



DataExtractor

Data Extractor - User Guide

Release 1.0.1

**Hester Lyons
Andy Foy**

Jan 05, 2018

List of Figures	1
1 Preface	3
1.1 Recommended User Knowledge	3
1.2 Reading Guide	4
1.3 Licensing	4
1.4 Useful links	4
1.5 Acknowledgements	5
1.6 Conventions used in this user guide	6
2 Introduction	7
2.1 Background	7
2.2 Tool overview	8
2.3 Benefits	9
3 Anatomy of data extraction	11
3.1 Data extraction process	11
3.2 The Data Extractor tool	12
3.3 Tool outputs	14
4 Setting up the tool	19
4.1 Configuring the tool	19
4.2 Setting up the SQL Server database	29
4.3 Installing the tool	30
5 Running the tool	47
5.1 Opening the form	48
5.2 Using the form	50
5.3 Extract results	54
6 Frequently Asked Questions	55
6.1 General questions	55
6.2 Operating the tool	55
6.3 Tool issues	56
7 Appendix	57
7.1 Example XML file	57

7.2 GNU Free Documentation License 63

List of Figures

2.1	The Data Extractor tool interface	8
3.1	A MapInfo workspace configured for using the Data Extractor tool	12
3.2	The Data Extractor tool workflow	13
3.3	Example of the Data Extractor tool output folder	14
3.4	Example of a text file output from the Data Extractor tool	15
3.5	Example of a compressed output file containing a single GIS layer (MapInfo)	16
3.6	Example of a Data Extractor tool log file	17
4.1	Example of a partner table loaded into MapInfo	22
4.2	Simplified example of an SQL table subset configuration	23
4.3	Example of a partner table	25
4.4	Simplified example of an SQL table subset configuration	27
4.5	The Tool Manager in MapInfo 12 or earlier	30
4.6	Adding a tool in Tool Manager	31
4.7	The Data Extractor tool is loaded	32
4.8	The Data Extractor tool menu	32
4.9	Installing the Data Extractor tool from Windows Explorer	34
4.10	Installation begins after clicking 'Install Add-in'	35
4.11	Starting the ArcGIS Add-In Manager	36
4.12	The 'Options' tab in the ArcGIS Add-In Manager	37
4.13	The ArcGIS Add-In Manager showing the Data Extractor tool	38
4.14	Starting Customize Mode in ArcGIS	39
4.15	Customising the document in ArcGIS	40
4.16	Adding a new toolbar in ArcGIS	41
4.17	Naming the new toolbar in ArcGIS	42
4.18	New toolbar added to the ArcGIS Interface	43
4.19	Finding the Data Extractor tool in the add-in commands	44
4.20	Adding the Data Extractor tool to the new toolbar	45
4.21	Locating the configuration file folder	46
5.1	Launching the Data Extractor tool	48
5.2	Example warning message when any GIS layers are not found	48
5.3	The form displaying the active partners and available data layers	49
5.4	Select the required options on the form	50

5.5	Progress window during the extract process	51
5.6	Prompt to close form when extract process completes	52
5.7	Example log file shown for review	53
5.8	Outputs are organised in partner-specific folders	54

Permission **is** granted to copy, distribute **and/or** modify this document under the terms of the GNU Free Documentation License, Version 1.3 **or** any later version published by the Free Software Foundation; **with** no Invariant Sections, no Front-Cover Texts **and** no Back-Cover Texts. A copy of the license **is** included **in** the Appendix section.

The most up to date version of this documentation can be found in **HTML** and **PDF** form on [ReadTheDocs¹](#).

1.1 Recommended User Knowledge

1.1.1 Users

This user guide assumes that users of the Data Extractor tool have:

- General IT experience including the use of Microsoft Windows.
- Experience in the use of a relevant GIS application supported by the tool (currently MapInfo and ArcGIS), including selecting and querying features and attributes.
- An understanding of the datasets that are used by the Data Extractor tool.

1.1.2 Administrators

It is recommended that a person within each organisation is designated as the tool and database administrator. This person should:

- Have an understanding and experience of IT systems management.
- Understand relational database structures.
- Be an expert user of SQL Server.
- Have certified training or equivalent experience in advanced features of the relevant GIS software.
- Become familiar with how the Data Extractor tool has been configured within the organisation.
- Have a good understanding of XML.

¹ <https://readthedocs.org/projects/dataextractor-userguide/>

1.2 Reading Guide

This Preface explains a little about the Data Extractor tool, the community of people who develop and use it, and the licensing conditions for using and distributing it. It also explains how to read this user guide.

Introduction (page 7) explains why the Data Extractor tool is needed, what it does and where it comes from.

Anatomy of data extraction (page 11) is a brief outline of the key stages of data extraction, and how the Data Extractor tool addresses these.

Setting up the tool (page 19) describes how to install and set up the Data Extractor tool.

Running the tool (page 47) describes how to run the Data Extractor tool.

FAQs (page 55) has a list of commonly asked questions and their answers.

Appendix (page 57) contains examples of the XML configuration files for MapInfo and ArcGIS, lists known issues with the tool and contains a copy of the GNU Free Documentation License v1.3 covering this guide.

1.3 Licensing

The code for the Data Extractor tool is ‘open source’ and is released under the [GNU General Public License \(GPL\) v3](#)². Users are free to install it on as many computers as they like, and to redistribute it according to the GPLv3 license.

This guide is released under the [GNU Free Documentation License \(FDL\) v1.3](#)³. Permission is granted to copy, distribute and/or modify this document under the terms of the license.

Please remember, however, that the tool cost a lot of money to develop and still requires further development and ongoing support. Hence any contributions towards costs would be gratefully received. Enquiries can be made via email to either [Hester](mailto:Hester@HesterLyonsConsulting.co.uk)⁴ or [Andy](mailto:Andy@AndyFoyConsulting.co.uk)⁵.

1.4 Useful links

Related community links:

- Administrators: ([MapInfo Installation](#)⁶) and ([ArcGIS Installation](#)⁷) - Release notes and installers.
- Developers ([MapInfo Source Code](#)⁸) and ([ArcGIS Source Code](#)⁹)- Source code for the Data Extractor tool.

² <http://www.gnu.org/licenses/gpl.html>

³ <http://www.gnu.org/licenses/fdl.html>

⁴ Hester@HesterLyonsConsulting.co.uk

⁵ Andy@AndyFoyConsulting.co.uk

⁶ <https://github.com/LERCAutomation/DataExtractor-MapInfo/releases/>

⁷ <https://github.com/LERCAutomation/Data-Extractor---ArcObjects/releases>

⁸ <https://github.com/LERCAutomation/DataExtractor-MapInfo>

⁹ <https://github.com/LERCAutomation/Data-Extractor---ArcObjects>

- Issues ([Known issues for MapInfo version¹⁰](#)) and ([Known issues for ArcGIS version¹¹](#)) - Details of known issues and existing change requests.

1.5 Acknowledgements

Many thanks are due to all the LERCs in the south-east of England and their staff who have, and continue to, fund and contribute to the Data Extractor tool. It takes many developers, testers and users to build a truly useful tool (especially users who care enough to test new releases, report bugs and discuss feature requests).

¹⁰ <https://github.com/LERCAutomation/DataExtractor-MapInfo/issues>

¹¹ <https://github.com/LERCAutomation/Data-Extractor---ArcObjects/issues>

1.6 Conventions used in this user guide

The following typographical conventions are used in this manual:

Ctrl-A Indicates a key, or combination of keys, to press.

Commit Indicates a label, button or anything that appears in user interfaces.

Tools... -> About Indicates a menu choice, or a combination of menu choices, tab selections or GUI buttons.

C:\Program Files (x86)\MapInfo\Professional Indicates a filename or directory name.

Tip: Tips can help save time or provide shortcuts.

Note: Notes explain things in more detail or highlight important points.

<p>Caution: Warnings where users should pay attention.</p>

2.1 Background

Carrying out data extractions (i.e. extracting records for partners on, for example, species or local nature reserves) is a routine task for many Local Environmental Record Centres (LERCs). The process is a repetitive one, with the same kind of extraction being carried out for what can be a considerable number of partners. Mostly, the only difference between partners is the geographical area they focus on, while the data they require can be represented by a few standard tables. Therefore this is a process that is ideally suited to being automated.

The Data Extractor tool was originally developed for Greenspace Information for Greater London (GiGL) and implemented in MapInfo. Currently the tool is used by a variety of LERCs and a version for ArcGIS is under consideration.

2.2 Tool overview

The Data Extractor tool is configurable in a flexible way according to the requirements of the LERC or individual user through a configuration document. It is integrated into the user interface of the GIS system and presented there as menu item. The tool itself has a simple interface (Fig. 2.1), requiring a minimum of input (the user is requested to select which partners an extract will be created for, and from which tables). There are some additional options to tailor the extracts to include confidential records, and some simple output options. Once set up, the tool communicates with both the GIS system and an associated SQL database to extract the required data.

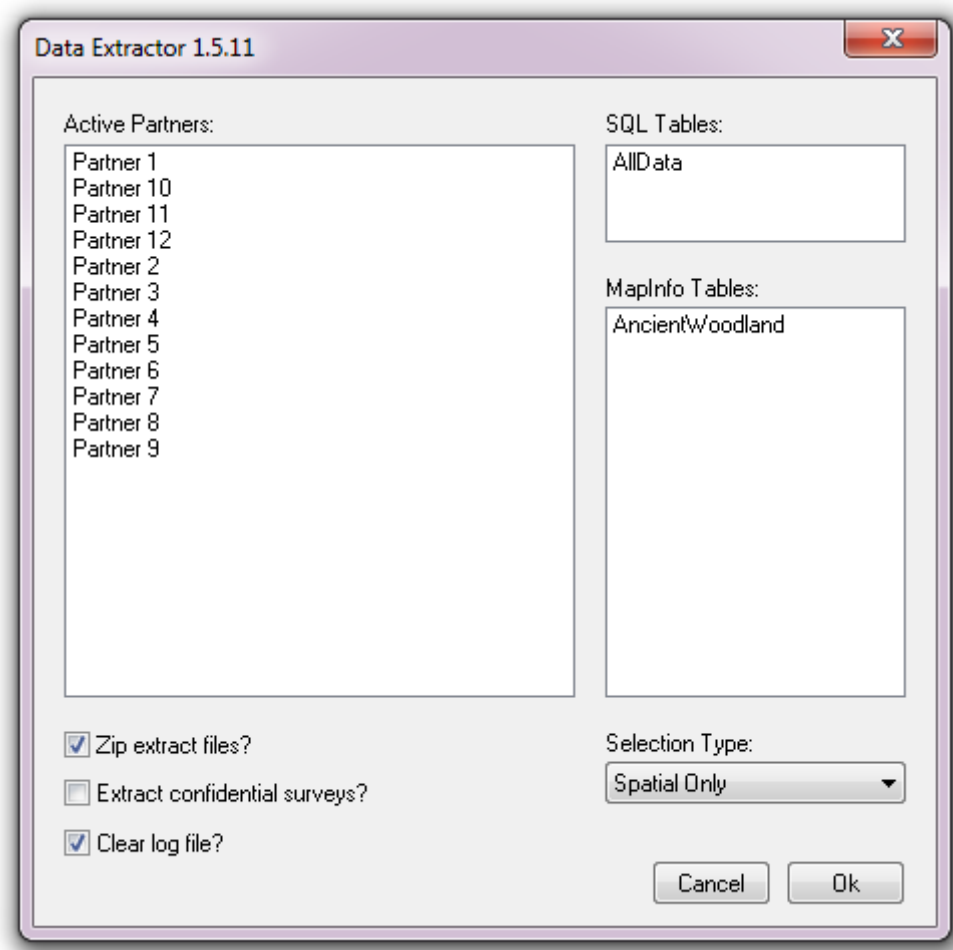


Fig. 2.1: The Data Extractor tool interface

Data layers for the tool can be contained in an SQL Server database or as GIS layers loaded in the GIS application. When running an extraction the tool uses a GIS layer, which is also present in the associated SQL database, to find the boundary for each partner and extract the records which fall within this boundary. The attributes for each partner in this GIS layer define which of the available data layers will be extracted and in which format. Extracts are saved in a predefined location, and a log file is kept that records the steps of each extraction. The process is discussed in this document in more detail in the section on *running the tool* (page 47).

Defining the way that extractions should be carried out, the output that they generate and the layers that can potentially be included is done via a configuration document written in XML. Using this document the user can configure all the parts of the extraction, for example:

- The name of the geographic layer containing the partner boundaries, and its key columns.
- The location of the SQL Server file DNS (defining the connection details for the database).
- The location of the output folder.
- For each data layer, a detailed definition of what information should be returned from it.
- Details on the display and labelling of output from individual data layers where relevant.

Using this configuration file, each individual LERC can tailor the Data Extractor tool to its individual requirements. Examples of the XML file are included in the [Appendix](#) (page 57), and the process of setting up this file is discussed in the section on [setting up the tool](#) (page 19).

Additionally to the XML file the SQL server database is set up via a number of auxiliary tables and stored procedures. Again, the process of configuring this is discussed in the section on [setting up the tool](#) (page 19).

2.3 Benefits

There are a number of clear benefits to using the Data Extractor tool for carrying out routine data extractions for partners.

1. The tool, by encapsulating and automating the process, saves considerable time over carrying out these extractions manually.
2. Both the process and the outputs of the extraction are standardised, therefore minimising the risk of user error that is present in a manual extraction.
3. By specifying the outputs of the tool centrally through the configuration file, the output for each extraction is consistent with all other extractions, regardless of the individual carrying out the extraction. This leads to comparability of results and a predictable experience for the users of a data extraction service.
4. The extractions are repeatable and, through the inclusion of the log file, automatically documented.

Anatomy of data extraction

This section describes how a typical data extraction might be carried out manually, and how the Data Extraction Tool automates this process. Please note that the examples used in this illustration are purely fictional and do not represent a real-world scenario.

3.1 Data extraction process

The process of a typical data extraction can be broken down into a number of distinct steps that are described here. In the next section the way that the Data Extraction Tool carries out these steps is explained.

Defining a partner boundary Before any extraction can be carried out, a polygon describing the area for which data will be extracted for a partner has to be entered into a GIS system. This would be held in a single GIS layer together with some associated attributes such as the name of the organisation represented by the boundary. Once this area has been entered into the data layer it can be used time and again.

Selecting the relevant data layers Using the boundary defined in the previous step, each of the data layers relating to the presence of protected sites, habitats, species, etc. is selected one by one. Where data is held within SQL Server this process is carried out within the SQL Server database.

Exporting the results The selected features are extracted in the format required by the user, containing only the relevant columns required from each data layer. Symbology may need to be defined at this step.

Repeating the process Where there is more than one partner for which data needs to be extracted, the process will be repeated for each partner boundary.

3.2 The Data Extractor tool

3.2.1 Tool components

There are four component parts to the Data Extractor tool that work together to automate the process described above:

1. A GIS layer that describes the boundaries of all relevant partners and stakeholders. This is held both within the GIS system and within SQL Server.
2. Spatial data held in an SQL database and / or in spatial data layers within the GIS system. Where data is held within SQL Server a stored procedure for its extraction is also required.
3. An XML configuration file that specifies how the extractions are set up and what data should be exported for each data layer.
4. The Data Extractor tool itself.

The Data Extractor tool is used within a GIS environment and requires all the required data layers to be preloaded in the GIS (see Fig. 3.1). Where data is to be extracted from SQL Server the partner boundary layer must also be preloaded into the SQL Server database.

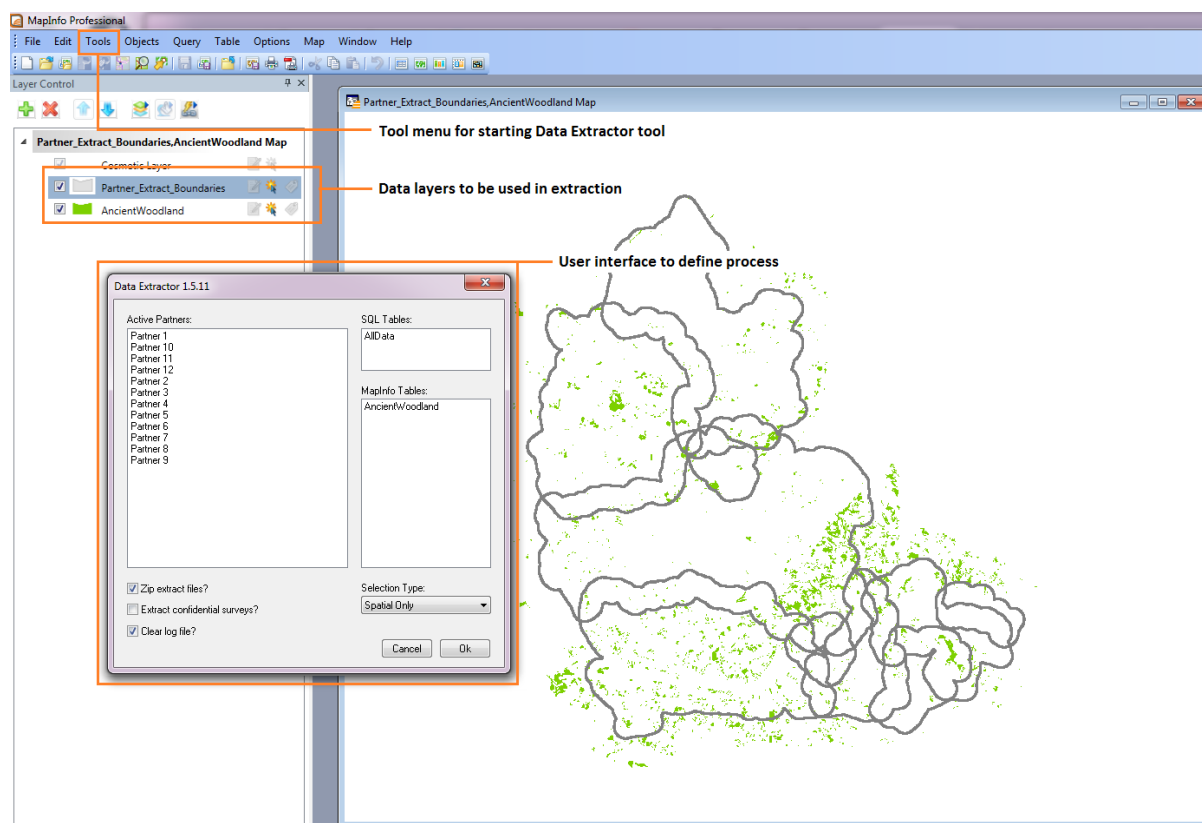


Fig. 3.1: A MapInfo workspace configured for using the Data Extractor tool

3.2.2 Tool workflow

The Data Extractor tool requires minimum user input in order to carry out its processes once it is configured. The simple workflow is as follows (see Fig. 3.2):

1. The user selects which partner(s) the extraction should be carried out.
2. The user specifies which data layers to extract from. Only layers that are loaded in the GIS or tables found in the SQL Server database are made available at this point.
3. The user selects whether the extracted files should be added to a zip file, whether confidential data from any SQL Server tables should be included, and whether the log file should be cleared before the process starts.
4. Finally, the user selects whether the selection of SQL Server data should be based on spatial location only, survey tags (names) only, or both. This allows for the inclusion of data relevant to a partner that is outside of that partner's boundary.
5. Once the user clicks **OK** the process starts.

