

– Opendata Euskadi –

Portal de Datos Abiertos del Gobierno Vasco

Denominación:	Ficha de DataSet
Documentación Técnica	
Autor	Departamento de Justicia y Administración Pública del Gobierno Vasco Dirección de Atención Ciudadana

Contenido

Capítulo/sección	Página
1 Introducción	3
2 Familia y Tipos OpenData	4
3 Ficha de DataSet	5
3.1 Visión General	5
3.2 Representación en XML del DataSet	7
3.3 MetaDatos	11

1 Introducción

El presente documento contiene una descripción de la ficha utilizada para describir los conjuntos de datos reutilizables (DataSets) de opendata.euskadi.net

El conocimiento de los meta-datos propios de esta ficha es importante de cara a conocer las características de cada uno de los conjuntos de datos.

Este documento es complementario al documento r01OpenDataEuskadi_Tech_MetaDatos_EstructuraDirectorios.doc donde se recoge:

- Una descripción general del modelo de contenido
- Los meta-datos comunes a todos los contenidos
- La estructura de los directorios de los contenidos

Audiencia: Analistas y Desarrolladores

Complejidad: Media / Alta

2 Familia y Tipos OpenData

Los contenidos en Euskadi.net se tipifican en tres niveles:

Cluster	División general de tipología de contenidos
Familia	<p>Una agrupación de tipos de contenidos con una naturaleza común:</p> <p>Ejemplos:</p> <ul style="list-style-type: none"> • Familia procedimientos Administrativos que contiene los tipos de contenido relacionados con el procedimiento administrativo (Ayudas, Contrataciones, Carnets, Autorizaciones, etc) • Familia OpenData que contiene los tipos de contenido relacionados con la apertura de datos (Estadística, Meteorología, etc)
Tipo de contenido	El tipo de contenido es una naturaleza de contenido como las Ayudas, la Información General, las Estadísticas, etc

Específicamente en el caso de la iniciativa OpenData Euskadi, la tipología de contenidos es la siguiente:

Familia	Tipo de Contenido	Nombre interno del tipo
OpenData	Estadísticas	estadisticas
	Datos Meteorológicos	ds_meteorologicos
	Datos Geográficos	ds_geograficos
	Datos de Recursos Turísticos	ds_recursos_turisticos
	Datos de Ayudas y Subvenciones	ds_ayudas_subvenciones
	Datos de Autorizaciones	ds_autorizaciones
	Datos de Carnés	ds_carnes
	Datos de Registros	ds_registros
	Datos Inspecciones	ds_inspecciones
	Datos de Multas y Sanciones	ds_multas_sanciones
	Datos de Arbitrajes, Denuncias y Reclamaciones	ds_arbitrajes_denuncias_reclamaciones
	Datos de Contrataciones	ds_contrataciones
	Datos de Premios y Concursos	ds_premios_concursos
	Datos de Otros Procedimientos	ds_procedimientos_otros
	Datos de Anuncios de Contrataciones	ds_anuncios_contratacion
	Datos Generales	ds_general

En la iniciativa opendata.euskadi todos estos tipos de contenido se describen utilizando una ficha común que se describe a continuación:

3 Ficha de DataSet

3.1 Visión General

FICHA

1 Comparativa del absentismo escolar entre la Comunidad Autónoma del País Vasco y Nueva Gales del Sur

2 Visualizar datos estadísticos: Eustat

3 Descargar datos: CSV (310 kB) XLS (450 kB)

4 Detalle

5 Tipo: Datos estadísticos, estadística oficial

6 Tema: Educación y Ciencia

7 Etiquetas: absentismo, gales, niños, fracasos varios

8 Ámbito geográfico: Gipuzkoa, Bizkaia, Araba, Comunidad Autónoma del País Vasco

9 Fuente de los datos: EUSTAT, Cuenta de la Protección Social

10 Licencia: CC-by-SA -- Información legal

11 Fecha de actualización: 18/05/2008

12 Frecuencia de actualización: Anual

13 Fecha de inicio de datos: 18/02/2008

14 Fecha de fin de datos: 18/03/2008

15 Descripción: The Toxics Release Inventory (TRI) is a publicly available EPA database that contains information on toxic chemical releases and waste management activities reported annually by certain industries as well as federal facilities.

