales para estandarizar los datos. https://catalogodatos.gub.uy/datos-referenciales-codigueras</i>
--------------------------------	--	-----------------	---	--

5.- Publicación

La publicación del conjunto de datos se debe realizar de alguna de las maneras siguientes.

- En el catálogo nacional de datos abiertos
- En los sitios web de los organismos
- o en ambos lugares

Agesic recomienda que los conjuntos de datos se publiquen en el catálogo nacional de datos abiertos; en el caso de que los conjuntos de datos se publiquen en los sitios web de los organismos, se deberían catalogar en el catálogo nacional (catalogodatos.gub.uy) enlazando los datos desde el sitio web institucional.

En el sitio web del organismo

Si se hace en los sitios web institucionales de los organismos, estos deberán contar con una sección propia para catalogar los datos abiertos siguiendo las recomendaciones de Agesic establecidas más adelante en esta guía.

En catalogodatos.gub.uy

Si se hace en el Catálogo Nacional de Datos Abiertos, se debe realizar siguiendo las recomendaciones de Agesic establecidas más adelante en esta guía.

Acceso a los conjuntos de datos publicados

Cada conjunto de datos permitirá acceder mediante una o más URLs a sus diferentes recursos:

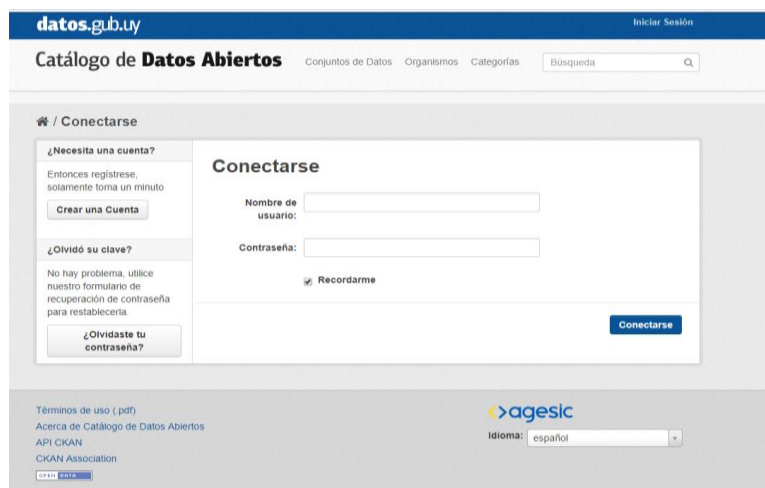
- Los datos que conforman el contenido del conjunto de datos, bajo la forma de uno o más archivos.
- Los metadatos que describen ese contenido.
- La licencia de acceso contenida en la Sección 4 del presente documento.

Para que un conjunto de datos publicado sea considerado como dato abiertos según la normativa vigente, debe usar tanto para el contenido como para los metadatos, formatos adecuados para publicaciones que tengan como mínimo el nivel de apertura 3 estrellas de la escala de Tim Berners-Lee.

Publicación en el catálogo nacional de datos abiertos

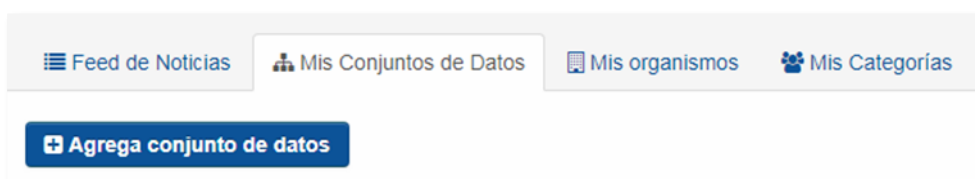
El catálogo de datos abiertos está implementado utilizando CKAN como software de base. CKAN es una plataforma de código abierto y gratuito.

Para comenzar la publicación de datos abiertos es necesario solicitar el alta de la organización y de los usuarios que serán responsables de las publicaciones. La solicitud se debe hacer por correo a la dirección datosabiertos@agesic.gub.uy. Otro procedimiento es que el propio usuario se registre y solicite por correo a la dirección datosabiertos@agesic.gub.uy que se lo asocie a una organización.



