

How to index online

Jan C. Wright

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 decisions that should be made and information-gathering that should be done before beginning an online project.

Many back-of-the-book indexers are saying yes to projects published on CD-ROM, accessed through a Web site or through an internal intranet. Agreeing to do a project for one of these environments can be overwhelming. There is so much to learn and figure out, and it seems that 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!

Scope of 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 create the index vary to meet that display. Several factors can affect the methods you use to index online, how the index is displayed, and how work will proceed. The first step is to define the project specifications just as 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.

The boundaries

The first question to ask is will this cosmos of information be infinitely expanded, 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 grow.

Authoring tools

Authoring tools are the software packages that compile 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 (Figure 1). These are usually the easiest to

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	

Figure 1
database-
style
indexing

<p>Monster Maker Pro</p> <p>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.</p> <p>Click a topic below for more information:</p> <p>Overview Creating a monster Animating a monster</p> <p>Topic Unique ID Monster01</p> <p>Monster Maker Pro: overview</p>	<p>Monster Animation</p> <p>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:</p> <p>Tell me about mad monsters Choosing an animation environment</p> <p>Topic Unique ID Monster07</p> <p>Monster Maker Pro: animation; Animation</p>
--	--

A typical topic with an index entry

Another typical topic with index entries

Figure 2
topic-
based
indexing

Index
Animation Monster Maker Pro: animation Monster Maker Pro: overview

The combined index entries in the index display

understand, and can be used with a thesaurus to make the indexing consistent. An example of this format would be a bibliographic database that includes the full text of the article.

- Tools that insert terms into each information topic. The terms 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 an attached field in the topic. When all the topics are compiled into the online product, the authoring tool builds the index display from

the codes found in each topic. The current version of Windows Help is an example of this format (Figure 2).

- 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 (Figure 3).

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

Figure 3
hard-coded indexing

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

Topic Unique ID
 MonsterIndex

Index file listing jumps in invisible codes

- Tools that take electronic files formatted for printing and convert embedded index entries and the files to an online format, with a clickable index. An example of this format

would be FrameMaker or PageMaker files. These software packages can convert embedded indexes in print-based files into an online file (Figure 4).

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

Figure 4
embedded indexing converted to online

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

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 influence how the indexing is done and must be established before beginning.

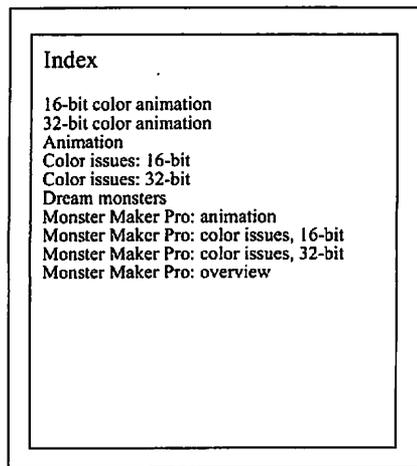
How the index is displayed

Another critical factor to be discussed before starting the indexing work is the index's display. You cannot begin to index effectively without knowing how the index will be displayed.

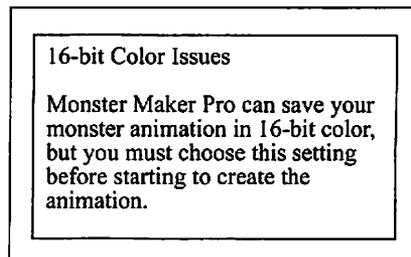
You need to know how your work will appear to the user. Some of the questions you need to have answered are:

How many levels of index will appear? Many tools allow only one level to appear. Other tools allow two level displays. You have to come up with workarounds in a one level index, such as using a colon to visually divide your entries into two levels:

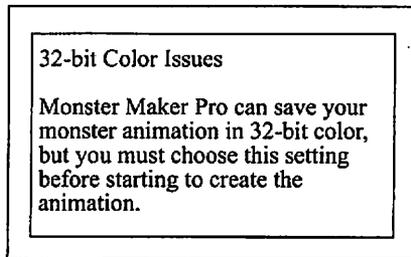
Monster Maker Pro
 Monster Maker Pro: opening
 Monster Maker Pro: quitting



Each entry can lead to only one topic



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

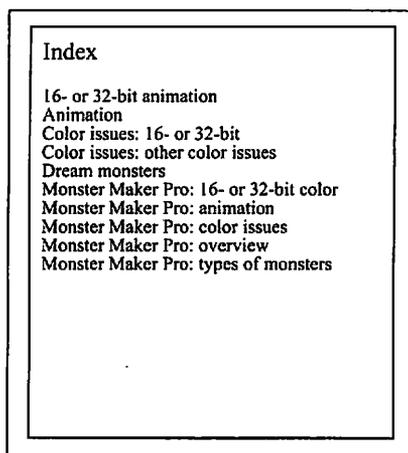


What you get if you click "32-bit color animation"

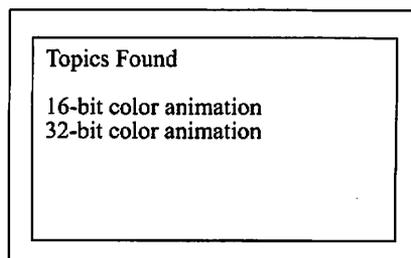
Figure 5
one-to-one indexing

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

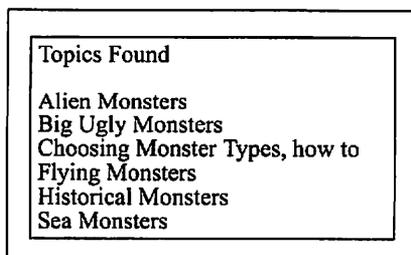
correspondence between index entry and topic? An example of a one-to-one index would be an HTML index like the ASI Web



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



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



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

Figure 6
one-to-many indexing

How to index online

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.

