



# n Interview with David Brown

**A to Z:** Introduction: David, tell us a little about yourself and your background, and the kinds of projects you do.

**DB:** With a degree in Computer Science and a couple of years experience, I was applying for sysadmin jobs when a recruiter asked if I was interested in tech writing. That was about fifteen years ago, and since then I've worked in software and hardware documentation almost exclusively.

About ten years ago, an employer wanted to move their DOS products to Windows, and I was asked to handle the technical aspects of developing the online help. I really enjoyed the fact that you could compile help as often as you liked and see the results of your changes immediately. I also liked the fact that customers actually used the help, and I've focused on online documentation of one kind or another ever since.

My indexing experience has been in Word documents to be printed or distributed as PDF files, Windows help systems, and HTML pages. In Word and WinHelp, you can embed index entries in the source files, a tremendous benefit in most software situations. That's what led me to develop HTML Indexer back in 1998.

**A to Z:** So you really have been focused on the online aspects and embedded aspects of indexing. Your path into indexing is much like many of the members of the Indexing SIG – coming in through tech writing, instead of the traditional route, from the library world. How do you think your online help experience shaped your approach to indexing?

**DB:** I index everything in pretty much the same way, regardless of whether I'm publishing in print (typically PDF) or online help. I usually know the material well; in fact, I'm usually the author. I go through the material, a page or a topic at a time, creating

every index entry that seems reasonable (flips, synonyms, and the like), and not worrying much about consistency beyond my memory of terms and styles I've decided to use in the project.

When I finish that pass, I compile the index and begin the refinement process. I combine subentries, fix stylistic inconsistencies, apply synonyms to related topics, create new entries that occur to me, and so on. I might do two or three such refinement passes, compiling the index as often as I like to see the result of my changes along the way.

Basically, I stop when I'm satisfied with the index. That's one of the luxuries of working on an hourly or salaried basis, as I usually do. As long as I deliver on time and stay within budget, I can spend as much time on the index as I need to get the level of quality I want.

**A to Z:** There's been a lot of discussion on Index-L about web indexing in the last few months. How viable do you think crafted indexes on the web are? By crafted, I mean something constructed carefully by humans, reflecting the *aboutness* of the content, rather than a search engine construct or taxonomic weighting.

**DB:** When we developed HTML Indexer, most web pages weren't generated automatically from databases, they

didn't have dynamic or personalized content. It's a lot harder (though not impossible, in most cases) to craft an index for pages that don't exist until they're displayed in someone's web browser.

As anyone who has tried to break into site indexing knows, site managers are lured by the ready availability of search engines. They're plentiful, free, and easy to put in place. It's very difficult to convince a site manager to spend the money and take the time to let a professional indexer create an index.



**Site managers are lured by the ready availability of search engines. They're plentiful, free, and easy to put in place. It's very difficult to convince a site manager to spend the money and take the time to let a professional indexer create an index."**

*Continued on next page*

---

**A # Z:** That's a good summary of what I see as well: the static page-based site is the one that we can really craft an index for, but the database-driven ones must be left to search, or the development of metadata. Have you considered developing a product like Metabrowser (metabrowser.spirit.net.au), an HTML Indexer for all the metadata fields? Metabrowser's free version only lets you manipulate metadata on one page at a time, but have you considered an adaptation of HTML Indexer for all the metadata across a site? That might be a viable migration of the technology you have worked on.

**DB:** Sure, we could store the index entries outside the source files, but then it wouldn't be embedded indexing. You'd lose the *certainty* of never creating a broken index entry, the ease of rebuilding an index when the filenames or directory structure changed. Henry Ford *could've* built sewing machines, but he was in the car business. (I haven't heard of the Metabrowser tool.)

**A # Z:** Could you elaborate on how HTML Indexer embeds its data and the protections that provides for indexing?

**DB:** HTML Indexer embeds any index entries you create in a special comment block in the invisible HEAD section of the HTML source files. Because the entries are embedded in the source files, you can move a file to another directory or another web site, even rename it, without losing any of your index entries. Just open your project in HTML Indexer, add the renamed or relocated file to your index project, and recompile the index—there's no need for laborious retyping of URLs to fix broken links.

**A # Z:** What would you consider to be the ideal web indexing project? What kind of sites are best for web indexing?

**DB:** For the reasons I mentioned earlier, a site comprising pages that are stored intact, rather than dynamically generated, is best suited for indexing—a site that changes, but not so quickly that the index *can't* be kept up to date. Assuming the content is at least mildly interesting, I guess that'd be the ideal site indexing project for me, too. (I'd like to point out that I'm talking about site indexing, rather than web indexing. It's an important distinction. Very few organizations have the resources even to attempt to index the web.)

**A # Z:** I recently used HTML Indexer to index a static site, just as you describe. It hasn't been posted yet, as the webmaster ignored my instructions to not change the temporary file content they had posted, and that the files I had were the "hot potatoes." So now we are faced with two sets of files: ones with indexing and bookmarks I inserted, and ones with content changes they made de-

spite all my warnings. Have you considered some way of storing the metadata outside of the files for situations like this, so you can "force" the indexing in? My best guess for these folks was to let me work with the changed files, since I understood the structure.