The screenshot shows the login page of the 'datos.gub.uy' portal. The header includes the site name 'datos.gub.uy' and a 'Iniciar Sesión' link. Below the header, there are navigation links for 'Conjuntos de Datos', 'Organismos', and 'Categorías', along with a search bar. The main content area is titled 'Conectarse' and contains a registration sidebar on the left with links for '¿Necesita una cuenta?', '¿Olvidó su clave?', and '¿Olvidaste tu contraseña?'. The main form has fields for 'Nombre de usuario' and 'Contraseña', a 'Recordarme' checkbox, and a 'Conectarse' button. The footer contains links for 'Términos de uso (.pdf)', 'Acerca de Catálogo de Datos Abiertos', 'API CKAN', and 'CKAN Association', along with the 'agesic' logo and a language dropdown set to 'español'.

Una vez que el usuario inicia sesión se le muestra la pizarra de actividades donde podrá ver alguna de estas opciones de acuerdo a los permisos que tenga asignados.



En la opción de Mis conjuntos de datos, tendrá acceso a todos los conjuntos de datos que haya publicado y podrá generar una nueva publicación.

La publicación del conjunto de datos consta de dos pasos, la creación del conjunto de datos y la publicación de los recursos.

Para la publicación de datos es recomendable se tengan en cuenta las buenas prácticas que se detallan en esta guía y el manual de usuario para la publicación de datos que se publica en el sitio de datos abiertos; datos.gub.uy.

Buenas prácticas para la publicación

Organización de la publicación

Es necesario para la correcta publicación de datos tener en cuenta algunos aspectos de organización de la publicación: la periodicidad de los datos, el tamaño, formato y ubicación de los recursos.

Teniendo en cuenta la periodicidad de los datos.

1- Si se publica un conjunto de datos cuyos recursos se actualizan una vez al año o tienen una frecuencia de actualización mayor a un año, se recomienda generar un conjunto de datos genérico y agregar recursos especificando en el nombre el año correspondiente.

Ejemplo:

Conjunto de datos: Rendiciones de cuentas 2010 - 2020

Recursos: Rendición de cuentas 2010

Rendición de cuentas 2011

Rendición de cuentas 2012

2- Si se publica un conjunto de datos cuyos recursos se actualizan mensualmente o con una periodicidad de entre un mes y un año se recomienda generar un conjunto de datos por año y recursos de acuerdo a la periodicidad.

Ejemplo:

Conjunto de datos: Remuneraciones 2017

Recursos: Escala de remuneraciones primer semestre

Escala de remuneraciones segundo semestre

3- Si se publican conjuntos de datos cuyos recursos tienen frecuencia de actualización menor a un mes es recomendable la utilización de API para la obtención de los datos.

Ejemplo:

Conjunto de datos: Campings habilitados

Recursos: Acceso a campings habilitados en línea

Este recurso tendría los campings habilitados en tiempo real y debería existir una API o URL que permita consultar los campings habilitados en cierto período de tiempo o para una región, etc.

Teniendo en cuenta el tamaño de los recursos

Actualmente el catálogo permite subir recursos de un tamaño de hasta 450Mb. Sin embargo es bueno tener algunas consideraciones al momento de realizar una publicación.

Si el recurso es superior a 100Mb se recomienda dividir el mismo en varios recursos de acuerdo a algún criterio establecido.

Ejemplo:

- Por rango de fechas.
- Por categorías

Si el recurso es superior a 450Mb se recomienda comprimir y dividir los recursos, la opción de compresión no es recomendable por lo que trate de evitar esta opción en los casos que considere se pueda hacer; teniendo en cuenta los tiempos de descarga y dificultades para su uso o publicación.

Sobre los formatos de los recursos

Tenga en cuenta que cada conjunto de datos puede contener los mismos datos en diferentes formatos. Si bien el formato obligatorio es el correspondiente al nivel 3 estrellas (CSV, XML, JSON, KML, etc.) existe un público que no está habituado o capacitado para trabajar con estos formatos, por lo que es una buena práctica incluir también los recursos en formatos que faciliten su uso a personas no técnicas, como por ejemplo la clásica planilla u hoja de cálculo.

