

Pautas de calidad para la publicación manual de datos abiertos

Agesic

Versión 1.0 - Agosto 2019

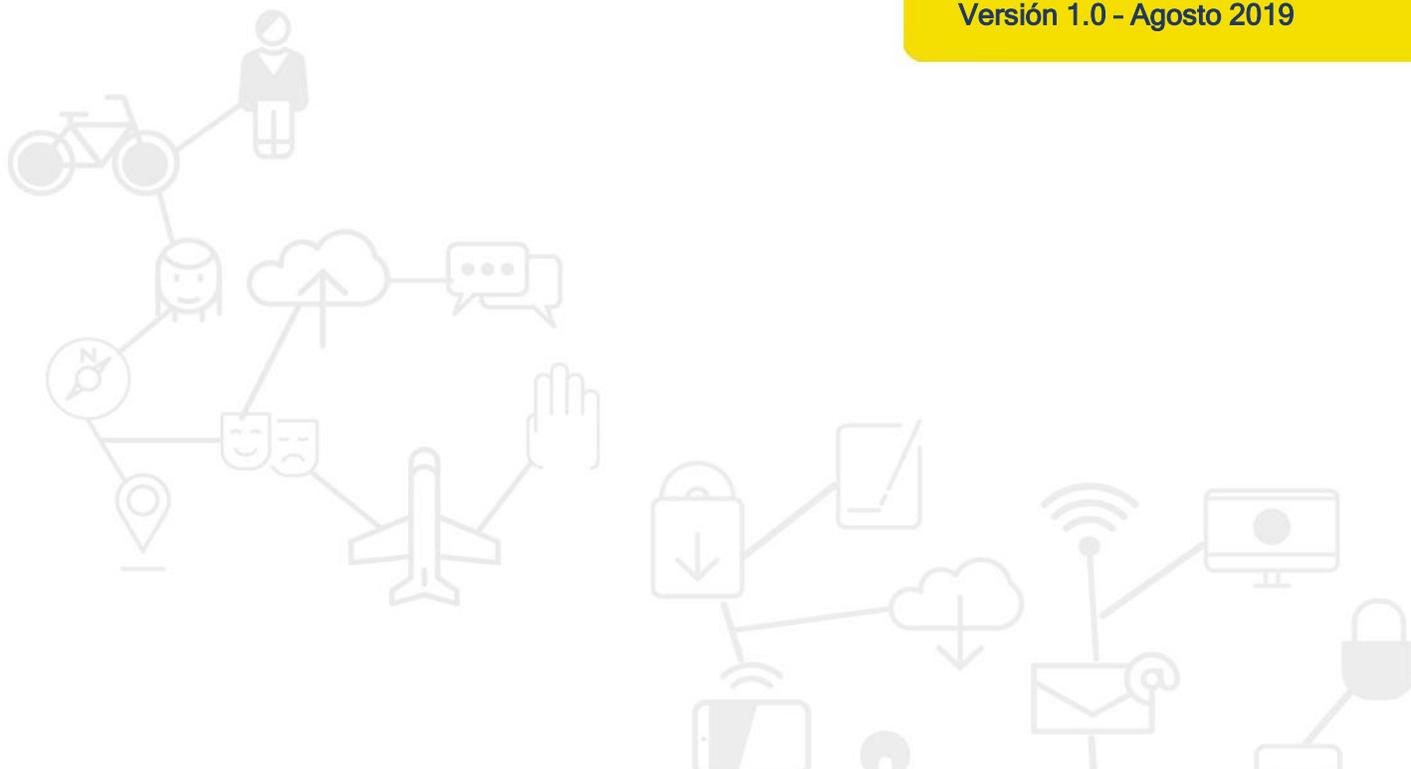


TABLA DE CONTENIDOS

| | |
|---|----|
| Introducción | 2 |
| Etapa 1 – Crear conjunto de datos | 2 |
| Etapa 2 – Agregar datos al conjunto de datos | 7 |
| Etapa 3 – Agregar metadatos al conjunto de datos | 10 |
| Etapa 4 – Agregar categorías al conjunto de datos | 12 |

Introducción

A continuación se presentan pautas y recomendaciones a tener en cuenta para llevar adelante el procedimiento de publicación manual de Datos Abiertos en el Catálogo Nacional de Datos Abiertos. Encontrará las pautas para cada una de las etapas del proceso de publicación.

