

# IndexConvert

## User Guide

### For Release 1.8.7

# Introduction

IndexConvert helps reverse engineer indexes into indexing software. It can also embed an index into Word when page numbering matches. Data presented as serial structured records can be converted to a database format using IndexConvert. IndexConvert will convert files with up to 32 heading levels. Cindex supports 15 levels, Sky Index and Word embedded support 7 levels. Export to spreadsheet or database supports 32 levels, enabling transformation of very complex data unlikely to be related to indexing.

Importing a back-of-book index into an indexing program or database requires the index to be restructured into a format that can be readily imported. This is normally a complex process that can be tedious and error prone when done manually. Some programs such as Macrex use clever software tricks that can make the process easier. The later more sophisticated index transfer file formats such as IXML and IDEX cannot be created without sophisticated computer support.

IndexConvert helps by regularizing and automating the most complex processes leaving relatively few manual adjustments to be made. This saves time and cost and improves quality. A simple strategy is used. Headings are labelled \$H1\_, \$H2\_ etc. and locator strings are labelled with a tab followed by \$L\_. Cross references are locators. Once completed, a variety of exports are available, up to 15 heading levels can be processed.

IndexConvert requires a set-out index. A user can insert labels manually or by search and replace, when converting a run-on index to a set-out layout, this makes indexConvert extremely versatile.

Preprocess looks for clues in the index which it uses to offer processes which help shape the file into a single column index for follow-on processing.

Headings and locator strings are then identified and labelled. Once a heading or locator string has been correctly labelled it is ignored in future labelling operations. This allows any combination of manual and automatic labelling to be used, depending on user preferences and the difficulties of the particular index.

Audit follows. This checks the integrity of the index and labels entries that need attention. The first stage identifies unlabelled headings, the second stage identifies missing locators. A locator is expected for every heading or subheading, this ensures that suspect entries are labelled for future attention. Legitimate headings can be ignored. Labels and punctuation are set to the regular style to prevent unwanted style overflow.

Where export is to a tab delimited file, Concatenate (prepend) and style labelling follows. Styles including bold and italic are labelled and can be restored using find/replace in the receiving software.

Finally the file is exported to the chosen file format.

IndexConvert has been used by indexers in a number of countries for converting indexes into Cindex, Macrex, Sky index and Index-Manager formats. To date, English, French, Dutch and German indexes have been converted. Back-of-book, legal and Biblical citation indexes have been converted. Indexes have been received as Word files, PDFs and IXML files. For IXML files, Cindex was used to add most of the labels, IndexConvert topped up those missed, with Audit giving a final check.

IndexConvert export formats include:

Tab Delimited Unicode UTF-8 file for Cindex v3.

IXML Unicode UTF-8 file for Cindex v3.

Expanded IXML file for Cindex v3 with a copy of the locator in the lowest subheading.

Tab delimited ANSI file for Sky Index SI7.

INDEX Unicode UTF-8 file for Sky Index SI8.

ANSI MBK file for Macrex.

Unicode UTF-8 file for Index-Manager

Tab delimited Unicode UTF-8 spreadsheet file with that may contain more than one locator in each record.

Tab delimited Unicode UTF-8 database file containing one locator in each record.

A Word AutoMark file can be created from an index.

IndexConvert can add an embedded index to a Word file where page numbers match.

IndexConvert includes **TextConvert** and **CharConvert** utilities that simplify entry of non-ANSI text and diacritic characters into Macrex and Sky Index v7. CharConvert can support development of translation tables by avoiding the need to use look-up tables when finding the decimal or hex value of a character.

**Elided 123-4** converts elided page numbers to an un-elided format needed by indexing programs when repaginating. Additional Elided conversions will be provided as required or can be supplied by third parties.

IndexConvert settings are remembered between sessions, this simplifies use following initial setup.

## How long will a conversion take?

Every index is different and offers a fresh set of challenges. If it's important to know how long processing will take, extract a single page of the index received. Run through the complete process and use the experience to estimate the processing of the entire document. It is likely that mistakes will be made, and tips and tricks learned which will speed processing of the whole document. A well-structured index with 10,000 entries has been converted in an hour. If exported from a PDF the same index may require a day or more depending on the user's familiarity with the conversion process.

## Disclaimer

Please read the [Disclaimer](#) before using IndexConvert.

## Contents

Introduction.....	ii
How long will a conversion take? .....	iii
Disclaimer.....	iii
Contents .....	iv
Tables.....	viii
Figures .....	viii
Concept.....	1
Compatibility.....	2
Microsoft Word .....	2
PDF .....	2
Exports from Indexing Software.....	2
Cindex Import.....	2
Macrex Import .....	2
Sky Index v7 Import .....	2
Sky Index v8 Import .....	2
Index-Manager Import.....	3
Word Embedded Index .....	3
Installation .....	3
Process Summary.....	4
Initial inspection.....	4
User Interface.....	4
Configuration.....	5
Preprocess .....	5
Locators L.....	5
Label Headings .....	5
Locators M.....	6
Audit .....	6
Elided .....	6
Concatenation and Style Labelling .....	7
Export .....	7
Process Commonality .....	7
Iteration .....	7
Quality Considerations.....	7
Index Preparation.....	8
Microsoft Word .....	8
Indexing Software .....	8

PDFs .....	8
Index Layout (IXML Export) .....	8
Index Layout (Delimited Export) .....	8
Locator Separators .....	9
Using IndexConvert .....	9
Start IndexConvert .....	9
Shortcuts .....	9
User Interface .....	10
Version .....	10
Configuration .....	10
Select Export Type .....	11
Progress Indicator .....	11
Ctrl-Break / Esc .....	11
Preprocess .....	12
Label Locators .....	14
Locator Options .....	14
Page Number Limit .....	15
Roman Numerals .....	15
Other Locators .....	15
Automatic Locator Search .....	15
Cross References .....	16
Locator Prefixes .....	17
Label Headings .....	18
Heading Options .....	18
Method A .....	19
Heading Levels .....	19
Method B .....	20
Method C .....	21
Audit .....	22
Audit Report .....	22
Manual Labelling .....	22
Headings .....	22
Locators .....	23
Elided Locators .....	23
Concatenate (Prepend) .....	23
Processing Speed and Timing .....	23
Unicode Characters: Sky Index v7 .....	24

Unicode Characters: Macrex .....	24
Label Styles.....	24
Export.....	24
File Extensions and Encoding .....	25
Exit.....	25
Metrics.....	25
Utilities .....	26
TextConvert.....	26
CharConvert.....	26
Measure Indent.....	27
Associated Add-Ins .....	28
Elided 123-4 .....	28
Data Structures .....	29
Cindex v3 Tab Delimited.....	29
Spreadsheet Tab Delimited .....	29
Macrex MBK.....	29
Sky Index v7Tab Delimited .....	30
Database Tab delimited .....	30
Cindex IXML.....	30
Cindex IXML Expanded .....	30
Sky Index SI8 IXML .....	30
Sky Index SI8 IDEX .....	31
Index-Manager .....	31
Word AutoMark File .....	31
Word Embedded .....	31
Labels used by IndexConvert .....	32
Headings and Locators .....	32
Styles.....	33
Error Codes .....	34
Error Code Definitions.....	34
Utilities - Examples .....	38
CharConvert.....	38
TextConvert.....	40
Macrex.....	40
Sky Index (SI7).....	40
a b c [c,] d e f g [gu] h i{\$Er305} i j k l m n o [o"] p r s [s,] t u [u"] v y z .....	40
Licences .....	41

Advice and Notes .....	42
Cindex .....	42
Cindex IXML Index Records .....	42
Cindex Tab Delimited File .....	42
Inspect File .....	42
Import .....	42
Remove Labels .....	42
Adjust Styles .....	42
Compare Files .....	43
Macrex .....	43
MBK, comma delimited .....	43
Unicode.ini .....	44
Macrex Import .....	45
SKY Index .....	45
IDEX Index Records .....	45
Tab Delimited .....	46
Concatenation .....	46
Import to Sky Index .....	47
Index-Manager .....	49
Spreadsheet, tab delimited .....	50
Database, tab delimited .....	50
Word AutoMark File .....	50
Word Embedded Using Page Numbers .....	50
Page Numbers .....	50
Options .....	51
Screen Flicker .....	51
Process .....	51
Locator Styles .....	52
Cross References (\t Field) .....	52
Styles .....	52
Error Codes .....	52
Word File Compare Tips .....	53
Alternative Installation .....	54
Trust Settings .....	54
Activate .....	54
AutoExec Word Macro .....	55
Test Indexes .....	56

Disclaimer.....	57
IndexConvert Copying, Licence Keys, and Redistribution.....	57
Enquiries .....	57
INDEX .....	58
Conventional Index .....	58
Linked Index.....	60

## Tables

Table 1 Navigation and Scrolling.....	12
Table 2 Example Cross Reference Terms (Lower Case).....	16
Table 3 Export file encoding and extensions.....	25
Table 4 Data Structures.....	29
Table 5 Content Labels.....	32
Table 6 Labelling example.....	32
Table 7 Style Labels .....	33
Table 8 Error Codes.....	34
Table 9 Licences .....	41
Table 10 Cindex Replace All Patterns to Recover Styles from Markup .....	43
Table 11 Macrex Unicode Table - Observations.....	44
Table 12 Embed Error Codes .....	53

## Figures

Figure 1 IndexConvert Process Summary .....	1
Figure 2 Security Warning .....	3
Figure 3 Installation of IndexConvert.....	4
Figure 4 IndexConvert User Interface .....	5
Figure 5 Manual Intervention .....	6
Figure 6 User Interface .....	10
Figure 7 Check for Updates.....	10
Figure 8 Configuration .....	10
Figure 9 Select Export Type .....	11
Figure 10 Progress Indicator .....	11
Figure 11 Example Dialog Box .....	12
Figure 12 Select Adjustment.....	13
Figure 13 Preprocess Error Message.....	13
Figure 14 Index Structure Dialog.....	14
Figure 15 Page Number Limit.....	15
Figure 16 Locator [ I ] Selected, Showing Roman Numerals Checked .....	15
Figure 17 Locators [ i ] Selected Showing Cross Reference Matrix .....	16

Figure 18 Locators [ i ] Selected Showing Optional Locator prefixes .....	17
Figure 19 Label Headings Options .....	18
Figure 20 Heading Level Error.....	18
Figure 21 Label Headings, Levels 1-3.....	19
Figure 22 Label Headings, Levels 4-6.....	19
Figure 23 Label Headings, Levels 7-10+.....	20
Figure 24 Label Headings Method B .....	20
Figure 25 Heading Dialog .....	20
Figure 26 Method B - Continue, Label or Cancel.....	21
Figure 28 Method C Heading Labelling.....	21
Figure 29 Concatenation Error .....	24
Figure 30 Metrics .....	25
Figure 31 TextConvert .....	26
Figure 32 CharConvert .....	27
Figure 33 Measure Indent.....	27
Figure 34 CharConvert – Macrex .....	39
Figure 35 CharConvert - Sky Index .....	39
Figure 36 Cindex Replace All Showing Recovery of Bold Font from Markup .....	43
Figure 37 Select Macrex Unicode file.....	44
Figure 38 Sky Index Translation File not found.....	46
Figure 39 Read Sky Index Translation File .....	47
Figure 40 Sky Index Replace.....	48
Figure 41 Sky Index Style Recovery .....	48
Figure 42 Embed Options.....	51
Figure 43 Microsoft Office Trusted Locations.....	54
Figure 44 Templates and Add-ins .....	55

## Concept

The principal part of IndexConvert processing is the identification and labelling of headings and locator strings. When elements are identified they can be manipulated for export to the chosen file format.

Headings are labelled \$H1\_ and subheadings are labelled \$H2\_. Lower level subheadings are labelled \$Hn\_ where n can be up to 15. Locators are labelled {tab}\$L\_.

Many checks are built into the process to enable accurate and consistent labelling. An audit process is the final check before export can take place. An audit report can be generated containing outstanding issues. This can be sent if support is needed.

Whilst IndexConvert is designed to do the labelling, if an initial inspection of the document suggests an easier alternative method, then use it. IndexConvert should be tried first as indexes generated from indexing software are extremely well structured and IndexConvert processing can be very fast. An index with a history of many updates to a wordprocessor file, or an export from PDF will present more challenges. For a names index where only Level 1 headings are bold capital, and all subheadings are a different style then Word find/replace may provide a fast solution. Where subentries are indicated by a comma, then replace the comma by a paragraph marker followed by a \$H2\_ label. If any other tricks present themselves, use them. A run-on index must be converted to a set-out style before using IndexConvert.

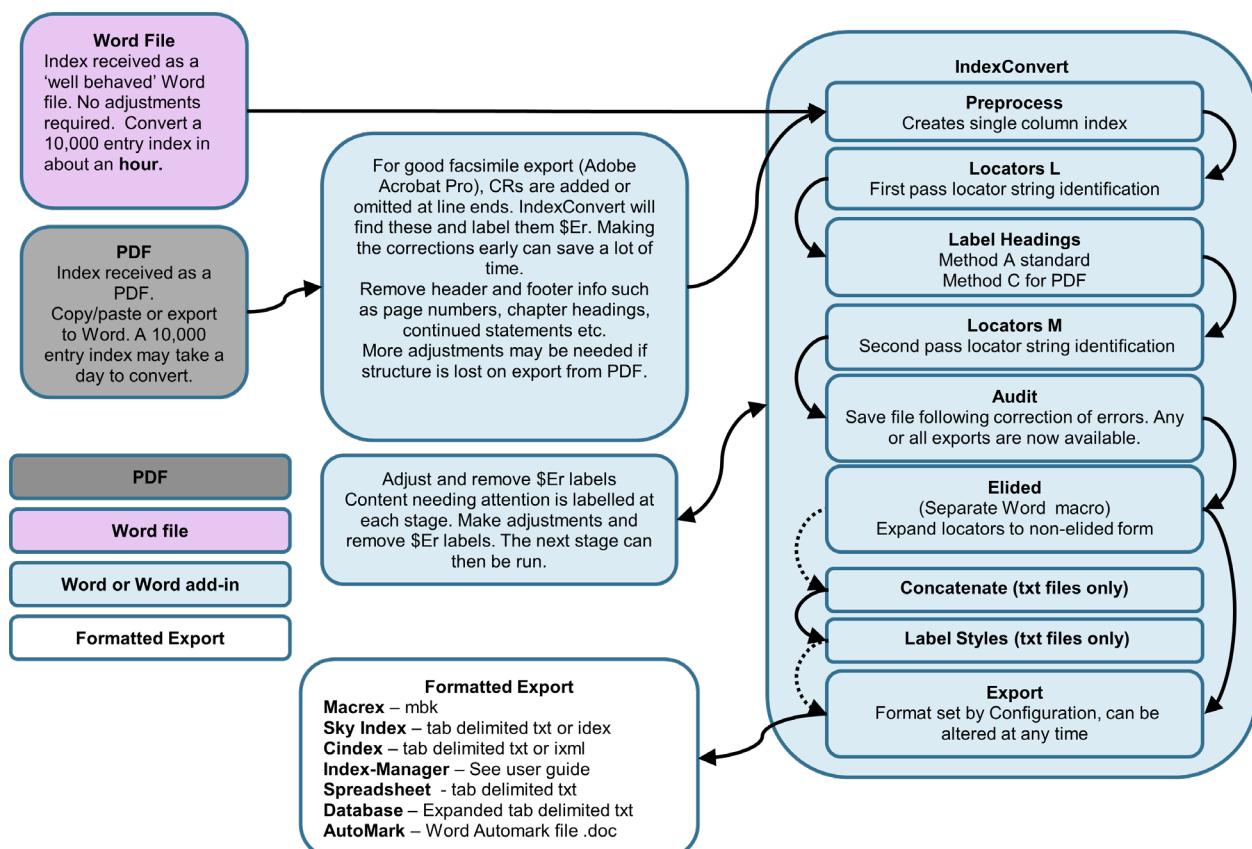


Figure 1 IndexConvert Process Summary

# Compatibility

## Microsoft Word

IndexConvert is a Word macro written in Visual Basic (VBA). VBA is closely integrated with Word supporting powerful and intricate processing. Use of IndexConvert requires the index to be in a .docm (macro enabled) format.

The add-in requires a 'desktop' version of Microsoft Word and will not run with an online version of Office 365. IndexConvert has been tested on Word 2007 through 2016 with the exception of Word 2008 which does not support Visual Basic (VBA), and Office 365.

Operation with Word for Macintosh was confirmed in early 2026. I was found that ixml, IDEX and Index Manager export will not work because they use an ActiveX component (ADODB) only available to Microsoft products. There are still plenty of tab delimited export options available.

## PDF

PDF files can be converted to Word files suitable for use with IndexConvert. Testing was carried out using Adobe Acrobat Pro DC version 2017.012.20098. More manual processing may be required if other Acrobat viewers are used.

## Exports from Indexing Software

Any indexing software can add IndexConvert labels to headings and locators. This provides a mechanism for converting between indexing programs.

## Cindex Import

Cindex version 3 and 4 can import tab delimited Spreadsheet, Database and IXML files generated by IndexConvert.

Locators must be separated by commas for correct processing within Cindex.

Cindex will not import ixml files with more than 15 heading levels.

## Macrex Import

Macrex can import comma delimited MBK files generated by IndexConvert. Diacritic and Unicode character support is provided by a Unicode.ini file supplied as part of Macrex. IndexConvert uses this to convert Unicode characters to printer replacement codes.

## Sky Index v7 Import

Sky Index v7 can import ANSI coded tab delimited files.

Diacritic and Unicode characters are supported by a translation file that must be created by the user. IndexConvert uses a copy of this to convert Unicode characters to keyboard entry codes.

Sky Index supports 7 heading levels.

## Sky Index v8 Import

Sky Index v8 should import database format tab delimited files.

SI8 supports Unicode UTF-8 characters, translation is not required.

The export is an IDEX file that can be imported directly to SI8.

Sky Index requires no more than one locator per record. Adjustments are needed before export from IndexConvert, these are explained in Section 14.

Sky Index supports 7 heading levels.

## Index-Manager Import

The Index-Manager import file is a Unicode UTF-8 file. By default Index-Manager supports 3 heading levels. More can be added for individual entries. Limits will be set by target software, 7 for Word, 4 for InDesign.

## Word Embedded Index

IndexConvert will export an index as embedded XE fields into a chosen file.

## Installation

Macros are supplied in Word docm (macro enabled document) files. The installation saves the file as a Word dotm (macro enabled template) file in the Word Startup folder.

