How to
Index Online

Jan C. Wright
Wright Information Indexing Services

Jan C. Wright is the owner of Wright Information Indexing Services, which has provided print and online indexing to companies such as Microsoft, Asymetrix, Visio, PeachPit Press, and Sakson and Taylor. She has served as a judge for the Society for Technical Communication's annual publications competition, has a Master's Degree in Library Science, and uses her many years of experience as a professional librarian to help make her indexes more user-friendly. She is a member of the American Society of Indexers, and recently co-developed the Society's World Wide Web home pages. She can be reached at JANCW@mindspring.com, or at 206-784-2895.

Jan C. Wright
Wright Information Indexing Services

206-784-2895
JANCW@mindspring.com

http://www.mindspring.com/~jancw

Your first online index

Abstract: The first online index project introduces an indexer to a whole new set of issues and concerns that affect good indexing. This article covers some of the groundwork of decisions and information-gathering that should be done before beginning an online project.

Many back-of-the-book indexers are saying yes projects to be printed on CD-ROM, accessed through a web site, or an internal Intranet. Agreeing to do a project for one of these environments can be overwhelming: there's so much to learn and figure out, and it seems like your indexing skills are the only thing you feel you understand. In this article we'll take a look at some of the decisions and questions that need to be made and asked at the beginning of an online project, so that you can get issues cleared up and concentrate on the thing you know the best—indexing!

What is online indexing?

Online indexing ranges from embedding key terms from a thesaurus in neatly-defined fields in databases, to a CD-ROM index with pictures of dinosaurs beside each letter section, to a Web-based HTML index jumping to other sites, to the Windows Help file index that came with your copy of Word or Excel. The display of each online index varies widely, and the techniques and tools you use to index vary to meet that display.

Several factors can affect the methods you use to index online, how the index is displayed, and how your work will proceed. Your first step is to define the project specifications the same way you do when you are discussing a print project. It's a different environment, with more unknowns, but laying out the specifications is still the first step.

What are the boundaries?

One beginning question to ask is will this cosmos of information be infinitely added on to, like a corporation's intranet, constantly adding new memos, presentations and budget reports, and therefore constantly changing the indexing? Or is it a confined environment, like a CD-ROM on dinosaurs, with an index that will be fixed in place when it is shipped, with no later additions? This will impact your index plan: you tend to leave more detail in place when the material will be added to.

What are authoring tools?

Authoring tools are the software packages that build raw information into a usable online resource. These tools can operate in several ways, and you will need to find out which way your project is going to work. There are:

- Tools that allow index terms to be placed into a database field. These are usually the easiest to understand, and can be used with a thesaurus to make the indexing in the fields consistent. An example of this format would be a bibliographic database with full text of the article.

| Title |
| Bibliographic information | Other info |
| Keywords | Unique ID |
| Full text of the item to be indexed can be available within the record or within the system |

Database-style indexing

- Tools that insert terms into each informational topic, and are picked up and compiled into an index. The indexing for each topic may be coded into the text of the topic itself, or live in a footnote or attached field in the topic. When all the topics are compiled into the online product, the authoring tools build the index display from the codes found in each topic. The current version of Windows Help is an example of this format.

Monster Maker Pro

Monster Maker Pro is a unique tool that not only allows you to create the monster of your dreams, but brings it to life with animation and sound effects.

Click a topic below for more information:

Overview
Creating a monster
Animating a monster

| Topic Unique ID Monster01 | Monster Maker Pro: overview |

A typical topic with an index entry

Monster Animation

Once you have created your monster, you can animate it and save the animation as a movie. Before you begin, you will need to understand some things about monsters and their reactions to being animated. Click below for more information:

Tell me about mad monsters

Choosing an animation environment

| Topic Unique ID Monster07 | Monster Maker Pro: animation; Animation |

Another typical topic with index entries

Index

Animation
Monster Maker Pro: animation
Monster Maker Pro: overview

Topic-based indexing   The combined index entries in the index display

- Tools that build hard-coded indexes from a set of entries and electronic locators. The best example of this form of index is the American Society of Indexer's web site index, where each entry jumps to a specific location in a specific file. The indexing itself is done in a word processor or indexing software package, and then coded to work with the rest of the online system's files.

