

METADATOR, HERRAMIENTA DE CREACIÓN AUTOMATIZADA DE METADATOS

La producción de los metadatos es algo indispensable para los que trabajan con datos geográficos, pero es una tarea bastante pesada y repetitiva que finalmente no está bien hecha generalmente. Específicamente para los miembros de la Unión Europea donde la [directiva INSPIRE](#) impone muchas normas.

En el marco de mi actividad profesional al [INDECI](#), tuve que desarrollar una herramienta para facilitar esta parte de nuestro trabajo.

SUMARIO

Presentación.....	2
Funcionamiento general	4
El programa se encuentra en 2 versiones:.....	5
1- Versión ejecutable Windows:	5
2 - Versión Python (script):.....	5
Crear sus propios perfiles de metadatos.....	6
Traducciones e internacionalización	7
Agregar un idioma:	7
Los campos atributarios repetitivos	8
Para editar los campos:.....	8
Adaptación a los sistemas operativos (Mac y Linux)	9
Requisitos:.....	9
Descargar lo necesario con el terminal:.....	9
Uso:.....	9
Por venir: perspectivas del desarrollo	10
Referencias	10
Descargas.....	10

[Julien Moura](#)

21 de enero del 2013

PRESENTACIÓN



Lo que hace Metadator:

- extraer las informaciones sobre los archivos shapefiles (.shp) contenidos a dentro de una carpeta y sus subcarpetas.
- basarse sobre perfiles personalizables para todas las informaciones demás: institución, contactos, etc.
- exportar los metadatos en 4 formatos: html, Word (.doc/.docx), Excel 2003 (.xls) y XML ISO19139 (solamente una parte) que después se puede importar en un sistema de catálogo tal como [GeoNetwork](#) o [GéoSource](#).

Metadatos de Airports	
Nombre del archivo	airports.shp
Palabras claves temáticas	decisión e intervención ; agua ;
Palabras claves geográficas	Lima, Callao
Descripción	
Marco de creación de la información	Programa PACIVUR (Programa Andino de Capacitación e Investigación sobre Vulnerabilidad Urbana y Riesgos)
Número de objetos	76
Número de atributos	11
Fecha de la información	18-11-2011
Fecha de la última actualización	21-9-2012
Fuente(s)	IRD - SIRAD 2010 - PAVICUR 2012 - INEI
Difusión	libre
Responsable general	Frère TUCK (Hadas Cie), tuck.bro@barnsdale.re
Contacto de la fuente	Robin Des BOIS (Hadas Cie), robin.hood@barnsdale.fr
Página web	Servidor SIRAD
Tipo de geometría	Punto
Escala	1: 500000
Precisión	40000 m
Sistema de coordenadas	Projection : Albers Código EPSG: 4267
Extensión espacial	Max Y : 6502586.83 Min X : -4480198.52 Max X : 4615124.91 Min Y : 1433525.8

LISTADO DE LOS ATRIBUTOS

Parte del metadato al formato Word

Airports	
Affichage Actions	
Sud 4636291.01700, -7357522.88800 1433525.8000	
Mots clés	accessibilité (theme)
Mots clés	mobilité (theme)
Mots clés	vulnérabilité urbaine (theme)
Mots clés	Callao (place)
Mots clés	Lima (place)
Nom du système de référence	Albers (EPSG : 4267) (EPSG)
Les contacts	
A l'origine de	Adresse : 54 forêt de Sherwood Ville Nottingham Code postal 97418 Pays La Réunion Adresse e-mail tuck.bro@barnsdale.re
Contacts pour les métadonnées	
Point de contact	Adresse : 54 forêt de Sherwood Ville Nottingham Code postal 33140 Pays France Adresse e-mail robin.hood@barnsdale.fr
Les informations techniques	
Contraintes d'accès et d'utilisation	Contraintes légales Contraintes d'accès Droit d'auteur / Droit moral (copyright): Restriction liée à l'exercice moral Contraintes d'utilisation otherRestrictions Autres contraintes
Type de représentation spatiale	Vecteur: Donnée vecteur (point, ligne, polygone)

Metadato como se ve una vez importado en GéoSource (auqí la version en francés)

