

- the ASI website (asindexing.org):
 - A three-part online webinar on Indexing in InDesign (full disclosure: I'm the instructor).
 - DTF (Digital Trends Task Force) maintains a plethora of information on trends in digital publishing.
 - Various chapter/regional/national conferences occasionally provide embedded indexing workshops.
- UC Berkeley had an Embedded Indexing course (but its status is unknown at this time).
- Resources pages on my (and other indexers') website provide step-by-step articles on indexing in FrameMaker, Word, or various plugins.
- SI embedded indexing course at www.indexers.org.uk/index.php?id=557
- Individual indexers known for their expertise might be amenable to providing one-on-one tutoring (it never hurts to ask!).

Conclusion

Remember, knowledge is power. I hope that now that you have more information, you can answer my original question: Should I jump on the embedded indexing bandwagon?

What do you think? Should you?

Lucie Haskins became an indexer in 2000 after a long career in corporate America, with roles spanning the computer industry and management consulting. She specializes in embedded indexing and in computer- and business-related topics, and has presented workshops on various indexing topics, instructed for UC Berkeley's online indexing course, and served ASI in chapter and national positions as well as on its webmaster team. Email: luciehaskins@gmail.com Website: luciehaskins.com

Does embedded indexing have a future?

Cheryl Landes

Cheryl Landes, an indexer who specializes in working with new technologies, discusses the ways in which indexing tools and techniques have changed over the years, and suggests, on the basis of her recent experience with clients, that traditional embedded indexing may be becoming a thing of the past.

While the concepts behind the craft of indexing have remained the same over the years, the techniques and tools we have to create indexes are changing. The advent of ebooks in the publishing industry accelerated this change. At the same time, the growing popularity of ebooks has forced publishers to look at new ways to produce books and indexes in an efficient, cost-effective manner. Embedded indexing is an effective way to tag content from an original source and create books in various formats that can be read on a variety of devices or in print. Single-sourcing content like this reduces labor and production costs and increases profits for the publisher.

Traditional embedded indexing meant tagging content in a word processing program, such as Microsoft Word, or a desktop publishing package, such as Adobe InDesign, QuarkXpress and Adobe PageMaker. However, when publishers want to single-source content, all of these programs are limited in their abilities to create the different types of formats readers want. As a result, publishers are turning to new, often custom, methods of creating and producing content, and indexers are also using different methods to create entries.

This article focuses on the future of embedded indexing, based on my experiences and observations as an indexer who works with new technologies. My primary indexing target market is technical materials, but I have also indexed hundreds of trade books since 1991, the year I started indexing. On the technical side, the majority of my work is embedded indexing, creating keywords for online help systems and web-based content, and optimizing content for search. In the 1990s and early to mid-2000s, my non-

technical (trade book) clients provided hard-copy (printed) manuscripts or PDFs, and I used my dedicated indexing software (Macrex) to create the indexes. I delivered my final indexes in either Microsoft Word or rich text format, and a graphic designer laid out the index in a desktop publishing program.

When ebooks became popular, my non-technical publishing clients began asking me to create embedded indexes in InDesign. They sent me the original manuscript files, created in InDesign, and I would insert the index entries in the files with InDesign's tagging features. When I finished, I would return the tagged files to the graphic designer, who would check the layout of each of the files in InDesign, make adjustments as necessary, and regenerate the index in InDesign before the files went to final production.

During the past five years, however, I have noticed a sharp decline in my clients requesting traditional embedded indexing. They are switching to content management systems, where the layout for the books is created separately from the content through templates, more commonly known as stylesheets. These stylesheets are applied when the content is ready to publish. With this process, the publisher receives the manuscript from the author in a word processing or text file, or more commonly during the past couple of years, the author enters the text directly into the content management system. Then the publisher tags the content with the appropriate style names to match the stylesheet, such as the book's title, main headings, subheadings, paragraphs, bulleted lists and other layout elements. Some of the newer content management systems I use have the tags set up so

that the author can even apply them, and then the publisher checks the layout before the content is published. Most of the systems I have used are either XML-based or in a wiki, but XML is the most popular because of its flexibility.

A simplified version of the publishing process in an XML-based content management system is shown in Figure 1.

When clients hire me to index books created in these content management systems, I create the entries either directly in the raw content by tagging or by a method outside the content, which is integrated into the content later. The two most common methods I use now are described below: directly tagging content in a repository, and indexing by paragraph numbers.

Directly tagging in content repositories

When a client asks me to create an index directly in a content repository, they give me a user name and password to access the content. By the time I start indexing the content, the client has added the tags for headings, paragraphs, bulleted and numbered lists, and other layout elements to match the stylesheets for each type of output (ebook, print and so on). The client provides special instructions for me to create tags, such as the format and name of the tags to use, separators for main entries and subentries, and formatting for cross-references.