**DB:** One solution would be to carefully copy the HTML Indexer comment block from the HEAD section of each file you worked on to the HEAD section of the corresponding changed file. You'd be totally on your own, of course—we simply cannot afford to support activities outside the HTML Indexer UI—but there's no reason for it not to work, if you were careful to copy and paste it exactly as you found it. Once you've done that, you'd just examine the differences between the old and new version of each file, and adjust the index entries as needed. (I'd be happy to subcontract to do the copying part of the task, if you can zip up the file hierarchy. It'd still be up to you to take it from there, of course.)

**A # Z:** What was your favorite index created with HTML Indexer, and is it up on the web? What did you especially like about it?

**DB:** Without question, it was the one I created for AboutMyHealth.net. Working with one other tech writer, I helped to design the look of the help system, I edited the help text, and I indexed the content with HTML Indexer. An oversized screen shot of the finished product is included in the Customer Examples area of our web site (www.html-indexer.com). (*See page 10.*)

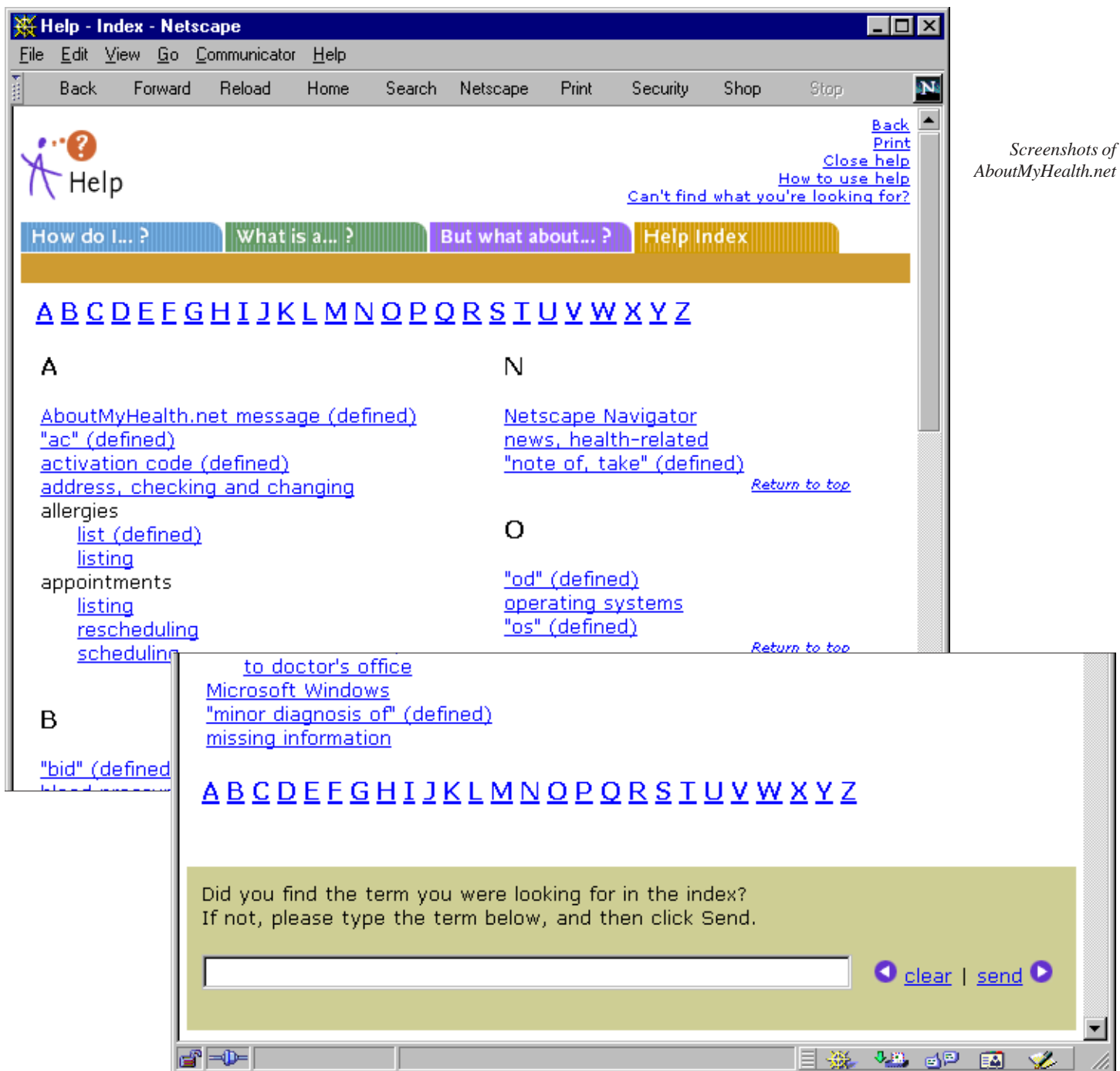
By the way, one of the things I like is how it's utterly integrated into the help system in a way that exactly matches the tabbed interface of the application itself.