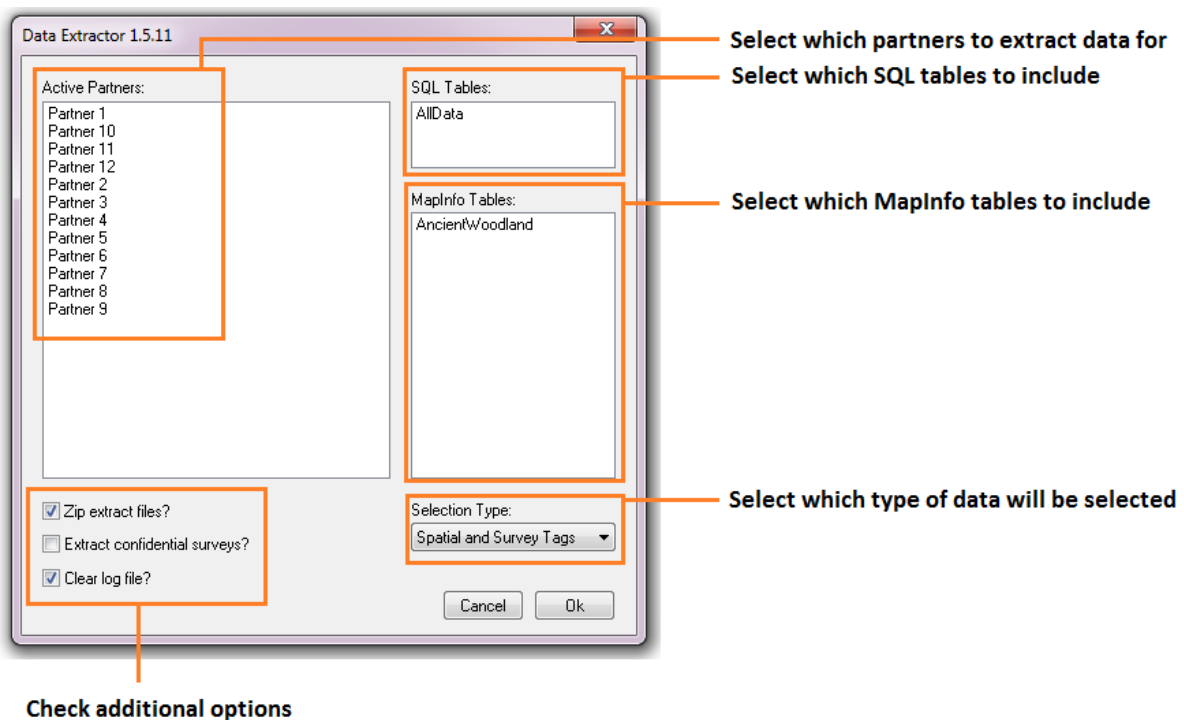


Fig. 3.2: The Data Extractor tool workflow

In essence, the process that the tool follows is identical to the manual process a user would perform:

1. The boundary of each selected partner is processed in sequence.
2. The specified SQL and GIS data layers are selected using the boundary (and/or the survey tags) for this partner.
3. The resulting selections are exported to the output folder as specified in the configuration file, using the columns and symbology specified in this configuration file (for SQL data) or defined within the GIS layer.
4. During the process the tool reports its progress to a log file and, when the process finishes, this log file is displayed allowing the user to assess the success of the data extraction.

3.3 Tool outputs

Below is a selection of outputs generated from an example data extraction using the data selections shown in figure Fig. 3.1. The extraction was carried out for all partners shown in the menu.

3.3.1 Output folder

The outputs are stored in a user defined folder (Fig. 3.3). These outputs may include a combination of GIS layers and text files in different formats, and the log file.

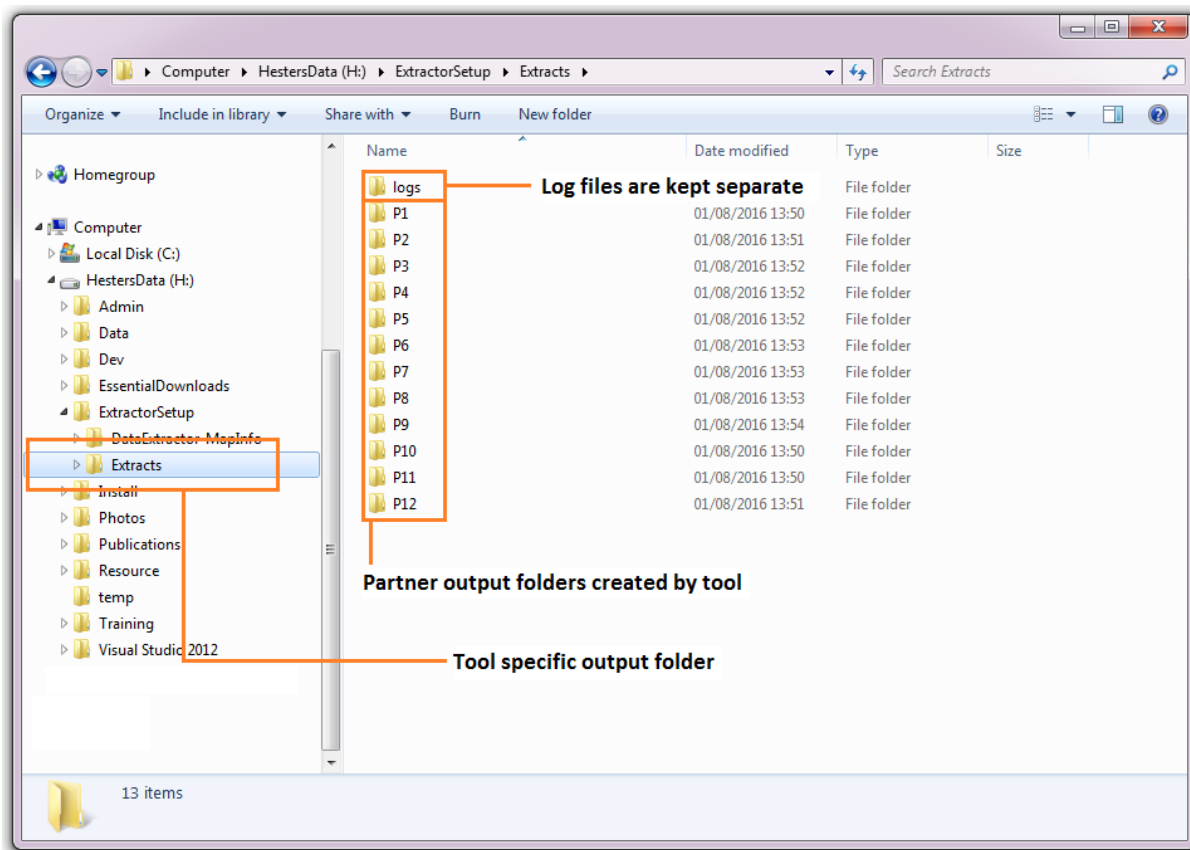


Fig. 3.3: Example of the Data Extractor tool output folder

3.3.2 Output files

Text file outputs can be produced in CSV format (Fig. 3.4). GIS layers can be output in MapInfo (.tab) format and / or converted into ArcGIS (.shp) format.

	A	B	C	D	E
1	Species	RecYear	GridRef		
2	Species1	2016	SU730870		
3	Species1	2015	SU730870		
4	Species1	2014	SU730870		
5	Species2	2016	SP433794		
6	Species2	2012	SP433794		
7	Species2	2010	SP433794		
8					
9					
10					
11					

Fig. 3.4: Example of a text file output from the Data Extractor tool

3.3.3 Output options

Options in the tool include compressing all outputs into a single zip file for each partner (Fig. 3.5; MapInfo only), including confidential records (defined in the configuration document) in any SQL table extracts, and clearing the log file before use.

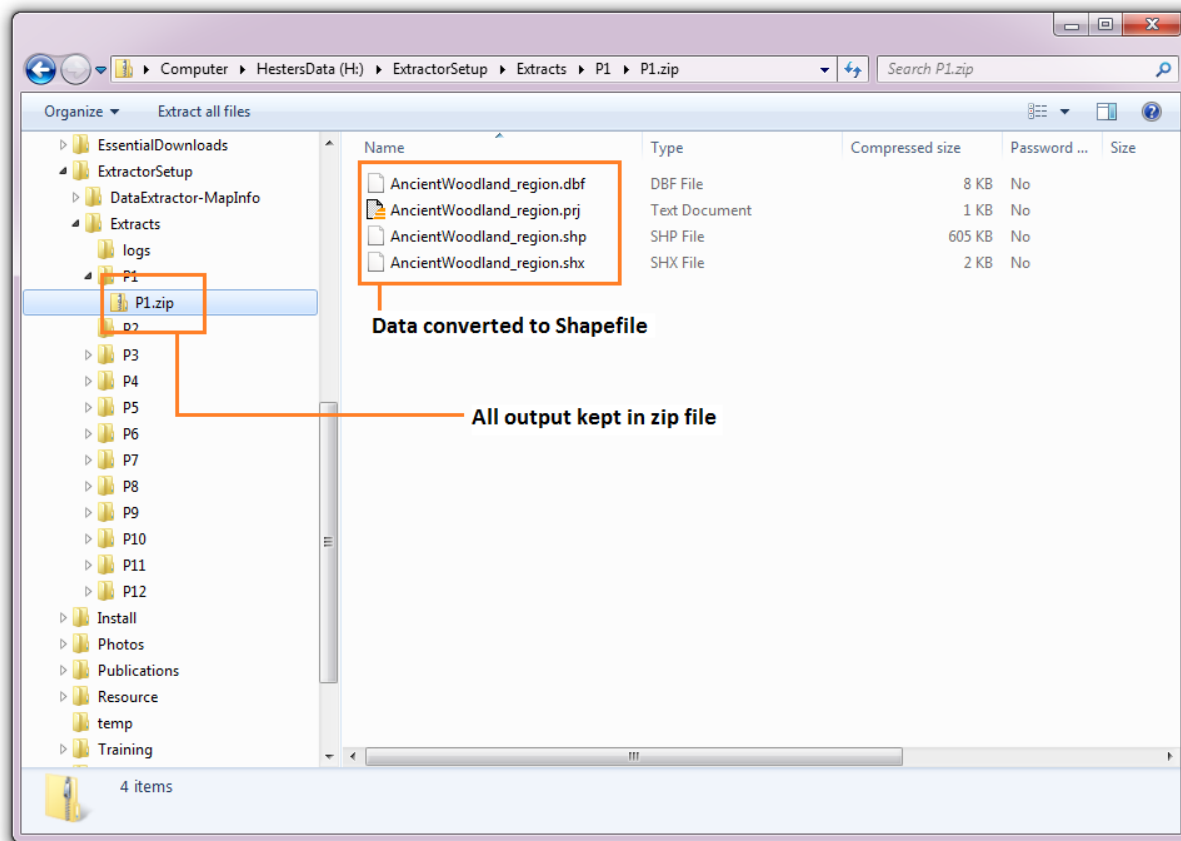


Fig. 3.5: Example of a compressed output file containing a single GIS layer (MapInfo)

Finally, the log file details each step that was taken during the process, and gives some feedback about the outcomes of the steps. This includes reporting on the input for the search, the number of features that were selected in each data layer, and which data layers did not return any features (Fig. 3.6).

```

DataExtractor_Sony.log - Notepad
File Edit Format View Help
-----
01/08/2016 13:50:12 : -----
01/08/2016 13:50:12 : Process started!
01/08/2016 13:50:12 : -----
01/08/2016 13:50:12 : -----
01/08/2016 13:50:12 : Processing partner 'Partner 1 (P1)' ...
01/08/2016 13:50:12 : Creating sub-directory 'H:\ExtractorSetup\Extracts\P1'
01/08/2016 13:50:12 : Starting process 1 of 12 ...
01/08/2016 13:50:12 : Processing table = 'Ancientwoodland' ...
01/08/2016 13:50:12 : Executing spatial selection ...
01/08/2016 13:50:16 : 317 records selected.
01/08/2016 13:50:16 : Saving local table ...
01/08/2016 13:50:16 : Translating to a SHP file ...
01/08/2016 13:50:16 : Switching to Universal Translator directory ... C:\Program Files (x86)\MapInfo\Professional\UT
01/08/2016 13:50:16 : Current MapInfo version is ... 1150
01/08/2016 13:50:16 : Creating Universal Translator mapping file ...
01/08/2016 13:50:17 : Executing Universal Translator ...
01/08/2016 13:50:19 : Translation to a SHP file complete.
01/08/2016 13:50:19 : Dropping the MapInfo table ...
01/08/2016 13:50:19 : Completed process 1 of 12.
01/08/2016 13:50:19 : -----
01/08/2016 13:50:19 : Zipping the extract files ...
01/08/2016 13:50:21 : Moving the zip file ...
01/08/2016 13:50:21 : Processing for partner complete.
01/08/2016 13:50:21 : -----
01/08/2016 13:50:21 : -----
01/08/2016 13:50:21 : Processing partner 'Partner 10 (P10)' ...
01/08/2016 13:50:21 : Creating sub-directory 'H:\ExtractorSetup\Extracts\P10'
01/08/2016 13:50:21 : Starting process 2 of 12 ...
01/08/2016 13:50:21 : Processing table = 'Ancientwoodland' ...
01/08/2016 13:50:21 : Executing spatial selection ...
01/08/2016 13:50:23 : 209 records selected.
01/08/2016 13:50:23 : Saving local table ...
01/08/2016 13:50:24 : Closing the MapInfo table ...
01/08/2016 13:50:24 : Completed process 2 of 12.
01/08/2016 13:50:24 : -----
01/08/2016 13:50:24 : Zipping the extract files ...
01/08/2016 13:50:26 : Moving the zip file ...
01/08/2016 13:50:26 : Processing for partner complete.
01/08/2016 13:50:26 : -----
01/08/2016 13:50:26 : -----
01/08/2016 13:50:26 : Processing partner 'Partner 11 (P11)' ...
01/08/2016 13:50:26 : Creating sub-directory 'H:\ExtractorSetup\Extracts\P11'
01/08/2016 13:50:26 : Starting process 3 of 12 ...
01/08/2016 13:50:26 : Processing table = 'Ancientwoodland' ...
01/08/2016 13:50:26 : Executing spatial selection ...
01/08/2016 13:50:26 : 84 records selected.
01/08/2016 13:50:26 : Saving local table ...
01/08/2016 13:50:26 : Closing the MapInfo table ...
01/08/2016 13:50:26 : Completed process 3 of 12.
01/08/2016 13:50:26 : -----
01/08/2016 13:50:26 : Zipping the extract files ...
01/08/2016 13:50:28 : Moving the zip file ...

```

Fig. 3.6: Example of a Data Extractor tool log file

The following chapters, *setting up the tool* (page 19) and *running the tool* (page 47), will guide you through setting up and operating the tool in such a way that these tool outputs meet the exact requirements of data extraction within your organisation.

Setting up the tool

Before the Data Extractor tool will function, it needs to be installed and configured. It is essential that the configuration is carried out first. There are some differences between the setup for MapInfo and ArcGIS, which are made clear below.

4.1 Configuring the tool

The configuration is stored in an XML file called 'DataExtractor.xml', an example of which can be found in the *Appendix* (page 57). Attributes and settings are presented as nodes (beginning with a start node, e.g. <example>, and finishing with an end node, e.g. <\example>), with the value for the setting held between the <value> and <\value> tag.

Caution: The name of the configuration file must be 'DataExtractor.xml'. The tool will not load if a different name is used.

The XML file can be edited in a text editor such as Notepad or Wordpad, or using a more feature rich XML editor such as *Sublime Text*¹². The configuration file is split into three sections:

General attributes General and default attributes for the tool.

SQL Tables Deals with how extracts from each SQL Server table should be handled.

Map Tables (MapInfo) or Map Layers (ArcGIS) Deals with how extracts from each GIS layer should be handled.

Caution: It is important that the structure of the file is maintained as it is presented in the *Appendix* (page 57). Any changes to the structure may result in the Data Extractor tool not loading, or not working as expected.

¹² <https://www.sublimetext.com/3>

Once editing has been completed and the edits have been saved, it is recommended that the configuration file is opened using an internet browser such as Internet Explorer which will help highlight any editing errors – only if the structure of the file is valid will the whole file be displayed in the Internet browser.

Note: It is recommended that the configuration file is kept in a central (network) location, so that all users use the same configuration. Additionally, it is essential that the configuration file is kept in the same folder as the compiled version of the tool.

4.1.1 Special characters in XML

The characters `&`, `<` and `>` are not valid within values and, so in order to be used, must be **escaped** with XML entities as follows:

< This must be escaped with `<` entity, since it is assumed to be the beginning of a tag. For example,
`RecYear < 2010`

> This should be escaped with `>` entity. It is not mandatory – it depends on the context – but it is strongly advised to escape it. For example, `RecYear > 1980`

& This must be escaped with `&` entity, since it is assumed to be the beginning of a entity reference. For example, `TaxonGroup = 'Invertebrates - Dragonflies & Damselflies'`

4.1.2 Setup for MapInfo

General attributes

The first section of the configuration file deals with a series of general attributes for the Data Extractor tool. Each node specifies where files are kept, how output files should be named, where the log file will be saved as well as other overall settings. Details on these attributes (and their typical values where known) are outlined below. The list follows the order within which the attributes are found in the configuration file. This version of the configuration details is valid for the MapInfo version 1.5.11 of the Data Extractor tool.

ToolTitle The title to use for the program in the MapInfo Tools menu.

LogFilePath The folder to be used for storing log files. This folder must already exist.

FileDSN The location of the file DSN which specifies the details of the connection to the SQL database.

DefaultPath The folder below which all partner folders will be created, and where extracts will be stored.

DatabaseSchema The schema used in the SQL database (typically `dbo`).

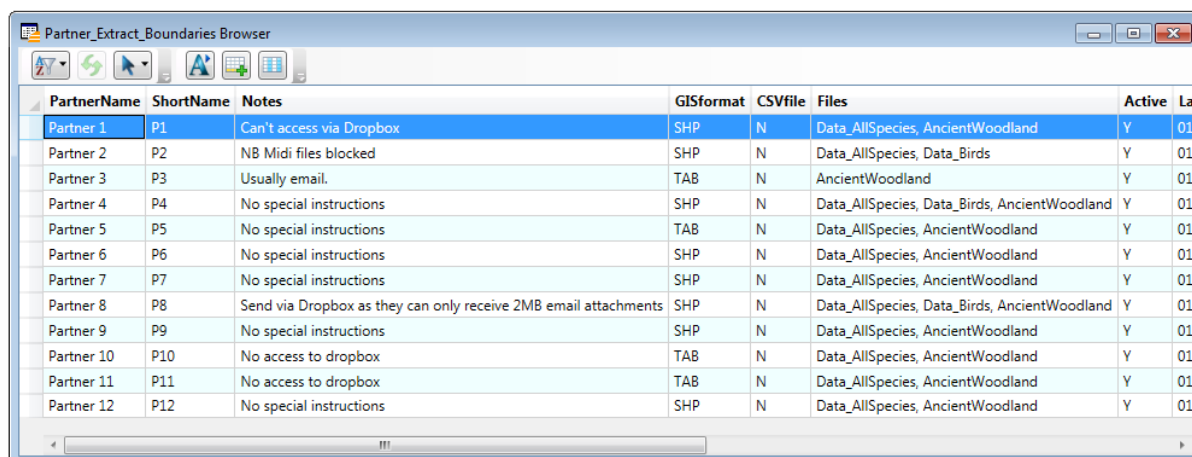
TableListSQL The SQL statement that is used to return the list of SQL tables which should be included in the user interface for selection by the user.