It is recommended that Word is closed and restarted before an installation is attempted.

To Install IndexConvert open the IndexConvert x.x.x.docm where x.x.x is the version number. On opening the file a security warning may be seen as below.

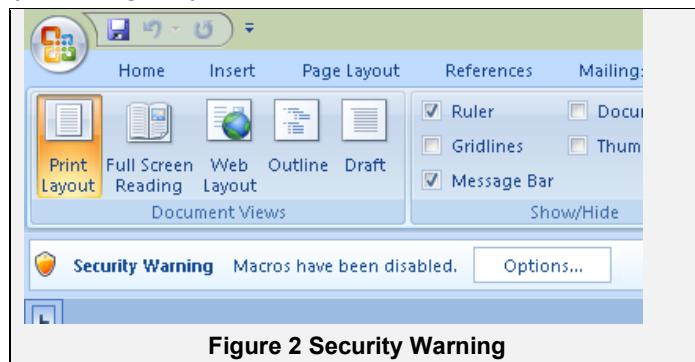
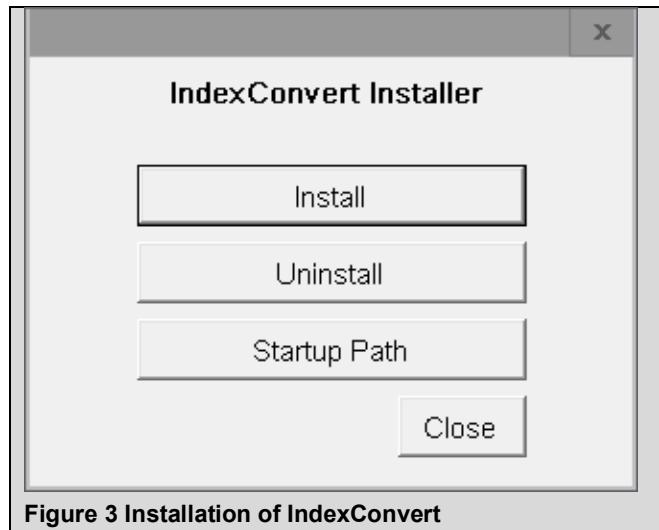


Figure 2 Security Warning

Select Options and allow macros to run.

The dialog box shown in Figure 3 appears.



**Figure 3 Installation of IndexConvert**

Select Install. If an earlier version of IndexConvert is present the installer will attempt to remove it. If it is not possible then an error message will appear stating that the earlier version of IndexConvert.dotm cannot be uninstalled.

Select Startup Path and make a note of the location

Close Word and follow the advice below.

Using file explorer, goto the startul path, an example is shown below.

c:\users\user\roaming\microsoft\Word\START

Delete IndexConvert.dotm

Open the file containing the IndexConvert macro and IndexConvert.dotm should install.

If you cannot do this then contact your system administrator."

Select Uninstall if you need to uninstall IndexConvert.

Installation of Elided 123-4 uses the same process.

Word should be restarted following an installation or re-installation.

If your IT organization prevents this installation or you prefer a different installation then see [Alternative Installation](#).

## Process Summary

### Initial inspection

Check that the file 'looks' like an index. Remove any introductory text. If group headings are longer than a single character then remove them.

If track changes is selected then ensure that No markup is selected as the additional characters will confuse IndexConvert.

### User Interface

The IndexConvert user interface is designed to support a process flow defined by controls in a column at the left. Start with Configuration and end with Export.

## Configuration

Select export format.

## Preprocess

An end of file marker idxEND is inserted to support termination of processing. This is automatically removed during Concatenate. Several internal parameters are set and the metrics are updated at this stage. This part of preprocess only occurs once.

Preprocess converts multiple column indexes to a single column, removes multiple paragraph marks, manual line feeds and single letter group headings. Tables are converted to text and any section, page and column breaks are removed. If commas are present preceded by two spaces then one space is removed. En dashes in locators can be replaced with hyphens and paragraph markers can be removed if it looks like locator strings have been broken.

The conversion of multi-column indexes to single column can result in entries being flowed together when a PDF to Word conversion doesn't judge that a CR is required. Preprocess can recognize these instances. A CR followed by a \$ErL0\_ label is inserted. Each labelled entry needs to be reviewed to ensure the entry levels are correct. Where an index is exported from PDF, manual checking and adjustment is recommended before running Preprocess. This will avoid having to remove labels inserted by Preprocess and ultimately save time.

Certain characters can confuse PreProcess. Error labels \$ErP01\_ are inserted and processing should continue. Adjust the labelled characters and remove the error label. Run Preprocess again.

## Locators L

Locators and see/see also cross references are identified as locators. Options exist to define alternative cross references such as s., s. a., voir, and locator prefixes for chapters, figures, plates, tables etc. The search is case sensitive.

Locators are identified by looking for characters, spaces and punctuation preceding the locator and the locator itself. Locators are evaluated; a maximum page number can be set in **Options** to prevent large numbers such as years being identified as locators. Many thousands of searches are performed automatically using an AI process.

A second pass identifies ambiguous entries and adds error labels beginning \$Er. These need to be corrected by searching for \$Er before labelling headings.

## Label Headings

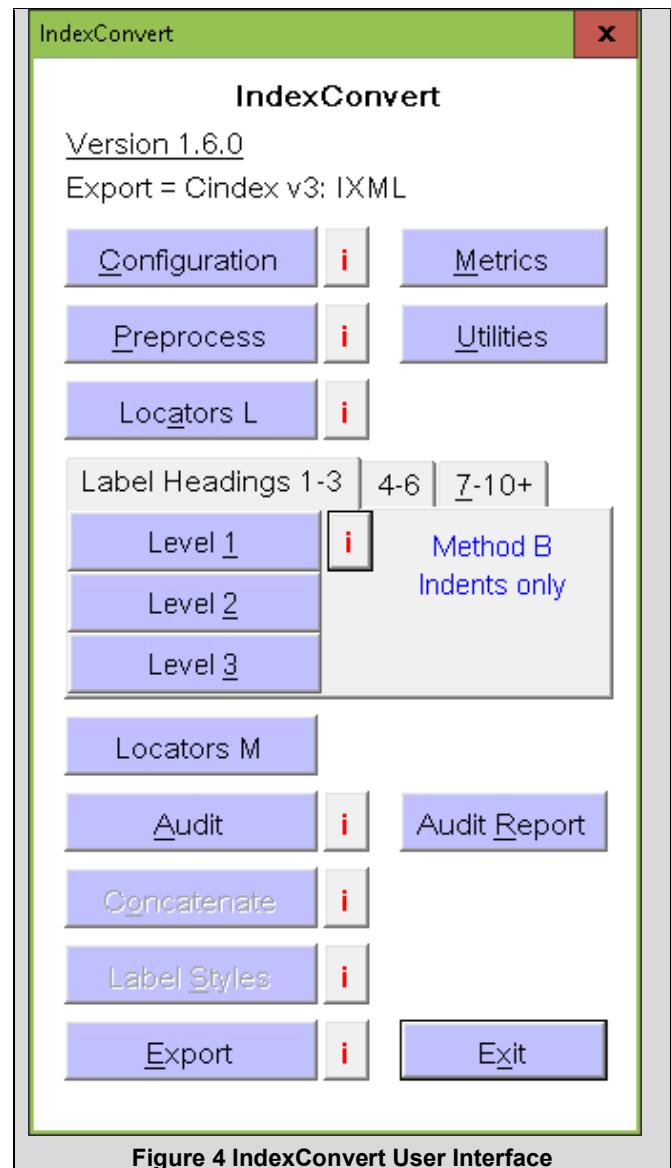


Figure 4 IndexConvert User Interface

Three methods are available for labelling headings.

**Method A** uses the settings of several style and indentation parameters. When a style is accepted all entries matching that style will be labelled. If a style is rejected it is remembered and will not be offered again. Method A is best for indexes exported from an indexing program. It should be tried if the index has been created manually or has been exported from a PDF. Multiple choice is available. When set to 1, labelling commences when that style is identified. If set to 10, labelling begins either when 10 heading styles have been identified, or no more styles can be matched.

**Method B** uses left indentations only. A large number of indentations can be accepted before labelling begins. This method is now rarely used.

**Method C** is interactive. A heading is selected in the Word document and the appropriate heading level is selected in the IndexConvert user interface. The heading is labelled and the next un-labelled heading is found. This method works for any index but is the best approach for really difficult indexes exported from PDF files.

Correct errors found by searching for \$Er before Locators M.

## Locators M

This is basically the same process as Locators L. Roman numerals are now numbered if any exist and cross references labelled as subheadings are labelled as locators. Options exist to include lower case or upper case Roman numeral locators.

Correct errors found by searching for \$Er before Audit.

## Audit

When heading and locator labelling is complete an audit process determines whether the labelling has been consistent and identifies entries that may have been labelled incorrectly.

Headings are audited first. If errors are found, these must be corrected before locators are audited. During locator audit, every line is checked for a locator, if one is not required, the line can be ignored. Lines containing suspect locators can be labelled and adjusted when audit is complete.

Commas in the locator string are automatically checked and adjusted to a normal style if necessary to prevent leakage of style markups during export.

An audit report can be created containing only the entries with \$Er error labels.

Correct errors found by searching for \$Er before selecting the next step.

## Elided

Elided is a separate utility which can be run following Audit. This converts elided page numbers to the un-elided form required by indexing programs. A dialog box with a red background acts as a reminder.

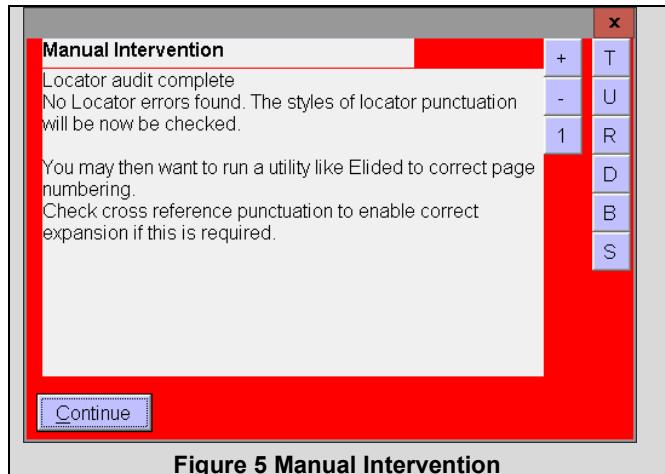


Figure 5 Manual Intervention

## Concatenation and Style Labelling

Concatenation and style labelling follows unless an IXML or IDEX export is chosen. For a tab delimited export such as for Cindex or Sky Index, each lowest level subheading is prepended by the parent headings, separated by tab characters to create a spreadsheet structure.

For a Macrex export underline characters “\_,” are inserted instead of the higher level headings to create the MBK structure.

Styles are labelled following concatenation. The styles identified are bold, italic, underline, small caps, subscript and superscript.

## Export

The file is exported to the selected file format. If a delimited export is selected the file is saved as a text file using suitable encoding. An IXML generator is used to create the IXML file. An IDEX generator is used to create the IDEX file. Index-Manager export is also available. An index can be embedded as XE fields into a chosen file. The user is notified when export is complete.

## Process Commonality

Preprocess, heading labelling, locator labelling and audit is the same for all exports.

Audit report, concatenate, label styles and export differ according to the selected export.

The export selected can be changed at any time up to and including the completion of the Audit process. If the file is saved at this point multiple exports are possible. The same index can be exported for Sky Index and Macrex or both an IXML file and a tab delimited file can be produced for Cindex.

## Iteration

It can be very useful to repeat some of the processes.

For an index with three heading levels the iteration could be as follows.

- a. Preprocess
- b. Locators L
- c. Label headings L1, Multiple choice set to 1, 5, or 10
- d. Label headings L2
- e. Label headings L3
- f. Repeat c,d,e
- g. Locators M

## Quality Considerations

IndexConvert contains many features designed to maximize the quality of the index structure in the Word document at the point of concatenation or export.

It is far easier to adjust the document before export than to adjust it once it is in an indexing program where errors will cause entries to appear in unexpected places. Keep a copy of the audited, corrected file and be prepared to make further adjustments and re-export if necessary.

# Index Preparation

## Microsoft Word

IndexConvert is designed to minimize the amount of preparation required. However some adjustments are needed.

Remove the title and any introductory comments to prevent them being labelled as headings.

If there are complex or numeric group headings then remove these.

If the document is exported from a PDF, check for layout problems and make corrections. This can be tedious but will save time later.

Run-on subheadings should be adjusted to set-out, take the opportunity to add labels at the same time.

## Indexing Software

Converting from one indexing software format to another is not always straightforward. IndexConvert may be the right solution when fonts and styles must be retained and has been used to convert from IXML (Cindex) to IDEX (Sky Index 8+). Adding the IndexConvert labels before export will save a lot of time. Use a set-out, single column layout, no continue clauses, no elision.

## PDFs

Every PDF seems to offer different challenges. Also, there is a wide range of PDF readers and exporters. Adobe Acrobat Pro DC has been used for testing IndexConvert. A few free exporters have been tried but found to be less satisfactory.

Adobe Acrobat Pro DC will export to a Word file. Two options to be selected are Retain Flowing Text and Recognize Text if Needed.

Retain Flowing Text attempts to keep text in the order it should be read.

Recognize Text if Needed invokes character recognition if required. A language should be selected.

The conversion may result in character sizes, indentations and other style features differing between entries by tiny fractions of a point (a point is 1/72 inch).

If the index is scanned then blots, smudges and dust can cause characters to be misread. These need to be adjusted manually.

During conversion, lines can be broken where index entries run on to multiple lines. Preprocess will rejoin lists of purely numeric locators.

Lines can also be joined resulting in multiple entries appearing as one.

## Index Layout (IXML Export)

The IndexConvert Heading and Locator labels are important for an IXML export. Deviation from a conventional back-of book style set-out index is unlikely to be successful if an IXML export is required.

## Index Layout (Delimited Export)

A delimited file is imported by an indexing program using the tab or comma delimiter to determine how each field is read. A good understanding of both IndexConvert and the Indexing program import process is recommended.

Label Locators is intended to process back-of-the-book indexes. There are certain indexes where Label Locators should not be performed. These include legal indexes where the locators appear as a set-out column of single locators. Labelling those as subheadings is likely to produce the required spreadsheet layout for export.

Biblical references such as

Genesis 6:1–4 **73**

and legal references such as

Abbey National Building Society v Cann [1990] 1 All ER 1085, HL **41**

present a problem for IndexConvert. To ensure IndexConvert is robust, first locators will not normally be recognized if they follow text that looks like a locator. Manual labelling of locators may be required.

## Locator Separators

IndexConvert and indexing programs require commas as locator separators. Other separators such as spaces or semicolons should be converted but check your software. The adjustment can be made at any stage before Concatenate for a tab delimited export and before Export for an IXML or IDEX export.

Cross references may need separating out. See instructions for your indexing program.

# Using IndexConvert

## Start IndexConvert

To start IndexConvert the file containing the index must be saved as a docm (macro enabled) file. This ensures that the advanced features used by IndexConvert will function correctly. An error message will be seen if an incorrect file format is used.

Select View>Macros>View Macros, select IndexConvert and Run.

## Shortcuts

Any macro can be assigned to a custom key combination by using Office Button>Word Options>Customise>Customise. Find Macros towards the bottom of the list. Select IndexConvert and enter a key combination to assign to the macro.

Some versions of Word allow a shortcut to be added to the quick access toolbar. Later versions allow them to be added to the ribbon.

Underlined characters in the user interface are shortcut keys used in combination with the Alt key. Use Alt-c for configuration. There are shortcut keys for responses in many of the dialogs. For these to work the pointer must be placed on the form.

## User Interface

The main user interface is shown in Figure 6. The **[i]** buttons supply further information and if appropriate, additional configuration settings.

## Version

Clicking on Version opens the Check for Updates dialog. If you are connected to the Internet, it opens the IndexConvert download page. Do this occasionally to check for updates.



Figure 7 Check for Updates

## Configuration

When IndexConvert is run for the first time a configuration screen appears. You must enter the licence key, name, email address and if appropriate, company name.

Choose the export appropriate to you. If your details have been entered correctly IndexConvert will run.

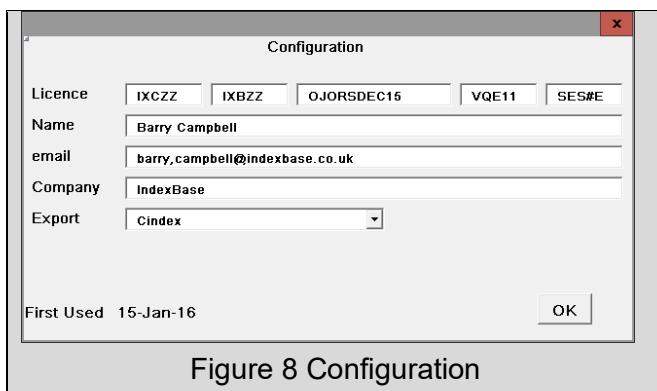


Figure 8 Configuration

If you have been supplied with a new or special licence then it can be entered by selecting Configuration.