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 wordprocessor, 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. In many cases, you will not be able to use cross-references at all. This will impact how much double-posting you include in the index.

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

How the indexing gets 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 and are not designed with indexing needs in mind. You are working in the dark. An example of this would be a database that has fields for keywords where 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 that you work at the client's worksite so 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 entries into a data file, and then paste it in place. These open tools are very rare, and usually are written in-house.

It is critical to identify someone in the project who can give you technical support for the authoring tool, especially if this is the first time the tool is used. If there is no one to help you, be sure to budget time for debugging problems and figuring out solutions.

Kinds of files 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. They should tell you about 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-post each individual monster directly under its name, of course.

Another type of topic structure you need to ask about is decision trees. Often, you can index to the top of a decision tree to force the user to go through steps in a particular order. 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.

Responsibility for inserting the indexing and maintaining it

Some projects will have you mark indexing on hard copies for someone else to input. 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 consideration with a project that expands and adds new materials after the initial indexing has been completed, is personnel. Who will be maintaining the indexing efforts and how will that process work? Expanding indexes usually require a master list of indexing terms used, or a complete thesaurus. Otherwise, later efforts will not match what is already in the system.

Document 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. Knowing these handoffs and stages 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 should not 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 & 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 & 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 determine the equipment needed to do the project before beginning the actual work. 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(s) 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 with stand-alone indexing software and then embed or insert the index into the authoring tool, be sure to 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; this frees 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. Instead of using disks, consider a Zip drive or a transportable hard drive, or even a laptop for transporting files. Receiving files by modem 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 design 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 anticipated, unless you are working for a group that has established its procedures and has used the authoring tools several times before. Even so, charging by the topic or file could leave you losing money, as files may become 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 is 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 the Help system, 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.

Jan C Wright is the owner of Wright Information Indexing Services, which provides print and online indexing services to software and publishing companies. She was confronted with a CD-ROM index as her very first paid project years ago, and has gone on to do many types of online indexing. She can be reached at JanCW@aol.com.

Indexing Children's Books: Summary of Research

A joint Liverpool John Moores University/ Society of Indexers project, *An investigation into the indexing of children's information books*, began in June 1996, with financial support from the British Library.

The Project Head is Professor Ken Bakewell, Chairman of the Society of Indexers Research Committee and the research is based at the Liverpool Business School.

The research is focusing on books aimed at children aged 7-11 years, National Curriculum Key Stage 2. The aims are to identify the quality of indexes in children's books and, if necessary, to make proposals for improving the effectiveness of such indexes.

Promotion of the project was done through the distribution of press releases to journals, schools of information and library studies in the UK and other relevant organisations. A visit made to the Learning Resources Exhibition in Birmingham in June 1996 also helped to promote the project and to develop useful contacts. Further contacts were made through letters and telephone calls with teachers, librarians, parents, publishers, indexers, school library services and relevant associations such as the School Library Association and the Youth Libraries Group of the Library Association, and the School Libraries Group. A visit was also made to an Anglo-German children's book exhibition at Manchester Central Library, arranged by The British Council and The Goethe-Institut, which gave a helpful list of children's book publishers.

Literature searching, involving the use of online and hard-copy sources, produced a number of relevant items. It was also a priority at this stage to verify the National Curriculum guidelines at Key Stage 2 with regard to indexes. The Curriculum states that 'pupils should be taught how to find information in books and computer-based sources by using organisational devices to help them decide which parts of the material to read more closely' (Department for Education, 1995). It goes on to state a number of organisational devices that should be taught at Key Stage 2 and among these stipulations are the use of indexes.

Questionnaires were designed and discussed with the Society of Indexers Research Committee and with staff at Loughborough University, Department of Information and Library Studies. The questionnaires were piloted with staff at Liverpool John Moores University and with two experienced children's librarians.

A pilot project was carried out with 35 children in Year 3 (aged 7-8) at Frodsham Church of England Primary School, in order to ascertain the range of information books available in schools and to assess the level of children's skills in using indexes. The pilot project allowed for

the testing of the exercise sheets used, which were modified accordingly to be used in other schools.

The next stage of the research was to distribute the questionnaires to teachers, librarians, parents, publishers and indexers, conduct interviews and visit schools to work with children. Responses to the questionnaires have been steady and interviews have been conducted with various experts in children's literature. Work with children in schools has been very successful and informative, with a total of 70 children involved to date. The majority of the children had some idea of what a book index was and how to use it and this allowed for the completion of exercises in index use, classifying information, alphabetisation and cross-references.

The main focus of the research at present is to collate the information from the questionnaires, interviews and children's work sheets for statistical analysis. Work also continues in data collection, with interviews and school visits to be completed shortly. The research is due to be completed by July 1997.

Paula L. Mathews, Research Assistant

References

Department for Education (1995) Key stages 1 & 2 of the national curriculum.

Indexing for the people

"The annual report is to contain:

- a table of contents
- an alphabetical index — covering the contents of any appendixes as well as the contents of the main body of the report;
- a compliance index — showing where in the report the content relating to each requirement is located ...; and
- a glossary — where necessary — to make clear the meanings of any abbreviations and acronyms used."

Government of Australia. *Requirements for departmental annual reports*. Department of the Prime Minister and Cabinet, March 1994 (updated February 1996)