

DATOS E INFORMACIÓN PÚBLICA

Guía Práctica

Para el uso y la publicación de metadatos

ÍNDICE

P3. Glosario

P4. Introducción

Estado y alcance de esta guía

¿Qué son los metadatos?

¿Cómo se publican?

Público objetivo de esta guía

P6. Portal Andino

Catálogo

Dataset

Distribución

Campo

Tema

P14. Otros catálogos

Estándar utilizado

Campos del perfil

- Catálogo
- Dataset
- Distribución
- Campo
- Tema

P29. Anexo I - Taxonomía temática global de la APN para los datasets (tabla)

P31. Anexo II - Pautas para la selección de etiquetas

P32. Anexo III - Especificación de frecuencias (según ISO-8601)

P33. Anexo IV - Ejemplo de data.json

P36. Anexo V - Taxonomía temática global de la APN para los datasets (JSON)

P38. Anexo VI - Ejemplo de metadatos como texto

Glosario

Activo de Datos: Es una colección de datos con valor informativo, de propiedad de una entidad u organismo, que reflejan su actividad, son relevantes para el desarrollo de sus misiones y funciones, o para los actores de su ecosistema.

Distribución o Recurso: Una distribución o recurso es la unidad mínima de un Catálogo de Datos, se trata de los archivos que se publican allí y que pueden ser descargados y reutilizados por un usuario. Los recursos pueden tener diversos formatos (.csv, .shp, etc.) y a su vez están acompañados de información contextual asociada (“metadata”) que describe el tipo de información que se publica, el proceso por el cual se obtiene, la descripción de los campos del recurso y cualquier información extra que facilite su correcta interpretación, procesamiento y lectura.

Dataset: Un conjunto de datos o dataset agrupa recursos referidos a un mismo tema que respetan una estructura de la información, pero difieren en el formato en que se los presenta (Ej. .csv, .json, .xls, etc.), la fecha a la que se refieren, el área geográfica cubierta o están separados bajo algún otro criterio. Al igual que los recursos, cada dataset también debe incluir una documentación contextual anexa referida a los recursos que lo componen.

Catálogo de Datos: Es una colección curada de metadatos sobre datasets.

Perfil de Metadatos: Una lista curada de campos de metadatos, junto con las instrucciones de cómo deben utilizarse.

Paquete para Apertura de Datos de la República Argentina: Conjunto de herramientas y guías de buenas prácticas diseñadas para facilitar la publicación de datos abiertos en la Argentina.

Portal Andino: Portal empaquetado y distribuible desarrollado sobre la base de la plataforma CKAN (portal de datos abiertos) diseñado para facilitar la apertura y federación de Activos de Datos en la Argentina.

Introducción

Este documento forma parte de la publicación de “guías de buenas prácticas” que lleva adelante la Dirección Nacional de Datos e Información Pública para instrumentar la Política de Datos Abiertos impulsada desde el Gobierno Nacional, a partir del Decreto N° 117/2016 del 12 de enero de 2016.

El presente documento tiene por objetivo ofrecer al lector una guía de recomendaciones y buenas prácticas para el uso y la publicación de metadatos.

Estado y alcance de esta guía

Las recomendaciones que aquí pueden encontrarse están basadas en estándares y vocabularios controlados utilizados a nivel nacional e internacional y en la experiencia de trabajo del equipo de la Dirección Nacional de Datos e Información Pública del Ministerio de Modernización de la Nación.

Es nuestra intención que este documento sea considerado como un trabajo colaborativo y en progreso. Por lo dicho, alentamos a organizaciones y ciudadanos a considerarlo de esa manera, y plantear ideas, sugerencias y comentarios que contribuyan a un mejor documento. De ningún modo entendemos a este trabajo como una versión final y completa.

Este documento es un complemento a la “Guía para la publicación de datos en formatos abiertos” y a la “Guía para la identificación y uso de entidades interoperables”, puestos a consideración por la Dirección Nacional de Datos e Información Pública.

¿Qué son los metadatos?

Los metadatos son los “datos sobre los datos”. Son elementos descriptivos que le dan contexto a un conjunto de datos, proveen al usuario de la información necesaria para entenderlos y utilizarlos eficazmente.

Un **título** y una breve **descripción** son los metadatos más fundamentales que cualquier conjunto de datos a publicar debería tener. Luego existen otros muchos elementos que ayudan al lector a hacer un buen uso de los datos, por ejemplo:

- **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?)

Una lista curada de campos de metadatos, junto con las instrucciones de cómo deben utilizarse, configura un **perfil de metadatos**.

En esta guía se describe el **Perfil de Metadatos recomendado para los catálogos de la Administración Pública Nacional** y cómo publicarlo.

¿Cómo se publican?

La publicación de los metadatos puede ser muy diversa en detalle, calidad y forma. Una publicación muy elemental consistiría en un **documento de texto** que ofrezca una descripción del dataset y de cada uno de los recursos que lo componen. Ver un [ejemplo textual](#) de los metadatos de un Catálogo de Datos en el [Anexo VI - Ejemplo de metadatos como texto](#).

Sin embargo, las computadoras no pueden leer fácilmente documentos de texto. La organización sistemática de colecciones de datasets (es decir, la creación de un **catálogo** de datos) exige un nivel de complejidad mayor para facilitar su descubrimiento, indexación, y reutilización por parte de scripts y aplicaciones de todo tipo.

La potencial reutilización de los conjuntos de datos y la posibilidad de que los mismos sean correctamente federados desde el Portal Nacional de Datos Abiertos (datos.gob.ar) dependerá de la calidad de sus metadatos. Adoptar y/o desarrollar estándares y vocabularios controlados es importante para la lectura e interpretación de los conjuntos de datos por personas y por aplicaciones informáticas.

Para esto, los catálogos de datos publican sus metadatos en un formato estructurado (JSON) respetando un determinado perfil estandarizado. Ver un [ejemplo en JSON](#) de los metadatos de un Catálogo de Datos en el [Anexo IV - Ejemplo de data.json](#).

En el resto de este documento se detallan las características de los estándares y vocabularios controlados adoptados para catálogos de datos, datasets y distribuciones.

Público objetivo de esta guía

Esta guía apunta a dos públicos distintos, quienes:

- **Publican sus catálogos de datos mediante un Portal Andino:** Generan una instancia del Portal Andino ofrecido en el Paquete para Apertura de Datos de la República Argentina. El Portal Andino es un paquete que pueden adoptar las dependencias de la Administración Pública Nacional (así como también de otros niveles gubernamentales sub-nacionales o incluso entidades del público general) para gestionar la publicación centralizada de sus Activos de Datos. El Portal Andino incluye formularios web para la carga de datos y metadatos. La publicación de metadatos mediante este Portal cumple con el perfil de metadatos recomendado en esta guía. El Portal debe publicarse en un dominio de la forma [http://datos.\[entidad\].gob.ar](http://datos.[entidad].gob.ar). Esto asegura la publicación automática de los metadatos en formato data.json en [http://datos.\[entidad\].gob.ar/data.json](http://datos.[entidad].gob.ar/data.json)
- **Publican sus catálogos de datos mediante otros medios:** Los Activos de Datos se publican mediante otra herramienta, en forma directa en el dominio digital de la entidad responsable o de alguna forma alternativa en la red. Esas entidades deberán publicar su catálogo de datos en un archivo estructurado (JSON) llamado *data.json* alojado en un dominio de la forma [http://datos.\[entidad\].gob.ar/data.json](http://datos.[entidad].gob.ar/data.json) siguiendo las especificaciones del perfil de metadatos de esta guía.

En [entidad] se recomienda utilizar un nombre simple y breve que represente a la organización que publica el catálogo (Ejemplos: datos.jus.gob.ar, datos.tucuman.gob.ar, datos.pilar.gob.ar)

Los **primeros** deberían leer la sección “*Portal Andino*” para comprender cómo completar cada campo en el formulario web del Portal Andino.

Los **segundos** deberían leer la sección “*Otros catálogos*” para comprender cómo deben generar y publicar un archivo *data.json* que cumpla con el perfil de metadatos de la Administración Pública Nacional.

Portal Andino

El Portal Andino es una herramienta especialmente diseñada para facilitar la publicación y apertura de datos. En este sentido, ha sido diseñado de forma tal que cumpla con el Perfil de Metadatos de la Administración Pública Nacional.

El usuario de este Portal carga datasets y distribuciones a través de formularios que generan los metadatos del perfil y los publican en un archivo *data.json* en el formato correcto.

El archivo *data.json* de un Portal Andino puede encontrarse en el directorio raíz del Portal: [http://datos.\[entidad\].gob.ar/data.json](http://datos.[entidad].gob.ar/data.json).

Catálogo

Como se verá más adelante, el perfil de metadatos propuesto se conforma de tres componentes principales, la información del catálogo, los datasets y las distribuciones.

El *data.json* de quienes utilicen el Portal Andino contará con los metadatos correspondientes a los datasets y distribuciones. Los metadatos a nivel de catálogo deberán ser completados en una planilla de cálculo y publicados en formato CSV en el directorio raíz del Portal, quedando disponibles para su descarga en una URL como [http://datos.\[entidad\].gob.ar/catalog.csv](http://datos.[entidad].gob.ar/catalog.csv). El resto de los metadatos generados al cargar o actualizar un dataset o una distribución en el portal, se generan a través de los formularios completados y se publican automáticamente en [http://datos.\[entidad\].gob.ar/data.json](http://datos.[entidad].gob.ar/data.json).