Sobre la ubicación de los recursos

Los recursos de datos pueden ser subidos al catálogo o enlazados a recursos publicados en otros sitios. En el caso de que los recursos sean enlazados a sitios externos hay que tener la precaución de que estos recursos estén publicados en una URL fija y que no sufra modificaciones en el tiempo; ya que de ocurrir cambios en el sitio fuente se perdería el enlace a todos los recursos que allí estuvieran alojados. Una solución a esto es publicar los recursos en un sitio FTP público no dependiente de la herramienta de gestión de contenidos del organismo.

Descripción de la publicación

Metadatos del conjunto de datos

Al catalogar un conjunto de datos es necesario describirlo mediante una serie de metadatos que provee el catálogo de datos abiertos. Sea claro en los conceptos y términos utilizados; el correcto uso de los datos depende en gran parte de esto.

Título del conjunto de datos y descripción. Sea breve y explícito al momento de elegir un nombre para el conjunto de datos. Podrá detallar sobre el contenido en la descripción del conjunto de datos.

No utilice textos solo en mayúsculas o solo minúsculas, respete las formas normales de redacción.

Etiquetas. Las etiquetas son útiles para que los usuarios puedan realizar búsquedas avanzadas mediante filtros previamente definidos. Para agregar etiquetas escriba una o varias palabras para la etiqueta y finalice la etiqueta con punto y coma. Agregue etiquetas siempre que sea posible y tenga en cuenta los términos que contribuyan a filtrar sus datos; esto facilitará la búsqueda para quien utilice los datos.

Licencia. Todo conjunto de datos abiertos debe especificar bajo que licenciamiento se liberan los datos. Seleccione una de las licencias válidas establecidas para el catálogo de datos abiertos de Uruguay. En caso de ser un organismo público la licencia a utilizar es la Licencia de Datos Abiertos de Uruguay.

Visibilidad. Este campo permite publicar los datos de forma privada o pública. La primera solo permite el acceso a los datos a los usuarios del organismo registrados en el catálogo, la segunda deja al conjunto de datos accesible para todo público.

La **fuentes**, se refiere al origen de los datos.

En caso de que los datos hayan salido de un sistema indicar desde donde se obtienen originalmente los datos.

Ejemplos:

- El organismo X publica sus compras, pero los datos se extraen desde el sitio de la Agencia de Compras y Contrataciones Estatales (ACCE), por lo tanto la fuente es la publicación que hace ACCE.
- Lo mismo aplica en caso de que los datos sean los utilizados para la confección de un anuario o informe. En este caso la fuente podría ser la URL de publicación del informe.

El conjunto de datos permite asignar una **versión** a la publicación de manera de facilitar el acceso a datos históricos. El control de versiones depende del publicador y cada versión se debe generar ante cambios en los recursos publicados o en los datos de los recursos.

Algunos ejemplos para generar una nueva versión son:

- Se debe tener una nueva versión si un conjunto de datos pasa de tener N recursos a tener M recursos
- Se debe tener una nueva versión cuando un recurso de un conjunto de datos varía en la cantidad de columnas del archivo.
- Se debe tener una nueva versión cuando cambia el sistema de referencia.
- Se podría tener una nueva versión si cambia la frecuencia de actualización.

En este caso se debe hacer un análisis del impacto, por ejemplo si cambia de mensual a anual no es lo mismo que si cambia de anual a una API de consulta en tiempo real.

En cuanto al **autor**, se debe considerar como autor al área del organismo que se encarga de la publicación. Mientras que el **mantenedor** hace referencia al equipo o persona que se encarga de mantener la publicación de los datos.

Para ambos campos es necesario incluir las **direcciones de correo** de manera de facilitar la interacción con los usuarios.

En cuanto a cada recurso de datos se deben considerar los siguientes metadatos.

La **frecuencia de actualización** indica cada cuanto tiempo se actualiza un recurso.

El **sistema de referencia** permite indicar el sistema de referencia geodésico utilizado en el caso de que los datos contengan coordenadas geográficas.

Algunos de estos sistemas son los siguientes:

- WGS84
- PSAD56
- SIRGAS

La **cobertura temporal y espacial** indica en que período y territorio son válidos los datos.

Nombre y descripción del recurso, aplica el mismo criterio que para el conjunto de datos. No utilice textos solo en mayúsculas o solo minúsculas, respete las formas normales de redacción.

En el caso del nombre, no es necesario que el mismo coincida con el nombre del archivo que se va a publicar. No utilice “_” para unir palabras en el nombre del archivo ni nomenclatura técnica.