Monster Maker Pro

Monster Maker Pro is a unique tool that not only allows you to create the monster of your dreams, but brings it to life with animation and sound effects.

Click a topic below for more information:

Overview
Creating a monster
Animating a monster

Topic Unique ID
Monster01

A typical topic with its unique code

Monster Animation

Once you have created your monster, you can animate it and save the animation as a movie. Before you begin, you will need to understand some things about monsters and their reactions to being animated. Click below for more information:

Tell me about mad monsters

Choosing an animation environment

Topic Unique ID
Monster07

Another typical topic with its unique code

Animation (Monster07)
Monster Maker Pro
    animation (Monster07)
    overview (Monster01)

Topic Unique ID
MonsterIndex

Hard-coded indexing

Index file listing jumps in invisible codes

- Tools that take print electronic files, and convert embedded index entries and the files to an online format, with a clickable index. An example of this format would be Frame or PageMaker files. These software packages can convert embedded indexes in print-based files into an online file.

Monster Maker Pro

Monster Maker Pro (index:Monster Maker Pro:overview) is a unique tool that not only allows you to create the monster of your dreams(index:Dream monsters), but brings it to life with animation and sound effects.

Creating a monster

Before you begin, you must let Monster Maker Pro know which kind of color environment (index:Color issues) you are

5

A typical chapter with embedded indexing

Monster Maker Pro

Monster Maker Pro is a unique tool that not only allows you to create the monster of your dreams, but brings it to life with animation and sound effects.

Creating a monster

Before you begin, you must let Monster Maker Pro know which kind of color environment you are working in.
To do this, choose "Color" from the Setup menu, and then choose the

5

File converted to online screen with invisible anchors for index entries

Index
Color issues, 5
Dream monsters, 5
Monster Maker Pro, overview, 5

Embedded indexing
converted to online

Index screen: clicking on the number takes you to the right page

You will need to know where the indexing lives: is it one long hard-coded file like web-page indexes, or does a search engine find the indexing in each topic and present it in a display as a coordinated whole? This will effect how you do the indexing, and must be established before beginning.
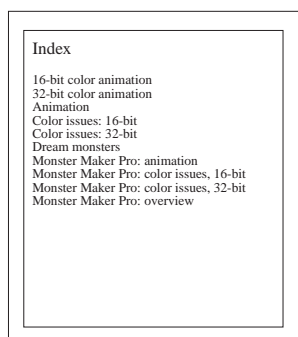
How will the index be displayed?

Another critical factor to be discussed before starting the indexing work is the index's display. You cannot begin to index effectively without having the index display settled, because you need to know how your work will appear to the user. Some of the questions you need to have answered are:

- How many levels of index will appear? Many tools allow only one level to appear, some allow two. You have to come up with workarounds in a one level index, such as using a semicolon to visually divide your entries into two levels:
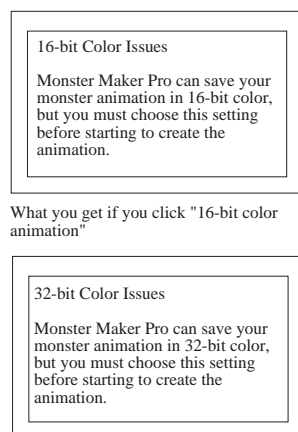
  Monster Maker Pro
  Monster Maker Pro: opening
  Monster Maker Pro: quitting

- When the user clicks on an entry, do they get a list of topics that have that term applied, or is there only a one-to-one correspondence between index entry and topic? An example of a one-to-one index would be an HTML index like the ASI web site: each click takes you to only one location. This type of index usually has more subheads under each topic than a one-to-many index, like the Windows Help index format.
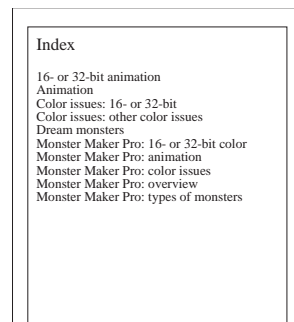
Index

16-bit color animation
32-bit color animation
Animation
Color issues: 16-bit
Color issues: 32-bit
Dream monsters
Monster Maker Pro: animation
Monster Maker Pro: color issues, 16-bit
Monster Maker Pro: color issues, 32-bit
Monster Maker Pro: overview

Each entry can lead to only one topic

One-to-one indexing

16-bit Color Issues

Monster Maker Pro can save your monster animation in 16-bit color, but you must choose this setting before starting to create the animation.

