eSoterParser utility v1.4 – Manual and Instructions

Introduction

One of the key issues for the e-SOTER project WP6 is the transferring of any of the large set of legacy data files from the old SOTER format to the new SOTER-ML Soter XML-based mark-up language. The eSoterParser tool documented here has been written by the Cranfield University team as a common utility able to do this by taking an existing standard SOTER data source and transforming it.

Prerequisites

In order to run the utility, you should have a 32bit Windows operating system, having a valid ODBC driver for Access (this would be the default if you have MS Office installed).

Your computer must have the Java SE Runtime, or the Java SDK already installed and correctly configured on the PC.

To test that Java is correctly installed (e.g. all the paths and environment variables are set up correctly) - type at the command prompt "java". If you then receive a page full of java help text, then all is well, if not consult your java installation notes.

This code has been tested and runs successfully on standard Windows XP and Vista computers.

Note, should you wish to *edit* the code etc., we would recommend using the free IDE toolkit 'NetBeans' (http://netbeans.org/).

Instructions on use

1. Download and unpack the compressed folder "eSoterParser.zip", say to c:\eSoterParser\.

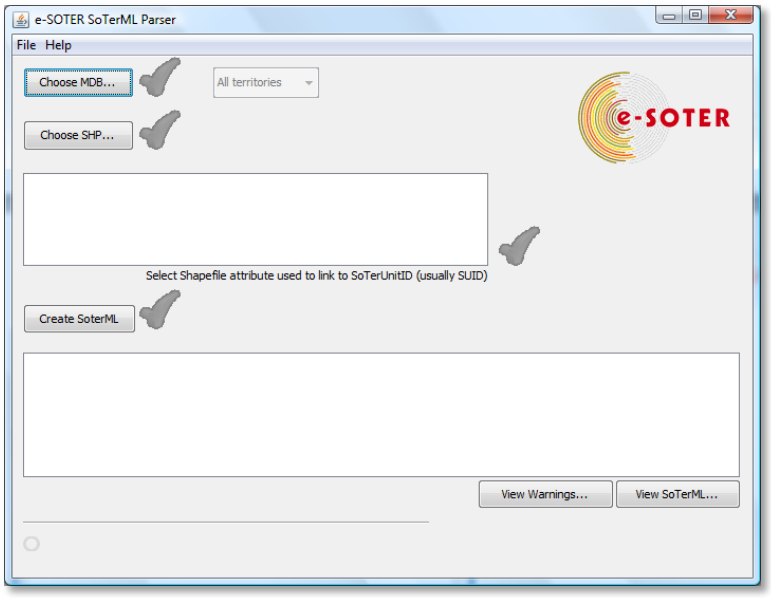
2. All compiled binary files required to run the tool are contained within this folder (the distribution code was compiled in Java). The source code can also be provided if required.

3. Sample SOTER databases are provided in the "SampleData" subfolder – for the purposes of these instructions, to get this working the first time, we will use the Kenya legacy dataset (".\eSoterParser\SampleData\Kenya"). These datasets are available as downloads on the ISRIC website (http://www.isric.org/UK/About+Soils/Soil+data/Geographic+data/).

4. Run eSoterParser by typing at the prompt:

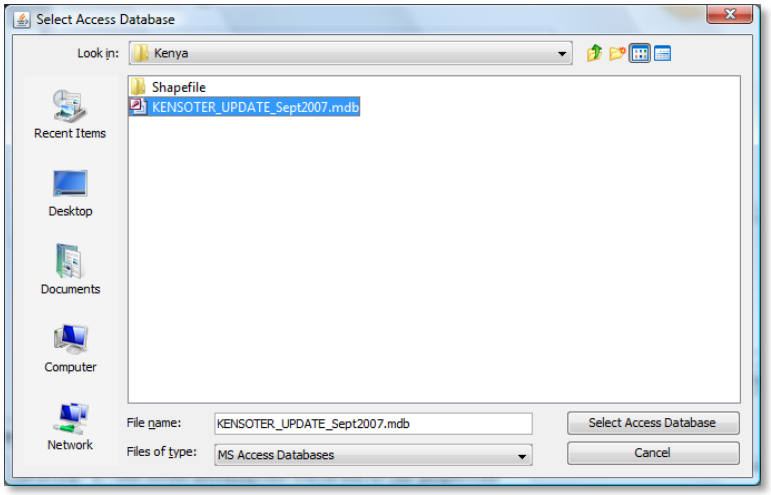
```
java -jar "eSoterParser.jar"
```

5. This starts the GUI for the tool, allowing you to select the relevant legacy SOTER database files. There are three distinct actions required to select these files. The GUI guides you through with 'tick' marks that go green when selections are made successfully.



Starting screen of the utility.

The first step is to select the Microsoft Access (version 2007) containing the non-spatial attributes. Use the first button to open a dialogue box to navigate to and select the MDB database file.