PartnerTable The name of the partner GIS layer (and SQL Server table) used to select records. The tool expects this layer to be present in the active MapInfo workspace and already present in the SQL Server database. A snapshot of a partner table is shown in [Fig. 4.1](#).

PartnerColumn The column in the *PartnerTable* (page 21) containing the partner name, which is passed to SQL Server by the tool to use the partner's boundary for selecting the records.

ShortColumn The name of the column in the partner GIS layer containing the abbreviated name to use as the sub-folder name for the destination of extracted records. The sub-folder is created in the *DefaultPath* (page 21) during extraction if it does not already exist.

NotesColumn The name of the column in the partner GIS layer containing any notes text relating to the partner.



PartnerName	ShortName	Notes	GISformat	CSVfile	Files	Active	La
Partner 1	P1	Can't access via Dropbox	SHP	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 2	P2	NB Midi files blocked	SHP	N	Data_AllSpecies, Data_Birds	Y	01
Partner 3	P3	Usually email.	TAB	N	AncientWoodland	Y	01
Partner 4	P4	No special instructions	SHP	N	Data_AllSpecies, Data_Birds, AncientWoodland	Y	01
Partner 5	P5	No special instructions	TAB	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 6	P6	No special instructions	SHP	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 7	P7	No special instructions	SHP	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 8	P8	Send via Dropbox as they can only receive 2MB email attachments	SHP	N	Data_AllSpecies, Data_Birds, AncientWoodland	Y	01
Partner 9	P9	No special instructions	SHP	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 10	P10	No access to dropbox	TAB	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 11	P11	No access to dropbox	TAB	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 12	P12	No special instructions	SHP	N	Data_AllSpecies, AncientWoodland	Y	01

Fig. 4.1: Example of a partner table loaded into MapInfo

Note: The partner GIS layer can be uploaded to SQL Server from MapInfo using the 'EasyLoader' tool.

Tip: Any notes for a partner can be displayed by 'double-clicking' the partner name in the list of partners in the tool interface.

ActiveColumn The name of the column in the partner GIS layer containing the Y/N flag to indicate if the partner is currently active. Only active partners will appear in the tool interface and be available for processing. The values in this column should be Y or N.

FormatColumn The name of the column in the partner GIS layer containing the GIS format required for the output records. The values in the column should be Shp, Tab or Both.

ExportColumn The name of the column in the partner GIS layer indicating whether an export should also be created as a CSV file. The values in this column should be Y or N.

FilesColumn The name of the column in the partner GIS layer indicating which SQL tables and map layers should be extracted for each partner. The entry in this column should be a comma-delimited list of the names of the layers (as defined in the XML file under *SQLTables* (page 23) and *MapTables* (page 24)) that should be included for each partner.

TagsColumn The name of the column in the partner GIS layer indicating which survey tags, if any, should be included in the export. The survey tags should be a comma-delimited list.

SelectTypeOptions The option list for the selection types to be included in the 'Selection Type' drop-down box on the tool interface. This attribute should not be changed. The options are *Spatial Only* (records are purely selected on whether they are inside or outside the partner boundary), *Survey tags only* (records are purely selected on the survey tags included in the *TagsColumn* (page 22)), and *Spatial and Survey Tags*, where both a spatial intersection and any records with the relevant survey tags are included in the extraction.

Note: The 'Selection Type' option in the tool interface **only** relates to extracts from SQL tables and **not** to extracts from GIS layers (which are always spatial).

DefaultSelectType The selection type that should be shown by default in the 'Selection Type' drop-down list. This attribute is the index number of the selection type options in the drop-down list,

with 1 being the first option.

RecMax The maximum number of records that will be extracted in any one partner extract.

DefaultZip The default value for zipping the extract files. This attribute should be set to `Yes` or `No`.

ConfidentialClause The SQL criteria for excluding any confidential surveys. The clause is appended to any SQL criteria already defined against each file under *SQLTables* (page 23).

DefaultConfidential Yes/No attribute, defining whether the check box for ‘Extract confidential surveys?’ will be set to checked (`Yes`) or unchecked (`No`) when the form is opened.

Note: The ‘ConfidentialClause’ and ‘Extract confidential surveys?’ option in the tool interface **only** relates to extracts from SQL tables and **not** to extracts from GIS layers.

UTPath The path to the Universal Translator program. The path will usually be `C:\Program Files (x86)\MapInfo\Professional\UT` (64 bit operating system) or `C:\Program Files\MapInfo\Professional\UT` (32 bit operating system) but it is dependent on the location of the MapInfo installation directory.

UTCommand The command to run the Universal Translator program. Unless the program has been renamed, this will usually be `Imut.exe` (MapInfo 11.5 or earlier) or `Fme.exe` (MapInfo 12 or later).

SQL table attributes

While the spatial selection that the tool carries out is over the entirety of the SQL table selected by the user, subsets of this data can be written out using the SQL table attributes. The details of these subsets are defined in the `<SQLTables>` node.

For each subset that may be included in the extracts a new child node must be created. The node name (e.g. `<AllSpecies>`) is a user-defined name used to identify an individual subset - the same name should be used in the *FilesColumn* (page 22) in the partner layer to indicate that this subset should be extracted for a partner. A simple example of an SQL layer definition with limited attributes is shown in Fig. 4.2.

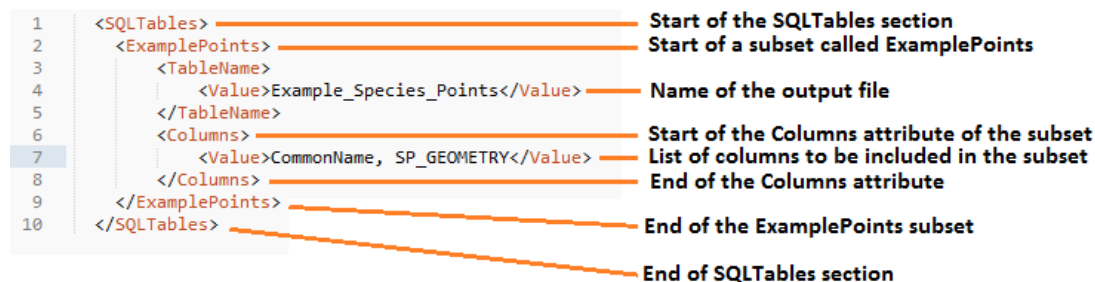


Fig. 4.2: Simplified example of an SQL table subset configuration

The attributes that are required for each SQL table are as follows:

TableName The name of the output GIS layer or text file that will be created for this subset.

Columns A comma-separated list of columns that should be included in the data exported for this subset during the extraction. The column names (not case sensitive) should match the column names in

the source table.

Clauses The SQL clause that should be used to select the data for this subset from the SQL table. This clause could, for example, ensure records are only included that have been entered after a certain date, are verified, are presence (not absence) records, or are a subset for particular taxon groups or protected species. Leave this entry blank to export the entire SQL table.

Note: Clauses specified here must adhere to SQL Server syntax as the clause will be run within SQL Server.

Symbology The symbology definition for this subset. Multiple symbols can be specified for use in the symbology using clauses. Each symbol is specified between `<Symbol>` and `</Symbol>` tags and is defined by the following child nodes:

Clause The clause that defines the records which will be assigned this symbol.

Object The object type that is symbolised using this symbol (e.g. `Point`)

Type The type of symbol to be used, usually 'Symbol'

Style The style of the symbol to be used.

Tip: In order to find the syntax for the Style attribute, set the desired symbol through Options => Symbol style, then write the following statement in the MapBasic window and hit enter: `Print CurrentSymbol()`. Then the full symbol definition (e.g. `137,255,12,"MapInfo Miscellaneous",256,0`) can be used in this attribute.

Map table attributes

All map layer attributes are found within the `<MapTables>` node. For each data layer that can be included in the extractions a new child node must be created. The node name (e.g. `<SSSIs>`) is a user-defined name used to identify the layer - the same name should be used in the *FilesColumn* (page 22) in the partner layer to indicate that this layer should be extracted for a partner. The attributes that are required for each map layer are as follows:

TableName The name of the source GIS layer as it is known in the active MapInfo workspace. This is also the name that will be used for the extracted file.

Columns A comma-separated list of columns that should be included in the data exported from this GIS layer during the extraction. The column names (not case sensitive) should match the column names in the source GIS layer.

Clause The SQL clause that should be used to select the data for this layer from the source GIS layer. Leave this entry blank to export the entire source GIS layer.

Note: Any clause specified here must adhere to MapInfo SQL syntax as the clause will be run within MapInfo.

Any exports from map layers will use the same symbology as the source layer.

4.1.3 Setting up the tool for ArcGIS

General attributes

The first section of the configuration file deals with a series of general attributes for the Data Extractor tool. Each node specifies where files are kept, how output files should be named, where the log file will be saved as well as other overall settings. Details on these attributes (and their typical values where known) are outlined below. The list follows the order within which the attributes are found in the configuration file. This version of the configuration details is valid for the ArcGIS version 1.0 of the Data Extractor tool.

LogFilePath The folder to be used for storing log files. This folder must already exist.

FileDSN The location of the file DSN which specifies the details of the connection to the SQL database.

ConnectionString In addition to a file DSN, the ArcGIS tool requires a connection string for the SQL database.

TimeoutSeconds The number of seconds before the connection to the database times out. If left blank this will default to 4,000 seconds.

DefaultPath The folder below which all partner folders will be created, and where extracts will be stored.

DatabaseSchema The schema used in the SQL database (typically dbo).

IncludeWildcard The wildcard for table names to list all the species tables in SQL Server that can be selected by the user to extract from. This might look like *LERC_Spp_*

ExcludeWildcard The wildcard for table names that will be excluded from the list of species tables. Intended to exclude temporary tables, this might take the form LERC_Spp_*.*

PartnerTable The name of the partner GIS layer (and SQL Server table) used to select records. The tool expects this layer to be present in the ArcMap Table of Contents and also present in the SQL Server database. A snapshot of a partner table is shown in Fig. 4.3.

PartnerName	ShortName	Notes	GISformat	CSVfile	Files	Active	La
Partner 1	P1	Can't access via Dropbox	SHP	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 2	P2	NB Midi files blocked	SHP	N	Data_AllSpecies, Data_Birds	Y	01
Partner 3	P3	Usually email.	TAB	N	AncientWoodland	Y	01
Partner 4	P4	No special instructions	SHP	N	Data_AllSpecies, Data_Birds, AncientWoodland	Y	01
Partner 5	P5	No special instructions	TAB	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 6	P6	No special instructions	SHP	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 7	P7	No special instructions	SHP	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 8	P8	Send via Dropbox as they can only receive 2MB email attachments	SHP	N	Data_AllSpecies, Data_Birds, AncientWoodland	Y	01
Partner 9	P9	No special instructions	SHP	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 10	P10	No access to dropbox	TAB	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 11	P11	No access to dropbox	TAB	N	Data_AllSpecies, AncientWoodland	Y	01
Partner 12	P12	No special instructions	SHP	N	Data_AllSpecies, AncientWoodland	Y	01

Fig. 4.3: Example of a partner table

Note: The partner GIS layer can be uploaded to SQL Server by right-clicking on the layer, then selecting Data => Export Data. In the resulting menu choose Database Feature Classes as the file type, and use the FileDSN as the location to save the data to.

PartnerColumn The column in the PartnerTable containing the partner name, which is passed to SQL Server by the tool to use the partner's boundary for selecting the records.

ShortColumn The name of the column in the partner GIS layer containing the abbreviated name to use as the sub-folder name for the destination of extracted records. The sub-folder is created in the DefaultPath during extraction if it does not already exist.

NotesColumn The name of the column in the partner GIS layer containing any notes text relating to the partner.

Tip: Any notes for a partner can be displayed by 'double-clicking' the partner name in the list of partners in the tool interface.

ActiveColumn The name of the column in the partner GIS layer containing the Y/N flag to indicate if the partner is currently active. Only active partners will appear in the tool interface and be available for processing. The values in this column should be Y or N.

FormatColumn The name of the column in the partner GIS layer containing the GIS format required for the output records. The values in the column should be SHP or GDB.

ExportColumn The name of the column in the partner GIS layer indicating whether an export should also be created as a CSV file. The values in this column should be CSV or TXT. If it is left blank no text export will be generated.

SQLFilesColumn The name of the column in the partner GIS layer indicating which SQL tables should be extracted for each partner. The entry in this column should be a comma-delimited list of the names of the layers (as defined in the XML file under *SQLTables* (page 27)) that should be included for each partner.

MapFilesColumn The name of the column in the partner GIS layer indicating which ArcGIS layers should be extracted for each partner. The entry in this column should be a comma-delimited list of the names of the layers (as defined in the XML file under *MapLayers* (page 28)) that should be included for each partner.

TagsColumn The name of the column in the partner GIS layer indicating which survey tags, if any, should be included in the export. The survey tags should be a comma-delimited list.

SelectTypeOptions The option list for the selection types to be included in the 'Selection Type' drop-down box on the tool interface. This attribute should not be changed. The options are *Spatial Only* (records are purely selected on whether they are inside or outside the partner boundary), *Survey tags only* (records are purely selected on the survey tags included in the TagsColumn), and *Spatial and Survey Tags*, where both a spatial intersection and any records with the relevant survey tags are included in the extraction.

Note: The 'Selection Type' option in the tool interface **only** relates to extracts from SQL tables and **not** to extracts from GIS layers (which are always spatial).

DefaultSelectType The selection type that should be shown by default in the 'Selection Type' drop-down list. This attribute is the index number of the selection type options in the drop-down list, with 1 being the first option.

DefaultZip The default value for zipping the extract files. This attribute is not currently used in ArcGIS.

ConfidentialClause The SQL criteria for excluding any confidential surveys. The clause is appended to any SQL criteria already defined against each file under *SQLTables* (page 27).

DefaultConfidential Yes/No attribute, defining whether the check box for ‘Extract confidential surveys?’ will be set to checked (Yes) or unchecked (No) when the form is opened.

Note: The ‘ConfidentialClause’ and ‘Extract confidential surveys?’ option in the tool interface **only** relates to extracts from SQL tables and **not** to extracts from GIS layers.

SQL table attributes

While the spatial selection that the tool carries out is over the entirety of the SQL table selected by the user, subsets of this data can be written out using the SQL table attributes. The details of these subsets are defined in the `<SQLTables>` node.

For each subset that may be included in the extracts a new child node must be created. The node name (e.g. `<AllSpecies>`) is a user-defined name used to identify an individual subset - the same name should be used in the *FilesColumn* (page 22) in the partner layer to indicate that this subset should be extracted for a partner. A simple example of an SQL layer definition with limited attributes is shown in Fig. 4.4.

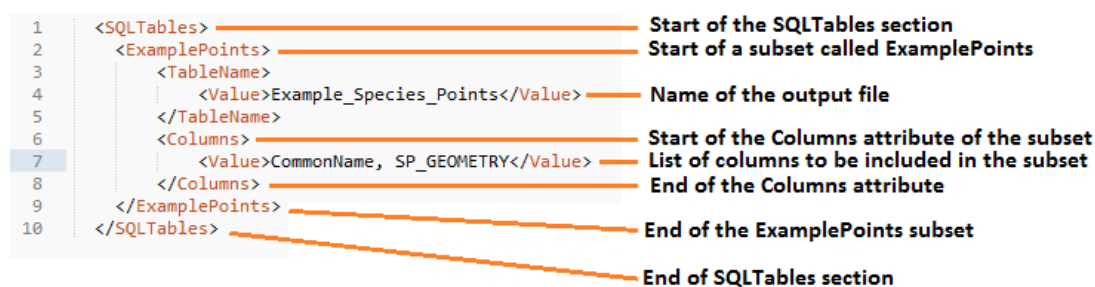


Fig. 4.4: Simplified example of an SQL table subset configuration

The attributes that are required for each SQL table are as follows:

TableName The name of the output GIS layer or text file that will be created for this subset.

Columns A comma-separated list of columns that should be included in the data exported for this subset during the extraction. The column names (not case sensitive) should match the column names in the source table.

Clauses The SQL clause that should be used to select the data for this subset from the SQL table. This clause could, for example, ensure records are only included that have been entered after a certain date, are verified, are presence (not absence) records, or are a subset for particular taxon groups or protected species. Leave this entry blank to export the entire SQL table.

Note: Clauses specified here must adhere to SQL Server syntax as the clause will be run within SQL Server.

Map layer attributes

All map layer attributes are found within the `<MapLayers>` node. For each data layer that can be included in the extractions a new child node must be created. The node name (e.g. `<SSSIs>`) is a user-defined name used to identify the layer - the same name should be used in the *FilesColumn* (page 22) in the partner layer to indicate that this layer should be extracted for a partner. The attributes that are required for each map layer are as follows:

LayerName The name of the source GIS layer as it is known in the ArcGIS Table of Contents. This is also the name that will be used for the extracted file.

Columns A comma-separated list of columns that should be included in the data exported from this GIS layer during the extraction. The column names (not case sensitive) should match the column names in the source GIS layer.

Clause The SQL clause that should be used to select the data for this layer from the source GIS layer. Leave this entry blank to export the entire source GIS layer.

Note: Any clause specified here must adhere to ArcGIS SQL syntax as the clause will be run within ArcGIS.

4.2 Setting up the SQL Server database

In addition to any SQL tables containing records to be extracted using the Data Extractor tool, two auxiliary tables must also be present in the SQL Server database in order for the tool to be able to extract data from tables held in SQL Server. These are as follows:

Survey table The Survey table is a standard table in the Recorder6 database. It is used to identify any records tagged with any survey tags listed in the *TagsColumn* (page 22) column in the partner GIS layer.

Spatial_Tables table This table contains information about any SQL data tables that may be used by the tool. The table has the following columns:

Table 4.1: Format of the Spatial_Tables table

Column	Description
TableName	The name of the data table
OwnerName	The database owner, usually dbo
XColumn	The name of the column holding the X coordinates of the record
YColumn	The name of the column holding the Y coordinates of the record
SizeColumn	The name of the column holding the grid size of the record (in metres)
IsSpatial	Bitwise column (1 = Yes, 0 = No) defining whether the table is spatially enabled
SpatialColumn	If the table is spatially enabled, the name of the geometry column (e.g. SP_GEOMETRY)
SRID	The name of the spatial reference system used to plot the records
CoordSystem	The coordinate system of the spatial data in the table
SurveyKeyColumn	The column containing the survey key for each record

Note: The British National Grid *SRID* value is Earth Projection 8, 79, "m", -2, 49, 0.9996012717, 400000, -100000 Bounds (-7845061.1011, -15524202.1641) (8645061.1011, 4470074.53373)

Caution: This table must be filled out correctly for each table that is included in the Data Extractor tool.

Note: A number of stored procedures that are used by the tool for extracting the required records must also be present in the SQL Server database. To obtain copies of these procedures please contact [Hester¹³](mailto:Hester@HesterLyonsConsulting.co.uk) or [Andy¹⁴](mailto:Andy@AndyFoyConsulting.co.uk).

¹³ Hester@HesterLyonsConsulting.co.uk

¹⁴ Andy@AndyFoyConsulting.co.uk

4.3 Installing the tool

Installation in MapInfo and ArcGIS is different. Please refer to the relevant section.