Etapa 1 - Crear conjunto de datos

A continuación se brinda un conjunto de pautas a tener en cuenta para completar cada uno de los campos, en el momento de crear los conjuntos de datos en el catálogo:

Título:

- Colocar un título que permita entender quien es el publicador, y los datos que contendrá el conjunto de datos.
- No utilizar tildes.
- No incluir la letra ñ.
- No incluir caracteres especiales (*, #,!,\$, _).
- No escribir todo el título en mayúsculas.

Ejemplos:

| Correcto ✓ | Incorrecto ✗ |
|---|---|
| AGESIC - Presupuesto para el año 2019 | Presupuesto |
| MTOP - Cantidad de plazas en el departamento de Montevideo por año entre el 2000 y 2019 | Cantidad de plazas entre el 2000 y 2019! |
| IM - Localización de organismos públicos en el departamento de Montevideo | CANTIDAD DE PLAZAS EN EL DEPTO DE MVO (2000 - 2019) |

Título URL Descripción:

En el catálogo nacional no pueden existir dos conjuntos con igual información en este campo, dado que la URL del conjunto debe ser única.

En caso de existir una URL con idéntica descripción, no se permitirá la creación y guardado del conjunto, deberá realizar modificaciones en el título.

Ejemplos:

| Correcto ✓ | Incorrecto ✗ |
|-----------------------------|---------------------------------|
| .../agesic-presupuesto-2019 | .../presupuesto_agesic_2019 |
| | .../presupuesto-agesic-año-2019 |
| | .../presupuestoagesic2019! |

Descripción:

Es importante considerar que a través de este campo los usuarios de los datos pueden entender que datos son publicados, si corresponden a un período y para que pueden utilizarse. Buscar que la descripción sea lo más detallada possible y sin utilizar términos técnicos. Cualquier persona debería comprenderla.

Ejemplos:

| Correcto ✓ | Incorrecto ✗ |
|---|---|
| <p>Este juego de datos contiene los horarios de ómnibus del transporte colectivo urbano de Montevideo para cada una de las paradas. Estos son los horarios teóricos estimados en los que pasará una determinada línea de ómnibus por una cierta parada a lo largo de su recorrido. Estos datos se obtienen a partir de los datos de Horarios y Paradas estimando los horarios de pasada por cada parada según la velocidad promedio de las unidades de transporte y la distancia entre paradas.</p> <p>Este juego se actualiza diariamente por lo que siempre se dispone de las últimas actualizaciones que se pudieran haber realizado el día anterior</p> <p>Cambios recientes</p> <ul style="list-style-type: none"> ● 7/7/2014 : Se agregan horarios por parada para variantes circulares. <p>Juegos de datos relacionados:</p> <ul style="list-style-type: none"> ● Transporte colectivo: paradas y puntos de control ● Líneas de ómnibus, origen y destino. ● Horarios de ómnibus urbanos - Sistema de Transporte Metropolitano | <p>Información de horarios y sus metadatos asociados.</p> |

Es recomendable agregar información en la **descripción** sobre los cambios de cada versión:

Cambios recientes

- **03/07/2014:** Se adjunta a este juego de datos las variantes circulares.
- **17/09/2013:** Se adjunta a este juego de datos las variantes no maximales. Es decir se unifica este juego de datos con **Líneas de ómnibus (variantes no maximales)** para simplificar el acceso y uso de los mismos.