Forma incorrecta	Forma correcta
datos_presupuesto_2019	
datosPres2019	Presupuesto para el año 2019
Presupuesto 2019.xls	

Formato indica en que formato se publica el recurso, este texto es libre pero debe indicar la extensión del archivo de datos; XML, JSON, KML, XLS, etc.

En caso de que el recurso de datos sea un archivo comprimido y su contenido se corresponda a un solo tipo de archivos de datos, el formato a especificar será

Ejemplo:

- ZIP
- RAR

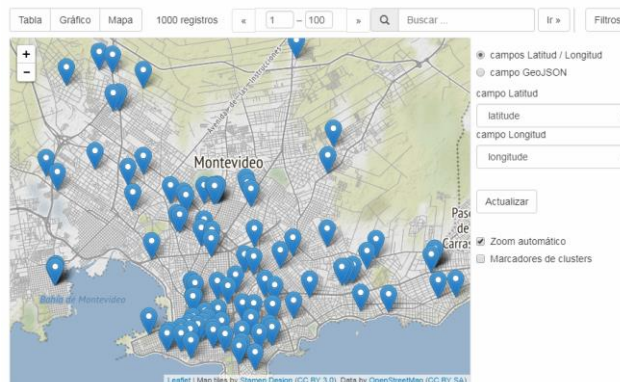
En el caso de que el recurso de datos sea un archivo comprimido conteniendo varios archivos de distintos formatos, utilice la extensión ZIP o RAR u otra de forma genérica sin especificar el formato de los archivos contenidos.

Si este campo se deja sin completar, el mismo se completa automáticamente.

Visibilidad de los datos

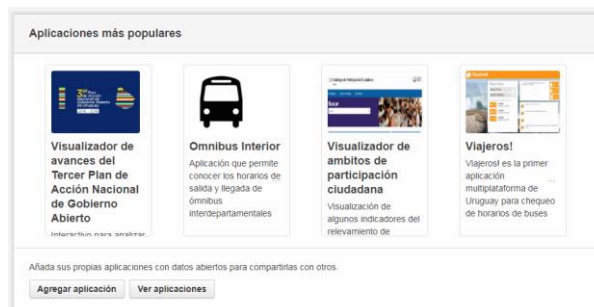
Dentro de las funcionalidades mejoradas del catálogo está la posibilidad de crear vistas para cada recurso de datos. Esto posibilita que los usuarios puedan hacer una visualización previa de los datos, verlos en gráficos o en mapas. Además da la posibilidad de enlazar a visualizaciones realizadas con herramientas de terceros en caso de querer presentar visualizaciones más complejas.

Se considera una buena práctica acompañar cada publicación con alguna visualización de los datos ya que de esta forma se facilita el entendimiento de los mismos para un público no experto, lo que permite una mayor difusión de los datos.



Ejemplo de visualización de datos en un mapa

El catálogo permite asociar visualizaciones, aplicaciones y artículos de investigación acompañando cada publicación de datos. En portada se pueden apreciar cuales son las aplicaciones más populares.



Capítulo III

Normativa de Datos Abiertos

Artículo 82 de Ley N° 19.355

La Ley N° 19.355, del 19 de diciembre de 2015, estableció en su artículo 82 la obligación para las entidades públicas de publicar en formato abierto la información preceptuada por los artículos 5° y 8° de la Ley N° 18.381, del 17 de octubre de 2008, y los artículos 38 y 40 de su Decreto Reglamentario N° 232/010, del 2 de agosto de 2010.

Este artículo establece lo siguiente:

Las entidades públicas deberán como mínimo publicar en formato abierto, la información preceptuada por el artículo 5 de la Ley N° 18.381, de 17 de octubre de 2008, y por los artículos 38 y 40 del Decreto N° 232/010, de 2 de agosto de 2010, según corresponda en el ámbito de su competencia.

Los datos y sus metadatos asociados deberán cumplir con las normas técnicas que determine la Agencia para el Desarrollo del Gobierno de Gestión Electrónica y la Sociedad de la Información y del Conocimiento. La publicación de estos datos deberá realizarse en el marco de lo dispuesto por la Ley N° 19.179, de 27 de diciembre de 2013 ("software" libre y formatos abiertos en el Estado).