16 Documentación asociada: Calendario (.txt, 34 kB), Informes y documentos (.txt, 34 kB), Definiciones (.txt, 34 kB), Notas de prensa (.txt, 34 kB), Metodología y Calidad (.txt, 34 kB), Principales Indicadores Estructurales (.txt, 34 kB), Sitio web de los datos públicos de Agricultura en Gipuzkoa

DCR: Meta-Datos de la Ficha

	Etiqueta DCR	Descripción
1	datos_generales/denominacion	Título de la estadística
2	datos_generales/es_estadistica	Indica si los datos son estadísticos o no
3	datos_eustat/imagen_eustat	URL del icono del EUSTAT
4	documentos_descarga/documento_descarga/enlace_descarga/target_descarga/formato_descarga/tamaño_descarga/unidad_descarga	URL del dataset Indica si el dataset se abre en una nueva ventana Formato del dataset Tamaño del dataset Unidad del tamaño de dataset
5	datos_eustat/tipo_estadistica	Tipología y indicador de si la estadística es oficial
6	detalle_datos/tema	Catalogación del eje Tema/ Subtema
7	detalle_datos/etiquetas	Palabras claves de la ficha separadas por comas
8	detalle_datos/ámbito_geografico	Catalogación del Eje ámbito geográfico
9	detalle_datos/fuente_departamentos	Fuente de los datos
10	detalle_datos/licencia	Texto de la licencia
11	detalle_datos/fecha_difusion	Fecha de actualización
12	detalle_datos/periodicidad	Frecuencia de actualización
13	detalle_datos/periodo_inicio	Periodo desde el cual se han tenido en cuenta los datos para realizar la estadística
14	detalle_datos/periodo_fin	Periodo hasta el cual se han tenido en cuenta los datos para realizar la estadística
15	datos_generales/descripción	Descripción del DataSet
16	enlaces_documentacion_es_asciadas/enlace_documento_asociada/texto/enlace_target	Texto del documento asociado Enlace del documento asociado Indica si el documento se abre en una nueva ventana
17	datos_generales/orden	Campo no visible para ordenar estadísticas
18	datos_generales/Datasets	Campo no visible que recoge toda la información de los datasets

La Ficha OpenData presentada en la anterior tabla, muestra la ficha adaptada a los contenidos de tipología **Estadística**.

Para el resto de contenidos, se usará la misma ficha, pero sin especificar los datos relativos a Estadística.:.

Comparativa del absentismo escolar entre la Comunidad Autónoma del País Vasco y Nueva Gales del Sur

Descargar datos

 CSV (310 Kb.)  XLS (450 Kb.)

Detalle

 **2**  Tipo: Datos euskadi.net
 Tema: Educación y Ciencia
 Etiquetas: **absentismo** **gales** **ninos** **fracasos varios**
 Ámbito geográfico: Gipuzkoa, Bizkaia, Araba, Comunidad Autónoma del País Vasco

Fuente de los datos:
EUSTAT, Cuenta de la Protección Social  Fecha de actualización: 18/05/2008
 Frecuencia de actualización: Anual
 Fecha de inicio de datos: 18/02/2008
 Fecha de fin de datos: 18/03/2008

Licencia:  CC-by-SA → Información legal

Descripción

The Toxics Release Inventory (TRI) is a publicly available EPA database that contains information on toxic chemical releases and waste management activities reported annually by certain industries as well as federal facilities.

Envíanos tu opinión

 Comenta este conjunto de datos. Tu opinión nos ayuda a seguir mejorando.

Documentación asociada

-  [Calendario \(.txt, 34 kb.\)](#)
-  [Informes y documentos \(.txt, 34 kb.\)](#)
-  [Definiciones \(.txt, 34 kb.\)](#)
-  [Notas de prensa \(.txt, 34 kb.\)](#)
-  [Metodología y Calidad \(.txt, 34 kb.\)](#)
-  [Principales Indicadores Estructurales \(.txt, 34 kb.\)](#)
-  [Sitio web de los datos públicos de Agricultura en Gipuzkoa](#)

Marzo 2010

3.2 Representación en XML del DataSet

Todos los contenidos capturados en euskadi.net tienen una representación XML que recoge TODOS sus metadatos que pueden ser de dos tipos:

MetaDatos Comunes	Aquellos que tienen cualquier tipo de contenido independientemente de su tipo
MetaDatos Específicos	Aquellos que tienen los contenidos en función de su tipo

Desde el punto de vista de la reutilización estos XMLs se encuentran en dos localizaciones:



Por lo tanto, los XMLs reutilizables se encuentran en:

MetaDatos Comunes	Directorio raíz del contenido con el nombre: [contentName]-content.xml [contentName]-idxContent.xml
MetaDatos Específicos	En el directorio data de cada uno de los documentos (versiones) del contenido