4.3.1 Installing the tool in MapInfo

To install the tool in MapInfo, make sure that the configuration of the XML file as described above is complete, that the XML file is in the same directory as the tool MapBasic application (.MBX) and that all required GIS layers are loaded in the current workspace. Then, open *Tool Manager* in MapInfo by selecting *Tools --> Tool Manager...* in the menu bar (Fig. 4.5).

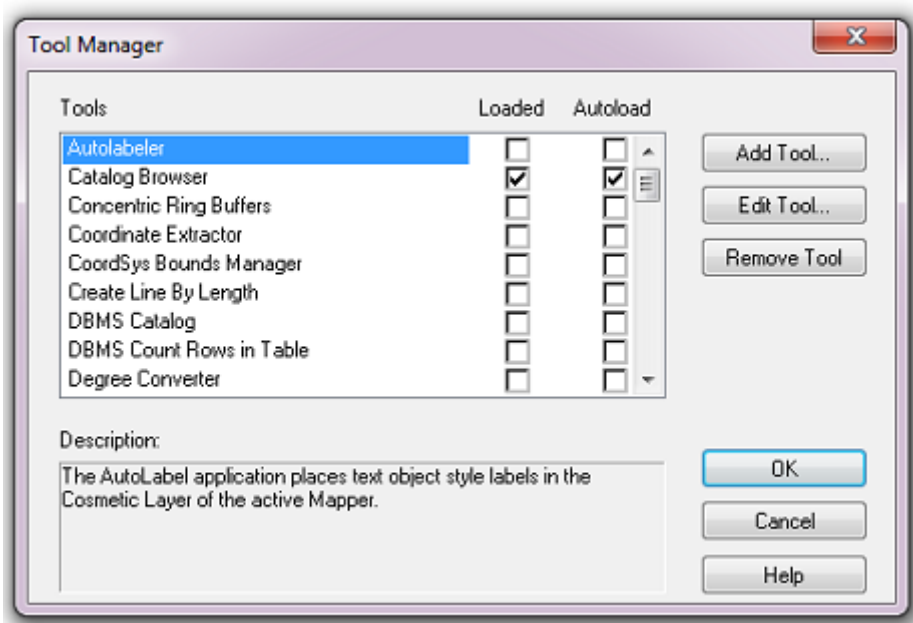


Fig. 4.5: The Tool Manager in MapInfo 12 or earlier

In the *Tool Manager* dialog, click **Add Tool...**, then locate the tool using the browse button ... on the *Add Tool* dialog (Fig. 4.6). Enter a name in the **Title** box (e.g. 'DataExtractor'), and a description if desired. Then click **Ok** to close the *Add Tool* dialog.

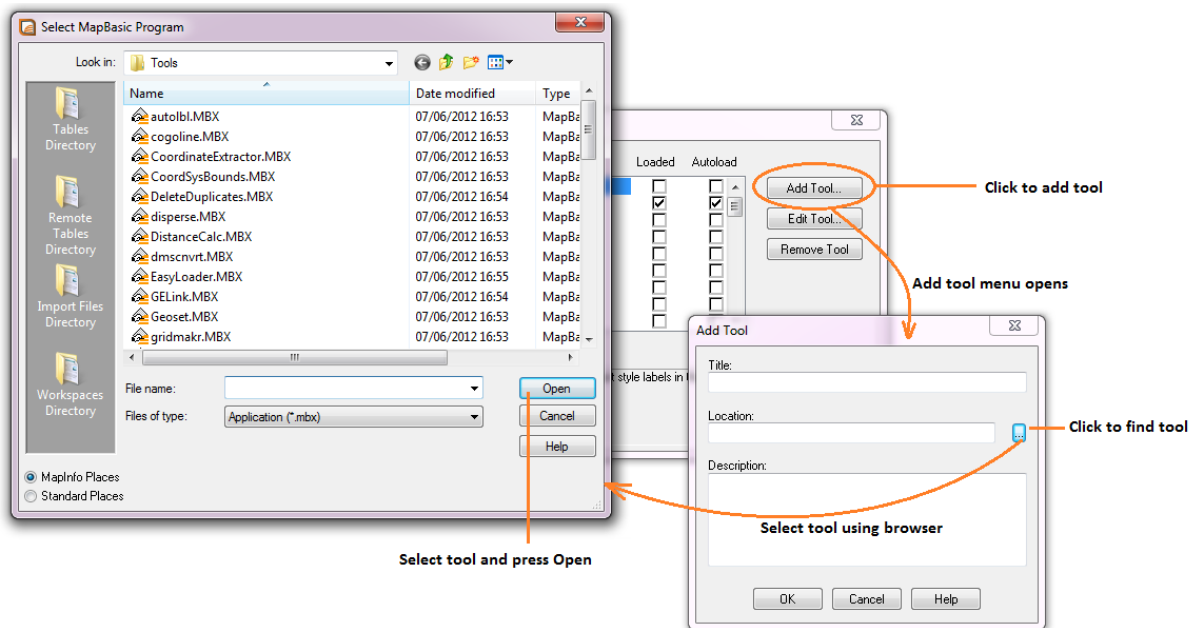


Fig. 4.6: Adding a tool in Tool Manager

The tool will now show in the *Tool Manager* dialog (Fig. 4.7) and the **Loaded** box will be checked. To load the tool automatically whenever MapInfo is started check the **AutoLoad** box. Then click **Ok** to close the *Tool Manager* dialog.

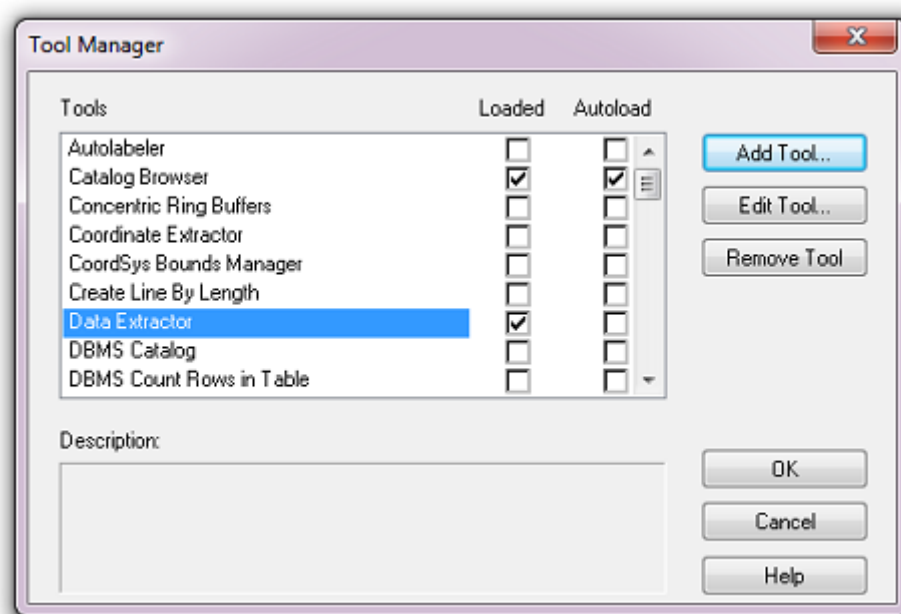


Fig. 4.7: The Data Extractor tool is loaded

The tool will now appear as a new entry in the *Tools* menu (Fig. 4.8).

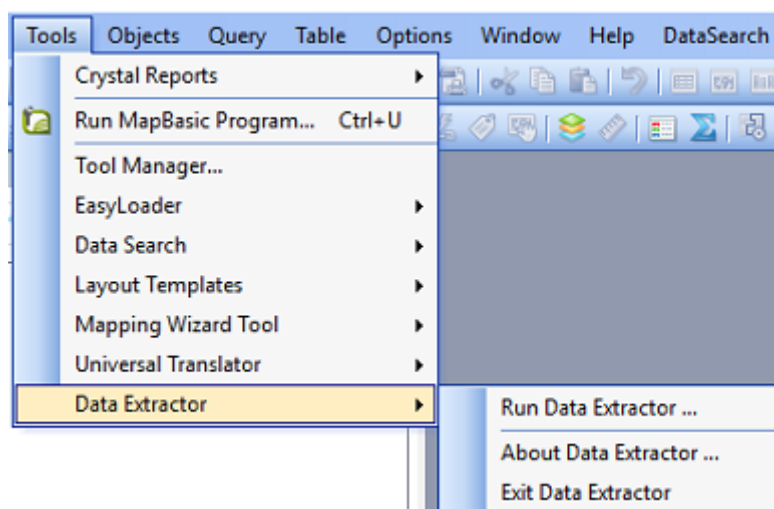


Fig. 4.8: The Data Extractor tool menu

Note: The name that will appear in the *Tools* menu is dependent on the *ToolTitle* (page 21) value in the configuration file, **not** the name given when adding the tool using the Tool Manager.

Tip: It is recommended that a MapInfo Workspace is created that contains all the required GIS layers to run the tool. Once this workspace has been set up and the tool has been configured and installed,

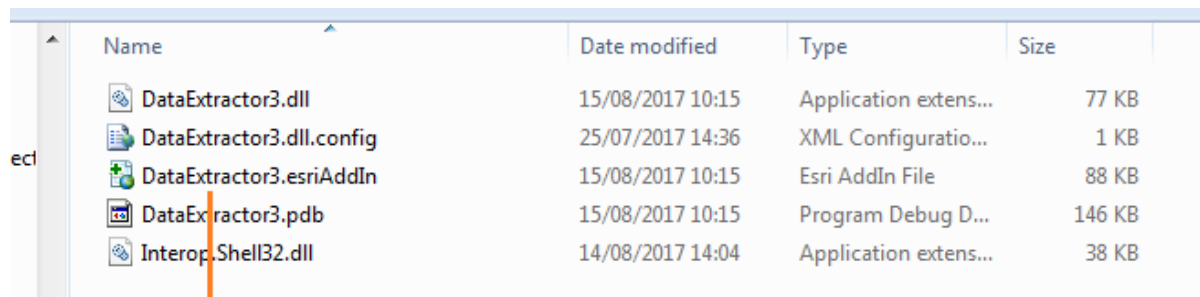
running the Data Extractor tool becomes a simple process.

4.3.2 Installing the tool in ArcGIS

Installing the tool in ArcGIS is straightforward. There are a few different ways it can be installed:

Installation through Windows Explorer

Open Windows Explorer and double-click on the ESRI Add-in file for the Data Extractor tool (Fig. 4.9).



The Data Extractor Add-in File. Double click to install.

Fig. 4.9: Installing the Data Extractor tool from Windows Explorer

Installation will begin after confirming you wish to install the tool on the dialog that appears (Fig. 4.10).

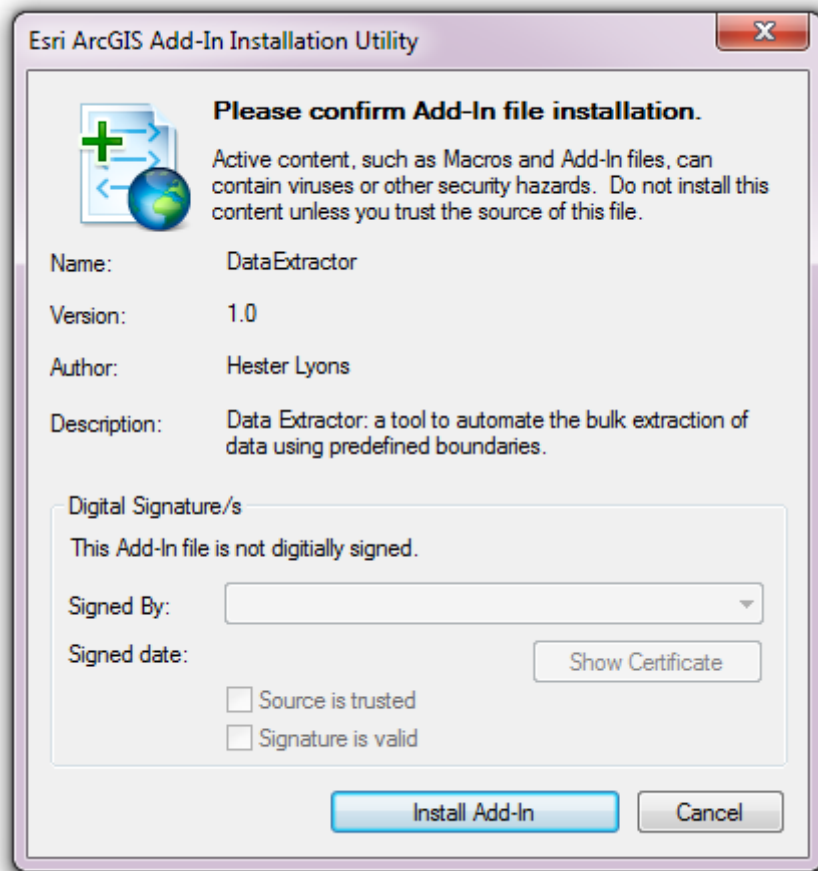


Fig. 4.10: Installation begins after clicking 'Install Add-in'

Once it is installed, it will become available to add to the ArcGIS interface as a button (see *Customising toolbars* (page 38)).

Note: In order for this process to work all running ArcMap sessions must be closed. The tool will not install or install incorrectly if there are copies of ArcMap running.

Installation from within ArcMap

Firstly, open the Add-In Manager through the Customize menu (Fig. 4.11).

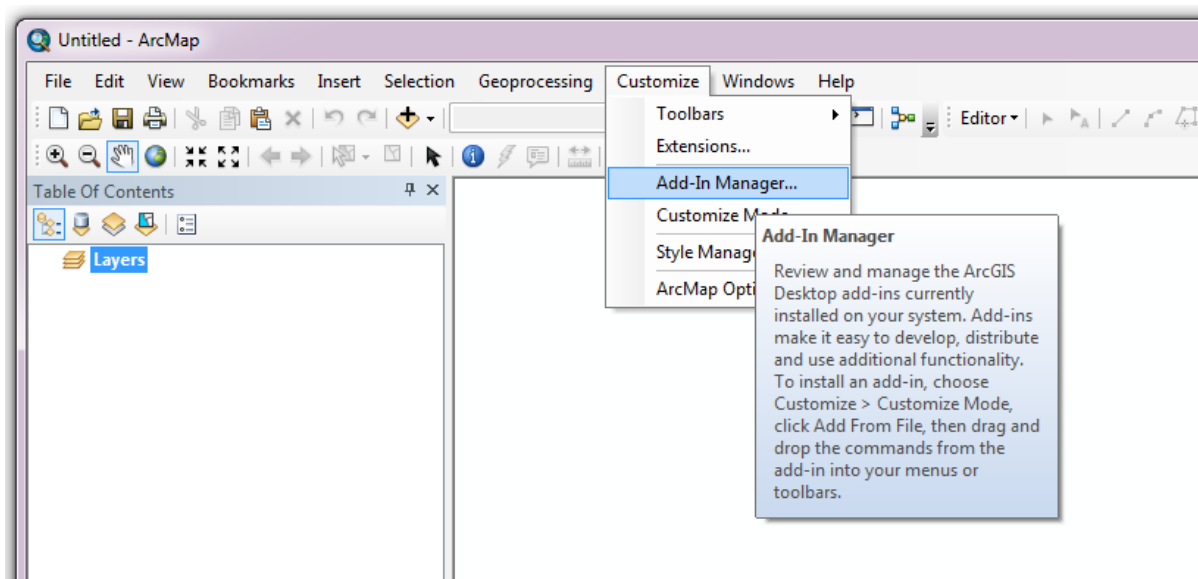


Fig. 4.11: Starting the ArcGIS Add-In Manager

If the Data Buffer tool is not shown, use the **Options** tab to add the folder where the tool is kept (Fig. 4.12). The security options should be set to the lowest setting as the tool is not digitally signed.

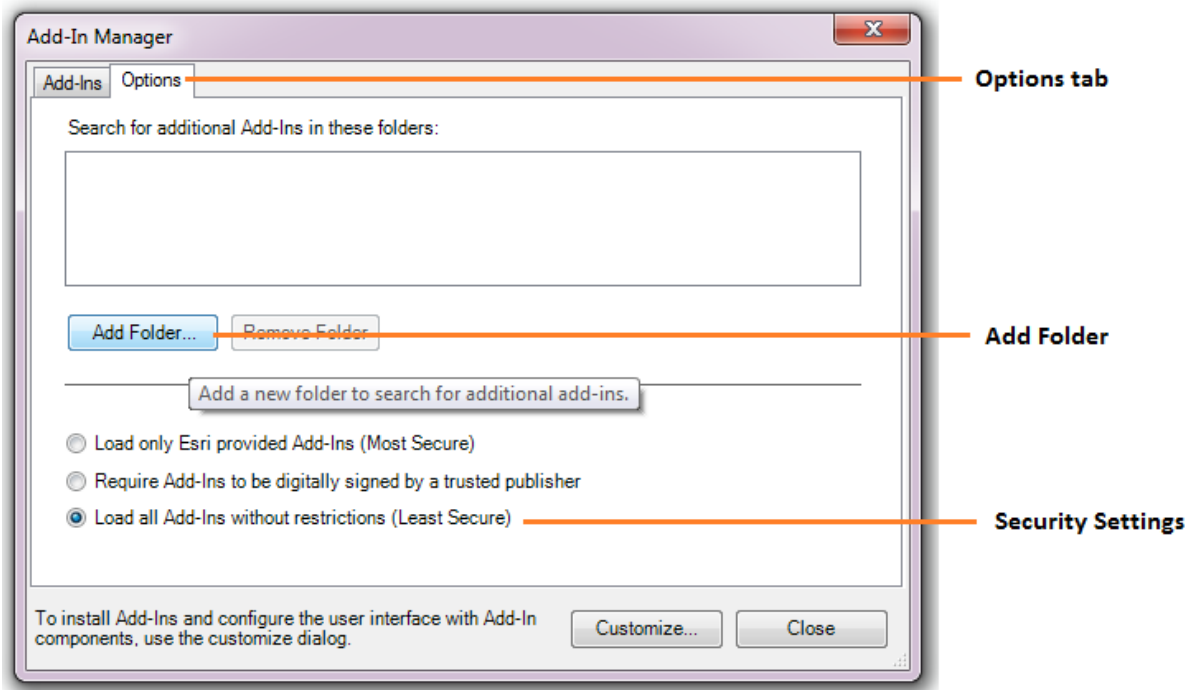


Fig. 4.12: The 'Options' tab in the ArcGIS Add-In Manager

Once the tool shows in the Add-In Manager (Fig. 4.13), it is available to add to the ArcGIS interface as a button (see *Customising toolbars* (page 38)).