Ver un [ejemplo textual](#) de metadatos de un **catálogo**.

A continuación se describen los metadatos necesarios para describir el Catálogo:

| Nombre | Requerido | Descripción | Ejemplo |
|--|-------------|---|---|
| Nombre | Si | Nombre dado al catálogo. Debe ser claro, breve y lo suficientemente abstracto como para abarcar la multiplicidad de datasets que contiene. | Datos Argentina |
| Descripción | Si | Descripción del contenido del catálogo. | Portal de Datos Abiertos del Gobierno de la República Argentina |
| Autor | Si | Responsable de la publicación del catálogo. | Ministerio de Modernización |
| Correo electrónico del autor | Si | Correo electrónico de contacto del responsable de la publicación del catálogo. | datos@modernizacion.gob.ar |
| Fecha de creación o publicación | Recomendado | Fecha de creación o publicación del catálogo. Se escribe según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset. | "2016-04-14T19:48:05.433640" para especificar fecha y hora "2016-04-14" para especificar fecha únicamente |
| Fecha de última actualización/modificación | Recomendado | Fecha de última actualización/modificación del catálogo. Se escribe según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset. | "2016-04-19T19:48:05.433640" para especificar fecha y hora "2016-04-19" para especificar fecha únicamente |
| Idioma(s) | Recomendado | Lenguaje para la descripción de los metadatos de los datasets contenidos en el catálogo. Hay 2 estándares ISO que pueden ser utilizados para este campo: (a) ISO 639-1 (2 letras) (b) ISO 639-2/T (3 letras) es el más recomendado. Puede definirse 1 o más lenguajes en una lista. (Link a los estándares ISO: https://www.loc.gov/standards/iso639-2/php/code_list.php) | ["es"] para un lenguaje ISO 639-1 ["spa", "eng"] para dos lenguajes ISO 639-2 |
| Taxonomía temática global | Si | Es el sistema de clasificación temática global de la Administración Pública Nacional. Compone una lista de temas globales y está publicada en "http://datos.gob.ar/superThemeTaxonomy.json". | http://datos.gob.ar/superThemeTaxonomy.json |
| Licencia | Recomendado | Indica la licencia bajo la cual todos los datasets y distribuciones del catálogo están disponibles mediante un enlace a la licencia o documento de la licencia seleccionada, o mediante el título textual de la licencia tal como aparece en la lista de http://opendefinition.org/licenses/ . Se recomienda utilizar la licencia "Open Database License (ODbL) v1.0". Un dataset o distribución que especifique una licencia diferente, sobreescribe a la licencia general del catálogo. | "http://opendatacommons.org/licenses/dbcl/1-0/" si se utiliza un enlace "Open Database License (ODbL) v1.0" si se consigna el nombre de la licencia a utilizar |

| | | | |
|-----------------------------------|-------------|---|--|
| Página web del catálogo | Recomendado | Dirección web de acceso a la página principal del catálogo. Enlace a la página principal del catálogo. | http://datos.gob.ar |
| Derechos sobre el catálogo | No | Información sobre derechos aplicables al catálogo en el caso que no se hayan especificado en la licencia. Los datasets y sus distribuciones heredan la información sobre derechos aplicables al catálogo, a menos que especifiquen unos derechos distintos. | |
| Cobertura geográfica | No | <p>El área geográfica cubierta por el catálogo. Una región o un lugar determinado. Puede tomar valores:</p> <p>a) de países y, provincias y municipios argentinos, según las recomendaciones de la "Guía para la identificación y uso de entidades interoperables".</p> <p>b) un área de coordenadas representada por latitud/ longitud en el orden: mínima longitud, mínima latitud, máxima longitud, máxima latitud.</p> <p>c) un punto geográfico representado por latitud/longitud.</p> <p>d) Si la referencia geográfico no está identificada en la "Guía para la identificación y uso de entidades interoperables" indicar la URIs según geonames.org; ej : http://sws.geonames.org/6255146</p> | <p>"ARG" es el código para la República Argentina.</p> <p>"06007" es el código de un departamento</p> <p>[-58.111111, -35.111111, -57.111111, -33.111111] es un bounding box</p> <p>[-58.111111, -35.111111] es un punto geográfico</p> <p>"http://sws.geonames.org/6255146"</p> |

Dataset

Ver un [ejemplo textual](#) de metadatos de un **dataset**.

A continuación se describen los metadatos que se deben completar para describir un Dataset a la hora de su carga o actualización en el Catálogo:

| Nombre | Requerido | Descripción | Ejemplo |
|--------------------------------|-----------|---|--|
| Título del dataset | Si | Nombre asignado al dataset tal como será publicado. Debe ser claro y lo suficientemente abstracto como para abarcar la multiplicidad de distribuciones que contiene. Se recomienda no exceder los 100 caracteres en la mayoría de los casos. En caso de que un título más largo se juzgue necesario o relevante, este no deberá exceder los 200 caracteres. | Sistema de contrataciones electrónicas |
| Descripción del dataset | Si | Descripción del contenido del dataset de un modo claro y conciso. Se recomienda no exceder los 500 caracteres en la mayoría de los casos. En caso de que una descripción más larga se juzgue | Datos correspondiente al Sistema de Contrataciones Electrónicas (Argentina Compra) |

| | | | |
|-------------------------------|-------------|--|---|
| | | necesaria o relevante, esta no deberá exceder los 1500 caracteres. | |
| Responsable | Si | Responsable de la publicación del dataset. En el caso de organizaciones, detallar la estructura jerárquica separada por puntos, de manera jerárquicamente descendiente. Si la organización es parte de la Administración Pública Nacional y está listada en el dataset llamado "Estructura Organica del Poder Ejecutivo Nacional" (http://datos.gob.ar/dataset/estructura-organica-pen), deberá utilizarse la denominación allí documentada. | Ministerio de Modernización. Secretaría de Modernización Administrativa. Oficina Nacional de Contrataciones. |
| E-mail del responsable | Recomendado | Correo electrónico de contacto del responsable de la publicación del dataset. | onc@modernizacion.gob.ar |
| Mantenedor | Recomendado | Área/persona de contacto que puede brindar información relevante sobre el dataset. | Ministerio de Modernización. Secretaría de Modernización Administrativa. Oficina Nacional de Contrataciones. Dirección de Compras Electrónicas. |
| E-mail del mantenedor | Recomendado | Correo electrónico del área/persona de contacto que puede brindar información relevante sobre el dataset. | onc-compraselectronicas@modernizacion.gob.ar |
| Temas globales | Si | Temática/s o categoría/s globales a la/s que se refiere el dataset al ser publicado. Un dataset puede pertenecer a más de una categoría global, de manera que el tipo de valor de este campo es una lista de categorías. La/s categoría/s o temática/s globales deben adoptarse según el campo "Código (authority code)" del <u>Anexo "Taxonomía temática para los datasets"</u> , que contiene una lista predefinida de temática/s globales. | ["ECON"] |
| Temas específicos | Recomendado | Temática/s o categoría/s específica/s a la/s que se refiere el dataset al ser publicado. Un dataset puede pertenecer a más de una categoría específica, de manera que el tipo de valor de este campo es una lista de categorías. La taxonomía a utilizar debe ser creada por la autoridad responsable del Catálogo. | ["Contrataciones", "Compras", "Macroeconomía"] |
| Etiquetas | Recomendado | Palabras que describen el título o el contenido del recurso. Es necesario que las etiquetas se encuentren correctamente escritas, en plural y respetando la existencia de tags anteriores. Etiquetas que colaboran en la búsqueda de los usuarios. Cuanto más amplia y uniforme sea la lista de tags mayor será su eficiencia. A tales | ["bienes", "compras", "contrataciones"] |

| | | | |
|------------------------------------|----|---|--|
| | | <p>fines se recomienda ver el <u>Anexo "Pautas para la selección de etiquetas"</u>.</p> | |
| Frecuencia de actualización | Si | <p>Frecuencia con la que se actualiza el dataset. Se recomienda especificar períodos normalizados con formato ISO-8601, agregando el campo "eventual" para datasets que se publican con una frecuencia eventual o no especificada. <u>Anexo "Especificación de frecuencias según ISO-8601"</u>.</p> | <p>"R/P1Y" para datasets que se actualizan anualmente</p> |
| Fecha de publicación | Si | <p>Fecha de publicación del dataset. Según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset.</p> | <p>"2016-04-14T19:48:05.433640" para especificar fecha y hora</p> <p>"2016-04-14" para especificar fecha únicamente</p> |
| Idioma | No | <p>Lenguaje para la descripción de los metadatos del dataset. Hay 2 estándares ISO que pueden ser utilizados para este campo:</p> <p>(a) ISO 639-1 (2 letras)</p> <p>(b) ISO 639-2/T (3 letras) es el más recomendado.</p> <p>Puede definirse 1 o más lenguajes en una lista. Los lenguajes especificados para un dataset, sobreescriben a los del catálogo. Si este campo está vacío el dataset hereda los lenguajes del catálogo. (Link a los estándares ISO: https://www.loc.gov/standards/iso639-2/php/code_list.php)</p> | <p>["es"] para un lenguaje ISO 639-1</p> <p>["spa", "eng"] para dos lenguajes ISO 639-2</p> |
| Cobertura geográfica | No | <p>Una región o lugar determinado al que el dataset haga referencia.</p> <p>Una región o un lugar determinado. Puede tomar valores:</p> <p>a) de países y, provincias y municipios argentinos, según las recomendaciones de la "Guía para la identificación y uso de entidades interoperables".</p> <p>b) un área de coordenadas representada por latitud/ longitud en el orden: minima longitud, minima latitud, maxima longitud, maxima latitud.</p> <p>c) un punto geográfico representado por latitud/longitud.</p> <p>d) Si la referencia geográfico no está identificada en la "Guía para la identificación y uso de entidades interoperables" indicar la URIs según geonames.org; ej : http://sws.geonames.org/6255146</p> | <p>"ARG" es el código para la República Argentina.</p> <p>"06007" es el código de un departamento</p> <p>[-58.111111, -35.111111, -57.111111, -33.111111] es un bounding box</p> <p>[-58.111111, -35.111111] es un punto geográfico</p> <p>"http://sws.geonames.org/6255146"</p> |

