



# The **Indexing** Revival

BY SETH A. MAISLIN, *Senior Member, Boston Chapter*

Over the past few years, the technical communication job market has undergone quite a few changes, and most have been negative. For some reason, however, my indexing career hasn't changed much.

No, actually that's wrong. It's changed a lot. It's gotten *busier*.

And it's not just me. The newest indexers are as worried, as always, about finding their first contracts, but the many established indexers I've spoken with are rolling in work. Business is good. They've never been in greater demand. That demand is coming from everywhere, from publishing houses and educational institutions to companies of all kinds, large and small, domestic and overseas.

It's not hard to understand why.

### The Struggle to Publish

When the economy or the marketplace gets tight, struggling publishers don't wait for the right moment to release new titles. They just go for it, hoping to maintain a consistent and reliable cash flow, anxious not to disappear even momentarily from their customers' view. Publishers cheaply rebuild their old books into marketable newer editions, touting the newest utilities and versions. Software companies are quicker to post bug fixes and updates on the Internet. And everyone is looking to the international market as a new source of cash.

The result is more documentation, faster. And every new chunk of documentation needs indexing.

Even as the economy starts to move in a positive direction, companies aren't quite ready to return to their hurry-up-and-wait approach to business. The financial benefits of *hurry* far outweigh the quality ideals of *wait*, at least for now. This is the search-engine ethic: get it on the Web now, worry about fixing it later. Even some book publishers have experimented with this, distributing beta versions of their textbooks at low cost, intending to sell final versions later on.



Documentation is cheap. It's not enough to have a Web site any more; now you have to be a Web logger, writing a Web site *about* Web sites. Unfortunately, blogs are like diaries, often written in a stream of consciousness. Have you ever tried searching one for content? Sure enough, indexers are actually positioning themselves in the blogging market, ready to turn chronologically ordered entries into topically accessible knowledge.

As people find more reasons to publish content and more ways to get their content published, indexers will have more content to work with.

### The Struggle to Find (or Not Find)

"Information overload" has become a catch phrase. Undoubtedly there's a lot of information out there these days, and the quantity of available information is growing at unprecedented speed. Consumers are continually challenged to find better ways to filter out the information they don't need. These days, hating the World Wide Web for its overwhelming piles of unwanted content is as trendy as using the word *Google* as a verb.

Surviving an information deluge requires better metadata, better structure, and better keywords. In short, it requires indexing.

Writing indexes for individual documents isn't enough when hundreds more documents are competing for the user's attention. Whether a company is producing volumes and volumes of technical manuals, uploading hundreds of white papers to its Web site, attempting to retain and organize employee knowledge and activity on its intranet, or competing for attention on the World Wide Web, someone has to organize all of the documentation as a whole.

A number of indexers are finding themselves involved in taxonomy building and information architecture. Others are returning to the indexing profession's roots in library science, constructing controlled vocabularies and thesauri. In the end, we need indexers



on both sides of the communication process: at the point of information creation and at the point of retrieval. Producers of *unsolicited* content—spammers, for example—can't survive in a marketplace where we improve our ability to analyze retrieved content. To some, that's the best reason to index right there.

### The Struggle to Keep Up

Looking back over this article, you might think there's no trend here. Indexes and metadata have been around for a long time, so there's nothing really new going on, right? In actuality, however, the indexing industry took a huge dive ten years ago. Can you guess why? Two words: *search engines*.

The first impressive demonstrations of Web-based search engine technology were simply boasts about quantity. Search engine performance depended on finding as much of the Web as possible. If a good search engine could give you ten thousand hits, an amazing search engine could give you a million. Technical communicators relied on both search engine technology and the user's expertise with search engines to make their documentation usable. "Throw it together, people will find it." Besides, working with search engines is cheap and scalable, whereas writing indexes—well, let's just say there are a lot of people who would rather sweep the dust off the Mojave Desert.

Thankfully for professional indexers, that's still true. Few people want our jobs.