Etiquetas:

Las etiquetas permiten buscar los conjuntos de datos por temáticas diferentes a las categorías.

Al etiquetar los conjuntos, se recomienda:

- Previamente visualizar las etiquetas existentes en el catálogo. Para acceder a las mismas utilice el siguiente link <https://catalogodatos.gub.uy/tag>.
- Verificar la ortografía de la etiqueta agregada.
- Evitar el uso de caracteres especiales o combinaciones de palabras que dificulten la búsqueda del conjunto de datos, como por ejemplo Tecnología:Informática, \Salud\.
- Una Publicación de datos sobre costos de un organismo gubernamental podría etiquetarse como Economía, Gobierno.

Ejemplo:

- Costos, Salud, Tecnología, Población, Vivienda, Energía, Medio Ambiente.

Licencia:

La licencia es un campo obligatorio para la correcta publicación del conjunto de datos. Y según la normativa, todos los organismos públicos deben seleccionar la Licencia de Datos Abiertos de Uruguay

Organismo:

Verificar que el organismo seleccionado sea correcto.

Visibilidad:

Los conjuntos de datos publicados serán sometidos a un control de calidad por parte de AGESIC que será el encargado de validar que el conjunto de datos pueda permanecer como público para todos los usuarios que visiten el catálogo.

En caso de desear subir un conjunto de datos como una prueba, se debe utilizar el ambiente de test del catálogo y no publicar el conjunto de datos como privado en el catálogo.

Utilice la lista de control de calidad para verificar todos los requisitos para publicar en el catálogo.

Nota: Mientras se encuentre en estado Privado, los datos serán únicamente visibles para los usuarios de la organización que lo publicó.

Fuente:

Es importante indicar la fuente de datos, esto permite garantizar la confiabilidad de los datos publicados.

Ejemplo:

- 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 (URL a la publicación).

Información Adicional

| Campo | Valor |
|----------------------|---|
| Fuente | https://www.comprasestatales.gub.uy/comprasenlinea/jboss/generarReporte |
| Autor | ACCE |
| Mantenedor | Mesa de ayuda |
| Versión | 1.0 |
| Última actualización | 26 de febrero de 2019, 12:34 (UTC-03:00) |
| Creado | 27 de agosto de 2018, 12:29 (UTC-03:00) |
| Licencia | Licencia de DAG de Uruguay |

Versión:

Al igual que en la gestión de los documentos, la versión permite conocer al usuarios si hubo modificaciones en la estructura de los datos

Ejemplo de versionado:

| Versión | Ejemplo |
|---------|--|
| 1.1 | Si el conjunto de datos sufrió un cambio menor. Ejemplo: Se corrigió un error de tipeo, el nombre del conjunto o se agregó información en la descripción. |
| 2.0 | Si el conjunto de datos sufrió un cambio en sus recursos (se agregó uno nuevo o modificó). |

Es recomendable agregar información en la **descripción** sobre los cambios de cada versión:

Cambios recientes

- **03/07/2014:** Se adjunta a este juego de datos las variantes circulares.
- **17/09/2013:** Se adjunta a este juego de datos las variantes no maximales. Es decir se unifica este juego de datos con **Líneas de ómnibus (variantes no maximales)** para simplificar el acceso y uso de los mismos.

Autor:

Se recomienda agregar el nombre del área responsable de la publicación.

Email del Autor (correo electrónico):

Agregar la dirección de correo electrónico del área.

Ejemplo de autor y email de autor:

| | |
|-------------------|------------------------------|
| Empresa en el Día | empresaeneldia@agesic.gub.uy |
| Carlos Pérez | empresaeneldia@agesic.gub.uy |

Mantenedor:

Se recomienda agregar el nombre de la persona o equipo responsable de mantener los datos actualizados.

Email de Mantenedor (correo electrónico):

Agregar el correo electrónico de la persona o equipo responsable del mantenimiento de los datos.