En el anexo I encontrará lo establecido en el artículo 5° de la Ley N° 18.381 y los artículos 38 y 40 del Decreto N° 232/010 a los que se hace referencia.

Decreto Reglamentario

En el marco del proceso de fortalecimiento de la transparencia de la información pública, el Grupo de Trabajo de Datos Abiertos elaboró la redacción del Decreto Reglamentario, aprobado por el Poder Ejecutivo el 22 de febrero.

El decreto establece la obligatoriedad para los organismos del Estado de publicar sus datos abiertos bajo las pautas establecidas en la licencia de datos abiertos de Uruguay. Los datos deberán publicarse de acuerdo a lo que establece la norma técnica sobre los formatos admitidos y formas de publicación incluidos como documentación anexa al decreto.

Tanto el decreto reglamentario como sus anexos: la licencia de datos abiertos de gobierno de Uruguay y las directrices técnicas se pueden descargar desde el sitio de Agesic o desde el portal de datos datos.gub.uy

Condiciones establecidas para la publicación de datos abiertos

FORMATOS

En la siguiente tabla se puede ver una lista de formatos generales para la publicación de datos abiertos. Estos formatos son considerados adecuados de acuerdo a lo establecido en la **primera versión** de las Directrices técnicas para la publicación de Datos Abiertos.

Nombre	Uso esperado	Referencias
CSV	Datos con estructura relacional (Tablas)	RFC4180
XML	Datos con estructura jerárquica (Arboles)	XML 1.0 XML 1.1
JSON	Datos con estructura jerárquica (Arboles)	
RDF	Datos con cualquier estructura. (Grafos)	RDF 1.1

LOCALIZACIÓN

Los datos abiertos podrán ser publicados en el sitio institucional del organismo, en el catálogo nacional de datos abiertos o en ambos.

LICENCIA DE DATOS ABIERTOS - URUGUAY

Los organismos del Estado deberán publicar estos datos bajo la licencia de datos abiertos de Uruguay incluida en el decreto reglamentario; para ello deberán hacer referencia al texto de la licencia publicado en el portal datos.gub.uy

A modo de información a continuación se transcribe la Licencia de Datos Abiertos de Uruguay incluida en el punto 4 del decreto reglamentario al que se hace mención en este capítulo.

Licencia de Datos Abiertos - Uruguay

Versión 0.1

Autorización. Los datos abiertos y los metadatos abiertos proporcionados, para uso comercial y no comercial podrán ser:

- enajenados, reproducidos, distribuidos, publicados, traducidos, adaptados, transformados, comunicados o puestos a disposición del público mediante cualquier forma o procedimiento;
- fusionados con datos propios y de terceros, pudiendo combinarse para formar conjuntos de datos nuevos e independientes;
- integrados en los procesos de negocio, productos y aplicaciones internas y externas en las redes electrónicas públicas y privadas.

Nota de Origen. A los efectos de utilizar los datos abiertos y metadatos abiertos en los términos referidos en el artículo anterior se deberá en todos los casos citar la siguiente información:

- el nombre del proveedor,
- la referencia a la "Licencia de Datos Abiertos - Uruguay".
- la referencia al conjunto de datos que está siendo utilizado.

Los cambios, las ediciones, los nuevos diseños o modificaciones deben ser mencionados en la nota de origen.

Implementación de la normativa de datos abiertos

La implementación de la normativa de acuerdo al decreto reglamentario exige a los organismos contar con un plan de trabajo para la publicación de los datos de transparencia activa.

Este plan debe establecer claramente los objetivos, alcance, restricciones y un cronograma de las actividades que se van a desarrollar. También debe incluir la conformación del equipo de trabajo y los roles de cada integrante.

Para facilitar la creación de este documento Agesic ha generado un modelo de plan de trabajo que se encuentra disponible en el anexo II y publicado en datos.gub.uy.

En la primera etapa de implementación se recomienda comenzar por los ítems relativos al Artículo 5 de la Ley de Acceso a la Información Pública y continuar por el resto en etapas más avanzadas del proyecto. Otra forma es comenzar por los datos que son más fáciles de publicar tal como se indicó previamente en el capítulo dos de esta guía.