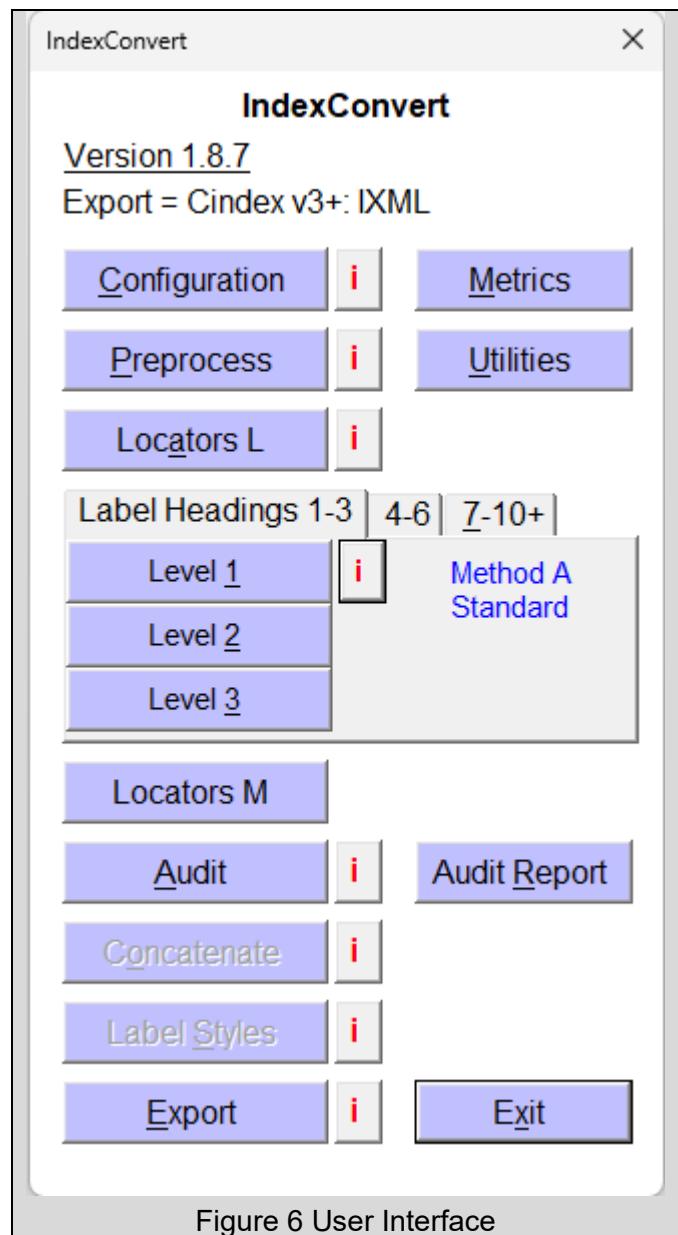


Figure 6 User Interface

## Select Export Type

The Configuration dialog is used to select the export type. The export type can be altered at any time. When an index is fully labelled it can be saved. Multiple exports can be achieved by returning to the saved file and altering the export type.

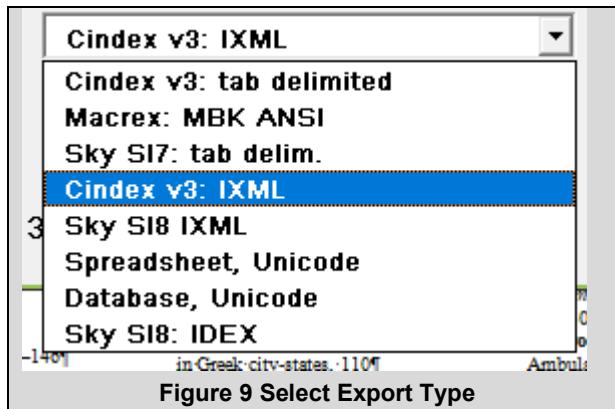


Figure 9 Select Export Type

## Progress Indicator

For complex processing likely to take time a progress indicator may appear showing a spot travelling round a circle.

The text at the top indicates the process; the text at the bottom reminds the user that interaction with the Word text during processing can cause processing to appear to stop. The numbers show progress. For Label Locators the dot completes a circle and the counter increments after 1000 Finds.

Status is also shown in the Word status bar at the bottom left of the screen. Note that Microsoft withdrew the ability of VBA to write to the status bar with Word 2010 but it seems to have reappeared with Word 2016.

**Warning:** If you interact with the text or swap applications when processing is taking place, screen updates may halt. Processing normally continues but progress is not obvious.

This can happen without any apparent cause but processing normally continues.

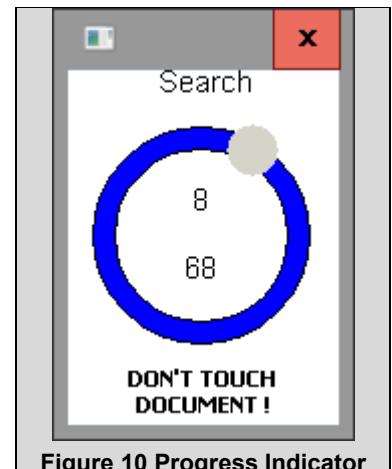


Figure 10 Progress Indicator

## Ctrl-Break / Esc

If you make a wrong choice and processing is already under way, Ctrl-Break will normally halt IndexConvert. Where there is no Break Key, Esc may work.

Apparently innocuous features in the text can cause processing to appear to halt. Pressing Ctrl-Break or Esc and Cancel will leave the cursor on the troublesome line. This has been experienced with IXML export.

## Preprocess

Before running Preprocess ensure your index is saved as a docm file.

If the index is run-on then look for rules you can use to alter the layout to set-out. IndexConvert tags can be added at the same time using carefully crafted Find and Replace statements.

Correct any obvious layout flaws. This is especially important when the index is exported from a PDF and can save a lot of time later.

Preprocess removes page, section and column breaks, double paragraph markers and group heading letters. It can also remove headers and footers and will convert a multi-column index to a single column.

Tables are converted to text. Page, section, and column breaks are removed.

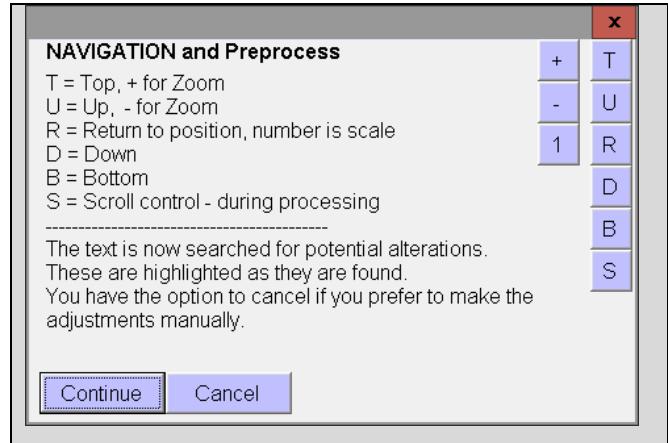
Preprocess can join broken locator strings to the previous line and entries run together can be separated. An error label \$ErP01\_ identifies these as they may end up at the wrong heading level and other entries may remain un-separated. Search for \$Er and correct errors before progressing to the next stage. These adjustments require time, many can be avoided by adjusting the layout before running Preprocess.

Preprocess is controlled by a dialog box like Figure 11. The control buttons at the bottom of the dialog box are the main process control buttons. Small buttons clustered to the right support navigation, screen updates and zoom.

**Table 1 Navigation and Scrolling**

+	Increase Word document zoom
-	Decrease Word document zoom
1	Increase dialog size (1 to 1.6, effective next time it is used)
T	Top of file
U	Up 5 lines (depends on zoom)
R	Return to original position
D	Down 5 lines (depends on zoom)
B	Bottom of file
S	Control screen updates

Scrolling (screen updates) can be turned off during processing using the S button. This will speed up the processing until the next user interaction is required.



**Figure 11 Example Dialog Box**

During operations taking a long time, if screen updates have not already been cancelled, a dialog box appears after a few seconds and allows S to be selected.

**Warning** When screen updates are prevented the progress meter is also disabled!

An example alteration is shown in Figure 12.

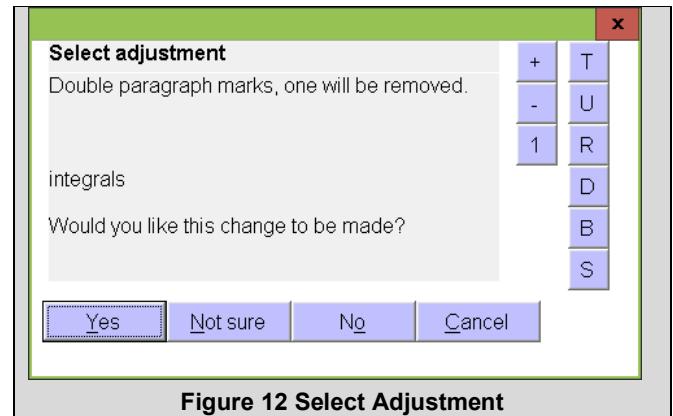


Figure 12 Select Adjustment

If you encounter a message similar to that shown in Figure 13 then select **Label** to mark the location in the document. If a \$ErP02\_ label is already present then select **Cancel**, make the correct adjustments manually and remove the \$ErP02\_ error label. Repeat Preprocess until no errors are encountered.

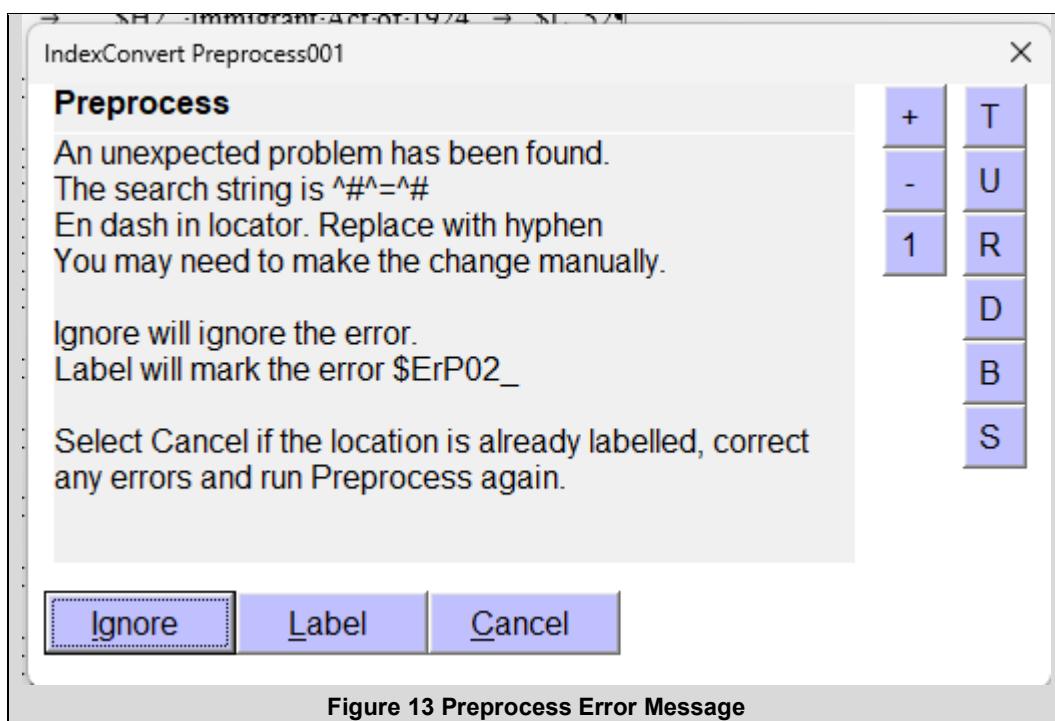


Figure 13 Preprocess Error Message

When running Preprocess for the first time a highlighted dialog box like Figure 14 may appear as a reminder that a pure set-out structure is required for successful processing. This is shown as a result of a Word search for

$([a-z]\{3\})(, )([a-z]\{3\})$

with wildcards selected. This searches for three lower case letters a-z (group /1) followed by a comma and space (group /2) and another three characters (group /3). This is looking for commas that may indicate a subheading requiring adjustment.

Replace string

$\backslash1^p^t\$H2\_3$

is preloaded by IndexConvert but not automatically used. This replaces group 2 with a new line followed by a tab and a level 2 heading label. If suitable then modify the find and replace strings as necessary.

## Label Locators

Locators are labelled before and after Headings are labelled.

Processing can be stopped at many stages. The IndexConvert dialog can be closed and restarted when required. This allows the index to be adjusted if necessary.

The first pass (Locators L) excludes Roman numeral locators and cross references shown as subheadings. The second pass (Locators M) labels cross references shown as subheadings, and Roman numeral locators if selected.

Two passes are performed with each locator definition. The first pass labels the first candidate in the entry with \$L\_. The second pass determines whether another candidate in the entry could be the first locator. If this is possible then the entry is labelled \$Er01\_. This notifies the user that the {tab}\$L\_ locator label may be in the wrong position and should be checked. The \$Er labels must be removed before continuing.

This feature identifies many run-on subheadings and joined entries resulting from PDF export.

## Locator Options

Selecting the [ i ] button opens a dialog box allowing the processing of Roman numeral locators, additional cross references and locator prefixes.

Cross references and locator prefixes will be found only if they match a value entered in the grid.

**Search times:** Search times increase as additional locator options are selected. Minimize the Roman numeral range or deselect Roman numerals and minimize cross reference types to speed processing.

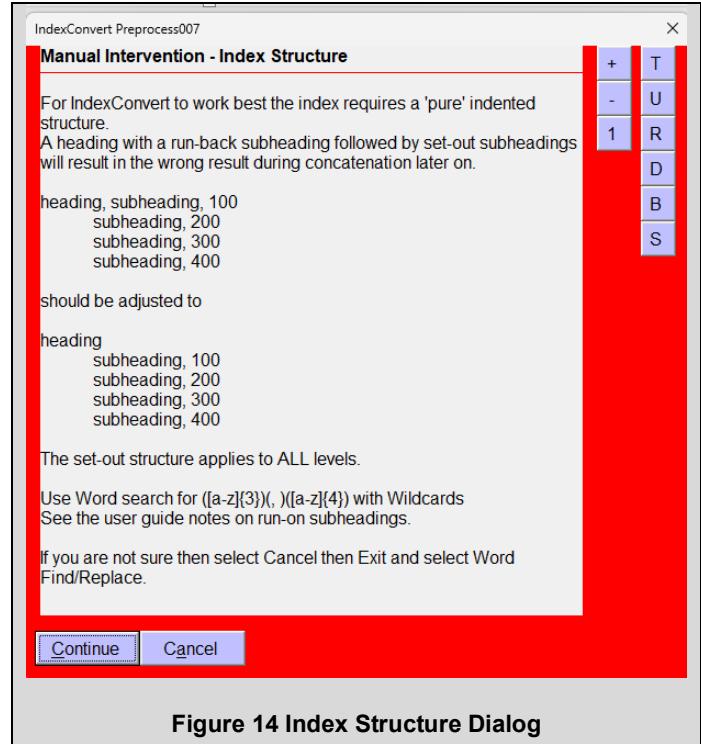


Figure 14 Index Structure Dialog

## Page Number Limit

Numbers within headings present particular difficulties because they may be identified as locators. A page number limit can be set to prevent large numbers being identified as locators. If no value is entered the default is 999. When Label Locators is selected the dialog Figure 15 appears as a reminder to set the correct value. If the value set is less than the maximum page number in the index then higher pages will not be labelled. Reset the page number limit and run again.

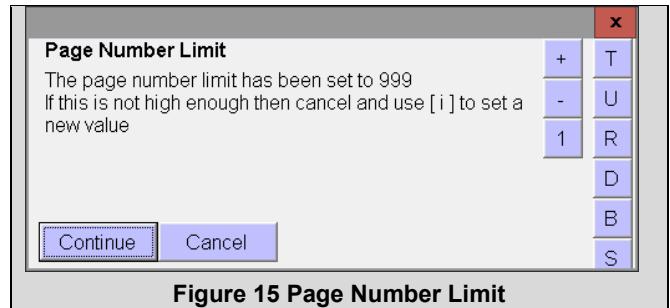


Figure 15 Page Number Limit

## Roman Numerals

If Roman numerals are not used or are unusual then do not select this option. The default configuration at installation is Roman numerals unselected.

Roman numerals in the range i to l (1 to 50) may be selected. The value may be expressed as either a Roman or decimal number. Don't enter a number higher than needed because Roman numeral identification is very complex and search times are extended.

Select **Upper Case** if upper case Roman numerals are used in the index. this selection is not remembered between sessions.

## Other Locators

IndexConvert will label the following locator types:

www.

http:

https:

## Automatic Locator Search

Selection of Automatic search causes IndexConvert to answer Yes automatically to the question 'Is this a First Locator'. Use with care! The default is interactive.

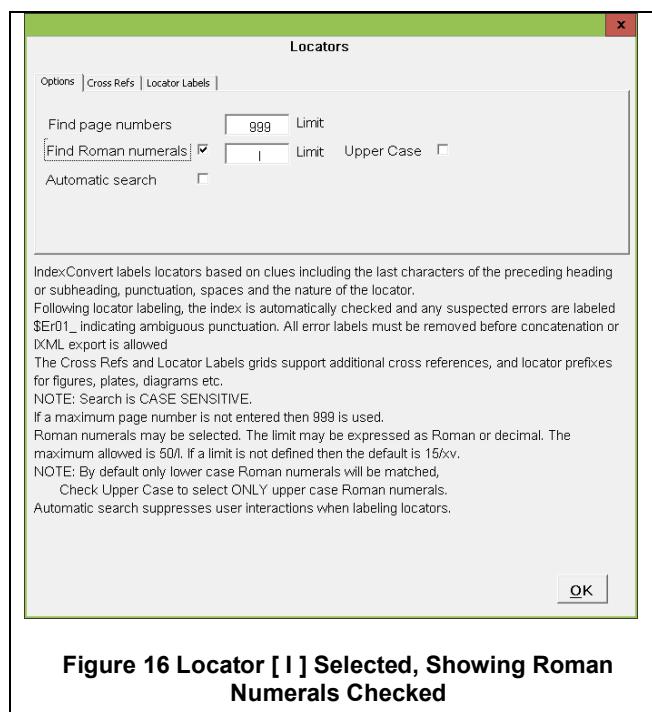


Figure 16 Locator [ I ] Selected, Showing Roman Numerals Checked

## Cross References

Cross references are considered to be locators and are processed during Label locators. The search is case sensitive.

For run-on cross references such as

**Click to Look Inside. See Search Inside**

The cross reference term will be offered as a first locator and should be accepted.

Cross reference terms must be defined in the Cross Refs grid. The large panel to the right is a scratch pad. This can be used for transferring a list of terms from another document. It is only a storage area and is not read during processing. Values entered here are retained between sessions. Unicode font is supported. Table 2 includes possible cross reference terms. For English language indexes it is recommended that see and (see are always used. When (see is used, IndexConvert will recognize (see as a cross reference but does not remove the enclosing (...)).

The grid contains two columns, one reserved for 'see' type references, the other for 'see also'. Cross reference processing is case sensitive.

