IndexConvert: what does it do and why and how would you use it?

Lucie Haskins

Most indexers occasionally and some indexers frequently are faced with the need to convert an existing PDF or RTF/doc index into a format that can be read into their indexing software. Easier said than done, and for the most part until now it has been a case of doing the job manually. Will the advent of IndexConvert in late 2015, designed specifically for the purpose, change all that? Lucie Haskins, an established indexer and former computer programmer, has been putting the program through its paces. She finds it to be a valuable program which does what it sets out to do well, but it calls for considerable ongoing interaction and decision making on the indexer's part, which makes it a program for indexers who like peeking under the covers, rather than for someone looking for a quick turnkey solution to a one-off problem.

Why would you need to convert an existing index?

Imagine the real-life scenario I periodically encounter where a publishing client needs an index to the new edition of their math book but, because the authors were so pleased with the index in the previous edition and because they had made minimal structural changes to the new edition's content, they want to use the previous edition's index and modify it as little as possible: repaginating as needed, and adding, deleting, or editing the content changes as little as possible. They have the previous edition's index for me to work with.

It sounds easy enough, but there are some practical issues to consider. Because this index is in PDF or RTF/DOC format, it isn't possible to just load it as is into my dedicated indexing software. I'll explain why.

Figure 1 (left-hand column) displays a sample index with 11 index entries. This is a typical indented index where each main entry appears on one line and each associated subentry is written on a separate line. The human eye recognizes the links between the flush left lines (main entries) and the indented lines (their subentries). But the relationships between these pieces of information (on separate lines) are all implied links[1] when they have been generated in RTF/DOC or PDF format.

There is no real connection between these separate lines in the left column as they exist in these formats. Neither Microsoft Word nor Adobe Acrobat recognizes these implied relationships; they each treat each line as a separate and complete record unto itself. In other words, the subentries are no longer associated with the main entries and all connections between the records have been severed.
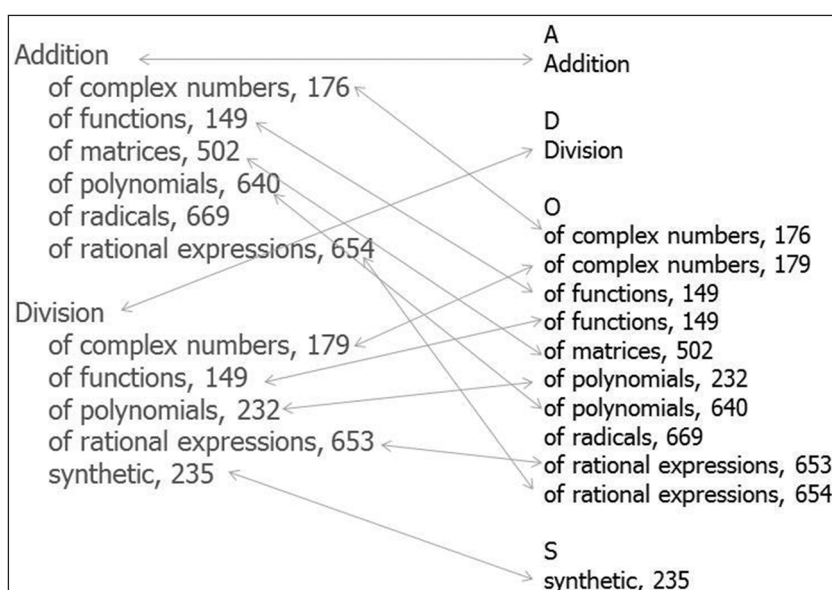
If the index entries were imported as-is, the entries would appear in our dedicated indexing software as 13 separate index records shown in sorted order in the right-hand column. Notice that the 'Addition' and 'Division' (main entry) lines are now completely separate and exist completely independently of their former subentries (which appear as main entries under the O and S alpha groups).