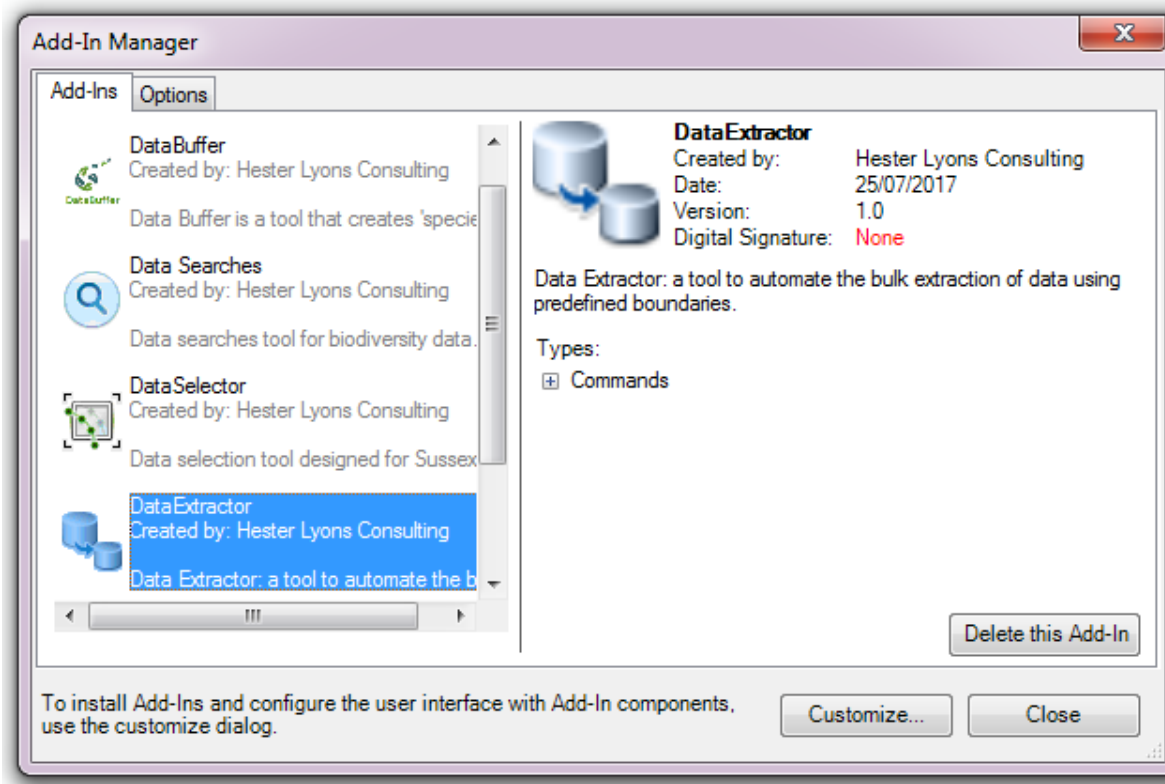


Fig. 4.13: The ArcGIS Add-In Manager showing the Data Extractor tool

Customising toolbars

In order to add the Data Buffer tool to the user interface, it needs to be added to a toolbar. It is recommended that this is done inside a document that has already been loaded with all the data layers that are required for the tool to run. The tool should then be saved with this document (see [Fundamentals of Saving your Customizations](#)¹⁵ for an explanation of how customisations are stored within ArcGIS).

Customising toolbars is done through the Customize dialog, which can be started either through the Add-In Manager (by clicking **Customize**, see Fig. 4.13), or through choosing the 'Customize Mode...' option in the Customize Menu (Fig. 4.14).

¹⁵ <http://desktop.arcgis.com/en/arcmap/10.3/guide-books/customizing-the-ui/fundamentals-of-saving-your-customizations.htm>

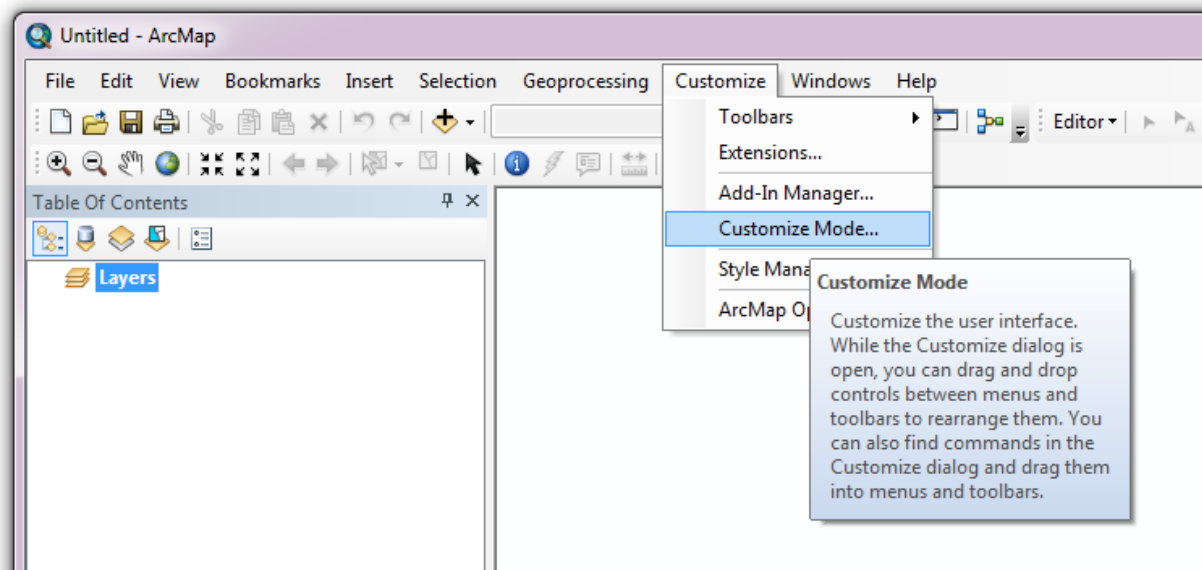


Fig. 4.14: Starting Customize Mode in ArcGIS

Once this dialog is open, ensure that the check box 'Create new toolbars and menus in the document' is checked in the **Options** tab (Fig. 4.15).

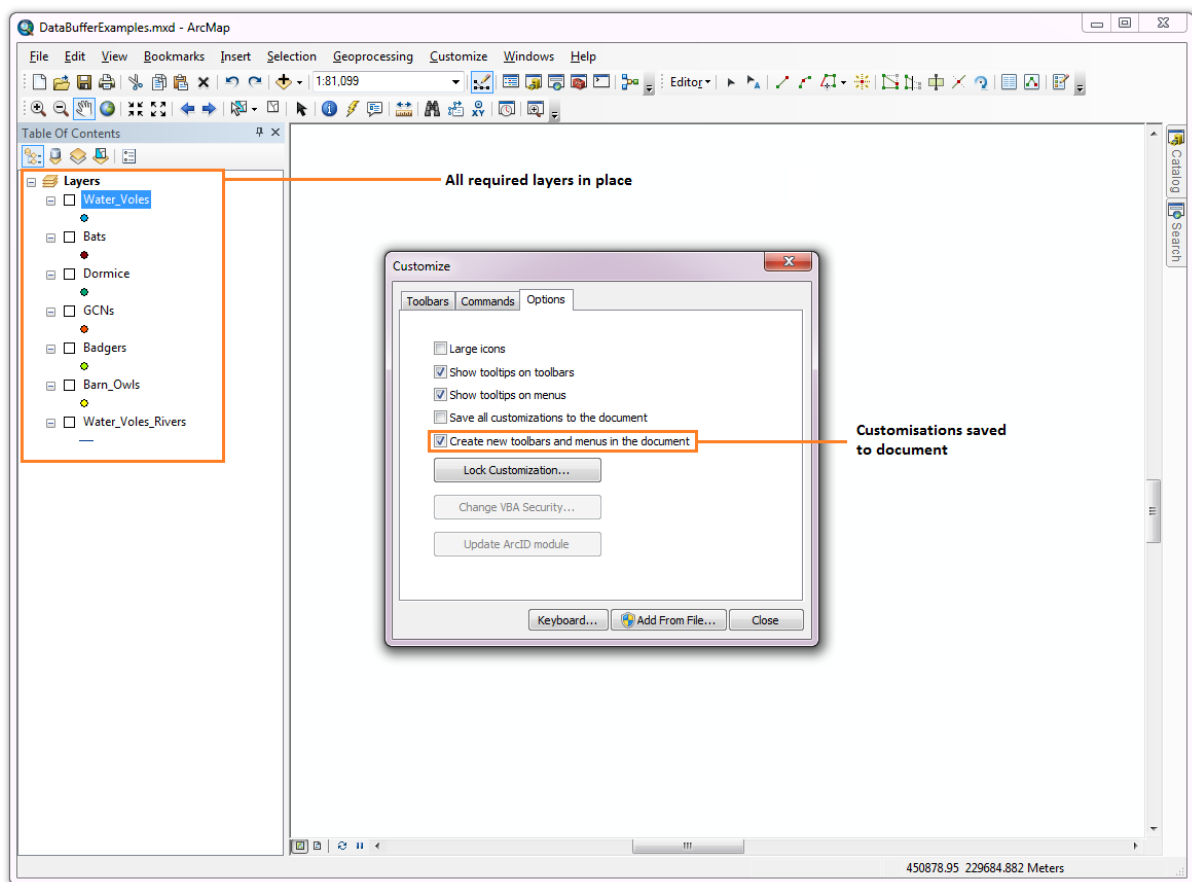


Fig. 4.15: Customising the document in ArcGIS

It is recommended that the button for the Data Extractor tool is added to a new toolbar. Toolbars are created through the **Toolbars** tab in the Customize dialog, as shown in figures Fig. 4.16 and Fig. 4.17.

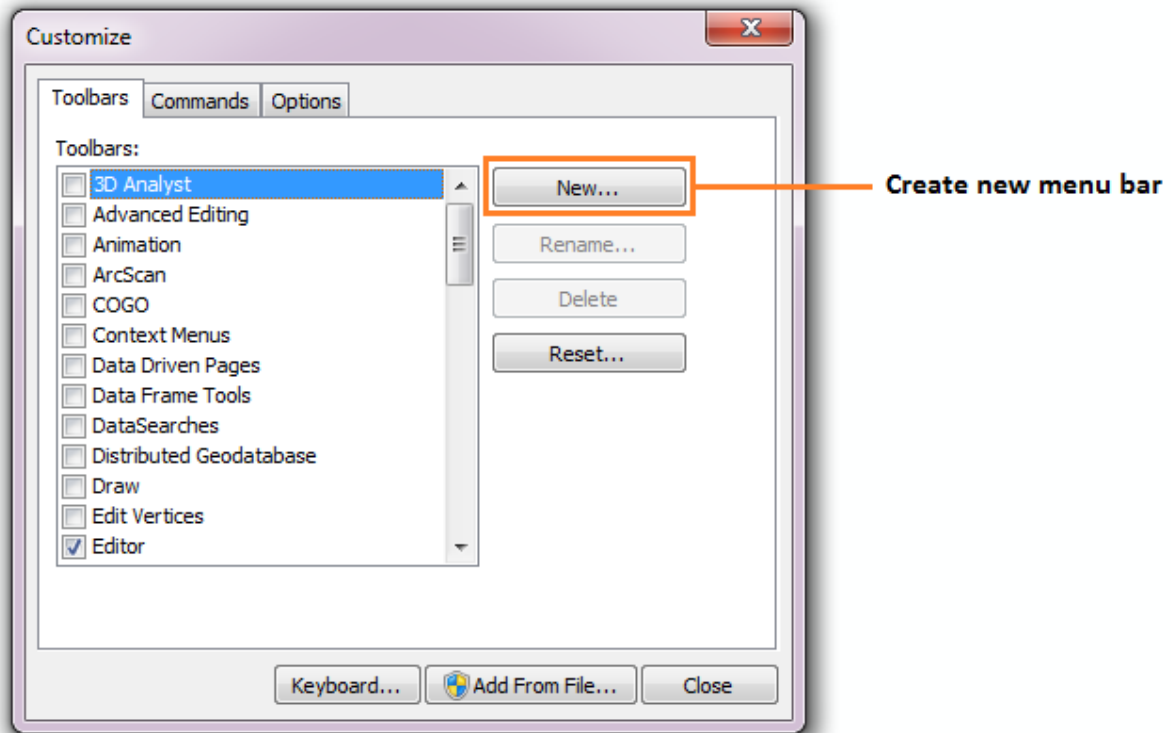


Fig. 4.16: Adding a new toolbar in ArcGIS

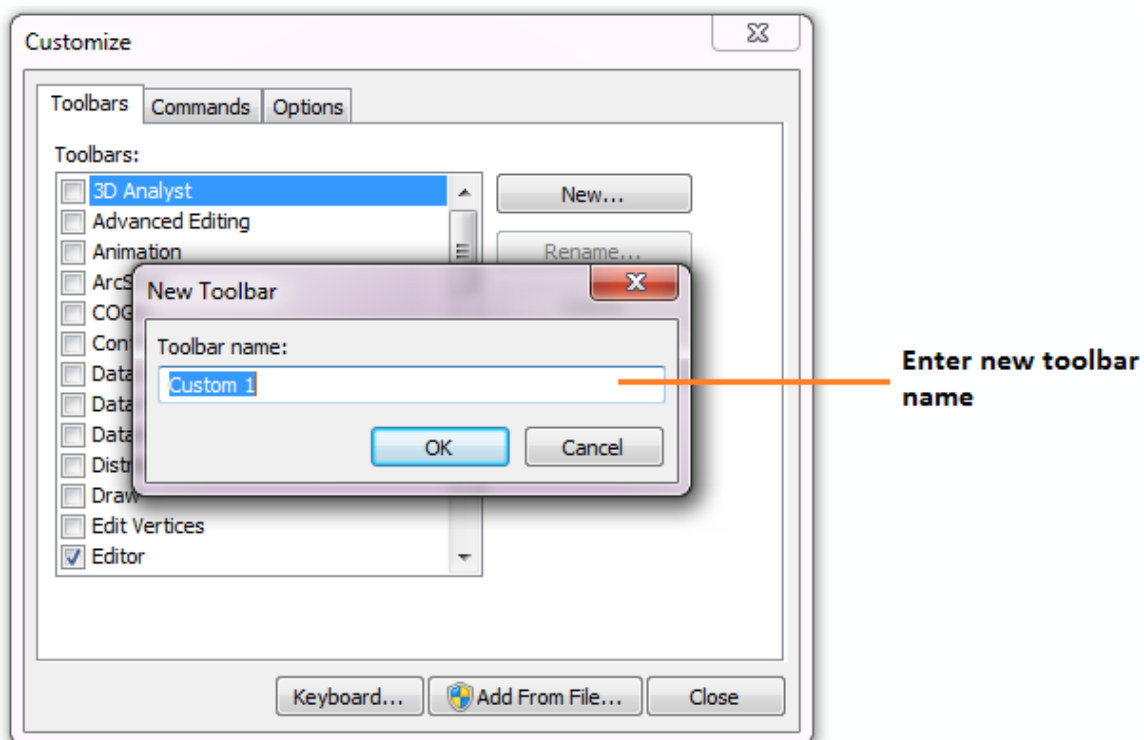


Fig. 4.17: Naming the new toolbar in ArcGIS

Once a new toolbar is created and named, it is automatically added to the ArcMap interface as well as to the Customize dialog (Fig. 4.18. In this case the toolbar was named 'TestToolbar').

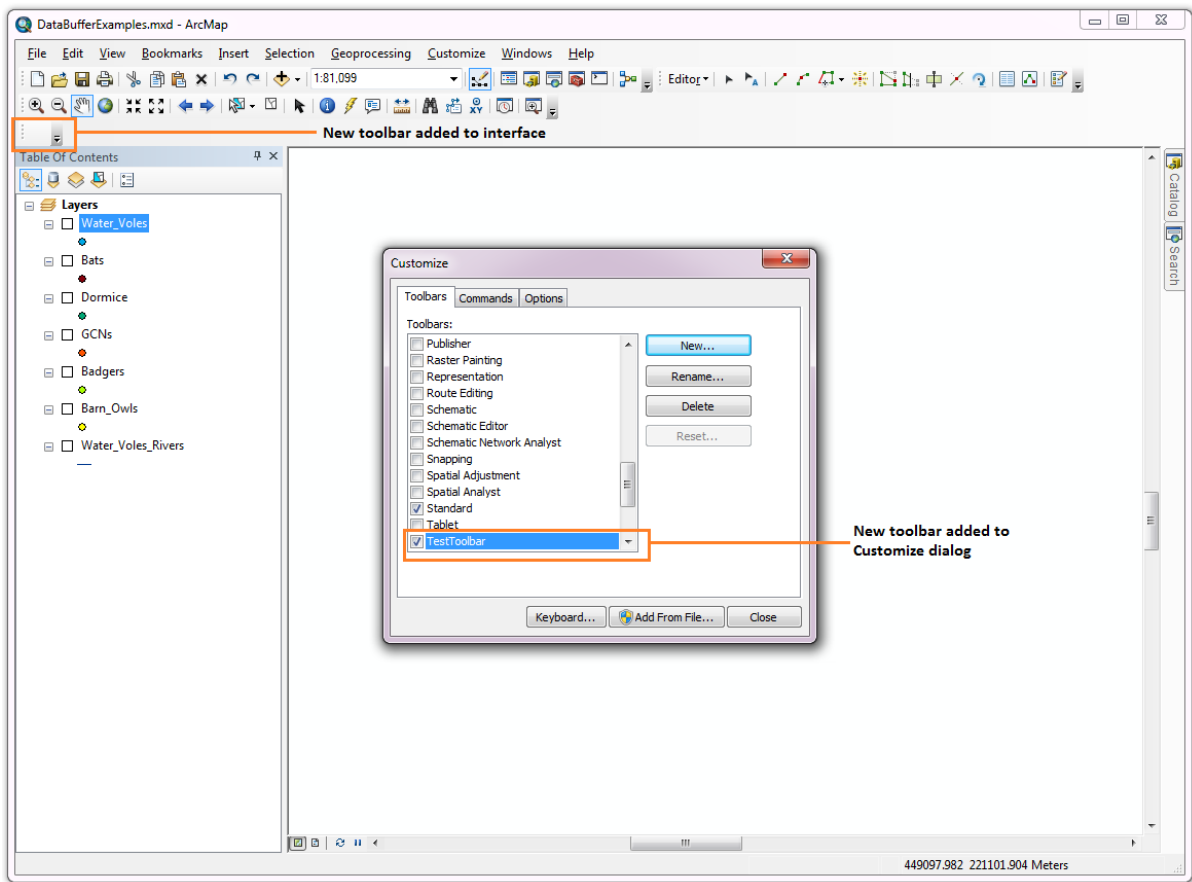


Fig. 4.18: New toolbar added to the ArcGIS Interface

As a final step the Data Extractor tool is added to the toolbar. This is done from the **Command** tab in the Customize dialog (Fig. 4.19). Click on **Add-In Controls** and the Data Extractor tool will be shown in the right-hand panel.