| | | | |
|---------------------------|-------------|--|---|
| Cobertura temporal | Recomendado | Período de tiempo cubierto por el dataset. El intervalo de tiempo está formado por una fecha de inicio y una de fin separadas por “/”, en formato ISO 8601, con el nivel de especificidad requerido por el dataset. | 2015-01-01/2015-12-31 2015-01-01T00:45:00Z/2016-01-15T00:06:00Z |
| Fuente | No | URL de una página web a través de la cual se puede acceder al dataset, sus recursos o información adicional sobre el mismo. | http://datos.gob.ar/dataset/sistema-de-contrataciones-electronicas-argentina-compra |
| Licencia | Recomendado | Indica la licencia bajo la cual el dataset y todas sus distribuciones están disponibles mediante un enlace a la licencia o documento de la licencia seleccionada, o mediante el título textual de la licencia tal como aparece en la lista de http://opendefinition.org/licenses/ . Se recomienda utilizar la licencia "Open Database License (ODbL) v1.0". Un dataset hereda por default la licencia general del catálogo salvo que se especifique una licencia diferente en este campo. Las distribuciones del dataset heredan esta licencia salvo que especifiquen una diferente. | "http://opendatacommons.org/licenses/dbcl/1-0/" si se utiliza un enlace "Open Database License (ODbL) v1.0" si se consigna el nombre de la licencia a utilizar |

Distribución

Ver un [ejemplo textual](#) de metadatos de una **distribución**.

A continuación se describen los metadatos que se deben completar al cargar o actualizar una Distribución de un Dataset en el Catálogo para describir a la misma:

| Nombre | Requerido | Descripción | Ejemplo |
|--------------------------------|-------------|---|--|
| Descripción del recurso | Recomendado | Breve descripción de la distribución. Se recomienda no escribir más de una línea. Toda información adicional puede ser incluida en la descripción del dataset. | Listado de las convocatorias abiertas durante el año 2015 en el sistema de contrataciones electrónicas |
| Formato | Recomendado | Indica el formato del archivo de la distribución. Si el tipo de la distribución está definido por IANA (http://www.iana.org/assignments/media-types/media-types.xml), debe usarse esa definición. En caso contrario deberán ponerse los caracteres después del punto final del archivo, que determinan el formato (cuando no está definido por IANA). | "text/csv" definición de IANA "csv" caracteres finales después del punto |
| Título del recurso | Si | Nombre asignado a la distribución. | Convocatorias abiertas durante el año 2015 |
| Licencia | Recomendado | Indica la licencia bajo la cual la distribución está disponible mediante un enlace a la licencia o documento de la licencia seleccionada, o mediante el | "http://opendatacommons.org/licenses/dbcl/1-0/" si se utiliza un enlace |

| | | | |
|---------------------------------------|-------------|--|--|
| | | título textual de la licencia tal como aparece en la lista de http://opendefinition.org/licenses/ . Se recomienda utilizar la licencia "Open Database License (ODbL) v1.0". Una distribución hereda por default la licencia del dataset al que pertenece, salvo que se especifique una licencia diferente en este campo. | "Open Database License (ODbL) v1.0" si se consigna el nombre de la licencia a utilizar |
| Derechos sobre la distribución | No | Información sobre derechos aplicables a la distribución que no se hayan especificado en la licencia. Si se especifican, estos derechos sobreesciben a los del catálogo. En caso contrario, las distribuciones heredan los derechos especificados para el catálogo. | |
| Campos de la distribución | Recomendado | Lista de campos que contiene una distribución tabular (no aplica para aquellas distribuciones que no sean tablas) y sus metadatos. Cada campo se representa con un objeto ("{}") donde se describen los metadatos especificados para la clase "field" de este perfil de metadatos (como "nombre", "tipo" y "descripción"). | [{...}, {...}] |

Campo

Como se ve en la tabla anterior (metadatos de una Distribución), se recomienda que las distribuciones tabulares incluyan algunos metadatos que ayuden a entender la información que contiene cada campo. Estos metadatos se completan en el mismo formulario que se utiliza para cargar o actualizar una distribución.

A continuación se describen los metadatos que el responsable de cargar o actualizar una Distribución de un Dataset en el Catálogo debe completar para describir los campos de la misma, en caso de corresponder:

| Nombre | Requerido | Descripción | Ejemplo |
|-------------------------|-------------|--|---|
| Nombre del campo | Recomendado | El nombre del campo tal como se denomina en el encabezado de la distribución. Véase la "Guía para la publicación de datos en formatos abiertos" para una adecuada elección del nombre de un campo. | Ejemplo para el cuarto campo de la distribución "Convocatorias abiertas durante el año 2015", valor para el nombre: "unidad_operativa_contrataciones_desc" |

| | | | |
|-------------------------|-------------|---|---|
| Tipo de datos del campo | Recomendado | <p>El tipo de dato contenido en el campo según la lista utilizada por la librería recline.js (http://okfnlabs.org/recline/docs/models.html#types). Los tipos incluidos en esta lista son:</p> <p>string (text): Valores de texto.</p> <p>number (double, float, numeric): Números que puedan no ser enteros (incluyen decimales).</p> <p>integer (int): Números que siempre son enteros.</p> <p>date: Fecha simple expresada según el estándar ISO 8601 incluyendo únicamente año, mes y día (YYYY-MM-DD) como en "2016-02-01".</p> <p>time: Tiempo expresado según el estándar ISO 8601 incluyendo únicamente horas, minutos y segundos (hh:mm:ss) como en "10:05:00".</p> <p>date-time (datetime, timestamp): Fecha completa expresada según el estándar ISO 8601 incluyendo año, mes, día, horas, minutos y segundos (YYYY-MM-DDThh:mm:ssZ) como en "2016-02-01T10:05:00+03:00"</p> <p>boolean (bool): Valores verdadero/falso.</p> <p>binary: Representación de datos binarios base64.</p> <p>geo_point: Ver estructura en http://www.elasticsearch.org/guide/reference/mapping/geo-point-type.html.</p> <p>geojson: Ver en http://geojson.org/</p> <p>array: Lista de valores.</p> <p>object (json): Objeto de JSON.</p> <p>any: Campo que puede contener valores de cualquier tipo.</p> | Ejemplo para el campo "unidad_operativa_contrataciones_desc" de la distribución "Convocatorias abiertas durante el año 2015", valor para tipo: "string" |
| Descripción del campo | Recomendado | La descripción de la información que contiene el campo. | Ejemplo para el campo "unidad_operativa_contrataciones_desc" de la distribución "Convocatorias abiertas durante el año 2015", valor para descripción: "Organismo que realiza la convocatoria. Organismo de máximo nivel jerárquico al que pertenece la unidad operativa de contrataciones." |

Tema

Cada Catálogo de Datos puede tener su propia taxonomía temática que permite clasificar a los datasets como pertenecientes a una o más categorías temáticas. Se recomienda que los temas tengan algunos metadatos que ayuden a un usuario a entenderlos mejor.

A continuación se describen los metadatos que el responsable de cargar o actualizar la taxonomía temática de un Catálogo debe completar para describir los temas de la misma:

| Nombre | Requerido | Descripción | Ejemplo |
|-------------------------------|-------------|---|---|
| Código identificador del tema | Recomendado | El identificador del tema. | AGRI |
| Etiqueta o título del tema | Recomendado | La etiqueta o título de un tema. | Agricultura, pesca, silvicultura y alimentación |
| Descripción del tema | Recomendado | Una breve y concisa descripción del tema. | Bajo este concepto se incluyen datasets que cubren dominios tales como agricultura, pesca, forestación o alimentos. |

Otros catálogos

Los usuarios que opten por una solución alternativa al Portal Andino para publicar sus datos, deberán publicar los metadatos de su Catálogo en un archivo *data.json* en una URL como la siguiente: [http://datos.\[entidad\].gob.ar/data.json](http://datos.[entidad].gob.ar/data.json).