Table 1 shows an example of tags used in indexing wiki content for one client.

After I finish inserting the index tags in the content, I notify my client. Then the client compiles the tags into a formatted index and sends the index to me to review. Then I edit the index by making changes in the tags. After I finish,

I notify the client again, and the client compiles the index again for a final edit pass. After I finish the final edits, the book goes into production.

Indexing by paragraph numbers

Indexing by paragraph numbers is becoming the most popular method of creating indexes for single-sourced content. With this method, my clients number each paragraph with XML tags in the content repository and generate PDFs of each chapter. I index the chapters from printouts of the PDF using Macrex. Instead of using page numbers, I use paragraph numbers. If a discussion spans more than one paragraph, I index the paragraph numbers as a page range. A detailed example of this method is shown in Figure 2.

My index for this chapter, using paragraph numbers as reference locators, was created in Macrex:

Baker City, Indian raid near, 52–58
 Carlisle farm, Indian raid on, 52–58
 cats, as target of jokes, 52–58
 feathers on a cat, 56–57
 felines. See cats
 Indian raids, jokes pulled during, 52–58
 jokes, cat as target of, 52–58
 molasses, cat dunked in, 56–57
 Mouser (cat), 52–58
 Native American raids. See Indian raids
 Oregon, Indian raids in, 52–58
 raids by Indians, jokes pulled during, 52–58
 Upper Willow Creek, Indian raids at, 52–58

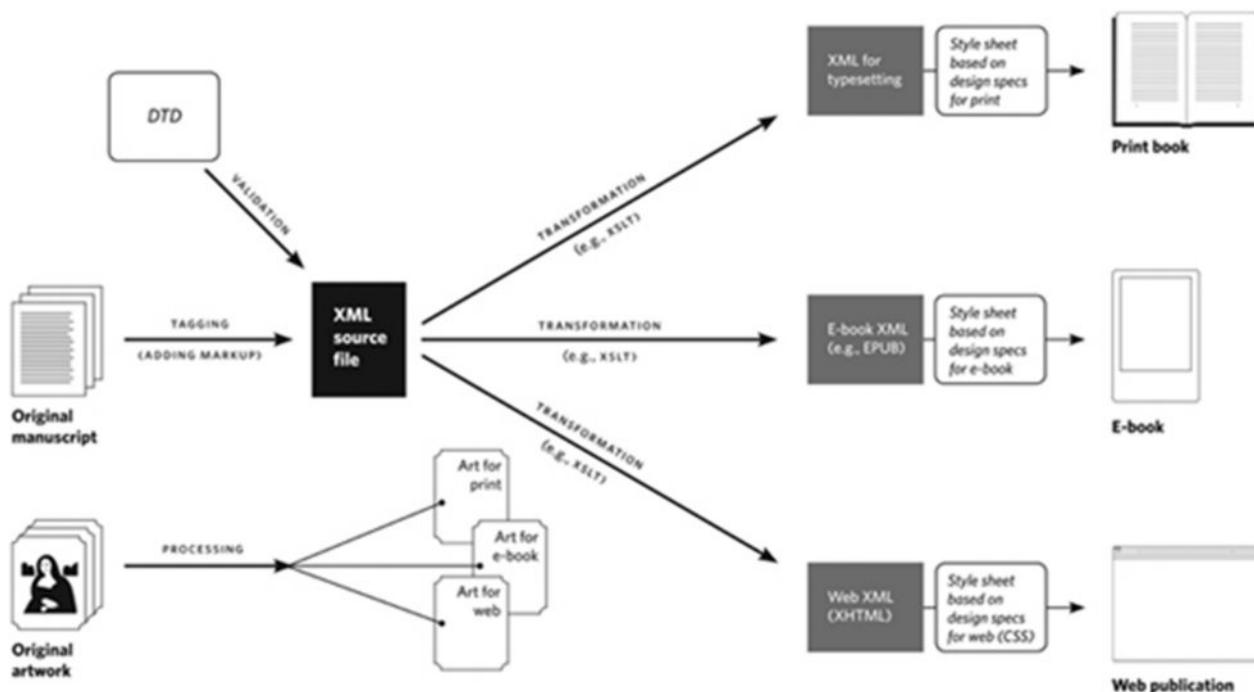


Figure 1 A simplified version of the publishing process in an XML-based content management system

Source: Chicago Manual of Style Online, 16th edition, www.chicagomanualofstyle.org/tools_workflow.html