What you get if you click "16-bit color animation"

32-bit Color Issues

Monster Maker Pro can save your monster animation in 32-bit color, but you must choose this setting before starting to create the animation.

What you get if you click "32-bit color animation"

Index

16- or 32-bit animation
Animation
Color issues: 16- or 32-bit
Color issues: other color issues
Dream monsters
Monster Maker Pro: 16- or 32-bit color
Monster Maker Pro: animation
Monster Maker Pro: color issues
Monster Maker Pro: overview
Monster Maker Pro: types of monsters

Each entry brings up a choice of topics, which necessitates one more click for the user to choose the wanted topic

One-to-many indexing

Topics Found

16-bit color animation
32-bit color animation
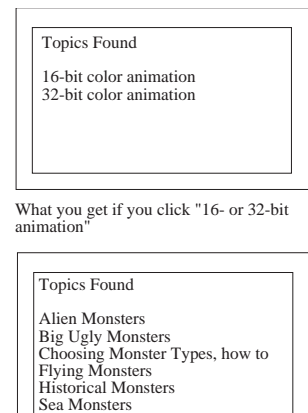
What you get if you click "16- or 32-bit animation"

Topics Found

Alien Monsters
Big Ugly Monsters
Choosing Monster Types, how to
Flying Monsters
Historical Monsters
Sea Monsters

What you get if you click "Monster Maker Pro: types of monsters"

- Are there any controls over the sorting of the entries? Often, the answer is "no" in compiled indexes or embedded indexes converted from print files. If you are hard-coding an index using a word-processor, you can control the sorting.

- How will you put in cross-references, and if the user clicks on them, what happens? You will often have to explain what cross-references are to the developers of the project, and in many cases, you won't be able to use them. This will impact how much double-posting you do.

- How will the user get to different sections of the index, by clicking a button or typing a letter? How the user navigates in the index is a concern: if the user types in plural words, and you have indexed with singular entries, they may scroll past the terms they need and never find them.

How does the indexing get into the files?

Once you find out where the indexing lives, and how it will be displayed, you need to settle the question of how it gets into the files. There are several ways of getting indexing into place, depending on the authoring tool for the project. You can divide authoring tools into closed systems and open systems.

Closed systems force you to work within a closed software package, where you can only use that tool to input entries into the system. These are the hardest to work with, as they usually do not show you an overall view of the index, are not designed with indexing needs in mind, and you are working in the dark. An example of this would be a database that has fields for keywords: you can see only one record at a time as you index. If you must use this type of blind tool, establish a thesaurus list, and decide whether working outside the files to build your index might save you from headaches before using the tool to place the codes in the files. You could build the entire index in your indexing program, using topic IDs or other electronic locators instead of traditional page numbers. Then sort by locators and enter the indexing into the authoring tool. Otherwise, your editing will be hampered by the inability to see what the index looks like.

Closed authoring tools usually will require you to work at the client's worksite, so that you can access their software and files. They may have additional constraints, such as only allowing one person to work in the files at a time. Make sure you find out about such constraints.

An open authoring tool will allow you to import and export indexing into and out of its file structures, usually with a spreadsheet-like data file. This openness allows you to work in a conventional indexing program, convert the index data to a data file, and then paste it into place. These open tools are very rare, and usually are written in-house.

Identifying someone in the project who can give you technical support for the authoring tool is critical, especially if this will be the first time the tool is used. If

there is no one to help you, budget time for debugging problems and figuring out solutions.

What kinds of files are included in the project?

An online index can jump a user to a topic, to photographs or sounds, to an animation file, to a menu of choices the user needs to make, to a Web site, or the index jump can start a macro running, opening files or formatting text. Interview the online system's writers and editors to get a full picture of what kinds of files are included, and any special details that you will be expected to work into the index.