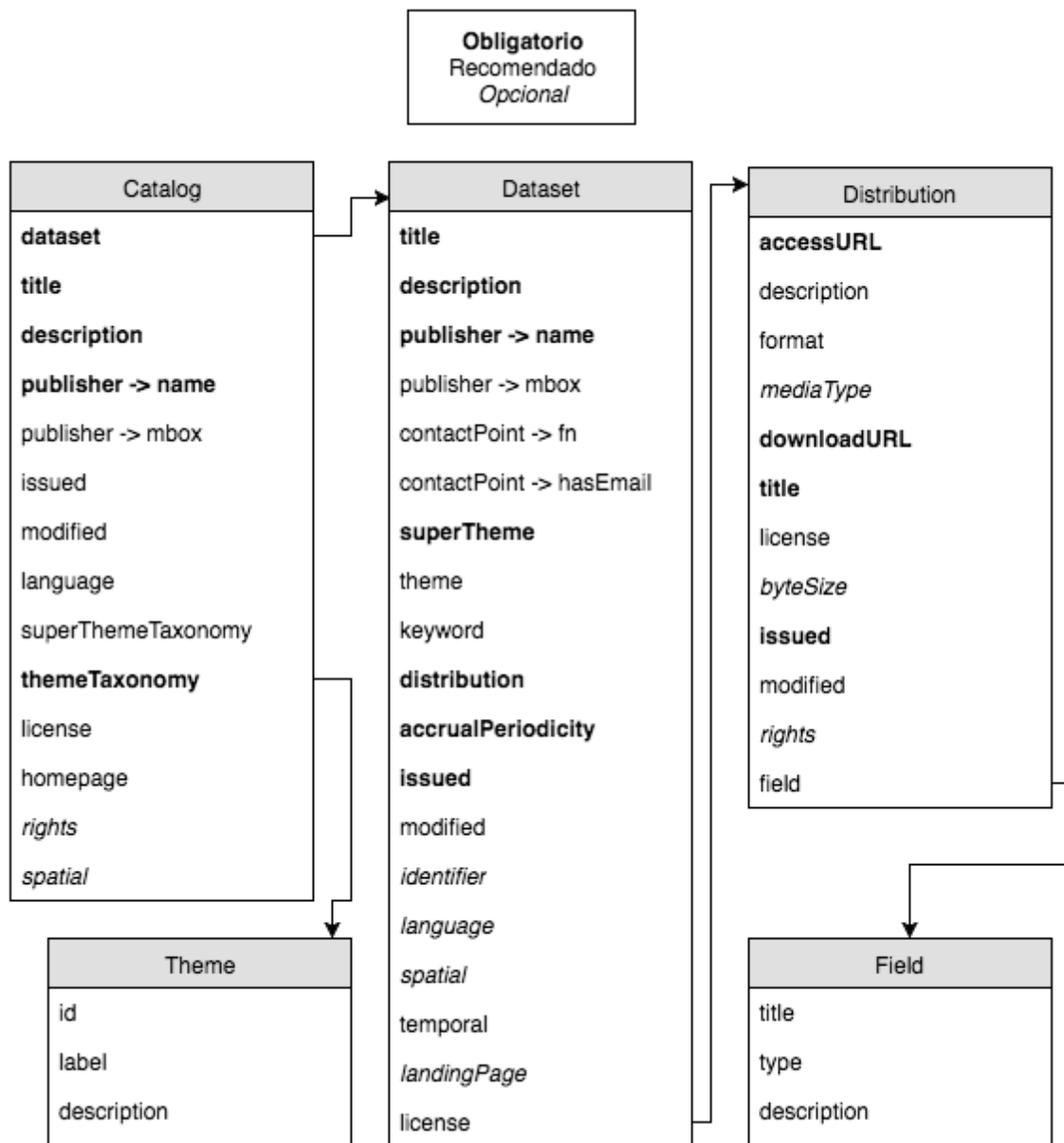
Este archivo deberá estar construido tal como se puede ver en el ejemplo del **Anexo IV - Ejemplo de data.json**, respetando el Perfil de Metadatos de la Administración Pública Nacional tal como se lo describe más adelante en la sección “*Campos del perfil*”.

Estándar utilizado

El Perfil de Metadatos recomendado en esta guía es una adaptación del estándar DCAT - AP, utilizado por los países de la Unión Europea. DCAT es un vocabulario controlado definido por la W3C, ampliamente utilizado a nivel global para la descripción de Catálogos de Datos.

Según la W3C: “Mediante la utilización de DCAT para describir datasets en catálogos de datos, quienes publican aumentan la posibilidad de descubrimiento (discoverability) y permiten a aplicaciones informáticas consumir metadatos de manera simple desde múltiples catálogos. Además permite la publicación descentralizada de catálogos y favorece la búsqueda “federada” de datasets a través de varios sitios.”

El Perfil de Metadatos propuesto para la Administración Pública Nacional se compone de 3 clases principales (Catalog, Dataset y Distribution) y 2 auxiliares (Field y Theme) que se relacionan según el siguiente esquema:



En las secciones subsiguientes, se describen en detalle los metadatos que el *data.json* debe contener, para cada una de estas clases.

Campos del perfil

Catálogo

[Link](#) para ver un ejemplo de metadatos de un catálogo en JSON.

A continuación se describen los metadatos que el *data.json* debe contener, a nivel de Catálogo:

| Nombre | Requerido | Descripción | Ejemplo | Variable (data.json) | Tipo (data.json) |
|---|-------------|--|--|--------------------------|------------------|
| Datasets | Si | Contiene una lista de los datasets que forman parte del catálogo. | [[...], {...}] | dataset | Array |
| Nombre | Si | Nombre dado al catálogo. Debe ser claro, breve y lo suficientemente abstracto como para abarcar la multiplicidad de datasets que contiene. | Datos Argentina | title | String |
| Descripción | Si | Descripción del contenido del catálogo. | Portal de Datos Abiertos del Gobierno de la República Argentina | description | String |
| Autor | Si | Responsable de la publicación del catálogo. | Ministerio de Modernización | publisher name -> | String |
| Correo electrónico del autor | Si | Correo electrónico de contacto del responsable de la publicación del catálogo. | datos@modernizacion.gob.ar | publisher mbox -> | String |
| Fecha de creación o publicación | Recomendado | Fecha de creación o publicación del catálogo. Se escribe según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset. | "2016-04-14T19:48:05.433640" para especificar fecha y hora "2016-04-14" para especificar únicamente fecha | issued | String |
| Fecha de última actualización / modificación | Recomendado | Fecha de última actualización/modificación del catálogo. Se escribe según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset. | "2016-04-19T19:48:05.433640" para especificar fecha y hora "2016-04-19" para especificar únicamente fecha | modified | String |
| Idioma(s) | Recomendado | Lenguaje para la descripción de los metadatos de los datasets contenidos en el catálogo. Hay 2 | ["es"] para un lenguaje ISO 639-1 | language | Array |

| | | | | | |
|--------------------------------------|-------------|---|--|---------------------------|--------|
| | | <p>estándares ISO que pueden ser utilizados para este campo:</p> <p>(a) ISO 639-1 (2 letras)</p> <p>(b) ISO 639-2/T (3 letras) es el más recomendado.</p> <p>Puede definirse 1 o más lenguajes en una lista. (Link a los estándares ISO: https://www.loc.gov/standards/iso639-2/php/code_list.php)</p> | ["spa", "eng"] para dos lenguajes ISO 639-2 | | |
| Taxonomía temática global | Si | Es el sistema de clasificación temática global de la Administración Pública Nacional. Compone una lista de temas globales y está publicada en " http://datos.gob.ar/superThemeTaxonomy.json ". | http://datos.gob.ar/superThemeTaxonomy.json | superThemeTaxonomy | String |
| Taxonomía temática específica | Recomendado | Es el sistema de clasificación temática específica, creado por la organización responsable del catálogo. Compone una lista de temas específicos a los datasets del catálogo. Si se clasifica algún dataset del catálogo como perteneciente a uno o más temas, este campo es obligatorio ya que se debe explicitar una taxonomía temática para poder usar sus temas. | [{...}, {...}] | themeTaxonomy | Array |
| Licencia | Recomendado | Indica la licencia bajo la cual todos los datasets y distribuciones del catálogo están disponibles mediante un enlace a la licencia o documento de la licencia seleccionada, o mediante el título textual de la licencia tal como aparece en la lista de http://opendefinition.org/licenses/ . Se recomienda utilizar la licencia "Open Database License (ODbL) v1.0". Un dataset o distribución que especifique una licencia diferente, sobreescribe a la licencia general del catálogo. | <p>"http://opendatacommons.org/licenses/dbcl/1-0/" si se utiliza un enlace</p> <p>"Open Database License (ODbL) v1.0" si se consigna el nombre de la licencia a utilizar</p> | license | String |
| Página web del catálogo | Recomendado | Dirección web de acceso a la página principal del catálogo. Enlace a la página principal del catálogo. | http://datos.gob.ar | homepage | String |

| | | | | | |
|----------------------------|----|---|--|---------|--------|
| Derechos sobre el catálogo | No | Información sobre derechos aplicables al catálogo en el caso que no se hayan especificado en la licencia. Los datasets y sus distribuciones heredan la información sobre derechos aplicables al catálogo, a menos que especifiquen unos derechos distintos. | | rights | String |
| Cobertura geográfica | No | <p>El área geográfica cubierta por el catálogo.</p> <p>Una región o un lugar determinado. Puede tomar valores:</p> <p>a) de países y, provincias y municipios argentinos, según las recomendaciones de la “Guía para la identificación y uso de entidades interoperables”.</p> <p>b) un área de coordenadas representada por latitud/ longitud en el orden: minima longitud, minima latitud, maxima longitud, maxima latitud.</p> <p>c) un punto geográfico representado por latitud/longitud.</p> <p>d) Si la referencia geográfico no está identificada en la “Guía para la identificación y uso de entidades interoperables” indicar la URIs según geonames.org; ej :</p> <p>http://sws.geonames.org/6255146</p> | <p>"ARG" es el código para la República Argentina.</p> <p>"06007" es el código de un departamento</p> <p>[-58.111111, -35.111111, -57.111111, -33.111111] es un bounding box</p> <p>[-58.111111, -35.111111] es un punto geográfico</p> <p>"http://sws.geonames.org/6255146"</p> | spatial | String |

Debe prestarse especial atención a los dos campos que contienen una lista de objetos: **dataset** y **themeTaxonomy**.