	A	B
1		Metadatos de Airports
2	Nombre del archivo	airports.shp
3	Palabras claves temáticas	decisión e intervención ; agua ;
4	Palabras claves geográficas	Lima, Callao
5	Descripción	DESCRIPCIÓN A COMPLETAR
6	Marco de creación de la información	Programa PACIVUR (Programa Andino de Capacitación e Investi
7	Número de objetos	76
8	Número de atributos	11
9	Fecha de la información	18/11/2011
10	Fecha de la última actualización	21/9/2012
11	Fuente(s)	IRD - SIRAD 2010 - PAVICUR 2012 - INEI
12	Difusión	libre
13	Responsable general	Frère TUCK (Hadas Cie), tuck.bro@barnsdale.re
14	Contacto de la fuente	Robin Des BOIS (Hadas Cie), robin.hood@barnsdale.fr
15	Página web	Servidor SIRAD
16	Tipo de geometría	Punto
17	Escala	500000
18	Precisión	40000
19	Sistema de coordenadas	Projection : Albers - Code EPSG : 4267
20	Extensión espacial	Min X : -4480198.52 - Max X : 4615124.98 // Min Y : 1433525.8 - Iv
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		

Parte de la primera pestaña del metadato en formato Excel

Desarrollado en Python, esta herramienta no hace milagro ni no evita el trabajo del profesional. Permite automatizar lo que puede serlo y así facilitar el trabajo para que el usuario solo tenga que focalizarse sobre lo que ninguna máquina puede hacer en su lugar: descripción de los atributos, pertinencia de las informaciones, precisiones, etc.

Todavía falta para que sea perfecto, pero he pensado que podría interesar varias personas.

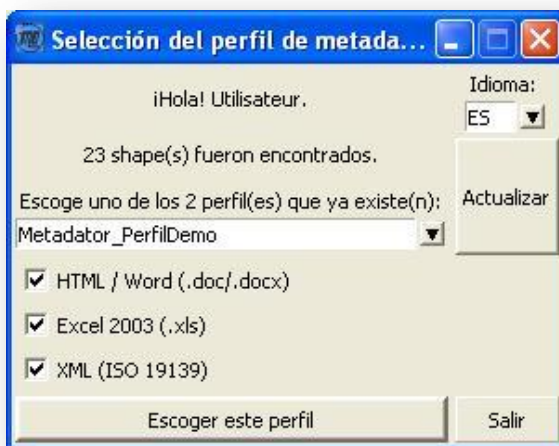
FUNCIONAMIENTO GENERAL

Metadator busca los shapes presentes dentro de la carpeta indicada por el usuario (también en las subcarpetas);

1. Aplica el [perfil de metadatos](#) elegido por el usuario para las informaciones generales;
2. Extrae las características de cada shape via OGR;
3. Calcula estadísticas básicas para describir los campos atributarios (suma, promedio, mediana, frecuencia de las ocurrencias, etc.);
4. Aplica las descripciones de los [campos atributarios repetitivos](#);
5. Crea los metadatos finales en el idioma y los formatos elegidos por el usuario.



Escoger la carpeta donde se encuentran los shapes



Escoger un perfil ya creado sino crear un nuevo.

EL PROGRAMA SE ENCUENTRA EN 2 VERSIONES:

1- VERSIÓN EJECUTABLE WINDOWS:

Uso más fácil pero menos flexible solo necesitando Windows y Microsoft Word.

Una vez el archivo zip extracto, dobla clic sobre **Metadator.exe**.

2 - VERSIÓN PYTHON (SCRIPT):

Uso más flexible pero necesitando unos requisitos y al menos un conocimiento mínimo del lenguaje Python:

- Windows XP SP3 o mas recién
- Microsoft Word 2003 o mas recién
- Python 2.7.x : <http://www.python.org/download/releases/2.7>
- Módulos python instalables desde [PyPi](#):
 - GDAL/OGR : <http://www.gdal.org/ogr>
 - Pmw : <http://pmw.sourceforge.net/doc/starting.html>
 - Dateutil : <http://pypi.python.org/pypi/python-dateutil/>
 - xlwt : <https://github.com/python-excel/xlwt>
 - numpy : <http://numpy.scipy.org/>
 - Pywin32 : <http://sourceforge.net/projects/pywin32/>