Ejemplo de mantenedor y email de mantenedor:

| Correcto ✓ | Incorrecto ✗ |
|--|---|
| Carlos Pérez empresaeneldia@agesic.gub.uy | Empresa en el Día empresaeneldia@agesic.gub.uy |

Etapa 2 - Agregar datos al conjunto de datos

Dato:

Los datos pueden ser cargados de una sola vez utilizando el botón “Guardar y Agregar otro”.

Nombre:

| Correcto ✓ | Incorrecto ✗ |
|--|------------------------------|
| Datos de personas del censo nacional de psicología para el año 2014. | Datos censo psicología 2014. |

Descripción:

En la descripción de los datos incorporados al conjunto de datos, intentar realizar una descripción detallada y clara, incorporando información de los datos que contiene, e información sobre la fuente de origen

Formato:

A continuación se presentan ejemplos de formatos incorrectos y correctos¹.

| Correcto ✓ | Incorrecto ✗ |
|---|---|
| <p>★★★ Los datos cumplen con el nivel de 2 estrellas, pero además el formato utilizado es no propietario y su acceso y procesamiento no dependen de ninguna organización pública o privada.</p> <p>Ejemplo: Archivo csv, json, xml</p> | <p>★ Los datos están publicados en cualquier formato y con una licencia abierta asociada.</p> <p>Ejemplo: Escaneo de una imagen, o PDF con una imagen</p> |
| <p>★★★★ Los datos cumplen con el nivel de 3 estrellas, y además se utilizan estándares abiertos para identificar elementos abstractos o del mundo real y posibilitar que puedan ser referenciados mediante URIs.</p> <p>Estándares abiertos de W3C (RDF y SPARQL)</p> | <p>★★ Los datos están publicados con una licencia abierta y en un formato estructurado que puede ser interpretado y/o procesado por una computadora.</p> <p>Ejemplo: Hoja de cálculo en formato Excel (.xls, .xlsx)</p> |
| <p>★★★★★ Los datos cumplen con el nivel de 4 estrellas, y además están relacionados (enlazados) con datos de otros conjuntos de datos a través de sus URIs.</p> <p>Estándares abiertos de W3C (RDF y SPARQL) y enlaces URIs</p> | |

¹ Por más información al respecto, se puede consultar la siguiente página:

<http://5stardata.info/es/> en donde se encontrarán ejemplos, ventajas y desventajas de cada nivel de la clasificación.

Sistema de referencia:

Estos son algunos ejemplos de los sistemas geodésicos más utilizados:

- WGS84: Sistema geodésico mundial que data de 1984.
- ED50: Datum europeo de 1950.
- ETRS89: Sistema de referencia terrestre europeo de 1989 muy similar al WGS84.
- NAD83: Datum estadounidense de 1983 el cual es muy similar al WGS84.
- PSAD56: Datum provisional sudamericano de 1956.
- SIRGAS: Sistema de Referencia Geocéntrico para las Américas.

Cobertura Espacial:

Indica que región cubren los datos publicados o su influencia. Por ejemplo: Uruguay, departamento, municipio, barrio, etc.

| Correcto ✓ | Incorrecto ✗ |
|---|--------------------------------------|
| Zona Metropolitana de Montevideo y Costa de Canelones | Todo Montevideo y parte de Canelones |

Cobertura Temporal:

Período de tiempo que cubren los datos publicados. Por ejemplo: 2015, Enero 2016 - Marzo 2017.

| Correcto ✓ | Incorrecto ✗ |
|-------------------|---------------------------|
| 2015 - 2016 | Desde 2015 al 2016 |
| 01/2018 - 01/2019 | Enero 2018 a Febrero 2019 |

Etapa 3 - Agregar metadatos al conjunto de datos

Metadatos:

Los metadatos explican el significado de cada campo o columna de los datos y son tan importantes como los datos mismos, ya que permiten su interpretación a los usuarios reutilizadores. Se debe incorporar un metadato por cada dato cargado.