**Table 2 Example Cross Reference Terms (Lower Case)**

English	see	(see
German	sehen	(sehen
French	voir	(voir
Spanish	ver	(ver
Portuguese	vejo	(vejo
Norwegian	Se	(se
Danish	se	(se
Swedish	se	(se
Finnish	nähdä	(nähdä
Dutch	Zie	(zie

For set-out cross references such as

Architecture

**See also Housings; Pyramids**

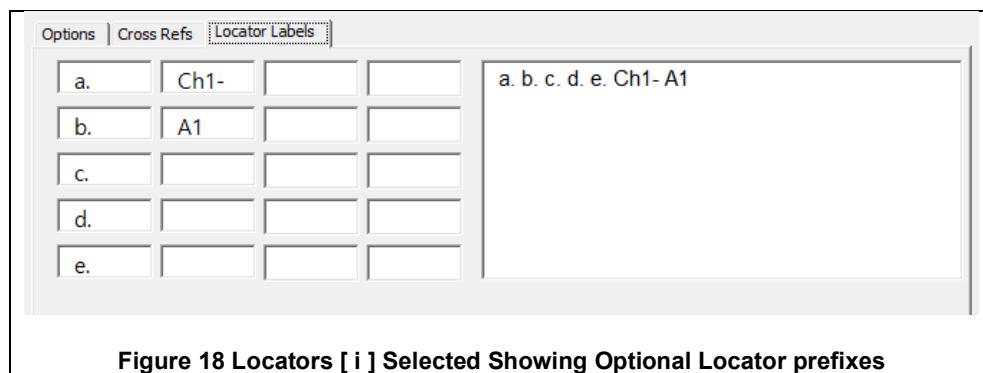
The See also will first be labelled as a subheading during heading labelling then labelled as a locator during locator labelling. There should be no space in the grid following a cross reference entry.

'see'	'see also'	
see	see also	siech auch
(see	(see also	
See	See also	
(See	(See also	

**Figure 17 Locators [ i ] Selected Showing Cross Reference Matrix**

## Locator Prefixes

Locator prefixes can be used to address chapters, diagrams, figures, plates etc. The search is case sensitive, this should be considered when setting values. A scratchpad is available in the panel at the right. If there are spaces following the prefix then these must be included in the grid. Values are retained between sessions. If more prefixes are required, maybe for a book with more than 20 chapters, then repeat label locators with a new set of prefixes. Do not run Audit until all prefixes have been processed. Where a prefix is followed by a space, this must be included in the grid entry.



The example below shows the result of locator labelling.

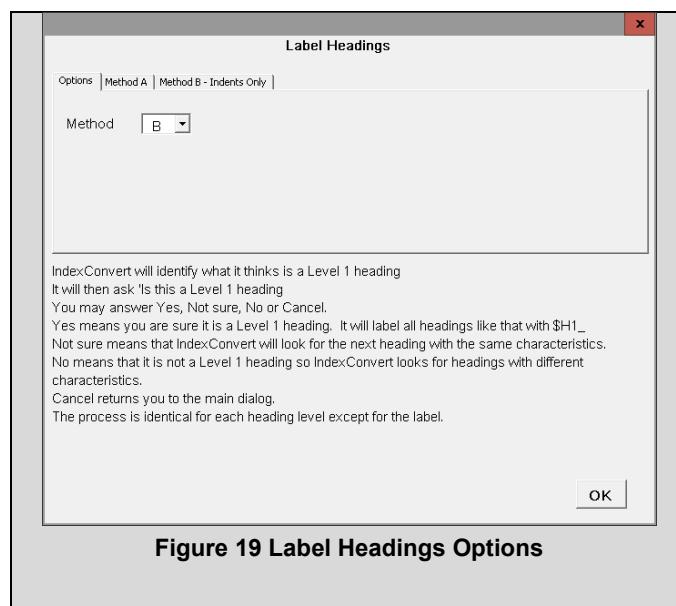
```
$H1_heading $L_1
      $H2_subheading      $L_2
      $H2_subheading      $L_3
      $H2_subheading      $L_4
```

## Label Headings

IndexConvert will label up to 32 heading levels. Cindex supports 15 levels, Sky Index and Word embedded support 7 levels. Export to spreadsheet or database supports 32 levels, enabling transformation of very complex data unlikely to be related to indexing.

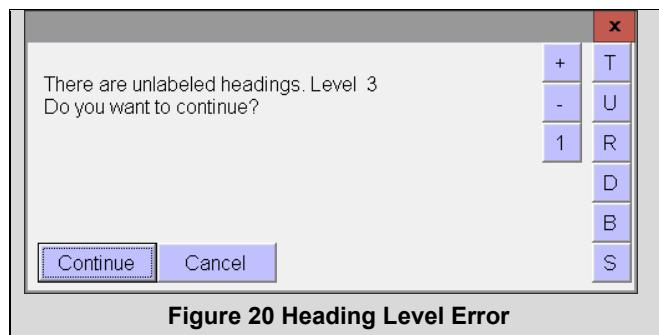
### Heading Options

There are three approaches to labelling headings, Method A, Method B, and Method C. Select Label Headings [1], Options to choose a method.



When labelling headings, groups of 10 are identified with the option to double the number at each user intervention, run to the end, or cancel.

If a Heading level is selected and higher levels have not been labelled then an error message is displayed. See Figure 24.



The result of label headings is indicated below.

\$H1\_heading

\$H2\_subheading a

\$H2\_subheading b

\$H2\_subheading c

## Method A

Method A uses several text formatting clues, these include tabs, spaces, indents and character size. First, select a level to label. When Level 1 is selected the first unlabelled entry is found. The user is then asked 'Is this a Level 1 Heading'. If the answer is Yes the style information is stored and entries matching the stored data will be labelled.

Up to 100 styles at the same level can be identified before labelling begins. Use Headings [1] and select Method A to set the multiple choice limit. Convenient selections are 1, 5, and 10, depending on the index.

If multiple choice is set to 1 then labelling begins immediately after the style is chosen.

When all Level 1 entries have been labelled proceed to Level 2.

Processing can be stopped at many stages. The IndexConvert dialog can remain visible or closed. This allows the index to be adjusted if necessary.

## Heading Levels

Select the heading level to be labelled. Figure 20, Figure 21 and Figure 22 show available selections.

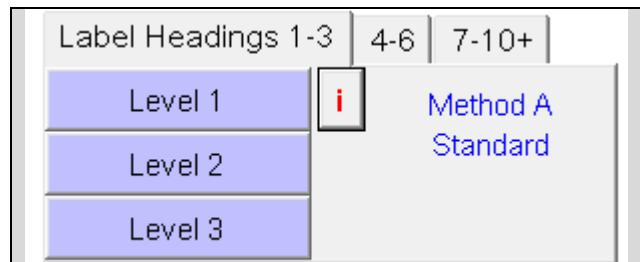


Figure 21 Label Headings, Levels 1-3

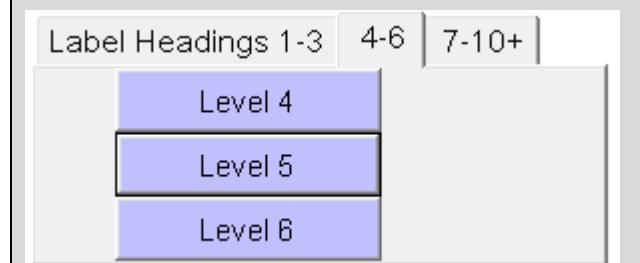
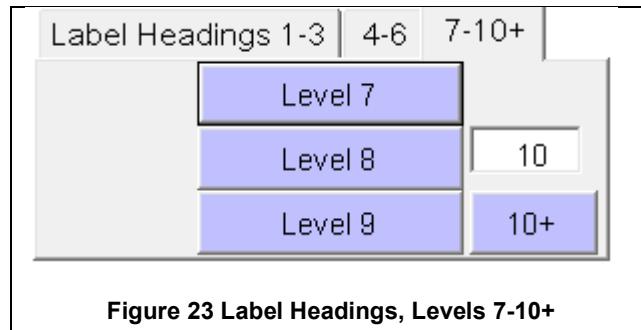


Figure 22 Label Headings, Levels 4-6

Lower level headings can be entered as a number and the 10+ button selected. Maximum 15 is recommended for Cindex, 7 for Sky.



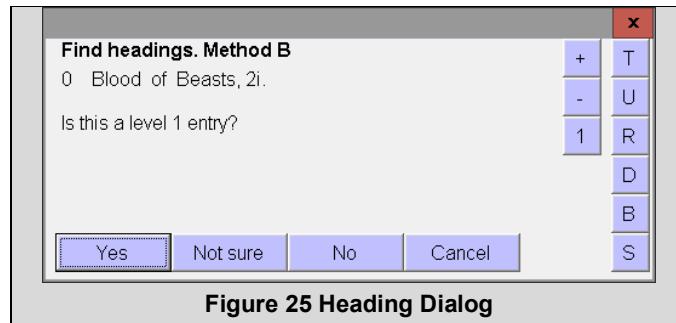
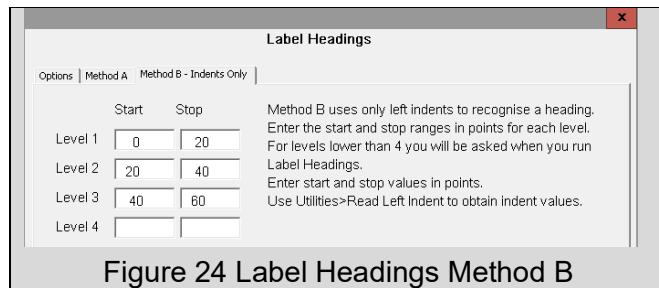
If Level 1 is selected, when a match is found the user is asked 'Is this a Level 1 Heading'. If Yes is chosen then headings with that style will be labelled with \$H1\_. If Not Sure is chosen then the entry is ignored. If No is chosen then the style is stored and a new style is sought.

## Method B

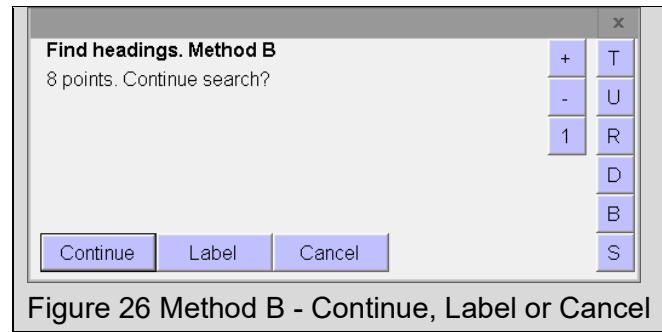
Method B searches only for left Indents within a predefined range, stepping gradually from the minimum to the maximum. When a match is found the user is asked 'Is this a Level 1 Heading?' If the answer is Yes then this is remembered and another indent is sought. Up to 100 different indents can be identified in a single pass.

Label Headings [i] allows left Indent ranges to be set. If they are not set then a default of 0 to 200 points is selected.

This method is used less now, as Method A has been made more capable and Method C has been introduced.



At intervals of approximately 10 seconds the user is given the option to continue searching, label headings or cancel. The selected indent value is shown for information.



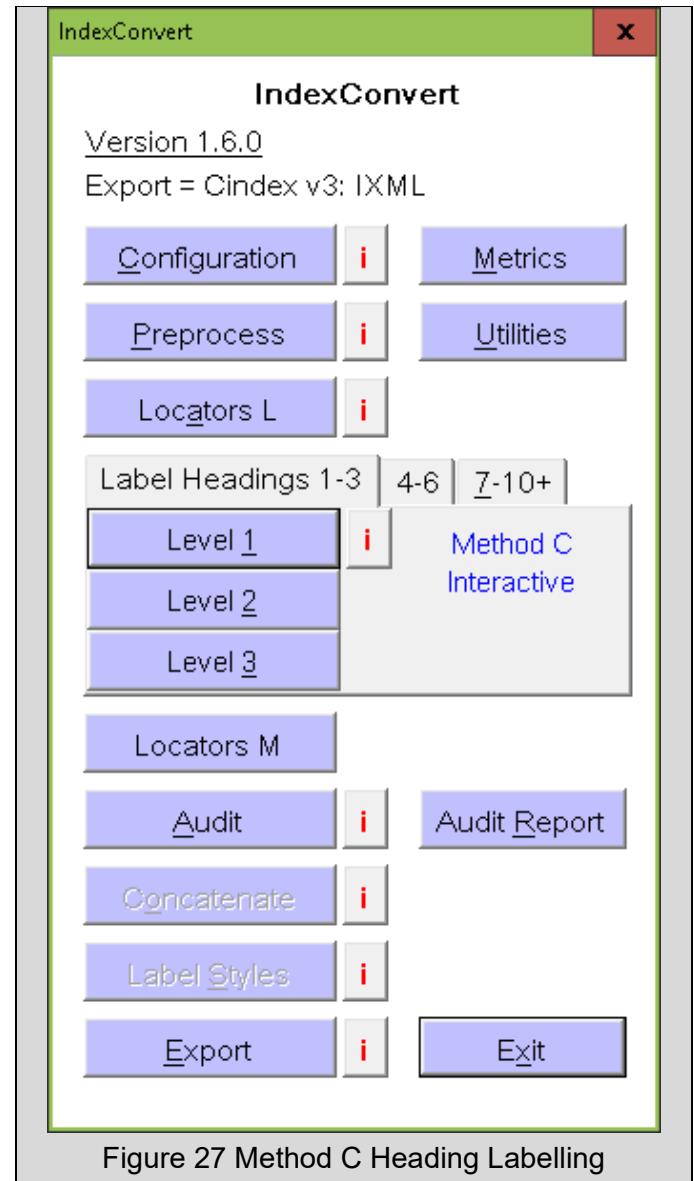
Processing can be stopped at many stages. The IndexConvert dialog can remain visible or closed. This allows the index to be adjusted if necessary.

### Method C

Method C labels one heading at a time. Select an entry in the Word document and a level at the IndexConvert user interface and it is labelled. The next unlabelled entry is then chosen. If a heading is already labelled, select any level at the IndexConvert user interface and the next unlabelled heading is found.

Method C is recommended for very complex and difficult indexes, for instance, indexes exported from a PDF file. It is also recommended where a forensic analysis and familiarization with the index is desired.

Processing can be stopped at many stages. The IndexConvert dialog can be closed and restarted when required.



## Audit

Audit checks consistency of the index labelling.

Headings are audited first. If no errors are found, then locators are audited.

Heading audit looks for:

- a. Unlabelled headings
- b. Broken hierarchy like a level1 heading followed by a level 3 heading
- c. Entries labelled twice

If errors are found during heading labelling then locator audit will not take place.

Locator audit looks for:

- a. Unlabelled locators
- b. Incorrectly labelled locators (missing tab)
- c. Unlabelled cross references
- d. Locators containing unexpected spaces
- e. Locators containing semicolons

Audit is interactive. If there is no fault present then it should not be labelled. Audit expects a locator with every entry. Where locators exist against lower level headings a locator may not be present for the main heading. The entry can be ignored during audit with the result that no error label is added. This prevents unnecessary rework later.

Audit adjusts commas in locator strings to normal text to prevent leakage when style labelling takes place.

All errors should be corrected following audit.

## Audit Report

The audit report contains all the entries labelled with a \$Er label. This can be useful for reviewing the status of the index and when support is required.

The last entry processed by the Label Headings, Label Locators, Concatenate and export processing is included in case there is a break in processing.

Roman numeral locator, custom cross reference and custom locator prefix settings are included.

A special audit report containing the translation codes is created following Macrex and Sky Index (SI7) concatenation. See the Macrex and Sky Index sections.

**Warning:** The audit report uses the Windows clipboard. Do not perform any copy/paste operations in other applications while audit report is in progress.

If the index file is read-only then the audit report will not complete.

## Manual Labelling

### Headings

There may be occasions when manual labelling of headings is a preferred option.

As an example a user had an index with two levels of heading. When copied and pasted from a PDF the level 2 headings collapsed to the same level as the level 1 headings. Fortunately, all level 1 headings began with a capital.

The approach suggested was to label the level 1 headings manually with help from a carefully constructed Word find/replace statement. Anything left will be level 2 headings. IndexConvert can then be run to label the level 2 headings.

If there are three levels of heading the best approach will be to label level 1s (capitals) then level 3s (not too many) then use IndexConvert to label level 2s.

Some users format the index before using IndexConvert using styles \$H1\_, \$H2\_ etc or equivalent.

## Locators

Most labels are added automatically. However, not all first locators will be recognized. This is deliberate and prevents too many features being wrongly labelled as locators. Biblical references and legal tables may cause difficulty.

Headings and locators can be labelled manually and this will be required normally following Audit but can be done at any stage if convenient.

## Elided Locators

If elided page ranges are present then use utility Elided 123-4 to expand them to un-elided values. See [Installation](#).

## Concatenate (Prepend)

Concatenate works through the audited hierarchies of headings, concatenating Level 1, Level 2 etc. Tabs are inserted between levels for Cindex, Sky Index, spreadsheet and database export. For Macrex higher level headings are represented by underscores, commas are inserted to separate the levels. IXML, IDEX and Index-Manager export do not require concatenation and the option is greyed out.

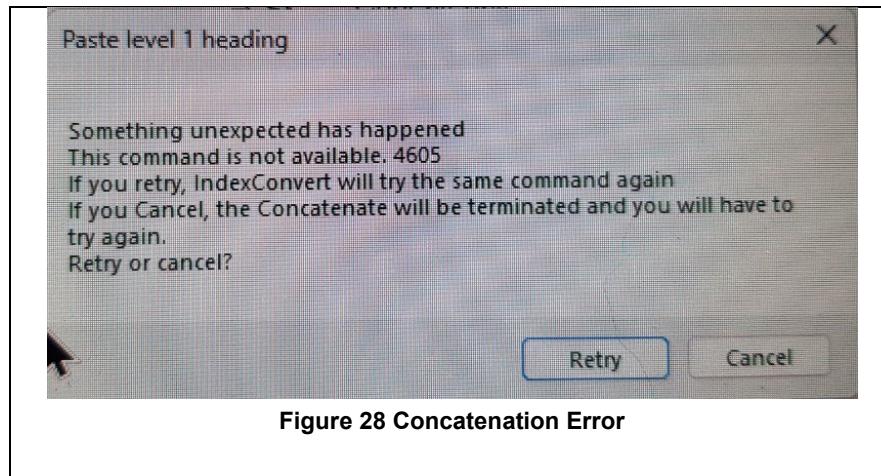
During concatenation the user is asked at regular intervals whether screen updates should be prevented (option S in the dialog). When S is selected processing proceeds faster and the dialog no longer appears.

The end of file label (idxEND) is removed at the end of Concatenate

If processing is interrupted for any reason concatenation can be restarted and will pick up where it left off.

## Processing Speed and Timing

IndexConvert can process data faster than Word expects. Testing in January 2024 by Walter Greulich, an indexer based in Germany, exposed a run-time timing issue with Concatenate which caused IndexConvert to terminate. This was addressed for IndexConvert version 1.8.5. Various provisions were included to try and find this error which had only ever been reported on Walter's computer. Enhanced debugging was introduced. Debug code can be activated by inserting \$DebugOn\_ into an entry. Debugging is stopped when \$DebugOff\_ is encountered. The debug code lists the steps taken during processing. A report is created when Concatenate is finished. During this testing the error was found to be caused by IndexConvert working faster than Word expects. The code was modified to address this. An error message like Figure 28 will be displayed when the timing problem occurs. Select Retry. The time between the error being displayed and Retry being selected will allow Word to catch up.