Es recomendado de instalar los módulos con setup ejecutables que se pueden encontrar en <http://www.lfd.uci.edu/~gohlke/pythonlibs> y también el paquete "Base". Una vez extracto el archivo zip descargado, iniciar Metadator.py desde el IDE Python o el terminal Windows (puede ser más lento):

```
cd ruta/donde_se_extractó/el_zip_descargado/Metadator
python Metadator.py
```

CREAR SUS PROPIOS PERFILES DE METADATOS

Para rellenar las informaciones generales sobre los datos, Metadator se base sobre perfiles de metadatos, aplicables después a varios shapes. El príncipe es que generalmente varias informaciones (sumario, contactos, palabras claves, etc.) son comunes a varios shapes. El programa permite crear los perfiles con la ayuda un formulario. Así el usuario puede crear tantos perfiles como le gusta o lo necesita.

Creación de un nuevo perfil de metadatos

PARAMETROS BÁSICOS

iHola! Utilisateur.

Nombre del fichero del perfil : 2013-01-16_ => O llenar los campos con este perfil:

DESCRIPCIÓN

Descripción: descripción general y común del marco en el cual fueron creados los datos.

Fuentes y histórico: origen e modalidades de la creación de los datos.

TEMÁTICAS INSPIRE

Aspectos geográficos de carácter meteorológico
Condiciones atmosféricas
Cubierta terrestre
Direcciones
Distribución de la población — demografía
Distribución de las especies
Edificios
Elevaciones
Geología
Hidrografía
Hábitats y biotopos
Instalaciones agrícolas y de acuicultura
Instalaciones de observación del medio ambiente
Instalaciones de producción e industriales
Lugares protegidos
Nombres geográficos
Ortoimágenes
Parcelas catastrales

=>

<=

[Vista a una parte del formulario](#)

El programa viene con dos perfiles de demostración: uno en castellano, el otro en francés.