En el documento de norma técnica encontrará información sobre aspectos técnicos de la publicación de datos que se aplican a esta normativa. Algunos de estos aspectos se trataron en el capítulo dos de esta guía.

Agesic definió para la publicación de datos de transparencia activa los campos mínimos obligatorios que deben contener cada recurso de datos abiertos. Esto permitirá que todos los conjuntos de datos de transparencia activa sean interoperables y fáciles de utilizar. Para estos datos se recomienda que al menos una versión de los mismos esté publicada en formato CSV.

Los formatos a utilizar para estas publicaciones se encuentran publicados en datos.gub.uy.

Agesic provee el apoyo técnico desde el proyecto de datos abiertos y desde la Unidad de acceso a la información pública para los organismos que necesiten ayuda en la implementación de este proceso.

Capítulo IV

Glosario de términos

Conjunto de datos - Agrupamiento de recursos de datos sobre la misma temática.

CSV - Formato de archivos separados por coma, donde cada valor es delimitado por una coma o punto y coma; es una forma sencilla de publicar información pero necesariamente se debe publicar también información sobre los datos que contiene (meta data). De otra manera esta información será difícilmente interpretada.

DBF - Formato de archivos utilizado por dBase, generalmente es el formato del archivo de base de datos para computadores personales.

DBMS - Database Management System; sistema de gestión de bases de datos.

DOC - Formato de archivo creado por la herramienta de procesamiento de textos de Microsoft.

INE - Instituto Nacional de Estadísticas.

JSON - JavaScript Object Notation es un lenguaje muy sencillo de automatizar su procesamiento. Si bien no está tan expandido como XML su filosofía similar y su fácil interpretación lo hacen un lenguaje muy apropiado para el intercambio de información.

KML - Keyhole Markup Language, formato de archivo utilizado para la representación de datos geográficos en tres dimensiones, basado en XML.

LINKED DATA (Datos Relacionados) - Forma parte de la web semántica, a diferencia de la web actual donde los documentos de hipertexto se relacionan a diferentes puntos en el HTML, esta nueva forma enlaza los datos con cosas arbitrarias que se describen en formato RDF.

METADATOS - Datos que describen otros datos.

ODS - Formato de archivo de Open Document, es un estándar abierto para documentos de hojas de cálculo.

PDF - Formato portable de documentos, estándar para la publicación de documentos en la web.

Recursos de datos - o recursos, son archivos de datos que pertenecen a conjuntos de datos de un catálogo de datos abiertos basado en CKAN.

RDF - Es un marco para la descripción de recursos en la web. El formato RDF permite mediante un documento XML describir recursos publicados en Internet.

SHP - Formato de archivo informático de datos espaciales desarrollado por la empresa ESRI comercializadora de productos ArcGIS y ArcInfo.

SPARQL - Motor de consultas para la web semántica.

TXT - Formato de archivo de texto plano.

URI - Identificador único de recurso, permite identificar únicamente un recurso, pagina, documento, dirección de correo electrónico, etc.

XLS - Formato de archivo creado por Microsoft para su herramienta de hoja de cálculo.

XML - Formato extensible de marcas, lenguaje que permite la definición de etiquetas o marcas derivado del SGML.

W3C - Consorcio internacional que produce recomendaciones para la WWW.

WEB SEMÁNTICA - La web semántica es una visión futura de la web donde la información publicada debe tener un significado exacto, esta información debe ser entendida y procesada por computadoras. “Es una web de datos que pueden ser procesados directa o indirectamente por maquinas” según la definición de Tim Berners-Lee.

WEBSERVICES - En el caso que la información sea extraída de una fuente de información viva, en constante modificación; se recomienda la utilización de servicios web que expongan dicha información en tiempo real.

Anexo I

Normativa asociada

Ley N° 18.381 - artículo 5

Difusión de la información pública.

Los sujetos obligados deberán prever la adecuada organización, sistematización y disponibilidad de la información en su poder, asegurando un amplio y fácil acceso a los interesados.

