

Introducción

El catálogo de datos permite identificar los datos abiertos de una organización, describirlos, y hacerlos públicos en DATOS.GOB.MX.

Objetivos del catálogo

- **Descriptivo:** Capturar los metadatos mínimos requeridos de un conjunto de datos
 - Para mostrar descriptivamente una ficha bibliográfica en el catálogo de datos en <http://catalogo.datos.gob.mx> [ejemplo]
- **Extensibilidad:** permitir agregar campos adicionales para proveer información adicional al usuario.
- **Heterogeneidad:** ser agnóstico de formato (p.ej. XML, RDF, CSV).
- **Recursos:** permitir la descripción de archivos contenedores de los datos.
- **Estándar:** ser compatible con el estándar DCAT y DCT de metadatos para *datasets* de gobierno.

Glosario

DCAT: Es un vocabulario diseñado para facilitar la interoperabilidad entre catálogos de datos en la Web, utilizado por data.gov y data.gov.uk (entre otros).

Catálogo de datos: Un catálogo contiene conjuntos de datos (*datasets*) y recursos/descargables (*distributions*).

Conjunto de datos (*dataset*): Colección de datos, publicados o curados por un agente, y disponibles para acceso o descarga en uno o más formatos.

Escenarios

Hemos identificado una serie de escenarios en diferentes dependencias, sobre la manera en la que se publicarían y presentarían los *datasets* en DATOS.GOB.MX:

1. Los recursos (*distributions*)
 - a. tienen diferentes **formatos** [enlace]
 - b. son **series** de tiempo [enlace]
 - c. son **segmentos** filtrados [enlace]
 - d. requieren **diccionario** de datos único
2. Los **resultados de búsqueda se podrían llenar de ruido** enumerando mismos *datasets* con diferentes versiones del conjunto de datos, p. ej. fonden 2011, fonden 2012, fonden 2013; en lugar de simplemente “fonden”, y tener como descargables 2011.csv, 2012.csv y 2013.csv.

3. Para **evitar llenar formularios**, como lo hizo data.gov¹, optamos por usar una tradicional hoja de cálculo [3] que automáticamente se transformará en DCAT.

Propuesta

Llenar la hoja de cálculo que sirve de plantilla para el catálogo de datos [enlace] llenando los metadatos acorde a lo que se presenta a continuación:

Nota: con **negrita** los campos obligatorios, en *cursiva* los condicionales, y sin estilo los opcionales.

Metadatos (descriptores) del conjunto de datos (*dataset*)

ds:identifier	Identificador único del conjunto de datos. Utilizado para agrupar recursos dentro de un <i>dataset</i> , p. ej. “rezago-social”, “unidades_medicas”, “adquisiciones”. Utilizar caracteres ASCII (sin acentos).
ds:title	Título descriptivo del conjunto de datos, p.ej. “Programa de fomento a la agricultura”, “Vuelos comerciales”.
ds:description	Una explicación de los datos, con suficiente detalle para que los usuarios puedan entender si es de su interés, p.ej. “Apoyos otorgados a través del programa Opciones Productivas, desglosado a nivel localidad”.
ds:keyword	Lista de términos clave separados por coma, que facilitarán al usuario la búsqueda del conjunto de datos. Es importante considerar el uso de términos no técnicos. p.ej. “salud, medicinas, compras, agricultura”.
ds:modified	Fecha y hora de la última modificación del conjunto de datos; en formato ISO 8601 ² , p.ej. “2014-05-27T01:42:05-05:00”
ds:contactPoint	Nombre de la persona de contacto que atenderá dudas y comentarios sobre el conjunto de datos, p.ej. “Rodolfo Wilhelmy”.
ds:mbox	Correo electrónico de contacto para atender dudas y comentarios sobre el conjunto de datos, p.ej. “ rodolfo.wilhelmy@presidencia.gob.mx ”, “ datos@presidencia.gob.mx ”.

¹ <http://project-open-data.github.io/catalog-generator/>

² http://en.wikipedia.org/wiki/ISO_8601

ds:temporal	La fecha o fechas que cubren los datos, p.ej. “2013”, “2010/2012”, “2014-01/2014-04”. Si es un rango de fechas, deberán ordenarse ascendentemente.
ds:spatial	El espacio geográfico que cubre el conjunto de datos. Puede ser una región, el nombre de un lugar, una clave INEGI, un polígono o un cuadro delimitador de coordenadas geográficas (bounding box) en GML ³ . P.ej. “Baja California”, 002, http://www.geonames.org/4017700/baja-california.html , “estatal”, o “32.71,-112.32 27.99, -118.45”.
ds:dataDictionary	Dirección electrónica para descargar el diccionario de datos. Este documento sirve como guía adicional para que el usuario entienda con mayor detalle los datos.
ds:accrualPeriodicity	Frecuencia con la cual el conjunto de datos será publicado o actualizado, p. ej. “mensualmente”.