El **primero** contendrá una lista de objetos que describen (cada uno) los metadatos de los distintos datasets que componen el catálogo (en la próxima sección se describen los metadatos que debe contener cada uno de estos objetos).

El **segundo** también contiene una lista de objetos, que juntos definen una taxonomía temática para el Catálogo. Cada uno de estos objetos contiene los metadatos que describen a cada uno de los temas de esta taxonomía. Más adelante se describen estos metadatos en la sección [Tema](#).

Dataset

[Link](#) para ver un ejemplo de metadatos de un dataset en JSON.

A continuación se describen los metadatos que el *data.json* debe contener, para describir a un dataset dentro de la lista contenida en el campo **dataset** del Catálogo:

| Nombre | Requerido | Descripción | Ejemplo | Variable (data.json) | Tipo (data.json) |
|-------------|-----------|---|--|--------------------------|------------------|
| Título | Si | Nombre asignado al dataset tal como será publicado. Debe ser claro y lo suficientemente abstracto como para abarcar la multiplicidad de distribuciones que contiene. Se recomienda no exceder los 100 caracteres en la mayoría de los casos. En caso de que un título más largo se juzgue necesario o relevante, este no deberá exceder los 200 caracteres. | Sistema de contrataciones electrónicas | title | String |
| Descripción | Si | Descripción del contenido del dataset de un modo claro y conciso. Se recomienda no exceder los 500 caracteres en la mayoría de los casos. En caso de que una descripción más larga se juzgue necesaria o relevante, esta no deberá exceder los 1500 caracteres. | Datos correspondiente al Sistema de Contrataciones Electrónicas (Argentina Compra) | description | String |
| Autor | Si | Responsable de la publicación del dataset. En el caso de organizaciones, detallar la estructura jerárquica separada por puntos, de manera jerárquicamente descendiente. Si la organización es parte de la Administración Pública Nacional y está listada en el dataset llamado "Estructura Organica del Poder Ejecutivo Nacional" (http://datos.gob.ar/dataset/est | Ministerio de Modernización. Secretaría de Modernización Administrativa. Oficina Nacional de Contrataciones. | publisher name -> | String |

| | | | | | |
|--|-------------|---|---|---------------------------------|--------|
| | | ructura-organica-pen), deberá utilizarse la denominación allí documentada. | | | |
| Correo electrónico del autor | Recomendado | Correo electrónico de contacto del responsable de la publicación del dataset. | onc@modernizacion.gob.ar | publisher mbox -> | String |
| Área/Persona de contacto | Recomendado | Área/persona de contacto que puede brindar información relevante sobre el dataset. | Ministerio de Modernización. Secretaría de Modernización Administrativa. Oficina Nacional de Contrataciones. Dirección de Compras Electrónicas. | contactPoint -> fn | String |
| Correo electrónico del área/persona de contacto | Recomendado | Correo electrónico del área/persona de contacto que puede brindar información relevante sobre el dataset. | onc-compraselectronicas@modernizacion.gob.ar | contactPoint -> hasEmail | String |
| Temática(s) globales | Si | Temática/s o categoría/s globales a la/s que se refiere el dataset al ser publicado. Un dataset puede pertenecer a más de una categoría global, de manera que el tipo de valor de este campo es una lista de categorías. La/s categoría/s o temática/s globales deben adoptarse según el campo "Código (authority code)" del <u>Anexo "Taxonomía temática para los datasets"</u> , que contiene una lista predefinida de temática/s globales. | ["ECON"] | superTheme | Array |
| Temática(s) específicas | Recomendado | Temática/s o categoría/s específica/s a la/s que se refiere el dataset al ser publicado. Un dataset puede pertenecer a más de una categoría específica, de manera que el tipo de valor de este campo es una lista de categorías. La taxonomía a utilizar debe ser creada por la autoridad responsable del Catálogo. | ["Contrataciones", "Compras", "Macroeconomía"] | theme | Array |
| Etiqueta(s) | Recomendado | Palabras que describen el título o el contenido del recurso. Es necesario que las etiquetas se encuentren correctamente escritas, en plural y respetando la existencia de tags anteriores. | ["bienes", "compras", "contrataciones"] | keyword | Array |

| | | | | | |
|---|-------------|--|--|---------------------------|--------|
| | | Etiquetas que colaboran en la búsqueda de los usuarios. Cuanto más amplia y uniforme sea la lista de tags mayor será su eficiencia. A tales fines se recomienda ver el Anexo "Pautas para la selección de etiquetas" . | | | |
| Distribuciones | Si | Lista de distribuciones que pertenecen al dataset y sus metadatos. Cada distribución se representa con un objeto ("{}") donde se describen los metadatos especificados para la clase "distribution" de este perfil de metadatos. | [{...}, {...}] | distribution | Array |
| Frecuencia de actualización | Si | Frecuencia con la que se actualiza el dataset. Se recomienda especificar períodos normalizados con formato ISO-8601, agregando el campo "eventual" para datasets que se publican con una frecuencia eventual o no especificada. Anexo "Especificación de frecuencias según ISO-8601" . | "R/P1Y" para datasets que se actualizan anualmente | accrualPeriodicity | String |
| Fecha de publicación | Si | Fecha de publicación del dataset. Según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset. | "2016-04-14T19:48:05.433640" para especificar fecha y hora "2016-04-14" para especificar fecha únicamente | issued | String |
| Fecha de última actualización/modificación | Recomendado | Fecha de última actualización/modificación del dataset (ya sea de sus datos o de sus metadatos). Según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset. | "2016-04-19T19:48:05.433640" para especificar fecha y hora "2016-04-19" para especificar fecha únicamente | modified | String |
| Identificador | No | Identificador único del dataset, este identificador debe ser único para todo el catálogo. | Un identificador único para el dataset. La URI u otro identificador único en el contexto del catálogo, ejemplo: "dataset-ejemplo-35782" | identifier | String |

| | | | | | |
|----------------------|----|---|--|----------|-----------------|
| Idioma(s) | No | <p>Lenguaje para la descripción de los metadatos del dataset. Hay 2 estándares ISO que pueden ser utilizados para este campo:</p> <p>(a) ISO 639-1 (2 letras)</p> <p>(b) ISO 639-2/T (3 letras) es el más recomendado.</p> <p>Puede definirse 1 o más lenguajes en una lista. Los lenguajes especificados para un dataset, sobreescriben a los del catálogo. Si este campo está vacío el dataset hereda los lenguajes del catálogo. (Link a los estándares ISO: https://www.loc.gov/standards/iso639-2/php/code_list.php)</p> | <p>["es"] para un lenguaje ISO 639-1</p> <p>["spa", "eng"] para dos lenguajes ISO 639-2</p> | language | Array |
| Cobertura geográfica | No | <p>Una región o lugar determinado al que el dataset haga referencia.</p> <p>Una región o un lugar determinado. Puede tomar valores:</p> <p>a) de países y, provincias y municipios argentinos, según las recomendaciones de la "Guía para la identificación y uso de entidades interoperables".</p> <p>b) un área de coordenadas representada por latitud/longitud en el orden: mínima longitud, mínima latitud, máxima longitud, máxima latitud.</p> <p>c) un punto geográfico representado por latitud/longitud.</p> <p>d) Si la referencia geográfica no está identificada en la "Guía para la identificación y uso de entidades</p> | <p>"ARG" es el código para la República Argentina.</p> <p>"06007" es el código de un departamento</p> <p>[-58.111111, -35.111111, -57.111111, -33.111111] es un bounding box</p> <p>[-58.111111, -35.111111] es un punto geográfico</p> <p>"http://sws.geonames.org/6255146"</p> | spatial | Array or String |

| | | | | | |
|------------------------------|-------------|--|---|--------------------|--------|
| | | interoperables" indicar la URIs según geonames.org; ej : http://sws.geonames.org/6255146" | | | |
| Cobertura temporal | Recomendado | Período de tiempo cubierto por el dataset. El intervalo de tiempo está formado por una fecha de inicio y una de fin separadas por "/", en formato ISO 8601, con el nivel de especificidad requerido por el dataset. | 2015-01-01/2015-12-31 2015-01-01T00:45:00Z/2016-01-15T00:06:00Z | temporal | String |
| Página de referencias | No | URL de una página web a través de la cual se puede acceder al dataset, sus recursos o información adicional sobre el mismo. | http://datos.gob.ar/dataset/sistema-de-contrataciones-electronicas-argentina-compra | landingPage | String |
| Licencia | Recomendado | Indica la licencia bajo la cual el dataset y todas sus distribuciones están disponibles mediante un enlace a la licencia o documento de la licencia seleccionada, o mediante el título textual de la licencia tal como aparece en la lista de http://opendefinition.org/licenses/ . Se recomienda utilizar la licencia "Open Database License (ODbL) v1.0". Un dataset hereda por default la licencia general del catálogo salvo que se especifique una licencia diferente en este campo. Las distribuciones del dataset heredan esta licencia salvo que especifiquen una diferente. | "http://opendatacommons.org/licenses/dbcl/1-0/" si se utiliza un enlace "Open Database License (ODbL) v1.0" si se consigna el nombre de la licencia a utilizar | license | String |

Debe prestarse especial atención al campo **distribution** que contiene una lista de objetos que describen los metadatos de cada una de las distribuciones del dataset. En la próxima sección se abordan estos metadatos.