Los organismos públicos, sean o no estatales, deberán difundir en forma permanente, a través de sus sitios web u otros medios que el órgano de control determine, la siguiente información mínima:

- A) Su estructura orgánica.
- B) Las facultades de cada unidad administrativa.
- C) La estructura de remuneraciones por categoría escalafonaria, funciones de los cargos y sistema de compensación.
- D) Información sobre presupuesto asignado, su ejecución, con los resultados de las auditorías que en cada caso corresponda.
- E) Concesiones, licitaciones, permisos o autorizaciones otorgadas, especificando los titulares o beneficiarios de éstos.
- F) Toda información estadística de interés general, de acuerdo a los fines de cada organismo.
- G) Mecanismos de participación ciudadana, en especial domicilio y unidad a la que deben dirigirse las solicitudes para obtener información.

Decreto 232/010

Art. 38.- Información que debe ser difundida por todos los sujetos obligados.- Los sujetos obligados deberán difundir en sus sitios web la siguiente información, que deberá ser actualizada mensualmente:

1. Creación y evolución histórica del sujeto obligado conjuntamente con sus cometidos.
2. Estructura orgánica en un formato que habilite la vinculación de cada oficina con ésta, las facultades y responsabilidades que le corresponden de conformidad con la normativa vigente.
3. El marco jurídico aplicable a cada organismo.
4. Programas operativos de largo y corto plazo y mecanismos que permitan visualizar metas y cumplimiento de éstas.

5. El Diario de Sesiones.

6. Los anteproyectos y proyectos de ley que se presenten, así como cualquier otro tipo de comunicación legislativa, la indicación del trámite de que ha sido objeto y las resoluciones que sobre éstos recaigan.

7. Listado con los servicios que ofrece y los programas que administra incluyendo los trámites para acceder a ellos y la población objetivo a que están dirigidos.

8. El listado de los funcionarios, a partir del jefe de departamento, gerente, director o equivalente hasta el jerarca máximo, incluyendo: nombre, domicilio postal, electrónico y números telefónicos oficiales. En caso de tratarse de órganos legislativos, las comisiones que se integren, estableciendo el número de convocatorias, presencias y ausencias en los diferentes órganos, inclusive de los suplentes proclamados por la Corte Electoral.

10. Nómina de los funcionarios que no perteneciendo al organismo cumplen funciones en el mismo, sea por ser contratados, sea por estar en comisión provenientes de otro organismo, con indicación de compensaciones recibidas con cargo a las partidas asignadas al funcionamiento de los organismos correspondientes.

11. Perfil de los diferentes puestos de trabajo y currículum actualizado de quienes ocupan aquéllos a partir del jefe de departamento, gerente, director o equivalente hasta el jerarca máximo.

12. Convocatorias a concursos de ingreso para ocupar cargos y resultados de los mismos.

13. Remuneración mensual nominal de todos los funcionarios incluyendo todas las prestaciones en razón de los diferentes sistemas de compensación que se aplicaren.

14. Indicación de los viáticos recibidos y la determinación de su utilización.

15. Listado de comisiones de servicio en el exterior de los funcionarios, viáticos percibidos, razón del viaje y resultados del mismo, incluyendo a todas las personas que integren la delegación sin excepción alguna.

16. Presupuesto y ejecución del mismo:

a. Ingresos recibidos por cualquier concepto, con indicación del responsable en la recepción, administración y ejecución.

b. Ingresos asignados por el presupuesto nacional.

c. Estados financieros y balances generales de gastos.

d. Auditorías.

i. Número y tipo de auditorías.

ii. Número de observaciones realizadas por rubro de auditoría sujeto a revisión.

iii. Total de aclaraciones efectuadas por el sujeto obligado.

17. En caso de los pliegos de bases y condiciones particulares de los procedimientos licitatorios que representen gastos de funcionamiento o de inversión y las resoluciones que

dispongan la adjudicación en dichos procedimientos, las que declaren desiertas o dispongan el rechazo de todas las ofertas presentadas, deberá establecerse un vínculo electrónico con el sitio www.comprasestatales.gub.uy, en cumplimiento de los requerimientos establecidos por las Leyes Nos. 16.736, de 5 de enero de 1996, art. 694, 17.060, de 23 de diciembre de 1998, art. 5 y 17.556, de 18 de setiembre de 2002, art. 163 y los Decretos Nos. 66/002, de 26 de febrero de 2002, 232/003, de 9 de junio de 2003, 393/004 de 3 de noviembre de 2004 y 191/007, de 28 de mayo de 2007.