## Unicode Characters: Sky Index v7

For Sky Index v7, if a Translation table is present, and has been exported, and is selected during Concatenate, then any characters defined in the Translation table are replaced by their keyboard entry codes. If a character is encountered that should be in the translation table then the first occurrence is labelled {\$Ernnnn} where nnnn is the decimal value of the Unicode character. The character and its hex value are included automatically in an Audit report.

## Unicode Characters: Macrex

The process is similar for Macrex except that the file referenced is the Unicode.ini file.

## Label Styles

Bold, italic, underline, small caps, subscript and superscript strings are labelled. \$BA\_....\$BZ\_ labels mark the beginning and end of bold text.

This step is not necessary if an IXML, IDEX or Index-Manager export is selected, the option is greyed out.

## Export

This saves or exports the processed index to a file with encoding suitable for your indexing program. The file name will be the same as the last saved name but with an extension depending on the export format chosen. An exception to this rule is for Word embedding where a target file is chosen.

Before importing a concatenated file into your indexing program inspect it and remove any blank lines from the end. If you have not removed headers and footers completely then they will appear at the end of the file and should be removed.

If you are doing a conversion for a client who uses indexing software different from your own you can run an export for your own software or a spreadsheet as a check.

## File Extensions and Encoding

The table below shows the file extensions and encoding for the different exports.

Table 3 Export file encoding and extensions

Type	Export	File Extension	Encoding
Cindex	Txt		UTF-8 Unicode
Spreadsheet	Txt		UTF-8 Unicode
Macrex	Mbk		Windows ANSI
Sky Index v7	Txt		Windows ANSI
Database	Txt		UTF-8 Unicode
Cindex IXML	ixml		UTF-8 Unicode
Cindex IXML expanded	ixml		UTF-8 Unicode
SI8 IXML expanded	ixml		UTF-8 Unicode
Sky Index v8 IDEX	idex		IDEX Unicode
Index-Manager	Txt		UTF-8 Unicode
Word AutoMark File	doc*		Word .doc
Word embedded	docx		Word docx

**Warning:** If any blank lines remain in the index then IXML and IDEX export will not complete. A warning is given and export can be repeated.

\*Word AutoMark files are created as docx files but must be saved as doc files for use by Word.

## Exit

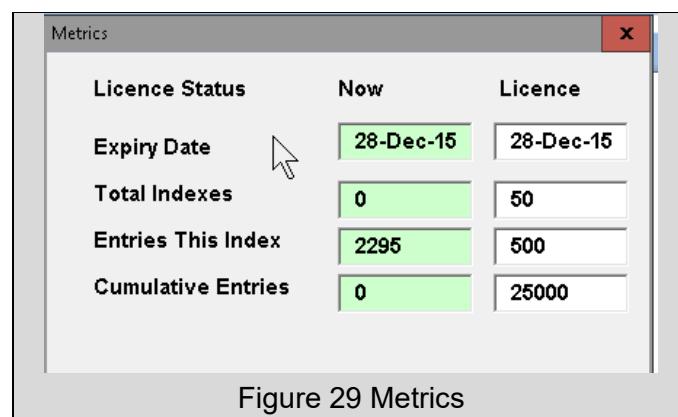
This exits IndexConvert. IndexConvert is designed to allow frequent starting and stopping as the processing continues. This allows for adjustment to the index as processing proceeds. Adjustments are required to address any error labels and may be required to address typos etc.

## Metrics

Licence keys are supplied to match user needs. They contain information controlling the number of indexes and the number of entries that can be processed. They also contain an expiry date. The values encoded by the license key are displayed by the Metrics dialog.

See the IndexConvert brochure for more details on license keys.

Select Configuration to enter a licence key.



Usage metrics are updated only during preprocessing. When Preprocess is run for the first time on a new index an end of file marker is added to the end of the file and several internal parameters are reset. It is important to save the file following preprocess to prevent metrics being incremented again. See Licences for more information.

## Utilities

IndexConvert includes TextConvert and CharConvert that simplify entry of non-ANSI text and diacritic characters into Macrex and Sky Index.

Conversion requires a valid Sky Index Translation file or a Macrex Unicode file to be referenced by the IndexConvert Configuration.

Sky Index v8 (SI8) supports Unicode so Translation is not required.

### TextConvert

**TextConvert** converts selected text to Macrex or Sky Index (SI7) format according to the selected export format. The file containing the text must be a docm file. The converted text can be copied and pasted into your indexing program.

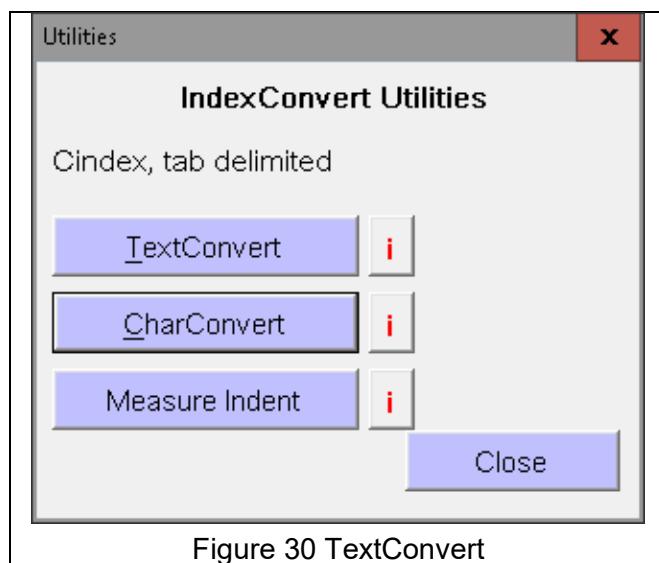


Figure 30 TextConvert

### CharConvert

CharConvert converts individual characters. Place the cursor to the left of, or over a single character and select CharConvert. The hex value, decimal value and converted character are displayed. Any of the results displayed can be copied and pasted to another document.

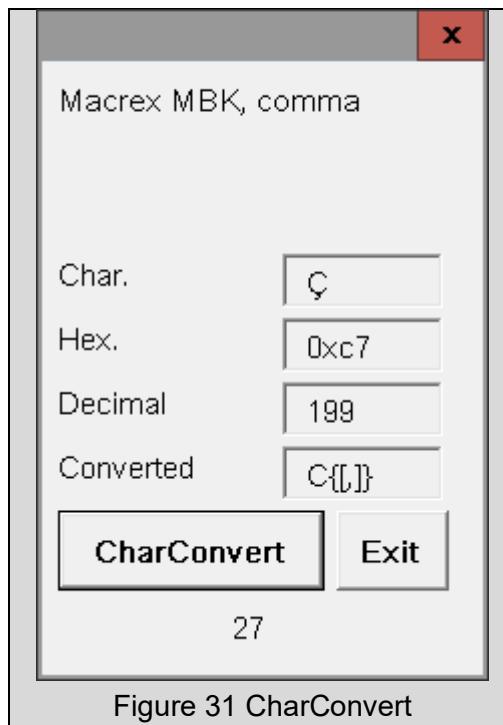


Figure 31 CharConvert

## Measure Indent

Measure Indent measures a left indent for use with setting limits when using Label Headings Method B. Place the cursor on a line and select Measure. The result is displayed as shown below.

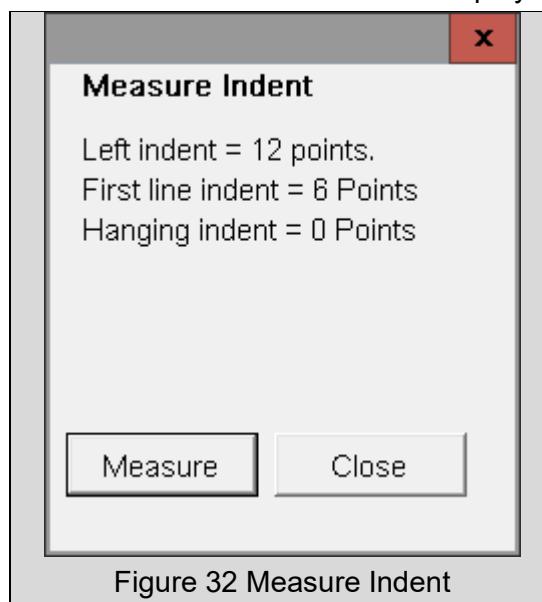


Figure 32 Measure Indent

## Associated Add-Ins

### Elided 123-4

IndexConvert does not expand elided ranges, this needs to be done either manually or using Elided 123-4. Elided 123-4 returns elided page ranges to their un-elided form. It expands page numbers in entries when a \$L\_ locator label is present and a \$Er error label is absent. Elided 123-4 processing is limited to 'pure' numbers. Locator ranges containing letter prefixes or suffixes will not be processed. 100-9 will be replaced by 100-109 but p100-9 will not be changed. Values to the left of the \$L\_ label are not processed.

A range of elided macros will be developed as the need emerges. Information can be supplied if you need to develop your own.

## Data Structures

The table below describes the data structures resulting from processing by IndexConvert.

**Table 4 Data Structures**

Selection	Export Format
<b>Cindex v3 Tab Delimited</b>	<p><b>Fields</b> are separated by tabs, the last field contains locators. The \$H and \$L labels are used to identify the fields. They are retained in the exported file to give extra assurance that all fields end up in the right place. Styles are labelled.</p> <p><b>Padding</b> There are no entries in the table for missing lower level subheadings.</p> <p><b>Styles</b> are identified by start and end labels, see Styles for more information.</p> <p><b>Fonts</b> are not identified</p> <p>Each record may contain more than one <b>locator</b>. Cindex always treats the last field as a locator.</p> <p>Export is a UTF-8 <b>Unicode</b> file.</p>
<b>Spreadsheet Tab Delimited</b>	<p><b>Fields</b> are separated by tabs, the last field contains locators. The \$H and \$L labels are used to identify the fields. They are retained in the exported file to give extra assurance that all fields end up in the right place. Styles are labelled.</p> <p><b>Padding</b> Missing lower level subheadings are padded with appropriate \$H labels.</p> <p><b>Styles</b> are identified by start and end labels, see Styles for more information.</p> <p><b>Fonts</b> are not identified</p> <p>Each record may contain more than one <b>locator</b>.</p> <p>Export is a UTF-8 <b>Unicode</b> file.</p>
<b>Macrex MBK</b>	<p>Heading and Locator <b>fields</b> are separated by commas. \$H Heading labels and \$L_ locator labels are removed. All style labels are replaced by Macrex equivalents. Where headings contain commas these are enclosed in braces {}, (soft commas).</p> <p>There is no <b>padding</b> for lower level subheadings.</p> <p><b>Styles</b> are identified by start and end labels, see Styles for more information.</p> <p><b>Fonts</b> are not identified</p> <p>Each record may contain more than one <b>locator</b>.</p> <p>Export is an <b>ANSI</b> file defined by the Windows code page. Unicode characters outside the ANSI set, and Greek characters may be defined as printer replacement codes including [alpha], a{[v]}. These are defined in a Unicode.ini table supplied with Macrex. IndexConvert uses this to expand the codes during the concatenation process.</p>

Selection	Export Format
<b>Sky Index v7Tab Delimited</b>	<p><b>Fields</b> are separated by tabs, the last field contains a single locator. The \$H and \$L labels are used to identify the fields. They are retained in the exported file to give extra assurance the all fields end up in the right place. Styles are labelled.</p> <p><b>Padding</b> Missing lower level subheadings are <b>padded</b> with \$H labels.</p> <p><b>Styles</b> are identified by start and end labels, see Styles for more information.</p> <p><b>Fonts</b> are not identified</p> <p>Each record contains one <b>locator</b>. Presence of commas within a locator string causes a new record to be created with the next locator. Headings from the previous record are replicated.</p> <p>Export is an <b>ANSI</b> file defined by the Windows code page. Unicode characters outside the ANSI set are replaced by translation codes using a copy of the user produced translation table.</p>
<b>Database Tab delimited</b>	<p><b>Fields</b> are separated by tabs, the last field contains locators. The \$H and \$L labels are used to identify the fields. They are retained in the exported file to give extra assurance the all fields end up in the right place. Styles are labelled.</p> <p><b>Padding</b> Missing lower level subheadings are <b>padded</b> with \$H labels.</p> <p><b>Styles</b> are identified by start and end labels, see Styles for more information.</p> <p><b>Fonts</b> are not identified</p> <p>Each record contains one <b>locator</b>. Presence of commas within a locator string causes a new record to be created with the next locator. Headings from the previous record are replicated.</p> <p>Export is a UTF-8 <b>Unicode</b> file.</p>
<b>Cindex IXML</b>	<p>An XML file is exported that is compatible with a IXML index records DTD supplied by Leverage Technologies. The DTD is supported by Indexing Research for Cindex.</p> <p><b>Fields</b> are identified by XML markup.</p> <p>There is no <b>padding</b> for lower level subheadings.</p> <p><b>Styles</b> are identified by XML markup.</p> <p><b>Fonts</b> are identified using XML markup.</p> <p>Each record may contain more than one <b>locator</b>.</p> <p>Export is a UTF-8 <b>Unicode</b> file.</p>
<b>Cindex IXML Expanded</b>	<p>As for Cindex IXML except that a record is written for each locator and a copy of the locator is placed in the lowest level subheading. If there is no locator an asterisk '*' is placed in the lowest level subheading.</p> <p>When Cindex is configured correctly a copy of the locator appears in the special subheading.</p>
<b>Sky Index SI8 IXML</b>	<p>As for Cindex IXML except that a record is written for each locator. Some versions of SI8 prior to May 2016 support this format.</p>

Selection	Export Format
<b>Sky Index SI8 INDEX</b>	<p>An IDEX file is exported that is compatible with the Sky Index IDEX specification. See <a href="http://www.sky-software.com/idex">www.sky-software.com/idex</a></p> <p><b>Fields</b> are identified by IDEX markup.</p> <p><b>Styles</b> are identified by IDEX markup.</p> <p><b>Fonts</b> are identified using IDEX markup.</p> <p><b>Locators</b> and <b>Cross references</b> are identified using IDEX markup.</p> <p>Each record may contain only one <b>locator</b>.</p> <p>Export is a UTF-8 <b>Unicode</b> file.</p>
<b>Index-Manager</b>	<p>A text file is exported that is compatible with Index-Manager.</p> <p><b>Locators</b> are identified by Index-Manager markup.</p> <p><b>Styles</b> are identified by using b for bold, I for italic, u for underline, s for small caps, + for superscript, - for subscript.</p> <p><b>Fonts</b> are not identified.</p> <p><b>Cross references</b> are identified as locators.</p> <p>Each record may contain more than one <b>locator</b>.</p> <p>Export is a UTF-8 <b>Unicode</b> file.</p> <p>See also p49</p>
<b>Word AutoMark File</b>	<p>A three column table is created if one doesn't already exist in the open Automark file. A row is created for each entry. Column 1 contains the lowest level heading, column 2 contains the heading and any subheadings, each separated by colons, locators are placed into column 3 for reference.</p> <p>Please note that very few indexes are compatible with successful AutoMark files. The facility is included in the rare event that this may prove useful.</p>
<b>Word Embedded</b>	<p>XE fields are created in the target document. Insertion point XE fields are placed close to the top of the page defined by the locator. Provision is included to minimise the risk of XE fields being placed in updateable fields. Locations of ranges are determined by the locators to recreate the source index.</p>

# Labels used by IndexConvert

## Headings and Locators

Index content is labelled to ensure that headings and locators are correctly identified. A common approach is employed independent of the final export format.

**Table 5 Content Labels**

Text	Label
Headings	\$H1_ to \$H15_
Locators	{tab}\$L_

Table 6 below shows a sample before and after labelling. No style or indentation adjustments are required. All spaces between a heading and the first locator are replaced by a tab character. En dashes can be replaced by hyphens. Once labelling is complete the file should be saved. Any one of the many export files can then be created. By returning to the saved file several export files can be created.

**Table 6 Labelling example**

Before labelling	After labelling
management accounting 3–4, 925 design of systems 13–16 emerging issues 774–7 force of change 18–21 information needs 18 and marketing management 757 and strategic decisions 5–6 management by exception 9, 494, 925	\$H1_management accounting \$L_3-4, 925 \$H2_design of systems \$L_13-16 \$H2_emerging issues \$L_774-7 \$H2_force of change \$L_18-21 \$H2_information needs \$L_18 \$H2_and marketing management \$L_757 \$H2_and strategic decisions \$L_5-6 \$H1_management by exception \$L_9, 494, 925

## Styles

This section does not apply to IXML, IDEX or Index-Manager exports.

Styles are labelled for Cindex, Sky Index, spreadsheet and database export using a similar style of labelling as is used for headings and locators. For Macrex export the Macrex labels are used.

**Table 7 Style Labels**

Text	Cindex Spreadsheet Sky Index Database	Macrex
Bold	\$BA_ <b>bold</b> \$BZ_	\b <b>bold</b> \
Italic	\$IA_ <i>italic</i> \$IZ_	^ <i>italic</i> ^
Underline	\$UA_ <u>underline</u> \$UZ_	{[UL]}underline{[ul]}
Superscript	\$+A_ <sup>superscript</sup> \$+Z_	{[S]} <sup>superscript</sup> {[s]}
Subscript	\$-A_ <sub>subscript</sub> \$-Z_	{[U]} <sub>subscript</sub> {[u]}
Small Caps	\$sA_SMALLCAPS\$sZ_	{[A]} SMALL CAPS {[A]}

## Error Codes

All error codes begin with \$Er. This pattern can be used to find them and to delete or alter them once they have been addressed. Some adjustments may be quite complex, and a good knowledge of the operation of the target indexing program is required to make the correct adjustment.

### Error Code Definitions

**Table 8 Error Codes**

Label	Description	Cindex, Sky Index, Spreadsheet, Database, IXML, IDEX, Index-Manager	Macrex
\$ErP01_	Label inserted during preprocess.	Where joined entries have been separated. Styling may be incorrect and the entry folded into the wrong heading level.	
\$ERP02	Unexpected error during preprocess	\$ERP02_ inserted where the error occurred.	
\$ErH01_	Hierarchy error. No level 1 heading label.	This indicates a higher level heading label is missing in the hierarchy.	
\$ErH02_	No level 2 heading label.	It could be a serious error and can be repaired either by restarting the labelling process or use find/replace to correct any errors.	
\$ErH03_	No level 3 heading label.		
\$ErH08_	No level 8 heading label.		