**A # Z:** What are you looking forward to in the next year, in terms of new directions for your own learning or for HTML Indexer? What areas of lexicography are you going towards next?

**DB:** You know, given the state of the high-tech industry in which I have cast my lot, I'm really glad to be working with a small group of talented people who trust me to know what I'm doing. They're so far ahead of me, in terms of the kind of programming they're doing, that I stay challenged just documenting the software they create. That's what I do, after all—I'm a tech writer.

As for HTML Indexer, I don't see big changes. Version 4 finally includes all the features I dreamed of when I first recognized the need for such a tool, and the market is too small to warrant another major push without a truly compelling case for additional features.

*Continued on next page*



Screenshots of AboutMyHealth.net

I guess if I won the lottery, though, I'd add some of the features professional indexers expect: automatic flips come to mind, for example, maybe search and replace, a spelling checker... If it were a really big jackpot, I might even add an option for lazy tech writers who want the software to *create* the index entries for them—I hate the idea, but we'd sure sell a lot more copies! One thing to keep in mind, though: I don't buy lottery tickets. :)

**A # Z:** Have lots of people asked for concordance abilities in HTML Indexer?

**DB:** Ha! After seeing my posts on Index-L and Techwr-L over the years, I assume they know better than to ask.

**A # Z:** RoboHELP has a continuum-style concordance engine for adding index entries. I actually like setting it on low-suggest mode and seeing what it comes up with for a topic, but then controlling and refining and adding entries to their suggestions.

**DB:** There's no harm in using such a tool in the way you describe. The problem is that harried tech writers (and their Bud-

*Continued on next page*



# Indexing Glossary Entries, Yea or Nay

compiled by Bill Collins

Guru Kamath posted a message on our list serv discussing the pros and cons of indexing glossary items. Guru also sent the post to Index-L, receiving there a considerable amount of feedback. However, for those on the Indexing SIG list who do not subscribe to Index-L, I thought I'd provide a few of the comments from indexers on that listserv. The consensus appeared to me to be that, barring client directions to the contrary, the indexer should index the glossary.

---

"Some clients want the glossary indexed. In that case, I index the term with a qualifier of 'defined' or 'definition.' some clients simply want the term 'glossary' indexed....If the client wants it, do it."

"If the book contains a definition of 'widget' and that term appears only in the glossary, why would you not want to point the reader there? Some readers turn first and only to the index to determine what's in the book."

"The index is the one place where everything is brought together. A glossary is one of those things. I generally index it unless there's some special reason not to."

"If the function of the index is to indicate where information on a subject is located, then it makes sense to index the glossary. It may well be the only concise definition of a term....I realize that it is not

the convention in some fields and at some publishing houses. That's all well and good, and that needs to be respected. The client gets the final word there. But if left to my own devices, I would index it."

---

References can be cited to support any position on this issue. *The Chicago Manual of Style* (at least up through the 14th edition), 17.27, says that the glossary should not be indexed. Nancy Mulvany, in *Indexing Books* (1994, p. 43) states that, "As a general rule, the contents of the bibliography and glossary are not indexed. However...it may be helpful to readers to cite in the index the presence of the glossary...."

I'll leave the last word to Hans Wellisch, from his classic *Indexing from A to Z*. Professor Wellisch, a true icon of the indexing profession, died recently at age 83. This citation also appeared in Guru's original post:

"... Glossaries are often included in books on topics whose terminology is still unstable or full of neologisms and specific usages of terms by the author. The presence of a glossary should be indexed, and the terms defined in it (which will almost inevitably be employed in the text and thus have index entries) should also be indexed, preferably by a subheading 'defined' ...."

---

## Interview with David Brown, continued

get-conscious managers) will be tempted to accept the results of that initial automated pass, and call it an index. *That* kind of thing we *have* gotten requests for, from time to time.

**A # Z:** What's your opinion on the "automatic indexing" software out there?

**DB:** I haven't seen anything I'd consider using on any project.

**A # Z:** If you were asked by tech writers what skills and technologies they need to know to effectively do online help and indexing these days, what would your list look like? What's important for someone new to the field to know?

**DB:** Well, I think it's something you just have to have a feel for. Some people just get it, some people pick it up through exposure to lots of good and bad examples, it may even be possible to learn it in a classroom. You definitely have to be an excellent

writer, but you also have to be able to write in different voices than the one with which you're most comfortable. You have to be an absolute lunatic for consistency and for accuracy, down to the tiniest detail. Maybe that's it—you just have to be a bit of a lunatic.

What you *shouldn't* do is worry about tools. If you have any chance of earning a living as a tech writer, you should be able to get productive with any text publishing tool in *very* little time. Don't get fooled by the marketing hype and the religious zealots telling you their tool is the only way to produce high-quality results.

David Brown can be reached at [dmbrown@brown-inc.com](mailto:dmbrown@brown-inc.com). Information about *HTML Indexer* and a free demo version are available at [www.html-indexer.com](http://www.html-indexer.com).