In order to import indexes with subentries into dedicated indexing software, conversion utilities must contain the intelligence to retain the implied links within the index entry groups. In addition, also consider formatting complexities potentially present in index entries:

- varying character styles (bold, italic, underline, SMALL-CAPS, subscript, superscript)
- diacritics and Unicode characters
- varying suppression (abbreviation) rules for page locator ranges (101–2, 345–57)
- multi-part page locators (A.1–A.2)
- cross-reference placement and formatting.

Any index conversion utility would also need to recognize such conditions and to provide a mechanism to import the

Figure 1 A sample index and the way it might conventionally be imported to an indexing program

index entries with their settings and formats intact. That's where third-party utilities like IndexConvert come in.

IndexConvert overview

IndexConvert is a Word macro developed by Barry Campbell, a member of SI.[2] It uses a multi-step methodology (see Table 1) to examine and prepare indented indexes (in DOCX/RTF format) prior to importing them into dedicated indexing software such as Cindex,[3] Macrex,[4] or Sky Index[5]).

Due to its complexity, IndexConvert requires significant interaction and decision making during setup and processing (Figure 2 shows the main menu). Users relishing hands-on control at such a detailed level will enjoy the options provided during each step of the process. However, it is not meant for indexers who want turnkey solutions or minimal involvement during processing.[6]
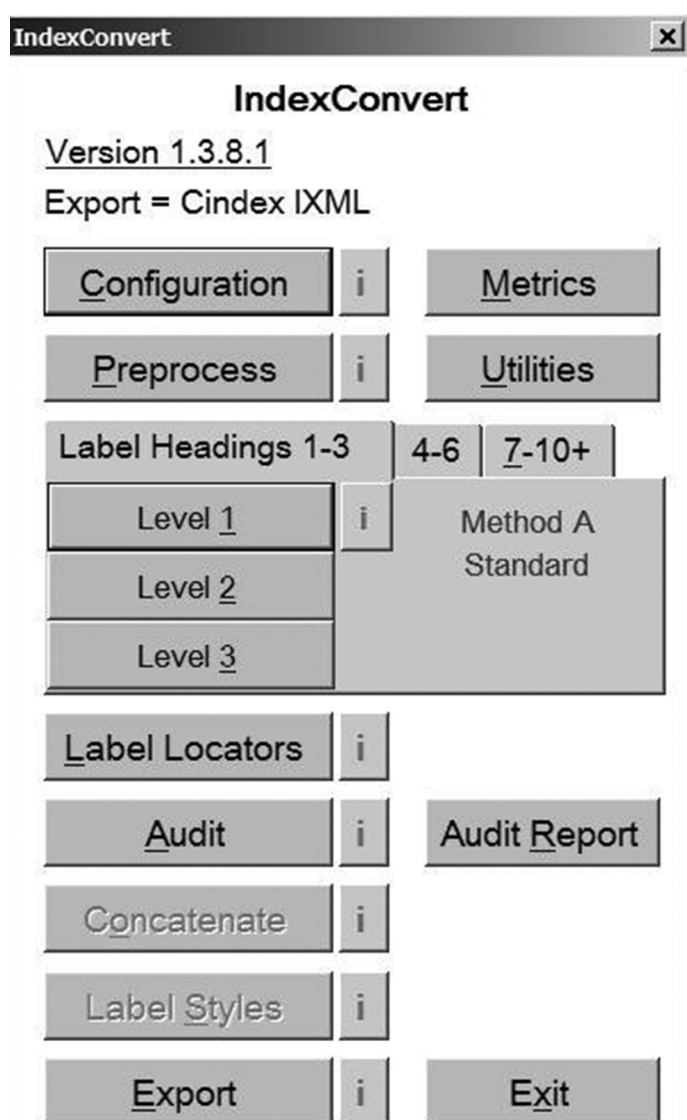
Licensing

Barry has developed a licensing structure based on the number of indexes and the total number of entries to be converted over the course of the license period. In order to select a licensing strategy would-be licensees need to estimate their expected 'conversion' workload. One-year, five-year, and three-year 'reserve' license periods are available. To clarify the options, the IndexConvert website provides a licensing spreadsheet that lists all the combination terms/entries and the prices for each. Options in the December 2016 pricing brochure included a three-year reserve license for 1,000 index entries for £25, and a five-year license for 10,000 entries for £150.