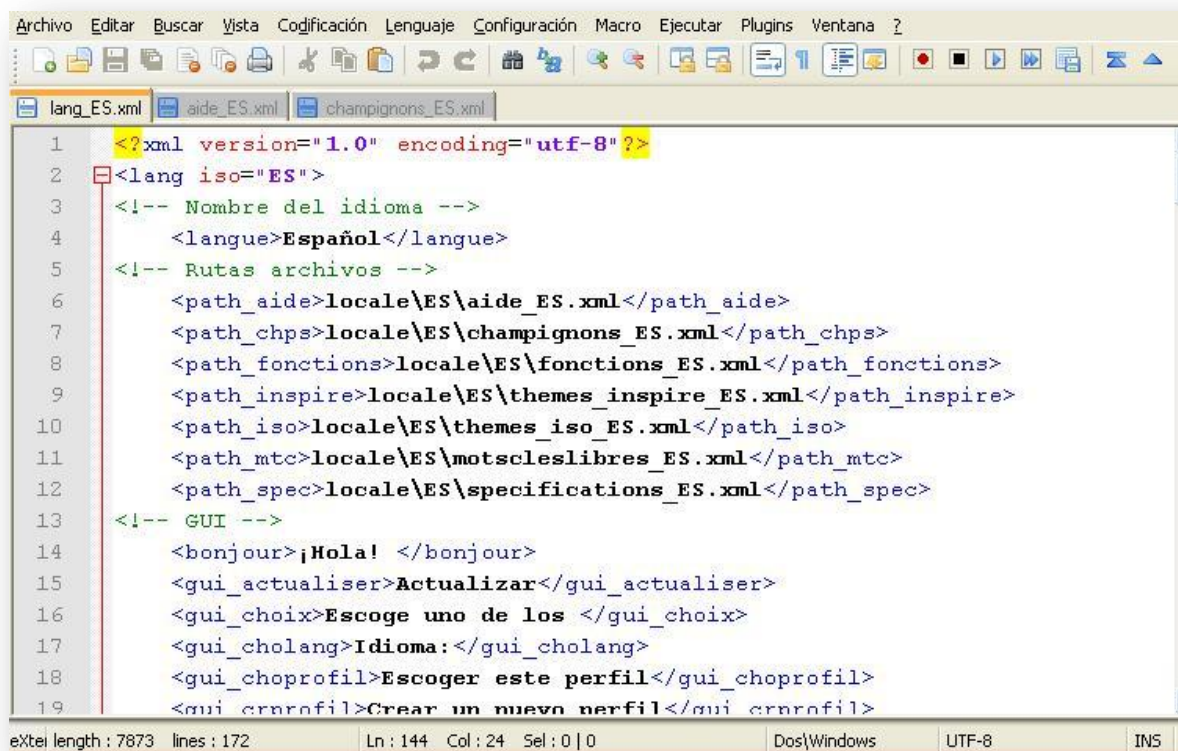
TRADUCCIONES E INTERNACIONALIZACIÓN

Inicialmente desarrollado en francés, el programa integra un sistema simple de internacionalización. También están incluidos 3 idiomas por defecto: francés, castellano e inglés.

Es muy fácil añadir un idioma o de personalizar los textos (de la interface como los de los archivos exportados) del programa.

AGREGAR UN IDIOMA:

- hacer una copia de la carpeta FR dentro de la carpeta locale;
- renombrar la carpeta creada con la abreviatura correspondiendo al nuevo idioma que se quiere crear;
- aplicar asimismo a los sufijos de los nombres de los archivos dentro de la carpeta creada;
- editar el archivo lang_XX.xml:



```
1 <?xml version="1.0" encoding="utf-8" ?>
2 <lang iso="ES">
3 <!-- Nombre del idioma -->
4 <langue>Español</langue>
5 <!-- Rutas archivos -->
6 <path_aide>locale\ES\aide_ES.xml</path_aide>
7 <path_chps>locale\ES\champignons_ES.xml</path_chps>
8 <path_fonctions>locale\ES\fonctions_ES.xml</path_fonctions>
9 <path_inspire>locale\ES\themes_inspire_ES.xml</path_inspire>
10 <path_iso>locale\ES\themes_iso_ES.xml</path_iso>
11 <path_mtc>locale\ES\motscleslibres_ES.xml</path_mtc>
12 <path_spec>locale\ES\specifications_ES.xml</path_spec>
13 <!-- GUI -->
14 <bonjour>¡Hola! </bonjour>
15 <gui_actualiser>Actualizar</gui_actualiser>
16 <gui_choix>Escoge uno de los </gui_choix>
17 <gui_cholang>Idioma:</gui_cholang>
18 <gui_choprofil>Escoger este perfil</gui_choprofil>
19 <gui_crnprofil>Crear un nuevo perfil</gui_crnprofil>
```