Un ejemplo de la versión XML del dataSet que se puede encontrar en el directorio **data** es la siguiente:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE opendata SYSTEM
"/iwmnt/euskadiplus/main/r01_system/WORKAREA/wr01_system_loc/templatedata/editorial/opendata/openda.dtd">
<opendata>
    <datos_generales>
        <denominacion>
            Comparativa entre el absentismo escolar de la CAV y Nueva Gales del sur
        </denominacion>
        <es_estadistica>0</es_estadistica>
        <datos_eustat>
            <texto_eustat>
                Comparativa entre el absentismo escolar de la CAV y Nueva Gales del sur
            </texto_eustat>
            <enlace_eustat>
                http://www.eustat.es/ele0000400/tbl0000486_c.html
            </enlace_eustat>
            <target_eustat>0</target_eustat>
            <imagen_eustat>/imagen.jpg</imagen_eustat>
            <tipo_estadistica>29</tipo_estadistica>
        </datos_eustat>
        <documentos_descarga>
            <documento_descarga>
                <enlace_descarga>
                    http://www.eustat.es/ele0000400/tbl0000486_c.csv
                </enlace_descarga>
                <target_descarga>0</target_descarga>
                <formato_descarga>5061</formato_descarga>
                <tamano_descarga>310</tamano_descarga>
                <unidad_descarga>5058</unidad_descarga>
            </documento_descarga>
            <documento_descarga>
                <enlace_descarga>
                    http://www.eustat.es/ele0000400/tbl0000486_c.xls
                </enlace_descarga>
                <target_descarga>1</target_descarga>
                <formato_descarga>5062</formato_descarga>
                <tamano_descarga>450</tamano_descarga>
                <unidad_descarga>5058</unidad_descarga>
            </documento_descarga>
        </documentos_descarga>
        <Datasets>
            [{"tipo":"CSV","path":"http://www.euskadi.net","tamanio":"123","unidades":"Kb"}, {"tipo":"XLS","path":"http://www.euskadi.net","tamanio":"12","unidades":"Gb"}]
        </Datasets>
        <formats>5058,5062</formats>
        <descripcion>
            &lt;p&gt;The Toxics Release Inventory (TRI) is a publicly available EPA database that contains information on toxic chemical releases and waste management activities reported annually by certain industries as well as federal facilities&lt;/p&gt;
        </descripcion>
    </datos_generales>

```

```
</descripcion>
<orden>1</orden>
</datos_generales>
<detalle_datos>
    <tema>Educación y Ciencia</tema>
    <ambito_geografico>
        Gipuzkoa, Bizkaia, Araba, Comunidad Autonóma del País Vasco
    </ambito_geografico>
    <etiquetas>
        absentismo, gales, ni&ntilde;os, fracasosvarios
    </etiquetas>
    <fuente_departamentos>
        EUSTAT, Cuenta de la Protección Social
    </fuente_departamentos>
    <licencia>&lt;p&gt;fdsafdsafads&lt;/p&gt;</licencia>
    <fecha_creacion>01/03/2010</fecha_creacion>
    <fecha_difusion>01/03/2010</fecha_difusion>
    <periodicidad>33</periodicidad>
    <periodo_inicio>01/03/2010</periodo_inicio>
    <periodo_fin>01/03/2010</periodo_fin>
</detalle_datos>
<enlaces_documentaciones_asociadas>
    <enlace_documentacion_asociada>
        <texto>Informes y documentos</texto>
        <enlace_relacionado>
            http://www.eustat.es/estad/temalista.asp?tema=159
        </enlace_relacionado>
        <target_documento>0</target_documento>
    </enlace_documentacion_asociada>
    <enlace_documentacion_asociada>
        <texto>Definiciones</texto>
        <enlace_relacionado>
            http://www.eustat.es/estad/temalista.asp?tema=159
        </enlace_relacionado>
        <target_documento>0</target_documento>
    </enlace_documentacion_asociada>
    <enlace_documentacion_asociada>
        <texto>Notas de prensa</texto>
        <enlace_relacionado>
            http://www.eustat.es/estad/temalista.asp?tema=159
        </enlace_relacionado>
        <target_documento>0</target_documento>
    </enlace_documentacion_asociada>
    <enlace_documentacion_asociada>
        <texto>Metodología y Calidad</texto>
        <enlace_relacionado>
            http://www.eustat.es/estad/temalista.asp?tema=159
        </enlace_relacionado>
        <target_documento>0</target_documento>
    </enlace_documentacion_asociada>
    <enlace_documentacion_asociada>
        <texto>Cuestionario</texto>
        <enlace_relacionado>
            http://www.eustat.es/estad/temalista.asp?tema=159
        </enlace_relacionado>
        <target_documento>0</target_documento>
    </enlace_documentacion_asociada>
    <enlace_documentacion_asociada>
        <texto>Códigos y nomenclaturas</texto>
        <enlace_relacionado>
            http://www.eustat.es/estad/temalista.asp?tema=159
        </enlace_relacionado>
        <target_documento>0</target_documento>
    </enlace_documentacion_asociada>
    <enlace_documentacion_asociada>
        <texto>Principales Indicadores Estructurales</texto>
        <enlace_relacionado>
            http://www.eustat.es/indic/indicadores.asp?idioma=c
        </enlace_relacionado>
    </enlace_documentacion_asociada>
</enlaces_documentaciones_asociadas>
```