The three-year reserve period seems to provide more flexibility in term and in number of entries/indexes. However, all licensing options selected can be changed if warranted as the term progresses. My suggestion would be to overestimate what you think you will need. Then contact Barry when purchasing IndexConvert to ask his opinion regarding the best option for your expected workload.

Table 1 IndexConvert checklist

Save index as DOCM file	
Verify version in use	
Configuration*	Enter license key (one-time only). Select export option.
Manual preparation	Manual steps to clean up extraneous information not needed for index conversion (headers, footers, etc.).
Preprocess*	Removes multiple paragraph marks, manual line feeds, and alpha group headings.
Label Headings 1–3* Label Headings 4–6* Label Headings 7–9*	Index entries are labeled regarding hierarchical level in the index. Up to 9 levels are possible ($H1_ to $H9_).
Label Locators*	$L_ label added to recognized page locators and cross-references.
Audit*	Inserts $Er labels where potential errors exist.
Audit Report*	Review and clean up any $Er… labels prior to proceeding.
Save File	Save updated DOCM file after labeling, audit review, and error correction steps have been completed.
Run Elided utility	Separate utility provided with IndexConvert (if necessary, expands abbreviated page ranges (examples: 133–34 to 133–134; 102–3 to 102–103; etc.).
Concatenate* (Not applicable for Cindex – IXML option)	For tab-delimited export options, prepends parent headings onto related subentries in a spreadsheet structure.
Label Styles* (Not applicable for Cindex – IXML option)	For tab-delimited export options, labels bold, italic, underline, subscript, superscript characters.
Export*	Tab-delimited files are exported using SAVE command. IXML files are saved using an IXML generator.
Import into dedicated indexing software	Import according to DIS instructions.
Remove labels and adjust styles for tab-delimited files	Follow the instructions in the back of the manual on the most efficient way to remove labels and adjust styles as needed.
File compare	Manually compare imported index with original index. Correct any inconsistencies.

* indicates IndexConvert menu item.

Figure 2 The IndexConvert main menu

Trial version

A trial version of IndexConvert is available for download. It provides the full functionality of the purchased product but is limited to ten indexes containing up to 100 entries each, with a maximum of 500 total entries for the entire trial period.

Installation

The IndexConvert solution consists of two macros: IndexConvert and Elided. The IndexConvert macro provides the complete conversion functionality. The Elided macro expands suppressed/elided page ranges into their full form prior to importing. Elided installs automatically but as a separate macro.

In addition, IndexConvert provides two functions (TextConvert and CharConvert) for use outside the main conversion process. These functions aid the entry of diacritics for Sky Index or Macrex indexes. IndexConvert's user manual describes how to use them. (They are not needed for the Cindex import process if the IXML option is selected.)

The installation process is quick and simple, but it doesn't create a custom tab for easy and repeated access to Index-Convert and Elided. To help indexers who might not be familiar with how to create custom tabs, I have prepared a detailed explanation that is available on my website resources page (`http://luciehaskins.com/resources.shtml`).

Documentation

The IndexConvert documentation consists of a user guide explaining how each component works and a checklist (Table 1) setting out the step-by-step processing. Both are available for download from the IndexConvert website. IndexConvert shares the problem of all technical documentation, that the more comprehensive and detailed it is, the more difficult it becomes to find specific pieces of information. But as I explain below, between them the user guide and the quick reference checklist do provide the essentials.

Support

While reviewing IndexConvert, I contacted Barry many times via email. He was always responsive to my questions and attentive to suggestions I made. However, as always with technical support, to provide the optimum chance of getting a timely solution it is wise to try to anticipate your needs and submit your request in as early in your workload as possible.

IndexConvert analysis and review

I started reviewing IndexConvert shortly after its release in 2016 and worked through many versions, starting with 0.8.2 and finishing with 1.3.8.1. I used Microsoft Word 2007 and 2010, and Cindex 3.0 on a Windows 7 Professional operating system.