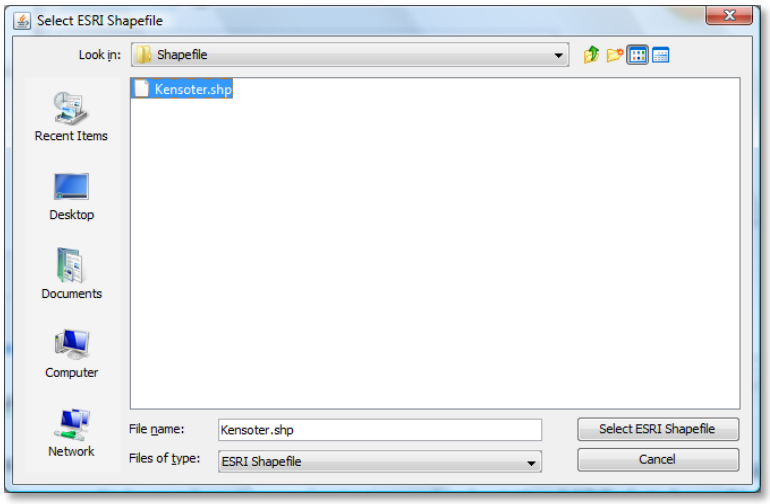


Selecting an Access database.

SOTER databses can hold data for several territories. By default all records are used, however the parser also allows you to filter out data for only one territory. Use the drop-down box to select.

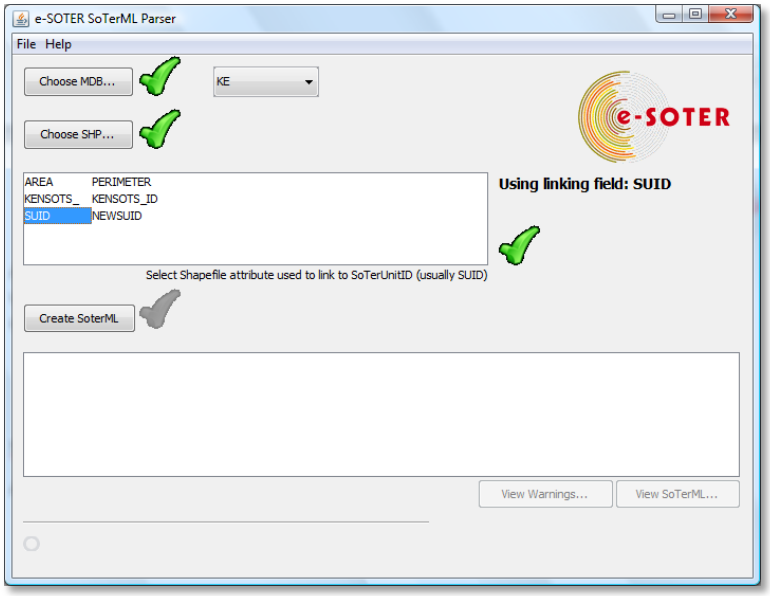


The second step is to select the GIS Shapefile database (DBF file) containing the spatial elements of the Soter Units. Use the second button to open a dialogue box to navigate to and select the DBF database file.



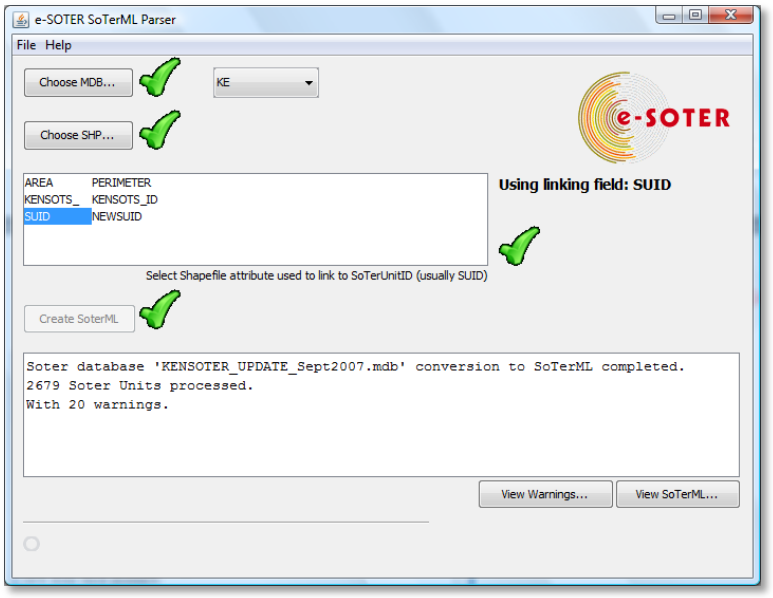
Selecting a Shapefile database.

The third step is to select the Shapefile attribute database field from the DBF file that will be used to connect the records in this table with the records in the MS Access table. This is how the legacy SOTER file data structure represents the link between the geometric and attribute features of the data. The utility tool presents a dynamic list of field names for selection. The field required is usually named "**SUID**" (Soter Unit Identifier).



Selecting the linking field.

The fourth step is to select the "CreatSoterML" button to start the conversion. Conversion may take some time, but eventually, a new output SOTERML file is created in the database folder "SoTerML.xml". This is a plain text file (althoughlikely a rather large one). You can then view this SoTerML file or also view any errors that were reported.



Running the conversion.

A message box provides feedback as to the success of the operation. Further buttons allow you to view in Notepad the output SoterML.xml file and the Errors log, both files being created in the folder containing the source data. Rerunning the conversion overwrites these files.

S.Hallett/A.Rayner/D.Simms 22/Oct/2011

Any queries, please contact s.hallett@cranfield.ac.uk