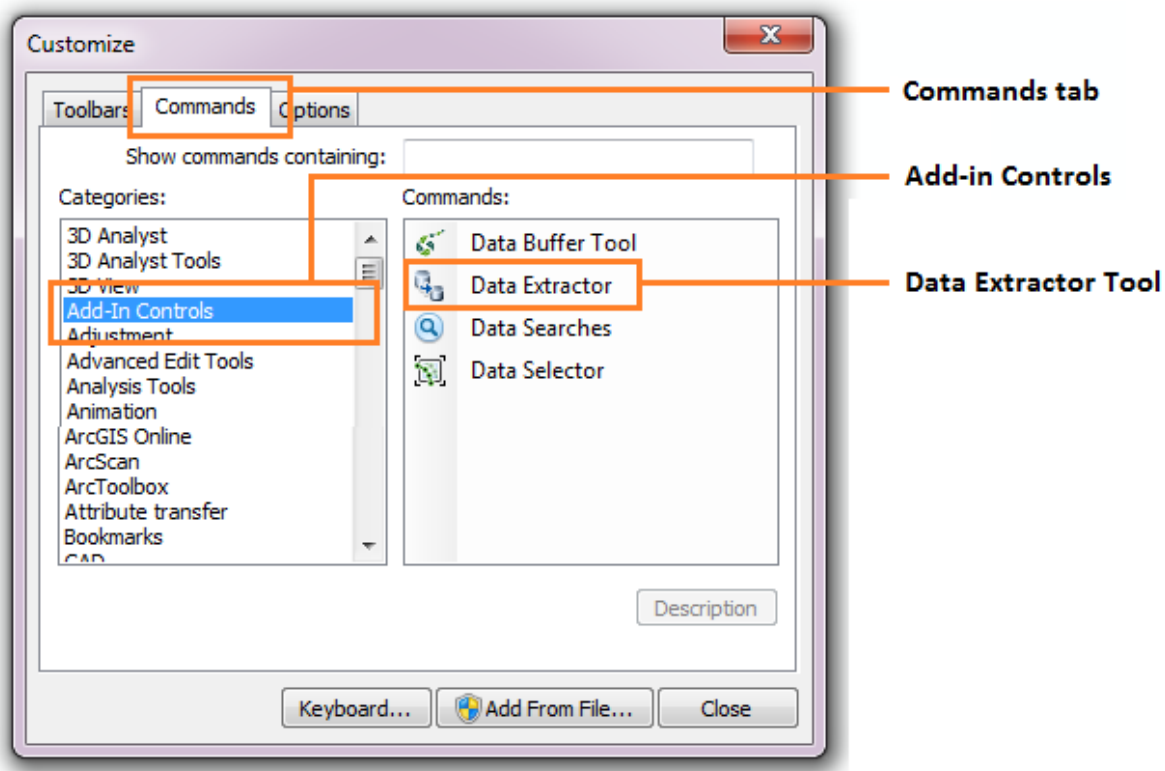


Fig. 4.19: Finding the Data Extractor tool in the add-in commands

To add the tool to the toolbar, simply drag and drop it onto it (Fig. 4.20). Close the Customize dialog and **save the document**. The Data Extractor tool is now ready for its final configuration and first use.

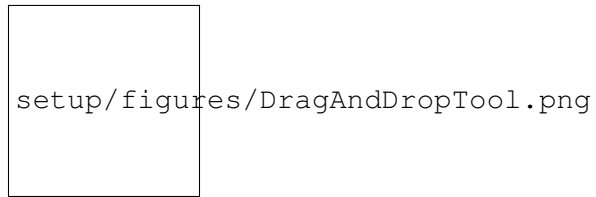


Fig. 4.20: Adding the Data Extractor tool to the new toolbar

In order to function, the tool needs to know the location of the XML configuration file. The first time the tool is run, or whenever the configuration file is moved, a dialog will appear asking for the folder containing the XML file (Fig. 4.21). Navigate to the folder where the XML file is kept and click **OK**. If the XML file is present and its structure is correct, the Data Searches form will be shown. Even if the tool is not run at this time, the location of the configuration file will be stored for future use.

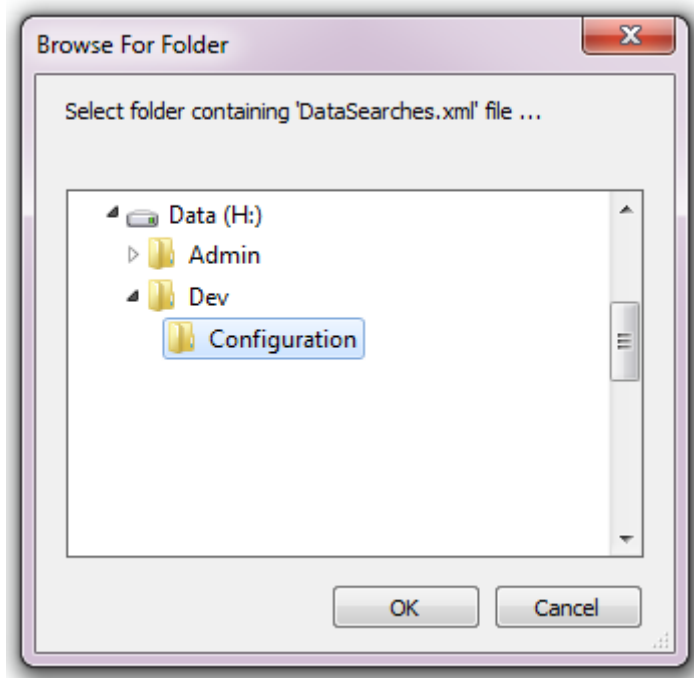


Fig. 4.21: Locating the configuration file folder

Running the tool

As discussed in the *Setting up the tool* (page 19) section, the Data Extractor tool is operated from a MapInfo workspace within which the GIS layers required to run the tool are already loaded. It also relies on an SQL database containing the boundaries of the relevant partners, the SQL tables containing any data that may be extracted, and a configuration document for setting up the tool. Therefore, before running the tool, ensure the following conditions are met:

- A MapInfo document has been created which contains both the partner boundary GIS layer and any MapInfo GIS layers that may be included in the extraction.
- An SQL database exists that contains the relevant partner boundaries, any data tables required, and the auxiliary tables and stored procedures as set out in the setup section.
- The file DSN for the SQL database has been created.
- The XML configuration document has been set up correctly, both for general settings and for each individual layer that will be queried. It is also named correctly and in the same directory as the tool MapBasic application (.MBX).
- The Data Extractor tool has been installed and is loaded in MapInfo using the MapInfo Tool Manager.

See also:

Please refer to the *setup* (page 19) section for further information about any of these requirements.

5.1 Opening the form

To open the Data Extractor tool, open the tool in the *Tools* menu (**Tools... -> Data Extractor**), as shown in Fig. 5.1.

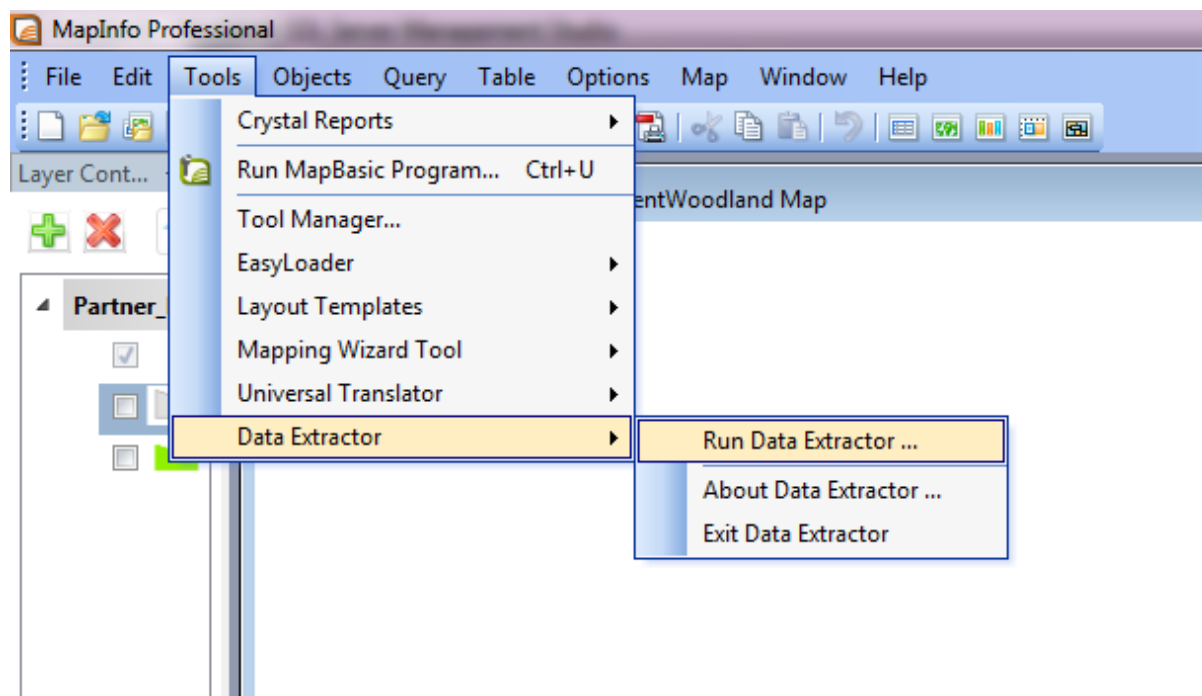


Fig. 5.1: Launching the Data Extractor tool

If there are any structural issues with the XML document, the tool will display a message that it has encountered an error, and not load any further. If any of the map layers that are listed in the configuration document are not present in the active workspace a warning will be shown (Fig. 5.2). The layers that are missing will not be loaded into the form and so cannot be included in the extracts.

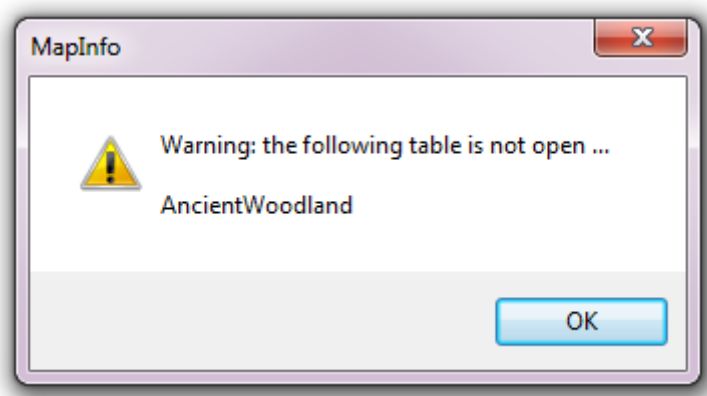


Fig. 5.2: Example warning message when any GIS layers are not found

Provided that the XML document is otherwise correct, the form will display (Fig. 5.3).

The screenshot shows a dialog box titled "Data Extractor 1.5.11". It has a standard Windows-style title bar with a close button (X) in the top right corner. The main content area is divided into several sections:

- Active Partners:** A list box containing 12 items: Partner 1, Partner 10, Partner 11, Partner 12, Partner 2, Partner 3, Partner 4, Partner 5, Partner 6, Partner 7, Partner 8, and Partner 9.
- SQL Tables:** A list box containing one item: AllData.
- MapInfo Tables:** A list box containing one item: AncientWoodland.
- Options:** Three checkboxes are present:
 - Zip extract files?
 - Extract confidential surveys?
 - Clear log file?
- Selection Type:** A dropdown menu currently showing "Spatial Only".
- Buttons:** "Cancel" and "Ok" buttons are located at the bottom right of the dialog.

Fig. 5.3: The form displaying the active partners and available data layers

5.2 Using the form

The form can be used to process as many partners and tables as required during a single execution. Select the partners for which you wish to run the extraction, and the SQL and / or MapInfo tables that you would like to include, what type of SQL table extraction you would like to carry out (spatial, survey tags, or both), make sure all other options are correct then press **OK**.

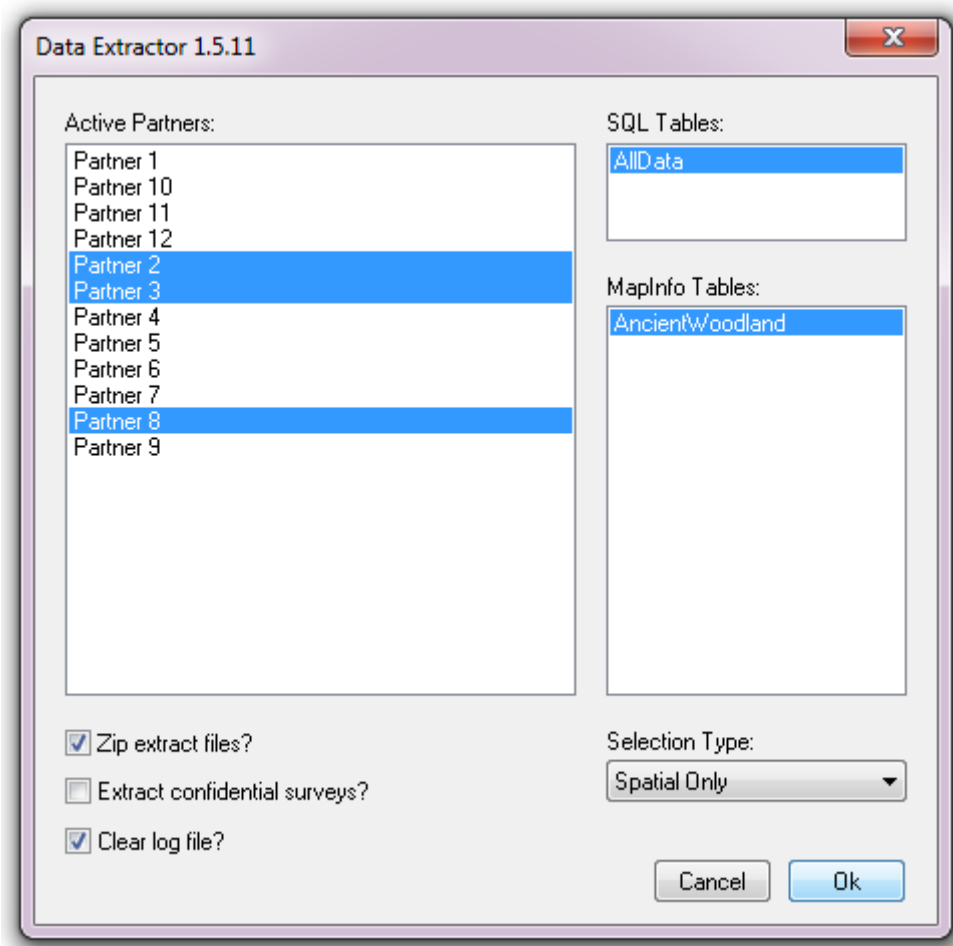


Fig. 5.4: Select the required options on the form

Caution: The tool can run for a considerable amount of time dependent on the number of records that are being selected.

Progress is shown in a progress window (Fig. 5.5).

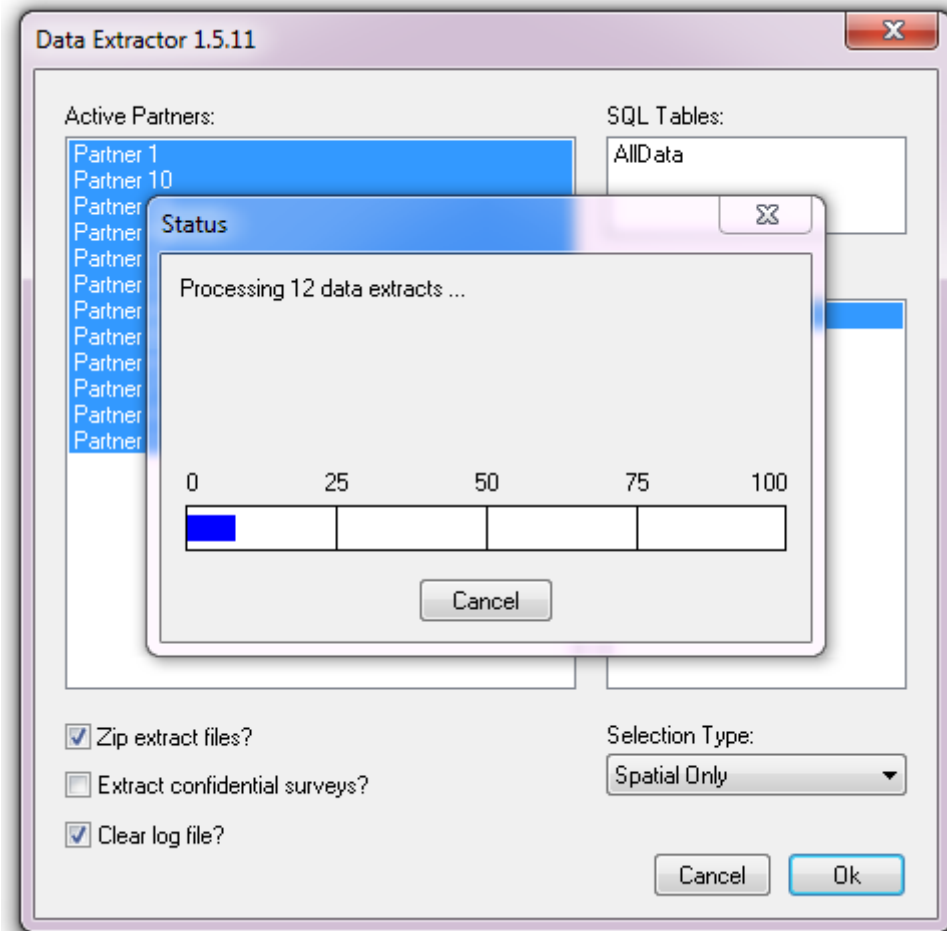


Fig. 5.5: Progress window during the extract process

When the extract process finishes it asks the user whether to close the form or keep it open for subsequent use (Fig. 5.6).