Distribución

[Link](#) para ver un ejemplo de metadatos de una distribución en JSON.

A continuación se describen los metadatos que el *data.json* debe contener, para describir a una distribución dentro de la lista contenida en el campo **distribution** de un dataset:

| Nombre | Requerido | Descripción | Ejemplo | Variable (data.json) | Tipo (data.json) |
|---------------------|-------------|---|---|----------------------|------------------|
| URL de acceso | Si | URL que permite el acceso a la distribución del dataset. Puede ser una página, feed u otro tipo de recurso que dé acceso indirecto a las distribuciones. Si las distribuciones son solo accesibles a través de la página de referencia del dataset, debe completarse el valor de la URL de acceso a la distribución con el mismo valor de la página de referencia del dataset. | http://datos.gob.ar/dataset/sistema-de-contrataciones-electronicas-argentina-compra/archivo/fa3603b3-0af7-43cc-9da9-90a512217d8a | accessURL | String |
| Descripción | Recomendado | Breve descripción de la distribución. Se recomienda no escribir más de una línea. Toda información adicional puede ser incluida en la descripción del dataset. | Listado de las convocatorias abiertas durante el año 2015 en el sistema de contrataciones electrónicas | description | String |
| Formato del archivo | Recomendado | Indica el formato del archivo de la distribución. Si el tipo de la distribución está definido por IANA (http://www.iana.org/assignments/media-types/media-types.xml), debe usarse esa definición. En caso contrario deberán ponerse los caracteres después del punto final del archivo, que determinan el formato (cuando no está definido por IANA). | "text/csv" definición de IANA "csv" caracteres finales después del punto | format | String |
| Tipo de archivo | No | Indica el tipo de archivo de la distribución, sólo si este está definido por IANA (http://www.iana.org/assignments/media-types/media-types.xml). En caso contrario este campo permanece vacío. | "text/csv" definición de IANA "" cuando el formato no tiene definición en IANA | mediaType | String |
| URL de descarga | Si | URL que permite la descarga directa de la distribución del dataset, vincula directamente a un archivo descargable en un formato dado. | http://datos.gob.ar/dataset/becaceb2-dbd0-4879-93bd-5f02bd3b8ca2/resource/bf2f67f4-9ab3-479b-a881-56b43565125e/download/contratos-2015.csv | downloadURL | String |
| Título | Si | Nombre asignado a la distribución. | Convocatorias abiertas durante el año 2015 | title | String |

| | | | | | |
|---|-------------|--|---|-----------------|--------|
| Licencia | Recomendado | Indica la licencia bajo la cual la distribución está disponible mediante un enlace a la licencia o documento de la licencia seleccionada, o mediante el título textual de la licencia tal como aparece en la lista de http://opendefinition.org/licenses/ . Se recomienda utilizar la licencia "Open Database License (ODbL) v1.0". Una distribución hereda por default la licencia del dataset al que pertenece, salvo que se especifique una licencia diferente en este campo. | " http://opendatacommons.org/licenses/dbcl/1-0/ " si se utiliza un enlace "Open Database License (ODbL) v1.0" si se consigna el nombre de la licencia a utilizar | license | String |
| Tamaño | No | Tamaño de la distribución en bytes. El tamaño puede ser aproximado cuando no se conozca el tamaño exacto. | Ejemplo para un archivo de 5Kb aproximadamente: "5120" | byteSize | String |
| Fecha de publicación | Si | Fecha de publicación de la distribución. Según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset. | "2016-04-14T19:48:05.433640" para especificar fecha y hora "2016-04-14" para especificar únicamente fecha | issued | String |
| Fecha de última actualización/modificación | Recomendado | Fecha de última actualización/modificación de la distribución. Según el formato ISO-8601, tipeado como fecha simple o fecha con hora, con el mayor detalle posible que sea relevante para el dataset. | "2016-04-19T19:48:05.433640" para especificar fecha y hora "2016-04-19" para especificar únicamente fecha | modified | String |
| Derechos sobre la distribución | No | Información sobre derechos aplicables a la distribución que no se hayan especificado en la licencia. Si se especifican, estos derechos sobreesciben a los del catálogo. En caso contrario, las distribuciones heredan los derechos especificados para el catálogo. | | rights | String |
| Campos de la distribución | Recomendado | Lista de campos que contiene una distribución tabular (no aplica para aquellas distribuciones que no sean tablas) y sus metadatos. Cada campo se representa con un objeto ("{}") donde se describen los | [{...}, {...}] | field | Array |

| | | |
|--|--|--|
| | metadatos especificados para la clase "field" de este perfil de metadatos (como "nombre", "tipo" y "descripción"). | |
|--|--|--|

Debe prestarse especial atención al campo **field** que contiene una lista de objetos que describen los metadatos de cada uno de los campos de la distribución (en el caso de distribuciones tabulares, únicamente). En la próxima sección se abordan estos metadatos.

Campo

[Link](#) para ver un ejemplo de metadatos de un campo en JSON.

A continuación se describen los metadatos que el *data.json* debe contener, para describir a un campo de una distribución tabular dentro de la lista contenida en el campo de metadatos **field** de una distribución:

| Nombre | Requerido | Descripción | Ejemplo | Variable (data.json) | Tipo (data.json) |
|--------|-------------|---|--|----------------------|------------------|
| Nombre | Recomendado | El nombre del campo tal como se denomina en el encabezado de la distribución. Véase la "Guía para la publicación de datos en formatos abiertos" para una adecuada elección del nombre de un campo. | Ejemplo para el cuarto campo de la distribución "Convocatorias abiertas durante el año 2015", valor para el nombre: "unidad_operativa_contrataciones_desc" | title | String |
| Tipo | Recomendado | <p>El tipo de dato contenido en el campo según la lista utilizada por la librería recline.js (http://okfnlabs.org/recline/docs/models.html#types).</p> <p>Los tipos incluidos en esta lista son:</p> <p>string (text): Valores de texto.</p> <p>number (double, float, numeric): Números que puedan no ser enteros (incluyen decimales).</p> <p>integer (int): Números que siempre son enteros.</p> <p>date: Fecha simple expresada según el estándar ISO 8601 incluyendo únicamente</p> | Ejemplo para el campo "unidad_operativa_contrataciones_desc" de la distribución "Convocatorias abiertas durante el año 2015", valor para tipo: "string" | type | String |

| | | | | |
|-------------|-------------|---|--|----------------------------------|
| | | <p>año, mes y día (YYYY-MM-DD) como en "2016-02-01".</p> <p>time: Tiempo expresado según el estándar ISO 8601 incluyendo únicamente horas, minutos y segundos (hh:mm:ss) como en "10:05:00".</p> <p>date-time (datetime, timestamp): Fecha completa expresada según el estándar ISO 8601 incluyendo año, mes, día, horas, minutos y segundos (YYYY-MM-DDThh:mm:ssZ) como en "2016-02-01T10:05:00+03:00"</p> <p>boolean (bool): Valores verdadero/falso.</p> <p>binary: Representación de datos binarios base64.</p> <p>geo_point: Ver estructura en http://www.elasticsearch.org/guide/reference/mapping/geo-point-type.html.</p> <p>geojson: Ver en http://geojson.org/</p> <p>array: Lista de valores.</p> <p>object (json): Objeto de JSON.</p> <p>any: Campo que puede contener valores de cualquier tipo.</p> | | |
| Descripción | Recomendado | <p>La descripción de la información que contiene el campo.</p> | <p>Ejemplo para el campo "unidad_operativa_contrataciones_desc" de la distribución "Convocatorias abiertas durante el año 2015", valor para descripción: "Organismo que realiza la convocatoria. Organismo de máximo nivel jerárquico al que pertenece la unidad operativa de contrataciones."</p> | <p>description</p> <p>String</p> |

Tema

[Link](#) para ver un ejemplo de metadatos de un campo en JSON.

A continuación se describen los metadatos que el *data.json* debe contener, para describir a un tema de la taxonomía temática de un Catálogo:

| Nombre | Requerido | Descripción | Ejemplo | Variable (data.json) | Tipo (data.json) |
|---------------|-------------|---|---|----------------------|------------------|
| Identificador | Recomendado | El identificador del tema. | AGRI | id | String |
| Etiqueta | Recomendado | La etiqueta o título de un tema. | Agricultura, pesca, silvicultura y alimentación | label | String |
| Descripción | Recomendado | Una breve y concisa descripción del tema. | Bajo este concepto se incluyen datasets que cubren dominios tales como agricultura, pesca, forestación o alimentos. | description | String |

Anexo I - Taxonomía temática global de la APN para los datasets (tabla)

El Portal Nacional de Datos utiliza la taxonomía temática definida por la Unión Europea, tal como se indica en el siguiente [enlace](#).

Cada Catálogo de Datos puede desarrollar su propia taxonomía, cuyo uso se expresa en los siguientes campos de metadatos (y sus equivalentes para el Portal Andino):

- **themeTaxonomy**: Es un campo de metadatos del Catálogo que define una lista de temas que se pueden utilizar para clasificar los datasets. Refiere al esquema completo de la taxonomía en sí, no a alguna de sus etiquetas en particular.
- **theme**: Es un campo de metadatos de un Dataset. Refiere a la/s etiqueta/s en particular bajos la/s cuales un Dataset es clasificado temáticamente. Sólo pueden usarse etiquetas que estén definidas en la taxonomía temática de *themeTaxonomy*.