18. Las partidas presupuestales provenientes de convenios con organismos internacionales o que se gestionen a través de éstos, deberán incluirse en la web del sujeto obligado.

19. Montos otorgados en carácter de financiamiento a los diferentes Partidos Políticos por parte del Estado.

20. Calendario de reuniones, citaciones de comisión, de directorios, de plenarios, de asambleas, que sean convocadas, así como presencias IM.P.O. IM.P.O. Nº 28.035 - Agosto 10 de 2010 301-A CARILLA Nº 9 y ausencias de los convocados, minuta de comunicación indicativa del orden del día de la convocatoria y resoluciones y resultados de las mismas.

21. Concesiones, licencias, permisos y autorizaciones debiendo publicarse el objeto, nombre o razón social del titular, así como si el contrato involucra el aprovechamiento de bienes, servicios o recursos públicos.

22. Programas educativos que se encuentran vigentes y correlaciones de adecuación en relación con los últimos dos planes de estudios inmediatos anteriores.

23. Listado de programas de capacitación, número de funcionarios capacitados así como sus evaluaciones.

24. Indicadores de gestión de evaluaciones al desempeño académico y/o administrativo.

25. Listados de partidos y agrupaciones políticas que se encuentran registrados.

26. Resultados totales de las elecciones y las discriminaciones que se consideren útiles para la ciudadanía.

27. La fecha de la última actualización.

28. Mapa del sitio.

29. Domicilio postal y electrónico y números telefónicos oficiales del sujeto obligado.

30. Información sobre la política de seguridad y protección de datos.

31. Cualquier otra información que pudiere ser de utilidad o relevante para el conocimiento y evaluación de las funciones y políticas públicas que son responsabilidad del sujeto obligado.

Art. 40.- Información adicional a ser presentada por el Poder Ejecutivo.- Sin perjuicio de la información requerida de acuerdo a lo dispuesto en el artículo 38, el Poder Ejecutivo deberá hacer pública en su sitio web la siguiente información:

1. El listado de expropiaciones que por razones de utilidad pública se cumplan.
2. La coordinación de proyectos con las Intendencias Municipales, Entes Autónomos y Servicios Descentralizados, Personas Jurídicas de Derecho Público no Estatal así como aquéllos con los sectores empresariales y sociales.
3. El presupuesto que haya sido aprobado por el Parlamento y las adecuaciones presupuestales que se sucedan en las diferentes rendiciones de cuentas.
4. Toda aquella información que se considere de utilidad o sea importante para el conocimiento y evaluación de las funciones y políticas públicas bajo responsabilidad de cada dependencia y entidad.

Anexo II

Modelo de presentación de proyecto para cumplimiento de normativa

PLAN PARA EL CUMPLIMIENTO DEL DECRETO REGLAMENTARIO DEL ART. 82 DE LA LEY N°19.355

Organismo:	
Referente:	

OBJETIVO

<Describir el objetivo del proyecto, puede ser el cumplimiento del decreto reglamentario del art.82 de la ley 19355, la publicación de un conjunto específico de datos para apoyar alguna política particular, etc.>

ALCANCE

<Describir el alcance de las acciones y conjuntos de datos priorizados por el organismo>

EXCLUSIONES

<Indicar los conjuntos de datos que queden excluidos del plan de trabajo>

EQUIPO DEL PROYECTO

<Identificar al equipo asegurándose de contar con las áreas de la organización relacionadas>

Se recomienda que el equipo esté integrado por referentes de diferentes áreas. Especialmente del área informática y legal complementando con otros perfiles que posean conocimiento de los datos que posee la organización.

Rol / Tareas en el proyecto	Nombre	Área donde se desempeña
Patrocinador del proyecto		
Coordinador del proyecto		
Responsable técnico TI		
Asesor legal		
Responsable Comunicaciones		
Responsable de Procesos y Calidad		
(otro que el organismo considere necesario para participar)		

CRONOGRAMA DE ACTIVIDADES

Fases	Etapas	Actividades	Enero				Febrero				
			S1	S2	S3	S4	S1	S2	S3	S4	
Fase1	Identificación										
		Actividad 1									
		Actividad 2									
	Preparación										
	Publicación										

APROBACIÓN DEL DOCUMENTO

Fecha	
Aprobado por	(nombre), (cargo o función)