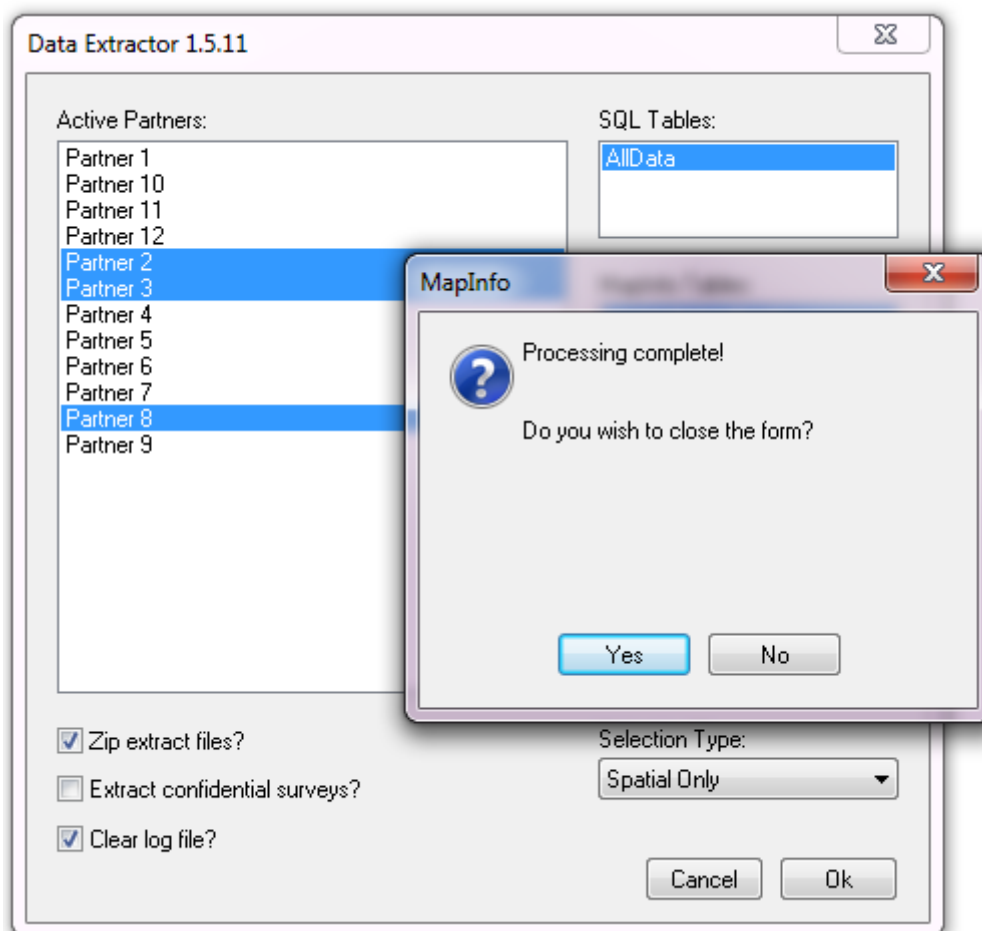
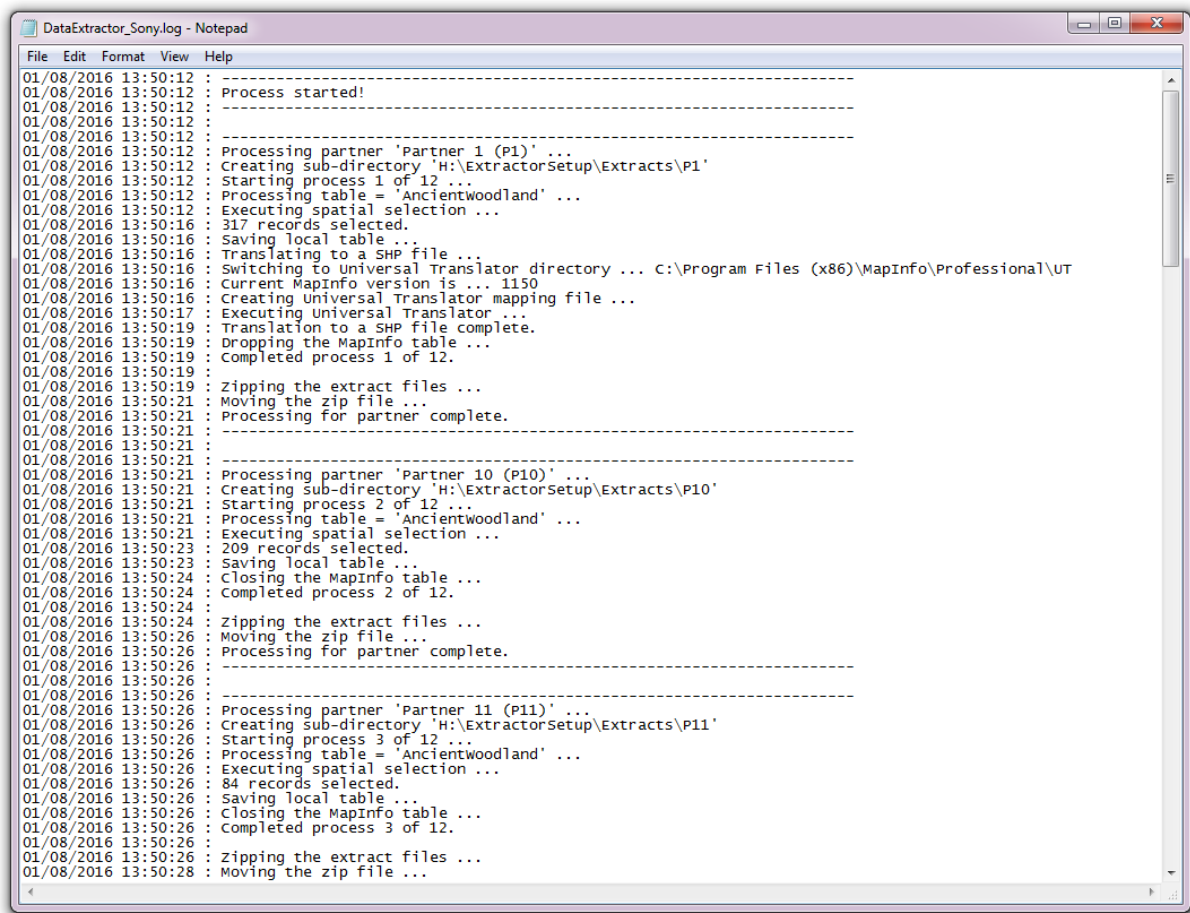


Fig. 5.6: Prompt to close form when extract process completes

Once the user makes a choice the log file is shown (Fig. 5.7). This should be checked thoroughly to ensure that all expected extracts have been generated, the format of the extract files is correct, and that the number of records included in each extract file is as expected.



```

DataExtractor_Sony.log - Notepad
File Edit Format View Help
-----
01/08/2016 13:50:12 : -----
01/08/2016 13:50:12 : Process started!
01/08/2016 13:50:12 : -----
01/08/2016 13:50:12 : -----
01/08/2016 13:50:12 : Processing partner 'Partner 1 (P1)' ...
01/08/2016 13:50:12 : Creating sub-directory 'H:\ExtractorSetup\Extracts\P1'
01/08/2016 13:50:12 : Starting process 1 of 12 ...
01/08/2016 13:50:12 : Processing table = 'Ancientwoodland' ...
01/08/2016 13:50:12 : Executing spatial selection ...
01/08/2016 13:50:16 : 317 records selected.
01/08/2016 13:50:16 : Saving local table ...
01/08/2016 13:50:16 : Translating to a SHP file ...
01/08/2016 13:50:16 : Switching to Universal Translator directory ... C:\Program Files (x86)\MapInfo\Professional\UT
01/08/2016 13:50:16 : Current MapInfo version is ... 1150
01/08/2016 13:50:16 : Creating Universal Translator mapping file ...
01/08/2016 13:50:17 : Executing Universal Translator ...
01/08/2016 13:50:19 : Translation to a SHP file complete.
01/08/2016 13:50:19 : Dropping the MapInfo table ...
01/08/2016 13:50:19 : Completed process 1 of 12.
01/08/2016 13:50:19 : -----
01/08/2016 13:50:19 : Zipping the extract files ...
01/08/2016 13:50:21 : Moving the zip file ...
01/08/2016 13:50:21 : Processing for partner complete.
01/08/2016 13:50:21 : -----
01/08/2016 13:50:21 : -----
01/08/2016 13:50:21 : Processing partner 'Partner 10 (P10)' ...
01/08/2016 13:50:21 : Creating sub-directory 'H:\ExtractorSetup\Extracts\P10'
01/08/2016 13:50:21 : Starting process 2 of 12 ...
01/08/2016 13:50:21 : Processing table = 'Ancientwoodland' ...
01/08/2016 13:50:23 : Executing spatial selection ...
01/08/2016 13:50:23 : 209 records selected.
01/08/2016 13:50:23 : Saving local table ...
01/08/2016 13:50:24 : Closing the MapInfo table ...
01/08/2016 13:50:24 : Completed process 2 of 12.
01/08/2016 13:50:24 : -----
01/08/2016 13:50:24 : Zipping the extract files ...
01/08/2016 13:50:26 : Moving the zip file ...
01/08/2016 13:50:26 : Processing for partner complete.
01/08/2016 13:50:26 : -----
01/08/2016 13:50:26 : -----
01/08/2016 13:50:26 : Processing partner 'Partner 11 (P11)' ...
01/08/2016 13:50:26 : Creating sub-directory 'H:\ExtractorSetup\Extracts\P11'
01/08/2016 13:50:26 : Starting process 3 of 12 ...
01/08/2016 13:50:26 : Processing table = 'Ancientwoodland' ...
01/08/2016 13:50:26 : Executing spatial selection ...
01/08/2016 13:50:26 : 84 records selected.
01/08/2016 13:50:26 : Saving local table ...
01/08/2016 13:50:26 : Closing the MapInfo table ...
01/08/2016 13:50:26 : Completed process 3 of 12.
01/08/2016 13:50:26 : -----
01/08/2016 13:50:26 : Zipping the extract files ...
01/08/2016 13:50:28 : Moving the zip file ...

```

Fig. 5.7: Example log file shown for review

You can now repeat the extract process for different partners and/or files as required.

5.3 Extract results

All results are written to the *DefaultPath* (page 21) folder as specified in the XML configuration document. As shown in Fig. 5.8 each partner has its own sub-folder where the extract files for that partner are stored in the formats specified in the partner GIS layer.

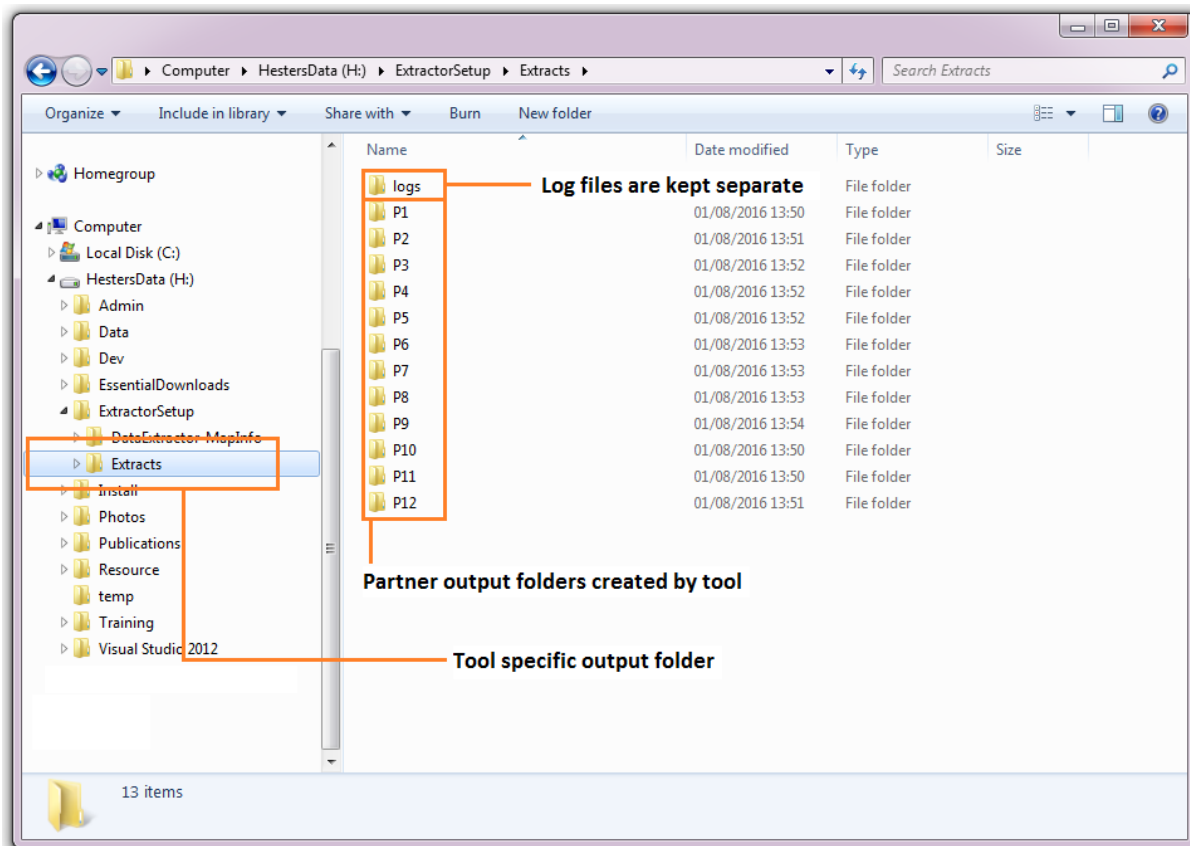


Fig. 5.8: Outputs are organised in partner-specific folders

The log file results of the process is saved in the *LogFilePath* (page 21) folder as specified in the XML configuration document.

Frequently Asked Questions

This is a list of Frequently Asked Questions about the Data Extractor tool. Feel free to suggest new entries!

6.1 General questions

How do I get a copy of the tool?

The latest version of the tool can be downloaded from [GitHub](#)¹⁶. Please ensure that you use the correct configuration file, an example copy of which is also included with the release.

Can several people use the tool at the same time?

Any number of users can use the tool simultaneously if they have a copy of it loaded in their own copy of MapInfo. The tool uses the data layers that are loaded in GIS in a read-only fashion, so there is no limit to the number of users of the tool. The stored procedures that are run store their temporary results in tables that carry the login name of the user, so as long as each user has a unique login ID no conflicts should arise. However, where results are written to a central (network) location, and the extraction is run for the same partner, conflicts may occur.

Does the tool work with QGIS or ArcGIS?

Currently only a MapInfo implementation of the tool exists. However, if funding was available the tool could be adapted to also support ArcGIS and/or QGIS.

6.2 Operating the tool

One of the data tables I want to use isn't showing in the form. How do I get it to show up?

This issue can arise in several ways:

¹⁶ <https://github.com/LERCAutomation/DataExtractor-MapInfo/releases>

- The table is a MapInfo GIS layer and isn't loaded in the active workspace. In this case, a *message will pop up* (page 48) before the form is shown telling you the layer isn't loaded. Add the layer to the workspace and the problem should be resolved.
- The table, either an SQL table or MapInfo GIS layer, isn't listed in the XML configuration document (or its name is not being obtained within the *TableListSQL node* (page 21)). Please refer to the *setup* (page 19) section and add it as a map layer or amend the SQL statement as appropriate.
- The table is a MapInfo GIS layer, and it is listed in the configuration document, but the *TableName* (page 24) is spelled incorrectly. Note that the name must follow the exact format of the name of the layer in the active workspace.

The partner I want to extract data for is not shown in the form. How do I get the name to show up?

This issue can arise in two ways:

- The partner boundary polygon does not exist in the partner boundary SQL table. Add it to the table.
- The partner's status is not **Active**. Change the value of the partner's status in the *ActiveColumn* (page 22) to Y.

The tool is not extracting the right data for the partner

This issue will arise if the names of files in the *FilesColumn* (page 22) are not correct. Check that the names match the table names in the XML configuration file. Alternatively, check the *FormatColumn* (page 22) and the *ExportColumn* (page 22) to ensure the correct format of data is requested.

6.3 Tool issues

How do I report a new bug or propose a change in the tool?

Please check the existing known issues and change requests on the LERCAutomation pages on [GitHub](#)¹⁷ before reporting/proposing new issues or changes. If you have a new issue or request you can submit it there and it will be picked up by the developers. Alternatively, you can email suggestions to [Hester](mailto:Hester@HesterLyonsConsulting.co.uk)¹⁸ or [Andy](mailto:Andy@AndyFoyConsulting.co.uk)¹⁹.

¹⁷ <https://github.com/LERCAutomation/DataExtractor-MapInfo>

¹⁸ Hester@HesterLyonsConsulting.co.uk

¹⁹ Andy@AndyFoyConsulting.co.uk

7.1 Example XML file

Below is an example of XML that might be used to set up the Data Extractor tool for MapInfo. Note, many of the settings have been included for illustration only and it is up to each user or LERC to ensure the system is configured to their requirements.

```
<?xml version="1.0" encoding="utf-8"?>

<!--
WARNING: This file should be changed carefully and a backup should be
taken before any changes so that they can be backed out.  Changed lines
can also be commented out as below.
-->

<!--
This config file contains all the variables used by the DataExtractor
MapBasic tool.

The 'configuration' node is the 'root' node and signifies the start of the
contents of the configuration file.

The 'DataExtractor' node contains all of the entries relating to the
MapBasic tool variables.

Each entry relates to a file, folder, table name, column name or SQL
statement used by the MapBasic tool to select and export data for the
GiGL partners.
-->

<configuration>
<DataExtractor>

  <!-- The existing file location where log files will be saved with output
  messages -->
```

```
<LogFilepath>
  <value>H:\ExtractorSetup\Extracts\logs</value>
</LogFilepath>

<!-- The location of the File DSN that specifies which SQL Server_
→database
  to connect to -->
<FileDSN>
  <value>H:\Data\Thames\ThamesNBNDData.dsn</value>
</FileDSN>

<!-- The existing file location under which all partner sub-folders will
  be created -->
<DefaultPath>
  <value>H:\ExtractorSetup\Extracts</value>
</DefaultPath>

<!-- The schema used in the SQL Server database -->
<DatabaseSchema>
  <value>dbo</value>
</DatabaseSchema>

<!-- The SQL statement used to list all the tables in SQL Server that can
  be selected
  by the user to extract for each partner -->
<TableListSQL>
  <value>Select table_name From information_schema.tables Where table_
→name
      Like 'AllData%' And table_name Not Like 'TVERC[_]Spp[_]%'
      And table_name Not Like '%Non%' Order By table_name</value>
</TableListSQL>

<!-- The name of the partner GIS layer in SQL Server used to select the
  records -->
<PartnerTable>
  <value>Partner_Extract_Boundaries</value>
</PartnerTable>

<!-- The name of the column in the partner GIS layer containing the
  partner name passed to SQL Server by the tool to use as the
  partner's boundary for selecting the records -->
<PartnerColumn>
  <value>PartnerName</value>
</PartnerColumn>

<!-- The name of the column in the partner GIS layer containing the
  abbreviated name passed to SQL Server by the tool to use as the
  sub-folder name for the destination of extracted records -->
<ShortColumn>
  <value>ShortName</value>
</ShortColumn>

<!-- The name of the column in the partner GIS layer containing any note
  text relating to the partner. -->
<NotesColumn>
  <value>Notes</value>
</NotesColumn>
```

```
<!-- The name of the column in the partner GIS layer containing the Y/N
      flag to indicate if the partner is currently active.  Only active
      partners will available for processing. -->
<ActiveColumn>
  <value>Active</value>
</ActiveColumn>

<!-- The name of the column in the partner GIS layer containing the GIS
      format required for the output records -->
<FormatColumn>
  <value>GISformat</value>
</FormatColumn>

<!-- The name of the column in the partner GIS layer indicating whether
      an export should also be created as a CSV file -->
<ExportColumn>
  <value>CSVfile</value>
</ExportColumn>

<!-- The name of the column in the partner GIS layer indicating which
      files should be created for each partner -->
<FilesColumn>
  <value>Files</value>
</FilesColumn>

<!-- The name of the column in the partner GIS layer indicating which
      survey tags, if any should be included in the export -->
<TagsColumn>
  <value>PartnerTags</value>
</TagsColumn>

  <!-- The options for the selection types -->
<SelectTypeOptions>
  <value>Spatial Only;Survey Tags Only;Spatial and Survey Tags</value>
</SelectTypeOptions>

<!-- The default selection type (1 = spatial, 2 = tags, 3 = both) -->
<DefaultSelectType>
  <value>1</value>
</DefaultSelectType>

<!-- The maximum number of records what will be extracted in any one
      partner extract -->
<RecMax>
  <value>10000000</value>
</RecMax>

<!-- The default value for zipping the extract files -->
<DefaultZip>
  <value>Yes</value>
</DefaultZip>

<!-- The SQL criteria for excluding any confidential surveys -->
<ConfidentialClause>
  <value></value>
</ConfidentialClause>
```

```

<!-- The default value for extracting confidential surveys -->
<DefaultConfidential>
  <value>No</value>
</DefaultConfidential>

<!-- The path to the Universal Translator program -->
<UTPath>
  <value>C:\Program Files (x86)\MapInfo\Professional\UT</value>
</UTPath>

<!-- The command to run the Universal Translator program -->
<UTCommand>
  <value>IMUT.exe</value>
</UTCommand>

<!-- The table columns and SQL where clauses used to select all the
required columns for the extract table subsets -->
<SQLTables>
  <Data_AllSpecies> <!-- The name of this subset as it is listed in the
partner table -->
    <TableName> <!-- The name of this subset as it will be exported -->
      <Value>Data_Spp_Full</Value>
    </TableName>
    <Columns> <!-- The columns that will be included in this subset -->
      <Value>TaxonName, CommonName, TaxonClass, TaxonGroup,
↳Abundance,
      AbundanceCount, RecDate, RecYear, VagueDateStart,
      VagueDateEnd, Recorder, Determiner, Gridref, RefSystem,
      Grid10K, GRPrecision, GRQualifier, Easting, Northing,
      Location, MoreInfo, RecType, StatusEuro, StatusUK,
      StatusNerc, StatusOther, StatusINNS, SurveyName,
      SurveyOrigin, SurveyRunBy, SurveyTags, Comments,
      Confidential, Sensitive, NegativeRec, HistoricRec,
      Verification, LastUpdated, SP_GEOMETRY</Value>
    </Columns>
    <Clauses> <!-- The SQL clause that should be used to extract this
subset from the SQL table -->
      <Value>RecYear &gt;= 1985 AND (NegativeRec &lt;&gt; 'Y' OR
      NegativeRec IS NULL) AND GRPrecision &lt;= 100 AND
      Gridref IS NOT NULL AND VagueDateStart IS NOT NULL AND
      Recorder IS NOT NULL AND TaxonName &lt;&gt; 'Homo
↳sapiens'
      AND Verification &lt;&gt; 'Considered incorrect'</Value>
    </Clauses>
    <Symbology> <!-- The symbology definition for this subset -->
      <Symbol> <!-- First symbol definition -->
        <Clause> <!-- The SQL clause that defines the records for
which this symbol will be used -->
          <Value>GRPrecision = 100</Value>
        </Clause>
        <Object> <!-- The type of object the symbol applies to -
↳-->
          <Value>Point</Value>
        </Object>
        <Type> <!-- The type of symbol -->
          <Value>Symbol</Value>