Además del uso de una taxonomía propia de cada Catálogo de Datos, esta guía recomienda la clasificación de los Datasets según la taxonomía del Portal Nacional de Datos. Esta es una *super taxonomía* a la que cada Catálogo hace referencia mediante los siguientes campos de metadatos (y sus equivalentes para el Portal Andino):

- **superThemeTaxonomy**: Es un campo de metadatos del Catálogo que apunta a la URL donde el Portal Nacional de Datos documenta la taxonomía temática de la Administración Pública Nacional.
- **superTheme**: Es un campo de metadatos de un Dataset. Refiere a la/s etiqueta/s en particular bajos la/s cuales un Dataset es clasificado temáticamente, según la *super taxonomía* que es la de la Administración Pública Nacional. Sólo pueden usarse etiquetas que estén definidas en la taxonomía temática de *superThemeTaxonomy*.

La ventaja de utilizar una super taxonomía temática es que facilita la clasificación de datasets por parte de un usuario según un conjunto de categorías temáticas más generales, que son interoperables con las utilizadas por otros países del mundo. Esto, a su vez, facilita la clasificación de datasets cosechados por el Portal Nacional de Datos.

| Código (authority code) | Estatus (Status) | Etiqueta (label) | Descripción (description) |
|-------------------------|------------------|---|--|
| AGRI | activo | Agricultura, pesca, silvicultura y alimentación | Bajo este concepto se incluyen <i>datasets</i> que cubren dominios tales como agricultura, pesca, forestación o alimentos. |
| ECON | activo | Economía y finanzas | Bajo este concepto se incluyen <i>datasets</i> que cubren dominios tales como economía o finanzas. |

| | | | |
|------|--------|--|---|
| EDUC | activo | Educación, cultura y deportes | Bajo este concepto se incluyen <i>datasets</i> que cubren dominios tales como educación, cultura o deporte. |
| ENER | activo | Energía | Bajo este concepto se incluyen <i>datasets</i> que cubre aspectos de energía. |
| ENVI | activo | Medio ambiente | Bajo este concepto se incluyen <i>datasets</i> que cubren aspectos de medioambiente. |
| GOVE | activo | Gobierno y sector público | Bajo este concepto se incluyen <i>datasets</i> que cubren dominios tales como gobierno y sector público. |
| HEAL | activo | Salud | Bajo este concepto se incluyen <i>datasets</i> que cubren aspectos relacionados a la salud. |
| INTR | activo | Asuntos internacionales | Bajo este concepto se incluyen <i>datasets</i> que cubren aspectos relacionados a asuntos internacionales. |
| JUST | activo | Justicia, sistema judicial y seguridad pública | Bajo este concepto se incluyen <i>datasets</i> que cubren dominios tales justicia, sistema legal o seguridad pública. |
| REGI | activo | Regiones y ciudades | Bajo este concepto se incluyen <i>datasets</i> que cubren dominios tales como regiones o ciudades. |
| SOCI | activo | Población y sociedad | This concept identifies datasets covering such domains as population or society. |
| TECH | activo | Ciencia y tecnología | Bajo este concepto se incluyen <i>datasets</i> que cubren dominios tales como ciencia y tecnología. |
| TRAN | activo | Transporte | Bajo este concepto se incluyen <i>datasets</i> que cubren aspectos relacionados al transporte. |

Anexo II - Pautas para la selección de etiquetas

En la definición de etiquetas para los datasets debemos contemplar que un etiquetado eficiente facilita la correcta adecuación de la información a los criterios de búsqueda. Cuanto más amplia y uniforme sea la lista de etiquetas mayor será su efectividad.

Las siguientes pautas para definir etiquetas son aplicables a la propiedad *keyword* de la clase Dataset.

- Escribir correctamente y en plural.
- Utiliza mayúsculas sólo donde corresponda.
- Identifica palabras claves.
- Respeta la existencia de etiquetas anteriores.
- Agrega sinónimos y emplea lenguaje natural.
- Etiqueta preferentemente compuesta por una palabra, aunque más de una es aceptable de ser necesario.
 - Si la etiqueta posee más de una palabra debe estar separada por un espacio, ej: “declaraciones juradas”.

Preguntas útiles a la hora de pensar los etiquetas:

- ¿Cuál es el tema?
- ¿Qué aspectos serán de interés para los usuarios?
- ¿De qué otro modo buscaría sobre esta información?
- ¿De qué tipo de información se trata?
- ¿Qué área la provee?

Anexo III - Especificación de frecuencias (según ISO-8601)

| Frecuencia | Valor según ISO-8601 |
|---------------------------|----------------------|
| Cada diez años | R/P10Y |
| Cada cuatro años | R/P4Y |
| Cada tres años | R/P3Y |
| Cada dos años | R/P2Y |
| Anualmente | R/P1Y |
| Cada medio año | R/P6M |
| Cuatrimensualmente | R/P4M |
| Trimestralmente | R/P3M |
| Bimestralmente | R/P2M o R/P0.5M |
| Mensualmente | R/P1M |
| Cada 15 días | R/P0.5M |
| Tres veces por mes | R/P0.33M |
| Semanalmente | R/P1W |
| Cada media semana | R/P3.5D |
| Dos veces a la semana | R/P2W or R/P0.5W |
| Tres veces a la semana | R/P0.33W |
| Diariamente | R/P1D |
| Cada hora | R/PT1H |
| Continuamente actualizado | R/PT1S |

Anexo IV - Ejemplo de data.json

[Link](#) al data.json de ejemplo.

```
{
  "title": "Datos Argentina",
  "description": "Portal de Datos Abiertos del Gobierno de la República Argentina",
  "publisher": {
    "name": "Ministerio de Modernización",
    "mbox": "datos@modernizacion.gob.ar"
  },
  "issued": "2016-04-14T19:48:05.433640-03:00",
  "modified": "2016-04-19T19:48:05.433640-03:00",
  "language": [
    "spa"
  ],
  "superThemeTaxonomy": "http://datos.gob.ar/superThemeTaxonomy.json",
  "themeTaxonomy": [
    {
      "id": "convocatorias",
      "label": "Convocatorias",
      "description": "Datasets sobre licitaciones en estado de convocatoria."
    },
    {
      "id": "compras",
      "label": "Compras",
      "description": "Datasets sobre compras realizadas."
    },
    {
      "id": "contrataciones",
      "label": "Contrataciones",
      "description": "Datasets sobre contrataciones."
    },
    {
      "id": "adjudicaciones",
      "label": "Adjudicaciones",
      "description": "Datasets sobre licitaciones adjudicadas."
    },
    {
      "id": "normativa",
      "label": "Normativa",
      "description": "Datasets sobre normativa para compras y contrataciones."
    },
    {
      "id": "proveedores",
      "label": "Proveedores",
      "description": "Datasets sobre proveedores del Estado."
    }
  ],
  "license": "Open Data Commons Open Database License 1.0",
  "homepage": "http://datos.gob.ar",
  "rights": "Derechos especificados en la licencia.",
  "spatial": "ARG",
  "dataset": [
    {
```

```

"title": "Sistema de contrataciones electrónicas",
"description": "Datos correspondientes al Sistema de Contrataciones Electrónicas (Argentina Compra)",
"publisher": {
  "name": "Ministerio de Modernización. Secretaría de Modernización Administrativa. Oficina Nacional de Contrataciones",
  "mbox": "onc@modernizacion.gob.ar"
},
"contactPoint": {
  "fn": "Ministerio de Modernización. Secretaría de Modernización Administrativa. Oficina Nacional de Contrataciones. Dirección de Compras Electrónicas.",
  "hasEmail": "onc-compraselectronicas@modernizacion.gob.ar"
},
"superTheme": [
  "ECON"
],
"theme": [
  "contrataciones",
  "compras",
  "convocatorias"
],
"keyword": [
  "bienes",
  "compras",
  "contrataciones"
],
"accrualPeriodicity": "R/P1Y",
"issued": "2016-04-14T19:48:05.433640-03:00",
"modified": "2016-04-19T19:48:05.433640-03:00",
"identifier": "99db6631-d1c9-470b-a73e-c62daa32c420",
"language": [
  "spa"
],
"spatial": "ARG",
"temporal": "2015-01-01/2015-12-31",
"landingPage": "http://datos.gob.ar/dataset/sistema-de-contrataciones-electronicas-argentina-compra",
"license": "Open Data Commons Open Database License 1.0",
"distribution": [
  {
    "accessURL":
      "http://datos.gob.ar/dataset/sistema-de-contrataciones-electronicas-argentina-compra/archivo/fa3603b3-0af7-43cc-9da9-90a512217d8a",
    "description": "Listado de las convocatorias abiertas durante el año 2015 en el sistema de contrataciones electrónicas",
    "format": "CSV",
    "mediaType": "text/csv",
    "downloadURL":
      "http://186.33.211.253/dataset/99db6631-d1c9-470b-a73e-c62daa32c420/resource/4b7447cb-31ff-4352-96c3-589d212e1cc9/download/convocatorias-abiertas-anio-2015.csv",
    "title": "Convocatorias abiertas durante el año 2015",
    "license": "Open Data Commons Open Database License 1.0",
    "byteSize": "5120",
    "issued": "2016-04-14T19:48:05.433640-03:00",
    "modified": "2016-04-19T19:48:05.433640-03:00",
    "rights": "Derechos especificados en la licencia.",
    "field": [
      {
        "title": "procedimiento_id",
        "type": "integer",
        "description": "Identificador único del procedimiento de contratación"
      }
    ]
  }
]