Table 1 An example of tags used in indexing wiki content

Single entry, no page range	<p><code>\{in main entry; subentry; sub-subentry}</code></p> <p>Example of a tagged single entry: <code>\{in cats; sleeping patterns of}\{in sleep patterns, cats}</code>Cats sleep an average of 15 hours a day.</p> <p>Displays in the book as: cats sleeping patterns of, 43 sleep patterns, cats, 43</p>
Page ranges	<p>Start range with <code>\{is id; main entry; subentry; sub-subentry}</code> End range with <code>\{ie id}</code></p> <p>The indexer needs to provide an ID number for each page range so that the correct ranges display in the index. My client uses a two-digit chapter number, followed by a two-digit entry number, such as: Start range - <code>\{is id ch01-01; main entry; subentry; sub-subentry}</code> End range - <code>\{ie id ch01-01}</code></p> <p>(bold shown above for emphasis – not used in the tagging)</p> <p>Example of tagged page range:</p> <p><code>\{is id ch06-01; Sand Island, fishing at}\{is id ch06-02; Columbia River; Sand Island, fishing at}\{is id ch06-03; Washington (state); Sand Island, fishing at}</code>Sand Island, a small piece of land in the mouth of the Columbia River, had been a menace to navigators but a blessing to <code>\{is id ch06-04; fishing, Sand Island}</code>fishermen for decades. Navigators dreaded sailing in this area, because constant erosion and silt deposits continually changed the shape and location of the island. Fishermen saw the waters surrounding Sand Island as a gold mine. Average <code>\{is id ch06-05; salmon fishing, Sand Island}</code> salmon catches ranged from 300 to 500 tons each season, the largest in the world.</p> <p>....</p> <p>Sand Island became a memory 36 years later. When the north and south jetties were built on the Columbia River in 1932, the island was completely covered with water. However, alluvial deposits eventually caused the island to resurface in a smaller form, closer to Cape Disappointment. The salmon runs have never vanished, and fishermen return for their share every year.<code>\{ie id ch06-01}\{ie id ch06-02}\{ie id ch06-03}\{ie id ch06-04}\{ie id ch06-05}</code></p> <p>Displays in the index as: Columbia River Sand Island, fishing at, 15-18 fishing, Sand Island, 15-18 salmon fishing, Sand Island, 15-18 Sand Island, fishing at, 15-18 Washington (state) Sand Island, fishing at, 15-18</p> <p>(Source of text: Cheryl Landes, <i>Those wild Northwest days</i>. Victoria, BC: Trafford Publishing, 2006.)</p>
Cross-reference: See	<p><code>\{ic term; term-to-see}</code></p> <p>Example of a tagged See cross-reference: <code>\{ic felines; cats}</code>Cats sleep an average of 15 hours a day.</p> <p>Displays in the book as: felines see cats</p>
Cross-reference: See also	<p><code>\{ia term; term1, term2, term3, ...}</code></p> <p>Example of a tagged cross-reference: <code>\{ia felines; tigers}</code>Tigers sleep from 16 to 20 hours a day. In the tropics, they often sleep in water holes to protect themselves from insects.</p> <p>Displays in the book as: felines see also tigers</p>

 **CINDEX™**
for professional indexers

**Create
a new
word order**

*Simplicity, power,
flexibility & sophistication*



Download your **FREE**
demo version and find
out all you need to know at

www.indexres.com

Student and Publishers' Editions are available. Please contact
us to find out how these might best suit your needs

CINDEX™ is proudly developed and supported by

 **Indexing Research**

technical or sales enquiries please email:
info@indexres.com

The compiled index for this chapter, which is published in the book:

Baker City, Indian raid near, 5–6
Carlisle farm, Indian raid on, 5–6
cats, as target of jokes, 5–6
feathers on a cat, 6
felines. See cats
Indian raids, jokes pulled during, 5–6
jokes, cat as target of, 5–6
molasses, cat dunked in, 6
Mouser (cat), 5–6

```
<H1 n='50'>A Cat of a Different Feather</H1>
<p n='51'>Thousands of stories exist about the lives
of American pioneers, but not many are told or written
about pioneer cats.</p>
<p n='52'>Mouser was a cat that lived on the Carlile
family farm at Upper Willow Creek near Baker City in
northeastern Oregon. His life was the same as that of any
average farm cat of 1878, until he became the target of a
joke during an Indian raid.</p>
<p n='53'>At that time, Indian raids were common in
northeastern Oregon. When tribes went on the warpath
and began to raid the countryside, riders on horseback
tried to warn their fellow settlers before the attack. The
homeowners then fled, taking refuge wherever they could
in the brush, timber, or mountains. They did not have
time to gather their belongings.</p>
<p n='54'>Therefore, when Indians moved in on Upper
Willow Creek, the Carliles left Mouser in the house as
they fled for safety. When they returned, they found the
house ransacked, as expected. But they also saw a small,
strange, four-footed creature dash past them. Feathers
were scattered everywhere.</p>
<p n='55'>When Mr. Carlile managed to get near the
animal, it meowed at him. It was Mouser!</p>
<p n='56'>Closer investigation unraveled the mystery
of Mouser's new attire. When the Carliles fled, they left
a sealed keg of molasses in the kitchen. It was not open.
Apparently, the taste did not appeal to the Indians.
Instead, they had grabbed Mouser, dunked him in the
molasses, and turned him loose. As an added touch, it
seemed, they ripped open the feather bed so the feathers
would stick to the cat's body.</p>
<p n='57'>Of course, it was impossible to salvage either
the molasses or the bed. But after a long, wild chase, the
Carliles finally succeeded in catching the frantic Mouser.
They sheared and bathed him, and eventually, he became
a normal-looking cat again.</p>
<p n='58'>Fortunately, Mouser witnessed no more
Indian raids. He lived out the rest of his life as a happy,
carefree cat on Upper Willow Creek.</p>
```