One critical thing to ask about is the structure of the project and its menu screens, if they are used. Menu screens are table of contents screens that present a user with choices and then lead to specified topics. For example, if a user wants to explore a software package's help system, the first menu screen may be:

Help for Monster Maker Pro:
    Overview
    Types of Monsters
    Monster Making Tutorial
    Monster Movement Animation
    Monster Speech
    Monster Maker Pro Incompatibilities

The user is interested in "Types of Monsters" and clicks on that topic. Another menu screen appears:

Types of Monsters
    Big Ugly Monsters
    Intelligent and Smooth Monsters
    Horned and Scaled Monsters
    Historical Monsters
    Sea Monsters
    Flying Monsters
    Alien Monsters

Each choice on the menu screen may lead to 5 or 6 more topics in a menu on each type of monster. For instance, Big Ugly Monsters may have a menu screen with choices for George the Big Ugly Monster, Hank the Big Ugly Monster, and their friends. If you are indexing this product, and there are hoards of monsters, you could consider adding an entry "Monsters: types" and have it jump to this "Types of Monsters" menu screen, rather than listing each type out. You would always double-posting each individual monster directly under its name, of course.

Another type of topic structure you need to ask about are decision trees. Often, you can index to the top of a decision tree to force the user to go through steps they need to go through. For example, if you want to create a monster movie, you may not

know at this point that it is very important for Monster Maker Pro to know what kind of environment your monster movie will run on. You will need to make a choice between 16-bit color monsters and 32-bit color monsters before beginning, so the help system has a decision tree to force that choice before offering any help. If you click on "Monster Movies: creating," you will get this decision screen:

Creating Monster Movies

Before we begin to make a movie,
what type of color environment will it run
on?

16-bit color, or
32-bit color

The user has to make a decision before any further help is offered, and often the help screens that are offered are geared to the specific decision. The indexer can choose to index to the start of these decision trees, and ensure that the user makes the choices, rather than to the bottom level of instructions. You will need to find out if anything similar to this structure is present in your project, and plan to use it after discussing the pros and cons with the project staff.

Who's inserting the indexing and maintaining it?

Some projects will have you mark indexing on hard copies for someone else to key in. Others have you doing the index insertion for the entire project. You will need to establish a process for getting indexing into place, and it is a good idea to write the procedures down once they are decided. One question to raise with a project that will be expanding and adding new materials after the initial indexing is personnel: who will be maintaining the indexing efforts, and how will that process work? Expanding indexes usually require a master indexing list of used terms to be maintained, or a complete thesaurus. Otherwise, later efforts will not match what is already in the system.

Document as well how the entire process will work - how do you know when files are ready, who tells you, who reviews and proofs the index, who compiles it, and who hands off changes and updates to you. Figuring these handoffs and stages out ahead of time will save a lot of headaches.

Time frames

Estimating the time it takes for an online project is complicated by all these different factors: authoring tool problems, debugging glitches, difficulties in editing, etc. One estimate that seems to work is to allow 10% to 15% of total project resources (total person hours) for indexing.

Working with a known tool that has been used before will reduce the schedule; unknown and new tools need more time in testing and compiling. Here is a sample

schedule for a known tool such as Windows Help, with a smallish project of 300 topics. (Keep in mind that topics can range from 2 sentences to 3 or 4 pages. They shouldn't be 3 or 4 pages in terms of good online design, but they can be that long.)

Sample schedule:	
Creating entries (36% of total time)	4.32 days*
Editing and embedding entries (24%)	2.88 days*
Compiling for review (5%)	0.60 days
Indexer's proofing (5%)	0.60 days*
Review and testing (8%)	0.96 days
Incorporating proof and test edits (4%)	0.48 days*
Handling updated/new topics (5%)	0.60 days*
Recompiling (5%)	0.60 days
Final review and final edits (8%)	0.96 days*
Total	12 days

*=Indexer's work stages

This schedule allows for every stage that needs to happen, but can be simplified, for example, if there are no new topics introduced after indexing, or if you decide to have only one review cycle. If you were working with an unknown tool for the first time, you would want to increase the amount of time devoted to embedding the index, compiling, and testing. If you need to build a thesaurus, add extra time for that as well.