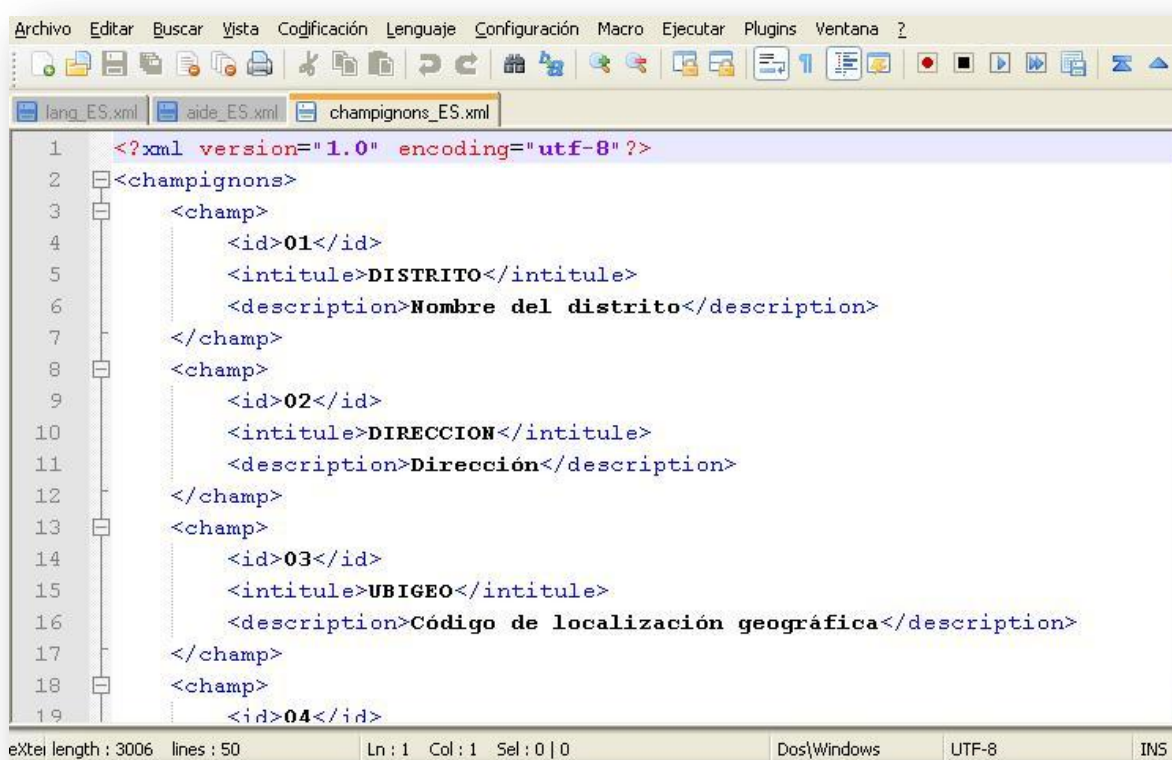
Vista a una parte del archivo lang_ES.xml

LOS CAMPOS ATRIBUTARIOS REPETITIVOS

Observando que dentro de la misma actividad profesional, unos campos atributarios están casi siempre presentes en las capas de información, el programa permite facilitar la descripción de estos campos y también la coherencia global de los metadatos a la escala de una empresa o institución. Por ejemplo, en mi trabajo en el Perú, siempre tengo el UBIGEO, los códigos administrativos del INEI para los lugares. También, cuando trabajo con datos extractos desde OpenStreetMap, siempre encuentro los mismos atributos. Es muy pesado de siempre re copiar las mismas descripciones, así que Metadator permite al usuario reducir el peso de esta tarea y personalizar los campos atributarios repetitivos.

PARA EDITAR LOS CAMPOS:

1. editar el archivo champignons_XX.xml, donde XX corresponde al idioma querido;
2. agregar/eliminar/modificar el ID, el nombre y la descripción de los atributos repetitivos.



Parte del archivo de los atributos repetitivos

Cuidado, los nombres de los campos tienen que hacer coincidir mayúsculas y minúsculas. Así id e ID son dos atributos diferentes para Metadator.

Es compatible con el código html para personalizar las letras: `negrita`, etc.

ADAPTACIÓN A LOS SISTEMAS OPERATIVOS (MAC Y LINUX)

REQUISITOS:

Python 2.7.x

Módulos indicados más arriba excepto Pywin32 obviamente.