```

```

        "title": "organismo_unidad_operativa_contrataciones_id",
        "type": "integer",
        "description": "Identificador único del organismo que realiza la convocatoria. Organismo de máximo nivel jerárquico al que pertenece la unidad operativa de contrataciones."
    },
    {
        "title": "unidad_operativa_contrataciones_id",
        "type": "integer",
        "description": "Identificador único de la unidad operativa de contrataciones"
    },
    {
        "title": "organismo_unidad_operativa_contrataciones_desc",
        "type": "string",
        "description": "Organismo que realiza la convocatoria. Organismo de máximo nivel jerárquico al que pertenece la unidad operativa de contrataciones."
    },
    {
        "title": "unidad_operativa_contrataciones_desc",
        "type": "string",
        "description": "Unidad operativa de contrataciones."
    },
    {
        "title": "tipo_procedimiento_contratacion",
        "type": "string",
        "description": "Tipo de procedimiento al que se adecua la contratación."
    },
    {
        "title": "ejercicio_procedimiento_anio",
        "type": "date",
        "description": "Año en el que se inició el proceso de la convocatoria."
    },
    {
        "title": "fecha_publicacion_convocatoria",
        "type": "date",
        "description": "Fecha de publicación de la convocatoria en formato AAAA-MM-DD, ISO 8601."
    },
    {
        "title": "modalidad_convocatoria",
        "type": "string",
        "description": "Modalidad bajo la cual se realiza la convocatoria."
    },
    {
        "title": "clase_convocatoria",
        "type": "string",
        "description": "Clase de la convocatoria."
    },
    {
        "title": "objeto_convocatoria",
        "type": "string",
        "description": "Objeto/objetivo de la convocatoria"
    }
}
]
}
]
}
}

```

Anexo V - Taxonomía temática global de la APN para los datasets (JSON)

[Link](#) a la taxonomía temática global.

```
[
  {
    "id": "AGRI",
    "label": "Agricultura, pesca, silvicultura y alimentación",
    "description": "Bajo este concepto se incluyen datasets que cubren dominios tales como agricultura, pesca, forestación o alimentos."
  },
  {
    "id": "ECON",
    "label": "Economía y finanzas",
    "description": "Bajo este concepto se incluyen datasets que cubren dominios tales como economía o finanzas."
  },
  {
    "id": "EDUC",
    "label": "Educación, cultura y deportes",
    "description": "Bajo este concepto se incluyen datasets que cubren dominios tales como educación, cultura o deporte."
  },
  {
    "id": "ENER",
    "label": "Energía",
    "description": "Bajo este concepto se incluyen datasets que cubre aspectos de energía."
  },
  {
    "id": "ENVI",
    "label": "Medio ambiente",
    "description": "Bajo este concepto se incluyen datasets que cubren aspectos de medioambiente."
  },
  {
    "id": "GOVE",
    "label": "Gobierno y sector público",
    "description": "Bajo este concepto se incluyen datasets que cubren dominios tales como gobierno y sector público."
  },
  {
    "id": "HEAL",
    "label": "Salud",
    "description": "Bajo este concepto se incluyen datasets que cubren aspectos relacionados a la salud."
  },
  {
    "id": "INTR",
    "label": "Asuntos internacionales",
    "description": "Bajo este concepto se incluyen datasets que cubren aspectos relacionados a asuntos internacionales."
  },
  {
    "id": "JUST",
```

```

    "label": "Justicia, sistema judicial y seguridad pública",
    "description": "Bajo este concepto se incluyen datasets que cubren dominios tales justicia, sistema legal o seguridad pública."
  },
  {
    "id": "REGI",
    "label": "Regiones y ciudades",
    "description": "Bajo este concepto se incluyen datasets que cubren dominios tales como regiones o ciudades."
  },
  {
    "id": "SOCI",
    "label": "Población y sociedad",
    "description": "Bajo este concepto se incluyen datasets que cubren dominios tales como población o sociedad"
  },
  {
    "id": "TECH",
    "label": "Ciencia y tecnología",
    "description": "Bajo este concepto se incluyen datasets que cubren dominios tales como ciencia y tecnología."
  },
  {
    "id": "TRAN",
    "label": "Transporte",
    "description": "Bajo este concepto se incluyen datasets que cubren aspectos relacionados al transporte."
  }
]

```

Anexo VI - Ejemplo de metadatos como texto

[Link](#) al ejemplo en markdown.

Catálogo: Datos Argentina

Portal de Datos Abiertos del Gobierno de la República Argentina

- Derechos sobre el catálogo: Derechos especificados en la licencia.
- Correo electrónico del autor: datos@modernizacion.gob.ar
- Autor: Ministerio de Modernización
- Licencia: Open Data Commons Open Database License 1.0
- Idioma(s): spa
- Fecha de creación o publicación: 2016-04-14T19:48:05.433640-03:00
- Taxonomía temática global: <http://datos.gob.ar/superThemeTaxonomy.json>
- Fecha de última actualización/modificación: 2016-04-19T19:48:05.433640-03:00
- Cobertura geográfica: ARG
- Página web del catálogo: <http://datos.gob.ar>

Taxonomía temática específica

- Convocatorias (convocatorias): Datasets sobre licitaciones en estado de convocatoria.
- Compras (compras): Datasets sobre compras realizadas.
- Contrataciones (contrataciones): Datasets sobre contrataciones.
- Adjudicaciones (adjudicaciones): Datasets sobre licitaciones adjudicadas.
- Normativa (normativa): Datasets sobre normativa para compras y contrataciones.
- Proveedores (proveedores): Datasets sobre proveedores del Estado.

Datasets

Dataset: Sistema de contrataciones electrónicas

Datos correspondientes al Sistema de Contrataciones Electrónicas (Argentina Compra)

- **Correo electrónico del autor:** onc@modernizacion.gob.ar
- **Autor:** Ministerio de Modernización. Secretaría de Modernización Administrativa. Oficina Nacional de Contrataciones
- **Página de referencias:**
<http://datos.gob.ar/dataset/sistema-de-contrataciones-electronicas-argentina-compra>
- **Temática(s) globales:** ECON
- **Fecha de publicación:** 2016-04-14T19:48:05.433640-03:00

- **Cobertura temporal:** 2015-01-01/2015-12-31
- **Fecha de última actualización/ modificación:** 2016-04-19T19:48:05.433640-03:00
- **Idioma(s):** spa
- **Temática(s) específicas:** contrataciones, compras, convocatorias
- **Etiqueta(s):** bienes, compras, contrataciones
- **Frecuencia de actualización:** R/P1Y
- **Cobertura geográfica:** ARG
- **Licencia:** Open Data Commons Open Database License 1.0
- **Correo electrónico del área/persona de contacto:**
onc-compraselectronicas@modernizacion.gob.ar
- **Área/Persona de contacto:** Ministerio de Modernización. Secretaría de Modernización Administrativa. Oficina Nacional de Contrataciones. Dirección de Compras Electrónicas.
- **Identificador:** 99db6631-d1c9-470b-a73e-c62daa32c420

Distribuciones

Distribución: Convocatorias abiertas durante el año 2015

Listado de las convocatorias abiertas durante el año 2015 en el sistema de contrataciones electrónicas

- **URL de acceso:**
<http://datos.gob.ar/dataset/sistema-de-contrataciones-electronicas-argentina-compra/archivo/fa3603b3-0af7-43cc-9da9-90a512217d8a>
- **Derechos sobre la distribución:** Derechos especificados en la licencia.
- **Licencia:** Open Data Commons Open Database License 1.0
- **Tamaño:** 5120
- **Formato del archivo:** CSV
- **Tipo de archivo:** text/csv
- **Fecha de última actualización/ modificación:** 2016-04-19T19:48:05.433640-03:00
- **URL de descarga:**
<http://186.33.211.253/dataset/99db6631-d1c9-470b-a73e-c62daa32c420/resource/4b7447cb-31ff-4352-96c3-589d212e1cc9/download/convocatorias-abiertas-anio-2015.csv>
- **Fecha de publicación:** 2016-04-14T19:48:05.433640-03:00

Campos de la distribución

- **procedimiento_id** (integer): Identificador único del procedimiento de contratación
- **organismo_unidad_operativa_contrataciones_id** (integer): Identificador único del organismo que realiza la convocatoria. Organismo de máximo nivel jerárquico al que pertenece la unidad operativa de contrataciones.
- **unidad_operativa_contrataciones_id** (integer): Identificador único de la unidad operativa de contrataciones
- **organismo_unidad_operativa_contrataciones_desc** (string): Organismo que realiza la convocatoria. Organismo de máximo nivel jerárquico al que pertenece la unidad operativa de contrataciones.
- **unidad_operativa_contrataciones_desc** (string): Unidad operativa de contrataciones.
- **tipo_procedimiento_contratacion** (string): Tipo de procedimiento al que se adecua la contratación.

- **ejercicio_procedimiento_anio** (date): Año en el que se inició el proceso de la convocatoria.
- **fecha_publicacion_convocatoria** (date): Fecha de publicación de la convocatoria en formato AAAA-MM-DD, ISO 8601.
- **modalidad_convocatoria** (string): Modalidad bajo la cual se realiza la convocatoria.
- **clase_convocatoria** (string): Clase de la convocatoria.
- **objeto_convocatoria** (string): Objeto/objetivo de la convocatoria