```

```

        </Type>
        <Style> <!-- The MapInfo style of the symbol -->
            <Value>2,64,255,14,MapInfo Dispersed Group,0,0</Value>
        </Style>
    </Symbol>
    <Symbol> <!-- Next symbol -->
        <Clause>
            <Value>GRPrecision &lt;= 10</Value>
        </Clause>
        <Object>
            <Value>Point</Value>
        </Object>
        <Type>
            <Value>Symbol</Value>
        </Type>
        <Style>
            <Value>2,65,255,12,MapInfo Dispersed Group,0,0</Value>
        </Style>
    </Symbol>
</Symbology>
</Data_AllSpecies>
<Data_Birds>
    <TableName>
        <Value>Data_Spp_Birds</Value>
    </TableName>
    <Columns>
        <Value>TaxonName, CommonName, TaxonClass, TaxonGroup,
↪Abundance,
            AbundanceCount, RecDate, RecYear, VagueDateStart,
            VagueDateEnd, Recorder, Determiner, Gridref, RefSystem,
            Grid10K, GRPrecision, GRQualifier, Easting, Northing,
            Location, MoreInfo, RecType, StatusEuro, StatusUK,
            StatusNerc, StatusOther, StatusINNS, SurveyName,
            SurveyOrigin, SurveyRunBy, SurveyTags, Comments,
            Confidential, Sensitive, NegativeRec, HistoricRec,
            Verification, LastUpdated, SP_GEOMETRY</Value>
    </Columns>
    <Clauses>
        <Value>RecYear &gt;= 1985 AND (NegativeRec &lt;&gt; 'Y' OR
            NegativeRec IS NULL) AND GRPrecision &lt;= 100 AND
            Gridref IS NOT NULL AND VagueDateStart IS NOT NULL AND
            Recorder IS NOT NULL AND TaxonName &lt;&gt; 'Homo
↪sapiens'
            AND Verification &lt;&gt; 'Considered incorrect' AND
            TaxonGroup = 'Birds'</Value>
    </Clauses>
    <Symbology>
        <Symbol>
            <Clause>
                <Value>GRPrecision = 100</Value>
            </Clause>
            <Object>
                <Value>Point</Value>
            </Object>
            <Type>
                <Value>Symbol</Value>
            </Type>
        </Symbol>
    </Symbology>
</Data_Birds>

```

```
        <Style>
          <Value>2,64,255,14,MapInfo Dispersed Group,0,0</Value>
        </Style>
      </Symbol>
    <Symbol>
      <Clause>
        <Value>GRPrecision &lt;= 10</Value>
      </Clause>
      <Object>
        <Value>Point</Value>
      </Object>
      <Type>
        <Value>Symbol</Value>
      </Type>
      <Style>
        <Value>2,65,255,12,MapInfo Dispersed Group,0,0</Value>
      </Style>
    </Symbol>
  </Symbology>
</Data_Birds>
</SQLTables>

<!-- The names and local names of the MapInfo tables and the required
      columns for the MapInfo tables -->
<MapTables>
  <AncientWoodland> <!-- The name of this MapInfo table as it is listed
                    in the partner table -->
    <TableName> <!-- The name of this MapInfo table as it is shown in
                the MapInfo interface and on the form -->
      <Value>AncientWoodland</Value>
    </TableName>
    <Columns> <!-- Columns that will be included in the extract -->
      <Value>NAME, THEMNAME, STATUS, x_COORD, y_COORD, AREA,
            PERIMETER</Value>
    </Columns>
  </AncientWoodland>
</MapTables>

</DataExtractor>
</configuration>
```

7.2 GNU Free Documentation License

GNU Free Documentation License
Version 1.3, 3 November 2008

Copyright (C) 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc.
<<http://fsf.org/>>

Everyone **is** permitted to copy **and** distribute verbatim copies of this license document, but changing it **is not** allowed.

0. PREAMBLE

The purpose of this License **is** to make a manual, textbook, **or** other functional **and** useful document "**free**" **in** the sense of freedom: to assure everyone the effective freedom to copy **and** redistribute it, **with or** without modifying it, either commercially **or** noncommercially. Secondly, this License preserves **for** the author **and** publisher a way to get credit **for** their work, **while not** being considered responsible **for** modifications made by others.

This License **is** a kind of "**copyleft**", which means that derivative works of the document must themselves be free **in** the same sense. It complements the GNU General Public License, which **is** a copyleft license designed **for** free software.

We have designed this License **in** order to use it **for** manuals **for** free software, because free software needs free documentation: a free program should come **with** manuals providing the same freedoms that the software does. But this License **is not** limited to software manuals; it can be used **for any** textual work, regardless of subject matter **or** whether it **is** published **as** a printed book. We recommend this License principally **for** works whose purpose **is** instruction **or** reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to **any** manual **or** other work, **in any** medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited **in** duration, to use that work under the conditions stated herein. The "**Document**", below, refers to **any** such manual **or** work. Any member of the public **is** a licensee, **and is** addressed **as** "**you**". You accept the license **if** you copy, modify **or** distribute the work **in** a way requiring permission under copyright law.

A "**Modified Version**" of the Document means **any** work containing the Document **or** a portion of it, either copied verbatim, **or with** modifications **and/or** translated into another language.

A "**Secondary Section**" **is** a named appendix **or** a front-matter section of the Document that deals exclusively **with** the relationship of the publishers **or** authors of the Document to the Document's **overall** subject (**or** to related matters) **and** contains nothing that could fall directly within that overall subject. (Thus, **if** the Document **is in** part a textbook of mathematics, a Secondary Section may **not** explain

any mathematics.) The relationship could be a matter of historical connection **with** the subject **or with** related matters, **or** of legal, commercial, philosophical, ethical **or** political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, **as** being those of Invariant Sections, **in** the notice that says that the Document **is** released under this License. If a section does **not** fit the above definition of Secondary then it **is not** allowed to be designated **as** Invariant. The Document may contain zero Invariant Sections. If the Document does **not** identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, **as** Front-Cover Texts **or** Back-Cover Texts, **in** the notice that says that the Document **is** released under this License. A Front-Cover Text may be at most 5 words, **and** a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented **in** a format whose specification **is** available to the general public, that **is** suitable **for** revising the document straightforwardly **with** generic text editors **or** (**for** images composed of pixels) generic paint programs **or** (**for** drawings) some widely available drawing editor, **and** that **is** suitable **for** input to text formatters **or** **for** automatic translation to a variety of formats suitable **for** input to text formatters. A copy made **in** an otherwise Transparent file format whose markup, **or** absence of markup, has been arranged to thwart **or** discourage subsequent modification by readers **is not** Transparent. An image format **is not** Transparent **if** used **for** any substantial amount of text. A copy that **is not** "Transparent" **is** called "Opaque".

Examples of suitable formats **for** Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML **or** XML using a publicly available DTD, **and** standard-conforming simple HTML, PostScript **or** PDF designed **for** human modification. Examples of transparent image formats include PNG, XCF **and** JPG. Opaque formats include proprietary formats that can be read **and** edited only by proprietary word processors, SGML **or** XML **for** which the DTD **and/or** processing tools are **not** generally available, **and** the machine-generated HTML, PostScript **or** PDF produced by some word processors **for** output purposes only.

The "Title Page" means, **for** a printed book, the title page itself, plus such following pages **as** are needed to hold, legibly, the material this License requires to appear **in** the title page. For works **in** formats which do **not** have any title page **as** such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person **or** entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either **is** precisely XYZ **or** contains XYZ **in** parentheses following text that translates XYZ **in** another language. (Here XYZ stands **for** a specific section name mentioned below, such **as** "Acknowledgements", "Dedications", "Endorsements", **or** "History".) To "Preserve the Title"

of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that

edition to the public.

It **is** requested, but **not** required, that you contact the authors of the Document well before redistributing **any** large number of copies, to give them a chance to provide you **with** an updated version of the Document.

4. MODIFICATIONS

You may copy **and** distribute a Modified Version of the Document under the conditions of sections 2 **and** 3 above, provided that you release the Modified Version under precisely this License, **with** the Modified Version filling the role of the Document, thus licensing distribution **and** modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things **in** the Modified Version:

- A. Use **in** the Title Page (**and** on the covers, **if any**) a title distinct **from that** of the Document, **and from those** of previous versions (which should, **if** there were **any**, be listed **in** the History section of the Document). You may use the same title **as** a previous version **if** the original publisher of that version gives permission.
- B. List on the Title Page, **as** authors, one **or** more persons **or** entities responsible **for** authorship of the modifications **in** the Modified Version, together **with** at least five of the principal authors of the Document (**all** of its principal authors, **if** it has fewer than five), unless they release you **from this** requirement.
- C. State on the Title page the name of the publisher of the Modified Version, **as** the publisher.
- D. Preserve **all** the copyright notices of the Document.
- E. Add an appropriate copyright notice **for** your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, **in** the form shown **in** the Addendum below.
- G. Preserve **in** that license notice the full lists of Invariant Sections **and** required Cover Texts given **in** the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, **and** add to it an item stating at least the title, year, new authors, **and** publisher of the Modified Version **as** given on the Title Page. If there **is** no section Entitled "History" **in** the Document, create one stating the title, year, authors, **and** publisher of the Document **as** given on its Title Page, then add an item describing the Modified Version **as** stated **in** the previous sentence.
- J. Preserve the network location, **if any**, given **in** the Document **for** public access to a Transparent copy of the Document, **and** likewise the network locations given **in** the Document **for** previous versions it was based on. These may be placed **in** the "History" section. You may omit a network location **for** a work that was published at least four years before the Document itself, **or if** the original publisher of the version it refers to gives permission.
- K. For **any** section Entitled "Acknowledgements" **or** "Dedications", Preserve the Title of the section, **and** preserve **in** the section **all** the substance **and** tone of each of the contributor acknowledgements **and/or** dedications given therein.
- L. Preserve **all** the Invariant Sections of the Document,

- unaltered **in** their text **and in** their titles. Section numbers **or** the equivalent are **not** considered part of the section titles.
- M. Delete **any** section Entitled "Endorsements". Such a section may **not** be included **in** the Modified Version.
- N. Do **not** retitle **any** existing section to be Entitled "Endorsements" **or** to conflict **in** title **with any** Invariant Section.
- O. Preserve **any** Warranty Disclaimers.

If the Modified Version includes new front-matter sections **or** appendices that qualify **as** Secondary Sections **and** contain no material copied **from the** Document, you may at your option designate some **or all** of these sections **as** invariant. To do this, add their titles to the **list** of Invariant Sections **in** the Modified Version's **license notice**. These titles must be distinct **from any** other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—**for** example, statements of peer review **or** that the text has been approved by an organization **as** the authoritative definition of a standard.

You may add a passage of up to five words **as** a Front-Cover Text, **and** a passage of up to 25 words **as** a Back-Cover Text, to the end of the **list** of Cover Texts **in** the Modified Version. Only one passage of Front-Cover Text **and** one of Back-Cover Text may be added by (**or** through arrangements made by) **any** one entity. If the Document already includes a cover text **for** the same cover, previously added by you **or** by arrangement made by the same entity you are acting on behalf of, you may **not** add another; but you may replace the old one, on explicit permission **from the** previous publisher that added the old one.

The author(s) **and** publisher(s) of the Document do **not** by this License give permission to use their names **for** publicity **for or** to **assert or** imply endorsement of **any** Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document **with** other documents released under this License, under the terms defined **in** section 4 above **for** modified versions, provided that you include **in** the combination **all** of the Invariant Sections of **all** of the original documents, unmodified, **and** **list** them **all as** Invariant Sections of your combined work **in** its license notice, **and** that you preserve **all** their Warranty Disclaimers.

The combined work need only contain one copy of this License, **and** multiple identical Invariant Sections may be replaced **with** a single copy. If there are multiple Invariant Sections **with** the same name but different contents, make the title of each such section unique by adding at the end of it, **in** parentheses, the name of the original author **or** publisher of that section **if** known, **or else** a unique number. Make the same adjustment to the section titles **in** the **list** of Invariant Sections **in** the license notice of the combined work.

In the combination, you must combine **any** sections Entitled "History" **in** the various original documents, forming one section Entitled "History"; likewise combine **any** sections Entitled "Acknowledgements",

and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document **and** other documents released under this License, **and** replace the individual copies of this License **in** the various documents **with** a single copy that **is** included **in** the collection, provided that you follow the rules of this License **for** verbatim copying of each of the documents **in** all other respects.

You may extract a single document **from such** a collection, **and** distribute it individually under this License, provided you insert a copy of this License into the extracted document, **and** follow this License **in** all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document **or** its derivatives **with** other separate **and** independent documents **or** works, **in or** on a volume of a storage **or** distribution medium, **is** called an "aggregate" **if** the copyright resulting **from the** compilation **is not** used to limit the legal rights of the compilation's users **beyond what the individual works permit**. When the Document **is** included **in** an aggregate, this License does **not** apply to the other works **in** the aggregate which are **not** themselves derivative works of the Document.

If the Cover Text requirement of section 3 **is** applicable to these copies of the Document, then **if** the Document **is** less than one half of the entire aggregate, the Document's Cover Texts **may be placed on** covers that bracket the Document within the aggregate, **or** the electronic equivalent of covers **if** the Document **is in** electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation **is** considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections **with** translations requires special permission **from their** copyright holders, but you may include translations of some **or all** Invariant Sections **in** addition to the original versions of these Invariant Sections. You may include a translation of this License, **and all** the license notices **in** the Document, **and any** Warranty Disclaimers, provided that you also include the original English version of this License **and** the original versions of those notices **and** disclaimers. In case of a disagreement between the translation **and** the original version of this License **or** a notice **or** disclaimer, the original version will prevail.

If a section **in** the Document **is** Entitled "Acknowledgements", "Dedications", **or** "History", the requirement (section 4) to Preserve

its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may **not** copy, modify, sublicense, **or** distribute the Document **except as** expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, **or** distribute it **is** void, **and** will automatically terminate your rights under this License.

However, **if** you cease **all** violation of this License, then your license **from a** particular copyright holder **is** reinstated (a) provisionally, unless **and** until the copyright holder explicitly **and finally** terminates your license, **and** (b) permanently, **if** the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license **from a** particular copyright holder **is** reinstated permanently **if** the copyright holder notifies you of the violation by some reasonable means, this **is** the first time you have received notice of violation of this License (**for any work**) **from that** copyright holder, **and** you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does **not** terminate the licenses of parties who have received copies **or** rights **from you** under this License. If your rights have been terminated **and not** permanently reinstated, receipt of a copy of some **or all** of the same material does **not** give you **any** rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License **from time** to time. Such new versions will be similar **in** spirit to the present version, but may differ **in** detail to address new problems **or** concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License **is** given a distinguishing version number. If the Document specifies that a particular numbered version of this License "**or any later version**" applies to it, you have the option of following the terms **and** conditions either of that specified version **or** of **any** later version that has been published (**not as** a draft) by the Free Software Foundation. If the Document does **not** specify a version number of this License, you may choose **any** version ever published (**not as** a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's **public statement of acceptance of a** version permanently authorizes you to choose that version **for** the Document.

11. RELICENSING

"Massive Multiauthor Collaboration Site" (**or** "MMC Site") means **any** World Wide Web server that publishes copyrightable works **and** also

provides prominent facilities **for** anybody to edit those works. A public wiki that anybody can edit **is** an example of such a server. A "Massive Multiauthor Collaboration" (**or** "MMC") contained **in** the site means **any set** of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a **not-for-profit** corporation **with** a principal place of business **in** San Francisco, California, **as well as** future copyleft versions of that license published by that same organization.

"Incorporate" means to publish **or** republish a Document, **in whole or in part, as part** of another Document.

An MMC **is** "eligible for relicensing" **if** it **is** licensed under this License, **and if** all works that were first published under this License somewhere other than this MMC, **and** subsequently incorporated **in whole or in part** into the MMC, (1) had no cover texts **or** invariant sections, **and** (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained **in** the site under CC-BY-SA on the same site at **any** time before August 1, 2009, provided the MMC **is** eligible **for** relicensing.

ADDENDUM: How to use this License **for** your documents

To use this License **in** a document you have written, include a copy of the License **in** the document **and** put the following copyright **and** license notices just after the title page:

Copyright (c) YEAR YOUR NAME.

Permission **is** granted to copy, distribute **and/or** modify this document under the terms of the GNU Free Documentation License, Version 1.3 **or any** later version published by the Free Software Foundation; **with** no Invariant Sections, no Front-Cover Texts, **and** no Back-Cover

→Texts.

A copy of the license **is** included **in** the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts **and** Back-Cover Texts, replace the "with...Texts." line **with** this:

with the Invariant Sections being LIST THEIR TITLES, **with** the Front-Cover Texts being LIST, **and with** the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, **or** some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples **in** parallel under your choice of free software license, such **as** the GNU General Public License, to permit their use **in** free software.

A

Acknowledgements, 5

B

Background, 7

Benefits, 9

C

Configuring the tool, 19

Conventions used in this user guide, 6

D

Data extract process, 11

E

Example XML file, 57

Extraction results, 54

F

FAQ, 54

Frequently asked questions, *see* FAQ

G

General attributes

 ArcGIS, 25

 MapInfo, 21

I

Installing the tool, 30

 ArcGIS, 34

 MapInfo, 30

L

Licensing, 4

M

Map table attributes

 ArcGIS, 27

 MapInfo, 24

O

Opening the form, 48

R

Reading guide, 4

Recommended user knowledge, 3

Running the tool, 46

S

Setting up the SQL database, 29

Setup for ArcGIS, 25

Setup for MapInfo, 21

Special characters in XML, 20

SQL table attributes

 ArcGIS, 27

 MapInfo, 23

T

Tool outputs, 14

Tool overview, 8, 12

U

Useful links, 4

Using the form, 50