Figure 2 Example of a tagged chapter by paragraph number

Source of text: Cheryl Landes, *Those wild Northwest days*, Victoria, BC: Trafford Publishing, 2006.)

Note: Paragraph numbering conventions vary by publisher.

Native American raids. See Indian raids
Oregon, Indian raids in, 5–6
raids by Indians, jokes pulled during, 5–6
Upper Willow Creek, Indian raids at, 5–6

After I finish indexing, I create a Word file of the index and send it to the client. Then the client runs a script that migrates the index to the content management system, embeds the index entries into the content by paragraph number, and translates the paragraph numbers into page numbers. When migration is complete, the client compiles the tags into a formatted index and sends the index to me to review. If I need to make any corrections, I edit my Macrex file and send the client an updated index in Word. The client updates the index using the same migration process, compiles the index, and the book goes into production.

Will embedded indexing survive?

Based on my experiences, embedded indexing is losing popularity for trade indexes. I have not received any requests to create an embedded index during the past two years. Clients now ask me to use the paragraph indexing method described above. I believe as technologies change, this process will become the preference for indexing. Indexers also find the process faster and more efficient, because they can use their dedicated software and index in an environment in which they are familiar.

*Cheryl Landes founded Tabby Cat Communications in Seattle in 1995. She has indexed since 1991, when she began her career as a technical writer at a marine transportation company in Seattle. Today, she is a technical and marketing communications consultant and specializes as a findability strategist. She has published five books and contributed chapters about findability to *The language of content strategy* (XML Press, 2014) and *The language of technical communication* (to be published in 2016 by XML Press). Cheryl currently lives in Vancouver, Washington. Email: clandes407@aol.com*

A look at classification and indexing practices for elementary school children: who are we really serving?

Valerie Nessel

As indicated in the findings of a larger study investigating the information-seeking behavior of Grade 3 students it is asserted that traditional classification and indexing methods used in school libraries and print reference materials targeted at young students often do not address young searchers' unique information needs and searching behaviors.

The findings from a broader study of the information-seeking behavior of Grade 3 students conducted in an elementary school located in a suburb of Montreal, Quebec, Canada indicate that the classification of print materials within an elementary school library may not be intuitive to young searchers. Furthermore, once a book was found, often the indexing practices did not facilitate information retrieval within the text itself. In order to address this problem, librarians and indexers need to be aware of the rapid changes in intellectual development that occur during a child's tenure in the first six grades of elementary school and how they can affect their information-seeking behavior.

The participants in this particular study were searching for information for a project on Canadian animals in winter using both print and electronic resources. This paper concentrates on the former. In the school library, it was observed that the students were not in the habit of consulting the card catalog before searching for materials but instead preferred to browse the shelves, which in the non-fiction section were identified with simplified labels (such as Animals and

Insects). As a result, they did not rely on the call numbers for locating suitable books; rather, they tended to use visual cues (such as book covers and shelf labels).

For this particular study, at the teacher's request, relevant books were removed from the school library, shelved by call number and placed on a small bookcase in the classroom. This resulted in a rather artificial environment as it meant that it was not possible to observe the students search for print resources in the actual library. (Cooper (2002a) employed a similar technique in her study of the information-seeking behavior of 7-year-old children.) Yet, although it was somewhat artificial in terms of a public library situation, in the classroom setting this was a common practice and occurred even within the school library. For example, sometimes a teacher would request the library staff to select and remove books on a specific subject from the regular collection and place them together for easy access by the students in the class.

Even though there were relatively few books on the subject (under 100), observation over a period of weeks