As with any new product I review, I put IndexConvert exhaustively through its paces. I provided feedback to Barry on functions and process steps I found difficult to understand, and made suggestions on product improvement. Over the course of the year, Barry implemented many enhancements (his own and some based on my suggestions), making IndexConvert increasingly robust and efficient with every update.

Because IndexConvert is a multifaceted program, my learning curve was steep, and it took me time to become proficient in how to apply each conversion step correctly. I depended heavily on the IndexConvert documentation throughout my testing and review period. I found the user guide to be thorough and the additional quick reference checklist an essential aid. Between these two sets of documentation, I retrieved all the information I needed: either to explain the process in general or to guide me quickly through the various steps.

As I am a former computer programmer, getting heavily involved with utility functions is well within my comfort zone, and I especially relished the deep level of control I had within each process step. But I wondered if perhaps IndexConvert might require more involvement than most

users would want. It therefore helped me immeasurably to query fellow indexers (and early IndexConvert users) Max McMaster and Pierke Bosschieter on their experiences. This is what they said.

Max McMaster:[7]

The program ran brilliantly. There were some 2,250 entries in the file (of a detailed gardening book). The whole conversion from start to finish took me around 50 minutes, which included fixing up a few gremlins once the file got to Macrex.

I was so impressed. Just to see how long it would have taken me to convert the index the old fashioned way, I converted a couple of pages of the index, then extrapolated. I estimate the conversion would have taken me around 3 hours. So a saving of around 70–75% of my time!

Pierke Bosschieter:[8]

I converted 5 indexes. The smallest was 910 entries and the largest was 3,100 entries. The shortest took 25 minutes, the longest took around 60 minutes.

Even diacritics in Arabic names went well, with the input of a translation file. This was important for me, as I almost always work with a lot of diacritics.

When an entry ended with a number or year, I got errors, but I cleaned this up before conversion (searching for the $Er in the Word file.)

I have also used IndexDeconstructor, but that only works for one level of subentries. When I am asked to work with an existing index it is almost always an index with 2 or more subentry levels.

I am very impressed. It took some getting used to, but once I knew the ropes, I felt really at home and I knew exactly what would be asked of me.

I do find the pricing structure very confusing. I have years in which I don't have to convert any indexes, while the next year has three such assignments. I'd rather pay for the whole package once, and pay for major updates once in a while.

Conclusion

Max and Pierke's feedback validated my impression that there were others who valued this comprehensive utility for its functionality. This is not a program for indexers looking for a one-off turnkey solution, but it would be a significant aid for the indexer who is called on quite frequently to convert an index, probably justifying the time and patience needed to master the process. It would particularly appeal to the indexer of a technical bent who wants to peek under the covers and likes being involved in all the decisions and in every facet of the conversion.

For anyone interested in IndexConvert, it costs nothing to download the trial version and try it out. Satisfy your curiosity and see what it can do.

Notes

1 I described implied links in my *KeyWords* article 'Demystifying IndexDeconstructor' (**18**(1), January–March 2010). Visit my website page: `http://luciehaskins.com/resources.shtml` for a link to the article.
2 Additional information for IndexConvert is available at `www.indexbase.co.uk`
3 Cindex Indexing Software for Windows and Mac: `www.indexres.com`
4 Macrex indexing software: `www.macrex.com`
5 Sky Index Professional indexing software: `www.sky-software.com`
6 Leverage Technologies (`www.levtechinc.com`) and Indexing Research (`www.indexres.com`) provide turnkey conversion services for converting existing indexes into a format that allows importing into dedicated indexing software. Additional information about IndexDeconstructor (a comparable-in-function but simpler conversion tool) can be found at `www.editorium.com/indexdeconstructor.htm`. See also note 1 for a link to my *KeyWords* article on IndexDeconstructor.
7 Email communication with Max McMaster, 15 April 2016.
8 Email communication with Pierke Bosschieter, 27 August 2016.

Lucie Haskins specializes in embedded indexing and in computer- and business-related topics, and enjoys discovering and applying tools that optimize her indexing process.
Email: `luciehaskins@gmail.com`