Label	Description	Cindex, Sky Index, Spreadsheet, Database, IXML, IDEX, Index-Manager	Macrex
\$ErL00_	No locator has been labelled for this entry.  Where the locator has been mis-labelled this can be corrected and the \$ErL00_ label removed, including the leading tab character.	Ensure that the first locator is labelled {tab}\$L_  For headings and subheadings without locators the \$ErL00_ should be removed.	Ensure the first locator is labelled {tab}\$L_ or a stopper ~!~,
\$ErL01_	Locator ambiguity. Two parts of an entry contain punctuation that is recognized as a locator. A \$ErL01_ label is added to the end of the entry.	Ensure that the first locator is labelled {tab}\$L_.	Ensure the first locator is labelled {tab}\$L_ or a stopper ~!~,  If there is a number in the heading then enclose in curly brackets {…}.
\$ErL02_	Macrex export only. A locator marker {tab}\$L_ and a stopper ~ ~ are both present in an entry.		Ensure the first locator is labelled {tab}\$L_ or a stopper ~!~.
\$ErL03_	Unexpected spaces or punctuation in the locator string. First locator may have been incorrectly labelled.	Check first locator labelling. Move if necessary, with associated tab character and remove \$ErL03_ label.  Close unexpected spaces. Colon or semicolon may require the entry to be reworked. see abc; def; ghi may need reworking to see abc, see def, see ghi.	As Cindex etc.

Label	Description	Cindex, Sky Index, Spreadsheet, Database, IXML, IDEX, Index-Manager	Macrex
\$ErL04_	Incorrect number of tabs in concatenated entry.	Adjust as appropriate.	N/A
\$ErL05	IndexConvert cannot determine how a locator should be styled. Error message is written into the export file.	Index-Manager only.  Adjust Index-Manager file before import.  Remove error label.	N/A
{\$Ernnnn}	An unknown Unicode value	Sky Index only. Change the character, if possible or update the Translation table.  Remember to re-export the modified translation table.	Change the character, if possible or update the Unicode table.
{\$Er}	Macrex only. A right to left character. Not currently supported by IndexConvert.		No current solution.
\$ErP_	Processing error. Processing may have terminated unexpectedly.	Something in the text causes a processing problem. Currently this only occurs for IXML export. IndexConvert will continue processing if possible. The error label marks the problem.	

Label	Description	Supported by	Macrex
\$ErX01_	Export error, an unexpected tab character was encountered.	Sky SI8 IDEX file only. This can upset the export with unexpected consequences. The exported file may be incorrect. Fix and repeat the export.	

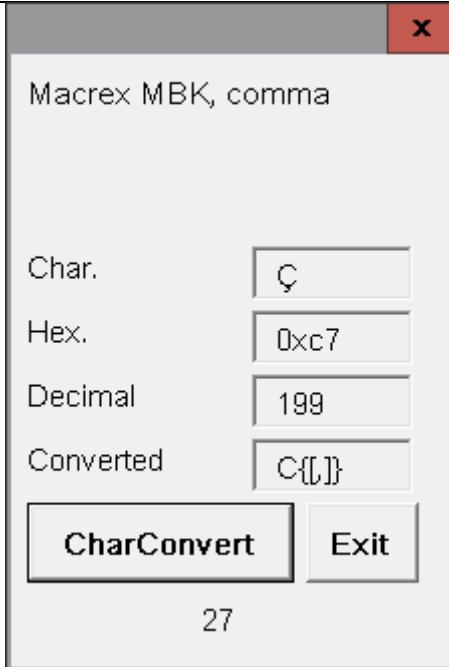
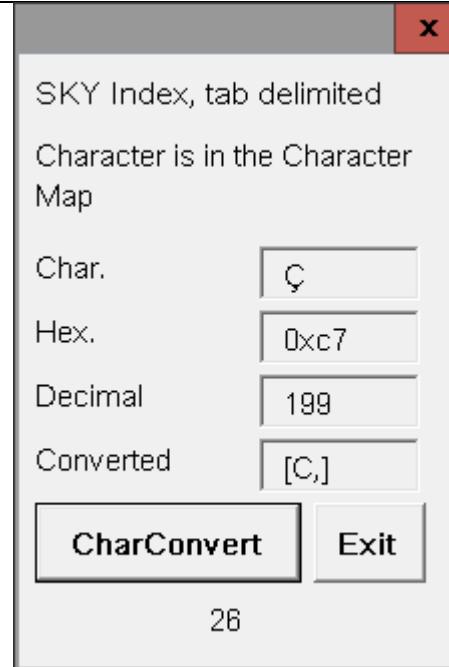
## Utilities - Examples

### CharConvert

Select Utilities and CharConvert. The CharConvert dialog appears.

Move the cursor to a character. The cursor can be either to the left of the character or over the character.

Select CharConvert. The original character, its hex and decimal equivalent and its conversion (if available) are displayed. Any of the results can be copied and pasted to another file or application.

Macrex	Sky Index (SI7)
 <p>Macrex MBK, comma</p> <p>Char. <input type="text" value="ç"/></p> <p>Hex. <input type="text" value="0xc7"/></p> <p>Decimal <input type="text" value="199"/></p> <p>Converted <input type="text" value="C[;]"/></p> <p><b>CharConvert</b> <b>Exit</b></p> <p>27</p>	 <p>SKY Index, tab delimited</p> <p>Character is in the Character Map</p> <p>Char. <input type="text" value="ç"/></p> <p>Hex. <input type="text" value="0xc7"/></p> <p>Decimal <input type="text" value="199"/></p> <p>Converted <input type="text" value="C[;]"/></p> <p><b>CharConvert</b> <b>Exit</b></p> <p>26</p>

## TextConvert

Select Utilities and TextConvert. The examples use the Turkish alphabet.

Capital letters A B C Ç D E F G Ğ H İ J K L M N O Ö P R S Ş T U Ü V Y Z

Lower case letters a b c ç d e f g ğ h i j k l m n o ö p r s ş t u ü v y z

Create a Docm document containing the alphabet above.

Run Preprocess to add the idxEND end of file label.

## Macrex

Known diacritics characters are substituted by their Macrex Printer Replacement Codes. Unknown characters are labelled with an error code containing the decimal Unicode value of the character.

Capital letters

A B C C{[,] D E F G G{[u]} H I I{[.]} J K L M N O O{["} P R S S{[,] T U U{["} V Y Z

Lower case letters

a b c c{[,] d e f g g{[u]} h i{\$Er305} i j k l m n o o{["} p r s s{[,] t u u{["} v y z

## Sky Index (SI7)

Known diacritics characters are substituted by their Sky Index Translation codes. Unknown characters are labelled with an error code containing the hex Unicode value of the character.<sup>1</sup>

Capital letters

A B C [C,] D E F G [Gu] H I [I.] J K L M N O [O"] P R S [S,] T U [U"] V Y Z

Lower case letters

a b c [c,] d e f g [gu] h i{\$Er305} i j k l m n o [o"] p r s [s,] t u [u"] v y z

---

<sup>1</sup> Example uses a Translation table used for testing IndexConvert with Sky Index.

## Licences

A range of licenses is available to suit individual needs. When first downloaded an internal licence is imposed allowing a million entries to be processed. This gradually decreases until a licence is required after about 100 days. An entry is a paragraph and may be several lines, especially for run-on indexes.

**Table 9 Licences**

	Parameter	Internal licence	Supplied licence	Note
a	Index size, number of entries	Time dependent	Defined by licence.2	The size of an individual index is imposed by the internal licence but not by a supplied licence.
b	Index quantity	Time dependent	Defined by licence	Limited by the licence for all licences. <sup>2</sup>
c	Total entries	=a x b, initially 1000000	=a x b	This limit is imposed by supplied licences. There is a 10% safety margin.
d	Expiry date	Set by software	Defined by licence	Licence expiry is defined by a month. Expiry date is the 28 <sup>th</sup> of the month with a 10 day safety margin.

One year, 5 year and perpetual licences are normally supplied. These limit the annual usage but metrics are reset for the first index in a new year. A Reserve license can also be supplied, this has a 3 year expiry date but no annual limit, the licence expiry is based on accumulated entry count only. Other licences are supplied by special request; these include short term licences for evaluation or training and reset licences that reset the usage metrics when loaded.

<sup>2</sup> Index quantity is normally set at a generous number for any licence, normally by setting index size to 1000.

# Advice and Notes

## Cindex

### Cindex IXML Index Records

When IndexConvert generates an IXML file the first pass determines the fonts used. The second pass creates the IXML records.

The Cindex IXML format can be imported by Cindex. Select import IXML or double click on the file. No post processing is required. If you use Cindex 3.0 or later this is highly recommended.

The Cindex expanded option places a copy of the locator in the lowest level subheading. If no locator is present an asterisk '\*' is placed in the lowest level subheading. To display records correctly in Cindex, adjust the Record Structure to display a minimum of 3 fields, check Subhead Preceding Locator is Required and optionally rename the lowest level subheading to Special. Using Style and Layout select Collapse Headings Below the lowest level available to suppress the special subheading in the formatted view.

### Cindex Tab Delimited File

The recommended process for importing an IndexConvert generated tab delimited file into Cindex is now described.

The process is similar for all programs where a delimited text file is imported but is only detailed for Cindex.

#### ***Inspect File***

Before importing the file into your indexing program inspect it and remove any blank lines from the end. If you have not removed headers and footers completely then they will appear at the end of the file and should be removed.

#### ***Import***

Create a new index file and import.

#### ***Remove Labels***

Always make a backup before beginning this editing.

First remove labels from the fields where they are expected. This allows remaining labels in the wrong field to be found rapidly using global search.

Remove \$H1\_ labels from the main heading.

Remove \$H2\_ labels from sub headings.

Similarly process any other subheadings.

Remove \$L\_ labels from locators.

Then look for any remaining labels. These will have ended up in the wrong fields and will require review and correction as necessary.

#### ***Adjust Styles***

This requires Cindex for Windows version 3.5.5.0 or later.

Save your file first. Use Find/Replace to restore bold, italic, underline, small caps, subscript and superscript to the original styles.

Figure 35 shows the Replace dialog using regular expressions to remove the bold markup and reinstate the original style. It uses three groups in the search window. The first contains \\$BA\_, the second contains .\*? to find any characters, the third contains \\$BZ\_. The \ is used because \$ has special significance.

Finally, check that all styles have been adjusted. Searching for A\_ will do this. If headings or subheadings have multiple clauses labelled with the same style then Find/Replace does not adjust these correctly and a manual adjustment will be required.

### Compare Files

Word compare files compares two files and displays the differences in a file that can be saved for future reference. Deselect Sort in Cindex, the index will be exported in the same order as the original making comparison easier.

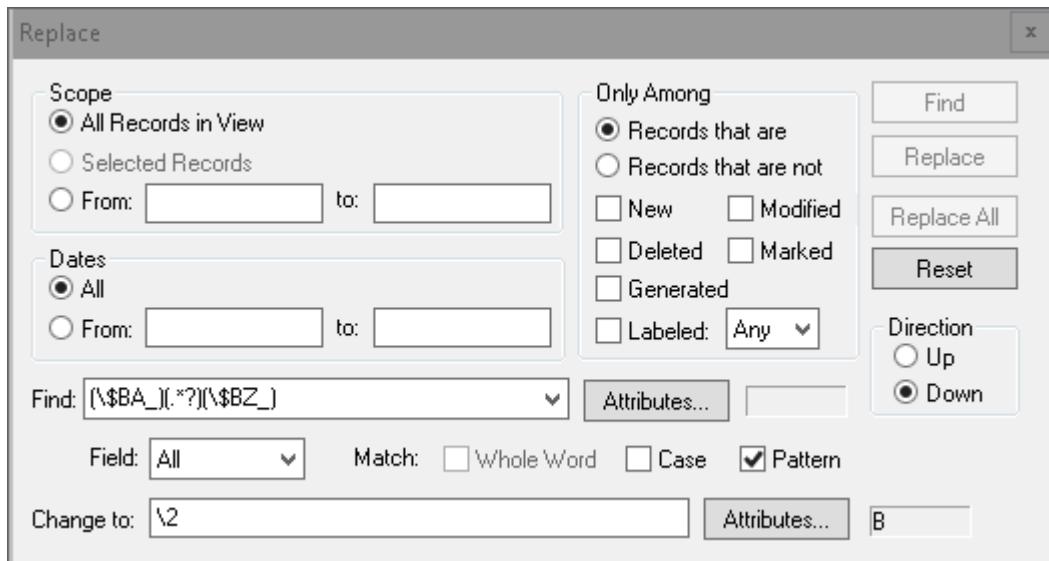


Figure 35 Cindex Replace All Showing Recovery of Bold Font from Markup

Table 10 Cindex Replace All Patterns to Recover Styles from Markup

Style	Find <sup>3</sup>	Replace
Italic	(\\$IA_)(.*?)(\\$IZ_)	\2, Attribute I
Bold	(\\$BA_)(.*?)(\\$BZ_)	\2, Attribute B
Smallcaps	(\\$sA_)(.*?)(\\$sZ_)	\2, Attribute S
Underline	(\\$UA_)(.*?)(\\$UZ_)	\2 Attribute U
Subscript	(\\$-A_)(.*?)(\\$-Z_)	\2 Attribute Sb
Superscript	(\\$+A_)(.*?)(\\$+Z_)	\2 Attribute Sp

+ is escaped

## Macrex

### MBK, comma delimited

Unicode characters are not supported directly by the Macrex mbk file but are replaced by printer replacement codes defined in a Unicode table.

Expansion of Unicode characters to printer replacement codes takes place during concatenation. An audit report is created automatically which lists all the printer replacement codes in the Unicode file and also those appearing in the index but not found in the Unicode table. The user may need to adjust the Unicode table to obtain the desired result.

<sup>3</sup> The Find string can be copied into the Cindex find field. Then you just need to adjust the attribute for each find/replace operation.

The \$H and \$L labels added during earlier stages of processing are replaced during concatenation to give a format suitable for Macrex import.

Heading levels are separated by commas. Commas within headings are enclosed in braces {}, and locators are separated from headings by commas.

Where the Unicode table defines printer replacement codes these replace the Unicode characters during concatenation. The first time an unrecognized code is encountered it is labelled {\$Lnnnn} where nnnn is the decimal code value of the Unicode character. The code is added to the internal copy of the Unicode table which is written to the audit report to support diagnosis. Right to left characters are not currently processed fully by IndexConvert. Every occurrence is labelled {\$Er}.

All \$Er codes must be removed before export.

Macrex expects non numeric characters in locator strings to be enclosed in braces, these are not added by IndexConvert. The user will need to address these either during Macrex editing or by editing the mbk file prior to import.

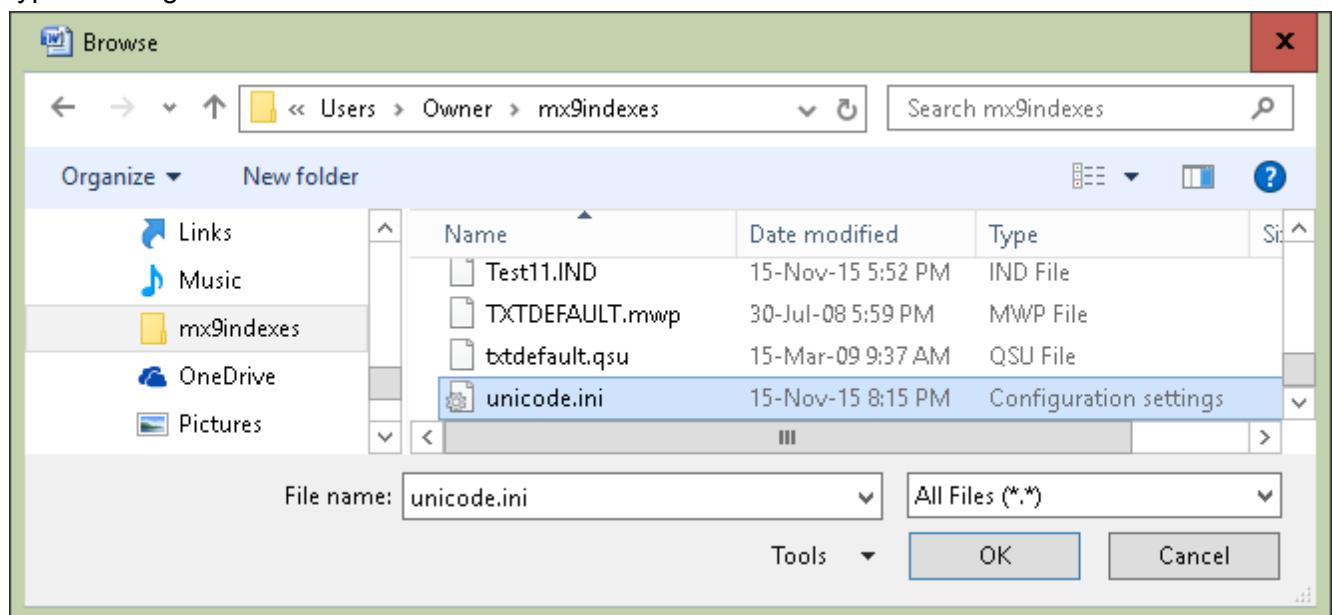
Styles are labelled \Bold\, ^italic^, {[S]}<sup>superscript</sup>{[s]}, {[U]}<sub>subscript</sub>{[u]}, {[UL]}underline{[ul]} and {[A]}SMALL CAPS{[a]}. These are interpreted by Macrex.

Nested braces may be present in the IndexConvert export file where complex combinations of diacritics and styles are present. Many adjustments have been included in the software to try and prevent their occurrence.

### Unicode.ini

Unicode.ini file is the Macrex default definition of the Unicode characters available and their expanded counterparts. IndexConvert uses the Unicode file to convert any Unicode characters to printer replacement codes recognized by Macrex.

The Unicode file is selected during configuration by selecting the Macrex Unicode File button. A typical dialog box is shown below.



**Figure 36 Select Macrex Unicode file**

The interaction of IndexConvert, the index being exported and the Unicode file can be very complex. The table below includes observations that may help but individual users will have to use their own knowledge and experience to make appropriate choices.