```
</enlace_relacionado>
<target_documento>0</target_documento>
</enlace_documentacion_asociada>
</enlaces_documentaciones_asociadas>
</opendata>
```

3.3 MetaDatos

Los Metadatos para los contenidos de OpenData son los siguientes:

Datos Generales			Tipo datos	de buscable?	Búsquedas			Ordenable?	Ficha?	Indexable
Nombre	Identificador	Descripción			Condiciones	Resultado Busq.?				
Denominación	DocumentName	Nombre Documento	Texto	✓	LIKE	✓	--	✓	✓	✓
Datos Estadísticos	OpendataEstadistic	Indica si son datos estadísticos	Texto	--	--	--	--	--	--	✓
Texto Eustat	OpendataTextHTML	Texto de la pagina Eustat con formato HTML	Texto	--	--	--	--	--	--	✓
URL Eustat	OpendataURLHTML	URL de la página EUSTAT con formato HTML	Texto	--	--	--	--	--	--	✓
Imagen Eustat	OpendataImageEustat	URL con el icono de EUSTAT	Texto	--	--	--	--	--	--	✓
Oficialidad	statisticOfficiality	Estadística oficial / no oficial	LISTA	✓	EQ NEQ	✓	--	✓	✓	✓
URL Descarga	OpendataURL	URL del Dataset	Texto	--	--	✓	--	✓	--	--
Formato Descarga	OpendataFormat	Formato del Dataset	Lista	✓	EQ	✓	--	✓	--	--
Tamaño Descarga	OpendataSize	Tamaño del Dataset	Numérico	--	--	✓	--	✓	--	--
Unidad Descarga	OpendataUnit	Unidad del tamaño del Dataset	Lista	--	--	✓	--	✓	--	--
Datasets	OpendataDatasets	Mapa con todos los Datasets	Mapa	✓	LIKE	--	--	--	--	✓
Etiquetas	OpendataLabels	Palabras clave del contenido	Texto	✓	LIKE	--	--	✓	✓	✓
Fuente Datos	OpendataSource	Fuente o fuentes propietarias de los datos	Texto	✓	LIKE	--	--	✓	✓	✓
Licencia	OpendataLicence	Tipo de licencia y enlace a la descripción de la misma	Texto	--	--	--	--	--	--	--
Fecha creación	OpendataReleaseDate	Fecha de liberación de los datos	Fecha	✓	EQ NEQ BET WE EN >= <=	✓	✓	✓	✓	✓
Fecha difusión	statisticDiffusionDate	Fecha de última actualización de los datos	Fecha	--	--	--	--	--	--	✓
Frecuencia de actualización	statisticRegularity	Frecuencia con la que se actualizan los datos	LISTA	✓	EQ	✓	✓	✓	✓	✓

Datos Generales			Tipo de datos	Buscable?	Búsquedas			Ficha?	Indexable
Nombre	Identificador	Descripción			Condiciones	Resultado Busq.?	Ordenable?		
Periodo inicio	OpendataStartDate	Fecha desde la cual se tienen en cuenta datos para el Dataset	Fecha	--	--	--	--	✓	--
Periodo fin	OpendataEndDate	Fecha hasta la cual se tienen en cuenta datos para el Dataset	Fecha	--	--	--	--	✓	--
Texto documento asociado	OpendataDocument	Descripción del documento asociado	Texto	--	--	--	--	✓	--
URL documento asociado	OpendataDocumentURL	URL del documento asociado	Texto	--	--	--	--	✓	--
Orden	OpendataOrder	Campo interno por el cual se ordenan por defecto las fichas	Numérico	--	--	--	✓	--	✓
Tema	OpendataTheme	Etiquetas Eje Catalogación Tema/Subtema Para datos estadísticos: Etiquetas Eje Catalogación propio de Temas de Estadística	Eje Catalogación	--	--	--	--	✓	✓
Ambito Geográfico	OpendataGeographical	Etiquetas Eje Catalogación ámbito geográfico	Eje Catalogación	--	--	--	--	✓	✓
Formatos	OpendataFormats	Id numéricos de los formatos	Texto	✓	LIKE	--	--	--	✓