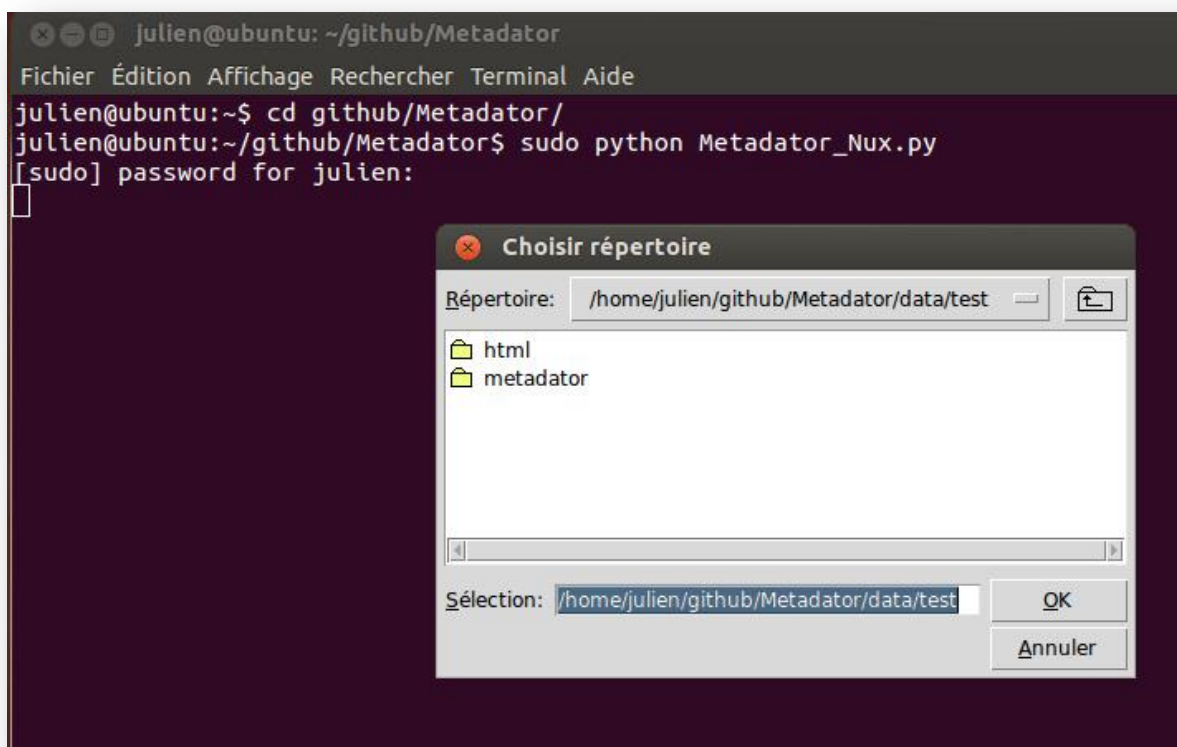
DESCARGAR LO NECESARIO CON EL TERMINAL:

```
cd ruta/donde_quiere_instalarlo  
git clone https://github.com/Guts/Metadator.git --branch  
alpha-unix
```

Uso:

Iniciar el archivo Metadator_Nux.py tomando en cuenta que no es disponible el formato Word para exportar.

Esta versión no está mantenida. Ha sido probada con Ubuntu 12.04.

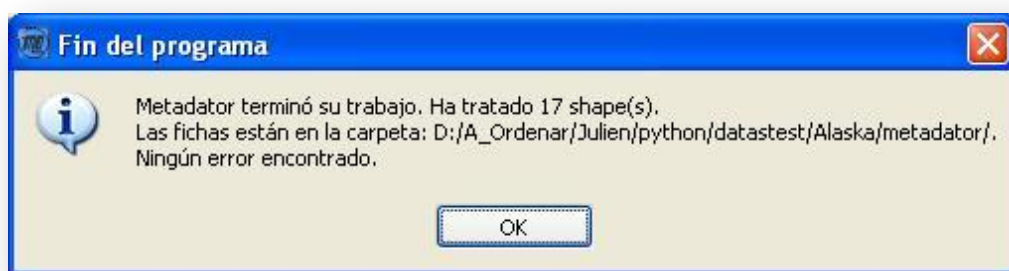


Ventana para elegir la carpeta en Ubuntu 12.04 (versión francesa)

POR VENIR: PERSPECTIVAS DEL DESARROLLO

A pesar de un ritmo de desarrollo lento, algunos mejoramientos son considerados:

- mejor tolerancia a los caracteres especiales
- exportar en .odt
- exportar en html con xslt
- exportar el catálogo de los atributos (ISO 19110)
- compatibilidad con archivos MapInfo
- terminar el sistema de ayuda para el formulario de creación de perfil



Notificación y informe finales del programa. El error indicado es voluntario para los necesitos del desarrollo.

REFERENCIAS

Gracias a Pierre por haber creado el embrión de Metadator (DataInspector) y prestarme su blog para presentarlo: <http://www.datamadre.com/posts/17>

Desarrollado con PyScripter : <https://code.google.com/p/pyscripter/>

DESCARGAS

=> [Versión estable en GitHub](#)

=> [Versión en desarrollo en GitHub](#)

=> [Versión Ubuntu en GitHub](#)