| 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. | Descripción detallada del significado del atributo orientada a las personas que pretenden utilizar el atributo. | Numérico, Texto, Fecha/Hora, Moneda | 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. | Cualquier información adicional sobre el atributo que complemente la información básica. |
| Ejemplo: Código Sexo | Código asociado al sexo de la persona | Numérico | https://catalogodatos.gub.uy/dataset/9225c6f8-ce7c-468c-8913-aebad8407dcc/resource/0ea23b5a-dd59-4928-a9f1-72f15d1af786/download/sexo_v1.xml | Agestic proporciona el conjunto de datos referenciales para estandarizar los datos. https://catalogodatos.gub.uy/dataset/datos-referenciales-codigueras |

Ejemplo de metadatos en JSON:

```
{
  "atributos": [
    {
      "descripcion": "año",
      "informacionAdicional": "",
      "tipoDeDato": "Number",
      "recursoRelacionado": "",
      "nombreDeAtributo": "año"
    },
    {
      "descripcion": "valor",
      "informacionAdicional": "",
      "tipoDeDato": "Number",
      "recursoRelacionado": "",
      "nombreDeAtributo": "valor"
    }
  ],
  "titulo": "Metadatos de Tasa de reclusos cada 100.000 habitantes. Total país.",
  "descripcion": "Cantidad de reclusos cada 100.000 habitantes.",
  "calculo": "Cantidad de reclusos/Población total*100 mil",
  "unidad": "Cantidad",
  "fuente": "Departamento de Estadísticas",
  "direccion": ""
}
```

Agregar un metadato a partir de un CSV utilizando el diccionario de datos

Si se carga un dato con formato .csv el catálogo brinda la posibilidad de agregar los metadatos asociados a través de una funcionalidad llamada diccionario de datos. La misma identifica los campos del .csv y a partir de ellos se solicita la información descriptiva a completar.

Para los campos identificados:

Clave

Subclave

Conteo

Al acceder al diccionario se debe completar la siguiente información:

[Editar recurso](#)

[DataStore](#)

[Diccionario de datos](#)

[Vistas](#)

Campo 1. Clave (text)

Redefinir Tipo:

Etiqueta:

Descripción:

Puede usar formato [Markdown](#) aquí

Los tipos posibles a completar son: Numeric, Text, Timestamp

Etapa 4 - Agregar categorías al conjunto de datos

Existen diversas categorías pre establecidas disponibles en el catálogo.

Se deberá seleccionar entre alguna de ellas de forma de poder identificar en el catálogo al conjunto de datos de forma más rápida.

En caso de no existir la categoría para su conjunto de datos contactarse con el equipo de datos abiertos de AGESIC por correo electrónico a datosabiertos@agesic.gub.uy.

| | | |
|--|---|---|
|  Alimentación 48 Conjuntos de datos |  Cultura 3 Conjuntos de datos |  Deporte 0 Conjuntos de datos |
|  Desarrollo Social 219 Conjuntos de datos |  Educación 308 Conjuntos de datos |  Economía 85 Conjuntos de datos |
|  Estadísticas 431 Conjuntos de datos |  Industria y energía 11 Conjuntos de datos |  Geográficos 7 Conjuntos de datos |
|  Infraestructura 21 Conjuntos de datos |  Salud 258 Conjuntos de datos |  Medio Ambiente 19 Conjuntos de datos |
|  Seguridad Ciudadana 20 Conjuntos de datos |  Trabajo 230 Conjuntos de datos |  Soc. Demográfico 283 Conjuntos de datos |
|  Transparencia 47 Conjuntos de datos |  Turismo 3 Conjuntos de datos |  Transporte 12 Conjuntos de datos |
|  Violencia 49 Conjuntos de datos | |  Vivienda 188 Conjuntos de datos |