Metadatos (descriptores) del recurso o descargable (*distribution*)

ds:identifier	La clave que identifica al conjunto de datos al que pertenece este recurso, p. ej. “rezago-social”, “unidades_médicas”, “adquisiciones”.
rs:title	Título descriptivo del recurso o descargable, p.ej. “Otorgamientos del 2013”, “Otorgamientos del 2014”, “Apoyos por municipio”, “Apoyos por localidad”.
rs:description	Ver ds:description. Esta explicación es adicional a la que existe en el conjunto de datos. En caso de que se requiera un diccionario de datos adicional (i.e. por recurso) se puede agregar la referencia en este espacio, p. ej. “Índice de rezago social por localidad, utilizando el catálogo de INEGI 2010. El diccionario de datos lo puedes descargar de http://web.coneval.gob.mx/Informes/Pobreza/Datos_abiertos/Pobreza%20por%20entidad%202010_DIC.txt ”.
rs:downloadURL	Dirección electrónica (enlace) para la descarga del recurso.
rs:mediaType	Formato de archivo del recurso a descargar, p. ej. “text/csv”, “application/rss+xml”. Ver [4].
rs:byteSize	El tamaño en bytes del recurso o descargable, p. ej. 3145728
rs:temporal	Ver ds:temporal

³ http://en.wikipedia.org/wiki/Geography_Markup_Language

rs:spatial	Ver ds:spatial
------------	----------------

Notas

- En caso de necesitar diccionario de datos por recursos, agregar la liga en su campo “rs:description”.
- El campo “dct:publisher” será agregado automáticamente por la herramienta ADELA, y corresponde al nombre corto de la dependencia.
- En lugar de llenar el campo “dct:license”, se asume que todo lo publicado bajo datos.gob.mx está bajo los términos de libre uso del sitio en datos.gob.mx/tos.

Trabajo futuro

- Agregar catálogo para campos [dct:temporal](#), [dct:spatial](#), [dct:accrualPeriodicity](#)

Referencias

- [1] “Common Core Metadata Schema”. Project Open Data.
- [2] “Data Catalog Vocabulary (DCAT)”. W3C Recommendation 16 January 2014.
- [3] “Catálogo de Datos v2.0”. DATOS.GOB.MX.
- [4] “Media Types”. IANA.

Relación entre metadatos y datos.gob.mx

En el listado de resultados de búsqueda se presentan los campos de título, descripción y formatos.

The screenshot displays the DATOS.GOB.MX website interface. At the top, there is a navigation bar with the logo, 'DATOS.GOB.MX', 'ABRIR PARA TRANSPARENCIA', and links for 'Datos', 'Historias', 'Herramientas', 'Avances', and 'Acerca'. Below this, a breadcrumb trail shows 'Inicio / conjuntos de datos'. A sidebar on the left includes a map titled 'Filtrar por localidad' with a 'Borrar' button, and lists for 'Instituciones' (Presidencia (17), SEDESOL (5), CONEVAL (5)) and 'Etiquetas'. The main content area features a search bar with the text 'programas sociales' and a magnifying glass icon. Below the search bar, it states '32 conjuntos de datos encontrados' and 'programas sociales'. Two green arrows point to the search results: one labeled 'ds:title' pointing to the title 'PROGRAMA DE COINVERSIÓN SOCIAL', and another labeled 'ds:description' pointing to the description 'Datos del ejercicio fiscal 2013 del Programa de Coinversión Social, que fortalece a los actores sociales a través de la promoción de proyectos de coinversión entre el gobierno y...'. A 'ZIP' button is visible below the description.

DATOS.GOB.MX

ABRIR PARA AHORRAR

[Datos](#)
[Historias](#)
[Herramientas](#)
[Avances](#)
[Acerca](#)

/ instituciones / coneval / indicadores de pobreza, ...

Conjuntos de datos Instituciones Grupos

Indicadores de pobreza, pobreza por ingresos, rezago social y gini 2010 (municipal)

Institución

CONEVAL
El Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) es un organismo autónomo con la capacidad técnica para generar información objetiva sobre la... leer más

Social

- Google+
- Twitter
- Facebook

Licencia

LIBRE USO MX

Administrar

Conjunto de datos Grupos Flujo de Actividad Relacionados

Indicadores de pobreza, pobreza por ingresos, rezago social y gini 2010 (municipal)

Este conjunto de datos contiene información de pobreza que de acuerdo con la Ley General de Desarrollo Social (LGDS), el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) debe medir la pobreza a nivel municipio cada cinco años y para ello debe utilizar información que genera el Instituto Nacional de Estadística y Geografía (INEGI). Además de esta información, se cuenta con información sobre pobreza por ingresos, rezago social y coeficientes de gini.

DATOS Y RECURSOS

Municipal 2010
Indicadores de pobreza, pobreza por ingresos, gini y rezago social...
 [Explorar](#)

Pobreza gini y rezago social pobreza por ingresos

INFORMACION ADICIONAL	
Campo	Valor
Estado	active
dcat_modified	2013-06-30T17:13:00
dcat_publisher_email	eeminor@coneval.gob.mx
dcat_publisher_name	CONEVAL
guid	Indicadores de pobreza, pobreza por ingresos, gini y rezago social
harvest_object_id	30e05e03-75e5-4317-a44f-a05dbb2327a5
harvest_source_id	e0cd65ee-daa8-4ccc-ad9d-2a6981cce5d1
harvest_source_title	CONEVAL
language	es