It's almost impossible to believe that we bought into this quantity-over-quality rhetoric. Ask people why they like a particular search technology and they'll answer, "Because what I'm looking for is usually in the first three or four results." Nobody wants a million hits any more. They want three. If you're still not convinced, count the number of educational courses and books dedicated to end-user search technique. My favorite book is *Google Hacks*, published by O'Reilly & Associates, which lists 100 dif-



ferent approaches to improving your Google.com-based research skills. Who could have predicted that the Google Web site would need a 352-page manual? With an index!

Free text searching is an inadequate tool for finding things when the quantities of information are as large and varied as they are. Information is coming too fast; search terms have become overloaded. The indexing industry is jumping in to compensate with concept analysis, term differentiation, and controlled vocabularies. No longer is it acceptable to publish without a good index, to build a database without editorially selected keywords, or to construct Web pages without intuitive keywords and a functional architecture. Even search engine algorithms themselves take advantage of dynamic, editorially created categorizations to help users browse results lists, or to filter the scope of their searches. It's all indexing.

#### Indexing Training

When I gave my first telephone seminar with STC, called "A Brief, Comprehensive Indexing Primer," it attracted a lot of attention among members. I gave the talk again this past October, as well as two other indexing seminars in between, and they all garnered large audiences. There are people who want these skills, *need* these skills, and they're paying to get them.

As the economy starts to return to normal, and as companies move away from their recession-based policies, the amount of money budgeted for training and consulting increases. During the recession, my consulting work faded while my indexing workload increased. Now, interest in my consulting services is returning to and exceeding normal, even while my indexing work remains stable. Other indexers are telling me the same thing. STC's Indexing special interest group and the American Society of Indexers (ASI) are both growing. ASI expects to launch a series of indexing courses in the summer of 2005.

Indexing has been on the back burner for too long. Consumers are



## Indexing has been on the back burner for too long.

complaining, and technical communicators are responding. Boosted by professional interest in information architecture and XML-based programming strategies, indexing and metadata construction are back with a vengeance. Despite innumerable attempts to automate the indexing process—most recently by harnessing the power of community-based behavior (e.g., if a site is linked to more, it's assumed to be more valuable)—the editorial nature of indexing continues to defy codification. Indexing documentation is even less automatic than copyediting it. Humans remain the most effective tool for pointing users to what they need. In other words, you can *try* to automate it, but you'll never truly succeed.

Non-English-speaking countries are expressing interest, too. So far, indexing is a phenomenon dominant in only the English language, although I don't know why. (Many French textbooks, for example, have simple tables of contents in the front and deep tables of contents in the back.) Additionally, English indexes are written well only by English

speakers, and indexes can't be translated; the brevity and context of index entries in any language demands a deeper understanding of word meanings and distinctions that is unavailable to nonnative speakers. Nevertheless, European and Asian countries are competing fiercely in global markets—as we all are—and so they are focusing more on that alphabetized list at the back of the book. Within the past year, Germany and the Netherlands began fledgling indexing societies. Technical communicators in India are motivated to learn about indexing because until they do, indexing work will remain primarily in the United States and Britain. They are anxious to learn the skills quickly, despite any language barriers.

#### Conclusion

In the information age, whoever controls the information wins. The only way to control the information is to know what you have and how to get it. Search engines are imperfect and insufficient; in our industry, perhaps they always were. Certainly they are valuable tools, but in more circumstances than ever before, they aren't good enough.

Last, it is worth mentioning that the role of the index within a documentation set is shifting in response to changes in technologies. What might appear as a *See also* cross-reference in a book is equivalent to a list of related topics in an online Help file, or a clickable link between hypertext documents. As more documentation is distributed online, indexes are growing more integrated into the documentation itself. For example, is there a difference between an index entry with multiple subentries and a Web page with multiple links? I don't think so. Perhaps the reason indexing is becoming more and more popular has to do with all the additional navigational tools available to writers of online documentation.

Regardless of the reasons, it is clear that indexing skills—and the indexers and writers who know how to use them—have never been more in demand. **1**