**Table 11 Macrex Unicode Table - Observations**

Unicode	Source	Observation	Recommendation
---------	--------	-------------	----------------

0x7b	Default Unicode table	Certain curly brackets placed by IndexConvert are replaced by the Unicode values.{{()}} and {{}}.	Either remove from the Unicode table or edit the file after export.
0x7d			

### ***Macrex Import***

Use the Load from Backup option.

Macrex parses the backup file as it is being read. The number of alterations required will depend on the attention paid to IndexConvert processing through the various stages. It will be more complex if a large number of Unicode characters are not defined in the Unicode table.

## **SKY Index**

### **INDEX Index Records**

When IndexConvert generates an IDEX file the first pass determines the fonts used. The second pass creates the IDEX records.

The IDEX format can be imported by Sky Index v8 (SI8). Select import IDEX.

Commas ',' are recognized as field separators in the locator field, records will be expanded before export.

See also entries containing several targets must be reorganized as below.

heading see also target a, target b, target c

must be reorganized to

heading

see also target a

see also target b

see also target c

heading, 1, 2, 3, 4 see also target

needs to be

heading, 1, 2, 3, 4

see also target

The IDEX exporter uses terms defined in the cross reference matrix at Locators [ I ] to identify cross references in the index being processed. Ensure that terms are filed accurately under 'see' or 'see also'. Case is unimportant for IDEX export but is important for locator labelling, careful selection of terms will avoid the need to alter them as processing proceeds.

If a see term is a syllable in another see term then the export may need adjusting.

see main

will export correctly

seeeee main

will export as see eee main

Spaces within see and see also terms will normally be Unicode decimal 32 characters. If other values are used then include them in the Cross refs grid.

see also with nonbreaking space

see also with em space

see also with en space

see also with  $\frac{1}{4}$  em space

The IDEX range separator is not supported by IndexConvert.

## Tab Delimited

### Concatenation

Concatenation expands the index to one locator per record based on the comma separators within the locator string. If separators are semicolons or colons then they should be altered to commas prior to Concatenation.

If the index contains Unicode characters outside the ANSI set then an exported copy of the Translation table is used by IndexConvert to expand Unicode characters to their keyboard text equivalents. For example Å will expand to [A`] if that entry exists in the Translation table. Use Options>Translation Manager in Sky Index to reach the Translation table and select Export to generate a file that is readable by IndexConvert. Any characters not supported by the translation table are recognized during Concatenate. They are labelled [\$Ernnnn] where nnnn is the hex value of the Unicode character required when creating a translation code. A copy of entries containing \$Er codes appears in an automatically generated audit report. The Translation table appears in the same audit report with all unsupported Unicode hex values appended. The Translation table must be updated and re-exported and Concatenate must be repeated until no \$Ernnnn error codes appear.

The dialog boxes that appear when Concatenation is selected are below.

The first appears if a Translation file has not been selected at the Configuration stage or one does not exist.



Figure 37 Sky Index Translation File not found

The second appears if a Translation file has been found.

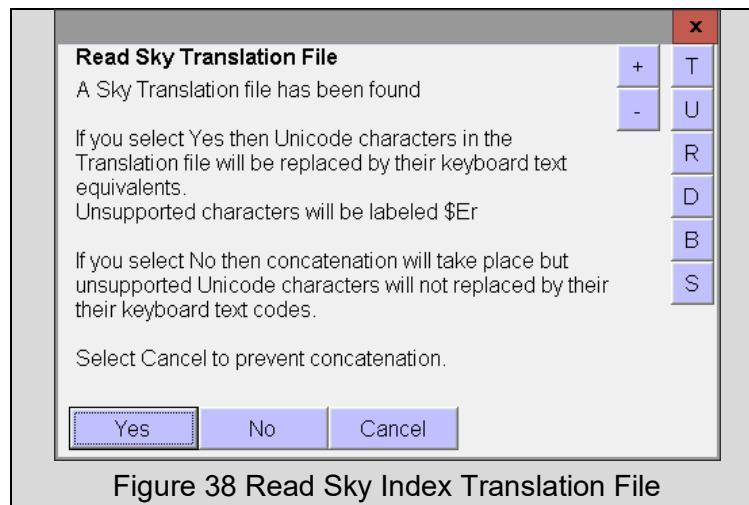


Figure 38 Read Sky Index Translation File

When export takes place, Unicode characters outside of the ANSI set and without an entry in the Sky Index Translation file will appear in the exported file as question marks (?).

#### ***Import to Sky Index***

Heading and locator labels are retained for quality purposes and will need to be removed.

Style labels are exported and need to be recovered using Sky Index find/replace. Always make a backup before commencing this editing.

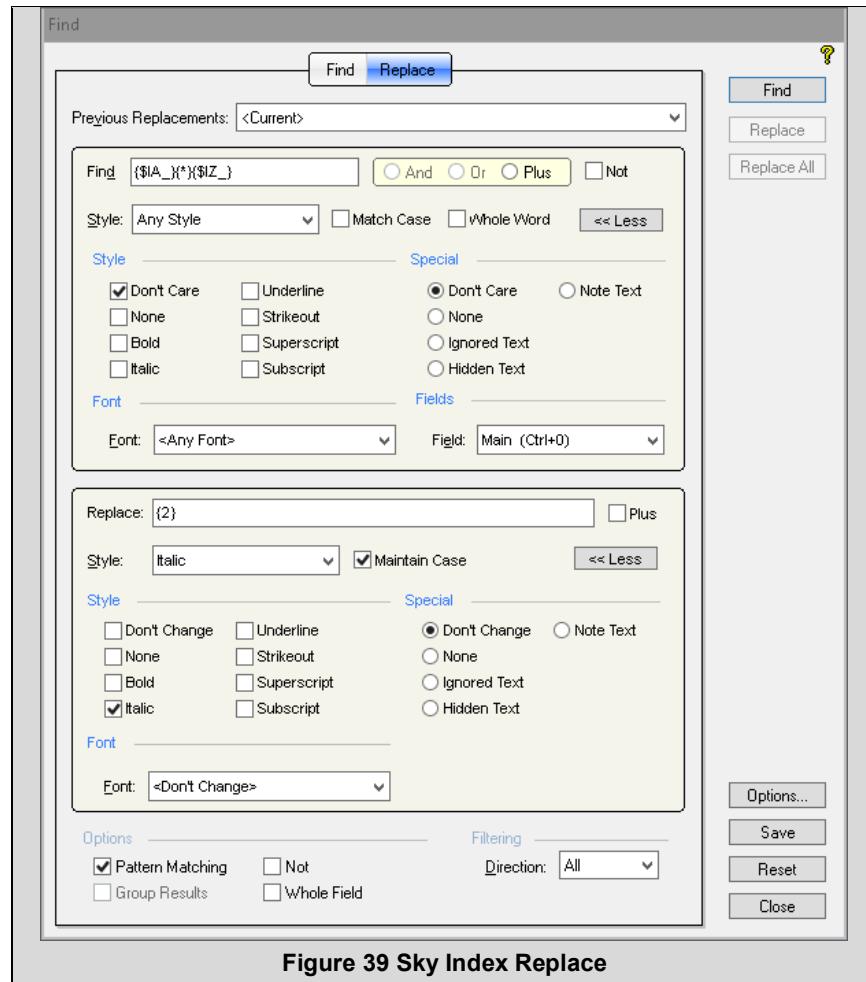


Figure 39 Sky Index Replace

The dialog box above shows Sky Index replace dialog set to restore the italic style. The Find strings in the table below require Pattern Matching to be selected.

Style	Find <sup>4</sup>	Replace
Italic	\${IA_}{*}\${IZ_}	{2}, Italic
Bold	\${BA_}{*}\${BZ_}	{2}, Bold
Underline	\${UA_}{*}\${UZ_}	{2}, Underline
SmallCaps	\${sA_}{*}\${sZ_}	Not supported by Sky Index.
Subscript	\${-A_}{*}\${-Z_}	{2}, Subscript
Superscript	\${+A_}{*}\${+Z_}	{2}, Superscript

Figure 40 Sky Index Style Recovery

<sup>4</sup> The Find string can be copied into the Sky Index find field. Then you just need to adjust the attribute for each find/replace operation.

## Index-Manager

A Unicode text file is exported compatible with Index-Manager.

### Headings

Each record contains one heading or subheading (etc.) only. Higher level headings are indicated by – (dash space).

heading

- subheading

-- sub-subheading

Styles are identified using a **<b>,,,</b>** convention.

- b bold i.e. **<b>bold</b>**
- I italic **<i>italic</i>**
- u underline **<u>underline</u>**
- + superscript **<+>superscript</+>**
- - subscript **<->subscript</->**
- s small caps **<s>small caps</s>**

Style markups are nested i.e. **<b><i><u>text sample</u></i></b>**.

When the index is embedded into a Word document styles can be recovered using find/replace.

Ensure that the embedded index entries are visible.

Example

Find what: **(\<b\>)(\<\*\>)(\</b\>)**

Options: Use Wildcards

Replace with \2

Format: Font: Bold

### Locators

Locator strings will have been labeled **{tab}\$L\_** in earlier processing.

During export, pure numeric locators are identified by evaluation. Cross references are identified by reference to user defined cross reference terms defined at Locators [i] in the IndexConvert user interface. Locators containing prefixes are identified by reference to user defined prefixes defined at Locators [i] and evaluation of the remnant.

Other locators {TBD more to be added when available}

### Locator Formatting

Each record may contain zero, one or more locators.

Locators are identified **<idx>locator</idx>**

Locator style markup will be that of the first character.

Styles are identified as follows:

- b bold i.e. **<idx format='b'>bold locator</idx>**
- I italic **<idx format='i'>italic locator</idx>**
- u underline **<idx format='u'>underline locator</idx>**
- bold italic underline **<idx format='bui'>bui locator</idx>**

Where the locator contains superscript, subscript and small caps formats these are ignored.

### Cross reference Formatting

Cross references are identified as locators but formatting will be as for headings.

See carbon dioxide (CO<sub>2</sub>) will become **<idx> </i>See</> carbon dioxide (CO<->2</>)</idx>**

**Export** is a UTF-8 Unicode file.

## Spreadsheet, tab delimited

For spreadsheet export empty headings are padded with \$Hn\_ labels to ensure that locators always end up in the locators column.

If IndexConvert guidelines are followed, import to a spreadsheet should be straightforward.

Heading (\$Hn\_) and locator (\$L\_) labels are retained for quality purposes and will need to be removed.

Locators are not separated into separate fields and the file is not expanded to provide one record per locator.

## Database, tab delimited

Database export is similar to Spreadsheet export except that each entry contains one locator only.

Heading and locator labels are retained for quality purposes and will need to be removed.

Cindex is able to import the Database Tab Delimited format. This may be useful if an expanded one locator per entry format is desired. Following Concatenate the file can be manipulated to copy the locator into a lowest level subheading which will then appear in Cindex as a Special Subheading. This is useful if a configuration record is wanted when re-indexing.

Insert a copy of the locator into a new field labelled \$S\_ before the locator. Begin by using the following Word find and replace with Patterns selected.

Find (^t\$L\_)(\*)(>)

Replace ^t\$S\_12\1\2

If it doesn't work you may need to adjust the patterns.

Export the file, configure Cindex to show \$S\_ in the special subheading and import to Cindex.

## Word AutoMark File

A Word AutoMark file contains 1, 2 or more columns laid out as text or a table that supports automated embedded indexing of matching phrases. At minimum a concordance is created but careful use of triggers (column 1) and index entries (column 2) makes this a very powerful process. Column 3 can be used for notes. Word processes an AutoMark file in a few seconds and can create hundreds or thousands of entries in this time. If the result is unsatisfactory Ctrl-Z removes all the insertions.

Successful operation of an AutoMark file will depend on each row creating no more than about 5 entries.

## Word Embedded Using Page Numbers

This option embeds the index into a target Word document using page numbers to determine XE field and bookmark locations.. See the IndexExploit user guide for details of XE fields.

### WARNING – FLASHING SCREEN

Please note that the application continually swaps between the file containing the index and the target document, causing the screen to flash.

### Page Numbers

The location of embedded entries is determined by page numbers. The index page numbers and content must match the target document. Unelided page numbers are required for ranges. Locator range separators must be hyphens. Simple page numbers are required, e.g.148-150. Chapter prefixes, although allowed in Word, are not supported by this process. It is possible to set up chapters and page formats following indexing, and the index will show the new styles.

If the index and target document are exports from PDF then special care is required. If the index is created using indexing software then configure it to add the \$H and \$L symbols expected. Run Preprocess, Audit and Elided before attempting to embed.

## Options

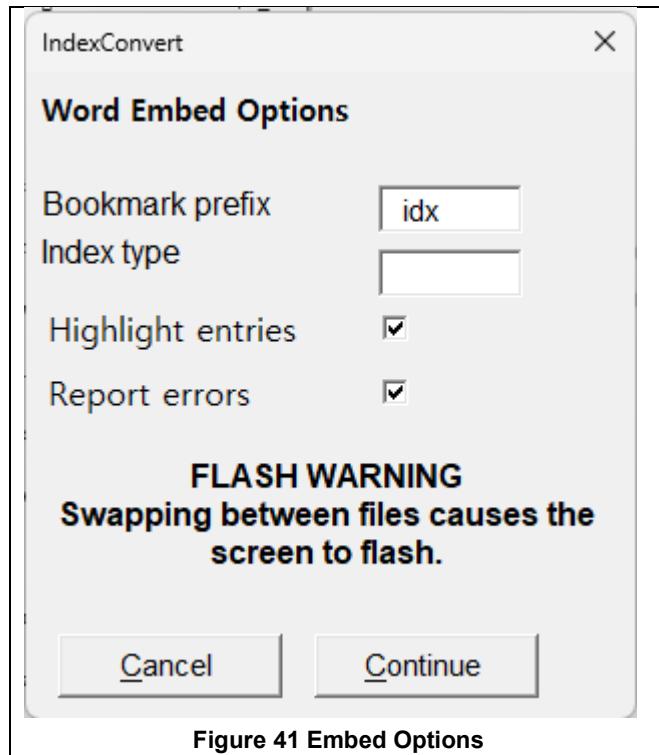


Figure 41 Embed Options

A **Bookmark prefix** can be chosen. The default is idx.

An **Index type** can be chosen to support multiple indexes. Type should be left blank for the main or general index. A single character is used for other indexes. Process the general index last. Adding the !f switch is complicated, any missed will end up as general index entries. \$ErY tokens will be written to the target document where the type code cannot be written successfully..

When **Highlight entries** is selected the ranges and XE field insertion points and referenced ranges are highlighted. If your document already contains highlighting that you don't want to disturb then don't use this option.

**Report errors** causes any errors to be reported as they occur. Regardless of the setting, errors will be indicated by \$Er tokens written into the source or target document. Certain errors in the source index will cause processing to stop and must be fixed before a repeat is allowed.

## Screen Flicker

IndexConvert switches rapidly between the source and target files, this can cause intense and disturbing flicker. It can be reduced by opening the two files and setting them side by side on the screen with no overlap before starting the embedding. The cursor must be in the index source document when starting the embed.

## Process

Every entry in the index document is parsed and inserted into the target document as an XE field. Heading and cross reference styles are copied across. Locator styles are limited by Word to normal, bold, and italic. Ranges and their names are determined from page number ranges.

Simple page numbers (1, 100, 301, 314-320 etc.) are accepted. Roman numeral locators will be processed as text, like cross references, and will need to be moved to the correct locations and adjusted.

## Locator Styles

The `\b` and `\i` switches are set by reading the locator styles in the source index.

## Cross References (`\t` Field)

Cross reference terms are checked against the cross reference terms in the Locators [ `I` ] form. Any not matching will be labelled with \$ErX tokens written to the cross reference field. Roman numerals and any other non-numeric locators will be treated as text.

## Styles

The VBA Mark Index Entry function is used for embedding entries. This does not support rich text. Styles are initially coded. Bold styles are encoded using `$BA_bold` text`$BZ_`. When the embedding is completed, these are automatically converted back to the original style in the target document using Find/Replace. If processing is interrupted this will need to be done manually. Coding used is summarized below.

Style	Code
AllCaps	A
DoubleStrikeThrough	D
Emboss	E
Engrave	G
SmallCaps	s
Bold	B
Italic	I
Outline	O
StrikeThrough	T
Shadow	H
Superscript	+
Subscript	-
Underline	U

## Error Codes

Error codes are written to the source file or the target file depending on the type. Errors in the source document may halt Embed but can be avoided by taking care during the earlier IndexConvert processes. If labelling is accomplished by indexing software then Preprocess and Audit should be run.

**Table 12 Embed Error Codes**

Code	Error	Location	Note
\$ErHnn_	Heading label	Source document	Embed will be abandoned.
\$ErSnn_	Heading or locator label error	Source document	Embed continues
\$ErXnn	Possible cross reference error	XE fields in target	Search for \$Er with XE fields visible
\$ErLnn_	Locator error	XE fields in target	Search for \$Er with XE fields visible
\$ErYnn_	Index type could not be written	Target	Search for \$Er in target

## Word File Compare Tips

The following may help if you need to perform a Word Compare Files operation to check the original index against the converted index re-exported from your indexing program

Description	Cause	Treatment
Sort order	Sort order of the original index may have been arranged manually or may have been as a result of complex sort algorithms.	For Cindex it is possible to disable sort. The index entries when exported will appear in the same order as your original index. This makes comparison much easier.
Commas may have different styling from the original index.	Dependent on the adjust punctuation setting of your indexing program.	Adjust your punctuation setting.
Commas or full stops within quotation marks.	These are present in the original index. When exported by your indexing program a second comma or full stop is inserted.	Edit as necessary or adjust punctuation setting.
Cross references within brackets.	Present in the original index.	Edit if necessary

## Alternative Installation

Installation instructions relate to Microsoft Office Word 2007 or 2010 running on Windows 10. Other installations may differ

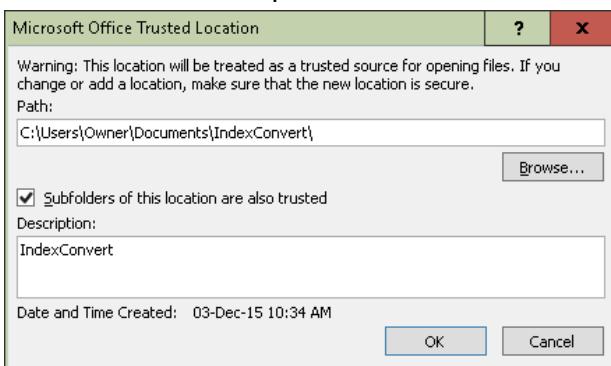
The macro is contained in a file named IndexConvert x.x.x.docm where x.x.x is the release. This can be installed as a Word add-In.