Equipment needs

You will need to settle the kinds of equipment needed to do the project before beginning. If you are working offsite, you must make sure your computer can handle the tools with no memory or incompatibility problems. If you are working onsite, you must be provided with a computer that is powerful enough to do the work. Your basic needs include:

- A computer with enough memory and storage
- Modem or network card
- Authorization and passwords to access files
- Ability to FTP or have remote access if offsite
- Authoring tool(s) installed correctly with all subsidiary software and utilities
- Capability to run authoring tool without conflicts

In some cases, the indexer's machine must also be set up to match the production environment, with fonts, font display utilities, printer drivers, and other special tweaks as needed. This is especially true if Windows Word or a page-layout program is one of the components of the indexing system, or if your project involves creating both print indexing and online indexing for the same files.

Availability of files

With most authoring tools, you often have to share access to the files being indexed with writers and editors. If you are working in-house, make sure that you will have access to the authoring tools and the files for enough time to get the work done.

If you are going to work in stand-alone indexing software, and then embedding or inserting the indexing into the authoring tool, get printouts or copies of the files to work from. With large pieces, rather than printing everything out, you may want to consider a second computer to display the files electronically, allowing you to index on your main computer without cluttered-up screen real estate. Make arrangements for getting the files or printouts efficiently.

Arrangements for transporting the files offsite could get complex, especially if they are big files and contain lots of embedded graphics. Rather than using disks, consider a zip drive or a transportable hard drive, or even a laptop for transporting files. Modeming files from an FTP site or being given remote access to the client's server system and network will help you retrieve files without resorting to disks.

Trial run indexing

Hopefully, you have settled all the issues and can now get down to the basic work of indexing. Whichever system you are using, do a few sample entries, and then do a trial run of the compiled index to see if everything is working the way you think. Make sure you include entries that could cause the compiler to hiccup: sorting problems, special characters, odd punctuation, cross-references, special jumps to animations or odd topics, things you think may not work as planned. You can often figure out work-arounds, and it is easier to implement them at the start.

Fees for online projects

Indexing for online projects often involves problems and delays that are not antici-pated, unless you are working for a group that has established its procedures and used the authoring tools several times before. Even so, charging by the topic or file could leave you losing money, as files get corrupted, indexing codes can get lost or deleted, or the tools you thought would work are incompatible. If you are working with a group that is developing the display for the index as you are indexing, you should allow extra time for problems and testing. Charge an hourly rate for these kinds of projects.

Notes for the next time

Once you have a feel for some of the things that affect getting good online indexing, you will know what questions to ask. Every online project winds up being different— it's always a new challenge.

Further Readings

There is not a lot of material about indexing online documents, but here's a starting list of readings.

American Society of Indexers Web Site, Web indexing page: http://www.well.com/ user/asi/webndx.htm. Provides links to different types of indexes to Web sites, some of which exemplify the hard-coded back-of-the-book approach. Good for getting ideas.

Koch, Traugott, "Literature about Search Services:" http://www.ub2.lu.se//desire/ radar/lit-about-search-services.html. An incredible bibliography about every facet of Internet searching and indexing.

Lathrop, Lori, "Considerations in Indexing Online Documents," *STC Intercom*, January 1996. A good introduction to style and editing issues in online document indexes, regardless of the authoring tool.

Microsoft HTML Help Web site: http://www.microsoft.com/workshop/author/ htmlhelp. Covers the newest version of Help online systems, and its indexing system, under development at Microsoft. The indexing is specified as being inserted in topics, and picked up and compiled into a multi-level display, with a one-to-many approach.

Netscape HTML Help Web site: http://home.netscape.com/eng/help. Covers Netscape's new online help and documentation systems. You will not find much about indexing, as they work with a hard-coded one to be written in HTML.

Wright, Jan C., "Working with Windows Help Keywords," *WinHelp Journal*, Winter 1996 and Spring/Summer 1996. Covers style issues and formatting work-arounds for common WinHelp 95 keywording problems.