The following paragraphs explain how to configure security settings to let IndexConvert run and then activate it.

For users who prefer not to have to activate IndexConvert for each Word session there is a description of how to configure an AutoExec macro to activate IndexConvert when Word starts.

### Trust Settings

The Trust settings can be adjusted to allow files in the chosen directory and subdirectory to be trusted. Use Word Options>Trust Center>Trust Center Settings.



**Figure 42 Microsoft Office Trusted Locations**

Select the View ribbon and Macros. A dialog box appears. Select IndexConvert and Run. The security warning will no longer appear.

### Activate

IndexConvert is active for a word session following activation. You will be able to avoid this step if you have an AutoExec macro configured to activate IndexConvert.

To activate IndexConvert select the Office Button>Word Options>Add-Ins>Word Add-Ins>Go. Select Add as shown in the dialog box below if IndexConvert is not present and add the file.

If IndexConvert is already present then check it to activate it.

If you do this with macro recording on the macro created will include an AddIns command that can be used in an AutoExec macro.

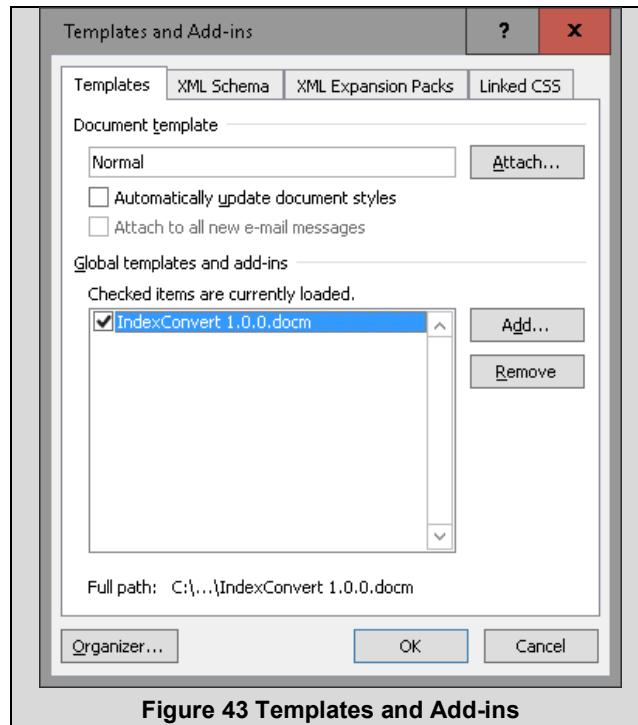


Figure 43 Templates and Add-ins

## AutoExec Word Macro

AutoExec is a Word macro that runs on commencement of Word. It is installed in 'All active templates and documents'.

AutoExec can perform chosen commands when Word starts which can include activating IndexConvert.

If an AutoExec file exists then include AddIns commands to activate selected add-ins when Word starts. If AutoExec does not exist it can be created using the Word Macros dialog.

The AddIns command syntax is

    AddIns("Path").Installed = True

    An example of Path is C:\Users\Owner\Documents\IndexConvert\IndexConvert.docm

If you change the file name to remove the version number and remember to do this whenever a new version is released then the AutoExec file will not need to be changed.

An example AutoExec is shown below.

```
Sub AutoExec()
AddIns("C:\Users\Owner\Documents\IndexConvert\IndexConvert.docm").Installed = True
AddIns("C:\Users\Owner\Documents\IndexConvert\Elided_123_4.docm").Installed = True
End Sub
```

## Test Indexes

Many test indexes have been used when developing IndexConvert. Some have been supplied by publishers, some by other indexers and many have been specially generated for specific test cases. Publicly available indexes are a great asset and some of these have been used.

*Index to West Plains Gazette* Issues 1 through 30 1978–1987 is made available as a pro-bono PDF download by the PNW chapter of ASI. It is three separate indexes that were separated out following conversion to Word.

[http://pnwasi.org/wp/wp-content/uploads/2016/09/Index\\_WPG\\_Vol\\_1-30.pdf](http://pnwasi.org/wp/wp-content/uploads/2016/09/Index_WPG_Vol_1-30.pdf)

## Disclaimer

The success of IndexConvert depends on an understanding of the structure of the index, an understanding of how IndexConvert operates and an understanding of the indexing software being used.

The process supported automates many of the tedious operations involved but full automation is not possible.

This software is provided 'as is' without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of fitness for a purpose.

The software and its documentation could include technical or other mistakes, inaccuracies or typographical errors.

In no event shall the author of this software be liable to you or any third parties for any special, punitive, incidental, indirect or consequential damages of any kind, or any damages whatsoever, including, without limitation, those resulting from loss of use, data or profits, whether or not the author has been advised of the possibility of such damages, and on any theory of liability, arising out of or in connection with the use of this software.

The use of the software is at your own discretion and risk and with agreement that you will be solely responsible for any damage to your computer system or loss of data that results from such activities. No advice or information, whether oral or written, obtained by you from the author shall create any warranty for the software.

The user is responsible for ensuring that any material processed by IndexConvert is done in a manner that does not infringe on the rights of any copyright owner.

This agreement is made in accordance with laws of the United Kingdom.

## IndexConvert Copying, Licence Keys, and Redistribution

You may make copies of IndexConvert for archiving. You may keep copies on multiple computers. One computer should be the principal computer for the purpose of index conversion. Any others should be for backup and familiarization only.

You may not redistribute IndexConvert. Users should always download IndexConvert from [www.indexbase.co.uk](http://www.indexbase.co.uk) to be sure the latest version is used.

You are not permitted to reverse engineer or decompile this software. You are not permitted to decode passwords used to protect this software.

Any licence keys supplied are for use by the end user only and are not transferable. They may not be redistributed.

From time to time special purpose licence keys will be supplied for special maintenance purposes. These are not to be redistributed.

## Enquiries

If you have queries about this software or require advice about its use, then please email [enquiries@indexbase.co.uk](mailto:enquiries@indexbase.co.uk) or use the enquiry form at [www.indexbase.co.uk](http://www.indexbase.co.uk) which will result in a faster response.

# INDEX

IndexConvert commands and settings are shown in **Bold**.

## Conventional Index

- \$DebugOff\_**, 23
- \$DebugOn\_**, 23
- Acrobat DC Pro
  - Recognise Text if Needed, 8
  - Retain Flowing Text, 8
- Audit**, 6, 17, 22
- Audit report**, 22
- Autoexec macro, 55
- AutoMark**, 25, 31, 50
- Automatic locator search**, 15
- blank lines, 25
- bold, 7
- Bookmark prefix** (Word embedded), 51
- Bottom of file (B)**, 12
- chapters, 5
- CharConvert**, 26–27, 38
- Cindex, 2
  - import, 42–43
  - IXML, 25, 30
  - IXML expanded, 25, 30–31
  - tab delimited, 29
- commas, as locator separators, 9
- compatibility, 2–3
- Concatenate** (prepend), 7, 23–24
- concepts, 1
- Configuration**, 10–11
  - conversion for a client, 24
- cross references, 5, 16
  - unlabeled, 22
- ctrl-break/Esc, 11
- data structures, 29–31
- database, 25, 30
  - import, 50
- delimited file, 8
- disclaimer, 57
- Down (D)**, 12
- Elided add-in**, 6, 28
- Elided page ranges, 23
- embedded indexing, 3, 50–53
- encoding, 25
- end of file label (**idxEND**), 5, 23
- enquiries, 57
- error codes, 34–37
  - Word embedded, 52
- Exit**, 25
- Export**, 7, 11, 24–25
- Word embedded, 50–53
- file extensions, 25
- group headings, 8
- heading hierarchy, 22
- heading levels, 19
- [ i ] buttons**, 10
- IDEX**, 2, 7–8, 25, 31, 45
- index conversion
  - between indexing programs, 8
- index layout, 8–9
- index structure
  - set-out, 14
- Index-Manager, 3, 7, 25, 31, 49
  - import, 49
- initial inspection, 4
- installation, 3–7
  - alternative, 54
- italic, 7
- iteration, 7
- IXML, 7, 8
- Label headings**, 6, 18–21
  - [ i ]**, 18, 20
  - Method A**, 6, 19–20
  - Method B**, 6, 20–21
  - Method C**, 6, 21
- label locators**
  - [ i ]**, 14
- Label locators, 14–17
  - first pass, 14
  - manual, 9
  - second pass, 14
- Label styles**, 7, 24
- labels, 32–33
  - example, 32
- last entry processed, 22
- legal indexes, 8
- licences, 41, 57
- Licences**, 10
- Locator prefixes**, 5, 17
- locator separators, 9
- locators, 9
  - incorrectly labeled, 22
  - unlabeled, 22
  - with unexpected spaces, 22
- Locators L**, 5
- Locators M**, 6

Macrex, 2, 25, 29  
 diacritics, 40  
 import, 43–45  
 manual labelling  
 headings, 22  
 locators, 23  
**Measure Indent**, 27  
**Method B**, 27  
**Metrics**, 25–26  
 multi-column indexes, 5  
 multiple exports, 7  
 navigation, 12  
**Page number limit**, 15  
 PDF files, 2, 8  
 layout, 8  
 preparation, 8  
 prepend. see concatenate  
**Preprocess**, 5–7, 12–17  
 process commonality, 7  
 process summary, 1  
 processing interrupted, 23–24  
 processing speed, 23  
 progress indicator, 11  
 quality, 7  
 quick access toolbar, 9  
**Return to original position (R)**, 12  
**Roman numerals**, 15  
 upper case, 15  
 run-back subheadings, 14  
**Scratchpad**, 16, 17  
 screen flicker  
 Word embedded, 51  
**Screen update control (S)**, 12  
 screen updates, 12, 23  
 scrolling, 12  
 search times, 14  
 security warning, 3  
 shortcuts, 9  
**Sky Index**, 2, 30  
 diacritics, 40  
**INDEX**, 25, 31  
 import, 45–48  
**SI8 IXML**, 30–31  
**SI8 IXML expanded**, 25  
**Sky Index v7**, 25  
 small caps, 7  
 spreadsheet, 25, 29  
 import, 50  
**start IndexConvert**, 9  
 startup path, 4  
 styles, 7  
 example, 33  
 leakage, 22  
 subscript, 7  
 superscript, 7  
 test indexes, 56  
**TextConvert**, 26, 40  
 time for conversion, iii  
**Top of file (T)**, 12  
**Translation table (Sky Index)**, 24, 26  
 underline, 7  
 underline characters (Macrex), 7  
 Unicode, 2  
**Unicode.ini file (Macrex)**, 24, 26  
 uninstall, 4  
 unlabeled headings, 22  
**Up (U)**, 12  
 user interface, 10  
 utilities, 26–27  
**Version**, 10  
 Windows clipboard, 22  
**Word**  
 docm (macro enabled) file, 9  
 File Compare, 53  
 Office 365, 2  
 ribbon, 9  
 status bar, 11  
**Word AutoMark**, 25, 31–32  
**Word embedded**, 25, 31  
 bookmark prefix, 51  
 cross references, 52  
 error codes, 52  
 errors, 51  
 highlight entries, 51  
 index type, 51  
 locator styles, 52  
 options, 51  
 page numbers, 50  
 process, 51  
 screen flicker, 51  
 styles, 52  
**Zoom (+/-)**, 12

## Linked Index

The linked index is created using IndexExploit.

Group headings are prefixed INDEX\_. Search for INDEX\_ to create a list of groups for quick return to the linked index.

### INDEX\_ \$

\$DebugOff\_, [23](#)

\$DebugOn\_, [23](#)

### INDEX\_ A

Acrobat DC Pro

    Recognise Text if Needed, [8](#)

    Retain Flowing Text, [8](#)

Audit, [6](#) , [17](#) , [21–22](#)

Audit report, [22](#)

Autoexec macro, [54–55](#)

AutoMark, [25](#) , [31](#) , [50](#)

Automatic locator search, [15](#)

### INDEX\_ B

blank lines, [25](#)

bold, [7](#)

Bookmark prefix (Word embedded), [51](#)

Bottom of file (B), [12](#)

### INDEX\_ C

chapters, [5](#)

CharConvert, [26–27](#) , [34–38](#)

Cindex, [2](#)

    import, [41–43](#)

    IXML, [25](#) , [30](#)

    IXML expanded, [25](#) , [30–31](#)

    tab delimited, [29](#)

commas, as locator separators, [9](#)

compatibility, [1–3](#)

Concatenate (prepend), [6–7](#) , [23–24](#)

concepts, [1–3](#)

Configuration, [10–11](#)

conversion for a client, [24](#)

cross references, [5](#) , [15–16](#)

    unlabeled, [22](#)

ctrl-break/Esc, [11](#)

### INDEX\_ D

data structures, [28–31](#)

database, [25](#) , [30](#)

Page 60

import, [50](#)

delimited file, [8](#)

disclaimer, [56–57](#)

Down (D), [12](#)

### INDEX\_ E

Elided add-in, [6](#) , [27–28](#)

Elided page ranges, [23](#)

embedded indexing, [3](#) , [50–53](#)

encoding, [25](#)

end of file label (idxEND), [5](#) , [23](#)

enquiries, [57](#)

error codes, [33–38](#)

    Word embedded, [52](#)

Exit, [25](#)

Export, [7](#) , [10–11](#) , [24–25](#)

    Word embedded, [50–53](#)

### INDEX\_ F

file extensions, [24–25](#)

### INDEX\_ G

group headings, [8](#)

### INDEX\_ H

heading hierarchy, [22](#)

heading levels, [19](#)

### INDEX\_ I

[ i ] buttons, [10](#)

INDEX, [2](#) , [7–8](#) , [25](#) , [31](#) , [45](#)

index conversion

    between indexing programs, [8](#)

index layout, [8–9](#)

index structure

    set-out, [12–14](#)

Index-Manager, [2–3](#) , [7](#) , [25](#) , [31](#) , [31](#) , [45–49](#)

    import, [49](#)

initial inspection, [4](#)

installation, [3–8](#)

    alternative, [53–54](#)

italic, [7](#)

iteration, [7](#)  
 IXML, [7](#), [8](#)  
**INDEX\_L**  
 Label headings, [5–6](#), [17–21](#)  
   [i], [18](#), [20](#)  
   Method A, [6](#), [18–20](#)  
   Method B, [6](#), [19–21](#)  
   Method C, [6](#), [20–21](#)  
 Label locators, [14–17](#)  
   [i], [14](#)  
   first pass, [14](#)  
   manual, [8–9](#)  
   second pass, [14](#)  
 Label styles, [7](#), [24](#)  
 labels, [31–33](#)  
   example, [32](#)  
 last entry processed, [22](#)  
 legal indexes, [8](#)  
 Licences, [10](#), [40–41](#), [57](#)  
 Locator prefixes, [5](#), [16–17](#)  
 locator separators, [9](#)  
 locators, [9](#)  
   incorrectly labeled, [22](#)  
   unlabeled, [22](#)  
   with unexpected spaces, [22](#)  
 Locators L, [5](#)  
 Locators M, [6](#)  
**INDEX\_M**  
 Macrex, [2](#), [25](#), [29](#)  
   diacritics, [40](#)  
   import, [42–45](#)  
 manual labelling  
   headings, [22](#)  
   locators, [22–23](#)  
 Measure Indent, [26–27](#)  
 Method B, [27](#)  
 Metrics, [25–26](#)  
 multi-column indexes, [5](#)  
 multiple exports, [7](#)  
**INDEX\_N**  
 navigation, [12](#)  
**INDEX\_P**  
 Page number limit, [14–15](#)  
 PDF files, [2](#), [8](#)  
   layout, [8](#)  
 preparation, [7–8](#)  
 prepend, see concatenate  
 Preprocess, [4–7](#), [11–17](#)  
 process commonality, [7](#)  
 process summary, [1](#)  
 processing interrupted, [23–24](#)  
 processing speed, [23](#)  
 progress indicator, [11](#)  
**INDEX\_Q**  
 quality, [7–8](#)  
 quick access toolbar, [9](#)  
**INDEX\_R**  
 Return to original position (R), [12](#)  
 Roman numerals, [15](#)  
   upper case, [15](#)  
 run-back subheadings, [14](#)  
**INDEX\_S**  
 Scratchpad, [16](#), [17](#)  
 screen flicker  
   Word embedded, [51](#)  
 Screen update control (S), [12](#)  
 screen updates, [12](#), [23](#)  
 scrolling, [12](#)  
 search times, [14](#)  
 security warning, [3](#)  
 shortcuts, [9](#), [9](#)  
 Sky Index, [2](#), [2](#), [29–30](#)  
   diacritics, [40](#)  
   INDEX, [25](#), [30–31](#)  
   import, [43–48](#)  
   SI8 IXML, [30–31](#)  
   SI8 IXML expanded, [25](#)  
 Sky Index v7, [25](#)  
 small caps, [7](#)  
 spreadsheet, [25](#), [29](#)  
   import, [49–50](#)  
 start IndexConvert, [9](#)  
 startup path, [3–4](#)  
 styles, [7](#)

example, [32–33](#)  
leakage, [22](#)  
subscript, [7](#)  
superscript, [7](#)  
**INDEX\_T**  
test indexes, [55–56](#)  
TextConvert, [26](#) , [26](#) , [38–40](#)  
time for conversion, [3](#)  
Top of file (T), [12](#)  
Translation table (Sky Index), [23–24](#) , [26](#)  
**INDEX\_U**  
underline, [7](#)  
underline characters (Macrex), [7](#)  
Unicode, [2](#)  
Unicode.ini file (Macrex), [24](#) , [26](#)  
uninstall, [4](#)  
unlabeled headings, [22](#)  
Up (U), [12](#)  
user interface, [9–10](#)  
utilities, [25–27](#)  
**INDEX\_V**  
Version, [10](#)  
**INDEX\_W**  
Windows clipboard, [22](#)  
Word  
    docm (macro enabled) file, [9](#)  
    File Compare, [52–53](#)  
    Office 365, [2](#)  
    ribbon, [9](#)  
    status bar, [11](#)  
Word AutoMark, [25](#) , [31–32](#)  
Word embedded, [25](#) , [31](#)  
    bookmark prefix, [51](#)  
    cross references, [52](#)  
    error codes, [52](#)  
    errors, [51](#) , [51](#)  
    highlight entries, [51](#)  
    index type, [51](#)  
    locator styles, [51–52](#)  
    options, [50–51](#)  
    page numbers, [50](#)  
    process, [51](#)